Extremal Hypergraph Problems

by

Xizhi Liu
B.S. University of Science and Technology of China, 2018
M.S. University of Illinois at Chicago, 2020

Thesis submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Mathematics
in the Graduate College of the
University of Illinois at Chicago, 2022

Chicago, Illinois

Defense Committee:
Dhruv Mubayi, Chair and Advisor
Marcus Michelen
Will Perkins
Oleg Pikhurko, University of Warwick
Aditya Potukuchi

ACKNOWLEDGMENTS

I am deeply grateful to my advisor Prof. Dhruv Mubayi for his invaluable advice, continuous support, and encouragement during my Ph.D. study. His immense knowledge and plentiful experience have encouraged me in all the time of my research. It is always a great pleasure and inspiration for me to work with him. I am also very grateful to Prof. Jie Ma for his support and patience during my undergraduate study at the University of Science and Technology of China. Many thanks to Marcus Michelen, Will Perkins, Oleg Pikhurko, and Aditya Potukuchi for serving on my thesis committee.

CONTRIBUTION OF AUTHORS

Chapter 1 contains a list of notations. Chapter 2 is a brief introduction to topics covered in this thesis. Chapter 3 contains a joint paper with Dhruv Mubayi which has been published in the Journal of Combinatorial Theory, Series B [167] and two papers by myself, one of which has been published in the European Journal of Combinatorics [163], and the other one has appeared on arXiv [159]. Chapter 4 contains a joint paper with Dhruv Mubayi and Christian Reiher that has appeared on arXiv [171]. Chapter 5 contains a joint paper with Dhruv Mubayi that has been published in Combinatorica [169] and two joint papers with Dhruv Mubayi and Christian Reiher, one of which has appeared on arXiv [173] and the other one is still in preparation. Chapter 6 contains a joint paper with Dhruv Mubayi that has been published in the European Journal of Combinatorics [168] and three papers by myself, one of which was published in the European Journal of Combinatorics [160], one was published in the Electronic Journal of Combinatorics [162], and another one has appeared on arXiv [161]. Chapter 7 contains a joint paper with Jie Ma thas has been published in the Journal of Graph Theory [164], and a joint paper with Dhruv Mubayi that has been published in the Journal of Graph Theory [166]. Chapter 8 contains a joint paper with Dhruv Mubayi and Christian Reiher that has appeared on arXiv [172]. Chapter 9 contains a joint paper with Tom Bohman that has been published in Random Structures & Algorithms [22], and a joint paper with Dhruv Mubayi that has been accepted for publication in the Electronic Journal of Combinatorics [165].

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

TABLE OF CONTENTS (Continued)

CHAPTER **PAGE**

TABLE OF CONTENTS (Continued)

CHAPTER **PAGE**

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES (Continued)

FIGURE **PAGE**

LIST OF ABBREVIATIONS

UIC	University of Illinois at Chicago

SUMMARY

This thesis studies extremal problems in hypergraph theory, set theory, and graph theory. Results in this thesis can be divided into seven parts.

In the first part, we study the feasible region $\Omega(\mathcal{F})$ of a hypergraph family $\mathcal{F}$, which is the set of points (x, y) so that there exists a sequence of $\mathcal{F}$-free r-graphs whose shadow densities approach x and whose edge densities approach y. We prove some general results about the shape of $\Omega(\mathcal{F})$, and study $\Omega(\mathcal{F})$ for some specific examples such as the cancellative hypergraphs and the expansion of cliques.

In the second part, we present a unified framework for proving stability theorems in graph and hypergraph theory. Our main result reduces stability for a large class of hypergraph Turán problems to the simpler question of checking that a hypergraph $\mathcal{H}$ with large minimum degree that omits the forbidden structures is vertex-extendable. We illustrate our method by giving new short proofs to many stability theorems.

In the third part, we provide a construction of finite hypergraph families that have arbitrarily (but finite) many extremal configurations. This is the first such construction. Before our work, every family of hypergraphs whose Turán density is known has a unique extremal configuration.

In the fourth part, we study problems that are generalizations of the celebrated Erdős–Ko–Rado theorem. We give the correct bounds for the size of a family that does not contain a d-cluster but contains at least two disjoint edges. This resolves a conjecture of Mammoliti and Britz. We also extend a structural theorem due to Frankl about conditionally intersecting

3-graphs to the general case, and use it to give new proofs to some theorems in Extremal set theory. Extending the celebrated Katona intersecting shadow theorem, we give tight bounds for the size of the shadow of t-intersecting families and families with a bounded matching number. Finally, resolving a conjecture of Mubayi and Verstraëte, we determine the maximum size of a family that does not contain nontrivial intersecting subgraphs.

In the fifth part, we study some extensions of the Turán theorem in Graph theory. We determine the minimum number of copies of a color-critical graph F in a graph with fixed number of edges and fixed F-covering number. We also consider the local density problem in K_4-free graphs and prove a conjecture of Chung and Graham, and independently, Erdős, Faudree, Rousseau, and Schelp for all almost-regular K_4-free graphs.

In the sixth part, we study the independence number of hypergraps that omit just one intersection. This is related to Rödl and Šiňajová's result on the independence number of (n, k, ℓ)-systems. We also show some explicit constructions of (n, k, ℓ)-systems with small independence number. Such constructions are usually very useful in theoretical computer science.

In the last part, we study the feasible region $\Omega_{\text{ind}}(F)$ of induced graphs F, which is the set of points (x, y) so that there exists a sequence of graphs whose edge density approaches x and whose induced F-density approaches y. We prove some general results about the shape of $\Omega_{\text{ind}}(F)$, and study $\Omega_{\text{ind}}(F)$ for some specific examples such as complete bipartite graphs and complete graphs minus one edge K_r^-. Our results on K_r^- sharpen those predicted by the edge-statistics conjecture of Alon et. al.

CHAPTER 1

LIST OF NOTATIONS

In this section, we list some notations and definitions that will be used frequently later.

- For a positive integer n let $[n] = \{1, \ldots, n\}$.
- For two positive integers m and n with $m \le n$ let $[m, n] = \{m, m+1, \ldots, n\}$.
- For a set V and an integer $r \ge 0$ let $\binom{V}{r}$ be the collection of all r-subsets of V.
- An r-uniform hypergraph (henceforth r-graph) $\mathcal{H}$ on V is a subset of $\binom{V}{r}$.
- A graph is a 2-graph.
- Given two r-graphs $\mathcal{H}$ and $\mathcal{H}'$ with the same number of vertices the edit-distance of $\mathcal{H}$ and $\mathcal{H}'$ is

  $$d_1(\mathcal{H}, \mathcal{H}') = \min\{|H \triangle H''| \colon V(H'') = V(H) \text{ and } \mathcal{H}'' \cong \mathcal{H}'\}.$$

  It is well known and easy to confirm that this distance satisfies the triangle inequality.
- For integers $\ell \ge r \ge 2$ let K_ℓ^r denote the complete r-graph on ℓ vertices. If $r = 2$, we simplify the notation by omitting the superscript r.
- For a hypergraph $\mathcal{H}$ we use $V(\mathcal{H})$ and $E(\mathcal{H})$ to denote the vertex set and edge set of $\mathcal{H}$, respectively.
- Let $v(\mathcal{H})$ and $e(\mathcal{H})$ denote the size of $V(\mathcal{H})$ and $E(\mathcal{H})$, respectively.
- The edge density of an r-graph $\mathcal{H}$ is $\rho(\mathcal{H}) = e(\mathcal{H})/\binom{v(\mathcal{H})}{r}$.
- For a vertex set $S \subset V(\mathcal{H})$ the induced subgraph of $\mathcal{H}$ on S is denoted by $\mathcal{H}[S]$.
- For two disjoint vertex sets $S, T \subset V(\mathcal{H})$ the hypergraph $\mathcal{H}[S, T]$ denotes the set of edges in $\mathcal{H}$ that have nonempty intersection with both S and T.

- Given an r-graph $\mathcal{H}$ the shadow $\partial\mathcal{H}$ of $\mathcal{H}$ is an $(r-1)$-graph which is defined as

$$\partial\mathcal{H} = \left\{ A \in \binom{V(\mathcal{H})}{r-1} : \exists B \in \mathcal{H} \text{ such that } A \subset B \right\}.$$

- For integers $r > i \geq 2$ define the i-th shadow of $\mathcal{H}$ as $\partial_i\mathcal{H} = \partial(\partial_{i-1}\mathcal{H})$, where $\partial_1\mathcal{H} = \partial\mathcal{H}$.
- For a vertex $v \in V(\mathcal{H})$ the link $L_{\mathcal{H}}(v)$ of v in $\mathcal{H}$ is

$$L_{\mathcal{H}}(v) = \{A \in \partial\mathcal{H} \colon A \cup \{v\} \in \mathcal{H}\}.$$

- For a vertex $v \in V(\mathcal{H})$ the neighborhood $N_{\mathcal{H}}(v)$ of v in $\mathcal{H}$ is

$$N_{\mathcal{H}}(v) = \{u \in V(\mathcal{H}) \setminus \{v\} \colon \exists E \in \mathcal{H} \text{ such that } \{u, v\} \subset E\}.$$

- The degree of a vertex $v \in V(\mathcal{H})$ is $d_{\mathcal{H}}(v) = |L_{\mathcal{H}}(v)|$.
- The maximum degree and the minimum degree of $\mathcal{H}$ are denoted by $\Delta(\mathcal{H})$ and $\delta(\mathcal{H})$, respectively.
- Given two r-graphs F and $\mathcal{H}$ let $N(F, \mathcal{H})$ denote the number of (not necessarily induced) copies of F in $\mathcal{H}$.
- Given two r-graphs F and $\mathcal{H}$ let $N_{\text{ind}}(F, \mathcal{H})$ denote the number of induced copies of F in $\mathcal{H}$.

- Let $\rho(F,G) = N(F,\mathcal{H})/\binom{v(\mathcal{H})}{v(F)}$ and $\rho_{\text{ind}}(F,G) = N_{\text{ind}}(F,\mathcal{H})/\binom{v(\mathcal{H})}{v(F)}$ denote the F-density and the induced F-density of $\mathcal{H}$, respectively.
- Given a family $\mathcal{F}$ of r-graphs we say an r-graph $\mathcal{H}$ is $\mathcal{F}$-free if $N(F,\mathcal{H}) = 0$ for all $F \in \mathcal{F}$.
- A vertex set $I \subset V(\mathcal{H})$ is independent if no edge of $\mathcal{H}$ is completely contained in I.
- The independence number $\alpha(\mathcal{H})$ is the size of the maximum independent set in $\mathcal{H}$.
- The covering number $\tau(\mathcal{H})$ of a hypergraph $\mathcal{H}$ is the minimum size of a vertex $S \subset V(\mathcal{H})$ such that every edge has at least one vertex in S.
- A graph G is ℓ-colorable (or ℓ-partite) if there exists a partition $V(G) = V_1 \cup \cdots \cup V_\ell$ such that V_i is independent in G for all $i \in [\ell]$.
- The chromatic number $\chi(G)$ of a graph G is the smallest integer ℓ such that G is ℓ-colorable.
- The Turán graph $T(n,\ell)$ is the complete ℓ-partite graph on n vertices with the most number of edges.
- The Turán r-graph $T_r(n,\ell)$ is the r-graph on $[n]$ such that every edge has at most one vertex in V_i for $i \in [\ell]$, where $[n] = V_1 \cup \cdots \cup V_\ell$ is a partition such that $||V_i| - |V_j|| \leq 1$ for all $i,j \in [\ell]$.
- The standard n-simplex Δ^n is defined as

$$\Delta^n = \left\{ x \in \mathbb{R}^{n+1} \colon \sum_{i=1}^{n+1} x_i = 1 \text{ and } x_i \geq 0 \text{ for all } i \in [n+1] \right\}.$$

- For an r-graph $\mathcal{H}$ on $[n]$ define the weight polynomial of a hypergraph $\mathcal{H}$ as

$$p_{\mathcal{H}}(x_1, \ldots, x_n) = \sum_{E \in \mathcal{H}} \prod_{i \in E} x_i.$$

- The Lagrangian of an n-vertex r-graph $\mathcal{H}$ is

$$\lambda(\mathcal{H}) = \max \left\{ p_{\mathcal{H}}(x) \colon x \in \Delta^{n-1} \right\}.$$

- The matching number $\nu(\mathcal{H})$ of a hypergraph $\mathcal{H}$ is the maximum number of disjoint edges in $\mathcal{H}$.

- For integers $n \geq k \geq \ell$ and $t \geq 1$ an (n, k, ℓ, t)-system is an r-graph on n vertices such that every ℓ-set is contained in at most t edges. We also use (n, k, ℓ)-system to represent an $(n, k, \ell, 1)$-system.

- For integers $n \geq k \geq \ell$ and $t \geq 1$ a Steiner (n, k, ℓ, t)-system is an r-graph on n vertices such that every ℓ-set is contained in exactly t edges. We also use Steiner (n, k, ℓ)-system to represent a Steiner $(n, k, \ell, 1)$-system.

- For integers $k \geq \ell \geq 0$ and $\lambda \geq 1$ the sunflower $S^k_\lambda(\ell)$ is a k-graph that consists of λ edges $E_1, \ldots, E_\lambda$ such that $E_i \cap E_j = S$ for $1 \leq i < j \leq \lambda$ and some fixed set S (called the center) of size ℓ. We will omit the superscript k if it is clear from the text.

- Given two r-graphs F and $\mathcal{H}$ a map $\phi \colon V(F) \to V(\mathcal{H})$ is said to be a homomorphism if it preserves edges, i.e., if $\phi(E) \in \mathcal{H}$ holds for all $E \in F$. If ϕ is surjective and every edge of $\mathcal{H}$ is an image of an edge of F, i.e., $\mathcal{H} = \{\phi(E) \colon E \in F\}$, we call $\mathcal{H}$ a homomorphic image of F.

- An r-graph $\mathcal{H}$ is said to be a blowup of another r-graph F if there exists a map $\psi \colon V(\mathcal{H}) \longrightarrow V(F)$ such every $E \in \binom{V(\mathcal{H})}{r}$ satisfies the equivalence $\psi(E) \in F \iff E \in \mathcal{H}$. If ψ is surjective, the blowup is called proper.

- Given two r-graphs F and $\mathcal{H}$ we say $\mathcal{H}$ is F-colorable if $\mathcal{H}$ is a subgraph of some blowup of F.

- An r-graph $\mathcal{S}$ is a star if all edges in $\mathcal{S}$ contain a fixed vertex v, which is called the center of $\mathcal{S}$.

- Given two positive functions $f(n)$ and $g(n)$ we write $f(n) = O\left(g(n)\right)$, or equivalently, $g(n) = \Omega\left(f(n)\right)$ if there exists a constant $C > 0$ such that $f(n) \le Cg(n)$ for all sufficiently large n, we write $f(n) = o\left(g(n)\right)$ if $\lim_{n\to\infty} f(n)/g(n) = 0$, and we write $f(n) = \Theta\left(g(n)\right)$ if both $f(n) = O\left(g(n)\right)$ and $f(n) = \Omega\left(g(n)\right)$ hold.

CHAPTER 2

INTRODUCTION

2.1 The feasible region of hypergraphs

Let $\mathcal{F}$ be a family of r-graphs. The Turán number $\mathrm{ex}(n, \mathcal{F})$ of $\mathcal{F}$ is the maximum number of edges in an $\mathcal{F}$-free r-graph on n vertices. Using a simple averaging argument, Katona, Nemetz, and Simonovits [133] showed that $\mathrm{ex}(n, \mathcal{F})/\binom{n}{r}$ is decreasing in n. Hence the limit $\lim_{n\to\infty} \mathrm{ex}(n, \mathcal{F})/\binom{n}{r}$ exists. This limit is called the Turán density of $\mathcal{F}$ and is denoted by $\pi(\mathcal{F})$.

A fundamental theorem in Extremal graph theory due to Turán [241] is that $\mathrm{ex}(n, K_{\ell+1}) = |T(n, \ell)|$ for all integers $\ell \geq 2$, and moreover, the Turán graph $T(n, \ell)$ is the unique $K_{\ell+1}$-free graph on n with $\mathrm{ex}(n, K_{\ell+1})$ edges. Erdős and Stone [74] extended Turán's result and proved that $\pi(F) = \frac{1}{\chi(F)-1}$ for every graph F. Later, Simonovits [71] observed that the Erdős–Stone theorem implies that $\pi(\mathcal{F}) = \frac{1}{\chi(\mathcal{F})-1}$ for all graph families $\mathcal{F}$, where $\chi(\mathcal{F}) = \min\{\chi(F) \colon F \in \mathcal{F}\}$. However, on the other hand, there is comparatively little understanding of the Turán problem for r-graphs when $r \geq 3$. Determining $\pi(\mathcal{F})$ for r-graph families with $r \geq 3$ is known to be notoriously hard in general, and it is a major open problem to determine $\pi(K_\ell^r)$ for any case with $\ell > r \geq 3$. A long-standing conjecture of Turán states as follows.

Conjecture 2.1.1 (Turán [241])**.** $\pi(K_\ell^3) = 1 - \left(\frac{2}{\ell-1}\right)^2$.

Despite lots of effort has been devoted to this area and successively better upper bounds for $\pi(K_4^3)$ were obtained by de Caen [49], Giraud (see [38]), Chung and Lu [38], and Razborov [214], Conjecture 2.1.1 still remains wide open. The current record is $\pi(K_4^3) \leq 0.561666$, which was obtained by Razborov [214] using flag algebra machinery. To generate more interest in this conjecture, it worth mentioning that Erdős [62] offered \$500 for the solution of Conjecture 2.1.1 for any case and \$1000 for a general solution. For more results on hypergraph Turán problems before 2011 we refer the reader to an excellent survey by Keevash [135].

To gain a better understanding of hypergraph Turán problems. We combine it with the Kruskal–Katona [154; 132] theorem, another seminal result in Combinatorics, which gives a tight upper bound for $|\mathcal{H}|$ as a function of $|\partial\mathcal{H}|$. More specifically, in Chapter 3, we consider the following more general question.

If $\mathcal{H}$ is $\mathcal{F}$-free, what are the possible values of $|\mathcal{H}|$ for fixed $|\partial\mathcal{H}|$?

We are interested in the asymptotical solution to the question above. This motivates us to study the feasible region $\Omega(\mathcal{F})$ of r-graph families $\mathcal{F}$, which is the set of points (x, y) in the unit square such that there exists a sequence of $\mathcal{F}$-free r-graphs whose shadow densities approach x and whose edge densities approach y.

The feasible region provides a lot of combinatorial information, for example, the supremum of y over all $(x, y) \in \Omega(\mathcal{F})$ is the Turán density $\pi(\mathcal{F})$, and $\Omega(\emptyset)$ gives the Kruskal–Katona theorem.

We undertake a systematic study of $\Omega(\mathcal{F})$, and prove that $\Omega(\mathcal{F})$ is completely determined by a left-continuous almost everywhere differentiable function $g(\mathcal{F})$; and moreover, there exists a finite family $\mathcal{F}$ for which $g(\mathcal{F})$ is not continuous. We show that for a finite family $\mathcal{F}$ the function $g(\mathcal{F})$ can have arbitrarily finitely many global maxima and countably many local maxima. We also extend some old theorems about various hypergraph Turán problems. For example, we completely determine the feasible region of the weak expansion of complete graphs and almost completely determine the feasible region for cancellative triple systems.

2.2 Hypergraph stability

Many families $\mathcal{F}$ have the property that there is a unique $\mathcal{F}$-free hypergraph $\mathcal{G}$ on n vertices achieving $\mathrm{ex}(n, \mathcal{F})$, and moreover, every $\mathcal{F}$-free hypergraph $\mathcal{H}$ of size close to $\mathrm{ex}(n, \mathcal{F})$ can be transformed to $\mathcal{G}$ by deleting and adding very few edges. Such a property is called stability of $\mathcal{F}$. The first stability theorem was proved independently by Erdős and Simonovits [233], which shows that if the number of edges in an n-vertex K_ℓ-free graph G is $(1 - o(1))\mathrm{ex}(n, K_\ell)$, then G can be transformed to the Turán graph $T(n, \ell - 1)$ by deleting and adding $o(n^2)$ edges.

The stability phenomenon has been used to determine $\mathrm{ex}(n, \mathcal{F})$ exactly in many cases. It was first used by Simonovits in [233] to determine $\mathrm{ex}(n, F)$ exactly for all edge-critical graphs F and large n, and then by several authors (e.g. see [113; 141; 142; 195; 209; 30; 205]) to prove exact results for hypergraphs.

However, there are many Turán problems for hypergraphs that (perhaps) do not have the stability property. The famous example K_4^3 we mentioned above was shown to have exponentially many extremal constructions in the number of vertices (e.g. see Kostochka [151] and

Brown [32]). These constructions can be used to show that K_4^3 does not have the stability property (assuming Turán's tetrahedron conjecture is true). For K_ℓ^3 with $\ell \geq 5$, different near-extremal constructions were given by Sidorenko [228], and Keevash and Mubayi [135]. These also show that K_ℓ^3 does not have stability (assuming Conjecture 2.1.1 is true).

The absence of stability seems to be a fundamental barrier in determining the Turán numbers of some families. Indeed, the Turán numbers of the examples we presented above are not known, even asymptotically, and in fact, no Turán number of a family without the stability property has been determined.

In Chapter 5, we provide the first such example. More specifically, we construct a family $\mathcal{M}$ of 3-graphs, prove that $\mathcal{M}$ has exactly two different extremal configurations, and hence, does not have the stability property, We also determine $\pi(\mathcal{M})$, and even $\mathrm{ex}(n, \mathcal{M})$ for infinitely many n.

Further, using some results from Design theory, for every positive integer t we construct a finite family of triple systems $\mathcal{M}_t$, determine its Turán number, and show that there are exactly t extremal $\mathcal{M}_t$-free configurations that are far from each other in edit-distance.

We also prove a strong stability theorem: every $\mathcal{M}_t$-free triple system whose size is close to the maximum size is a subgraph of one of these t extremal configurations after removing a small proportion of vertices. This is the first stability theorem for a hypergraph problem with an arbitrary (finite) number of extremal configurations. Moreover, the extremal hypergraphs have very different shadow sizes (unlike the case of the famous Turán tetrahedron conjecture).

Hence a corollary of our result is that the function $g(\mathcal{M}_t)$ has exactly t global maxima. We also extend the construction to the general case of $r \geq 4$.

To prove the stability theorem above, in Chapter 4, we develop a method which provides a unified framework for most stability theorems that have been proved in graph and hypergraph theory. Our main result reduces stability for a large class of hypergraph problems to the simpler question of checking that a hypergraph $\mathcal{H}$ with large minimum degree that omits the forbidden structures is vertex-extendable. This means that if v is a vertex of $\mathcal{H}$ and $\mathcal{H} - v$ is a subgraph of the extremal configuration(s), then $\mathcal{H}$ is also a subgraph of the extremal configuration(s). In many cases vertex-extendability is quite easy to verify.

We illustrate our approach by giving new short proofs of hypergraph stability results of Pikhurko [208], Hefetz–Keevash [122], Brandt–Irwin–Jiang [30], Bene Watts–Norin–Yepremyan [17] and others. Since our method always yields minimum degree stability, which is the strongest form of stability, in some of these cases our stability results are stronger than what was known earlier. Along the way, we clarify the different notions of stability that have been previously studied.

2.3 Extremal set theory

Many problems in Extremal set theory can be viewed as Turán-type problems, but these problems usually have very simple extremal constructions. We call a family $\mathcal{H}$ a star if every edge in it contains a fixed vertex, which is called the center of $\mathcal{H}$. A family is intersecting if every pair of edges in it has a nonempty intersection. The celebrated Erdős–Ko–Rado theorem [69] states that if $n \geq 2r$ and $\mathcal{H} \subset \binom{[n]}{r}$ is an intersecting family, then $|\mathcal{H}| \leq \binom{n-1}{r-1}$. Moreover, equality holds only if $\mathcal{H}$ is a star when $n > 2r$. There has been many extensions of the Erdős–Ko–Rado theorem since it was proved. One extension that was proposed by Mubayi [190] is as follows.

A collection of d different r-sets $A_1, \ldots, A_d$ is called a d-cluster if $|A_1 \cup \cdots \cup A_d| \leq 2r$ and $|A_1 \cap \cdots \cap A_d| = 0$. A family $\mathcal{F} \subset \binom{[n]}{r}$ is d-cluster-free if it does not contain d-clusters. Notice that an intersecting family is simply a 2-cluster-free family. Mubayi [190] conjectured that for every $n \geq dr/(d-1)$ with $r \geq d \geq 3$ every d-cluster-free family $\mathcal{H} \subset \binom{[n]}{r}$ has size at most $\binom{n-1}{r-1}$. Moreover, equality holds only if $\mathcal{H}$ is a star. Mammoliti and Britz [185] considered Mubayi's conjecture for a special kind of families, and they proposed the problem of determining the maximum size of an n-vertex d-cluster-free set with at least two disjoint edges. They conjectured that the extremal construction should be the disjoint union of a star and an edge. In Section 6.1, we show that their conjecture is true for $r = 3$ but false for $r \geq 4$. Moreover, we provide the correct bounds for the case $r \geq 4$, and show that they are related to the Hypergraph Turán problems that allow multiple edges.

A more general problem is to consider the so-called conditionally intersecting families. We say a family $\mathcal{H} \subset \binom{[n]}{r}$ is (d,s)-conditionally intersecting if it does not contain d sets with union of size at most s and empty intersection. In particular, a family $\mathcal{H} \subset \binom{[n]}{r}$ is $(d,2r)$-conditionally intersecting if it does not contain d-clusters. In [95], Frankl studied the structure of $(3,6)$-conditionally intersecting families and used it to give a new proof of the Erdős–Ko–Rado theorem for 3-graphs. In Section 6.2, we extended his result to (d,s)-conditionally intersecting families with $s \geq 2r + d - 3$ and $(r, 2r)$-conditionally intersecting families, and use them to give new proofs to some classical theorems in Extremal set theory and the Mammoliti–Britz conjecture mentioned above for $r = 3$.

A triangle is a family of three sets A, B, C such that $A \cap B, B \cap C, C \cap A$ are nonempty but $A \cap B \cap C$ is empty. Settling a longstanding conjecture of Erdős [77], Mubayi and Verstraëte [198] proved that for $r \geq 3$ and $n \geq 3r/2$ every triangle-free family $\mathcal{H} \subset \binom{[n]}{r}$ has size at most $\binom{n-1}{r-1}$, and moreover, equality holds only if $\mathcal{H}$ is a star. Let us call a hypergraph $\mathcal{H} \subset \binom{[n]}{r}$ non-trivial intersecting if every two edges in it have a nonempty intersection but no vertex is contained in all edges of $\mathcal{H}$. In fact, Mubayi and Verstraëte proved a stronger statement which says that for every $r \geq d+1 \geq 3$ and $n \geq (d+1)r/d$ every family $\mathcal{H} \subset \binom{[n]}{r}$ without a non-trivial intersecting subgraph of size $d + 1$ has at most $\binom{n-1}{r-1}$ edges. They conjectured that a similar result holds even for $d \geq r \geq 4$ when n is sufficiently large. We prove their conjecture in Section 6.4.

For $r = 3$, Mubayi and Verstraëte observed that a Steiner $(n, 3, 2, k-1)$-system does not contain a non-trivial intersecting family of size $3k + 1$ whenever $k \geq 2$. Hence they conjectured that for $k \geq 2$ and sufficiently large n the 3-graph $\mathcal{H} \subset \binom{[n]}{3}$ that does not contain a non-trivial intersecting family of size $3k + 1$ and with the most number of edges should be a Steiner $(n, 3, k-1)$-system (if there is such a Steiner $(n, 3, 2, k-1)$-system). In Section 6.4, we show this conjecture is false.

As we mentioned before, the seminal Kruskal–Katone theorem gives a tight upper bound for $|\mathcal{H}|$ as a function of $|\partial\mathcal{H}|$. This theorem was extended to many families with additional properties. One such central result is due to Katona [131] who proved the following theorem for t-intersecting families, which are families in which every two sets have at least t common elements.

Theorem 2.3.1 (Katona [131]). *Let $n \geq k > t \geq \ell \geq 1$. If $\mathcal{H} \subset \binom{[n]}{k}$ is t-intersecting, then*

$$|\partial_\ell \mathcal{H}| \geq \frac{\binom{2k-t}{k-\ell}}{\binom{2k-t}{k}} |\mathcal{H}|.$$

The only case of equality in Theorem 2.3.1 is when $n = 2k - t$ and $\mathcal{H} \cong \binom{[2k-t]}{k}$ (see [1]).

Theorem 2.3.1 is a foundational result in extremal set theory with many applications. Its first application was to prove a conjecture of Erdős–Ko–Rado on the maximum size of a t-intersecting family in $2^{[n]}$. It was used to obtain short new proofs for several classical results. For example, Frankl–Füredi [103] used it to give a short proof for the Erdős–Ko–Rado theorem, and Frankl–Tokushige [106] used it to obtain a short proof for the Hilton–Milner theorem. It also has many applications to Sperner families and other types of intersection problems [28; 42; 98; 100; 116; 185; 224; 250].

Section 6.3 is concerned with improving the bounds in Theorem 6.3.2 and related results about shadows of families with certain properties. Our proofs use some structural results about t-intersecting families and in many cases the bounds we proved are best possible.

2.4 Extension of Turán's theorem

Before the celebrated Turán theorem was proved, Mantel [186] showed that every graph on n vertices with $\lfloor n^2/4 \rfloor + 1$ edges contains at least one copy of K_3. Rademacher showed that there are actually at least $\lfloor n/2 \rfloor$ copies of K_3 in such graphs. Later, Erdős [56; **?**] proved that if $t \leq cn$ for some small constant $c > 0$, then every graph on n vertices with $\lfloor n^2/4 \rfloor + t$ edges contains at least $t \lfloor n/2 \rfloor$ copies of K_3. Erdős also conjectured that the same conclusion holds for all $t < n/2$. Later, Lovász and Simonovits [178] proved Erdős' conjecture and they also proved a similar result for K_r with $r \geq 4$. In [194], Mubayi extended their results by proving tight bounds on the number of copies of color critical graphs in a graph with a prescribed number of vertices and edges.

Xiao and Katona [249] proposed a generalized Erdős–Rademacher problem by putting constraints on $\tau_F(G)$. More precisely, they asked for the minimum value of $N(F, G)$ for graphs G with a fixed number of vertices and edges and a fixed F-covering number which is the minimum size of $S \subset V(G)$ such that every copy of F in G has at least one vertex in S.

In Section 7.2, we show the correct bound on the number of copies of K_3 for all s, t and sufficiently large n, and also prove several bounds, which is tight up to some error term, on the number of copies of K_r and r-critical graphs F in a graph G on n vertices with $t_{r-1}(n) + t$ edges and $\tau_r(G) = s$.

In Section 7.1, we consider the following generalization of Turán's theorem that was initialized by Erdős [61]: Given a constant $0 \leq \alpha \leq 1$, what is the minimum value $\beta = \beta(\alpha, r)$ such that every n-vertex K_r-free graph contains a vertex set of size $\lfloor \alpha n \rfloor$ which spans at most βn^2 edges? This is often referred as the local density problem. The case $\alpha = 1/2$ is of special interest. Erdős [63] offered \$250 for the first solution to the following long-standing conjecture on triangle-free graphs.

Conjecture 2.4.1 (Erdős [61])**.** *Every triangle-free graph on n vertices contains a vertex set of size $\lfloor n/2 \rfloor$ that spans at most $n^2/50$ edges.*

The balanced blow-ups of both the 5-cycle and the Petersen graph show that the bound $n^2/50$ would be best possible if this conjecture is true. Despite extensive research [153; 143; 203; 15], Conjecture 2.4.1 is still open.

A similar question also has been asked for K_4-free graphs. Chung and Graham [40], and Erdős, Faudree, Rousseau and Schelp [65] posted the following conjecture.

Conjecture 2.4.2 (Chung et al. [40], Erdős et al. [65])**.** *Every K_4-free graph on n vertices contains a vertex set of size $\lfloor n/2 \rfloor$ that spans at most $n^2/18$ edges.*

The Turán graph $T_3(n)$ shows that the bound $n^2/18$ in Conjecture 2.4.2 would be best possible if it is true. A closely related conjecture of Erdős (see [64]), which was proved by Sudakov [237], states that every K_4-free graphs on n vertices can be made bipartite by deleting at most $n^2/9$ edges. An interesting interplay between these problems for regular graphs was observed by Krivelevich [153], where he pointed out that a bound in the local density problem can imply a bound (doubled) in the problem of making a graph bipartite; also see [237] for an illustration.

In Section 7.1, we prove Conjecture 2.4.2 for almost-regular graphs. It is worth mentioning that utilizing some of our results and some ideas from the Ramsey–Turán theory, Reiher [218] completely resolved this conjecture very recently.

2.5 The feasible region of induced graphs

A quantum graph Q is a formal linear combination of finitely many graphs, i.e., an expression of the form

$$Q = \sum_{i=1}^{m} \lambda_i F_i \,,$$

where m is a nonnegative integer, the numbers $\lambda_1, \ldots, \lambda_m$ are real, and $F_1, \ldots, F_m$ are graphs. We call F_i a constituent of Q if $\lambda_i \neq 0$. Two quantum graphs Q, Q' are equal if they have the same constituents and the same (nonzero) coefficients for each constituent. The complement of Q is $\overline{Q} = \sum_{i=1}^{m} \lambda_i \overline{F}_i$, where $\overline{F}_i$ denotes the complement of F_i for each $i \in [m]$. A quantum graph Q is self-complementary if $Q = \overline{Q}$. Every graph parameter f can be extended linearly to quantum graphs by stipulating $f(Q) = \sum_{i=1}^{m} \lambda_i f(F_i)$. In particular,

$$N(Q, G) = \sum_{i=1}^{m} \lambda_i N(F_i, G) \quad \text{and} \quad \rho(Q, G) = \sum_{i=1}^{m} \lambda_i \rho(F_i, G) \,.$$

The feasible region $\Omega_{\text{ind}}(Q)$ of a quantum graph Q is the collection of points (x, y) in the unit square such that there exists a sequence of graphs whose edge densities approach x and whose induced Q-densities approach y.

For single graphs, a complete description of $\Omega_{\text{ind}}(F)$ is not known for any F with at least four vertices that is not a clique or an independent set (see [216; 201; 217]). The feasible region provides a lot of combinatorial information about F. For example, $\Omega_{\text{ind}}(K_r)$ yields the Kruskal–

Katona [154; 132] and clique density theorems, and the supremum of y over all $(x, y) \in \Omega_{\text{ind}}(F)$ is the inducibility of F, which is defined as

$$\text{ind}(F) = \lim_{n\to\infty} \max \{\rho(F, G)\colon v(G) = n\}.$$

In Chapter 8, we begin a systematic study of $\Omega_{\text{ind}}(Q)$. We prove some general results about the shape of $\Omega_{\text{ind}}(Q)$. Our main result here shows that $\Omega_{\text{ind}}(Q)$ is a closed set and is completely determined by two continuous and almost everywhere differentiable functions $I(Q, x)$ and $i(Q, x)$. We also study $\Omega_{\text{ind}}(F)$ for some specific choices of graphs F for which $\text{ind}(F)$ has been investigated by many researchers. We prove a general upper bound for $I(F, x)$ where F are complete multipartite graphs, this generalizes an old result of Bollobás [25] for the number of cliques in a graph with given edge density. Prior to our work, $\Omega_{\text{ind}}(F)$ for a single graph F was determined only when F is a clique or an independent set. Here we extend this to the case $F = K_{1,2}$ and also obtain results for complete bipartite graphs. Furthermore we study $\Omega_{\text{ind}}(K_r^-)$, where K_r^- arises from the clique K_r by the deletion of a single edge. As a consequence of our results, we determine the inducibility $\text{ind}(K_r^-)$, which is new for $r \geq 5$. Our results sharpen those predicted by the edge-statistics conjecture of Alon et. al. while also extending a theorem of Hirst [124] for K_4^- that was proved using computer aided techniques and flag algebras.

2.6 Independent sets in sparse hypergraphs

A central topic in Combinatorics is to study the independence number of various family of hypergraphs. The celebrated Turán theorem [241] implies that $\alpha(G) \geq n/(d+1)$ for every graph G on n vertices with average degree d. Later, Spencer [234] extended Turán's result and proved that for all $k \geq 3$ every n-vertex k-graph $\mathcal{H}$ with average degree d satisfies

$$\alpha(\mathcal{H}) \geq c_k \frac{n}{d^{1/(k-1)}} \tag{2.1}$$

for some constant $c_k > 0$.

The bound for $\alpha(\mathcal{H})$ can be improved if we forbid some family $\mathcal{F}$ of hypergraphs in $\mathcal{H}$. For $\ell \geq 2$ a (Berge) cycle of length ℓ in $\mathcal{H}$ is a collection of ℓ distinct edges $E_1, \ldots, E_\ell \in \mathcal{H}$ such that there exists ℓ distinct vertices $v_1, \ldots, v_\ell$ with $v_i \in E_i \cap E_{i+1}$ for $i \in [\ell - 1]$ and $v_\ell \in E_\ell \cap E_1$. A seminal result of Ajtai, Komlós, Pintz, Spencer, and Szemerédi [4] states that for every n-vertex k-graph $\mathcal{H}$ with average degree d that contains no cycles of length in $\{2, 3, 4\}$, there exists a constant $c'_k > 0$ such that

$$\alpha(\mathcal{H}) \geq c'_k \frac{n}{d^{1/(k-1)}} (\log d)^{1/(k-1)}. \tag{2.2}$$

Moreover, this is tight apart from c'_k.

Spencer [200] conjectured and Duke, Lefmann, and Rödl [53] proved that the same conclusion holds even if $\mathcal{H}$ just contains no cycles of length 2. Their result was further extended by Rödl and Šiňajová [222] to the larger family of (n, k, ℓ)-systems.

Closely related to the (n, k, ℓ)-systems is a family called (n, k, ℓ)-omitting systems, where a family $\mathcal{H} \subset \binom{[n]}{k}$ is an (n, k, ℓ)-omitting system if $|A \cap B| \neq \ell$ for every pair of edges $\{A, B\} \subset \mathcal{H}$.

One important difference between (n, k, ℓ)-systems and (n, k, ℓ)-omitting systems is their maximum sizes. By definition, every set of ℓ vertices in an (n, k, ℓ)-system is contained in at most one edge, thus every (n, k, ℓ)-system has size at most $\binom{n}{\ell}/\binom{k}{\ell} = O\left(n^{\ell}\right)$. However, this is not true for (n, k, ℓ)-omitting systems. Indeed, the following result of Frankl and Füredi [99] shows that the maximum size of an (n, k, ℓ)-omitting system can be much larger than that of an (n, k, ℓ)-system when $k > 2\ell + 1$.

Let $k > \ell \geq 1$ and $\lambda \geq 1$ be integers. Observe that an n-vertex k-graph is an (n, k, ℓ)-omitting system iff it is $S_2(\ell)$-free, and is an (n, k, ℓ)-system iff it is $\{S_2(\ell), \ldots, S_2(k-1)\}$-free.

Theorem 2.6.1 (Frankl–Füredi [99])**.** *Let $k > \ell \geq 1$ and $\lambda > 1$ be fixed integers and $\mathcal{H}$ be an $S_\lambda(\ell)$-free k-graph on n vertices. Then $|\mathcal{H}| = O\left(n^{\max\{\ell, k-\ell-1\}}\right)$. Moreover, the bound is tight up to a constant multiplicative factor.*

In Chapter 9, we study the independence number of (n, k, ℓ)-omitting systems. Our results for (n, k, ℓ)-omitting systems are divided into two parts. For $k \leq 2\ell + 1$, we believe that the behavior is similar to that of (n, k, ℓ)-systems and prove a nontrivial lower bound for the first open case $\ell = k - 2$. For $k > 2\ell + 1$ we give new lower and upper bounds which show that the minimum independence number of (n, k, ℓ)-omitting systems has a very different behavior than for (n, k, ℓ)-systems.

Explicit constructions of (n, r, s)-systems with certain properties are very useful in theoretical computer science. For example, in the seminal work of Nisan and Wigderson [202], dense

(n, r, s)-systems are used to construct pseudorandom generators (PRGs) (see also [240; 213] for more applications). More recently, explicit constructions of (n, r, s)-systems with small independence number were used to construct extractors for adversarial sources [36; 35].

Rödl and Šiňajová's proof of the existence of an (n, r, s)-system with small independence number uses the Lovász local lemma, and hence it does not provide an explicit way to construct them. Perhaps the first explicit construction of an $(n, 3, 2)$-system (also called a Steiner triple system) with independence number $O(n^{1-\epsilon})$ for some absolute constant $\epsilon > 0$ is due to Chattopadhyay, Goodman, Goyal, and Li [36]. Their proof uses results about cap sets (see [45; 55]).

Theorem 2.6.2 (Chattopadhyay–Goodman–Goyal–Li [36])**.** *There exists a constant $C \geq 1$ such that for every integer $n \geq 3$ there exists an explicit construction of an $(n, 3, 2)$-system with independence number at most $Cn^{0.9228}$.*

Later, using results about linear codes [125; 29] and Sidorenko's recent bounds on the size of sets in $\mathbb{Z}_2^n$ containing no r elements that sum to zero [229; 230], Chattopadhyay and Goodman [35] extended Theorem 2.6.2 to all integers $r > s \geq 2$ with $s \geq \lceil r/2 \rceil$.

Theorem 2.6.3 (Chattopadhyay–Goodman [35])**.** *There exists a constant $C \geq 1$ such that for every integer $s \geq 2$ and every even integer $r > s$ there exists an explicit construction of an (n, r, s)-system with independence number at most $Cr^4 n^{\frac{2(r-s)}{r}}$.*

For odd r they showed that there exists an explicit construction of an (n, r, s)-system with independence number at most $C(r+1)^4 n^{\frac{2(r+1-s)}{r+1}}$.

Our main results on this topic extend Theorem 2.6.2 for certain values of r and s in the range $s < \lceil r/2 \rceil$ which was not addressed by Theorem 2.6.3.

CHAPTER 3

THE FEASIBLE REGION OF HYPERGRAPHS

Previously published as X. Liu. New short proofs to some stability theorems. European J. Combin., 96:Paper No. 103350, 8, 2021, and X. Liu and D. Mubayi. The feasible region of hypergraphs. J. Comb. Theory, Ser. B,148:2359, 2021.

3.1 Introduction

In this chapter we consider the following question which combines the classical Kruskal–Katona [154; 132] Theorem and the hypergraphs Turán problem.

$$\text{If } \mathcal{H} \text{ is } \mathcal{F}\text{-free, what are the possible values of } |\mathcal{H}| \text{ for fixed } |\partial\mathcal{H}|? \qquad (*)$$

Note that for the case $\mathcal{F} = \emptyset$ the solution to problem $(*)$ is exactly the Kruskal–Katona Theorem. For the sake of simplicity we state the following technically simpler version due to Lovász.

Theorem 3.1.1 (see Lovász [175])**.** *Let $\mathcal{H}$ be an r-graph, and suppose that $|\partial\mathcal{H}| = \binom{z}{r-1}$ for some real number $z \geq r$. Then $|\mathcal{H}| \leq \binom{z}{r}$.*

If $\mathcal{F} \neq \emptyset$, then $(*)$ is closely related to the hypergraph Turán problem. In fact, $\mathrm{ex}(n, \mathcal{F})$ gives a universal upper bound for $|\mathcal{H}|$ no matter what $|\partial\mathcal{H}|$ is, and it is tight for some (at least one) values of $|\partial\mathcal{H}|$. However, the upper bound given by $\mathrm{ex}(n, \mathcal{F})$ gives us a rather limited picture of the relationship between the shadow and size of an $\mathcal{F}$-free hypergraph. Our objective in this chapter is to provide a much more detailed view of this relationship.

An analogous question has been studied extensively in extremal graph theory. For fixed graphs H_1 and H_2 and (large) graph G, the following problem is a cornerstone of extremal graph theory:

$$\text{What are the possible values of } \rho(H_2; G) \text{ if } \rho(H_1; G) \text{ is fixed?} \qquad (\star)$$

Even for $(H_1, H_2) = (K_2, K_t)$ with $t \geq 3$, question $(\star)$ is known to be highly nontrivial and was asymptotically solved for $t = 3$ by Razborov [216], $t = 4$ by Nikiforov [201], and for all t only recently by Reiher [217]. We refer the reader to [178; 25; 215] for the history of $(\star)$.

The main difficulty in $(\star)$ is to determine the lower bound for $\rho(H_2; G)$. However, it will be shown later that the main difficulty in $(*)$ is to determine the upper bound for $|\mathcal{H}|$. In order to state our results formally we need some definitions.

Definition 3.1.2 (Feasible region of hypergraphs)**.** *Fix $r \geq 3$.*

(a) Given an r-graph $\mathcal{H}$, its edge density is $d(\mathcal{H}) = |\mathcal{H}|/\binom{v(\mathcal{H})}{r}$ and its shadow density is $d(\partial\mathcal{H}) = |\partial\mathcal{H}|/\binom{v(\mathcal{H})}{r-1}$.

(b) An r-graph sequence $(\mathcal{H}_k)_{k=1}^{\infty}$ is good if $v(\mathcal{H}_k) \to \infty$ as $k \to \infty$ and both $\lim_{k\to\infty} d(\mathcal{H}_k)$ and $\lim_{k\to\infty} d(\partial\mathcal{H}_k)$ exist.

(c) Let $(\mathcal{H}_k)_{k=1}^{\infty}$ be a good sequence of $\mathcal{F}$-free r-graphs, and $(x, y) \in [0,1]^2$. Then $(\mathcal{H}_k)_{k=1}^{\infty}$ realizes (x, y) if $\lim_{k\to\infty} d(\partial\mathcal{H}_k) = x$ and $\lim_{k\to\infty} d(\mathcal{H}_k) = y$.

(d) The feasible region $\Omega(\mathcal{F})$ of $\mathcal{F}$ is the collection of all points $(x, y) \in [0,1]^2$ that can be realized by a good sequence of $\mathcal{F}$-free r-graphs.

As mentioned earlier, the upper bound given by $\mathrm{ex}(n, \mathcal{F})$ gives us a rather limited picture of $\Omega(\mathcal{F})$, since it only determines

$$\sup\{y\colon \exists x \in [0,1] \text{ such that } (x, y) \in \Omega(\mathcal{F})\}.$$

As indicated by $(*)$, we study $\Omega(\mathcal{F})$. Our results are of three flavors.

- We prove some general results about the shape of $\Omega(\mathcal{F})$. Our main results here are Theorems 3.1.11 and 3.1.12 which state that the boundary of $\Omega(\mathcal{F})$ is completely determined by a left-continuous almost everywhere differentiable function $g(\mathcal{F})$ with at most countably many jump discontinuities, and give examples showing that $g(\mathcal{F})$ can indeed be discontinuous.

- We study $\Omega(\mathcal{F})$ for some specific choices of $\mathcal{F}$ for which $\mathrm{ex}(n, \mathcal{F})$ has been investigated by many researchers. We focus on two specific families: cancellative hypergraphs and hypergraphs without expansions of cliques. Our results, which go beyond determining just the Turán density, are summarized in Corollaries 3.1.18 and 3.1.22 (see Figure 6 and Figure 7).

- We analyze the structure of $\mathcal{F}$-free hypergraphs $\mathcal{H}$ whose shadow density and edge density are close to the boundary of $\Omega(\mathcal{F})$ for cancellative hypergraphs (Theorem 3.1.26) and hypergraphs without expansions of cliques, which extends the classical stability theorems proved for $\mathcal{F}$-free hypergraphs. In particular, for cancellative 3-graphs we prove a stability theorem that connects Steiner triple systems with cancellative 3-graphs, and moreover, using this stability theorem we show that the function $g(\mathcal{F})$ for cancellative 3-graphs has countably many local minima (Theorem 3.1.28).

Regarding our results on the shape of $\Omega(\mathcal{F})$, there are (at least) two previous works of a similar flavor: Razborov [216] determined the closure of the set of points defined by the homomorphism density of the edge and the triangle in finite graphs (and showed that the boundary is almost everywhere differentiable) and Hatami–Norine [120] constructed examples which show that the restrictions of the boundary to certain hyperplanes of the region defined by the homomorphism densities of a list of given graphs can have nowhere differentiable parts.

Our work can be viewed as a continuation of a long line of research in asymptotic extremal combinatorics perhaps beginning with the seminal work of Erdős–Lovász–Spencer [70] and continuing today in different guises such as the graph limits paradigm of Lovász [177] or the method of Flag algebras of Razborov [215].

3.1.1 General results about $\Omega(\mathcal{F})$

In this section we state some general results about feasible regions.

Proposition 3.1.3. *The region $\Omega(\mathcal{F})$ is closed for all $r \geq 3$ and all (possibly infinite) families $\mathcal{F}$ of r-graphs.*

Definition 3.1.4 (Projection of the feasible region)**.** *The projection of $\Omega(\mathcal{F})$ on the x-axis is*

$$\mathrm{proj}\Omega(\mathcal{F}) = \{x \colon \exists y \in [0,1] \text{ such that } (x,y) \in \Omega(\mathcal{F})\}\,.$$

Note that it is not necessarily true that $\mathrm{proj}\Omega(\mathcal{F}) = [0,1]$ in general. Later we will present an example of $\mathcal{F}$, which shows $\mathrm{proj}\Omega(\mathcal{F}) = [0,(\ell)_{r-1}/\ell^{r-1}]$ for $\ell \geq 3$. On the other hand, by removing edges one by one from $\mathcal{H}$ one can reduce the edge density of $\partial\mathcal{H}$ continuously (in the limit sense) to 0. This yields the following observation.

Observation 3.1.5. *For every family $\mathcal{F}$ of r-graphs with $r \geq 3$ there exists $\hat{c} \in [0,1]$ such that* $\mathrm{proj}\Omega(\mathcal{F}) = [0,\hat{c}]$.

Proposition 3.1.3 enables us to define the following function.

Definition 3.1.6 (Boundary of the feasible region)**.** *Given a family $\mathcal{F}$ of r-graphs with $r \geq 3$, let $g(\mathcal{F}) : \mathrm{proj}\Omega(\mathcal{F}) \to [0,1]$ be defined by*

$$g(\mathcal{F})(x) = \max\left\{y\colon (x,y) \in \Omega(\mathcal{F})\right\},$$

for all $x \in \mathrm{proj}\Omega(\mathcal{F})$.

Here we abuse notation by writing $g(\mathcal{F}, x)$ for $g(\mathcal{F})(x)$. Our next result shows that $\Omega(\mathcal{F})$ is determined by $\mathrm{proj}\Omega(\mathcal{F})$ and $g(\mathcal{F})$.

Proposition 3.1.7. *Let $r \geq 3$ and let $\mathcal{F}$ be a family of r-graphs. If $(x_0, y_0) \in \Omega(\mathcal{F})$, then $(x_0, y) \in \Omega(\mathcal{F})$ for all $y \in [0, y_0]$.*

Combining the Kruskal–Katona theorem with some further observations yields the following universal upper bound for $g(\mathcal{F}, x)$.

Proposition 3.1.8. *Let $r \geq 3$ and $\mathcal{F}$ be a family of r-graphs. Then $g(\mathcal{F}, x) \leq x^{r/(r-1)}$ for all $x \in \mathrm{proj}\Omega(\mathcal{F})$ (see Figure 1). In particular, $\mathrm{proj}\Omega(\emptyset) = [0,1]$ and $g(\emptyset, x) = x^{r/(r-1)}$ for all $x \in [0,1]$.*

In [120], Hatami and Norin considered the region defined by the homomorphism densities of a list of given graphs, which is a more general version of ($\star$) (that generalizes ($\star$) from two graphs H_1, H_2 to more graphs). They constructed examples which show that the restrictions of the boundary to certain hyperplanes can have nowhere differential parts. However, we will show in the next result that $g(\mathcal{F})$ is well-behaved.

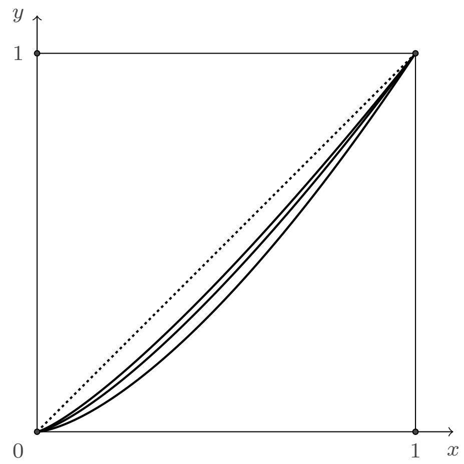


Figure 1. Upper bounds for $g(\mathcal{F}, x)$ when $r = 3, 4, 5$.

Definition 3.1.9 (Left/right continuity)**.** *Let $f : \mathbb{R} \to \mathbb{R}$. Then f is left-continuous (resp. right-continuous) at x if for any $\epsilon > 0$ there exists $\delta > 0$ such that $|f(x') - f(x)| < \epsilon$ for all $x' \in (x - \delta, x)$ (resp. $|f(x') - f(x)| < \epsilon$ for all $x' \in (x, x + \delta)$). If f is left-continuous (resp. right-continuous) at all $x \in \mathbb{R}$, then we say f is left-continuous (resp. right-continuous).*

Definition 3.1.10 (Types of discontinuities)**.** *Let $f : \mathbb{R} \to \mathbb{R}$ and $x \in \mathbb{R}$ be a discontinuity of f. If $\lim_{x \to x^-} f(x)$ and $\lim_{x \to x^+} f(x)$ exist, then f is said to have the discontinuity of the first kind at x. Otherwise, the discontinuity is said to be of the second kind. Furthermore, suppose that x is a discontinuity of the first kind of f. Then x is a removable discontinuity if $\lim_{x \to x^-} f(x) = \lim_{x \to x^+} f(x)$. Otherwise, x is a jump discontinuity.*

Theorem 3.1.11. *For any $r \geq 3$ and any family $\mathcal{F}$ of r-graphs, $g(\mathcal{F})$ is left-continuous, has at most countably many jump discontinuities, and is almost everywhere differentiable.*

Furthermore, the next result shows that $g(\mathcal{F})$ can indeed be discontinuous.

Theorem 3.1.12. *There exists a family $\mathcal{D}$ of 3-graphs with* $\mathrm{proj}\Omega(\mathcal{D}) = [0,1]$ *and* $g(\mathcal{D}, 2/3) = 2/9$*, but there exists an absolute constant* $\delta_0 > 0$ *such that* $g(\mathcal{D}, 2/3+\epsilon) < 2/9 - \delta_0$ *for all* $\epsilon \in (0, 10^{-8})$.

Actually, Theorem 3.1.12 can be extended to $r \geq 4$, and the condition that $\epsilon < 10^{-8}$ is not necessary (for all $r \geq 3$). The proof for these extensions can be found in the arXiv version of [167].

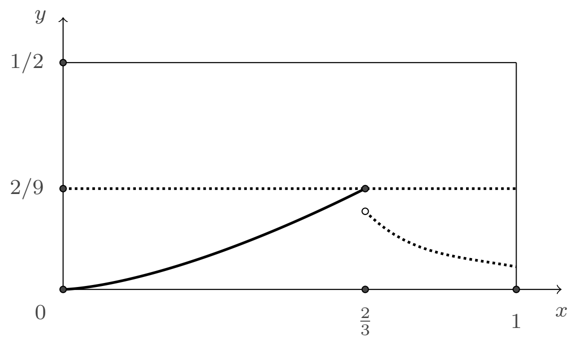


Figure 2. The function $g(\mathcal{D})$ is discontinuous at $x = 2/3$.

3.1.2 Cancellative hypergraphs

In this section we consider the feasible region of cancellative hypergraphs, which is perhaps the first example of an extremal hypergraph problem that was well understood. Our results are summarized in Corollary 3.1.18 stated at the end of this section.

Definition 3.1.13. *Let $\mathcal{T}_r$ be the collection of all r-graphs on at most $2r-1$ vertices with 3 edges A, B, C such that $A \triangle B \subset C$. An r-graph is cancellative iff it is $\mathcal{T}_r$-free.*

For $r = 2$ the family $\mathcal{T}_2$ comprises only one graph K_3. For $r = 3$ the family $\mathcal{T}_3$ comprises two hypergraphs K_4^{3-} and F_5, where K_4^{3-} is the 3-graph on 4 vertices with exactly 3 edges, and F_5 is the 3-graph on 5 vertices with edge set $\{123, 124, 345\}$. Let

$$t_r(n, \ell) = |T_r(n, \ell)| \approx \binom{\ell}{r} \left(\frac{n}{\ell}\right)^r$$

denote the size of the Turán r-graph $T_r(n, \ell)$. Motivated by Mantel's theorem, in the 1960's, Katona raised the question of determining the maximum size of a $\mathcal{T}_3$-free 3-graph and conjectured that the maximum size of a cancellative 3-graph is achieved by $T_3(n, 3)$. Katona's conjectured was proved by Bollobás in [24].

Theorem 3.1.14 (Bollobás [24])**.** *A $\mathcal{T}_3$-free 3-graph on n vertices has at most $t_3(n, 3)$ edges. Moreover, $T_3(n, 3)$ is the unique extremal construction.*

Consequently, $g(\mathcal{T}_3, x) \leq 2/9$ for all $x \in \text{proj}\Omega(\mathcal{T}_3)$. Later, Keevash and Mubayi [139] proved a stability theorem for $\mathcal{T}_3$-free hypergraphs.

Theorem 3.1.15 (Stability; Keevash–Mubayi [139]). *For every $\delta > 0$ there exists $\epsilon > 0$ and n_0 such that the following holds for all $n \geq n_0$. Every n-vertex $\mathcal{T}_3$-free 3-graph $\mathcal{H}$ with at least $(1-\epsilon)t_3(n,3)$ edges has a partition of its vertex set as $V_1 \cup V_2 \cup V_3$ such that all but at most δn^3 edges of $\mathcal{H}$ has one vertex in each V_i.*

In Section 3.2 we will present a new short proof to both the exact and the stability result for $\mathcal{T}_3$-free 3-graphs. In fact, our proof not only gives the exact value of $\mathrm{ex}(n,\mathcal{T}_3)$ but also shows a relation between $|\partial\mathcal{H}|$ and $|\mathcal{H}|$ for a $\mathcal{T}_3$-free 3-graph $\mathcal{H}$ on n-vertices. More specifically, it shows that

$$4\left(\frac{3|\mathcal{H}|/|\partial\mathcal{H}|}{n-3|\mathcal{H}|/|\partial\mathcal{H}|}\right)^2|\partial\mathcal{H}| \leq n^2 - 2|\partial\mathcal{H}|, \tag{3.1}$$

and it follows that (see Figure 3)

$$g(\mathcal{T}_3, x) \leq \frac{\sqrt{2(1-x)x^3}+x^2-x}{3x-1}, \text{ for all } x \in \mathrm{proj}\Omega(\mathcal{T}_3). \tag{3.2}$$

This serves as a motivation for us to study the feasible region of hypergraphs.

Our next result concerns cancellative r-graphs for $r \geq 3$, and improves the bound in Proposition 3.1.8 as well as that in Equation 3.2 for $x \in [0, 2/3]$.

Theorem 3.1.16. *Let $r \geq 3$ and $x \in \mathrm{proj}\Omega(\mathcal{T}_r)$. Then*

$$g(\mathcal{T}_r, x) \leq \left(\frac{x^r}{r!}\right)^{\frac{1}{r-1}}.$$

Moreover, equality holds for all $x \in [0, (r-1)!/r^{r-2}]$ (see Figure 4).

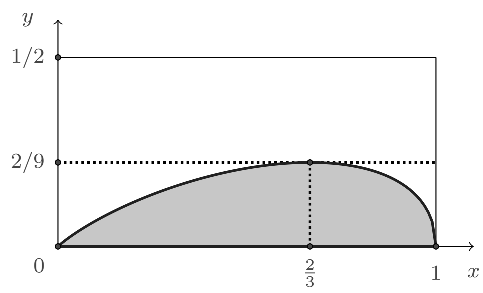


Figure 3. $\Omega(\mathcal{T}_3)$ is contained in the dark area above.

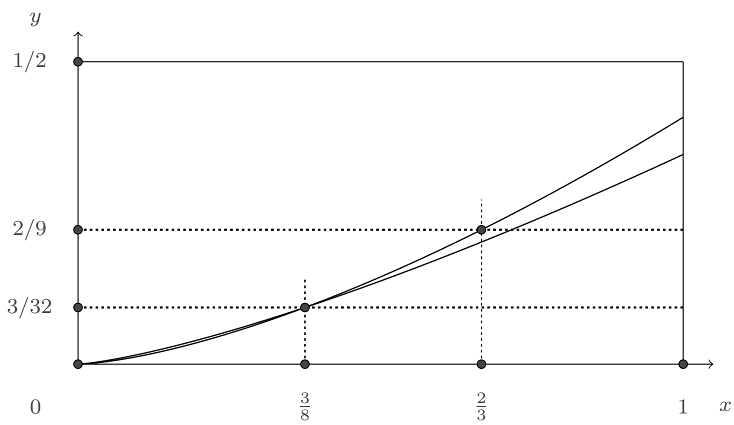


Figure 4. Upper bounds for $g(\mathcal{T}_r, x)$ when $r = 3, 4$.

For $r = 3$, the bound given by Theorem 3.1.16 is not tight for any $x \in (2/3, 1]$ according to Bollobás' theorem [24]. Our next result will present an improved bound for $g(\mathcal{T}_3, x)$ for $x \in (2/3, 1]$.

Theorem 3.1.17. *The inequality* $g(\mathcal{T}_3, x) \leq x(1-x)$ *holds for all* $x \in [0, 1]$. *Moreover,* $g(\mathcal{T}_3, (k-1)/k) = (k-1)/k^2$ *when* $k \equiv 1$ *or* 3 *(mod* 6*) (see Figure 5).*

Christian Reiher observed that the function $x(1-x)$ in Theorem 3.1.17 can be replaced by a piecewise linear function that always lies below $x(1-x)$ (see Section ? for details). The lower bound for $g(\mathcal{T}_3, (k-1)/k)$ when $k \equiv 1$ or 3 (mod 6) comes from the balanced blow up of Steiner triple systems on k vertices, this will be explained in more detail in Section ?. Combining

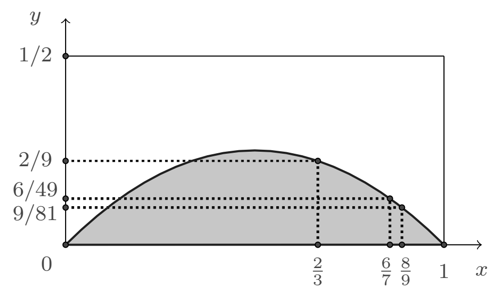


Figure 5. $\Omega(\mathcal{T}_3)$ is contained in the dark area.

Theorems 3.1.16 and 3.1.17 yields the following result for $g(\mathcal{T}_3, x)$, which provides a rather comprehensive picture of $\Omega(\mathcal{T}_3)$ (see Figure 6).

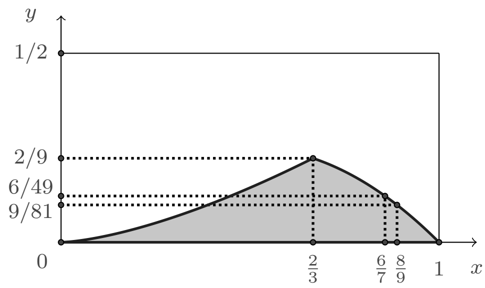


Figure 6. $\Omega(\mathcal{T}_3)$ is contained in the dark area.

Corollary 3.1.18. *We have* $g(\mathcal{T}_3, x) = x^{3/2}/\sqrt{6}$ *for all* $x \in [0, 2/3]$*, and* $g(\mathcal{T}_3, x) \leq x(1-x)$ *for all* $x \in (2/3, 1]$*. Moreover,* $g(\mathcal{T}_3, (k-1)/k) = (k-1)/k^2$ *for all integers* $k \equiv 1$ *or* 3 *(mod* 6*).*

3.1.3 Hypergraphs without an expansion of a large clique

In this section we consider the feasible region of hypergraphs without expansion of cliques. These hypergraphs were introduced by Mubayi in [191] as a way to generalize Turán's theorem to hypergraphs. Another reason for their importance is that they provide the first (and still the only) explicitly defined examples which yield an infinite family of numbers realizable as Turán densities for hypergraphs.

Let $\mathcal{K}^r_{\ell+1}$ be the collection of all r-graphs F with at most $\binom{\ell+1}{2}$ edges such that for some $(\ell+1)$-set S, which will be called the core of F, every pair $\{u, v\} \subset S$ is covered by an edge in F. Let the r-graph $H^r_{\ell+1}$ be obtained from the complete graph K_ℓ by adding $r-2$ new vertices into each edge. The graph $H^r_{\ell+1}$ is called the expansion of K_ℓ. It is an easy observation that $H^r_{\ell+1} \in \mathcal{K}^r_{\ell+1}$.

Mubayi introduced the notion of $\mathcal{K}_{\ell+1}^r$ and $H_{\ell+1}^r$ in [191] and proved both the exact and stability results for $\mathcal{K}_{\ell+1}^r$-free r-graphs.

Theorem 3.1.19 (Mubayi [191])**.** *Let $n \geq 1$ and $\ell \geq r \geq 2$ be integers. Then* $\mathrm{ex}(n, \mathcal{K}_{\ell+1}^{(r)}) = t_r(n,\ell)$*. Moreover, $T_r(n,\ell)$ is the unique extremal construction on n vertices.*

Theorem 3.1.20 (Stability; Mubayi [191])**.** *Fix $\ell \geq r \geq 2$. For every $\delta > 0$ there exists an $\epsilon > 0$ and an n_0 such that the following holds for all $n \geq n_0$. Let $\mathcal{H}$ be an n-vertex $\mathcal{K}_{\ell+1}^r$-free r-graph with at least $(1-\epsilon)t_r(n,\ell)$ edges. Then the vertex set of $\mathcal{H}$ has a partition $V_1 \cup \cdots \cup V_\ell$ such that all but at most δn^r edges in $\mathcal{H}$ have at most one vertex in each V_i.*

In [208], Pikhurko improved the result in [191] and proved that if n is sufficiently large then $\mathrm{ex}(n, H_{\ell+1}^r) = t_r(n,\ell)$ and $T_r(n,\ell)$ is the unique $H_{\ell+1}^r$-free r-graph on n vertices with exactly $t_r(n,\ell)$ edges.

In order to state our result, we need to extend the definition of shadows. Let $\mathcal{H}$ be an r-graph and $S \subset V(\mathcal{H})$. For $i \leq 0$ we extend the definition of the i-th shadow $\partial_i \mathcal{H}$ as follows:

$$\partial_i \mathcal{H} = \left\{ A \in \binom{V(\mathcal{H})}{r-i} : \mathcal{H}[A] \text{ is a complete } r\text{-graph} \right\}. \tag{3.3}$$

In particular, $\partial_1 \mathcal{H} = \partial \mathcal{H}$ and $\partial_0 \mathcal{H} = \mathcal{H}$. By definition, $\partial_{i+1}\mathcal{H} = \partial\left(\partial_i \mathcal{H}\right)$ for all $0 \leq i \leq r-2$, and $\partial\left(\partial_i \mathcal{H}\right) \subset \partial_{i+1}\mathcal{H}$ for all $i \leq -1$.

Our first result here relates the sizes of different shadows of a $\mathcal{K}_{\ell+1}^r$-free r-graph $\mathcal{H}$. This generalizes an important result of Fisher and Ryan [83] from graphs to hypergraphs.

Theorem 3.1.21. *Let $\ell \geq r \geq 2$ and $\mathcal{H}$ be a $\mathcal{K}_{\ell+1}^r$-free r-graph. Then*

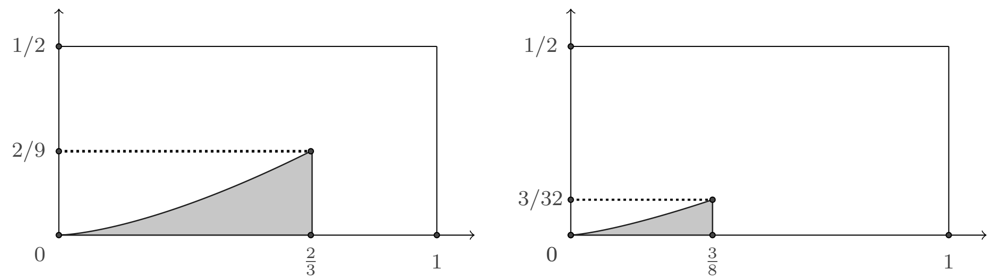
$$\left(\frac{|\partial_{r-\ell}\mathcal{H}|}{\binom{\ell}{\ell}}\right)^{\frac{1}{\ell}} \leq \cdots \leq \left(\frac{|\partial_{-1}\mathcal{H}|}{\binom{\ell}{r+1}}\right)^{\frac{1}{r+1}} \leq \left(\frac{|\mathcal{H}|}{\binom{\ell}{r}}\right)^{\frac{1}{r}} \leq \left(\frac{|\partial_{1}\mathcal{H}|}{\binom{\ell}{r-1}}\right)^{\frac{1}{r-1}} \leq \cdots \leq \left(\frac{|\partial_{r-1}\mathcal{H}|}{\binom{\ell}{1}}\right)^{\frac{1}{1}}.$$

Using Theorem 3.1.21 we are able to determine $g(\mathcal{K}_{\ell+1}^r)$ completely via the following result. We will use $(\ell)_r$ to denote $\ell(\ell-1)\cdots(\ell-r+1)$.

Corollary 3.1.22. *Let $\ell \geq r \geq 3$. Then* $\mathrm{proj}\Omega(\mathcal{K}_{\ell+1}^r) = [0, (\ell)_{r-1}/\ell^{r-1}]$ *and*

$$g(\mathcal{K}_{\ell+1}^r, x) = (\ell - r + 1)\left(\frac{x^r}{(\ell)_r}\right)^{\frac{1}{r-1}}$$

for all $x \in [0, (\ell)_{r-1}/\ell^{r-1}]$ (see Figure 7).



(a) $\ell = 3, r = 3$.

(b) $\ell = 4, r = 4$.

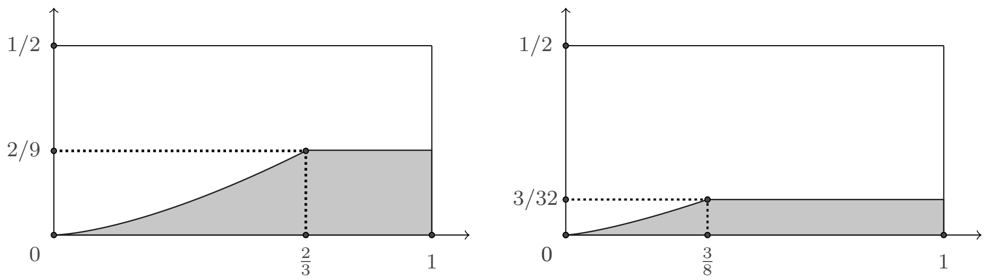
Figure 7. The region $\Omega(\mathcal{K}_{\ell+1}^r)$ for $(\ell, r) = (3, 3)$ and $(\ell, r) = (4, 4)$.

Determining $\Omega(H^r_{\ell+1})$ is much more difficult than determining $\Omega(\mathcal{K}^r_{\ell+1})$ because the shadow density of an $H^r_{\ell+1}$-free r-graph can be greater than $(\ell)_{r-1}/\ell^{r-1}$. An r-graph $\mathcal{S}$ is called a star if all edges in $\mathcal{S}$ contain a fixed vertex, which is called the center of $\mathcal{S}$. It is easy to see that a star does not contain $H^r_{\ell+1}$ as a subgraph, and the shadow density of a star can be arbitrarily close to 1. Still, we are able to determine $g(H^r_{\ell+1}, x)$ for all $x \in [0, (\ell)_{r-1}/\ell^{r-1}]$.

Theorem 3.1.23. *Let* $\ell \geq r \geq 3$. *Then* $\mathrm{proj}\Omega(H^r_{\ell+1}) = [0, 1]$ *and*

$$g(H^r_{\ell+1}, x) = (\ell - r + 1)\left(\frac{x^r}{(\ell)_r}\right)^{\frac{1}{r-1}}$$

for all $x \in [0, (\ell)_{r-1}/\ell^{r-1}]$ *(see Figure 8).*



(a) $\ell = 3, r = 3$. (b) $\ell = 4, r = 4$.

Figure 8. $\Omega(H^r_{\ell+1})$ for $(\ell, r) = (3, 3), (4, 4)$ are contained in the dark areas above, respectively.

3.1.4 Stability near the boundary

For a positive integer k a k-vertex Steiner triple system (STS) is a 3-graph on k vertices such that every pair of vertices is contained in exactly one edge. It is known that a k-vertex STS exists if and only if $k \in 6\mathbb{N} + \{1, 3\}$ (e.g. see [243]), where $6\mathbb{N} + \{1, 3\}$ is the set of integers that congruent to 1 or 3 modulo 6.

Let STS(k) denote the family of all Steiner triple systems on k vertices up to isomorphism. In particular, STS(3) contains only the 3-graph K_3^3, STS(6) contains only the Fano plane (see Figure 10), and STS(9) contains only the affine plane of order 3 (see Figure 11). For $k \in 6\mathbb{N} + \{1, 3\}$ let $s_k = |\mathrm{STS}(k)|$. Then $s_3 = s_7 = s_9 = 1$, $s_{13} = 2$ (see [43]), $s_{15} = 80$ (see [118]), and Keevash (see [137] and also [246; 54; 81]) proved that $s_k = \big(k/e^2 + o(k)\big)^{k^2/6}$.

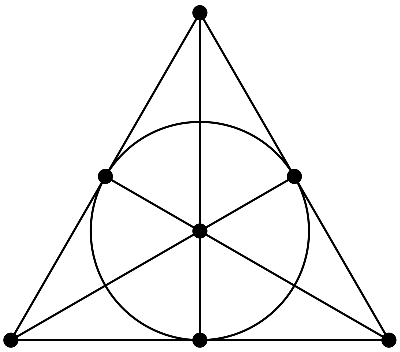


Figure 10. The Fano plane.

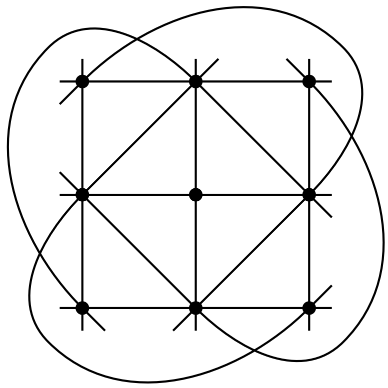


Figure 11. The affine plane of order 3.

Definition 3.1.24. *An r-graph $\mathcal{H}$ is a blowup of another r-graph $\mathcal{G}$ if there exists a map $\psi\colon V(\mathcal{H}) \to V(\mathcal{G})$ so that $\psi(E) \in \mathcal{G}$ if and only if $E \in \mathcal{H}$, and we say $\mathcal{H}$ is $\mathcal{G}$-colorable if there exists a map $\phi\colon V(\mathcal{H}) \to V(\mathcal{G})$ so that $\phi(E) \in \mathcal{G}$ for all $E \in \mathcal{H}$.*

Note that $\mathcal{H}$ is $\mathcal{G}$-colorable if and only if $\mathcal{H}$ occurs as a subgraph in some blowup of $\mathcal{G}$. The following easy observation relates cancellative 3-graphs to the Steiner triple systems.

Observation 3.1.25. *Suppose that $\mathcal{H}$ is a blowup of a Steiner triple system. Then $\mathcal{H}$ is cancellative. Moreover, if $\mathcal{H}$ is a 3-graph on n vertices that is a balanced blowup of a Steiner triple system on k vertices, then $|\partial\mathcal{H}| \sim \frac{k-1}{k}\frac{n^2}{2}$ and $|\mathcal{H}| \sim \frac{1}{6}\frac{k-1}{k^2}n^3$.*

The following stability theorem connects the cancelltive 3-graph with the Steiner triple systems.

Theorem 3.1.26. *Let $k \in 6\mathbb{N} + \{1,3\}$ and $k \geq 3$. For every $\delta > 0$ there exists $\epsilon > 0$ and n_0 such that the following holds for all $n \geq n_0$. Suppose that $\mathcal{H}$ is cancellative 3-graph on n vertices that satisfies*

$$\left(d(\partial\mathcal{H}) - \frac{k-1}{k}\right)^2 + \left(d(\mathcal{H}) - \frac{k-1}{k^2}\right)^2 \leq \epsilon.$$

Then $\mathcal{H}$ is $\mathcal{S}$-colorable for some $\mathcal{S} \in \mathrm{STS}(k)$ after removing at most δn^3 edges.

Remarks.

- Roughly speaking, Theorem 3.1.26 says if the shadow density and the edge density of a cancellative 3-graph $\mathcal{H}$ are close (in the sense of Euclidean distance in $\mathbb{R}^2$) to $\frac{k-1}{k}$ and $\frac{k-1}{k^2}$

respectively, then the structure of $\mathcal{H}$ is close to the balanced blowup of a Steiner triple system on k vertices.

- Theorem 3.1.26 contains Keevash and Mubayi's result (Theorem 3.1.15) as a special case (i.e. $k = 3$) because by Corollary 3.1.18, if $d(\mathcal{H})$ is close to $2/9$, then $d(\partial\mathcal{H})$ must be close to $2/3$. So by Theorem 3.1.26, $\mathcal{H}$ is structurally close to the balanced blowup of K_3^3, which is $T_3(n, 3)$.
- Our proof shows that the relation $\delta = O(\epsilon^{1/2})$ is sufficient for Theorem 3.1.26.

A more detailed analysis of the proof of Theorem 3.1.26 yields the following exact result.

Theorem 3.1.27. *Let $k \in 6\mathbb{N} + \{1, 3\}$, $k \geq 3$, and n be a sufficiently large integer. Suppose that $\mathcal{H}$ is a cancellative 3-graph on n vertices with $|\partial\mathcal{H}| = t_2(n, k)$. Then $|\mathcal{H}| \leq s(n, k)$, where*

$$s(n,k) = \max\left\{|\mathcal{G}| \colon \mathcal{G} \textit{ is a blowup of } \mathcal{S} \textit{ for some } \mathcal{S} \in \mathrm{STS}(k) \textit{ and } |V(\mathcal{G})| = n\right\}.$$

Moreover, equality holds only if $\mathcal{H}$ is a blowup of $\mathcal{S}$ for some $\mathcal{S} \in \mathrm{STS}(k)$.

As an application of Theorem 3.1.26 we show that the feasible region function $g(\mathcal{T}_3)$ has countably many local maxima. This is the first example showing that the feasible region function can have a local maximum.

Theorem 3.1.28. *Let $k \in 6\mathbb{N} + \{1,3\}$ and $k \geq 3$ be fixed. Then there exists an absolute constant $c > 0$ such that for every constant $\epsilon \leq c$ there exists another constant $\delta = \delta(\epsilon) > 0$ so that*

$$g(\mathcal{T}_3, (k-1)/k - \epsilon) \leq \frac{k-1}{k^2} - \delta \quad \text{and} \quad g(\mathcal{T}_3, (k-1)/k + \epsilon) \leq \frac{k-1}{k^2} - \delta.$$

In particular, the point $((k-1)/k, (k-1)/k^2)$ is a local maximum of $g(\mathcal{T}_3)$.

Remark. Our proof shows that a linear dependency between δ and ϵ is sufficient for Theorem 3.1.28.

3.2 Warm up

In this section we present new short proofs to some exact and stability theorems mentioned in the previous section.

3.2.1 Proofs of Theorems 3.1.14 and 3.1.15

In this section we prove Theorem 3.1.14 for the case 3 divides n and Theorem 3.1.15. Let us first present some preliminary definitions and results.

The following inequality will be used intensively in our proofs.

Lemma 3.2.1 (Jensen's inequality [129])**.** *Suppose that $f\colon I \to \mathbb{R}$ is a convex function on some interval I and $x_1, \ldots, x_n \in I$. Then*

$$\sum_{i\in[n]} f(x_i) \geq n \cdot f\left(\frac{\sum_{i\in[n]} x_i}{n}\right).$$

Let $\mathcal{H}$ be an r-graph on $[n]$. We abuse the use of notation in this section by calling a set $I \subset [n]$ independent if every edge in $\mathcal{H}$ contains at most one vertex in I.[1] For every nonempty set $S \subset [n]$ define the link $L_{\mathcal{H}}(S)$ of S in $\mathcal{H}$ as

$$L_{\mathcal{H}}(S) = \left\{A \in \partial\mathcal{H}\colon A \cup \{s\} \in \mathcal{H} \text{ for all } s \in S\right\},$$

and we will omit the subscript $\mathcal{H}$ if there is no cause of any ambiguity. For convenience we write $L(u)$ and $L(u, v)$ for $L(\{u\})$ and $L(\{u, v\})$, respectively.

---

[1] Note that this definition of independent set is stronger that the definition in Chapter 1.

For $T \in \partial\mathcal{H}$ the neighborhood of T in $\mathcal{H}$ is defined as

$$N_{\mathcal{H}}(T) = \{v \in V(\mathcal{H}) \colon T \cup \{v\} \in \mathcal{H}\},$$

and the degree of T is $d_{\mathcal{H}}(T) = |N_{\mathcal{H}}(T)|$. Also, we will omit the subscript $\mathcal{H}$ if there is no cause of any ambiguity. It follows from an easy double counting argument that

$$\sum_{T \in \partial\mathcal{H}} d(T) = r|\mathcal{H}|.$$

Given a graph G we define an auxiliary digraph $\vec{G}$ by letting $\vec{G} = \{(u,v) \colon \{u,v\} \in G\}$. Note that if $\{u,v\} \in G$, then the ordered pairs (u,v) and (v,u) are both contained in $\vec{G}$. So $|\vec{G}| = 2|G|$. For a set N we use N^2 to denote the cartesian product $N \times N$, which is the collection of all ordered pairs (u,v) with $u, v \in N$ (note that u and v can be the same vertex).

We have the following two lemmas concerning properties of cancellative 3-graphs.

Lemma 3.2.2. *Let $\mathcal{H}$ be a cancellative 3-graph and $v \in V(\mathcal{H})$ be a vertex. Then the link $L(v)$ is a triangle-free graph.*

Proof. Suppose that the set $\{x, y, z\}$ induces a copy of triangle in $L(v)$. Then the three edges $\{v,x,y\}, \{v,x,z\}, \{v,y,z\}$ are all contained in $\mathcal{H}$. This is a contradiction since

$$\{v,x,y\} \triangle \{v,x,z\} = \{y,z\} \subset \{v,y,z\}.$$

∎

Lemma 3.2.3. *Let $\mathcal{H}$ be a cancellative 3-graph, and $T \in \partial\mathcal{H}$. Then the set $N(T)$ is independent in $\mathcal{H}$.*

Proof. Let u, v be two vertices in $N(T)$, and let $A_1 = \{u\} \cup T$ and $A_2 = \{v\} \cup T$. Then the sets A_1 and A_2 are both contained in $\mathcal{H}$. Since $A_1 \triangle A_2 = \{u, v\}$ and by the assumption that there is no edge in $\mathcal{H}$ containing $\{u, v\}$, the set $N(T)$ is an independent set. ∎

In the proof of Theorem 3.1.15 we will need the following lemma, which is essentially the stability of triangle-free graphs. For the sake of completeness we include its proof here.

Lemma 3.2.4. *Let G be a triangle-free graph on $[n]$ with at least $(1-\epsilon)(n/2)^2$ edges. Then G contains two vertices v_1 and v_2 such that the sets $N_G(v_1)$ and $N_G(v_2)$ are disjoint and $|N_G(v_1)| + |N_G(v_2)| \geq (1-\epsilon)n$.*

Proof. Since G is triangle-free, the sets $N_G(u)$ and $N_G(v)$ are disjoint for all edges uv in G. So in order to prove this lemma it suffices to find an edge uv in G such that $d_G(u) + d_G(v) \geq (1-\epsilon)n$. Combining an easy counting argument with Jensen's inequality we obtain

$$\sum_{uv \in E(G)} (d_G(u) + d_G(v)) = \sum_{v \in V(G)} d_G^2(v) \geq \frac{\left(\sum_{v \in V(G)} d_G(v)\right)^2}{n} = \frac{4|G|^2}{n}.$$

It follows from an easy averaging argument that there exists an edge uv in G such that $d_G(u) + d_G(v) \geq 4|G|/n \geq (1-\epsilon)n$. ∎

The key step in our proofs of Theorems 3.1.14 and 3.1.15 is building a relation between $\mathcal{H}$ and $\partial\mathcal{H}$, which is Equation 3.4.

Proof of Theorem 3.1.14 for the case 3 divides n. Let $\mathcal{H}$ be a cancellative 3-graph on n vertices. Let us do a double counting on the number of ordered pairs (u,v) in $[n]^2 \setminus \overrightarrow{\partial\mathcal{H}}$. By Lemma 3.2.3, $N(T)^2 \subset [n]^2 \setminus \overrightarrow{\partial\mathcal{H}}$ for all $T \in \partial\mathcal{H}$. On the other hand, since every $\{u,v\} \subset [n]$ is contained in exactly $|L(u,v)|$ sets in $\{N(T) \colon T \in \partial\mathcal{H}\}$, we have

$$\sum_{T\in\partial\mathcal{H}} \sum_{(u,v)\in N(T)^2} \frac{1}{|L(u,v)|} = \sum_{\substack{(u,v)\in[n]^2\setminus\overrightarrow{\partial\mathcal{H}} \\ L(u,v)\neq\emptyset}} \sum_{\substack{T\in\partial\mathcal{H} \\ \{u,v\}\subset N(T)}} \frac{1}{|L(u,v)|} \leq n^2 - |\overrightarrow{\partial\mathcal{H}}|$$

$$= n^2 - 2|\partial\mathcal{H}|. \tag{3.4}$$

By Lemmas 3.2.2 and 3.2.3, $L(u,v)$ is a triangle-free graph on $[n] \setminus N(T)$ for all $(u,v) \in N(T)^2$. So, by Mantel's theorem, $|L(u,v)| \leq (n - d(T))^2/4$ for all $(u,v) \in N(T)^2$ and all $T \in \partial\mathcal{H}$. Therefore, it follows from Equation 3.4 that

$$n^2 - 2|\partial\mathcal{H}| \geq \sum_{T\in\partial\mathcal{H}} \sum_{(u,v)\in N(T)^2} \frac{1}{|L(u,v)|} \geq \sum_{T\in\partial\mathcal{H}} 4\left(\frac{d(T)}{n-d(T)}\right)^2.$$

Since the function $(x/(n-x))^2$ is convex on $(0,n)$, it follows from Jensen's inequality and $\sum_{T\in\partial\mathcal{H}} d(T) = 3|\mathcal{H}|$ that

$$4\left(\frac{3|\mathcal{H}|/|\partial\mathcal{H}|}{n - 3|\mathcal{H}|/|\partial\mathcal{H}|}\right)^2 |\partial\mathcal{H}| \leq n^2 - 2|\partial\mathcal{H}|.$$

Now let $z = \frac{3|\mathcal{H}|/|\partial\mathcal{H}|}{n-3|\mathcal{H}|/|\partial\mathcal{H}|}$. Then the inequality above implies that

$$|\partial\mathcal{H}| \leq \frac{n^2}{2(2z^2+1)}.$$

Substitute $|\mathcal{H}| = \frac{zn}{3(z+1)}|\partial\mathcal{H}|$ into the equation above we obtain that

$$|\mathcal{H}| \leq \frac{z}{6(z+1)(2z^2+1)}n^3.$$

Since the maximum of $\frac{z}{6(z+1)(2z^2+1)}$ is $1/27$, we obtain $|\mathcal{H}| \leq (n/3)^3$. This proves Theorem 3.1.14 for the case 3 divides n. ∎

Next we prove Theorem 3.1.15.

Proof of Theorem 3.1.15. Let $\delta > 0$ be a sufficiently small constant, $\epsilon = \delta/100$, and n be a sufficiently large integer. Let $\mathcal{H}$ be a cancellative 3-graph on $[n]$ with at least $(1-\epsilon)t_3(n,3) > (1-2\epsilon)(n/3)^3$ edges. First, we have the following claim.

Claim 3.2.5. *There exists $T \in \partial\mathcal{H}$ such that*

$$\sum_{(u,v)\in N(T)^2} |L(u,v)| \geq (1-100\epsilon)d^2(T)\left(\frac{n-d(T)}{2}\right)^2. \tag{3.5}$$

Proof. Suppose that Equation 3.5 is false for all $T \in \partial\mathcal{H}$. Since the function $1/x$ is convex on $(0, \infty)$, it follows from Jensen's inequality that

$$\sum_{(u,v)\in N(T)^2} \frac{1}{|L(u,v)|} \geq \frac{d^2(T)}{\sum_{(u,v)\in N(T)^2} |L(u,v)|/d^2(T)} > \frac{4d^2(T)}{(1-100\epsilon)\,(n-d(T))^2}. \tag{3.6}$$

Substitute Equation 3.6 into Equation 3.4 we obtain

$$\sum_{T\in\partial\mathcal{H}} \frac{4d^2(T)}{(1-100\epsilon)\,(n-d(T))^2} < n^2 - 2|\partial\mathcal{H}|.$$

Similar to the proof of Theorem 3.1.14, applying Jensen's inequality to the function $(x/(n-x))^2$ we obtain

$$n^2 - 2|\partial\mathcal{H}| \geq \frac{4}{1-100\epsilon}\left(\frac{3|H|/|\partial\mathcal{H}|}{n-3|\mathcal{H}|/|\partial\mathcal{H}|}\right)^2 |\partial\mathcal{H}| = \frac{4z^2}{1-100\epsilon}|\partial\mathcal{H}|.$$

Therefore,

$$|\partial\mathcal{H}| \leq \frac{n^2}{2\left(\frac{2z^2}{1-100\epsilon}+1\right)},$$

and hence

$$|\mathcal{H}| = \frac{zn}{3(z+1)}|\partial\mathcal{H}| \leq \frac{z}{6(z+1)\left(\frac{2z^2}{1-100\epsilon}+1\right)}n^3.$$

By assumption we have

$$\frac{3|\mathcal{H}|}{|\partial\mathcal{H}|} > \frac{3(1-2\epsilon)(n/3)^3}{n^2/2} \geq \frac{n}{9}.$$

Therefore, we may assume that $z > 1/8$. Then, it follows that $\frac{2z^2}{1-100\epsilon} + 1 > \frac{2z^2+1}{1-2\epsilon}$ (here we used the assumption that ϵ is sufficiently small, and in particular, $\epsilon < 1/100$). So

$$\frac{z}{6(z+1)\left(\frac{2z^2}{1-100\epsilon}+1\right)} < (1-2\epsilon)\frac{z}{6(z+1)\left(2z^2+1\right)} \leq \frac{1}{27}(1-2\epsilon),$$

which implies that $|\mathcal{H}| < (1-2\epsilon)\left(n/3\right)^3 < (1-\epsilon)t_3(n,3)$, a contradiction. This prove Claim 3.2.5. ∎

Now choose $T \in \partial\mathcal{H}$ such that Equation 3.5 holds for T. Let $V_1' = N(T)$. Then by the Pigeonhole principle, there exists $(u,v) \in N(T)^2$ such that

$$|L(u,v)| \geq \frac{\sum_{(u,v)\in N(T)^2}|L(u,v)|}{|N(T)|^2} \geq (1-100\epsilon)\left(\frac{n-d(T)}{2}\right)^2. \tag{3.7}$$

Let $L = L(u,v)$ and $U = [n] \setminus N(T)$. By Lemma 3.2.3, $N(T)$ is an independent set in $\mathcal{H}$, so L is a graph on the set U. Due to Lemma 3.2.2, L is triangle-free. So by Lemma 3.2.4 and Equation 3.7, there exist $x, y \in U$ such that $N_L(x)$ and $N_L(y)$ are disjoint and

$$|N_L(x)| + |N_L(y)| \geq (1-100\epsilon)(n-d(T)).$$

Let $V_2 = N_L(x)$ and $V_3 = N_L(y)$. Note that $N_L(x) = N(\{u, x\})$ and $N_L(y) = N(\{u, y\})$, so by Lemma 3.2.3, V_2 and V_3 are independent in H. Now we have three pairwise disjoint independent sets V_1', V_2 and V_3, and moreover,

$$|V_1'| + |V_2| + |V_3| \geq d(T) + (1 - 100\epsilon)(n - d(T)) > n - 100\epsilon n.$$

To finish the proof we let $V_1 = [n] \setminus (V_2 \cup V_3)$. Then the number of edges in $\mathcal{H}$ that have at least two vertices in some V_i is at most $\binom{100\epsilon n}{3} + \binom{100\epsilon n}{2}\binom{n}{1} + \binom{100\epsilon n}{1}\binom{n}{2} < 100\epsilon n^3 = \delta n^3$. This completes the proof of Theorem 3.1.15. ∎

3.2.2 Proofs of Theorems 3.1.19 and 3.1.20

In this section we prove Theorem 3.1.19 for ℓ divides n and Theorem 3.1.20. Our proofs are based on two results about $K_{\ell+1}$-free graphs, and the first one is a stability theorem for $K_{\ell+1}$-free graphs.

Theorem 3.2.6 (Füredi [110]). *Let $t \geq 0$ be an integer, and let G be an n-vertex $K_{\ell+1}$-free graph with $t_2(n, \ell) - t$ edges. Then G contains an ℓ-partite subgraph G' with at least $t_2(n, \ell) - 2t$ edges.*

The second result describes a relationship between the number of copies of K_{r_1} and K_{r_2} in a $K_{\ell+1}$-free graph, where r_1 and r_2 are two positive integers less than $\ell + 1$.

Theorem 3.2.7 (Fisher–Ryan [83])**.** *Let G be an n-vertex $K_{\ell+1}$-free graph. For every $i \in [\ell]$ let k_i denote the number of copies of K_i in G. Then*

$$\left(\frac{k_\ell}{\binom{\ell}{\ell}}\right)^{\frac{1}{\ell}} \leq \left(\frac{k_{\ell-1}}{\binom{\ell}{\ell-1}}\right)^{\frac{1}{\ell-1}} \leq \cdots \leq \left(\frac{k_2}{\binom{\ell}{2}}\right)^{\frac{1}{2}} \leq \left(\frac{k_1}{\binom{\ell}{1}}\right)^{\frac{1}{1}}.$$

Recall that for an r-graph $\mathcal{H}$ and an integer $1 \leq i \leq r-1$ the i-th shadow $\partial_i \mathcal{H}$ of $\mathcal{H}$ is

$$\partial_i \mathcal{H} = \left\{ A \in \binom{V(\mathcal{H})}{r-i} : \exists B \in \mathcal{H} \text{ such that } A \subset B \right\}.$$

In particular, $\partial_{r-2}\mathcal{H}$ is a graph on $V(\mathcal{H})$.

The following easy observation about $\mathcal{H}$ and $\partial_{r-2}\mathcal{H}$ is key to our proofs.

Observation 3.2.8. *(a) An r-graph $\mathcal{H}$ is $\mathcal{K}_{\ell+1}^{(r)}$-free iff $\partial_{r-2}\mathcal{H}$ is $K_{\ell+1}$-free.*

(b) The number of edges in $\mathcal{H}$ is at most the number of copies of K_r in $\partial_{r-2}\mathcal{H}$.

Now we prove Theorem 3.1.19 for the case ℓ divides n.

Proof of Theorem 3.1.19 for the case ℓ divides n. Let $\mathcal{H}$ be a $\mathcal{K}_{\ell+1}^{r}$-free r-graph with n vertices. Let k_r denote the number of copies of K_r in $\partial_{r-2}\mathcal{H}$. Combining Observation 3.2.8 with Theorem 3.2.7 and Turán's theorem we obtain

$$|\mathcal{H}| \leq k_r \leq \binom{\ell}{r}\left(\frac{|\partial_{r-2}\mathcal{H}|}{\binom{\ell}{2}}\right)^{r/2} \leq \binom{\ell}{r}\left(\frac{n}{\ell}\right)^{r}.$$

This proves Theorem 3.1.19 for the case ℓ divides n. ∎

Next, we prove Theorem 3.1.20.

Proof of Theorem 3.1.20. Let $\delta > 0$ be a sufficiently small constant, $\epsilon = (r-2)!\delta$, and n be a sufficiently large integer. Let $\mathcal{H}$ be a $\mathcal{K}_{\ell+1}^r$-free r-graph on n vertices with at least $(1-\epsilon)t_r(n,\ell) \geq (1-2\epsilon)\binom{\ell}{r}(n/\ell)^r$ edges. Let k_r denote the number of copies of K_r in $\partial_{r-2}\mathcal{H}$. Combining Observation 3.2.8 with Theorem 3.2.7 we obtain

$$|\partial_{r-2}\mathcal{H}| \geq \binom{\ell}{2}\left(\frac{k_r}{\binom{\ell}{r}}\right)^{2/r} \geq \binom{\ell}{2}\left(\frac{|\mathcal{H}|}{\binom{\ell}{r}}\right)^{2/r} \geq (1-2\epsilon)^{2/r}\binom{\ell}{2}\left(\frac{n}{\ell}\right)^2$$
$$\geq (1-2\epsilon)\binom{\ell}{2}\left(\frac{n}{\ell}\right)^2 \geq (1-2\epsilon)t_2(n,\ell).$$

Theorem 3.2.6 applied to $\partial_{r-2}\mathcal{H}$ implies that there exists a partition $V(\mathcal{H}) = V_1 \cup \cdots \cup V_\ell$ such that all but at most $2\epsilon t_2(n,\ell)$ edges in $\partial_{r-2}\mathcal{H}$ have at most one vertex in each V_i. It follows that all but at most $2\epsilon t_2(n,\ell)\binom{n}{r-2} \leq \epsilon n^r/(r-2)! \leq \delta n^r$ edges in $\mathcal{H}$ have at most one vertex in each V_i. This completes the proof of Theorem 3.1.20. ∎

3.2.3 Applications to the generalized Turán problems

In this section we present some applications of Theorem 3.2.7 for the generalized Turán problems.

Let T and H be two ordinary graphs. Denote by $\mathrm{ex}(n,T,H)$ the maximum possible number of copies of T in an ordinary H-free graph on n vertices. The function $\mathrm{ex}(n,T,H)$ is called the generalized Turán number.

Fix $\ell \geq r \geq 3$. In [57] Erdős proved that $\mathrm{ex}(n, K_r, K_{\ell+1}) \leq t_r(n, \ell)$. A similar argument as in the proof of Theorem 3.1.20 also gives the following stability result to $\mathrm{ex}(n, K_r, K_{\ell+1})$. For the sake of completeness let us include its short proof here.

Theorem 3.2.9. *Fix integers $\ell \geq r \geq 3$. Then for every $\delta > 0$ there exist $\epsilon > 0$ and n_0 such that the following holds for all $n \geq n_0$. If G is an n-vertex $K_{\ell+1}$-free graph containing at least $(1-\epsilon)t_r(n, \ell)$ copies of K_r, then the vertex set of G has a partition $V_1 \cup \cdots \cup V_\ell$ such that all but at most δn^2 edges in G have at most one vertex in each V_i.*

In fact, the following proof shows that it suffices to choose $\epsilon = \delta$ in Theorem 3.2.9.

Proof of Theorem 3.2.9. Let $\epsilon > 0$ be a sufficiently small constant and n be a sufficiently large integer. Let G be a $K_{\ell+1}$-free graph on n vertices with at least $(1-\epsilon)t_r(n, \ell) \geq (1-2\epsilon)\binom{\ell}{r}(n/\ell)^r$ copies of K_r. For $1 \leq i \leq \ell$ let k_i denote the number of copies of K_i in G. Then it follows from Theorem 3.2.7 that

$$\begin{aligned} |G| = k_2 \geq \binom{\ell}{2}\left(\frac{k_r}{\binom{\ell}{r}}\right)^{2/r} &\geq (1-2\epsilon)^{2/r}\binom{\ell}{2}\left(\frac{n}{\ell}\right)^2 \\ &\geq (1-2\epsilon)\binom{\ell}{2}\left(\frac{n}{\ell}\right)^2 \geq (1-2\epsilon)t_2(n, \ell). \end{aligned}$$

Theorem 3.2.6 applied to G implies that there exists a partition $V(G) = V_1 \cup \cdots \cup V_\ell$ such that all but at most $2\epsilon t_2(n, \ell) \leq \epsilon n^2$ edges in G have at most one vertex in each V_i. ∎

In [10], Alon and Shikhelman studied the function $\mathrm{ex}(n, T, H)$ for many other combinations of T and H. In particular they proved that $\mathrm{ex}(n, K_r, H) = (1 + o(1))t_r(n, \ell)$ for all graphs H

with chromatic number $\chi(H) = \ell + 1$. Later their result was improved by Ma and Qiu [184], who proved that $\text{ex}(n, K_r, H) = t_r(n, \ell) + \text{biex}(n, H) \cdot \Theta(n^{r-2})$, where $\text{biex}(n, H)$ is the Turán number of the decomposition family of H. Their proof is based on the following stability theorem for $\text{ex}(n, K_r, H)$.

Theorem 3.2.10 (Ma–Qiu [184]). *Fix $\ell \geq r \geq 3$, and $\delta > 0$. For every graph H with chromatic number $\ell + 1$, there exists an $\epsilon > 0$ and an n_0 such that the following holds for all $n \geq n_0$. If G is an n-vertex H-free graph containing at least $(1 - \epsilon)t_r(n, \ell)$ copies of K_r, then the vertex set of G has a partition $V_1 \cup \cdots \cup V_\ell$ such that all but at most δn^2 edges in G have at most one vertex in each V_i.*

Here we present a short proof to Theorem 3.2.10 using Theorem 3.2.9 and the Removal Lemma, and our proof shows that $\epsilon = \delta/3$ suffices for Theorem 3.2.10.

Theorem 3.2.11 (Removal Lemma, e.g. see [110; 85]). *Let H be a graph with chromatic number $\ell + 1$. For every $\delta > 0$ there exists an n_0 such that the following holds for all $n \geq n_0$. Every n-vertex H-free graph G can be made $K_{\ell+1}$-free by removing at most δn^2 edges.*

Proof of Theorem 3.2.10. Let $\epsilon = \delta/3$, and let n be sufficiently large. Let G be an n-vertex H-free graph containing at least $(1 - \epsilon)t_r(n, \ell)$ copies of K_r. By the Removal Lemma, G contains a $K_{\ell+1}$-free subgraph G' with at least $|G| - \epsilon n^2/\ell^r$ edges. Since every edge e in G is contained in at most $\binom{n}{r-2}$ copies of K_r in G, the number of copies of K_r in G' is at least $(1 - 2\epsilon)t_r(n, \ell)$. By Theorem 3.2.9, the vertex set of G' has a partition $V_1 \cup \cdots \cup V_\ell$ such that all but at most

$2\epsilon n^2$ edges in G' have at most one vertex in each V_i. Therefore, all but at most $3\epsilon n^2$ edges in G have at most one vertex in each V_i. ∎

3.2.4 Concluding Remarks

We showed that a linear dependence between δ and ϵ is sufficient for Theorems 3.1.20, 3.1.15, 3.2.9 and 3.2.10, and in [110] Füredi showed that a linear dependence between δ and ϵ is also sufficient for Theorem 3.2.6. It seems to be an interesting problem in general to determine the exact relationship between ϵ and δ in these stability theorems, and we refer the reader to [**?**; 13; 147] for related results on this topic.

3.3 Proofs for the general results

In this section we prove several general results about the feasible region $\Omega(\mathcal{F})$. First let us present a simple but useful idea that will be used in our proofs.

Fact 3.3.1. *Let $r \geq 2$ be an integer. Suppose that $\mathcal{H}$ is an r-graph on n vertices, and every edge in $\mathcal{H}$ contains an $(r-1)$-subset that is not covered by any other edge in $\mathcal{H}$. Then $|\mathcal{H}| \leq \binom{n}{r-1}$.*

Indeed, if every edge in $\mathcal{H}$ contains a unique $(r-1)$-subset, then we can map every edge $E \in \mathcal{H}$ to an $(r-1)$-subset of E that is not covered by any other edge in $\mathcal{H}$. This map is an injection from $\mathcal{H}$ to $\binom{[n]}{r-1}$ and it implies the upper bound in Fact 3.3.1. Actually, it was shown by Bollobás [23] that $|\mathcal{H}| \leq \binom{n-1}{r-1}$.

Algorithm 1 (Remove edges with the edge density threshold d)

Input: An r-graph $\mathcal{H}$ and the density threshold $d \in [0, 1]$.

Operation: If $d(\mathcal{H}) \leq d$ or $|\mathcal{H}| \leq \binom{n}{r-1}$, then do nothing and let $\mathcal{H}$ be the output. Otherwise, by Fact 3.3.1, there exists $E \in \mathcal{H}$ such that every $(r-1)$-subset of E is covered by another edge in $\mathcal{H}$. Remove E from the edge set of $\mathcal{H}$, and let $\mathcal{H}$ denote the resulting r-graph. Repeat this operation until $d - 1/\binom{n}{r} < d(\mathcal{H}) \leq d$.

Output: Either the original r-graph $\mathcal{H}$ or a subgraph $\mathcal{H}' \subset \mathcal{H}$ with $d - 1/\binom{n}{r} < d(\mathcal{H}') \leq d$, and $|\partial\mathcal{H}'| = |\partial\mathcal{H}|$.

Notice that the Operation above does not change $|\partial\mathcal{H}|$ since all $(r-1)$-subsets of the removed edge E are also covered by some other edge in $\mathcal{H}$. Therefore, the output r-graph $\mathcal{H}'$ satisfies

$|\partial\mathcal{H}'| = |\partial\mathcal{H}|$. On the other hand, since each step of the operation reduces $|\mathcal{H}|$ by exactly one, $d(\mathcal{H})$ can be reduced to some real number d' with $d - 1/\binom{n}{r} < d' \leq d$.

3.3.1 Basic properties

In this section we will prove Propositions 3.1.3, 3.1.7, and 3.1.8, and Theorem 3.1.11. First we prove Proposition 3.1.3.

Proof of Proposition 3.1.3. Let (x, y) be a limit point of $\Omega(\mathcal{F})$. For every positive integer k we will specify a hypergraph $\mathcal{H}_k$ with $v(\mathcal{H}_k) \geq k$, $|d(\partial\mathcal{H}_k) - x| \leq 1/k$ and $|d(\mathcal{H}_k) - y| \leq 1/k$. The resulting sequence $(\mathcal{H}_k)_{k=1}^{\infty}$ will be good and realize (x, y), so it will establish $(x, y) \in \Omega(\mathcal{F})$. For the construction of $\mathcal{H}_k$ we first take a point $(x_k, y_k) \in \Omega(\mathcal{F})$ such that $|x - x_k| \leq 1/(2k)$ and $|y - y_k| \leq 1/(2k)$. Every good sequence $(\mathcal{H}_{k,m})_{m=1}^{\infty}$ realizing (x_k, y_k) contains a hypergraph $\mathcal{H}_k$ with $v(\mathcal{H}_k) \geq k$, $|d(\partial\mathcal{H}_k) - x_k| \leq 1/k$ and $|d(\mathcal{H}_k) - y_k| \leq 1/(2k)$. By the triangle inequality, $\mathcal{H}_k$ has the desired properties. ∎

Next we prove Proposition 3.1.7. Its proof uses Algorithm 1.

Proof of Proposition 3.1.7. Since $(x_0, y_0) \in \Omega(\mathcal{F})$, there exists a good sequence of $\mathcal{F}$-free r-graphs $(\mathcal{H}_k)_{k=1}^{\infty}$ for which $\lim_{k\to\infty} d(\partial\mathcal{H}_k) = x_0$ and $\lim_{k\to\infty} d(\mathcal{H}_k) = y_0$. Now fix $y \in [0, y_0)$. For every $k \geq 1$ apply Algorithm 1 to $\mathcal{H}_k$ with edge density threshold y and let $\mathcal{H}'_k$ denote the r-graph that Algorithm 1 outputs. We claim that $(\mathcal{H}'_k)_{k=1}^{\infty}$ is a good sequence of $\mathcal{F}$-free r-graphs that realizes (x_0, y). Indeed, choose $\epsilon = (y_0 - y)/2 > 0$, by the assumption that $\lim_{k\to\infty} d(\mathcal{H}_k) = y_0$, there exists k_0 such that $d(\mathcal{H}_k) \in (y_0 - \epsilon, y_0 + \epsilon)$ for all $k \geq k_0$. Therefore, by Algorithm 1, $y - 1/\binom{v(\mathcal{H}_k)}{r} < d(\mathcal{H}'_k) \leq y$ for all $k \geq k_0$, and hence $\lim_{k\to\infty} d(\mathcal{H}'_k) = y$. On the other hand, since $|\partial \mathcal{H}'_k| = |\partial \mathcal{H}_k|$ for all $k \geq 1$, $\lim_{k\to\infty} d(\partial \mathcal{H}'_k) = x$. Therefore, $(\mathcal{H}'_k)_{k=1}^{\infty}$ is a good sequence of $\mathcal{F}$-free r-graphs that realizes (x_0, y), and hence $(x_0, y) \in \Omega(\mathcal{F})$. ∎

Recall that $\mathrm{ex}(n, \mathcal{F}_1) \leq \mathrm{ex}(n, \mathcal{F}_2)$ whenever $\mathcal{F}_2 \subset \mathcal{F}_1$. By the definition of $g(\mathcal{F})$, a similar inequality also holds for $g(\mathcal{F})$.

Observation 3.3.2. *Let $r \geq 3$. Suppose that $\mathcal{F}_1$ and $\mathcal{F}_2$ are two families of r-graphs with $\mathcal{F}_1 \subset \mathcal{F}_2$. Then $\Omega(\mathcal{F}_2) \subset \Omega(\mathcal{F}_1)$. In particular, $g(\mathcal{F}_2, x) \leq g(\mathcal{F}_1, x)$ for all $x \in \mathrm{proj}\Omega(\mathcal{F}_2)$.*

Now we are ready to prove Proposition 3.1.7.

Proof of Proposition 3.1.7. By Observation 3.3.2, it suffices to show that $\mathrm{proj}\Omega(\emptyset) = [0, 1]$ and $g(\emptyset, x) = x^{r/(r-1)}$ for all $x \in [0, 1]$. The first part is easy, since the complete r-graph on n vertices has shadow density 1, and it follows from Observation 3.1.5 that $\mathrm{proj}\Omega(\emptyset) = [0, 1]$.

Now we consider the second part. First we show that $g(\emptyset, x) \leq x^{r/(r-1)}$ for all $x \in [0, 1]$. Let $(\mathcal{H}_k)_{k=1}^{\infty}$ be a good sequence of r-graph that realizes (x, y). For every $k \geq 1$ let α_k denote the real number that satisfies $|\partial \mathcal{H}_k| = \binom{\alpha_k v(\mathcal{H}_k)}{r-1}$. By the Kruskal–Katona theorem, $|\mathcal{H}_k| \leq \binom{\alpha_k v(\mathcal{H}_k)}{r}$ for all $k \geq 1$. By assumption and $\lim_{k\to\infty} v(\mathcal{H}_k) = \infty$,

$$x = \lim_{k\to\infty} \frac{|\partial \mathcal{H}_k|}{\binom{v(\mathcal{H}_k)}{r-1}} = \lim_{k\to\infty} \frac{\binom{\alpha_k v(\mathcal{H}_k)}{r-1}}{\binom{v(\mathcal{H}_k)}{r-1}} = \lim_{k\to\infty} (\alpha_k)^{r-1},$$

which implies that $\lim_{k\to\infty} \alpha_k = x^{1/(r-1)}$. Therefore, by assumption,

$$y = \lim_{k\to\infty} \frac{|\mathcal{H}_k|}{\binom{v(\mathcal{H}_k)}{r-1}} \leq \lim_{k\to\infty} \frac{\binom{\alpha_k v(\mathcal{H}_k)}{r}}{\binom{v(\mathcal{H}_k)}{r}} = \lim_{k\to\infty} (\alpha_k)^r = x^{\frac{r}{r-1}},$$

and this proves that $g(\emptyset, x) \leq x^{r/(r-1)}$ for all $x \in [0,1]$.

Next we show that $g(\emptyset, x) \geq x^{r/(r-1)}$ for all $x \in [0,1]$. Choose an arbitray $x \in [0,1]$ and let $\alpha = x^{1/(r-1)}$. Let $\mathcal{H}_n(\alpha)$ denote the vertex disjoint union of a complete r-graph on αn vertices and a set of $(1-\alpha)n$ isolated vertices. Then we claim that $(\mathcal{H}_k(\alpha))_{k=1}^{\infty}$ is a good sequence of r-graphs that realizes $(x, x^{r/(r-1)})$. Indeed,

$$\lim_{k\to\infty} \frac{|\partial\mathcal{H}_k(\alpha)|}{\binom{n}{r-1}} = \lim_{k\to\infty} \frac{\binom{\alpha n}{r-1}}{\binom{n}{r-1}} = \alpha^{r-1} = x,$$

and

$$\lim_{k\to\infty} \frac{|\mathcal{H}_k(\alpha)|}{\binom{n}{r}} = \lim_{k\to\infty} \frac{\binom{\alpha n}{r}}{\binom{n}{r}} = \alpha^{r} = x^{\frac{r}{r-1}},$$

and it follows from the definition that $g(\emptyset, x) \geq x^{r/(r-1)}$ for all $x \in [0,1]$. ∎

3.3.2 Continuity and differentiability

In this section we will prove Theorem 3.1.11 and some other related corollaries. We will use the following theorem in our proofs.

Theorem 3.3.3 (see Section 3 of Chapter 3 in [235])**.** *Let $f : \mathbb{R} \to \mathbb{R}$ be a monotone function. Then f has at most countably many discontinuities of the first kind and no discontinuity of the second kind. Moreover, f is almost everywhere differentiable.*

The following lemma is the main tool in our proofs.

Lemma 3.3.4. *Let $r \geq 3$ and $\mathcal{F}$ be a family of r-graphs. Then*

$$(g(\mathcal{F}, x+h))^{\frac{r-1}{r}} \leq (g(\mathcal{F}, x))^{\frac{r-1}{r}} + \frac{(g(\mathcal{F}, x))^{\frac{r-1}{r}}}{x} h$$

for all $x \in \mathrm{proj}\Omega(\mathcal{F}) \setminus \{0\}$ and all $h \geq 0$ with $x + h \in \mathrm{proj}\Omega(\mathcal{F})$.

Proof. Suppose that $x + h \in \mathrm{proj}\Omega(\mathcal{F})$. Choose

$$\alpha = \left(\frac{x+h}{x}\right)^{\frac{1}{r-1}} - 1.$$

Let $(\mathcal{H}_k)_{k=1}^{\infty}$ be a good sequence of $\mathcal{F}$-free r-graphs that realizes $(x+h, g(\mathcal{F}, x+h))$. For every $k \geq 1$ let $n_k = v(\mathcal{H}_k)$ and let $\mathcal{H}'_k$ be obtained from $\mathcal{H}_k$ by adding a set of αn_k isolated vertices and let $n'_k = (1+\alpha)n_k$. Then,

$$\lim_{k\to\infty} \frac{|\partial \mathcal{H}'_k|}{\binom{n'_k}{r-1}} = \lim_{k\to\infty} \frac{|\partial \mathcal{H}_k|}{\binom{(1+\alpha)n_k}{r-1}} = \frac{x+h}{(1+\alpha)^{r-1}} = x,$$

and

$$\lim_{k\to\infty} \frac{|\mathcal{H}'_k|}{\binom{n'_k}{r}} = \lim_{k\to\infty} \frac{|\mathcal{H}_k|}{\binom{(1+\alpha)n_k}{r}} = \frac{g(\mathcal{F}, x+h)}{(1+\alpha)^r} = \left(\frac{x}{x+h}\right)^{\frac{r}{r-1}} g(\mathcal{F}, x+h).$$

Therefore, $(\mathcal{H}'_k)_{k=1}^{\infty}$ a good sequence of $\mathcal{F}$-free r-graphs that realizes

$$\left(x, \left(\frac{x}{x+h}\right)^{\frac{r}{r-1}} g(\mathcal{F}, x+h)\right).$$

Consequently,

$$g(\mathcal{F}, x) \geq \left(\frac{x}{x+h}\right)^{\frac{r}{r-1}} g(\mathcal{F}, x+h), \tag{3.8}$$

which gives

$$(g(\mathcal{F}, x+h))^{\frac{r-1}{r}} \leq (g(\mathcal{F}, x))^{\frac{r-1}{r}} + \frac{(g(\mathcal{F}, x))^{\frac{r-1}{r}}}{x} h.$$

∎

Corollary 3.3.5. *Let $r \geq 3$ and $\mathcal{F}$ be a family of r-graphs. Then for any $x \in \mathrm{proj}\Omega(\mathcal{F}) \setminus \{0\}$ and any $\delta > 0$, there exists $\epsilon > 0$ such that $g(\mathcal{F}, x') > g(\mathcal{F}, x) - \delta$ for all $x' \in (x - \epsilon, x)$.*

Proof. We may assume that $\delta < 1$. Choose $\epsilon = \delta x/3$ and let $x' \in (x - \epsilon, x)$. Then Equation 3.8 gives

$$\begin{aligned} g(\mathcal{F}, x') &\geq \left(\frac{x'}{x}\right)^{\frac{r}{r-1}} g(\mathcal{F}, x) \\ &= \left(1 - \frac{x - x'}{x}\right)^{\frac{r}{r-1}} g(\mathcal{F}, x) \\ &\geq \left(1 - \frac{2\epsilon}{x}\right) g(\mathcal{F}, x) = g(\mathcal{F}, x) - \frac{2g(\mathcal{F}, x)\epsilon}{x} > g(\mathcal{F}, x) - \delta, \end{aligned}$$

where the second inequality follows from the fact that $(1 - x)^a \geq 1 - ax$ for all $x \in [0, 1]$ and all $a \geq 1$. ∎

Proposition 3.1.3 together with Corollary 3.3.5 will show that $g(\mathcal{F})$ does not contain removable discontinuities.

Corollary 3.3.6. *Let $r \geq 3$ and $\mathcal{F}$ be a family of r-graphs. Then $g(\mathcal{F})$ does not contain removable discontinuities.*

Proof. Suppose that $x_0 \in \text{proj}\Omega(\mathcal{F})$ is a removable discontinuity of $g(\mathcal{F})$. Then $x_0 > 0$ and $\lim_{x\to x_0^-} g(\mathcal{F}, x) = \lim_{x\to x_0^+} g(\mathcal{F}, x) \neq g(\mathcal{F}, x_0)$. Let $y_0 = \lim_{x\to x_0^-} g(\mathcal{F}, x)$. By Proposition 3.1.3, $(x_0, y_0) \in \Omega(\mathcal{F})$, and by the definition of $g(\mathcal{F})$, $g(\mathcal{F}, x_0) > y_0$. Letting $\delta = (g(\mathcal{F}, x_0) - y_0)/2$ in Corollary 3.3.5, we obtain

$$y_0 = \lim_{x\to x_0^-} g(\mathcal{F}, x) > g(\mathcal{F}, x_0) - \delta = \frac{g(\mathcal{F}, x_0) + y_0}{2} > y_0,$$

a contradiction. ∎

Now we are ready to prove Theorem 3.1.11.

Proof of Theorem 3.1.11. First we show that $g(\mathcal{F})$ is almost everywhere differentiable. Let $f(x) = (g(\mathcal{F}, x))^{\frac{r-1}{r}} - x$. It follows from Lemma 3.3.4 and Proposition 3.1.8 that

$$\begin{aligned}
(g(\mathcal{F}, x+h))^{\frac{r-1}{r}} &\leq (g(\mathcal{F}, x))^{\frac{r-1}{r}} + \frac{(g(\mathcal{F}, x))^{\frac{r-1}{r}}}{x}h \\
&\leq (g(\mathcal{F}, x))^{\frac{r-1}{r}} + \frac{\left(x^{\frac{r}{r-1}}\right)^{\frac{r-1}{r}}}{x}h \\
&= (g(\mathcal{F}, x))^{\frac{r-1}{r}} + h,
\end{aligned}$$

which implies that f is decreasing on $\text{proj}\Omega(\mathcal{F})$. By Theorem 3.3.3, f is almost everywhere differentiable, and so is $g(\mathcal{F})$.

Next, we show that $g(\mathcal{F})$ has at most countably many jump discontinuities. By Theorem 3.3.3, f has at most countably many discontinuities of the first kind, and so does $g(\mathcal{F})$ since $g(\mathcal{F}, x) = (f(x) + x)^{r/(r-1)}$ for all $x \in \text{proj}\Omega(\mathcal{F})$. Corollary 3.3.5 shows that $g(\mathcal{F})$ does not have a removable discontinuity, therefore, $g(\mathcal{F})$ has at most countably many jump discontinuities.

Finally, we show that $g(\mathcal{F})$ is left-continuous. Let $x_0 \in \text{proj}\Omega(\mathcal{F})$ be a discontinuity of $g(\mathcal{F})$. By the previous result, x_0 can only be a jump discontinuity. Let $y_0^- = \lim_{x \to x_0^-} g(\mathcal{F}, x)$ and $y_0^+ = \lim_{x \to x_0^+} g(\mathcal{F}, x)$. By Proposition 3.1.3, $(x_0, y_0^-) \in \Omega(\mathcal{F})$ and $(x_0, y_0^+) \in \Omega(\mathcal{F})$. So, it suffices to show that $y_0^- > y_0^+$. Indeed, suppose that $y_0^+ > y_0^-$. Then, by the definition of $g(\mathcal{F})$ we would have $g(\mathcal{F}, x_0) = y_0^+$. Letting $\delta = (y_0^+ - y_0^-)/2$ in Corollary 3.3.5, we obtain

$$y_0^- = \lim_{x \to x_0^-} g(\mathcal{F}, x) > g(\mathcal{F}, x_0) - \delta = \frac{y_0^- + y_0^+}{2} > y_0^-,$$

a contradiction, and this completes the proof. ∎

The proof of Theorem 3.1.11 also gives the following corollary.

Corollary 3.3.7. *Let $r \geq 3$ and $\mathcal{F}$ be a family of r-graphs. Suppose that $x_0 \in \text{proj}\Omega(\mathcal{F})$ is a discontinuity of $g(\mathcal{F})$. Then both $\lim_{x \to x_0^-} g(\mathcal{F}, x)$ and $\lim_{x \to x_0^+} g(\mathcal{F}, x)$ exist and $\lim_{x \to x_0^-} g(\mathcal{F}, x) > \lim_{x \to x_0^+} g(\mathcal{F}, x)$. In particular, if $g(\mathcal{F})$ is increasing on $[c_1, c_2]$ for some $c_2 > c_1 \geq 0$, then $g(\mathcal{F})$ is continuous on $[c_1, c_2]$.*

3.4 A point of discontinuity

In this section we will prove Theorem 3.1.12 by defining a family $\mathcal{D}$ of 3-graphs, and showing that $g(\mathcal{D})$ is discontinuous at $x = 2/3$.

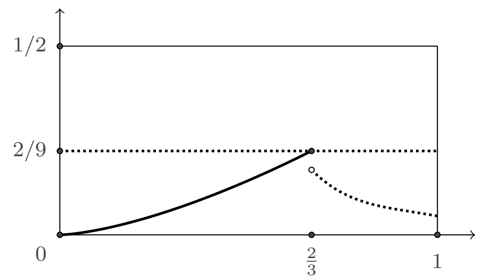


Figure 11. The function $g(\mathcal{D})$ is discontinuous at $x = 2/3$.

First we define a 3-graph $\mathcal{S}_n$ on $[n]$ as follows. Fix $u \in [n]$, let

$$\mathcal{S}_n = \left\{ uvw \colon vw \in \binom{[n] \setminus \{u\}}{2} \right\},$$

and note that $\mathcal{S}_n$ is a star with $|\mathcal{S}_n| = \binom{n-1}{2}$.

Definition 3.4.1. *Let $\mathcal{D}$ be the collection of all 3-graphs $F \in \mathcal{K}_4^3$ such that $F \not\subset \mathcal{S}_n$ for all $n \geq 4$.*

Note that $\mathcal{D} \neq \emptyset$ as $H_4^3 \in \mathcal{D}$. Since $\mathcal{S}_n$ is $\mathcal{D}$-free and $\lim_{n\to\infty} |\partial\mathcal{S}_n|/\binom{n}{2} = 1$, by Observation 3.1.5, $\mathrm{proj}\Omega(\mathcal{D}) = [0,1]$.

Since $T_3(n,3)$ is $\mathcal{K}_4^3$-free, $\mathrm{ex}(n,\mathcal{D}) \geq t_3(n,3)$. On the other hand, $\mathrm{ex}(n,\mathcal{D}) \leq \mathrm{ex}(n,H_4^3)$, which, by [209], is at most $t_3(n,3)$ when n is sufficiently large. Therefore, we obtain the following result.

Theorem 3.4.2. *Let n be sufficiently large. Then $ex(n,\mathcal{D}) = t_3(n,3)$ and $T_3(n,3)$ is the unique $\mathcal{D}$-free 3-graph with n vertices and $t_3(n,3)$ edges.*

Theorem 3.4.2 implies that $g(\mathcal{D},x) \leq 2/9$ for all $x \in [0,1]$ and equality holds for $x = 2/3$. Therefore, in order to prove Theorem 3.1.12 it suffices to prove the following result.

Theorem 3.4.3. *There exists an absolute constant $\delta_0 > 0$ such that the following is true for all $\epsilon \in (0, 10^{-8})$ and sufficiently large n. Suppose that $\mathcal{H}$ is a $\mathcal{D}$-free 3-graph on n vertices with $|\partial\mathcal{H}| = (1/3+\epsilon)n^2$. Then $|\mathcal{H}| \leq (1/27 - \delta_0)n^3$.*

The proof of Theorem 3.4.3 uses a stability result for $\mathcal{D}$-free 3-graphs, which can be easily obtained from a stability theorem for $H_{\ell+1}^r$-free r-graphs proved by Pikhurko [209].

Theorem 3.4.4 (Stability)**.** *For every $\xi > 0$ there exists $\delta > 0$ (we may assume that $\delta \leq \xi$) and $n_0 = n_0(\xi)$ such that the following holds for all $n \geq n_0$. Suppose that $\mathcal{H}$ is a $\mathcal{D}$-free 3-graph on n vertices with $|\mathcal{H}| \geq (1/27 - \delta)n^3$. Then $V(\mathcal{H})$ has a partition $V_1 \cup V_2 \cup V_3$ such that all but at most ξn^3 edges in $\mathcal{H}$ have exactly one vertex in each V_i.*

Now we are ready to prove Theorem 3.4.3.

Proof of Theorem 3.4.3. We prove Theorem 3.4.3 by contradiction. Suppose that for all constant $\delta > 0$ and all integers n_0 there exists $\epsilon = \epsilon(\delta) \in (0, 10^{-8})$ such that there exists a 3-graph $\mathcal{H}$ on $n > n_0$ vertices for some n with $|\partial\mathcal{H}| = (1/3 + \epsilon)n^2$ and $|\mathcal{H}| > (1/27 - \delta)n^3$.

Choose $\xi > 0$ to be sufficiently small, and let $\delta > 0$ and $n_0 = n_0(\xi)$ be given by Theorem 3.4.4 and note that we may assume that $\delta \leq \xi$. By assumption, there exists $\epsilon \in (0, 10^{-8})$ and a $\mathcal{D}$-free 3-graphs $\mathcal{H}$ on $n > n_0$ vertices with $|\partial\mathcal{H}| = (1/3 + \epsilon)n^2$ and $|\mathcal{H}| > (1/27 - \delta)n^3$. Apply Theorem 3.4.4 to $\mathcal{H}$. We obtain a partition $V(\mathcal{H}) = V_1 \cup V_2 \cup V_3$ such that all but at most ξn^3 edges in $\mathcal{H}$ have exactly one vertex in each V_i. Let $\mathcal{H}'$ denote the induced 3-partite 3-graph of $\mathcal{H}$ on $V_1 \cup V_2 \cup V_3$, that is,

$$\mathcal{H}' = \{E \in \mathcal{H} : |E \cap V_i| = 1 \text{ for all } i \in [3]\}.$$

Note that

$$|\mathcal{H}'| > \frac{n^3}{27} - \delta n^3 - \xi n^3. \tag{3.9}$$

Claim 3.4.5. $\left||V_i| - \frac{n}{3}\right| < 4(\delta + \xi)^{1/2} n$ *for all* $i \in [3]$.

Proof. Fix $1 \leq i \leq 3$ and let $\alpha = |V_i|$. Then $|\mathcal{H}'| \leq \alpha(n-\alpha)^2/4$ and Equation 3.9 gives

$$\frac{\alpha(n-\alpha)^2}{4} > \frac{n^3}{27} - \delta n^3 - \xi n^3,$$

which implies $n/3 - 4(\delta + \xi)^{1/2} n < \alpha < n/3 + 4(\delta + \xi)^{1/2} n$. ∎

Let $G = \partial\mathcal{H}$ and $G' = \partial\mathcal{H}'$. Note that $\mathcal{H}' \subset \mathcal{H}$, $G' \subset G$, and G' is 3-partite. Let K be a 3-partite subgraph of G with the maximum number of edges among all 3-partite subgraphs of G, and let X_1, X_2, X_3 denote the three parts of K.

Claim 3.4.6. $|K| \geq |G'| > \frac{n^2}{3} - 5(\delta+\xi)^{1/2}n^2$.

Proof. Counting the number of edges in $\mathcal{H}'$ we obtain

$$|G'|\left(\frac{n}{3} + 4(\delta+\xi)^{1/2}n\right) \overset{\text{Claim 3.4.5}}{>} 3|\mathcal{H}'| \overset{\text{Equation 3.9}}{>} \frac{n^3}{9} - 3\left(\delta+\xi\right)n^3,$$

which implies $|G'| > n^2/3 - 5(\delta+\xi)^{1/2}n^2$. Since G' is also a 3-partite subgraph of G, by the maximality of K, we obtain $|K| \geq |G'|$. ∎

Claim 3.4.7. $\left||X_i| - \frac{n}{3}\right| < 4\left(\delta+\xi\right)^{1/4} n$ *for all* $i \in [3]$.

Proof. Fix $i \in [3]$ and let $\alpha' = |X_i|$. By Claim 3.4.6,

$$\alpha'(n-\alpha') + \frac{(n-\alpha')^2}{4} \geq |K| \geq |G'| > \frac{n^2}{3} - 5(\delta+\xi)^{1/2}n^2,$$

which implies $n/3 - 4\left(\delta+\xi\right)^{1/4} n < \alpha' < n/3 + 4\left(\delta+\xi\right)^{1/4} n$. ∎

For $uv \in K$ the degree of uv in $\mathcal{H}$ is $d(uv) := |\{E \in \mathcal{H} \colon \{u,v\} \subset E\}|$. Our next claim shows that most edges in K have a large degree.

Claim 3.4.8. *The number of edges in K that have degree at most* 10 *in $\mathcal{H}$ is at most* $n^2/40000$.

Proof. Suppose not. Then the assumption that $|G| = |\partial\mathcal{H}| = (1/3+\epsilon)n^2$ together with Claims 3.4.6 and 3.4.7 imply

$$\begin{aligned}|\mathcal{H}| &\overset{\text{Claim 3.4.7}}{\leq} \frac{1}{3}\left(|K| - \frac{n^2}{40000}\right)\left(\frac{n}{3} + 4(\delta+\xi)^{1/4}n\right) + \frac{10n^2}{40000} + (|G| - |K|)\,n \\ &\overset{\text{Claim 3.4.6}}{\leq} \frac{1}{3}\left(\frac{n^2}{3} - \frac{n^2}{40000}\right)\left(\frac{n}{3} + 4(\delta+\xi)^{1/4}n\right) + \frac{n^2}{4000} + \epsilon n^3 + 5(\delta+\xi)^{1/4}n^3 \\ &< \frac{n^3}{27} - \frac{n^3}{500000},\end{aligned}$$

which contradicts the assumption that $|\mathcal{H}| > (1/27 - \delta)n^3$. Here we used the fact that δ, ξ are sufficiently small, n is sufficiently large, and $\epsilon < 10^{-8}$. ∎

The next claim shows that if G has a large complete 4-partite subgraph, then it contains many edges that have degree at most 10 in $\mathcal{H}$. This is the only place where we use the definition of $\mathcal{D}$.

Claim 3.4.9. *Let $v_1v_2 \in G$ and $U_1, U_2 \subset V(\mathcal{H}) \setminus \{v_1, v_2\}$. Let*

$$L = \{\{u_1, u_2\} \colon u_1 \in U_1, u_2 \in U_2 \text{ and } d(u_1u_2) \geq 10\}.$$

Suppose that v_1 and v_2 are adjacent to all vertices in $U_1 \cup U_2$. Then L is an intersecting family, and hence $|L| < n$.

Proof. Let $u_1u_2 \in L$ and

$$\mathcal{E}_{v_1v_2} = \{E \in \mathcal{H} : \{v_1, v_2\} \subset E\}.$$

We claim that every set $E \in \mathcal{E}_{v_1v_2}$ satisfies $E \cap \{u_1, u_2\} \neq \emptyset$. Indeed, suppose that there exists $E_{v_1v_2} \in \mathcal{E}_{v_1v_2}$ with $E_{v_1v_2} \cap \{u_1, u_2\} = \emptyset$. Since $d(u_1u_2) \geq 10$, there exists $E_{u_1u_2} \in \mathcal{H}$ such that $\{u_1, u_2\} \in E_{u_1u_2}$ and $E_{u_1u_2} \cap E_{v_1v_2} = \emptyset$. Let $E_{v_1u_1}, E_{v_1u_2}, E_{v_2u_1}$, and $E_{v_2u_2}$ be edges in $\mathcal{H}$ that cover $v_1u_1, v_1u_2, v_2u_1, v_2u_2$, respectively, and let F_1 denote the 3-graph with edge set

$$\{E_{v_1v_2}, E_{v_1u_1}, E_{v_1u_2}, E_{v_2u_1}, E_{v_2u_2}, E_{u_1u_2}\}.$$

Note that $F_1 \subset \mathcal{H}$ and $F_1 \in \mathcal{K}_4^3$. However, since $E_{u_1u_2} \cap E_{v_1v_2} = \emptyset$, $F_1 \not\subset \mathcal{S}_n$ for any n, and hence $F_1 \in \mathcal{D}$, which is a contradiction. Therefore, every set $E \in \mathcal{E}_{v_1v_2}$ satisfies $E \cap \{u_1, u_2\} \neq \emptyset$.

Suppose that L contains another edge w_1w_2 that is disjoint from u_1u_2. Then, the same argument as above implies that every set $E \in \mathcal{E}_{v_1v_2}$ satisfies $E \cap \{w_1, w_2\} \neq \emptyset$. Therefore, every set $E \in \mathcal{E}_{v_1v_2}$ satisfies $E \cap \{u_1, u_2\} \neq \emptyset$ and $E \cap \{w_1, w_2\} \neq \emptyset$, which is impossible since E is a 3-set. Therefore, L is intersecting and it follows from the Erdős–Ko–Rado theorem [69] that $|L| < n$. ∎

Our goal in the rest of the proof is to find $v_1v_2 \in G$ and $U_1, U_2 \subset V(\mathcal{H}) \setminus \{v_1, v_2\}$ with $|U_1||U_2|$ large, such that v_1 and v_2 are adjacent to all vertices in $U_1 \cup U_2$. Then, by Claim 3.4.9, many edges in the induced subgraph of K on $U_1 \cup U_2$ would have degree at most 10, which contradicts Claim 3.4.8.

Let

$$B = \{uv \in G \colon \{u, v\} \subset X_i \text{ for some } i \in [3]\},$$

and

$$M = \left\{ \{u,v\} \in \binom{V(\mathcal{H})}{2} \setminus K \colon u \in X_i, v \in X_j \text{ for some } i,j \in [3] \text{ and } i \neq j \right\}.$$

Sets in B are called bad edges of K and sets in M are called missing edges of K. For $v \in V(\mathcal{H})$ let $d_M(v)$ denote the number of missing edges that contain v. By Claim 3.4.6,

$$|M| \leq 5(\delta + \xi)^{1/2} n^2. \tag{3.10}$$

On the other hand, the assumption $|G| = n^2/3 + \epsilon n^2$ implies

$$|B| \geq |M| + \epsilon n^2. \tag{3.11}$$

Let B_i be the collection of bad edges in G that are completely contained in X_i for $i \in [3]$. Without loss of generality, we may assume that $|B_1| \geq |B|/3$. Let Δ denote the maximum degree of B_1.

Case 1: $\Delta < n/100$.

Then there exits a set M' of at least $|B_1|/(2\Delta) > 15|B|/n$ pairwise disjoint edges in B_1. Fix $uv \in B_1$. Let $U_i(uv) = N_K(u) \cap N_K(v) \cap X_2$ for $i \in \{2,3\}$ and let K_{uv} denote the induced subgraph of K on $U_2(uv) \cup U_3(uv)$. By Claim 3.4.9, all but at most n edges in K_{uv} have degree at most 10 in $\mathcal{H}$. It follows that

$$|U_2(uv)||U_3(uv)| \overset{\text{Claim 3.4.8}}{\leq} \frac{n^2}{40000} + n + |M| \overset{\text{Equation 3.10}}{\leq} \frac{n^2}{40000} + n + 5(\delta+\xi)^{1/2} n^2 < \frac{n^2}{30000}.$$

Therefore, by Claim 3.4.7,

$$|N_K(u) \cap N_K(v)| < \frac{n}{3} + 4(\delta+\xi)^{1/4}n + \frac{n^2/30000}{n/3 + 4(\delta+\xi)^{1/4}n} < \frac{n}{3} + 4(\delta+\xi)^{1/4}n + \frac{n}{10000},$$

and it follows from Inclusion-Exclusion and Claim 3.4.7 that

$$\begin{aligned} d_K(u) + d_K(v) &= |N_K(u) \cup N_K(v)| + |N_K(u) \cap N_K(v)| \\ &\leq 2\left(\frac{n}{3} + 4(\delta+\xi)^{1/4}n\right) + \frac{n}{3} + 4(\delta+\xi)^{1/4}n + \frac{n}{10000} \\ &< \frac{101n}{100}. \end{aligned} \tag{3.12}$$

Note that

$$d_K(u) + d_M(u) + d_K(v) + d_M(v) = 2\left(|X_2| + |X_3|\right),$$

which implies

$$\begin{aligned} |M| \geq \sum_{uv \in M'} (d_M(u) + d_M(v)) &\geq \frac{15|B|}{n}\left(2\left(|X_2| + |X_3|\right) - d_K(u) - d_K(v)\right) \\ &\overset{\text{Claim 3.4.7 and Equation 3.12}}{>} \frac{15|B|}{n}\left(\frac{4n}{3} - \frac{102n}{100}\right) \\ &> 4|B| \overset{\text{Equation 3.11}}{>} |M|, \end{aligned}$$

a contradiction.

Case 2: $\Delta \geq n/100$.

Then choose a vertex $v_1 \in X_1$ with degree Δ. Let $N_i = N_K(v_1) \cap X_i$ for $1 \leq i \leq 3$. The maximality of K implies that $|N_2| \geq \Delta$ and $|N_3| \geq \Delta$, since otherwise we could move v_1 into V_2 or V_3 to get a larger 3-partite subgraph of G. Choose $v_2 \in N_1$ and let $U_i(v_1v_2) = N_K(v_2) \cap N_i$ for $i \in \{2, 3\}$. Similar to Case 1, we have $|U_2(v_1v_2)||U_3(v_1v_2)| \leq n^2/30000$. Therefore, v_2 is not adjacent (in K) to at least $n/200$ vertices in $N_2 \cup N_3$, which implies

$$|M| \geq \sum_{u \in N_1} d_M(u) \geq \frac{n}{100} \times \frac{n}{200} = \frac{n^2}{20000} > 5(\delta + \xi)^{1/2} n^2 \overset{\text{Equation 3.10}}{\geq} |M|,$$

a contradiction. ∎

Christian Reiher pointed out that the conclusion in Theorem 3.4.3 still holds even if we replace the assumption $|\partial\mathcal{H}| = (1/3 + \epsilon)n^2, \epsilon \in (0, 10^{-8})$ by $|\partial\mathcal{H}| \geq t_2(n, 3) + 1$. In fact he proved the following stronger stability theorem for $\mathcal{D}$, which immediately implies the stronger version of Theorem 3.4.3.

Lemma 3.4.10 (Reiher)**.** *For every $\epsilon > 0$ there are $\delta > 0$ and n_0 such that every $\mathcal{D}$-free 3-graph $\mathcal{H}$ on $n \geq n_0$ vertices with $|\mathcal{H}| \geq (1/27 - \delta)n^3$ admits a partition $V(\mathcal{H}) = U_1 \cup U_2 \cup U_3 \cup U_4$ such that*

- *every edge $E \in \mathcal{H}$ not incident with U_4 has exactly one vertex in each of U_1, U_2, U_3,*
- *the sets U_1, U_2, U_3 are independent in $\partial\mathcal{H}$,*
- *every vertex in U_4 is incident with at most $(1/18 + \epsilon)n^2$ edges in $\mathcal{H}$ and at most $(1/2 + \epsilon)n$ edges in $\partial\mathcal{H}$,*
- *and $|U_4| \leq \epsilon n$.*

3.5 Cancellative hypergraphs

In this section we prove Theorems 3.1.16 and 3.1.17. First let us present some useful lemmas. For a subset $S \subset V(\mathcal{H})$ let

$$\sigma_{\mathcal{H}}(S) = \sum_{v \in S} d_{\mathcal{H}}(v).$$

When it is clear from context we will omit the subscript $\mathcal{H}$ from the notions above.

Lemma 3.5.1. *Let $r \geq 3$ and let $\mathcal{H}$ be a cancellative r-graph. Then, for every $v \in V(\mathcal{H})$ the link $L(v)$ is a cancellative $(r-1)$-graph.*

Proof. Suppose that there exist $A, B, C \in L(v)$ such that $A \triangle B \subset C$. Let $A' = A \cup \{v\}$, $B' = B \cup \{v\}$ and $C' = C \cup \{v\}$, and note that $A', B', C' \in \mathcal{H}$. Then, $A' \triangle B' \subset C'$, which is a contradiction. ∎

Lemma 3.5.2. *Let $r \geq 3$ and let $\mathcal{H}$ be a cancellative r-graph. Suppose that $\{u, v\} \subset V(\mathcal{H})$ is covered by an edge in $\mathcal{H}$. Then $L(u) \cap L(v) = \emptyset$.*

Proof. Suppose that there exists $E \in L(u) \cap L(v)$. Let $A = E \cup \{u\}$ and $B = E \cup \{v\}$, and note that $A, B \in \mathcal{H}$. Then $A \triangle B = \{u, v\}$, which by assumption is covered by another edge C in $\mathcal{H}$, a contradiction. ∎

Lemma 3.5.2 gives the following corollary.

Corollary 3.5.3. *Let $r \geq 3$ and $\mathcal{H}$ be a cancellative r-graph. Let $S \subset V(\mathcal{H})$ and suppose that $(\partial_{r-2}\mathcal{H})[S]$ is a complete graph. Then,*

$$\sigma_{\mathcal{H}}(S) = \sum_{v \in S} d_{\mathcal{H}}(v) \leq |\partial\mathcal{H}|.$$

Proof. Suppose that $S = \{v_1, \ldots, v_s\}$. Lemma 3.5.2 implies that the links $L(v_1), \ldots, L(v_s)$ are pairwise edge disjoint. Since $\bigcup_{i=1}^{s} L(v_i) \subset \partial\mathcal{H}$, we have $\sum_{v \in S} d_{\mathcal{H}}(v) \leq |\partial\mathcal{H}|$. ∎

3.5.1 Proof of Theorem 3.1.16

In this section we will prove Theorem 3.1.16, but instead of proving it directly we will prove the following stronger statement.

Theorem 3.5.4. *Let $r \geq 2$ and let $\mathcal{H}$ be a cancellative r-graph. Then*

$$|\mathcal{H}| \leq \left(\frac{|\partial\mathcal{H}|}{r}\right)^{\frac{r}{r-1}}.$$

First we show that Theorem 3.5.4 implies Theorem 3.1.16.

Proof of Theorem 3.1.16. Let us consider the lower bound first. Let $\alpha \in [0,1]$ and let $\mathcal{H}_n(\alpha)$ be the vertex disjoint union of $T_r(\alpha n, r)$ and a set of $(1-\alpha)n$ isolated vertices. It is clear that $\mathcal{T}_r \not\subset \mathcal{H}_n(\alpha)$. Let

$$x = \lim_{n\to\infty} \frac{|\partial\mathcal{H}_n(\alpha)|}{\binom{n}{r-1}} = \lim_{n\to\infty} \frac{r\left(\alpha n/r\right)^{r-1}}{\binom{n}{r-1}} = \frac{\alpha^{r-1}(r-1)!}{r^{r-2}},$$

and

$$y = \lim_{n\to\infty} \frac{|\mathcal{H}_n(\alpha)|}{\binom{n}{r}} = \lim_{n\to\infty} \frac{(\alpha n/r)^r}{\binom{n}{r}} = \frac{\alpha^r (r-1)!}{r^{r-1}}.$$

Then, $y = (x^r/r!)^{1/(r-1)}$. Letting α vary from 0 to 1, we obtain $g(\mathcal{T}_r, x) \geq (x^r/r!)^{1/(r-1)}$ for all $x \in [0, (r-1)!/r^{r-2}]$.

Next we prove the upper bound. Suppose that $(\mathcal{H}_k)_{k=1}^{\infty}$ is a good sequence of cancellative r-graphs that realizes (x, y). Let $x_k = (r-1)!|\partial\mathcal{H}_k| / (v(\mathcal{H}_k))^{r-1}$ and $y_k = r!|\mathcal{H}_k| / (v(\mathcal{H}_k))^r$ for all $k \geq 1$. Then Theorem 3.5.4 gives

$$\frac{y_k (v(\mathcal{H}_k))^r}{r!} \leq \left(\frac{x_k (v(\mathcal{H}_k))^{r-1}}{r(r-1)!}\right)^{\frac{r}{r-1}},$$

which implies

$$y_k \leq \left(\frac{(x_k)^r}{r!}\right)^{\frac{1}{r-1}}.$$

Letting $k \to \infty$, we obtain $y \leq (x^r/r!)^{1/r-1}$, and this completes the proof. ∎

Now we prove Theorem 3.5.4. We will use the following fact.

Fact 3.5.5. *Let X be a collection of non-negative real numbers and $a \in [0, 1]$. Then*

$$\sum_{x\in X} x^a \leq |X| \left(\frac{\sum_{x\in X} x}{|X|}\right)^a = |X|^{1-a} \left(\sum_{x\in X} x\right)^a, \tag{3.13}$$

and

$$\left(\sum_{x\in X} x\right)^2 \leq |X| \sum_{x\in X} x^2. \tag{3.14}$$

Proof of Theorem 3.5.4. We proceed by induction on r. When $r = 2$, this is just Mantel's theorem, so we may assume that $r \geq 3$.

By Lemma 3.5.1, $L(v)$ is a cancellative $(r-1)$-graph for all $v \in V(\mathcal{H})$. Therefore, by the induction hypothesis,

$$d(v) \leq \left(\frac{|\partial L(v)|}{r-1}\right)^{\frac{r-1}{r-2}}. \tag{3.15}$$

It follows that

$$\begin{aligned} |\mathcal{H}| = \frac{1}{r} \sum_{v \in V(\mathcal{H})} d(v) &= \frac{1}{r} \sum_{v \in V(\mathcal{H})} (d(v))^{\frac{1}{r-1}} (d(v))^{\frac{r-2}{r-1}} \\ &\overset{\text{Equation 3.15}}{\leq} \frac{1}{r(r-1)} \sum_{v \in V(\mathcal{H})} (d(v))^{\frac{1}{r-1}} |\partial L(v)|. \end{aligned} \tag{3.16}$$

Notice that

$$\begin{aligned} \sum_{v \in V(\mathcal{H})} (d(v))^{\frac{1}{r-1}} |\partial L(v)| &= \sum_{v \in V(\mathcal{H})} \sum_{\substack{S \in \partial\mathcal{H} \\ v \in S}} (d(v))^{\frac{1}{r-1}} \\ &= \sum_{S \in \partial\mathcal{H}} \sum_{v \in S} (d(v))^{\frac{1}{r-1}} \\ &\overset{\text{Equation 3.13}}{\leq} ((r-1)|\partial\mathcal{H}|)^{\frac{r-2}{r-1}} \left(\sum_{S \in \partial\mathcal{H}} \sum_{v \in S} d(v)\right)^{\frac{1}{r-1}} \\ &= ((r-1)|\partial\mathcal{H}|)^{\frac{r-2}{r-1}} \left(\sum_{S \in \partial\mathcal{H}} \sigma(S)\right)^{\frac{1}{r-1}}. \end{aligned} \tag{3.17}$$

Define $\hat{\sigma} = \max\{\sigma(H) : H \in \mathcal{H}\}$ and suppose that $E \in \mathcal{H}$ satisfies $\sum_{v\in E} d(v) = \hat{\sigma}$. Then,

$$\begin{aligned}
\sum_{S\in\partial\mathcal{H}} \sigma(S) &= \sum_{S\in\bigcup_{v\in E} L(v)} \sigma(S) + \sum_{S\in\partial\mathcal{H}\setminus\bigcup_{v\in E} L(v)} \sigma(S) \\
&\overset{\text{Lemma 3.5.2}}{=} \sum_{v\in E}\sum_{S\in L(v)} \sigma(S) + \sum_{S\in\partial\mathcal{H}\setminus\bigcup_{v\in E} L(v)} \sigma(S) \\
&\leq \sum_{v\in E} d(v)\left(\hat{\sigma} - d(v)\right) + \left(|\partial\mathcal{H}| - \hat{\sigma}\right)\hat{\sigma} \\
&\overset{\text{Equation 3.14}}{\leq} \left(\sum_{v\in E} d(v)\right)\left(\hat{\sigma} - \frac{\sum_{v\in E} d(v)}{r}\right) + \left(|\partial\mathcal{H}| - \hat{\sigma}\right)\hat{\sigma} \\
&= \hat{\sigma}\left(\hat{\sigma} - \frac{\hat{\sigma}}{r}\right) + \left(|\partial\mathcal{H}| - \hat{\sigma}\right)\hat{\sigma} \\
&= \left(|\partial\mathcal{H}| - \frac{\hat{\sigma}}{r}\right)\hat{\sigma}. \qquad (3.18)
\end{aligned}$$

Note that Corollary 3.5.3 gives $\hat{\sigma} \leq |\partial\mathcal{H}|$. On the other hand, since $\left(|\partial\mathcal{H}| - \hat{\sigma}/r\right)\hat{\sigma}$ is increasing in $\hat{\sigma}$ when $\hat{\sigma} \leq r|\partial\mathcal{H}|/2$, it follows from Equation 3.18 and $r \geq 3$ that

$$\sum_{S\in\partial\mathcal{H}} \sigma(S) \leq \left(|\partial\mathcal{H}| - \frac{\hat{\sigma}}{r}\right)\hat{\sigma} \leq \frac{r-1}{r}|\partial\mathcal{H}|^2. \qquad (3.19)$$

Plugging Equation 3.17 and Equation 3.19 into Equation 3.16, we obtain

$$|\mathcal{H}| \leq \frac{1}{r(r-1)}\left((r-1)|\partial\mathcal{H}|\right)^{\frac{r-2}{r-1}}\left(\frac{r-1}{r}|\partial\mathcal{H}|^2\right)^{\frac{1}{r-1}} = \left(\frac{|\partial\mathcal{H}|}{r}\right)^{\frac{r}{r-1}},$$

and this completes the proof. ∎

3.5.2 Proof of Theorem 3.1.17

In this section we will prove Theorem 3.1.17. As before, we will prove a stronger statement which implies Theorem 3.1.17.

Theorem 3.5.6. *Suppose that $\mathcal{H}$ is a cancellative 3-graph on n vertices. Then*

$$|\mathcal{H}| \leq \frac{\left(n^2 - 2|\partial\mathcal{H}|\right)|\partial\mathcal{H}|}{3n} + 3n^2.$$

First we show that Theorem 3.5.6 implies Theorem 3.1.17.

Proof of Theorem 3.1.17. Let us consider the lower bound first. Recall that a k-vertex Steiner triple system (STS for short) is a 3-graph on k vertices such that every pair of vertices is covered by exactly one edge. It is known that a k-vertex STS exists iff $k \equiv 1$ or $3 \pmod 6$ (e.g. see [243]). Let STS(k) denote the family of all Steiner triple systems on k vertices. Let $\mathcal{S}(n,k)$ denote the collection of all 3-graphs on n vertices that can be obtained from a 3-graph $H \in STS(k)$ by blowing up every vertex in H into a set of size either $\lfloor n/k \rfloor$ or $\lceil n/k \rceil$. It is easy to see that every 3-graph in $\mathcal{S}(n,k)$ is cancellative.

Fix an integer k with $k \equiv 1$ or $3 \pmod 6$. Let $\mathcal{H}_n \in \mathcal{S}(n,k)$ and in order to keep the calculations simple let us assume that k divides n. Then

$$\lim_{n\to\infty} \frac{|\partial\mathcal{H}_n|}{\binom{n}{2}} = \frac{(k-1)n^2/(2k)}{\binom{n}{2}} = \frac{k-1}{k},$$

and

$$\lim_{n\to\infty} \frac{|\mathcal{H}_n|}{\binom{n}{3}} = \frac{(k-1)n^3/(6k^2)}{\binom{n}{3}} = \frac{k-1}{k^2}.$$

Therefore, the sequence $(\mathcal{H}_n)_{n=1}^{\infty}$ realizes $\big((k-1)/k, (k-1)/k^2\big)$. So, $g(\mathcal{T}_3, (k-1)/k) \geq (k-1)/k^2$ for all integers k with $k \equiv 1$ or $3 \pmod 6$.

Next we prove the upper bound. Let $(\mathcal{H}_k)_{k=1}^{\infty}$ be a good sequence of cancellative 3-graphs that realizes (x, y). Let $x_k = 2|\partial\mathcal{H}_k|/\left(v(\mathcal{H})\right)^2$ and $y_k = 6|\mathcal{H}_k|/\left(v(\mathcal{H})\right)^3$ for $k \geq 1$. Then, it follows from Theorem 3.5.6 that

$$\frac{y_k \left(v(\mathcal{H}_k)\right)^3}{6} \leq \frac{\left(\left(v(\mathcal{H}_k)\right)^2 - x_k \left(v(\mathcal{H}_k)\right)^2\right) x_k \left(v(\mathcal{H}_k)\right)^2/2}{3v(\mathcal{H}_k)} + 3\left(v(\mathcal{H}_k)\right)^2,$$

which implies

$$y_k \leq x_k(1-x_k) + \frac{18}{v(\mathcal{H}_k)}.$$

Letting $k \to \infty$, we obtain $y \leq x(1-x)$, and this completes the proof. ∎

The idea of the proof of Theorem 3.5.6 is to first choose $S \subset V(\mathcal{H})$ such that $(\partial\mathcal{H})[S]$ is a clique. Then we apply the induction hypothesis to $V(\mathcal{H}) \setminus S$. However, in order to do the induction we need to prove a stronger statement which implies Theorem 3.5.6.

We will use G to denote the graph $\partial\mathcal{H}$. Let $U \subset V(\mathcal{H})$ and let $G_U = G[U]$ and $\mathcal{H}_U = \mathcal{H}[U]$.

Theorem 3.5.7. *Let $\mathcal{H}$ be a cancellative 3-graph on n vertices. Let $U \subset V(\mathcal{H})$ be a set of size m. Suppose that $|G_U| = xm^2/2$ for some real number x with $0 \leq x \leq (m-1)/m$. Then,*

$$|\mathcal{H}_U| \leq \frac{(1-x)x}{6}m^3 + 3m^2.$$

In particular, letting $U = V(\mathcal{H})$ in Theorem 3.5.7 we obtain

$$|\mathcal{H}| \leq \frac{\left(n^2 - 2|\partial\mathcal{H}|\right)|\partial\mathcal{H}|}{3n} + 3n^2,$$

which is exactly Theorem 3.5.6.

The proof of Theorem 3.5.7 is by induction on m. Note that Theorem 3.5.7 holds trivially for all $m \leq 20$ since $\binom{m}{3} \leq 3m^2$ for all $m \leq 20$. Also, by Theorem 3.5.4,

$$|\mathcal{H}_U| \leq \frac{|\partial\left(\mathcal{H}_U\right)|^{3/2}}{3\sqrt{3}} \leq \frac{|G_U|^{3/2}}{3\sqrt{3}} = \frac{x^{3/2}}{6\sqrt{6}}m^3,$$

which is less than $x(1-x)m^3/6 + 3m^2$ when $x \leq 2/3$. Therefore, Theorem 3.5.7 is true for all $x \leq 2/3$, and hence we may assume that $x > 2/3$ in the rest of the proof.

In the proof of Theorem 3.5.7 we need the following version of Turán's theorem. The clique number $\omega(G)$ of a graph G is the largest integer ω such that there is a copy of K_ω in G. Turán's theorem implies that any n-vertex graph with no $K_{\omega+1}$ has at most $(\omega-1)n^2/(2\omega)$ edges.

Theorem 3.5.8 ([241])**.** *Let G be an n-vertex graph with at least $xn^2/2$ edges for some real number $x \geq 0$. Then $\omega(G) \geq \lceil 1/(1-x) \rceil$.*

Proof. Let $\omega = \omega(G)$. By Turán's theorem, $xn^2/2 \le (\omega-1)n^2/(2\omega)$. Simplifying this inequality we obtain $\omega \ge 1/(1-x)$. Since ω is an integer, $\omega \ge \lceil 1/(1-x) \rceil$. ∎

The idea in the proof of Theorem 3.5.7 is to first apply Turán's theorem on G_U to find a large clique, say on S, and then apply the induction hypothesis to $T = U \setminus S$ to get an upper bound for $|\mathcal{H}_T|$. In order to get an upper bound for $|\mathcal{H}_U|$ we just need to apply Corollary 3.5.3 to $\mathcal{H}_U$ to get an upper bound for $|\mathcal{H}_U \setminus \mathcal{H}_T|$.

Proof of Theorem 3.5.7. Suppose that G_U contains a clique on ω vertices. We may assume that $\omega < m$ since otherwise by Corollary 3.5.3, we are done. Choose $S \subset U$ of size ω so that $G_S \cong K_\omega$. Let $T = U \setminus S$. Let e_s denote the number of edges in G_U that have nonempty intersection with S. Let $x' = \frac{xm^2 - 2e_s}{(m-\omega)^2}$.

First, notice a simple but crucial fact is that every vertex in T is adjacent to at most $\omega - 1$ vertices in S, since otherwise there would be a copy of $K_{\omega+1}$ in G_U, which contradicts the definition of ω. Therefore,

$$e_s \le (\omega - 1)(m - \omega) + \binom{\omega}{2}. \tag{3.20}$$

Applying the induction hypothesis to T we obtain

$$|\mathcal{H}_T| \le \frac{x'(1-x')}{6}(m-\omega)^3 + 3(m-\omega)^2.$$

On the other hand, Corollary 3.5.3 gives

$$|\mathcal{H}_U \setminus \mathcal{H}_T| \leq \sum_{v \in S} d(v) \leq |G_U| = \frac{x}{2}m^2.$$

Therefore,

$$|\mathcal{H}_U| = |\mathcal{H}_T| + |\mathcal{H}_U \setminus \mathcal{H}_T| \leq \frac{x'(1-x')}{6}(m-\omega)^3 + 3(m-\omega)^2 + \frac{x}{2}m^2. \tag{3.21}$$

Claim 3.5.9. *For $2/3 \leq x \leq 1$ and $0 \leq x' \leq 1$ we have*

$$\frac{x(1-x)}{6}m^3 + 3m^2 \geq \frac{x'(1-x')}{6}(m-\omega)^3 + 3(m-\omega)^2 + \frac{x}{2}m^2.$$

Proof. Notice that

$$\begin{aligned}(m-\omega)^2(x-x') = x\left((m-\omega)^2 - m^2\right) + 2e_s &\overset{\text{Equation 3.20}}{\leq} x\omega(\omega-2m) + (\omega-1)(2m-\omega) \\ &= (2m-\omega)\left(\omega(1-x)-1\right).\end{aligned}$$

Consequently,

$$\begin{aligned}(m-\omega)^2\left(x'(1-x') - x(1-x)\right) &= (m-\omega)^2(x-x')(x+x'-1) \\ &\leq (2m-\omega)\left(\omega(1-x)-1\right)x. \end{aligned} \tag{3.22}$$

Indeed, if $x \geq x'$ this follows from the previous estimate and $x + x' - 1 \leq x$. If $x < x'$, then $x + x' - 1 \geq 1/3$ and the left side of Equation 3.22 is negative, while the right side of Equation 3.22 is nonnegative. Multiplying Equation 3.22 by $m - \omega$ and taking the identity $\omega(m-\omega)(2m-\omega) = m^3 - (n-\omega)^3 - m^2\omega$ into account we obtain

$$\begin{aligned}(m-\omega)^3 &\left(x'(1-x') - x(1-x)\right) \\ &\leq \left(m^3 - (m-\omega)^3\right) x(1-x) - m^2\omega x(1-x) - x(m-\omega)(2m-\omega),\end{aligned}$$

which due to $\omega(1-x) \geq 1$ implies

$$(m-\omega)^3 x'(1-x') \leq m^3 x(1-x) - x\left(m^2 + (m-\omega)(2m-\omega)\right).$$

Adding $3xm^2$ on both sides and using

$$2m^2 - (m-\omega)(2m-\omega) = 3m\omega - \omega^2 \leq 18\left(m^2 - (m-\omega)^2\right)$$

we reach

$$(m-\omega)^3 x'(1-x') + 3xm^2 \leq m^3 x(1-x) + 18x\left(m^2 - (m-\omega)^2\right).$$

Due to $x \leq 1$ this implies the claim. ∎

Finally, $|\mathcal{H}_U| \le x(1-x)m^3/6 + 3m^2$ is an immediate consequence of Claim 3.5.9 and Equation 3.21 and this completes the proof of the theorem. ∎

3.6 Hypergraphs without the expansion of cliques

In this section we consider the feasible region of hypergraphs without expansion of cliques. First we will prove the following result, from which Theorem 3.1.21 can be easily obtained.

Theorem 3.6.1. *Let $\ell \geq r \geq 2$. Let $\mathcal{H}$ be a $\mathcal{K}_{\ell+1}^{r}$-free r-graph. Then*

$$\left(\frac{|\mathcal{H}|}{\binom{\ell}{r}}\right)^{1/r} \leq \left(\frac{|\partial \mathcal{H}|}{\binom{\ell}{r-1}}\right)^{1/(r-1)}.$$

In order to derive Theorem 3.1.21 from Theorem 3.6.1 we need an easy observation.

Observation 3.6.2. *Let $r \geq 3$ and $\mathcal{H}$ be an r-graph. If $0 \leq i \leq r-2$, then $\mathcal{H}$ is $\mathcal{K}_{\ell+1}^{r}$-free iff $\partial_i \mathcal{H}$ is $K_{\ell+1}^{r-i}$-free. In particular, $\mathcal{H}$ is $\mathcal{K}_{\ell+1}^{r}$-free iff $\partial_{r-2}\mathcal{H}$ is $K_{\ell+1}$-free. If $i \leq -1$, then $\mathcal{H}$ is $\mathcal{K}_{\ell+1}^{r}$-free implies that $\partial_i \mathcal{H}$ is $\mathcal{K}_{\ell+1}^{r-i}$-free.*

Now we show how to prove Theorem 3.1.21 using Theorem 3.6.1.

Proof of Theorem 3.1.21. Fix $r-\ell \leq i \leq r-2$. Then by Observation 3.6.2, $\partial_i \mathcal{H}$ is $\mathcal{K}_{\ell+1}^{r-i}$-free. Since $\partial\left(\partial_i \mathcal{H}\right) \subset \partial_{i+1}\mathcal{H}$, it follows from Theorem 3.6.1 that

$$\left(\frac{|\partial_i \mathcal{H}|}{\binom{\ell}{r-i}}\right)^{1/(r-i)} \leq \left(\frac{|\partial(\partial_i \mathcal{H})|}{\binom{\ell}{r-i-1}}\right)^{1/(r-i-1)} \leq \left(\frac{|\partial_{i+1} \mathcal{H}|}{\binom{\ell}{r-i-1}}\right)^{1/(r-i-1)},$$

and this completes the proof. ∎

To show that all inequalities in Theorem 3.1.21 are tight, consider the following construction. Fix $\alpha \in [0,1]$ and let $\mathcal{H}_n(\alpha)$ be the vertex disjoint union of $T_r(\alpha n, \ell)$ and a set of $(1-\alpha)n$

isolated vertices. It is clear that $\mathcal{H}_n(\alpha)$ is $\mathcal{K}^r_{\ell+1}$-free. In order to keep the calculations simple, let us assume that αn is an integer that is a multiple of ℓ. For fixed $\ell - r \le i \le r-1$,

$$|\partial_i \mathcal{H}_n(\alpha)| = \binom{\ell}{r-i}\left(\frac{\alpha n}{\ell}\right)^{r-i},$$

and hence

$$\left(\frac{|\partial_i \mathcal{H}_n(\alpha)|}{\binom{\ell}{r-i}}\right)^{\frac{1}{r-i}} = \frac{\alpha n}{\ell}.$$

Therefore, all inequalities in Theorem 3.1.21 are tight.

Notice that the construction above also proves the lower bound in Corollary 3.1.22 and we omit the calculations here.

The proof of Theorem 3.6.1 uses some ideas in Fisher and Ryan's proof [83]. However we need to translate their proof into the language of hypergraphs, since an edge in $\partial_i\mathcal{H}$ might not be equivalent to a copy of K_{r-i} in $\partial_{r-2}\mathcal{H}$ for $-\ell \le i \le r-3$. Define the clique set $\mathcal{K}_{\mathcal{H}}$ of $\mathcal{H}$ as

$$\mathcal{K}_{\mathcal{H}} = \left\{A \subset V(\mathcal{H})\colon (\partial_{r-2}\mathcal{H})[A] \cong K_{|A|}\right\}.$$

For every $E \in \partial\mathcal{H}$ let $N(E) = \{v \in V(\mathcal{H})\colon \{v\} \cup E \in \mathcal{H}\}$. Recall that $\sigma(S) = \sum_{v\in S} d(v)$. We first prove a lemma that will be used in the proof of Theorem 3.6.1.

Lemma 3.6.3. $\sum_{E\in\partial\mathcal{H}} \sigma(E) \le \frac{(\ell-r+1)(r-1)}{\ell}|\partial\mathcal{H}|^2$.

Proof. Let $S \subset V(\mathcal{H})$. For every $v \in V(\mathcal{H})$ we have $d(v) = \sum_{E\in\partial\mathcal{H}} |N(E) \cap \{v\}|$. So,

$$\sigma(S) = \sum_{v\in S} d(v) = \sum_{v\in S} \sum_{E\in\partial\mathcal{H}} |N(E) \cap \{v\}| = \sum_{E\in\partial\mathcal{H}} |N(E) \cap S|. \tag{3.23}$$

On the other hand,

$$\begin{aligned}(\sigma(S))^2 = \left(\sum_{v\in S} d(v)\right)^2 &\overset{\text{Equation 3.14}}{\leq} |S| \sum_{v\in S} (d(v))^2 = |S| \sum_{v\in S} \sum_{E\in L(v)} d(v) \\ &= |S| \sum_{v\in S} \sum_{\substack{E\in\partial\mathcal{H} \\ v\in N(E)}} d(v) = |S| \sum_{E\in\partial\mathcal{H}} \sum_{v\in S\cap N(E)} d(v) \\ &= |S| \sum_{E\in\partial\mathcal{H}} \sigma\left(N(E) \cap S\right),\end{aligned}$$

which implies

$$\sum_{E\in\partial\mathcal{H}} \sigma\left(N(E) \cap S\right) \geq \frac{(\sigma(S))^2}{|S|}. \tag{3.24}$$

Now suppose that $S \in \mathcal{K}_{\mathcal{H}}$. Since $\mathcal{H}$ is $\mathcal{K}_{\ell+1}^r$-free, $|E| + |N(E) \cap S| \leq \ell$ for all $E \in \partial\mathcal{H}$. It follows from Equation 3.23 that

$$\sigma(S) = \sum_{T\in\partial\mathcal{H}} |N(T) \cap S| \leq (\ell - r + 1)|\partial\mathcal{H}|. \tag{3.25}$$

Let z be the largest real number such that $\sigma(R) \leq (\ell - r + 1)|\partial\mathcal{H}| - (\ell - |R|)z$ for all $R \in \mathcal{K}_{\mathcal{H}}$. Let $R_0 \in \mathcal{K}_{\mathcal{H}}$ such that

$$\sigma(R_0) = (\ell - r + 1)|\partial\mathcal{H}| - (\ell - |R_0|)z. \tag{3.26}$$

For every $E \in \partial\mathcal{H}$, $E \cup (N(E) \cap R_0) \in \mathcal{K}_{\mathcal{H}}$, therefore,

$$\begin{aligned}
\sum_{E\in\partial\mathcal{H}} \sigma(E) &= \sum_{E\in\partial\mathcal{H}} \left(\sigma(E \cup (N(E) \cap R_0)) - \sigma(N(E) \cap R_0)\right) \\
&\leq \sum_{E\in\partial\mathcal{H}} \left((\ell - r + 1)|\partial\mathcal{H}| - (\ell - |E \cup (N(E) \cap R_0)|)z - \sigma(N(E) \cap R_0)\right) \\
&\leq \sum_{E\in\partial\mathcal{H}} \left((\ell - r + 1)\left(|\partial\mathcal{H}| - z\right) + |N(E) \cap R_0|z - \sigma(N(E) \cap R_0)\right) \\
&= (\ell - r + 1)(|\partial\mathcal{H}| - z)|\partial\mathcal{H}| + z \sum_{E\in\partial\mathcal{H}} |N(E) \cap R_0| - \sum_{E\in\partial\mathcal{H}} \sigma(N(E) \cap R_0) \\
&\overset{\text{Equation 3.24, Equation 3.25}}{\leq} (\ell - r + 1)(|\partial\mathcal{H}| - z)|\partial\mathcal{H}| + z\sigma(R_0) - \frac{(\sigma(R_0))^2}{|R_0|} \\
&\overset{\text{Equation 3.26}}{=} (\ell - r + 1)(|\partial\mathcal{H}| - 2z)|\partial\mathcal{H}| + z^2\ell - \frac{((\ell - r + 1)|\partial\mathcal{H}| - z\ell)^2}{|R_0|}.
\end{aligned} \tag{3.27}$$

Since $|R_0| \leq \ell$, we may plug $|R_0| = \ell$ into Equation 3.27 and z will be cancelled in the calculation and hence

$$\sum_{E\in\partial\mathcal{H}} \sigma(E) \leq \frac{(\ell - r + 1)(r - 1)}{\ell}|\partial\mathcal{H}|^2.$$

∎

Now we are ready to prove Theorem 3.6.1.

Proof of Theorem 3.6.1. We proceed by induction on r. The case $r = 2$ is just Turán's theorem, so we may assume that $r \geq 3$.

For every $v \in V(\mathcal{H})$ the link $L(v)$ is a $\mathcal{K}_\ell^{r-1}$-free $(r-1)$-graph, therefore, by the induction hypothesis,

$$d(v) \leq \binom{\ell-1}{r-1} \left(\frac{|\partial L(v)|}{\binom{\ell-1}{r-2}} \right)^{\frac{r-1}{r-2}}. \tag{3.28}$$

It follows that

$$\begin{aligned} |\mathcal{H}| = \frac{1}{r} \sum_{v \in V(\mathcal{H})} d(v) &= \frac{1}{r} \sum_{v \in V(\mathcal{H})} (d(v))^{\frac{1}{r-1}} (d(v))^{\frac{r-2}{r-1}} \\ &\overset{\text{Equation 3.28}}{\leq} \frac{\binom{\ell-1}{r-1}^{\frac{r-2}{r-1}}}{r\binom{\ell-1}{r-2}} \sum_{v \in V(\mathcal{H})} (d(v))^{\frac{1}{r-1}} |\partial L(v)|. \end{aligned} \tag{3.29}$$

Similar to Equation 3.17, we have

$$\begin{aligned} \sum_{v \in V(\mathcal{H})} (d(v))^{\frac{1}{r-1}} |\partial L(v)| &= \sum_{E \in \partial\mathcal{H}} \sum_{v \in E} (d(v))^{\frac{1}{r-1}} \\ &\overset{\text{Equation 3.13}}{\leq} ((r-1)|\partial\mathcal{H}|)^{\frac{r-2}{r-1}} \left(\sum_{E \in \partial\mathcal{H}} \sum_{v \in E} d(v) \right)^{\frac{1}{r-1}} \\ &= ((r-1)|\partial\mathcal{H}|)^{\frac{r-2}{r-1}} \left(\sum_{E \in \partial\mathcal{H}} \sigma(E) \right)^{\frac{1}{r-1}} \\ &\overset{\text{Lemma 3.6.3}}{\leq} (r-1) \left(\frac{\ell - r + 1}{\ell} \right)^{\frac{1}{r-1}} |\partial\mathcal{H}|^{\frac{r}{r-1}}. \end{aligned} \tag{3.30}$$

It follows from Equation 3.29 and Equation 3.30 that

$$|\mathcal{H}| \le \binom{\ell}{r}\left(\frac{|\partial\mathcal{H}|}{\binom{\ell}{r-1}}\right)^{\frac{r}{r-1}}.$$

∎

Now we show how to prove Corollary 3.1.22 using Theorem 3.1.21.

Proof of Corollary 3.1.22. Let $(\mathcal{H}_k)_{k=1}^{\infty}$ be a good sequence of $\mathcal{K}_{\ell+1}^r$-free r-graphs that realizes (x, y). Let $x_k = (r-1)!|\partial\mathcal{H}_k|/(v(\mathcal{H}_k))^{r-1}$ and $y_k = r!|\mathcal{H}_k|/(v(\mathcal{H}_k))^r$. First, we show that $\text{proj}\Omega(\mathcal{K}_{\ell+1}^r) = [0, (\ell)_{r-1}/\ell^{r-1}]$.

It follows from Theorem 3.1.21 that

$$\frac{x_k\,(v(\mathcal{H}_k))^{r-1}}{(r-1)!} \le \binom{\ell}{r-1}\left(\frac{v(\mathcal{H}_k)}{\ell}\right)^{r-1},$$

which implies $x_k \le (\ell)_{r-1}/\ell^{r-1}$. Letting $k \to \infty$, we obtain $x \le (\ell)_{r-1}/\ell^{r-1}$. Therefore, $\text{proj}\Omega(\mathcal{K}_{\ell+1}^r) \subset [0, (\ell)_{r-1}/\ell^{r-1}]$. On the other hand, $(T_r(k,\ell))_{k=1}^{\infty}$ shows that $(\ell)_{r-1}/\ell^{r-1} \in \text{proj}\Omega(\mathcal{K}_{\ell+1}^r)$ and it follows from Observation 3.1.5 that $\text{proj}\Omega(\mathcal{K}_{\ell+1}^r) = [0, (\ell)_{r-1}/\ell^{r-1}]$.

Next, we show the upper bound for $g(\mathcal{K}_{\ell+1}^r, x)$. It follows from Theorem 3.1.21 that

$$\left(\frac{y_k\,(v(\mathcal{H}_k))^r}{r!\binom{\ell}{r}}\right)^{\frac{1}{r}} \le \left(\frac{x_k\,(v(\mathcal{H}_k))^{r-1}}{(r-1)!\binom{\ell}{r-1}}\right)^{\frac{1}{r-1}},$$

which implies $y_k \le (\ell - r + 1)\,(x_k^r/(\ell)_r)^{1/(r-1)}$. Letting $k \to \infty$, we obtain $y \le (\ell - r + 1)\,(x^r/(\ell)_r)^{1/(r-1)}$. Therefore, $g(\mathcal{K}_{\ell+1}^r, x) \le (\ell - r + 1)\,(x^r/(\ell)_r)^{1/(r-1)}$ for all $x \in \text{proj}\Omega(\mathcal{K}_{\ell+1}^r)$.

The construction for the lower bound is exactly the same as the construction for Theorem 3.1.21, and it shows that $g(\mathcal{K}^r_{\ell+1}, x) \ge (\ell - r + 1)\,(x^r/(\ell)_r)^{1/(r-1)}$ for all $x \in \mathrm{proj}\Omega(\mathcal{K}^r_{\ell+1})$. Therefore, $g(\mathcal{K}^r_{\ell+1}, x) = (\ell - r + 1)\,(x^r/(\ell)_r)^{1/(r-1)}$ for all $x \in \mathrm{proj}\Omega(\mathcal{K}^r_{\ell+1})$. ∎

Let us present a lemma before proving Theorem 3.1.23.

Lemma 3.6.4. *Let $r \ge 3$ and $\mathcal{F}_1, \mathcal{F}_2$ be two families of r-graphs with $\mathcal{F}_1 \subset \mathcal{F}_2$. Suppose that every n-vertex $\mathcal{F}_1$-free r-graph can be made $\mathcal{F}_2$-free by removing at most $o(n^r)$ edges, and $g(\mathcal{F}_2, x)$ is increasing on $[0, c]$ for some $c > 0$. Then $g(\mathcal{F}_1, x) = g(\mathcal{F}_2, x)$ on $[0, c]$.*

Proof. Since $\mathcal{F}_1 \subset \mathcal{F}_2$, it follows from Observation 3.3.2 that $g(\mathcal{F}_2, x) \le g(\mathcal{F}_1, x)$ for all $x \in \mathrm{proj}\Omega(\mathcal{F}_2)$. So it suffices to show that $g(\mathcal{F}_2, x) \ge g(\mathcal{F}_1, x)$ for all $x \in [0, c]$. Let $(x_0, y_0) \in \Omega(\mathcal{F}_1)$ with $x_0 \in [0, c]$ and $y_0 = g(\mathcal{F}_1, x_0)$. By definition, there exists a sequence of $\mathcal{F}_1$-free r-graphs $(\mathcal{H}_k)_{k=1}^{\infty}$ with $\lim_{k\to\infty} d(\partial\mathcal{H}_k) = x_0$ and $\lim_{k\to\infty} d(\mathcal{H}_k) = y_0$.

For every $k \ge 1$ let $\mathcal{H}'_k$ be a subgraph of $\mathcal{H}_k$ that is $\mathcal{F}_2$-free and of maximum size, and let $x'_k = d(\partial\mathcal{H}'_k)$ and $y'_k = d(\mathcal{H}'_k)$. By the Bolzano–Weierstrass theorem, $(x'_k, y'_k)_{k=1}^{\infty}$ contains a convergent subsequence $\left(x'_{t_k}, y'_{t_k}\right)_{k=1}^{\infty}$. Let $x'_0 = \lim_{k\to\infty} x'_{t_k}$ and $y'_0 = \lim_{k\to\infty} y'_{t_k}$, and it is easy to see from the definition of $\mathcal{H}'_k$ that $x'_0 \le x_0$ and $y'_0 \le y_0$. Since $\left(\mathcal{H}'_{t_k}\right)_{k=1}^{\infty}$ is a good sequence of $\mathcal{F}_2$-free r-graphs that realizes (x'_0, y'_0), we obtain $(x'_0, y'_0) \in \Omega(\mathcal{F}_2)$.

By assumption, for every $\epsilon > 0$ there exists $n(\epsilon)$ such that $\mathcal{H}_k$ can be made $\mathcal{F}_2$-free by removing at most $\epsilon\,(v(\mathcal{H}_k))^r$ edges whenever $v(\mathcal{H}_k) \ge n(\epsilon)$. Since $\lim_{k\to\infty} v(\mathcal{H}_k) = \infty$, there exists $k(\epsilon)$ such that $v(\mathcal{H}_k) \ge n(\epsilon)$ for all $k \ge k(\epsilon)$, and hence $|\mathcal{H}'_k| \ge |\mathcal{H}_k| - \epsilon\,(v(\mathcal{H}_k))^r$ for all

$k \geq k(\epsilon)$. Therefore, $y_0' \geq y_0 - r!\epsilon$. Letting $\epsilon \to 0$, we obtain $y_0' \geq y_0$, and hence $y_0' = y_0$. Therefore, $(x_0', y_0) \in \Omega(\mathcal{F}_2)$. By the assumption that $g(\mathcal{F}_2)$ is increasing on $[0, c]$, we obtain

$$g(\mathcal{F}_2, x_0) \geq g(\mathcal{F}_2, x_0') \geq y_0 = g(\mathcal{F}_1, x_0).$$

Since x_0 was chosen arbitrarily from $[0, c]$, $g(\mathcal{F}_2, x) \geq g(\mathcal{F}_1, x)$ for all $x \in [0, c]$, and this completes the proof. ∎

Now we prove Theorem 3.1.23 using Corollary 3.1.22.

Proof of Theorem 3.1.23. It was shown by Pikhurko (see the proof of Lemma 3 in [209]) that every $H_{\ell+1}^r$-free r-graph on n-vertices can be made $\mathcal{K}_{\ell+1}^r$-free by removing at most $o(n^r)$ edges. On the other hand, Corollary 3.1.22 shows that $g(\mathcal{K}_{\ell+1}^r)$ is increasing on $[0, (\ell)_{r-1}/\ell^{r-1}]$. So, it follows from Lemma 3.6.4 that

$$g(H_{\ell+1}^r, x) = g(\mathcal{K}_{\ell+1}^r, x) = (\ell - r + 1)\left(\frac{x^r}{(\ell)_r}\right)^{\frac{1}{r-1}}$$

for all $x \in [0, (\ell)_{r-1}/\ell^{r-1}]$. ∎

3.7 Countably many local maxima

In this section we prove Theorem 3.1.26. Let us present some preliminary results first.

3.7.1 Preliminaries

For a pair of vertices $u, v \in V(\mathcal{H})$ the neighborhood of uv (we use uv as a shorthand for $\{u, v\}$) is

$$N_{\mathcal{H}}(uv) = \{w \in V(\mathcal{H}) \setminus \{u, v\} \colon \exists A \in \mathcal{H} \text{ such that } \{u, v, w\} \subset A\},$$

and the size of $N_{\mathcal{H}}(uv)$ is called the codegree of uv. Denote by $\Delta_2(\mathcal{H})$ and $\delta_2(\mathcal{H})$ the maximum codegree and the minimum codegree of $\mathcal{H}$, respectively.

For a graph G the clique number $\omega(G)$ of G is the largest integer ω such that $K_\omega \subset G$.

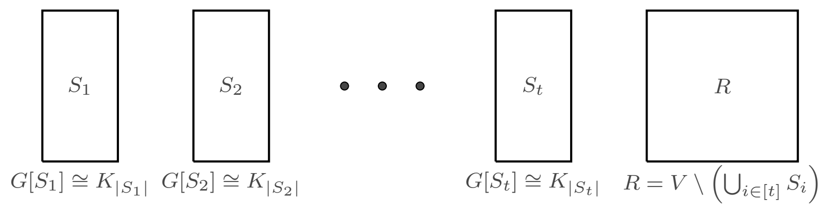


Figure 12. A clique expansion of graph G.

Definition 3.7.1 (Clique expansion)**.** *Let $t \geq 1$, $\kappa \geq 1$ be positive integers and G be a graph on the set V.*

(a) *A $(t+1)$-tuple $(S_1, \dots, S_t, R)$, where $S_1, \dots, S_t$ are pairwise disjoint subsets of $V(G)$ and $R = V \setminus (S_1 \cup \dots \cup S_t)$, is a clique expansion of G if $G[S_i]$ is a complete graph for every $i \in [t]$ (see Figure 12).*

(b) *A clique expansion $(S_1, \dots, S_t, R)$ is maximal if the size of S_i equals the clique number of the induced subgraph of G on $V \setminus (S_1 \cup \dots \cup S_{i-1})$ for $i \in [t]$.*

(c) *We say a clique expansion $(S_1, \dots, S_t, R)$ has a threshold κ if $|S_i| \geq \kappa$ for every $i \in [t]$ but $\omega(G[R]) < \kappa$.*

The following observation is immediate from the definition.

Observation 3.7.2. *Suppose that $(S_1, \dots, S_t, R)$ is a maximal clique expansion of G. Then*

(a) *$|S_1| \geq \dots \geq |S_t|$, and*

(b) *every vertex in $V \setminus (S_1 \cup \dots \cup S_i)$ is adjacent to at most $|S_i| - 1$ vertices in S_i for $i \in [t]$.*

We will need the following classical result due to Andrásfai, Erdős, and Sós [11].

Theorem 3.7.3 (Andrásfai–Erdős–Sós [11])**.** *Let $k \geq 2$ and $n \geq 1$ be positive integers. Then every K_{k+1}-free graph on n vertices with minimum degree greater than $\frac{3k-4}{3k-1}n$ is a k-partite graph.*

Sometimes it will be convenient to consider $\mathcal{H}$ and $\partial\mathcal{H}$ separately.

Definition 3.7.4 (Cancellative pair)**.** *Let G be a graph on V and $\mathcal{H}$ be a 3-graph on the same vertex set V. We say the pair $(G, \mathcal{H})$ is cancellative if $\partial\mathcal{H} \subset G$ and it does not contain three distinct sets $A, B \in \mathcal{H}$ and $C \in G \cup \mathcal{H}$ such that $A \triangle B \subset C$. We call V the vertex set of the pair $(G, \mathcal{H})$.*

Let $\mathcal{H}$ be a cancellative 3-graph and $U \subset V(\mathcal{H})$. Then it is easy to see that the pair $((\partial\mathcal{H})[U], \mathcal{H}[U])$ is cancellative (note that $(\partial\mathcal{H})[U]$ and $\partial(\mathcal{H}[U])$ are not necessarily the same).

Observation 3.7.5. *Suppose that $(G, \mathcal{H})$ is a cancellative pair. Then $N_{\mathcal{H}}(uv)$ is an independent set in G for every $uv \in \partial\mathcal{H}$.*

The following results concerning cancellative pairs were proved in Section 3.5.

Theorem 3.7.6. *Let $m \geq 1$ be an integer and $(G, \mathcal{H})$ be a cancellative pair on a set V of size m. Suppose that $|G| = xm^2/2$ for some real number $x \in [0, 1]$. Then*

$$|\mathcal{H}| \leq \frac{x(1-x)}{6}m^3 + 3m^2.$$

Lemma 3.7.7. *Let $(G, \mathcal{H})$ be a cancellative pair on a set V. Suppose that $G[S]$ is a complete graph for some set $S \subset V$. Then*

$$\sum_{v \in S} d_{\mathcal{H}}(v) \leq |\partial\mathcal{H}|.$$

Lemma 3.7.7 yields the following result.

Lemma 3.7.8. *Let $t \geq 1$ be a positive integer and $(G, \mathcal{H})$ be a cancellative pair on n vertices. Suppose that $(S_1, \ldots, S_t, R)$ is a clique expansion of G. Then*

$$|\mathcal{H}| \leq |\mathcal{H}[R]| + t|\partial\mathcal{H}|.$$

Proof. Notice that every edge in $\mathcal{H}$ either contains at least one vertex in $S_1 \cup \cdots \cup S_t$ or is completely contained in R. So,

$$|\mathcal{H}| \leq |\mathcal{H}[R]| + \sum_{i \in [t]} \sum_{v \in S_i} d_{\mathcal{H}}(v).$$

For every $i \in [t]$ since $G[S_i]$ is a complete graph, it follows from Lemma 3.7.7 that $\sum_{v \in S_i} d_{\mathcal{H}}(v) \leq |\partial\mathcal{H}|$. Therefore, $|\mathcal{H}| \leq |\mathcal{H}[R]| + t|\partial\mathcal{H}|$. ∎

3.7.2 Proof of Theorem 3.1.26

In this section we prove the following statement that implies Theorem 3.1.26.

Theorem 3.7.9. *Let $k \in 6\mathbb{N} + \{1, 3\}$ and $k \geq 3$. For every $\delta > 0$ there exists an $\epsilon > 0$ and n_0 such that the following holds for all $n \geq n_0$. Suppose that $\mathcal{H}$ is a cancellative 3-graph on n vertices with*

$$|\partial\mathcal{H}| \geq (1 - \epsilon)\frac{k-1}{2k}n^2 \quad \text{and} \quad |\mathcal{H}| \geq (1 - \epsilon)\frac{k-1}{6k^2}n^3. \tag{3.31}$$

Then, $\mathcal{H}$ is $\mathcal{S}$-colorable for some $\mathcal{S} \in \mathrm{STS}(k)$ after removing at most δn^3 edges.

Remarks.

(a) Our proof shows that $\delta = 20000k^6\epsilon^{1/2}$ is sufficient for Theorem 3.7.9.

(b) It is easy to see from Corollary 3.1.18 (also see Figure 6) that if $\mathcal{H}$ satisfies Equation 3.31, then the Euclidean distance between $(d(\partial\mathcal{H}), d(\mathcal{H}))$ and $\left(\frac{k-1}{k}, \frac{k-1}{k^2}\right)$ is bounded by some constant $\zeta = \zeta(\epsilon)$ that is linear in ϵ, and vice versa (we omit the detailed calculations here).

The technical parts of the proof of Theorem 3.1.26 are contained in proofs of Lemma 3.7.10 and Lemma 3.7.11. In Lemma 3.7.10 we will show that every 3-graph $\mathcal{H}$ that satisfies assumptions in Theorem 3.7.9 contains a small set U of vertices such that the induced subgraph of $\partial\mathcal{H}$ on $V(\mathcal{H}) \setminus U$ is k-partite. In Lemma 3.7.11 we will show that if a cancellative pair $(G, \mathcal{H})$ satisfies similar assumptions in Theorem 3.7.9 and G is k-partite, then $\mathcal{H}$ is $\mathcal{S}$-colorable for some $\mathcal{S} \in \mathrm{STS}(k)$ after removing very few number of edges.

Lemma 3.7.10. *Let $k \in 6\mathbb{N}+\{1,3\}$ and $k \geq 3$. There exists an absolute constant $c_1 = c_1(k) > 0$ such that for every constant ϵ satisfying $0 \leq \epsilon \leq c_1$ there exists $n_0 = n_1(k, \epsilon)$ such that the following holds for all $n \geq n_1$. Suppose that $\mathcal{H}$ is a cancellative 3-graph on n vertices that satisfies Equation 3.31. Then there exists a set $U \subset V(\mathcal{H})$ of size at most $130\epsilon k^4 n$ such that the induced subgraph of $\partial\mathcal{H}$ on $V(\mathcal{H}) \setminus U$ is k-partite.*

Lemma 3.7.11. *Let $k \in 6\mathbb{N}+\{1,3\}$ and $k \geq 3$. There exists an absolute constant $c_2 = c_2(k) > 0$ such that for every constant ϵ' satisfying $0 \leq \epsilon' \leq c_2$ there exists $n_2 = n_2(k, \epsilon')$ such that the following holds for all $n \geq n_2$. Suppose that $(G, \mathcal{H})$ is a cancellative pair on a set V of size n, G is k-partite,*

$$|G| \geq (1-\epsilon')\frac{k-1}{2k}n^2 \quad \text{and} \quad |\mathcal{H}| \geq (1-\epsilon')\frac{k-1}{6k^2}n^3. \tag{3.32}$$

Then, $\mathcal{H}$ is $\mathcal{S}$-colorable for some $\mathcal{S} \in \mathrm{STS}(k)$ after removing at most $600(\epsilon')^{1/2}k^3n^3$ edges.

First let us show that Lemma 3.7.10 and Lemma 3.7.11 imply Theorem 3.7.9.

Proof of Theorem 3.7.9 assuming Lemmas 3.7.10 and 3.7.11. Let $\epsilon > 0$ be a sufficiently small constant such that $\epsilon' = 800\epsilon k^6 n$ satisfies $\epsilon' \leq c_2$. Let $\delta = 20000\epsilon^{1/2}k^6$. Suppose that $\mathcal{H}$ is a cancellative 3-graph on a set V of n vertices and $\mathcal{H}$ satisfies assumptions in Theorem 3.7.9. By Lemma 3.7.10, there exists a set $U \subset V$ of size at most $130\epsilon k^4 n$ such that the induced subgraph of $\partial\mathcal{H}$ on $V \setminus U$ is k-partite. Let $G' = (\partial\mathcal{H})[V \setminus U]$ and $\mathcal{H}' = \mathcal{H}[V \setminus U]$. Then it is easy to see that $(G', \mathcal{H}')$ is a cancellative pair, and moreover,

$$|G'| \geq |\partial\mathcal{H}| - |U| \cdot n \geq (1-\epsilon)\frac{k-1}{2k}n^2 - 130\epsilon k^4 n^2 \geq (1-\epsilon')\frac{k-1}{2k}n^2,$$

and

$$|\mathcal{H}'| \geq |\mathcal{H}| - |U| \cdot n^2 \geq (1-\epsilon)\frac{k-1}{6k^2}n^3 - 130\epsilon k^4 n^3 \geq (1-\epsilon')\frac{k-1}{6k^2}n^3.$$

Therefore, by Lemma 3.7.11, $\mathcal{H}'$ contains subgraph $\mathcal{H}''$ of size at least $|\mathcal{H}'| - 600(\epsilon')^{1/2}k^3n^3$ such that $\mathcal{H}''$ is $\mathcal{S}$-colorable for some $\mathcal{S} \in \mathrm{STS}(k)$. Note that $|\mathcal{H}| - |\mathcal{H}''| \leq 600(\epsilon')^{1/2}k^3n^3 + 130\epsilon k^4 n^3 < 20000\epsilon^{1/2}k^6n^3$. This completes the proof of Theorem 3.7.9. ∎

3.7.3 Proof of Lemma 3.7.10

Proof of Lemma 3.7.10. Fix $k \in 6\mathbb{N} + \{1, 3\}$ and $k \geq 3$. Let $\epsilon > 0$ be a sufficiently small constant and n be a sufficiently large integer. Let $\mathcal{H}$ be a cancellative 3-graph on n vertices and assume that $\mathcal{H}$ satisfies assumptions in Lemma 3.7.10.

Claim 3.7.12. *We have* $|\partial\mathcal{H}| \le \left(\frac{k-1}{2k} + \epsilon\right) n^2$.

Proof. Let $x = 2|\partial\mathcal{H}|/n^2$ and suppose to the contrary that $x \ge ((k-1)/k + 2\epsilon)$. Then it follows from Theorem 3.7.6 that

$$
\begin{aligned}
|\mathcal{H}| \le \frac{(1-x)x}{6} n^3 + 3n^2 &\le \frac{1}{6}\left(\frac{k-1}{k} + 2\epsilon\right)\left(\frac{1}{k} - 2\epsilon\right) n^3 + 3n^2 \\
&\le \frac{k-1}{6k^2}\left(1 - \frac{2k(k-2)}{k-1}\epsilon\right) n^3 + 3n^2 < (1-\epsilon)\frac{k-1}{6k^2},
\end{aligned}
$$

which contradicts Equation 3.31. ∎

Claim 3.7.13. *The clique number* $\omega(\partial\mathcal{H})$ *of* $\partial\mathcal{H}$ *satisfies* $\omega(\partial\mathcal{H}) \le 10k\epsilon n$.

Proof. Suppose to the contrary that there exists a set $S \subset V(\mathcal{H})$ of size $\lceil 10k\epsilon n \rceil$ such the the induced subgraph of $\partial\mathcal{H}$ on S is complete. To keep the calculations simple, let us assume that $10k\epsilon n$ is an integer. Let $a = 10k\epsilon$, $R = V(\mathcal{H}) \setminus S$, and $e = |\partial\mathcal{H}|$. Let e_s be the number of edges in $\partial\mathcal{H}$ that have at least one vertex in S, and set $x' = (xn^2 - e_s)/(n-an)^2$. Notice that $e_s \le an^2$.

It follows from Lemma 3.7.7 and Theorem 3.7.6 that

$$
\begin{aligned}
|\mathcal{H}| &\le |\mathcal{H}[T]| + \sum_{v \in S} d(v) \\
&\le \frac{(1-x')x'}{6}(1-a)^3 n^3 + 3(1-a)^2 n^2 + e \\
&= \frac{((1-a)^2 n^2 - 2(e-e_s))(e-e_s)}{3(1-a)n} + 3(1-a)^2 n^2 + e \\
&= \frac{-2e_s^2 + \left(4e - (1-a)^2 n^2\right) e_s + (1-a)^2 n^2 e - 2e^2}{3(1-a)n} + 3(1-a)^2 n^2 + e. \qquad (3.33)
\end{aligned}
$$

Since $-2e_s^2 + \left(4e - (1-a)^2 n^2\right) e_s$ is increasing in e_s when $e_s \leq e - (1-a)^2 n^2/4$ and

$$\begin{aligned} e - \frac{(1-a)^2 n^2}{4} &> (1-\epsilon)\frac{(k-1)n^2}{2k} - \frac{(1-a)^2 n^2}{4} \\ &\geq (1-\epsilon)\frac{n^2}{3} - \frac{(1-a)^2 n^2}{4} > an^2, \end{aligned}$$

we may substitute $e_s = an^2$ into Equation 3.33 and obtain

$$|\mathcal{H}| \leq \frac{-2e^2 + \left((1+a)^2 n^2 + 3(1-a)n\right) e - (1+a^2)an^4}{3(1-a)n} + 3(1-a)^2 n^2. \tag{3.34}$$

Since $-2e^2 + \left((1+a)^2 n^2 + 3(1-a)n\right) e$ is decreasing in e when $e \geq (1+a)^2 n^2/4 + 3(1-a)n/4$ and

$$\frac{(1+a)^2 n^2}{4} + \frac{3(1-a)n}{4} < \frac{n^2}{4} + \frac{n^2}{100} < (1-\epsilon)\frac{n^2}{3} < (1-\epsilon)\frac{(k-1)n^2}{2k},$$

we may substitute $e = (1-\epsilon)(k-1)n^2/(2k)$ into Equation 3.34 and obtain

$$\begin{aligned} |\mathcal{H}| &< \frac{((1-\epsilon)(k-1) - 2ka)(1 + ka^2 + (k-1)\epsilon)}{6(1-a)k^2} n^3 + \left((1-\epsilon)\frac{(k-1)}{2k} + 3(1-a)^2\right) n^2 \\ &\leq \frac{((k-1)(1-a) - ka)\,(1+k\epsilon)}{6(1-a)k^2} n^3 + 4n^2 \\ &\leq \frac{k-1}{6k^2} n^3 + \frac{\epsilon}{6} n^3 - \frac{a}{6k} n^3 + 4n^2 < \frac{k-1}{6k^2} n^3 - \epsilon n^3, \end{aligned}$$

which contradicts Equation 3.31. ∎

Claim 3.7.14. *Suppose that $(S_1, \dots, S_t, R)$ is a maximal clique expansion of $\partial\mathcal{H}$ with $|S_t| \geq k+1$ for some positive integer t. Then $\sum_{i\in[t]} |S_i| < 30k^2\epsilon n$.*

Proof. Suppose to the contrary that there exist some positive integer t and an maximal clique expansion $(S_1, \dots, S_t, R)$ of $\partial\mathcal{H}$ with $|S_t| \geq k+1$ such that $\sum_{i\in[t]} |S_i| \geq 30k^2\epsilon n$. Let $\Sigma_t = \sum_{i\in[t]} |S_i|$.

Let $\beta = 20k^2\epsilon$. By Claim 3.7.13, $|S_i| < 10k\epsilon n$ for $i \in [t]$. So there exists $t' \leq t$ such that $\beta n - 10k\epsilon n \leq W_{t'} \leq \beta n + 10k\epsilon n < 30k^2\epsilon n$. Without loss of generality we may assume that $W_t = \lceil \beta n \rceil$ (since otherwise we may replace W_t by $W_{t'}$, and the exact value of β is not crucial in the proof as long as $20k^2 \leq \beta \leq 30k^2$). To keep the calculations simple, let us assume that βn is an integer.

Let E_t denote the number of edges in $\partial\mathcal{H}$ that have at least one vertex in $S_1 \cup \dots \cup S_t$ and let $x' = 2(e - E_t)/(n - \Sigma_t)^2$. Notice from Observation 3.7.2 (b) that

$$E_t \leq \sum_{i\in[t]} (|S_i| - 1)\, n = (\Sigma_t - t)\, n.$$

It follows from Theorem 3.7.6 and Lemma 3.7.8 that

$$\begin{aligned} |\mathcal{H}| &\leq \frac{x'(1-x')}{6}(n-\Sigma_t)^3 + 3(n-\Sigma_t)^2 + te \\ &= \frac{-2E_t^2 + \left(4e - (n-\Sigma_t)^2\right)E_t + (n-\Sigma_t)^2 e - 2e^2}{3(n-\Sigma_t)} + 3(n-\Sigma_t)^2 + te. \end{aligned} \tag{3.35}$$

Similar to the proof of Claim 3.7.13, we may substitute $E_t = (\Sigma_t - t)\, n$ into Equation 3.35 and obtain

$$|\mathcal{H}| \leq \frac{-2n^2t^2 + \big(n(n+\Sigma_t)^2 - (n+3\Sigma_t)e\big)\, t - (e-\Sigma_t n)(2e - n^2 - \Sigma_t^2)}{3(n-\Sigma_t)} + 3(n-\Sigma_t)^2. \quad (3.36)$$

Since $|S_i| \geq k+1$ for $i \in [t]$, we have $t \leq W_t/(k+1)$. Since $-2n^2t^2 + \big(n(n+\Sigma_t)^2 - (n+3\Sigma_t)e\big)\, t$ is increasing in t when

$$t \leq \big(n(n+\Sigma_t)^2 - (n+3\Sigma_t)e\big)\,/(4n^2)$$

and $\big(n(n+\Sigma_t)^2 - (n+3\Sigma_t)e\big)\,/(4n^2) \geq \Sigma_t/(k+1)$, we may substitute $t = \Sigma_t/(k+1)$ into Equation 3.36 and obtain

$$\begin{aligned} |\mathcal{H}| \leq{} & \frac{(k+1)\left(-2(k+1)e^2 + \big((k+1)n^2 + (2k+1)\Sigma_t n + (k-2)\Sigma_t^2\big)\, e\right)}{3(k+1)^2(n-\Sigma_t)} \\ & - \frac{\big((k+1)(n^2+\Sigma_t^2) - 2\Sigma_t n\big)\, k\Sigma_t n}{3(k+1)^2(n-\Sigma_t)} + 3(n-\Sigma_t)^2. \end{aligned} \quad (3.37)$$

Since $-2(k+1)^2e^2 + (k+1)\left((k+1)n^2 + (2k+1)\Sigma_t n + (k-2)\Sigma_t^2\right) e$ is decreasing in e when

$$e \geq \frac{(k+1)n^2 + (2k+1)\Sigma_t n + (k-2)\Sigma_t^2}{4(k+1)}$$

and

$$(1-\epsilon)\frac{k-1}{2k}n^2 \geq \frac{(k+1)n^2+(2k+1)\Sigma_t n+(k-2)\Sigma_t^2}{4(k+1)},$$

we may substitute $e = (1-\epsilon)(k-1)n^2/(2k)$ into Equation 3.37 and obtain

$$\begin{aligned}
|\mathcal{H}| &\leq (1-\epsilon)\frac{k-1}{6k^2}n^3 - \frac{(k+1)^2\Sigma_t n^3 - k(k^3+2k^2-k+2)\Sigma_t^2 n^2 + 2k^3(k+1)\Sigma_t^3 n}{6k^2(k+1)^2(n-\Sigma_t)} \\
&\quad + \epsilon n^3 + 3(n-\Sigma_t)^2 \\
&< (1-\epsilon)\frac{k-1}{6k^2}n^3 - \frac{(k+1)^2\Sigma_t n^3}{12k^2(k+1)^2 n} + +\epsilon n^3 + 3(n-\Sigma_t)^2 \\
&< (1-\epsilon)\frac{k-1}{6k^2}n^3 - \frac{\beta n^3}{12k^2} + +\epsilon n^3 + 3(n-\Sigma_t)^2 < (1-\epsilon)\frac{k-1}{6k^2}n^3
\end{aligned}$$

contradicting Equation 3.31. Here we used $\beta = 20k^2\epsilon$. ∎

Now let $(S_1,\ldots,S_t,R)$ be a maximal clique expansion of $\partial\mathcal{H}$ with threshold $k+1$ for some positive integer t. Let $\tilde{n} = |R|$ and $G = (\partial\mathcal{H})[R]$. Notice that by the definition of threshold, G is K_{k+1}-free. It follows from Claim 3.7.14 that $\tilde{n} = n - \sum_{i\in[t]}|S_i| \leq n - 30k^2\epsilon n$ and

$$|G| > |\partial\mathcal{H}| - 30k^2\epsilon n\cdot n \geq (1-\epsilon)\,\frac{k-1}{2k}n^2 - 30k^2\epsilon n^2 > \frac{k-1}{2k}n^2 - 31k^2\epsilon n^2. \tag{3.38}$$

Define

$$Z(G) = \left\{v\in R\colon d_G(v) \leq \frac{3k-4}{3k-1}\tilde{n} + 100k^4\epsilon\tilde{n}\right\}.$$

Claim 3.7.15. *We have $|Z(G)| < 100k^4\epsilon\tilde{n}$.*

Proof. Let $z = |Z(G)|$ and suppose to the contrary that $z \geq 100k^4\epsilon\tilde{n}$. To keep the calculations simple, let us assume that $100k^4\epsilon\tilde{n}$ is an integer. We may assume that $z = 100k^4\epsilon\tilde{n}$ since otherwise we replace $Z(G)$ by a subset of size $100k^4\epsilon\tilde{n}$. Let $R' = R \setminus Z(G)$. Since $G[R']$ is K_{k+1}-free, by Turán's theorem $|G[R']| \leq \frac{k-1}{2k}(\tilde{n}-z)^2$. Therefore,

$$\begin{aligned}
|G| &\leq \frac{k-1}{2k}(\tilde{n}-z)^2 + \left(\frac{3k-4}{3k-1}\tilde{n} + 100k^4\epsilon\tilde{n}\right) z \\
&\leq \frac{k-1}{2k}\tilde{n}^2 - \frac{1}{3k^2-k} z\tilde{n} + z^2 + 100k^4\epsilon\tilde{n}z \\
&< \frac{k-1}{2k}\tilde{n}^2 - \frac{100k^4\epsilon}{3k^2-k}\tilde{n}^2 + 2\left(100k^4\epsilon\tilde{n}\right)^2 < \frac{k-1}{2k}\tilde{n}^2 - 31k^2\epsilon\tilde{n}^2,
\end{aligned}$$

which contradicts Equation 3.38. ∎

Let $U = S_1 \cup \cdots \cup S_t \cup Z(G)$. Then by Claims 3.7.14 and 3.7.15, $|U| \leq 30k^2\epsilon n + 100k^4\epsilon\tilde{n} < 130k^4\epsilon n$. On the other hand, by Equation 3.38 and Claim 3.7.15 the induced subgraph of G on $V(\mathcal{H}) \setminus U$ has minimum degree at least

$$\frac{3k-4}{3k-1}\tilde{n} + 100k^4\epsilon\tilde{n} - |Z(G)|\tilde{n} > \frac{3k-4}{3k-1}\tilde{n}.$$

So by Theorem 3.7.3, the induced subgraph $G[R \setminus Z(G)]$ is k-partite, that is, the induced subgraph of $\partial\mathcal{H}$ on $V(\mathcal{H}) \setminus U$ is k-partite. This completes the proof of Lemma 3.7.10. ∎

3.7.4 Proof of Lemma 3.7.11

We prove Lemma 3.7.11 in this section. The following observation about the Steiner triple systems will be helpful to understand the proof (starting from Claim 3.7.23) of Lemma 3.7.11.

Observation 3.7.16. *Suppose that $\mathcal{H}$ is $\mathcal{S}$-colorable for some $\mathcal{S} \in \mathrm{STS}(k)$. Then for every $v \in V(\mathcal{H})$ the link $L_{\mathcal{H}}(v)$ consists of $(k-1)/2$ pairwise vertices disjoint complete graphs.*

Proof of Lemma 3.7.11. Fix $k \in 6\mathbb{N} + \{1,3\}$ and $k \geq 3$. Let $\epsilon' > 0$ be a sufficiently small constant and n be a sufficiently large integer. Let $(G, \mathcal{H})$ be a cancellative pair on n vertices that satisfies assumptions in Lemma 3.7.11. Let $V = V(G) = V(\mathcal{H})$ and suppose that $V = V_1 \cup \cdots \cup V_k$ is a partition such that every edge in G contains at most one vertex from each V_i. Let $\widehat{G}$ denote the complete k-partite graph with k-parts $V_1, \ldots, V_k$. Let $M_G = \widehat{G} \setminus G$ and call members in M_G missing edges of G. Notice that

$$|M_G| = \sum_{1 \leq i < j \leq k} |V_i||V_j| - |G| \leq \frac{k-1}{2k}n^2 - \left(1-\epsilon'\right)\frac{k-1}{2k}n^2 = \frac{k-1}{2k}\epsilon' n^2 < \frac{\epsilon' n^2}{2}. \tag{3.39}$$

The following claim can be proved easily using the following inequality (see [171])

$$\sum_{1 \leq i < j \leq k} x_i x_j + \frac{1}{2}\sum_{i \in [k]} \left(x_i - \frac{1}{k}\right)^2 \leq \frac{k-1}{2k},$$

where $x_1, \ldots, x_k \in [0,1]$ are real numbers satisfying $x_1 + \cdots + x_k = 1$.

Claim 3.7.17. *We have $||V_i| - n/k| \leq 2(\epsilon')^{1/2} n$ for every $i \in [k]$.*

The next claim gives an upper bound for the maximum codegree of $\mathcal{H}$.

Claim 3.7.18. *We have* $\Delta_2(\mathcal{H}) \leq n/k + 2(\epsilon')^{1/2}n + k\epsilon' n \leq n/k + 3(\epsilon')^{1/2}n$.

Proof. Suppose to the contrary that there exists $uv \in \partial\mathcal{H}$ with $N_{\mathcal{H}}(uv) > n/k + 2(\epsilon')^{1/2}n + k\epsilon n$. Let $V_i' = N_{\mathcal{H}}(uv) \cap V_i$ and $x_i = |V_i'|$ for $i \in [k]$. By Observation 3.7.5, $N_{\mathcal{H}}(uv)$ is independent in G. Therefore, every pair $\{u_i, u_j\}$ with $u_i \in V_i'$, $u_j \in V_j'$, and $i \neq j$, is a member in M_G. In particular, $\sum_{1\leq i<j\leq k} x_i x_j \leq |M_G|$. Claim 3.7.17 implies that $x_i \leq n/k + 2(\epsilon')^{1/2}n$ for $i \in [k]$, which combined with $\sum_{i\in[k]} x_i = |N_{\mathcal{H}}(uv)| \geq n/k + 2(\epsilon')^{1/2}n + k\epsilon' n$ imply that

$$\sum_{1\leq i<j\leq k} x_i x_j \geq \left(\frac{n}{k} + 2(\epsilon')^{1/2}n\right)\left(\frac{n}{k} + 2(\epsilon')^{1/2}n + k\epsilon' n - \left(\frac{n}{k} + 2(\epsilon')^{1/2}n\right)\right) \geq \epsilon' n^2 > |M_G|$$

contradicting Equation 3.39. ∎

Define the set of edges in G with small codegree in $\mathcal{H}$ as

$$G_s = \left\{uv \in G \colon |N_{\mathcal{H}}(uv)| \leq \frac{n}{2k}\right\}.$$

Claim 3.7.19. *We have* $|G_s| < 4k(\epsilon')^{1/2}n^2$.

Proof. Suppose to the contrary that $|G_s| \geq 4k(\epsilon')^{1/2}n^2$. Then it follows from $\sum_{uv\in G} |N_{\mathcal{H}}(uv)| \geq 3|\mathcal{H}|$, Claims 3.7.12 and 3.7.18 that

$$\begin{aligned}
3|\mathcal{H}| &\leq \frac{n}{2k}|G_s| + \left(\frac{n}{k} + 3(\epsilon')^{1/2}n\right)(|G| - |G_s|) \\
&\leq \left(\frac{n}{k} + 3(\epsilon')^{1/2}n\right)|G| - \frac{n}{2k}|G_s| \\
&\leq \left(\frac{n}{k} + 3(\epsilon')^{1/2}n\right)\left(\frac{k-1}{2k} + \epsilon'\right)n^2 - 2(\epsilon')^{1/2}n^3 \leq \frac{k-1}{2k^2} - \frac{(\epsilon')^{1/2}}{4}n^3
\end{aligned}$$

contradicting Equation 3.32. ∎

The following claim shows that for every $uv \in G$ most vertices in $N_{\mathcal{H}}(uv)$ will be contained in some V_i. It will be used intensively in the remaining part of the proof.

Claim 3.7.20. *For every $uv \in G$ and every $i \in [k]$ either*

$$|N_{\mathcal{H}}(uv) \cap V_i| < \frac{\epsilon' n^2}{|N_{\mathcal{H}}(uv)|} \quad \text{or} \quad |N_{\mathcal{H}}(uv) \cap V_i| > |N_{\mathcal{H}}(uv)| - \frac{\epsilon' n^2}{|N_{\mathcal{H}}(uv)|}.$$

In particular, if $|N_{\mathcal{H}}(uv)| > (\epsilon' k)^{1/2} n$, then there exists a unique $i \in [k]$ such that $|N_{\mathcal{H}}(uv) \cap V_i| > |N_{\mathcal{H}}(uv)| - \epsilon' n^2 / |N_{\mathcal{H}}(uv)|$.

Proof. Fix $uv \in G$ and $i \in [k]$. We may assume that $|N_{\mathcal{H}}(uv)| \geq (2\epsilon')^{1/2} n$ since otherwise we would have

$$\frac{\epsilon' n^2}{|N_{\mathcal{H}}(uv)|} \geq |N_{\mathcal{H}}(uv)| - \frac{\epsilon' n^2}{|N_{\mathcal{H}}(uv)|},$$

and there is nothing to prove.

Suppose that there exists a set V_i contradicting the assertion of Claim 3.7.20. Let $\alpha = |N_{\mathcal{H}}(uv)|$ and $\beta = |N_{\mathcal{H}}(uv) \cap V_i|$. Then similar to the proof of Claim 3.7.18 we obtain

$$|M_G| \geq \beta(\alpha - \beta) \geq \frac{\epsilon' n^2}{\alpha}\left(\alpha - \frac{\epsilon' n^2}{\alpha}\right) = \epsilon' n^2 - \left(\frac{\epsilon' n^2}{\alpha}\right)^2 \geq \frac{\epsilon'}{2} n^2,$$

which contradicts Equation 3.39.

Now, suppose that $|N_{\mathcal{H}}(uv)| > (\epsilon' k)^{1/2} n$. Since

$$k \times \frac{\epsilon' n^2}{|N_{\mathcal{H}}(uv)|} < \frac{\epsilon' k n^2}{(\epsilon' k)^{1/2} n} = (\epsilon' k)^{1/2} n < |N_{\mathcal{H}}(uv)|,$$

there exists $i \in [k]$ such that $|N_{\mathcal{H}}(uv) \cap V_i| > |N_{\mathcal{H}}(uv)| - \epsilon' n^2 / |N_{\mathcal{H}}(uv)|$. Since

$$|N_{\mathcal{H}}(uv)| - \frac{\epsilon' n^2}{|N_{\mathcal{H}}(uv)|} > |N_{\mathcal{H}}(uv)| - \frac{\epsilon' n^2}{(\epsilon' k)^{1/2} n} = |N_{\mathcal{H}}(uv)| - \frac{(\epsilon')^{1/2}}{k^{1/2}} n > \frac{|N_{\mathcal{H}}(uv)|}{2},$$

such i is unique. ∎

Claim 3.7.21. *We have* $\Delta(\mathcal{H}) < \frac{k-1}{2k^2} n^2 + 3(\epsilon')^{1/2} n^2$.

Proof. Fix $v \in V$ and it suffices to show that $d_{\mathcal{H}}(v) < \frac{k-1}{2k^2} n^2 + 3(\epsilon')^{1/2} n^2$. Without loss of generality, we may assume that $v \in V_1$. For every vertex $w \in N_{\mathcal{H}}(v) \subset \bigcup_{i=2}^{k} V_i$ let $d_v(w)$ denote the degree of w in the link graph $L_{\mathcal{H}}(v)$. It follows from Claim 3.7.18 that $d_v(w) = |N_{\mathcal{H}}(vw)| < n/k + 3(\epsilon')^{1/2} n$. Therefore, by Claim 3.7.17,

$$\begin{aligned} d_{\mathcal{H}}(v) = |L_{\mathcal{H}}(v)| = \frac{1}{2} \sum_{w \in \bigcup_{i=2}^{k} V_i} d_v(w) &< \frac{1}{2}(k-1) \left(\frac{n}{k} + 2(\epsilon')^{1/2} n\right) \left(\frac{n}{k} + 3(\epsilon')^{1/2} n\right) \\ &< \frac{k-1}{2k^2} n^2 + 3(\epsilon')^{1/2} n^2. \end{aligned}$$

∎

Define the set of vertices with small degree in $\mathcal{H}$ as

$$V_s = \left\{ v \in V : d_{\mathcal{H}}(v) < \frac{k-1}{2k^2}n^2 - 18(\epsilon')^{1/2}kn^2 \right\}.$$

Claim 3.7.22. *We have* $|V_s| < \frac{n}{6k}$.

Proof. Suppose to the contrary that $|V_s| \geq \frac{n}{6k}$. Then by Claim 3.7.21, we have

$$\begin{aligned} 3|\mathcal{H}| = \sum_{v \in V} d_{\mathcal{H}}(v) &\leq \left(\frac{k-1}{2k^2}n^2 + 3(\epsilon')^{1/2}n^2 \right) \left(n - \frac{n}{6k} \right) + \left(\frac{k-1}{2k^2}n^2 - 18(\epsilon')^{1/2}kn^2 \right) \frac{n}{6k} \\ &\leq \left(\frac{k-1}{2k^2} + 3(\epsilon')^{1/2} \right) n^3 - \left(18(\epsilon')^{1/2}k + 3(\epsilon')^{1/2} \right) \frac{n^3}{6k} \\ &\leq \frac{k-1}{2k^2}n^3 - \frac{(\epsilon')^{1/2}}{2k}n^3 \end{aligned}$$

contradicting Equation 3.32. ∎

First we show that if $v \in V \setminus V_s$, then the structure of $L_{\mathcal{H}}(v)$ is close to what we expect to see in a blowup of Steiner triple systems on k vertices.

Claim 3.7.23. *For every* $v \in V \setminus V_s$ *there exists a subgraph of* $L_{\mathcal{H}}(v)$ *of size at least* $d_{\mathcal{H}}(v) - 243(\epsilon')^{1/2}k^3n^2$ *that consists of* $(k-1)/2$ *pairwise vertex disjoipnt bipartite graphs.*

Proof. Fix $v \in V \setminus V_s$ and without loss of generality we may assume that $v \in V_1$. Note that $L_{\mathcal{H}}(v)$ is a graph on $N_{\mathcal{H}}(v) \subset \bigcup_{i=2}^{k} V_i$. For every $u \in N_{\mathcal{H}}(v)$ let $d_v(u)$ denote the degree of u in the link $L_{\mathcal{H}}(v)$. Define the set of vertices with small degree in $L_{\mathcal{H}}(v)$ as

$$N_s = \left\{ u \in \bigcup_{i=2}^{k} V_i \colon d_v(u) \leq \frac{n}{k} - 240(\epsilon')^{1/2}k^2 n \right\}.$$

We claim that $|N_s| \leq \frac{n}{6k}$ since otherwise we would have

$$\begin{aligned} 2|L_{\mathcal{H}}| &= \sum_{v \in V \setminus V_1} d_v(w) \\ &\leq \left(\frac{n}{k} + 3(\epsilon')^{1/2} n\right) (|V \setminus V_1| - |N_s|) + \left(\frac{n}{k} - 240(\epsilon')^{1/2}k^2 n\right) |N_s| \\ &\leq \left(\frac{n}{k} + 3(\epsilon')^{1/2} n\right) \left(\frac{k-1}{k} n + 2(\epsilon')^{1/2} n\right) - \left(3(\epsilon')^{1/2} n + 240(\epsilon')^{1/2} kn\right) \frac{n}{6k} \\ &< \frac{k-1}{k} n^2 - 37(\epsilon')^{1/2} kn^2, \end{aligned}$$

which contradicts the assumption that $v \in V \setminus V_s$. Therefore, by Claim 3.7.17, for every $i \in [2, k]$ we have

$$|V_i \setminus N_s| > \left(\frac{n}{k} - 2(\epsilon')^{1/2} n\right) - \frac{n}{6k} > \frac{2n}{3k}. \tag{3.40}$$

Fix $i \in [2,k]$ and a vertex $u \in V_i \setminus N_s$. By Claim 3.7.20, there exists a unique $\psi(u,i) \in [2,k]$, $\psi(u,i) \neq i$, such that

$$\begin{aligned}|N_{\mathcal{H}}(uv) \cap V_{\psi(u,i)}| > |N_{\mathcal{H}}(uv)| - \frac{\epsilon' n^2}{|N_{\mathcal{H}}(uv)|} &= d_v(u) - \frac{\epsilon' n^2}{d_v(u)} \\ &> \left(\frac{n}{k} - 240(\epsilon')^{1/2}k^2 n\right) - \frac{\epsilon' n^2}{n/(2k)} \\ &> \frac{n}{k} - 241(\epsilon')^{1/2}k^2 n.\end{aligned}$$

Since $\psi(u,i) \in [2,k]$ for every $u \in V_i \setminus N_s$, it follows from Equation 3.40 and the Pigeonhole principle that there exists a set $U_i \subset V_i \backslash B_v$ with $|U_i| > 2n/(3k(k-2))$ such that $\psi(u,i) = \psi(u',i)$ for every pair $u, u' \in U_i$. We abuse notation by letting $\psi(i) = \psi(u,i)$ for $u \in U_i$.

Define the bipartite graph $G_{i,\psi(i)}$ as

$$G_{i,\psi(i)} = \left\{uw \in L_{\mathcal{H}}(v) \colon u \in U_i, w \in V_{\psi(i)}\right\},$$

and notice that

$$|G_{i,\psi(i)}| > |U_i| \left(\frac{n}{k} - 241(\epsilon')^{1/2}k^2 n\right). \tag{3.41}$$

Let

$$V'_{\psi(i)} = \left\{w \in V_{i'} \colon d_{G_{i,\psi(i)}}(w) > |U_i|/2 > n/(3k(k-2))\right\}.$$

Then it follows from $\sum_{w\in U_i} d_{G_{i,\psi(i)}}(w) = |G_{i,\psi(i)}| = \sum_{w\in V_{\psi(i)}} d_{G_{i,\psi(i)}}(w)$ that

$$|U_i|\left(\frac{n}{k} - 241(\epsilon')^{1/2}k^2n\right) < |V'_{\psi(i)}||U_i| + \left(|V_{\psi(i)}| - |V'_{\psi(i)}|\right)\frac{|U_i|}{2} = \frac{|U_i|}{2}\left(V_{\psi(i)} + V'_{\psi(i)}\right),$$

which with Claim 3.7.17 imply that

$$\begin{aligned}|V'_{\psi(i)}| &> 2\left(\frac{n}{k} - 241(\epsilon')^{1/2}k^2n\right) - |V'_{\psi(i)}| \\ &> 2\left(\frac{n}{k} - 241(\epsilon')^{1/2}k^2n\right) - \left(\frac{n}{k} + 2(\epsilon')^{1/2}n\right) > \frac{n}{k} - 242(\epsilon')^{1/2}k^2n.\end{aligned}$$

For every $w \in V'_{\psi(i)}$ since

$$|N_{\mathcal{H}}(vw) \cap V_i| \geq d_{G_{i,\psi}}(w) > \frac{n}{3k(k-2)} > \frac{\epsilon' n^2}{|N_{\mathcal{H}}(vw)|},$$

it follows from Claim 3.7.20 that, in fact,

$$\begin{aligned}|N_{\mathcal{H}}(vw) \cap V_i| > |N_{\mathcal{H}}(vw)| - \frac{\epsilon' n^2}{|N_{\mathcal{H}}(vw)|} &> |N_{\mathcal{H}}(vw)| - \frac{\epsilon' n^2}{n/(3k(k-2))} \\ &> |N_{\mathcal{H}}(vw)| - 3\epsilon' k^2 n.\end{aligned}$$

Consequently, $\psi(w, \psi(i)) = i$ for every $w \in V'_{\psi(i)}$, and we abuse notation by writing it as $\psi^2(i) = i$.

Repeating the argument above we obtain a set $V'_i \subset V_i$ of size at least $n/k - 242(\epsilon')^{1/2}k^2n$ such that $|N_{\mathcal{H}}(vw) \cap V_i| \geq |N_{\mathcal{H}}(vw)| - 3\epsilon' k^2 n$ for every $w \in V'_i$.

View ψ as a map from $[2,k]$ to $[2,k]$. Then due to $\psi^2(i)=i$ for $i\in[2,k]$ the map ψ defines a perfect matching, namely $\{\{i,\psi(i)\}\colon i\in[2,k]\}$ (note that $\{i,\psi(i)\}$ and $\{\psi(i),\psi(\psi(i))\}$ are the same), on $[2,k]$.

To keep the notations simple, let us assume that $\psi(i)=(k-1)/2+i$ for $2\le i\le(k+1)/2$. Then the argument above implies that the number of edges in $L_{\mathcal{H}}(v)$ that are not contained in the union of the induced bipartite subgraphs $\bigcup_{i=2}^{(k+1)/2} L_{\mathcal{H}}(v)[V_i,V_{\psi(i)}]$ is at most

$$\sum_{i=2}^{(k+1)/2}\left(3\epsilon' k^2 n\left(|V_i'|+|V_{\psi(i)}'|\right)+n\left(|V_i\setminus V_i'|+|V_{\psi(i)}\setminus V_{\psi(i)}'|\right)\right)$$
$$\le\frac{k-1}{2}\left(3\epsilon' k^2n^2+2\left(242(\epsilon')^{1/2}k^2n+2(\epsilon')^{1/2}n\right)n\right)\le 243(\epsilon')^{1/2}k^3n^2.$$

∎

The proof of Claim 3.7.23 implies that for every $i\in[k]$ and every $v\in V_i\setminus V_s$ there is a bijection $\psi_v\colon[k]\setminus\{i\}\to[k]\setminus\{i\}$ with $\psi_v^2(j)=j$ for every $j\in[k]\setminus\{i\}$ such that all but at most $243(\epsilon')^{1/2}k^3n^2$ edges in $L_{\mathcal{H}}(v)$ is contained in the union of the induced bipartite subgraphs $\bigcup_{j\in[k]\setminus\{i\}} L_{\mathcal{H}}(v)[V_j,V_{\psi_v(j)}]$.

Now fix $i\in[k]$ and without loss of generality we may assume that $i=1$. By the Pigeonhole principle, there exists a set $W_1\subset V_1\setminus V_s$ of size at least $|V_1\setminus V_s|/(k-1)!>n/(2k!)$ (here $(k-1)!$ is an upper bound for the number of bijections between $[2,k]$ and $[2,k]$) such that $\psi_v\equiv\psi_{v'}$ for every pair $v,v'\in W_1$. To keep the notations simple, let ψ_1 be the bijection that satisfies $\psi_1\equiv\psi_v$ for $v\in W_1$, and further assume that $\psi_1(i)=(k-1)/2+i$ for $2\le i\le(k+1)/2$.

Define an auxiliary bipartite graph M with two parts $P_1 = W_1$ and $P_2 = \bigcup_{i=2}^{(k+1)/2} V_i \times V_{\psi_1(i)}$ such that a pair $\{v, (u, w)\}$ is an edge in M if and only if $\{u, w\} \in L_{\mathcal{H}}(v)$. It follows from Claim 3.7.23 that for every $v \in P_1$ we have

$$\begin{aligned} d_M(v) \geq d_{\mathcal{H}}(v) - 243(\epsilon')^{1/2}k^3n^2 &> \left(\frac{k-1}{2k^2}n^2 - 18(\epsilon')^{1/2}kn^2\right) - 243(\epsilon')^{1/2}k^3n^2 \\ &> \frac{k-1}{2k^2}n^2 - 270(\epsilon')^{1/2}k^3n^2. \end{aligned}$$

On the other hand, notice from Claim 3.7.17 that

$$|P_2| \leq \frac{k-1}{2}\left(\frac{n}{k} + 2(\epsilon')^{1/2}n\right)^2 < \frac{k-1}{2k^2}n^2 + 2(\epsilon')^{1/2}n^2.$$

Let P_2' denote the set of vertices in P_2 that have degree at least $|P_1|/2$ in M. Then it follows from $\sum_{v\in P_1} d_M(v) = |G| = \sum_{e\in P_2} d_M(e)$ that

$$|P_1|\left(\frac{k-1}{2k^2}n^2 - 270(\epsilon')^{1/2}k^3n^2\right) \leq |P_2'||P_1| + \left(|P_2| - |P_2'|\right)\frac{P_1}{2} = \frac{|P_1|}{2}\left(|P_2| + |P_2'|\right),$$

which implies that

$$\begin{aligned} |P_2'| &\geq 2\left(\frac{k-1}{2k^2}n^2 - 270(\epsilon')^{1/2}k^3n^2\right) - |P_2| \\ &> 2\left(\frac{k-1}{2k^2}n^2 - 270(\epsilon')^{1/2}k^3n^2\right) - \left(\frac{k-1}{2k^2}n^2 + 2(\epsilon')^{1/2}n^2\right) \\ &> \frac{k-1}{2k^2}n^2 - 541(\epsilon')^{1/2}k^3n^2. \end{aligned}$$

For a pair $(u, w) \in P_2'$ since

$$|N_{\mathcal{H}}(uw) \cap V_1| \geq |N_{\mathcal{H}}(uw) \cap W_1| \geq d_M((u, w)) \geq \frac{|W_1|}{2} \geq \frac{n}{4k!},$$

it follows from Claim 3.7.20 that, in fact,

$$|N_{\mathcal{H}}(uw) \cap V_1| \geq |N_{\mathcal{H}}(uw)| - \frac{\epsilon' n^2}{|N_{\mathcal{H}}(uw)|} \geq |N_{\mathcal{H}}(uw)| - \frac{\epsilon' n^2}{n/(4k!)} \geq |N_{\mathcal{H}}(uw)| - 4\epsilon' k! n.$$

Let

$$\mathcal{S} = \{\{i, j, \psi_i(j)\} \colon i \in [k] \text{ and } j \in [k] \setminus \{i\}\}$$

(every edge in $\mathcal{S}$ appeared six times in the definition above but we only keep one of them). Notice that $\mathcal{S}$ is a Steiner triple system on $[k]$. Let $\widehat{\mathcal{S}}$ be the blowup of $\mathcal{S}$ obtained by replacing each vertex i by the set V_i and replacing each edge by a corresponding complete 3-partite 3-graph. Then the argument above implies that one can delete at most

$$\begin{aligned}
&k \times \left(|P_2 \setminus P_2'| \times n + |P_2'| \times 4\epsilon' k! n\right) \\
&\leq k \times \left(\left(2\epsilon^{1/2} n + 541(\epsilon')^{1/2} k^3 n^2\right) \times n + \left(\frac{k-1}{2k^2} n^2 - 541(\epsilon')^{1/2} k^3 n^2\right) \times 4\epsilon' k! n\right) \\
&\leq 600(\epsilon')^{1/2} k^3 n^3,
\end{aligned}$$

to transform $\mathcal{H}$ into a subgraph of $\widehat{S}$. This completes the proof of Lemma 3.7.11. ∎

3.7.5 Proof of Theorem 3.1.28

In this section we prove the following statement which implies Theorem 3.1.28.

Theorem 3.7.24. *There exists an absolute constant $c_3 > 0$ such that for every constant ϵ satisfying $0 \le \epsilon \le c_3$ there exists n_0 such that the following holds for all $n \ge n_0$. Every cancellative 3-graph $\mathcal{H}$ on n vertices with $|\partial\mathcal{H}| = (1-\epsilon)\frac{k-1}{2k}n^2$ satisfies $|\mathcal{H}| \le \frac{k-1}{6k^2}n^3 - \frac{k-1}{4k^2}\epsilon n^3 + O(\epsilon^{3/2}n^3)$.*

Remarks.

(a) If $\epsilon > 0$ is sufficiently small, then $|\mathcal{H}| \le \frac{k-1}{6k^2}n^3 - \frac{k-1}{4k^2}\epsilon n^3 + O(\epsilon^{3/2}n^3) < \frac{k-1}{6k^2}n^3 - \frac{k-1}{6k^2}\epsilon n^3$. This shows that $g(\mathcal{T}_3, (k-1)/k - \epsilon) \le (k-1)/k^2 - \epsilon(k-1)/k^2$ for sufficiently small $\epsilon > 0$.

(b) The other part of Theorem 3.1.28, namely $g(\mathcal{T}_3, (k-1)/k + \epsilon) \le (k-1)/k^2 - \delta$, follows from Corollary 3.1.18 and the fact that $x(1-x)$ is decreasing when $x > 1/2$ (see Figure 6).

Before proving Theorem 3.1.28 let us present some useful lemmas.

Lemma 3.7.25. *Let $k \in 6\mathbb{N} + \{1,3\}$ and $k \ge 3$. Let $\epsilon > 0$ be a sufficiently small constant and n be a sufficiently large constant. Suppose that $\mathcal{H}$ is a 3-graph on n vertices with $|\partial\mathcal{H}| = (1-\epsilon)\frac{k-1}{2k}n^2$, and $\mathcal{H}$ is a blowup of $\mathcal{S}$ for some $\mathcal{S} \in \mathrm{STS}(k)$. Then $|\mathcal{H}| \le \frac{k-1}{6k^3}n^3 - \frac{k-1}{2k^2}\epsilon n^3 + 9\epsilon^{3/2}k^3n^3$.*

Proof. Let $V = V(\mathcal{H})$ and $V = V_1 \cup \dots \cup V_k$ be a partition such that $\mathcal{H}$ equals the blowup $\mathcal{S}[V_1, \dots, V_k]$ of $\mathcal{S}$. Without loss of generality we may assume that $|V_1| \ge \dots \ge |V_k|$. Let $\delta = |V_k|/n - 1/k \ge 0$ and notice from Claim 3.7.17 that $\delta < 2\epsilon^{1/2}$. For every $i \in [k]$ fix a subset $V_i' \subset V_i$ of size exactly $(1/k - \delta)n$ (note that $V_k' = V_k$). Let $V' = V_1' \cup \dots \cup V_k'$ and $R = V \setminus V'$.

Notice that $|V'| = k(1/k - \delta)n = n - k\delta n$ and $|R| = k\delta n$. Since the induced subgraph $\mathcal{H}[V']$ is a balanced blowup of $\mathcal{S}$, we have

$$|\mathcal{H}[V']| = \frac{1}{3}\binom{k}{2}\left(\frac{1}{k} - \delta\right)^3 n^3.$$

Since the induced subgraph of $\partial\mathcal{H}$ on V' has size $\binom{k}{2}(1/k - \delta)^2 n^2$ and every vertex in R has exactly $(k-1)(1/k - \delta)\, n$ neighbors in V', the size of the induced subgraph of $\partial\mathcal{H}$ on R satisfies

$$|(\partial\mathcal{H})[R]| = |\partial\mathcal{H}| - \frac{k-1}{2k}\left(\frac{1}{k} - \delta\right)^2 n^2 - k\delta n \cdot (k-1)\left(\frac{1}{k} - \delta\right) n = \binom{k}{2}\delta^2 n^2 - \frac{k-1}{2k}\epsilon n^2.$$

For $i \in \{1, 2\}$ let $\mathcal{E}_i$ denote the set of edges in $\mathcal{H}$ that have exactly i vertices in V'. Notice that for every vertex $v \in R$ the induced subgraph of $L_{\mathcal{H}}(v)$ on V' consists of $(k-1)/2$ balanced complete bipartite graphs and each of them has $2(1/k - \delta)n$ vertices. Therefore,

$$|\mathcal{E}_1| = |R| \cdot \frac{k-1}{2}\left(\frac{1}{k} - \delta\right)^2 n^2 = \binom{k}{2}\left(\frac{1}{k} - \delta\right)^2 \delta n^3.$$

On the other hand, since every pair $uv \in (\partial\mathcal{H})[R]$ satisfies $|N_{\mathcal{H}}(uv) \cap V'| = |V_i'| = (1/k - \delta)n$ for some unique $i \in [k]$, we obtain

$$|\mathcal{E}_2| = |(\partial\mathcal{H})[R]| \cdot \left(\frac{1}{k} - \delta\right) n = \left(\binom{k}{2}\delta^2 - \frac{k-1}{2k}\epsilon\right)\left(\frac{1}{k} - \delta\right) n^3.$$

Therefore,

$$
\begin{aligned}
|\mathcal{H}| &= |\mathcal{H}[V']| + |\mathcal{E}_1| + |\mathcal{E}_2| + |\mathcal{H}[R]| \\
&= \frac{1}{3}\binom{k}{2}\left(\frac{1}{k}-\delta\right)^3 n^3 + \binom{k}{2}\left(\frac{1}{k}-\delta\right)^2 \delta n^3 + \left(\binom{k}{2}\delta^2 - \frac{k-1}{2k}\epsilon\right)\left(\frac{1}{k}-\delta\right) n^3 + (\delta k n)^3 \\
&= \frac{k-1}{6k^3}n^3 - \frac{k}{6}\delta^3 n^3 - \frac{k-1}{2k^2}\epsilon n^3 + \frac{k-1}{2k}\epsilon\delta n^3 + \delta^3 k^3 n^3.
\end{aligned}
$$

Since $\delta < 2\epsilon^{1/2}$, we obtain $|\mathcal{H}| \le \frac{k-1}{6k^3}n^3 - \frac{k-1}{2k^2}\epsilon n^3 + 9\epsilon^{3/2}k^3 n^3$. ∎

The next lemma extends Lemma 3.7.25 from blowups of $\mathcal{S}$ to $\mathcal{S}$-colorable 3-graphs.

Lemma 3.7.26. *Let $k \in 6\mathbb{N} + \{1,3\}$ and $k \ge 3$. Let $\epsilon > 0$ be a sufficiently small constant and n be a sufficiently large constant. Suppose that $(G,\mathcal{H})$ is a cancellative pair on a set V of size n, $V = V_1 \cup \cdots \cup V_k$ is a partition such that G is a k-partite graph with parts $V_1,\dots,V_k$, and $\mathcal{H}$ is a subgraph of the blowup $\mathcal{S}[V_1,\dots,V_k]$ of some $\mathcal{S} \in \mathrm{STS}(k)$. If $|G| = (1-\epsilon)\frac{k-1}{2k}n^2$, then $|\mathcal{H}| \le \frac{k-1}{6k^2}n^3 - \frac{k-1}{4k^2}\epsilon n^3 + 2\epsilon^{3/2}n^3$.*

Proof. Let $\widehat{\mathcal{S}} = \mathcal{S}[V_1,\dots,V_k]$. Since $|G| = (1-\epsilon)\frac{k-1}{2k}n^2$, it follows from Claim 3.7.17 that $|V_i| \ge n/k - 2\epsilon^{1/2}n$ for every $i \in [k]$. So $N_{\widehat{\mathcal{S}}}(uv) \ge n/k - 2\epsilon^{1/2}n$ for every $uv \in \partial\widehat{\mathcal{S}}$. Let $\delta \ge 0$ be the real number that satisfies

$$|\partial\widehat{\mathcal{S}}| = (1-\delta)\frac{k-1}{2k}n^2.$$

Note that $\delta \leq \epsilon$ and $|\partial\widehat{\mathcal{S}} \setminus G| = (\epsilon - \delta)\frac{k-1}{2k}n^2$. Define

$$\mathcal{E} = \left\{E \in \widehat{\mathcal{S}} \colon \exists\{u, v\} \in \partial\widehat{\mathcal{S}} \setminus G \text{ such that } \{u, v\} \subset E\right\}.$$

Let us partition $\mathcal{E}$ into three subsets $\mathcal{E}_1$, $\mathcal{E}_2$, and $\mathcal{E}_3$, where every edge $E \in \mathcal{E}_i$ satisfies that $|\binom{E}{2} \setminus G| = i$ for $i \in [3]$. Since every pair $uv \in \partial\widehat{\mathcal{S}}$ satisfies $N_{\widehat{\mathcal{S}}}(uv) \geq n/k - 2\epsilon^{1/2}n$, we have

$$\begin{aligned}|\mathcal{E}_1| + 2|\mathcal{E}_2| + 3|\mathcal{E}_3| = \sum_{uv \in \partial\widehat{\mathcal{S}} \setminus G} d_{\widehat{\mathcal{S}}}(uv) &\geq |\partial\widehat{\mathcal{S}} \setminus G| \left(\frac{n}{k} - 2\epsilon^{1/2}n\right) \\ &= (\epsilon - \delta)\frac{k-1}{2k}\left(\frac{1}{k} - 2\epsilon^{1/2}\right)n^3.\end{aligned}$$

On the other hand, notice that

$$|\mathcal{E}| = |\mathcal{E}_1| + |\mathcal{E}_2| + |\mathcal{E}_3| \geq \frac{1}{2}\left(|\mathcal{E}_1| + 2|\mathcal{E}_2| + 3|\mathcal{E}_3|\right) - \mathcal{E}_3,$$

and by Theorem 3.1.16, we have $|\mathcal{E}_3| \leq \left(\partial\widehat{\mathcal{S}} \setminus G\right)^{3/2} \leq \epsilon^{3/2}n^3$. Therefore,

$$|\mathcal{E}| \geq \frac{1}{2}(\epsilon - \delta)\frac{k-1}{2k}\left(\frac{1}{k} - 2\epsilon^{1/2}\right)n^3 - \epsilon^{3/2}n^3 \geq (\epsilon - \delta)\frac{k-1}{4k^2}n^3 - 2\epsilon^{3/2}n^3.$$

We may view $\mathcal{H}$ as a 3-graph obtained from $\widehat{\mathcal{S}}$ by removing some edges. In particular, since $\mathcal{E} \subset \widehat{\mathcal{S}} \setminus \mathcal{H}$, we have $|\mathcal{H}| \leq |\widehat{\mathcal{S}}| - |\mathcal{E}|$. It follows from Lemma 3.7.25 that $|\widehat{\mathcal{S}}| \leq \frac{k-1}{6k^2}n^3 - \frac{k-1}{2k^2}\delta n^3 + 9\delta^{3/2}k^3n^3$. Therefore,

$$\begin{aligned}
|\mathcal{H}| \leq |\widehat{\mathcal{S}}| - |\mathcal{E}| &\leq \frac{k-1}{6k^2}n^3 - \frac{k-1}{2k^2}\delta n^3 + 9\delta^{3/2}k^3n^3 - \left((\epsilon-\delta)\frac{k-1}{4k^2}n^3 - 2\epsilon^{3/2}n^3\right) \\
&= \frac{k-1}{6k^2}n^3 - \frac{k-1}{4k^2}\epsilon n^3 + 2\epsilon^{3/2}n^3 - \left(\frac{k-1}{4k^2}\delta n^3 - 9\delta^{3/2}k^3n^3\right) \\
&\leq \frac{k-1}{6k^2}n^3 - \frac{k-1}{4k^2}\epsilon n^3 + 2\epsilon^{3/2}n^3.
\end{aligned}$$

∎

The next lemma is the key in the proof of Theorem 3.7.24.

Lemma 3.7.27. *Let $k \in 6\mathbb{N} + \{1,3\}$ and $k \geq 3$. Let $\epsilon > 0$ be a sufficiently small constant and n be a sufficiently large constant. Suppose that $(G, \mathcal{H})$ be a cancellative pair on n vertices, $\mathcal{G}$ is a k-partite graph, and $|\mathcal{G}| = (1-\epsilon)\frac{k-1}{2k}n^2$. Then $|\mathcal{H}| \leq \frac{k-1}{6k^2}n^3 - \frac{k-1}{4k^2}\epsilon n^3 + 10^9\epsilon^{3/2}k^9n^3$.*

Proof. Let $V = V(G) = V(\mathcal{H})$. Let $V = V_1 \cup \cdots \cup V_k$ be a partition such that G is a subgraph of the complete k-partite graph $\widehat{G}$ with parts $V_1, \ldots, V_k$. Suppose to the contrary that $|\mathcal{H}| \geq \frac{k-1}{6k^2}n^3 - \frac{k-1}{4k^2}\epsilon n^3 + 10^9\epsilon^{3/2}k^9n^3 > (1-2\epsilon)\frac{k-1}{6k^2}n^3$. Then by Lemma 3.7.11, $\mathcal{H}$ contains a subgraph $\mathcal{H}'$ of size at least $|\mathcal{H}| - \delta n^3$, where $\delta = 600(2\epsilon)^{1/2}k^3n^3$, such that $\mathcal{H}'$ is a subgraph of the blowup $\widehat{\mathcal{S}} = \mathcal{S}[V_1, \ldots, V_k]$ of some $\mathcal{S} \in \mathrm{STS}(k)$. We may assume that $\mathcal{H}' = \mathcal{H} \cap \widehat{\mathcal{S}}$ (otherwise we can replace $\mathcal{H}'$ by $\mathcal{H} \cap \widehat{\mathcal{S}}$). Let $M_G = \widehat{G} \setminus G$ and call members in M_G missing edges of G. Note that

$$|M_G| = |\widehat{G}| - |G| \leq \frac{k-1}{2k}n^2 - (1-\epsilon)\frac{k-1}{2k}n^2 < \frac{\epsilon}{2}n^2.$$

Define

$$G'_\ell = \left\{uv \in \partial\mathcal{H}' \colon d_{\mathcal{H}'}(uv) \geq \frac{n}{100k}\right\}, \quad \text{and} \quad \mathcal{H}'' = \left\{E \in \mathcal{H}' \colon \binom{E}{2} \subset G'_\ell\right\}.$$

Let $\epsilon'_1 \geq 0$ be the real number such that $|G'_\ell| = (1-\epsilon'_1)\frac{k-1}{2k}n^2$. Note that $G'_\ell \subset G \subset \widehat{G}$ and $\epsilon'_1 \geq \epsilon$.

Claim 3.7.28. *We have $|G'_\ell| \geq \frac{k-1}{2k}n^2 - 8\delta kn^2$. In other words, $\epsilon'_1 \leq 16\frac{k^2}{k-1}\delta$.*

Proof. Suppose to the contrary that $|G'_\ell| < \frac{k-1}{2k}n^2 - 8\delta kn^2$. Then

$$\begin{aligned}
3|\mathcal{H}'| = \sum_{uv \in \partial\mathcal{H}'} N_{\mathcal{H}'}(uv) &\leq |G'_\ell|\left(\frac{n}{k} + 3\epsilon^{1/2}n\right) + |\partial\mathcal{H}' \setminus G'_\ell|\frac{n}{100k} \\
&\leq \left(\frac{k-1}{2k}n^2 - 8\delta kn^2\right)\left(\frac{n}{k} + 3\epsilon^{1/2}n\right) + 8\delta kn^2 \cdot \frac{n}{100k} \\
&\leq \frac{k-1}{2k^2}n^3 + 3\epsilon^{1/2}n^3 - 8\delta kn^2\left(\frac{n}{k} - \frac{n}{100k}\right) \\
&\leq \frac{k-1}{2k^2}n^3 - 5\delta n^3 + 3\epsilon^{1/2}n^3 < \frac{k-1}{2k^2}n^3 - 4\delta n^3,
\end{aligned}$$

which contradicts $|\mathcal{H}'| \geq |\mathcal{H}| - \delta n^3$ and $|\mathcal{H}| \geq \frac{k-1}{6k^2}n^3 - \frac{k-1}{4k^2}\epsilon n^3 + 10^9\epsilon^{3/2}k^9n^3$. ∎

We will consider two cases depending on the value of ϵ'_1.

Case 1: $\epsilon'_1 > 400k^3\epsilon$.

Define

$$\mathcal{B}_1 = \left\{ E \in \mathcal{H} \setminus \mathcal{H}' \colon \binom{E}{2} \cap G'_\ell \neq \emptyset \right\} \quad \text{and} \quad \mathcal{B}_2 = \left\{ E \in \mathcal{H} \setminus \mathcal{H}' \colon \binom{E}{3} \subset G \setminus G'_\ell \right\}.$$

Claim 3.7.29. *We have $|\mathcal{B}_1| \leq 50\epsilon k n^3$.*

Proof. For every $uv \in G'_\ell$ let $\varphi(uv) \in [k]$ denote the index such that $N_{\mathcal{H}'}(uv) \subset V_{\varphi(uv)}$. Suppose that $E = \{u, v, w\}$ is contained in $\mathcal{B}_1$ and $uv \in G'_\ell$. Then $w \notin V_{\varphi(uv)}$, since otherwise E would be contained in $\mathcal{H}'$. Notice that every vertex $w \in N_{\mathcal{H}}(uv) \setminus V_{\varphi(uv)}$ cannot be adjacent to vertices in $V_{\varphi(uv)}$ (in G), since by Observation 3.7.5 the set $N_{\mathcal{H}}(uv)$ is independent in G. Therefore, $d_{M_G}(w) \geq V_{\varphi(uv)} \geq n/(100k)$ for every vertex $w \in N_{\mathcal{H}}(uv) \setminus V_{\varphi(uv)}$. Let

$$B_V = \bigcup_{uv \in G'} (N_{\mathcal{H}}(uv) \setminus V_{\varphi(uv)}).$$

Since $|M_G| \leq \frac{k-1}{2k}\epsilon n^2$, it follows from $2|M_G| \geq \sum_{w \in B_V} d_{M_G}(w)$ and $d_{M_G}(w) \geq n/(100k)$ for every $w \in B_V$ that $|B_V|\frac{n}{100k} \leq 2 \times \frac{k-1}{2k}\epsilon n^2 < \epsilon n^2$. Therefore, $|B_V| \leq 100\epsilon k n$.

Since every edge in $\mathcal{B}_1$ contains at least one vertex in B_V, it follows that $|\mathcal{B}_1| \leq |B_V|\binom{n}{2} \leq 50\epsilon k n^3$. ∎

Since $\partial\mathcal{B}_2 \subset G \setminus G'_\ell$, we have $|\partial\mathcal{B}_2| \le \frac{k-1}{2k}(\epsilon'_1 - \epsilon)n^2 < \epsilon'_1 n^2$. So, by inequality (**??**), $|\mathcal{B}_2| \le (\partial\mathcal{B}_2)^{3/2} < (\epsilon'_1)^{3/2}n^3$. Therefore, Lemma 3.7.26 applied to $(G'_\ell, \mathcal{H}'')$ yields

$$
\begin{aligned}
|\mathcal{H}| &= |\mathcal{H}''| + |\mathcal{H}' \setminus \mathcal{H}''| + |\mathcal{B}_1| + |\mathcal{B}_2| \\
&\le \frac{k-1}{6k^2}n^3 - \frac{k-1}{4k^2}\epsilon'_1 n^3 + 2(\epsilon'_1)^{3/2}n^3 + \frac{k-1}{2k}\epsilon'_1 n^2 \cdot \frac{n}{100k} + 50\epsilon k n^3 + (\epsilon'_1)^{3/2}n^3 \\
&\le \frac{k-1}{6k^2}n^3 - \left(\frac{k-1}{4k^2}\epsilon'_1 n^3 - \frac{k-1}{200k^2}\epsilon'_1 n^3 - (\epsilon'_1)^{3/2}n^3 + 50\epsilon k n^3\right) \\
&\le \frac{k-1}{6k^2}n^3 - \frac{k-1}{4k^2}\epsilon n^3.
\end{aligned}
$$

where the last inequality follows from $\epsilon'_1 > 400k^3\epsilon$ and $\epsilon'_1 \le 16\frac{k^2}{k-1}\delta \ll 1$.

Case 2: $\epsilon'_1 \le 400k^3\epsilon$.

Define

$$
G_\ell = \left\{uv \in \partial\mathcal{H} \colon d_{\mathcal{H}}(uv) \ge \frac{n}{50k}\right\}, \quad \text{and} \quad \mathcal{H}_\ell = \left\{E \in \mathcal{H} \colon \binom{E}{2} \subset G_\ell\right\}.
$$

Let $\epsilon_1 \ge 0$ be the real numbers such that $|G_\ell| = (1-\epsilon_1)\frac{k-1}{2k}n^2$. Note that $G_\ell \subset G \subset \widehat{G}$ and $\epsilon_1 \ge \epsilon$.

We claim that it suffices to show that

$$
|\mathcal{H}_\ell| \le \frac{k-1}{6k^2}n^3 - \frac{k-1}{4k^2}\epsilon_1 n^3 + 10^8\epsilon_1^{3/2}k^9 n^3. \tag{3.42}
$$

Indeed, suppose that Equation 3.42 holds. Then

$$\begin{aligned}|\mathcal{H}| &\le |\mathcal{H}_\ell| + |G \setminus G_\ell| \cdot \frac{n}{50k} \\ &\le \frac{k-1}{6k^2}n^3 - \frac{k-1}{4k^2}\epsilon_1 n^3 + 10^8 \epsilon_1^{3/2} k^9 n^3 + (\epsilon_1 - \epsilon)\frac{k-1}{2k}n^2 \cdot \frac{n}{50k} \\ &\le \frac{k-1}{6k^2}n^3 - \frac{k-1}{4k^2}\epsilon n^3 - \left((\epsilon_1 - \epsilon)\frac{k-1}{5k^2} - 10^8 \epsilon_1^{3/2} k^9\right) n^3.\end{aligned}$$

If $\epsilon_1 \ge 2\epsilon$, then $(\epsilon_1 - \epsilon)\frac{k-1}{5k^2} - 10^8\epsilon_1^{3/2}k^9 > 0$. If $\epsilon_1 \le 2\epsilon$, then $10^8\epsilon_1^{3/2}k^9n^3 \le 10^9\epsilon^{3/2}k^9n^3$. In either case we are done.

Let $\mathcal{H}'_\ell = \mathcal{H}_\ell \cap \mathcal{H}'$ and $\mathcal{H}''_\ell = \mathcal{H}_\ell \cap \mathcal{H}''$. Define

$$\begin{aligned}\mathcal{B}'_1(a) &= \left\{E \in \mathcal{H}_\ell \setminus \mathcal{H}'_\ell \colon \left|\binom{E}{2} \cap G'_\ell\right| \ge 2\right\}, \\ \mathcal{B}'_1(b) &= \left\{E \in \mathcal{H}'_\ell \setminus \mathcal{H}''_\ell \colon \left|\binom{E}{2} \cap G'_\ell\right| = 2\right\}, \\ \text{and} \quad \mathcal{B}'_2 &= \left\{E \in \mathcal{H}_\ell \setminus \mathcal{H}''_\ell \colon \left|\binom{E}{2} \setminus G'_\ell\right| \ge 2\right\}.\end{aligned}$$

A crucial observation is that if an edge $E \in \mathcal{H}_\ell \setminus \mathcal{H}''_\ell = \mathcal{H}_\ell \setminus \mathcal{H}''$ satisfies $\left|\binom{E}{2} \cap G'_\ell\right| = 3$, then $E \in \mathcal{H}_\ell \setminus \mathcal{H}'_\ell = \mathcal{H}_\ell \setminus \mathcal{H}'$. Indeed, if $E \in \mathcal{H}'_\ell \subset \mathcal{H}'$ and $\binom{E}{3} \subset G'_\ell$, then by the definition of $\mathcal{H}''$, we would have $E \in \mathcal{H}''$, which implies that $E \in \mathcal{H}_\ell \cap \mathcal{H}'' = \mathcal{H}''_\ell$.

Therefore, $\mathcal{B}'_1(a) \cup \mathcal{B}'_1(b) \cup \mathcal{B}'_2$ is a partition of $\mathcal{H}_\ell \setminus \mathcal{H}''_\ell$.

Claim 3.7.30. *We have* $|\mathcal{B}'_1(a)| \le 10^4\epsilon^2k^2n^2$.

Proof. Similar to the proof of Claim 3.7.29, for every $uv \in G'_\ell$ let $\varphi(uv) \in [k]$ denote the index such that $N_{\mathcal{H}'}(uv) \subset V_{\varphi(uv)}$. Let

$$B_V = \bigcup_{uv \in G'} (N_{\mathcal{H}}(uv) \setminus V_{\varphi(uv)}),$$

and recall from the proof of Claim 3.7.29 that $|B_V| \leq 100\epsilon k n$.

Suppose that $E = \{u, v, w\}$ is contained in $\mathcal{B}'_1(a)$ and assume that $uv, uw \in G'_\ell$. Then by the definition of B_V, we have $w \in B_V$ (because $uv \in G'_\ell$) and $v \in B_V$ (because $uw \in G'_\ell$). Therefore, E has at least two vertices in B_V. It follows that $|\mathcal{B}'_1(a)| \leq \binom{|B_V|}{2} \cdot n < 10^4 \epsilon^2 k^2 n^2$. ∎

Let $\epsilon_2 \geq 0$ be the real number such that $|G_\ell \cap G'_\ell| = (1 - \epsilon_2)\frac{k-1}{2k} n^2$. Notice that $|G_\ell \cap G'_\ell| \geq |G_\ell| - |G \setminus G'_\ell| \geq |G_\ell| - \epsilon'_1 \frac{k-1}{2k} n^2$, thus $\epsilon_2 \leq \epsilon_1 + \epsilon'_1 < \epsilon_1 + 400k^3\epsilon < 401k^3\epsilon_1$.

Now suppose that $E = \{u, v, w\}$ is contained in $\mathcal{B}'_1(b)$. By definition, $E \cap (G_\ell \setminus G'_\ell) \neq \emptyset$, and without loss of generality we may assume that $uv \in G_\ell \setminus G'_\ell$. Since $uv \in G_\ell \subset G'_\ell$, it follows from definitions of G_ℓ and G'_ℓ that $|N_{\mathcal{H}}(uv)| \geq \frac{n}{50k}$ and $|N_{\mathcal{H}'}(uv)| < \frac{n}{100k}$. This implies that there exists $i \in [k] \setminus \{\varphi(uv)\}$ such that $|N_{\mathcal{H}}(uv) \cap V_i| \geq \frac{n}{50k} - \frac{n}{100k} = \frac{n}{100k}$. By Claim 3.7.20, we actually have $|N_{\mathcal{H}}(uv) \cap V_i| \geq |N_{\mathcal{H}}(uv)| - \frac{\epsilon n^2}{|N_{\mathcal{H}}(uv)|} \geq |N_{\mathcal{H}}(uv)| - 100\epsilon k n^2$, which in turn implies that $|N_{\mathcal{H}'}(uv)| \leq |N_{\mathcal{H}}(uv) \setminus V_i| \leq 100\epsilon k n^2$. Therefore,

$$|\mathcal{B}'_1(b)| \leq \sum_{uv \in G_\ell \setminus G'_\ell} |N_{\mathcal{H}'}(uv)| \leq 100\epsilon k n^2 \cdot |G \setminus G'_\ell| \leq 100\epsilon\epsilon'_1 k n^3.$$

Now suppose that E is an edge in $\mathcal{B}_2'$. By definition $\left|\binom{E}{2} \cap (G_\ell \setminus G_\ell')\right| \geq 2$. So

$$
\begin{aligned}
|\mathcal{B}_2'| \leq \frac{1}{2} \sum_{uv \in G_\ell \setminus G_\ell'} N_{\mathcal{H}_\ell}(uv) &\leq \frac{1}{2}|G_\ell \setminus G_\ell'| \left(\frac{n}{k} + 3\epsilon^{1/2} n\right) \\
&= \frac{1}{2}|G_\ell \setminus (G_\ell \cap G_\ell')| \left(\frac{n}{k} + 3\epsilon^{1/2} n\right) \\
&= \frac{1}{2}(\epsilon_2 - \epsilon_1)\frac{k-1}{2k} n^2 \left(\frac{n}{k} + 3\epsilon^{1/2} n\right) \\
&\leq \frac{k-1}{4k^2}(\epsilon_2 - \epsilon_1) n^3 + \epsilon^{1/2}\epsilon_2 n^3
\end{aligned}
$$

(the factor $1/2$ is due to the fact that every edge in $\mathcal{B}_2'$ is counted at least twice in the summation $\sum_{uv \in G_\ell \setminus G_\ell'} N_{\mathcal{H}_\ell}(uv)$).

Now Lemma 3.7.26 applied to $(G_\ell \cap G_\ell', \mathcal{H}_\ell'')$ yields

$$
\begin{aligned}
|\mathcal{H}_\ell| &= |\mathcal{H}_\ell''| + |\mathcal{B}_1'(a)| + |\mathcal{B}_2'(b)| + |\mathcal{B}_2'| \\
&\leq \frac{k-1}{6k^2} n^3 - \frac{k-1}{4k^2}\epsilon_2 n^3 + 2\epsilon_2^{3/2} n^3 + 10^4 \epsilon^2 k^2 n^3 + \frac{k-1}{4k^2}(\epsilon_2 - \epsilon_1) n^3 + \epsilon^{1/2}\epsilon_2 n^3 \\
&\leq \frac{k-1}{6k^2} n^3 - \frac{k-1}{4k^2}\epsilon_1 n^3 + \left(2\epsilon_2^{3/2} n^3 + 10^4 \epsilon^2 k^2 n^3 + \epsilon^{1/2}\epsilon_2 n^3\right) \\
&\leq \frac{k-1}{6k^2} n^3 - \frac{k-1}{4k^2}\epsilon_1 n^3 + \left(2(401k^3\epsilon_1)^{3/2} + 10^4(401k^3\epsilon_1)^2 k^2 + \epsilon_1^{1/2}(401k^3\epsilon_1)\right) n^3 \\
&\leq \frac{k-1}{6k^2} n^3 - \frac{k-1}{4k^2}\epsilon_1 n^3 + 10^8 \epsilon_1^{3/2} k^9 n^3,
\end{aligned}
$$

this proves Equation 3.42. ∎

Now we are ready to prove Theorem 3.7.24.

Proof of Theorem 3.7.24. Fix C be a sufficiently large constant (the exact value of C can be obtained from the last inequality of the proof). Let $\epsilon > 0$ be a sufficiently small constant and let n be a sufficiently large integer. Let $\mathcal{H}$ be a cancellative 3-graph on n vertices with $|\partial\mathcal{H}| = (1-\epsilon)\frac{k-1}{2k}n^2$. Suppose to the contrary that

$$|\mathcal{H}| \geq \frac{k-1}{6k^2}n^3 - \frac{k-1}{4k^2}\epsilon n^3 + C\epsilon^{3/2}n^3 \geq (1-2\epsilon)\frac{k-1}{6k^2}n^3.$$

Let $V = V(\mathcal{H})$. By Lemma 3.7.10, there exists a set $U \setminus V$ of size at most $130(2\epsilon)k^4 n$ such that the induced subgraph of $\partial\mathcal{H}$ on $V \setminus U$ is k-partite. Viewing $\mathcal{H}[V \setminus U]$ as a subgraph on V (so U is a set of isolate vertices in $\mathcal{H}[V \setminus U]$). Let $\epsilon_1 \geq 0$ be the real number such that $|(\partial\mathcal{H})[V \setminus U]| = (1-\epsilon_1)\frac{k-1}{2k}n^2$. Notice that $|(\partial\mathcal{H})[V \setminus U]| \geq |\partial\mathcal{H}| - |U|n$, so $\epsilon_1 \leq 4 \times 130(2\epsilon)k^4 \leq 1100k^4\epsilon$. Since the pair $((\partial\mathcal{H})[V \setminus U], \mathcal{H}[V \setminus U])$ is cancellative, it follows from Lemma 3.7.27 that

$$|\mathcal{H}[V \setminus U]| \leq \frac{k-1}{6k^2}n^3 - \frac{k-1}{4k^2}\epsilon_1 n^3 + 10^9\epsilon_1^{3/2}k^9 n^3.$$

For $i \in [3]$ let $\mathcal{E}_i$ denote the set of edges in $\mathcal{H}$ that have exactly i vertices in U. For every $u \in U$ let $d_{V\setminus U}(u) = |N_{\mathcal{H}}(u) \setminus U|$. Similar to Claim 3.7.18, we have $|N_{\mathcal{H}}(uv) \setminus U| \leq \left(1/k + 3\epsilon_1^{1/2}\right) n$ for every $v \in N_{\mathcal{H}}(u)$. In other words, every vertex in $L_{\mathcal{H}}(v)[V \setminus U]$ has degree at most $\left(1/k + 3\epsilon_1^{1/2}\right) n$. Therefore,

$$\begin{aligned} |\mathcal{E}_1| \le \sum_{u \in U} |L_{\mathcal{H}}(v)[V \setminus U]| &\le \frac{1}{2} \sum_{u \in U} d_{V \setminus U}(u) \times \left(\frac{1}{k} + 3\epsilon_1^{1/2}\right) n \\ &= \frac{1}{2} |(\partial \mathcal{H})[V, U]| \left(\frac{1}{k} + 3\epsilon_1^{1/2}\right) n. \end{aligned}$$

Therefore,

$$\begin{aligned} |\mathcal{H}| &\le |\mathcal{H}[V \setminus U]| + |\mathcal{E}_1| + |\mathcal{E}_2| + |\mathcal{E}_3| \\ &\le \frac{k-1}{6k^2} n^3 - \frac{k-1}{4k^2} \epsilon_1 n^3 + 10^9 \epsilon_1^{3/2} k^9 n^3 + \frac{1}{2} |(\partial \mathcal{H})[V, U]| \left(\frac{1}{k} + 3\epsilon_1^{1/2}\right) n + \binom{|U|}{2} n \\ &\le \frac{k-1}{6k^2} n^3 - \frac{k-1}{4k^2} \epsilon_1 n^3 + \frac{1}{2} \frac{k-1}{2k} (\epsilon_1 - \epsilon) n^2 \left(\frac{1}{k} + 3\epsilon_1^{1/2}\right) n + \left(10^9 \epsilon_1^{3/2} k^9 + (260)^2 \epsilon^2 k^8\right) n^3 \\ &\le \frac{k-1}{6k^2} n^3 - \frac{k-1}{4k^2} \epsilon_1 n^3 + \frac{k-1}{4k^2} (\epsilon_1 - \epsilon) n^3 + \left(\epsilon_1^{3/2} + 10^9 \epsilon_1^{3/2} k^9 + (260)^2 \epsilon^2 k^8\right) n^3 \\ &< \frac{k-1}{6k^2} n^3 - \frac{k-1}{4k^2} \epsilon n^3 + C \epsilon^{3/2} n^3 \end{aligned}$$

contradicting our assumption (the last inequality used $\epsilon_1 \le 1100 k^4 \epsilon$). This completes the proof of Theorem 3.7.24. ∎

3.8 Concluding remarks

In this chapter we proved that for any $r \ge 3$ and any family $\mathcal{F}$ of r-graphs the function $g(\mathcal{F})$ has at most countably many discontinuities. We also constructed a family $\mathcal{D}$ of 3-graphs such that $g(\mathcal{D})$ is discontinuous at $x = 2/3$. It seems natural to ask the following question.

Problem 3.8.1. *Can $g(\mathcal{F})$ have infinitely many discontinuities?*

In Section 3.5 we proved several results about $g(\mathcal{T}_r)$ for $r \geq 3$. Even for $r = 3$ the function $g(\mathcal{T}_3)$ is already shown to have many intersecting properties, and is closely related to Steiner triple systems. The following question seems difficult for x not of the form $(k-1)/k$ with $k \equiv 1$ or 3 (mod 6).

Problem 3.8.2. *Determine $g(\mathcal{T}_3, x)$ for all $x \in (2/3, 1]$.*

Reiher observed that the function $x(1-x)$ in Theorem 3.5.7 can be replaced by the piecewise linear function $p(x) = \frac{k-1}{k+1} - \frac{k^2-k-1}{k(k+1)}x$ for all $k \in \mathbb{N}^+$ and $\frac{k-1}{k} \leq x \leq \frac{k}{k+1}$, which implies that $g(\mathcal{T}_3, x) \leq p(x)$ for all $\frac{2}{3} \leq x \leq 1$. This can be shown by redoing the proof of Theorem 3.5.7 and taking into account that instead of $\omega \geq \frac{1}{1-x}$ one can directly use $\omega \geq k+1$, unless $x = \frac{k-1}{k}$, but this case is already understood.

Now let us show a lower bound for $g(\mathcal{T}_3, x)$ for $x \in (2/3, 6/7]$. Let $\mathbb{F}$ denote the Fano Plane, i.e., $\mathbb{F}$ is a 3-graph on 7 vertices with edge set

$$\{123, 345, 561, 174, 275, 376, 246\}.$$

Let $\alpha \in [1/7, 1/3]$ and $\beta = (1-3\alpha)/4$. Let $\mathcal{H}_n(\alpha)$ be obtained from $\mathbb{F}$ by blowing up each vertex in $\{1, 2, 3\}$ into a set of size of αn and blowing up each vertex in $\{4, 5, 6, 7\}$ into a set of size of βn (note that these weights for blowing up the Fano plane are optimal). Let

$$x = \lim_{n\to\infty} \frac{|\partial \mathcal{H}_n(\alpha)|}{\binom{n}{2}} = 6\alpha^2 + 12\beta^2 + 24\alpha\beta = \frac{3}{4}(1 + 2\alpha - 7\alpha^2), \tag{3.43}$$

and

$$y = \lim_{n\to\infty} \frac{|\mathcal{H}_n(\alpha)|}{\binom{n}{3}} = 6\alpha^3 + 36\alpha\beta^2 = \frac{3}{4}\alpha(3 - 18\alpha + 35\alpha^2). \tag{3.44}$$

Then, Equation 3.43 and Equation 3.44 give

$$y = \frac{1}{147}\left(-70\sqrt{18x^2 - 21x^3} + 63x + 60\sqrt{18 - 21x} - 36\right), \tag{3.45}$$

which implies

$$g(\mathcal{T}_3, x) \geq \frac{1}{147}\left(-70\sqrt{18x^2 - 21x^3} + 63x + 60\sqrt{18 - 21x} - 36\right)$$

for all $x \in [2/3, 6/7]$ (see Figure 13).

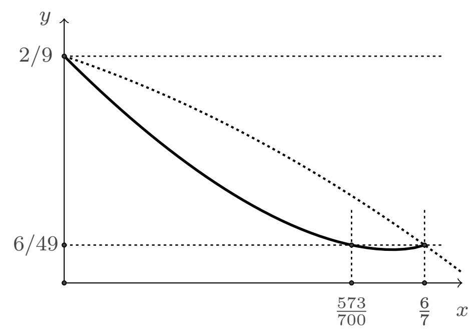


Figure 13. The lower bound for $g(\mathcal{T}_3, x)$ given by Equation 3.45.

The construction above gives an algebraic curve connecting $(2/3, 2/9)$ and $(6/7, 6/49)$. Using a similar method, one can construct an algebraic curve defined by

$$y = \frac{2\sqrt{3}(k+3)(k-1-kx)^{\frac{3}{2}}}{3k^2\sqrt{k-3}} + \frac{3kx-2k+2}{k^2} \tag{3.46}$$

to connect $(2/3, 2/9)$ and $((k-1)/k, (k-1)/k^2)$ for all $k \equiv 1$ or $3 \pmod 6$. However, we do not know how to construct curves to connect $((k-1)/k, (k-1)/k^2)$ and $((k'-1)/k', (k'-1)/k'^2)$ for consecutive $k, k' \geq 7$ and $k, k' \equiv 1$ or $3 \pmod 6$.

For $r \geq 4$ there is very little known about upper and lower bounds for $g(\mathcal{T}_r, x)$ for $x > (r-1)!/r^{r-2}$. We pose the following question.

Problem 3.8.3. *Let $r \geq 4$ and $x > (r-1)!/r^{r-2}$. Improve the upper bound for $g(\mathcal{T}_r, x)$, and construct cancellative r-graphs to give good lower bounds for $g(\mathcal{T}_r, x)$.*

Given our poor understanding of hypergraph Turán problems, determining the feasible region of other families of hypergraphs would also be of interest. In particular, we pose the following two questions.

Problem 3.8.4. *Determine the feasible region of $H^r_{\ell+1}$ for $r \geq 3$ and $\ell \geq r$.*

Problem 3.8.5. *Determine the feasible region of the Fano Plane.*

In Section 3.7, we showed that the cancellative triple systems and the Steiner triple systems are closely related to each other by proving that every cancellative triple system whose shadow density is close to $(k-1)/k$ and edge density is close to $(k-1)/k^2$ is structurally close to a balanced blowup of some Steiner triple system on k vertices for every $k \in 6\mathbb{N}+\{1,3\}$ and $k \geq 3$. Moreover, using this stability result we proved that the feasible region function $g(\mathcal{T}_3)$ of $\mathcal{T}_3$ has infinitely many local maxima.

• It would be interesting to explore whether there are similar relations between cancellative r-graphs and r-uniforms designs for $r \geq 4$. The case $r = 4$ is of particular interest since the stability property and the maximum size of an n-vertex cancellative 4-graph are already well studied [**?**; 208; 171].

• Theorem 3.1.28 shows that the point $\left(\frac{k-1}{k}, \frac{k-1}{k^2}\right)$ is a local maximum of $g(\mathcal{T}_3)$ for every $k \in 6\mathbb{N} + \{1,3\}$ and $k \geq 3$. In particular, $g(\mathcal{T}_3)$ of $\mathcal{T}_3$ has infinitely many local maxima. This indicates that the shape of the feasible region $\Omega(\mathcal{F})$ might be quite wild for general families $\mathcal{F}$. On the other hand, a complete determination of $g(\mathcal{T}_3)$ seems hard and it would be interesting to see whether methods that are used to solve the clique density problem [216; 201; 217] can be applied here.

• A nontrivial (i.e. the first coordinate is not on the boundary of the interval $\mathrm{proj}\Omega(\mathcal{F})$) local minimum (if it exists) of a feasible region function seems more mysterious. It would be interesting to find an explicit nontrivial local minimum of $g(\mathcal{F})$ for some family $\mathcal{F}$.

CHAPTER 4

A UNIFIED APPROACH TO HYPERGRAPH STABILITY

4.1 Introduction

We study the stability property of hypergraph Turán problems in this chapter. All families $\mathcal{F}$ of forbidden hypergraphs considered in this chapter will satisfy that $\pi(\mathcal{F}) > 0$, and such families are called nondegenerate.

Many families $\mathcal{F}$ have the property that there is a unique $\mathcal{F}$-free hypergraph $\mathcal{G}$ on n vertices achieving $\mathrm{ex}(n, \mathcal{F})$, and moreover, every $\mathcal{F}$-free hypergraph $\mathcal{H}$ of size close to $\mathrm{ex}(n, \mathcal{F})$ can be transformed to $\mathcal{G}$ by deleting and adding very few edges. Such a property is called stability of $\mathcal{F}$. The first stability theorem was proved independently by Erdős and Simonovits [233].

Theorem 4.1.1 (Erdős–Simonovits [233]). *Let $\ell \geq 2$ and let $\mathcal{F}$ be a family of graphs with $\chi(\mathcal{F}) = \ell + 1$. Then for every $\delta > 0$ there exist $\epsilon > 0$ and $N_0 \in \mathbb{N}$ such that every $\mathcal{F}$-free graph on $n \geq N_0$ vertices with at least $(1 - \epsilon)\mathrm{ex}(n, \mathcal{F})$ edges can be transformed to the Turán graph $T(n, \ell)$ by deleting and adding at most δn^2 edges.*

The stability phenomenon has been used to determine $\mathrm{ex}(n, \mathcal{F})$ exactly in many cases. It was first used by Simonovits in [233] to determine $\mathrm{ex}(n, F)$ exactly for all edge-critical graphs F and large n, and then by several authors (e.g. see [113; 141; 142; 195; 209; 30; 205]) to prove exact results for hypergraphs. In this chapter, stability will always mean stability relative to some intended class $\mathfrak{H}$ of 'almost extremal' $\mathcal{F}$-free graphs and we distinguish the following types of stability that have been studied in the literature.

Definition 4.1.2. *Let $\mathcal{F}$ be a nondegenerate family of r-graphs, where $r \geq 2$, and let $\mathfrak{H}$ be a class of $\mathcal{F}$-free r-graphs.*

(a) *If for every $\delta > 0$ there exist $\epsilon > 0$ and $N_0 \in \mathbb{N}$ such that every $\mathcal{F}$-free r-graph $\mathcal{H}$ on $n \geq N_0$ vertices with $|\mathcal{H}| \geq (\pi(\mathcal{F})/r! - \epsilon)n^r$ becomes a subgraph of some member of $\mathfrak{H}$ after removing at most $\delta|\mathcal{H}|$ edges, then $\mathcal{F}$ is said to be edge-stable with respect to $\mathfrak{H}$.*

(b) *If for every $\delta > 0$ there exist $\epsilon > 0$ and $N_0 \in \mathbb{N}$ such that every $\mathcal{F}$-free r-graph $\mathcal{H}$ on $n \geq N_0$ vertices with $|\mathcal{H}| \geq (\pi(\mathcal{F})/r! - \epsilon)n^r$ becomes a subgraph of some member of $\mathfrak{H}$ after removing at most $\delta|V(\mathcal{H})|$ vertices, then $\mathcal{F}$ is said to be vertex-stable with respect to $\mathfrak{H}$.*

(c) *If there exist $\epsilon > 0$ and N_0 such that every $\mathcal{F}$-free r-graph $\mathcal{H}$ on $n \geq N_0$ vertices with $\delta(\mathcal{H}) \geq \big(\pi(\mathcal{F})/(r-1)! - \epsilon\big)n^{r-1}$ is a subgraph of some member of $\mathfrak{H}$ we say that $\mathcal{F}$ is degree-stable with respect to $\mathfrak{H}$.*

As a trivial example, every nondegenerate family $\mathcal{F}$ is stable in all three senses with respect to the class $\mathfrak{Forb}(\mathcal{F})$ of all $\mathcal{F}$-free r-graphs. More interestingly, Theorem 4.1.1 tells us that every family $\mathcal{F}$ of graphs with $\chi(F) = \ell + 1 \geq 3$ is edge-stable with respect to the class $\mathfrak{T}_\ell = \{T(n, \ell) \colon n \in \mathbb{N}\}$ of ℓ-partite Turán graphs.

In general, if a family $\mathcal{F}$ of r-graphs is degree-stable with respect to some class $\mathfrak{H}$, then a standard vertex deletion argument (see e.g. Fact 4.2.1 (a)) shows that $\mathcal{F}$ is vertex-stable with respect to $\mathfrak{H}$ as well. Moreover, since any $\delta v(\mathcal{H})$ vertices of an r-graph $\mathcal{H}$ cover at most $\delta v(\mathcal{H})^r$ edges of $\mathcal{H}$, it is in all interesting examples the case that if $\mathcal{F}$ is vertex-stable with respect to $\mathfrak{H}$, then it is edge-stable with respect to $\mathfrak{H}$ as well.

The goal of this chpater is to provide a unified framework for the stability of certain classes of graph and hypergraph families. Our main result (Theorem 4.1.7) reduces the stability of many problems to the much simpler task of checking that $\mathcal{F}$-free graphs or hypergraphs with large minimum degree have a property we call vertex-extendability (see Definition 4.1.6). The approach is designed for degree-stability and thus it not only simplifies the proofs of many known stability theorems but also gives stronger forms of these stability theorems.

4.1.1 Main result

Our results can be regarded as adding a new ingredient to Zykov's symmetrization method [252] and we commence by describing an 'axiomatic' framework for the determination of extremal numbers by means of symmetrization.

Given two r-graphs F and $\mathcal{H}$ we say $\mathcal{H}$ is F-hom-free if there is no homomorphism from F to $\mathcal{H}$. For a family $\mathcal{F}$ of r-graphs, we say that $\mathcal{H}$ is $\mathcal{F}$-hom-free if it is F-hom-free for every $F \in \mathcal{F}$. The forbidden families $\mathcal{F}$ studied in this article have the following property.

Definition 4.1.3 (Blowup-invariance)**.** *A family $\mathcal{F}$ of r-graphs is blowup-invariant if every $\mathcal{F}$-free r-graph is $\mathcal{F}$-hom-free as well.*

For instance, for every $\ell \geq 2$ the families of graphs $\{K_\ell\}$ and $\{C_3, \dots, C_{2\ell-1}\}$ are blowup invariant, whilst $\{C_5\}$ is not blowup-invariant. In the graph case one can easily check that a one-element family $\{F\}$ is blowup-invariant if and only if F is a clique, but for hypergraphs blowup-invariant families consisting of a single hypergraph F are much more common. In fact, if every pair of vertices of F is covered by an edge of F, then $\{F\}$ is blowup-invariant. One confirms easily that every family $\mathcal{F}$ closed under taking homomorphic images is blowup-invariant.

We say two vertices $u, v \in V(\mathcal{H})$ are equivalent if $L_{\mathcal{H}}(u) = L_{\mathcal{H}}(v)$. Evidently, equivalence is an equivalence relation. We say that $\mathcal{H}$ is symmetrized if for any two non-equivalent vertices $u, v \in V(\mathcal{H})$ there is an edge $E \in \mathcal{H}$ containing both of them. For instance, a symmetrized graph is the same as a complete multipartite graph. We shall prove the following result by means of Zykov's symmetrization method.

Theorem 4.1.4. *Suppose that $\mathcal{F}$ is a blowup-invariant family of r-graphs. If $\mathfrak{H}$ denotes the class of all symmetrized $\mathcal{F}$-free r-graphs, then* $\mathrm{ex}(n, \mathcal{F}) = \mathfrak{h}(n)$ *holds for every $n \in \mathbb{N}$, where* $\mathfrak{h}(n) = \max\{|\mathcal{H}| \colon \mathcal{H} \in \mathfrak{H}$ *and* $v(\mathcal{H}) = n\}$.

Let us observe that this statement is very similar to the Lagrangian method developed and utilised by Motzkin–Straus [189], Sidorenko [232], Frankl–Füredi [102], and many others. Preparing the statement of our main result, we introduce some further notions. Recall that a class $\mathfrak{H}$ of r-graphs is called hereditary if it is closed under taking induced subgraphs.

Definition 4.1.5 (Symmetrized-stability)**.** *Let $\mathcal{F}$ be a family of r-graphs and let $\mathfrak{H}$ be a class of $\mathcal{F}$-free r-graphs. We say that $\mathcal{F}$ is symmetrized-stable with respect to $\mathfrak{H}$ if there exist* $\epsilon > 0$ *and* N_0 *such that every symmetrized $\mathcal{F}$-free r-graphs $\mathcal{H}$ on $n \geq N_0$ vertices with* $|\mathcal{H}| \geq (\pi(\mathcal{F})/r! - \epsilon)n^r$ *is a subgraph of a member of $\mathfrak{H}$.*

The next definition might be the most important one in this article.

Definition 4.1.6 (Vertex-extendibility)**.** *Let $\mathcal{F}$ be a family of r-graphs and let $\mathcal{H}$ be a class of $\mathcal{F}$-free r-graphs. We say that $\mathcal{F}$ is* vertex-extendable *with respect to $\mathfrak{H}$ if there exist* $\zeta > 0$ *and* $N_0 \in \mathbb{N}$ *such that for every $\mathcal{F}$-free r-graph $\mathcal{H}$ on $n \geq N_0$ vertices satisfying* $\delta(\mathcal{H}) \geq$

$\big(\pi(\mathcal{F})/(r-1)! - \zeta\big)n^{r-1}$ *the following holds: if* $\mathcal{H} - v$ *is a subgraph of a member of* $\mathfrak{H}$ *for some vertex* $v \in V(\mathcal{H})$*, then* $\mathcal{H}$ *is a subgraph of a member of* $\mathfrak{H}$ *as well.*

We can now state our sufficient conditions for degree-stability.

Theorem 4.1.7 (Main result)**.** *Suppose that* $\mathcal{F}$ *is a blowup-invariant nondegenerate family of* r*-graphs and that* $\mathfrak{H}$ *is a hereditary class of* $\mathcal{F}$*-free* r*-graphs. If* $\mathcal{F}$ *is symmetrized-stable and vertex-extendable with respect to* $\mathfrak{H}$*, then* $\mathcal{F}$ *is degree-stable with respect to* $\mathfrak{H}$ *as well.*

In practice the assumptions on $\mathfrak{H}$ are often easy to verify but it may happen that the family $\mathcal{F}$ we want to study fails to be blowup-invariant. If in such a situation we know for any reason that $\mathcal{F}$ is vertex-stable with respect to $\mathfrak{H}$, we can improve this information to degree-stability.

Theorem 4.1.8. *Suppose that* $\mathcal{F}$ *is a nondegenerate family of* r*-graphs and that* $\mathcal{H}$ *is a hereditary class of* $\mathcal{F}$*-free* r*-graphs. If* $\mathcal{F}$ *is vertex-stable and vertex-extendable with respect to* $\mathfrak{H}$*, then it is degree-stable with respect to* $\mathfrak{H}$ *as well.*

4.1.2 Further results and applications

For integers $\ell \geq r \geq 2$ let $\mathfrak{K}_\ell^r$ be the class of all blowups of K_ℓ^r. If $r = 2$ we omit the superscript and just write $\mathfrak{K}_\ell$ for the class of complete ℓ-partite graphs (whose vertex classes are allowed to be empty). Most but not all stability results described below are with respect to classes of the form $\mathfrak{K}_\ell^r$.

4.1.2.1 Graphs

The classical stability theorem of Erdős and Simonovits (Theorem 4.1.1) informs us that every family $\mathcal{F}$ of graphs with $\chi(\mathcal{F}) = \ell+1 \geq 3$ is edge-stable with respect to $\mathfrak{K}_\ell$. Complementing this result one can also characterise the families of graphs which are degree-stable and vertex-stable with respect to $\mathfrak{K}_\ell$. To this end we recall that a graph F is said to be edge-critical if it has an edge $e \in F$ such that $\chi(F-e) < \chi(F)$ and matching-critical if there exists a matching $M \subseteq F$ such that $\chi(F-M) < \chi(F)$. More generally, we call a family $\mathcal{F}$ of graphs edge-critical or matching-critical if there exists a graph $F \in \mathcal{F}$ with $\chi(F) = \chi(\mathcal{F})$ that is edge-critical or matching-critical. In the result that follows, part (b) is due to Erdős and Simonovits [72], while part (a) might very well be new.

Theorem 4.1.9. *A family $\mathcal{F}$ of graphs with $\chi(\mathcal{F}) = \ell + 1 \geq 3$ is*

(a) vertex-stable with respect to $\mathfrak{K}_\ell$ if and only if it is matching-critical

(b) and degree-stable with respect to $\mathfrak{K}_\ell$ if and only if it is edge-critical.

4.1.2.2 Cancellative hypergraphs and generalized triangles

Recall that an r-graph $\mathcal{H}$ is cancellative if $A \cup B = A \cup C$ implies that $B = C$ for all $A, B, C \in \mathcal{H}$. Since $A \cup B = A \cup C$ is equivalent to $B \triangle C \subseteq A$, an r-graph $\mathcal{H}$ is cancellative if and only if it is $\mathcal{T}_r$-free, where $\mathcal{T}_r$ denotes the family consisting of all r-graphs with three edges one of which contains the symmetric difference of the two other ones.

It was conjectured by Katona and proved by Bollobás [24] that the maximum number of edges in an n-vertex $\mathcal{T}_3$-free 3-graph is uniquely achieved by the balanced complete 3-partite 3-graph. Keevash and the second author [139] proved that $\mathcal{T}_3$ is edge-stable with respect to $\mathfrak{K}_3^3$, and the first author [160] discovered another short proof of the edge-stability of $\mathcal{T}_3$ giving a linear dependency between the error parameters. Sidorenko [232] proved that the maximum number of edges in an n-vertex $\mathcal{T}_4$-free 4-graph is uniquely achieved by the balanced complete 4-partite 4-graph. Later, Pikhurko [208] proved that $\mathcal{T}_4$ is vertex-stable with respect to $\mathfrak{K}_4^4$ using a sophisticated variation of Zykov symmetrization. For $r \geq 5$ the value of $\pi(\mathcal{T}_r)$ is unknown.

Cancellative hypergraphs are closely related to the Turán problem for generalized triangles. For $r \geq 2$ let Σ_r be the collection of all r-graphs with three edges A, B, C such that $|B \cap C| = r-1$ and $B \triangle C \subseteq A$. The unique r-graph $T_r \in \Sigma_r$ with $v(T_r) = 2r-1$ is called the *generalized triangle*. It is easy to see that $\Sigma_2 = \mathcal{T}_2 = \{K_3\}$, $\Sigma_3 = \mathcal{T}_3$, and $\Sigma_r \subset \mathcal{T}_r$ for $r \geq 4$.

The results on $\mathcal{T}_4$ due to Sidorenko [232] and Pikhurko [208] quoted earlier hold for Σ_4 instead of $\mathcal{T}_4$ as well. In particular, Σ_4 is known to be vertex-stable with respect to $\mathfrak{K}_4^4$. For $r = 5, 6$ Frankl and Füredi [102] proved that the extremal numbers $\mathrm{ex}(n, \Sigma_r)$ are only realized by balanced blowups of the famous Witt designs [248] with parameters $(11, 5, 4)$ and $(12, 6, 5)$, respectively. Norin and Yepremyan [204] proved that Σ_5 and Σ_6 are edge-stable with respect to blowups of these Witt-designs, but Pikhurko showed [208] that they fail to be vertex-stable. For $r \geq 7$ it is an open problem to determine $\pi(\Sigma_r)$.

Theorem 4.1.10. *For $r \in \{3, 4\}$ the family Σ_r is degree-stable with respect to $\mathfrak{K}_r^r$.*

4.1.2.3 Hypergraph expansions

A set $X \subset V(F)$ is called 2-covered in a hypergraph F if it induces a clique in $\partial_{r-2}F$. If $V(F)$ itself is 2-covered in F we simply say that F is 2-covered.

For an r-graph F with ℓ vertices we define $\mathcal{K}_\ell^F$ to be the set of all r-graphs of the form $F \cup \left\{S_{uv} \colon uv \in \binom{V(F)}{2} \setminus \partial_{r-2}F\right\}$, where for every pair of vertices $uv \in \binom{V(F)}{2} \setminus \partial_{r-2}F$ not covered by an edge of F the edge S_{uv} contains u and v. We write H_ℓ^F for the unique member of $\mathcal{K}_\ell^F$ having the largest number of vertices, namely

$$v(H_\ell^F) = \ell + (r-2)\left(\binom{\ell}{2} - |\partial_{r-2}F|\right).$$

The r-graphs in $\mathcal{K}_\ell^F$ are called weak expansions of F while H_ℓ^F is called the *expansion* of F. If F has no edges (and thus consists of ℓ isolated vertices) we write $\mathcal{K}_\ell^r$ and H_ℓ^r instead of $\mathcal{K}_\ell^F$ and H_ℓ^F.

The notion of hypergraph expansions was first introduced by the second author in [191] to extend Turán's theorem to hypergraphs. In [191] it was proved that for every $n \ge \ell \ge r \ge 2$ the maximum number of edges in an n-vertex $\mathcal{K}_{\ell+1}^r$-free r-graph is uniquely achieved by $T_r(n,\ell)$, the balanced complete ℓ-partite r-graph on n vertices. In addition, [191] proved that $\mathcal{K}_{\ell+1}^r$ is edge-stable with respect to $\mathcal{K}_{\ell+1}^r$. Later de Oliveira Contiero, Hoppen, Lefmann, and Odermann [51], and independently, the first author [160] improved the edge-stability result by showing that a linear dependence between δ and ϵ is sufficient. Pikhurko [209] refined [191] by showing that $T_r(n,\ell)$ is also the unique $H_{\ell+1}^r$-free r-graph on n vertices with the maximum number of edges for sufficiently large n.

Keevash [135] observed a generalization of these results to expansions of a large class of hypergraphs F. Let us write $\lambda(\mathcal{G})$ for the Lagrangian of a hypergraph $\mathcal{G}$ and set $\pi_\lambda(F) = \sup\{\lambda(\mathcal{G})\colon \mathcal{G} \text{ is } F\text{-free}\}$ for every r-graph F.

Theorem 4.1.11 (Keevash)**.** *Let F be an r-graph with $v(F) = \ell + 1$. If $\pi_\lambda(F) \le \binom{\ell}{r}/\ell^r$, then*

$$\mathrm{ex}(n, \mathcal{K}^F_{\ell+1}) \le \binom{\ell}{r} n^r/\ell^r$$

holds for all positive integers n, and equality occurs whenever n is divisible by r. In particular,

$$\pi(H^F_{\ell+1}) = \pi(\mathcal{K}^F_{\ell+1}) = (\ell)_r/\ell^r.$$

In the special case where F has an isolated vertex and $\pi_\lambda(F) < \binom{\ell}{r}/\ell^r$ Brandt, Irwin, and Jiang [30], and independently, Norin and Yepremyan [205] proved a stability theorem for the family $\mathcal{K}^F_{\ell+1}$ and used it to determine $\mathrm{ex}(n, H^r_\ell)$ exactly for all sufficiently large integers n. More specifically, [30] shows that $\mathcal{K}^F_{\ell+1}$ is vertex-stable, and [205] shows that $\mathcal{K}^F_{\ell+1}$ is edge-stable. Our result below shows the stronger fact that $\mathcal{K}^F_{\ell+1}$ is degree-stable.

Moreover, we prove degree-stability in many cases where F has no isolated vertices but is contained instead in the hypergraph $B(r, \ell+1)$ with vertex set $[\ell+1]$ and edge set

$$\{[r]\} \cup \{E \subseteq [2, \ell+1] \colon |E| = r \text{ and } |[2, r] \cap E| \le 1\}.$$

Theorem 4.1.12. *Let $\ell \ge r \ge 2$ and suppose that F is an r-graph satisfying $v(F) = \ell + 1$ and*

$$\sup\{\lambda(\mathcal{G}) \colon \mathcal{G} \textit{ is } F\textit{-free and not } K^r_\ell\textit{-colorable}\} < \binom{\ell}{r}/\ell^r. \tag{4.1}$$

If either F has an isolated vertex or $F \subseteq B(r, \ell+1)$, then the family $\mathcal{K}^F_{\ell+1}$ is degree-stable with respect to $\mathfrak{K}^r_\ell$.

There are several natural examples of hypergraphs F which have been proved to satisfy condition in the theorem above in the literature but whose families of weak expansions $\mathcal{K}^F_{\ell+1}$ were not known to be degree-stable before. For instance, Hefetz and Keevash [122] studied the case that $F = M^3_2$ is a 3-uniform matching with two edges and six vertices. More generally Jiang, Peng, and Wu [130] proved the assumption if $F = \{M^3_t, L^3_t, L^4_t\}$ holds for some $t \geq 2$; here M^3_t denotes the 3-uniform matching of size t and for $r \geq 2$ the r-graph L^r_t consists of t edges having one vertex v in common such that $E \cap E' = \{v\}$ holds for all distinct $E, E' \in L^r_t$. Brandt, Irwin, and Jiang [30] proved that in these cases the families $\mathcal{K}^F$ are vertex-stable. By combining the results in [130] on Lagrangians with Theorem 4.1.12 one immediately obtains the following strengthening of this fact.

Corollary 4.1.13. *For $t \geq 2$ the families $\mathcal{K}^{M^3_t}_{3t}$, $\mathcal{K}^{L^3_t}_{2t+1}$, $\mathcal{K}^{L^4_t}_{3t+1}$ are degree-stable with respect to $\mathfrak{K}^3_{3t-1}$, $\mathfrak{K}^3_{2t}$, $\mathfrak{K}^4_{3t}$, respectively.*

4.1.2.4 Expansions of matchings of size 2

For $r \geq 3$ let M^r_2 be the r-graph on $2r$ vertices consisting of two disjoint edges. The trivial observation that no r-graph in $\mathfrak{K}^r_{2r-1}$ contains a weak expansion of M^r_2 yields the lower bound $\pi(\mathcal{K}^{M^r_2}_{2r}) \geq (2r-1)_r/(2r-1)^r$. In their work [122] establishing equality for $r = 3$ Hefetz and Keevash also observed that for $r \geq 4$ there is a denser construction of $\mathcal{K}^{M^r_2}_{2r}$-free r-graphs.

Call an r-graph $\mathcal{H}$ semibipartite if there exists a partition $V(\mathcal{H}) = A \cup B$ such that $|A \cap E| = 1$ holds for every $E \in \mathcal{H}$. If $\mathcal{H}$ contains all $|A|\binom{|B|}{r-1}$ such edges we say that $\mathcal{H}$ is a complete semibipartite r-graph. It is easy to see that semibipartite r-graphs cannot contain weak expansions of M_2^r and that $(1-1/r)^{r-1}$ is the supremum of the edge densities of semibipartite r-graphs. A straightforward calculation shows that for $r \geq 4$ this number is indeed larger than the lower bound $(2r-1)_r/(2r-1)^r$ mentioned before. In fact Hefetz and Keevash [122] conjectured $\pi(\mathcal{K}_{2r}^{M_2^r}) = (1-1/r)^{r-1}$ for every $r \geq 4$. This was proved by Bene Watts, Norin, and Yepremyan [17] who also showed that $\mathcal{K}_{2r}^{M_2^r}$ is edge-stable with respect to the class $\mathcal{S}^r$ of all complete semibipartite r-graphs. Combining a substantial result on Lagrangians from their work with our Theorem 4.1.7 we strengthen this to degree-stability.

Theorem 4.1.14. *For every $r \geq 4$ the family of weak expansions of M_2^r is degree-stable with respect to $\mathcal{S}^r$.*

4.2 Proofs

4.2.1 Proof of the main result

Fact 4.2.1. *Let $\mathcal{F}$ be a family of r-graphs and let $\mathcal{H}$ be an $\mathcal{F}$-free r-graph on n vertices. If $\mathcal{H}$ has at least $(\pi(\mathcal{F})/r! - \epsilon)\, n^r$ edges, then*

(a) the set $Z_\epsilon(\mathcal{H}) = \{u \in V(\mathcal{H})\colon d_{\mathcal{H}}(u) \le (\pi(\mathcal{F})/(r-1)! - 2\epsilon^{1/2})\, n^{r-1}\}$ has size at most $\epsilon^{1/2} n$,

(b) and the r-graph $\mathcal{H}' = \mathcal{H} - Z_\epsilon(\mathcal{H})$ satisfies $\delta(\mathcal{H}') > (\pi(\mathcal{F})/(r-1)! - 3\epsilon^{1/2})\, n^{r-1}$.

We prove Theorems 4.1.4, 4.1.7, and 4.1.8 in this section. To this end the following pieces of notation will be convenient: If C denotes an equivalence class of some hypergraph $\mathcal{H}$, we write $d_{\mathcal{H}}(C)$ for the common degree of the vertices in C and $L_{\mathcal{H}}(C)$ for their common link. Given a class of hypergraphs $\mathfrak{H}$ we denote the class of spanning subgraphs of members of $\mathfrak{H}$ by $\mathfrak{H}^+$, i.e. we set

$$\mathfrak{H}^+ = \{\mathcal{H}\colon \text{there is } \mathcal{H}' \in \mathfrak{H} \text{ with } V(\mathcal{H}) = V(\mathcal{H}') \text{ and } \mathcal{H} \subseteq \mathcal{H}'\}.$$

If $\mathfrak{H}$ is hereditary, this is the same as the class of (not necessarily spanning) subgraphs of members of $\mathfrak{H}$.

Proof of Theorem 4.1.4. Fix $n \in \mathbb{N}$. The lower bound $\mathrm{ex}(n, \mathcal{F}) \ge \mathfrak{h}(n)$ is an immediate consequence of the fact that all members of $\mathfrak{H}$ are $\mathcal{F}$-free. So it remains to establish the upper bound $\mathrm{ex}(n, \mathcal{F}) \le \mathfrak{h}(n)$.

Suppose that it is not true and let $\mathcal{H}$ be an $\mathcal{F}$-free r-graph on n vertices with more than $\mathfrak{h}(n)$ edges chosen in such a way that the number m of its equivalence classes is minimum. Let $C_1, \dots, C_m$ denote the equivalence classes of $\mathcal{H}$.

Due to $|\mathcal{H}| > \mathfrak{h}(n)$ we know that $\mathcal{H}$ cannot be symmetrized. In other words, there exist $i, j \in [m]$ such that the graph $\partial_{r-2}\mathcal{H}$ is not complete between C_i and C_j. Without loss of generality we may assume that $\{i, j\} = \{1, 2\}$ and $d_{\mathcal{H}}(C_1) \leq d_{\mathcal{H}}(C_2)$. In view of the definition of equivalence there are actually no edges between C_1 and C_2 in $\partial_{r-2}\mathcal{H}$.

Now let $\mathcal{H}'$ be the unique r-graph satisfying $V(\mathcal{H}') = V(\mathcal{H})$, $\mathcal{H}' - C_1 = \mathcal{H} - C_1$, and $L_{\mathcal{H}'}(v) = L_{\mathcal{H}}(w)$ for all $v \in C_1$ and $w \in C_2$. Observe that $\{C_1 \cup C_2, C_3, \ldots, C_m\}$ is a refinement of the partition of $V(\mathcal{H}')$ into equivalence classes of $\mathcal{H}'$, for which reason $\mathcal{H}'$ has fewer than m equivalence classes. Together with

$$|\mathcal{H}'| = |\mathcal{H}| + |C_1| \left(d_{\mathcal{H}}(C_2) - d_{\mathcal{H}}(C_1)\right) \geq |\mathcal{H}| > \mathfrak{h}(n)$$

and our minimal choice of m this implies that $\mathcal{H}'$ cannot be $\mathcal{F}$-free. As there exists a homomorphism from $\mathcal{H}'$ to $\mathcal{H}$, it follows that $\mathcal{H}$ fails to be $\mathcal{F}$-hom-free. But, as $\mathcal{F}$ is blowup-invariant, this contradicts the assumption that $\mathcal{H}$ be $\mathcal{F}$-free. ∎

For the proof of Theorem 4.1.7 it will be convenient to say for $\zeta > 0$ and $N_0 \in \mathbb{N}$ that a family $\mathcal{F}$ of r-graphs is (ζ, N_0)-vertex-extendable with respect to a class of r-graphs $\mathfrak{H}$ if using the notation of Definiton 4.1.6 ζ and N_0 exemplify the vertex-extendibility of $\mathcal{F}$ with respect to $\mathfrak{H}$. The next lemma shows that vertex-extendibility can be used iteratively.

Lemma 4.2.2. *Suppose that the nondegenerate family $\mathcal{F}$ of r-graphs is $(2\epsilon, N_0)$-vertex-extendable with respect to a class $\mathfrak{H}$ of $\mathcal{F}$-free r-graphs, where $\epsilon \in (0, 1/2)$ and $N_0 \in \mathbb{N}$. Let $\mathcal{H}$ be an $\mathcal{F}$-free*

r-graph on $n \geq 2N_0$ vertices satisfying $\delta(\mathcal{H}) \geq \big(\pi(\mathcal{F})/(r-1)! - \epsilon\big)n^{r-1}$. If there exists a set $S \subseteq V(\mathcal{H})$ with $|S| \leq \epsilon n$ and $(\mathcal{H} - S) \in \mathfrak{H}^+$, then $\mathcal{H} \in \mathfrak{H}^+$.

Proof. Choose a minimal set $S' \subseteq S$ with $(\mathcal{H} - S') \in \mathfrak{H}^+$. If $S' = \emptyset$ we are done, so suppose for the sake of contradiction that there exists a vertex $v \in S'$. Setting $S'' = S' \setminus \{v\}$ and $\mathcal{H}'' = \mathcal{H} - S''$ we have $v(\mathcal{H}'') \geq n - |S| \geq (1-\epsilon)n \geq n/2 \geq N_0$ and

$$\begin{aligned}\delta(\mathcal{H}'') &\geq \delta(\mathcal{H}) - |S''|n^{r-2} > \big(\pi(\mathcal{F})/(r-1)! - \epsilon\big)n^{r-1} - \epsilon n^{r-1} \\ &\geq \big(\pi(\mathcal{F})/(r-1)! - 2\epsilon\big)v(\mathcal{H}'')^{r-1}.\end{aligned}$$

Moreover, we are assuming that $\mathcal{H}'' - v = \mathcal{H} - S'$ is in $\mathfrak{H}^+$. So by vertex-extendibility $\mathcal{H}''$ belongs to $\mathfrak{H}^+$ as well and S'' contradicts the minimality of S'. ∎

Next we shall show the following strengthening of Theorem 4.1.7 which also allows vertices of low degree in the almost extremal $\mathcal{F}$-free graphs. Recall that the sets $Z_\epsilon(\mathcal{H})$ appearing below were defined in Fact 4.2.1 (a).

Theorem 4.2.3. *Let $\mathcal{F}$ be a blowup-invariant nondegenerate family of r-graphs and let $\mathfrak{H}$ be a hereditary class of $\mathcal{F}$-free r-graphs. If $\mathcal{F}$ is symmetrized-stable and vertex-extendable with respect to $\mathfrak{H}$, then there are $\epsilon > 0$ and $N_0 \in \mathbb{N}$ such that every $\mathcal{F}$-free r-graph $\mathcal{H}$ on $n \geq N_0$ vertices with $|\mathcal{H}| > (\pi(\mathcal{F})/r! - \epsilon)\, n^r$ satisfies $\mathcal{H} - Z_\epsilon(\mathcal{H}) \in \mathfrak{H}^+$.*

The proof involves the following invariant of hypergraphs: If $C_1, \ldots, C_m$ are the equivalence classes of an r-graph $\mathcal{H}$, we set $\Psi(\mathcal{H}) = \sum_i |C_i|^2$.

Proof of Theorem 4.2.3. Choose $\epsilon \in (0, 1/36)$ so small and $N_0 \in \mathbb{N}$ so large that

(1) the symmetrized stability of $\mathcal{F}$ with respect to $\mathfrak{H}$ is exemplified by ϵ and N_0

(2) and $\mathcal{F}$ is $(6\epsilon^{1/2}, N_0/3)$-vertex-extendable with respect to $\mathfrak{H}$.

Now we fix $n \geq N_0$ and, assuming that the conclusion fails for n-vertex hypergraphs $\mathcal{H}$, we pick a counterexample $\mathcal{H}$ with $v(\mathcal{H}) = n$ such that the pair $(|\mathcal{H}|, \Psi(\mathcal{H}))$ is lexicographically maximal (which makes sense, as n is fixed). Let $C_1, \ldots, C_m$ be the equivalence classes of $\mathcal{H}$.

Setting $Z = Z_\epsilon(\mathcal{H})$ we have $(\mathcal{H} - Z) \notin \mathfrak{H}^+$ and, as $\mathfrak{H}$ is hereditary, $\mathcal{H} \notin \mathfrak{H}^+$ follows. Now (1) informs us that $\mathcal{H}$ is not symmetrized. So without loss of generality we may suppose that $\partial_{r-2}\mathcal{H}$ has no edges between C_1 and C_2 and that $(d_\mathcal{H}(C_1), |C_1|) \leq_{\text{lex}} (d_\mathcal{H}(C_2), |C_2|)$, where $\leq_{\text{lex}}$ means lexicographic ordering.

Now we pick two arbitrary vertices $v_1 \in C_1$ and $v_2 \in C_2$ and symmetrize only them, i.e., we let $\mathcal{H}'$ be the r-graph with $V(\mathcal{H}') = V(\mathcal{H})$, $\mathcal{H}' - v_1 = \mathcal{H} - v_1$ and $L_{\mathcal{H}'}(v_1) = L_\mathcal{H}(v_2)$. Clearly, if $d_\mathcal{H}(v_1) < d_\mathcal{H}(v_2)$, then $|\mathcal{H}'| > |\mathcal{H}|$. Moreover, if $d_\mathcal{H}(v_1) = d_\mathcal{H}(v_2)$, then $|\mathcal{H}'| = |\mathcal{H}|$, $|C_1| \leq |C_2|$, and

$$\Psi(\mathcal{H}') - \Psi(\mathcal{H}) \geq (|C_1| - 1)^2 + (|C_2| + 1)^2 - |C_1|^2 - |C_2|^2 = 2(|C_2| - |C_1| + 1) \geq 2.$$

In both cases $(|\mathcal{H}'|, \Psi(\mathcal{H}'))$ is lexicographically larger than $(|\mathcal{H}|, \Psi(\mathcal{H}))$ and our choice of $\mathcal{H}$ implies $\mathcal{H}' - Z_\epsilon(\mathcal{H}') \in \mathfrak{H}^+$. By Fact 4.2.1 (a) the set $Q = Z_\epsilon(\mathcal{H}') \cup \{v_1\}$ satisfies $|Q| \leq \epsilon^{1/2}n + 1 < 2\epsilon^{1/2}n$. Since $\mathfrak{H}$ is hereditary, the r-graph $\mathcal{H} - Q = \mathcal{H}' - Q$ belongs to $\mathfrak{H}^+$. Now Fact 4.2.1 (b) and (2) show that Lemma 4.2.2 applies to $3\epsilon^{1/2}$, $2N_0/3$, $\mathcal{H} - (Q \cap Z)$, and $Q \setminus Z$

here in place of ϵ, N_0, $\mathcal{H}$, and S there. Thus $\mathcal{H} - (Q \cap Z)$ belongs to $\mathfrak{H}^+$ and, since $\mathfrak{H}$ is hereditary, this yields the contradiction $(\mathcal{H} - Z) \in \mathfrak{H}^+$. ∎

4.2.2 Proof for graphs

In this subsection we prove Theorem 4.1.9. As we have already mentioned, its part (b) is due to Erdős and Simonovits [72], who proved that for every edge-critical graph F with $\chi(F) = \ell + 1 \geq 3$ there exists some $N_0 \in \mathbb{N}$ such that every graph G on $n \geq N_0$ vertices whose minimum degree is larger than $\frac{3\ell-4}{3\ell-1}n + O(1)$ either contains F or is ℓ-colorable. Consequently, all edge-critical families of graphs are degree-stable with respect to $\mathfrak{K}_\ell$.

Now suppose, conversely, that some graph family $\mathcal{F}$ is degree-stable with respect to the class $\mathfrak{K}_\ell$, where $\ell \geq 2$. This means, in particular, that for every $n \geq \ell + 1$ the graph $T^+(n, \ell)$ obtained from the n-vertex ℓ-partite Turán graph by inserting an additional edge into one of its vertex classes cannot be $\mathcal{F}$-free. Moreover, there cannot exist a graph $F' \in \mathcal{F}$ with $\chi(F') \leq \ell$, for then some member of $\mathfrak{K}_\ell$ would fail to be $\mathcal{F}$-free (as demanded by Definition 4.1.2). So altogether, $\mathcal{F}$ needs to contain an edge-critical graph F with $\chi(F) = \chi(\mathcal{F}) = \ell + 1$. In other words, $\mathcal{F}$ is indeed edge-critical.

We are left with proving part (a) of Theorem 4.1.9. The forward implication from vertex stability to matching-criticality is very similar to the argument in the previous paragraph, but instead of the graphs $T^+(n, \ell)$ one considers graphs obtained from Turán graphs by inserting (almost) perfect matchings into one of their partition classes. Omitting further details we proceed to the backwards implication. It clearly suffices to treat families consisting of a single graph.

Lemma 4.2.4. *Let F be a graph with $\chi(F) = \ell + 1 \geq 3$. If F is matching-critical, then F is vertex-stable with respect to $\mathfrak{K}_\ell$.*

Proof. Given $\delta > 0$ we choose $\epsilon, \eta > 0$ and $N_0 \in \mathbb{N}$ obeying the hierarchy $N_0^{-1} \ll \epsilon \ll \eta \ll \delta$. Suppose that G is an F-free graph on $n \geq N_0$ vertices with at least $\left(\frac{\ell-1}{2\ell} - \epsilon\right) n^2$ edges. We are to prove that G can be made ℓ-partite by deleting at most δn vertices. Theorem 4.1.1 yields a partition $V(G) = \bigcup_{i \in [\ell]} V_i$ such that $\sum_{i \in [\ell]} |G[V_i]| \leq \eta n^2$. Set

$$X_i = \left\{ x \in V_i \colon \text{ there is } j \in [\ell] \setminus \{i\} \text{ such that } |V_j \setminus N(x)| \geq \frac{n}{3\ell v(F)} \right\}$$

for every $i \in [\ell]$. Since $|G[V_i, V_j]| \geq |V_i||V_j| - 2\eta n^2$ holds for all distinct $i, j \in [\ell]$, we have $|X_i| \leq 6(\ell - 1)\ell v(F)\eta n \leq \delta n / 2\ell$ for every $i \in [\ell]$.

Recall that there is a matching M such that $\chi(F - M) \leq \ell$. If for some $i \in [\ell]$ there are $|M|$ independent edges $e_1, \ldots, e_{|M|}$ in $G[V_i \setminus X_i]$ we can find a copy of F in G where these edges $e_1, \ldots, e_{|M|}$ play the rôle of M. So by $F \nsubseteq G$ such matchings do not exist and it follows that for every $i \in [\ell]$ there is a set $Y_i \subseteq V_i \setminus X_i$ of size $|Y| \leq 2|M|$ covering all edges. Now the set $Q = \bigcup_{i \in [\ell]} (X_i \cup Y_i)$ has size at most $\delta n/2 + 2\ell|M| \leq \delta n$ and $G - Q$ is ℓ-partite. ∎

4.2.3 Proof for cancellative hypergraphs and generalized triangles

The goal of this subsection is to deduce Theorem 4.1.10 from Theorem 4.1.7. We commence by introducing a class $\mathfrak{T}_r$ of Σ_r-free r-graphs which is larger than $\mathfrak{K}_r^r$.

For integers $n \geq r \geq \ell \geq 1$ we call an r-graph $\mathcal{G}$ on n vertices an (n, r, ℓ)-system if every ℓ-subset of $V(\mathcal{G})$ is contained in at most one edge. As shown in [232; 102; 208; 167; **?**], the Turán problem for Σ_r is closely related to $(n, r, r-1)$-systems. Given any $r \geq 3$ we write $\mathfrak{T}_r$ for the class of all blowups of 2-covered $(n, r, r-1)$-systems. Since K_r^r is a 2-covered $(r, r, r-1)$-system, we have $\mathfrak{K}_r^r \subseteq \mathfrak{T}_r$. Perhaps at first sight surprisingly, we shall apply Theorem 4.1.7 to $\mathcal{F} = \Sigma_r$ and $\mathfrak{H} = \mathfrak{T}_r$. This choice of $\mathfrak{H}$ is forced upon us due to the symmetrized stability assumption and the following fact.

Lemma 4.2.5. *For $r \geq 3$ a Σ_r-free r-graph is symmetrised if and only if it is a proper blowup of some 2-covered $(n, r, r-1)$-system.*

Proof. Suppose first that $\mathcal{H}$ is a symmetrised Σ_r-free r-graph. Being a symmetrised hypergraph, $\mathcal{H}$ is a proper blowup of some 2-covered r-graph $\mathcal{T}$. If $\mathcal{T}$ fails to be a $(v(\mathcal{T}), r, r-1)$-system, then there are edges $B, C \in \mathcal{T}$ such that $|B \cap C| = r - 1$. Since $\mathcal{T}$ is 2-covered, some edge $A \in \mathcal{T}$ contains the two-element set $B \triangle C$. Now $\{A, B, C\}$ is a subgraph of $\mathcal{T}$ belonging to Σ_r, contrary to $\mathcal{H}$ being Σ_r-free. This proves that $\mathcal{T}$ is indeed a $(v(\mathcal{T}), r, r-1)$-system.

In the converse direction, proper blowups of 2-covered $(n, r, r-1)$-systems are clearly symmetrised and an argument similar to the previous paragraph shows that they are Σ_r-free as well. ∎

Proceeding with our intended application of Theorem 4.1.7 we observe that due to being closed under the formation of homomorphic images Σ_r is blow-up invariant. Moreover, $\mathfrak{T}_r$ is clearly hereditary and and the previous lemma shows that Σ_r is symmetrized-stable with respect to $\mathfrak{T}_r$. So it remains to verify vertex-extendibility for $r \in \{3, 4\}$. As the following lemma demonstrates, for this task we may restrict our attention to $\mathfrak{K}_r^r$ rather than $\mathfrak{T}_r$.

Lemma 4.2.6. *For $r \in \{3,4\}$ there exists $\epsilon_r > 0$ such that every $\mathcal{H} \in \mathfrak{T}_r$ with minimum degree $\delta(\mathcal{H}) > (r^{1-r} - \epsilon_r)n^{r-1}$ belongs to $\mathfrak{K}_r^r$.*

Proof. Choose $\epsilon_3, \epsilon_4 > 0$ sufficiently small and suppose that for some $r \in \{3,4\}$ an r-graph $\mathcal{H} \in \mathfrak{T}_r$ has n vertices and minimum degree at least $\left(r^{1-r} - \epsilon_r\right) n^{r-1}$. Without loss of generality we can suppose that $\mathcal{H}$ is a proper blowup of some (not necessarily 2-covered) $(m, r, r-1)$-system $\mathcal{T}$ with $V(\mathcal{T}) = [m]$. Write $\mathcal{H} = \mathcal{T}[V_1, \ldots, V_m]$ and set $x_i = |V_i|/n$ for every $i \in [m]$.

Since $d_{\mathcal{H}}(v) = L_{L_{\mathcal{T}}(i)}(x_1, \ldots, x_m)n^2$ holds for all $v \in V_i$ and $i \in [m]$, the minimum degree assumption yields $L_{L_{\mathcal{T}}(i)}(x_1, \ldots, x_m) \geq r^{1-r} - \epsilon_r$ for every $i \in [m]$. On the other hand, as every $(r-1)$-subset of $V(\mathcal{T})$ in contained in at most one edge of $\mathcal{T}$, we have $\sum_{i \in [m]} L_{L_{\mathcal{T}}(i)}(x_1, \ldots, x_m) \leq L_{K_m^{r-1}}(x_1, \ldots, x_m)$. It follows that

$$\left(r^{1-r} - \epsilon_r\right) m \leq \sum_{i \in [m]} L_{L_{\mathcal{T}}(i)}(x_1, \ldots, x_m) \leq \lambda(K_m^{r-1}) = \binom{m}{r-1} / m^{r-1}. \tag{4.2}$$

Now for $r = 3$ a sufficiently small choice of ϵ_3 guarantees $m \in \{2,3\}$; so $\mathcal{T}$ consists of a single edge and $\mathcal{H} \in \mathfrak{K}_3^3$. In the 4-uniform case Equation 4.2 leads to $m \in \{4,5\}$; since there exists no 2-covered $(5,4,3)$-system, the case $m = 5$ is impossible and thus we have indeed $\mathcal{H} \in \mathfrak{K}_4^4$. ∎

Due to the lower bound $\pi(\Sigma_r) \geq r!/r^r$, which follows from the fact that r-graphs in $\mathfrak{K}_r^r$ are Σ_r-free, the next lemma will imply that for $r \in \{3,4\}$ the family Σ_r is vertex-extendable with respect to $\mathfrak{T}_r$.

Lemma 4.2.7. *For every integer $r \geq 2$ there exist $\zeta > 0$ and $N_0 \in \mathbb{N}$ such that every Σ_r-free r-graph $\mathcal{H}$ on $n \geq N_0$ vertices which has minimum degree $\delta(\mathcal{H}) \geq (r^{1-r} - \zeta)n^{r-1}$ and possesses a vertex v such $\mathcal{H} - v$ is K_r^r-colorable is K_r^r-colorable itself.*

Proof. Given $r \geq 2$ we choose appropriate constants $\zeta > 0$ and $N_0 \in \mathbb{N}$ fitting into the hierarchy $N_0^{-1} \ll \zeta \ll r^{-1}$. Now let $\mathcal{H}$ be a Σ_r-free r-graph on $n \geq N_0$ vertices whose minimum degree is at least $(r^{1-r} - \zeta)n^{r-1}$. Set $V = V(\mathcal{H})$ and suppose that some vertex $v \in V$ has the property that $\mathcal{H}_v = \mathcal{H} - v$ is K_r^r-colorable. Fix a K_r^r-coloring $V(\mathcal{H}_v) = \bigcup_{i \in [r]} V_i$ of $\mathcal{H}_v$. Clearly

$$\delta(\mathcal{H}_v) \geq \left(r^{1-r} - \zeta\right) n^{r-1} - n^{r-2} \geq \left(r^{1-r} - 2\zeta\right) n^{r-1} \tag{4.3}$$

and some easy calculations yield

$$|V_i| = \left(1/r \pm \zeta^{1/3}\right) n \quad \text{for all } i \in [r]. \tag{4.4}$$

Claim 4.2.8. *Every edge of $\mathcal{H}$ intersects every vertex class V_i in at most one vertex.*

Proof. By symmetry it suffices to show $|E \cap V_1| \leq 1$ for every $E \in \mathcal{H}$. Assume for the sake of contradiction that there exist distinct vertices $w_1, w_1' \in E \cap V_1$. The $(r-1)$-graphs $G_1 = L_{\mathcal{H}_v}(w_1)$ and $G_1' = L_{\mathcal{H}_v}(w_1')$ are $(r-1)$-partite with vertex partition $V_2 \cup \cdots \cup V_r$ and by Equation 4.3 both of them have at least the size $(1/r^{r-1} - 2\zeta)n^{r-1}$. Due to Equation 4.4 this implies $|G_1 \cap G_1'| \geq n^{r-1}/2r^{r-1}$ and, in particular, there exists an edge $e \in G_1 \cap G_1'$. Now $\{E, e \cup \{w_1\}, e \cup \{w_1'\}\} \in \Sigma_r$ contradicts the assumption that $\mathcal{H}$ is Σ_r-free. ∎

Since no edge of $L_{\mathcal{H}}(v)$ can intersect all the partition classes $V_1, \dots, V_r$ we may assume without loss of generality that at least $d(v)/r$ edges of $L_{\mathcal{H}}(v)$ are contained in $V_2 \cup \dots \cup V_r$.

Claim 4.2.9. *We have $N_{\mathcal{H}}(v) \cap V_1 = \emptyset$.*

Proof. Suppose to the contrary that there exists a vertex $u \in N_{\mathcal{H}}(v) \cap V_1$ and consider an edge $E \in \mathcal{H}$ containing $\{u, v\}$. Let G_u and G_v be the subgraphs of $L_{\mathcal{H}}(u)$ and $L_{\mathcal{H}}(v)$ induced by $\bigcup_{j \in [2,r]} V_j$ respectively. Clearly, G_u is $(r-1)$-partite and by Claim 4.2.8 G_v is $(r-1)$-partite as well. Moreover, Equation 4.3 yields $|G_u| \geq \left(1/r^{r-1} - 2\zeta\right) n^{r-1}$. Together with $|G_v| \geq d(v)/r \geq \left(1/r^{r-1} - \zeta\right) n^{r-1}/r$ and Equation 4.4 this implies

$$|G_u \cap G_v| \geq \frac{1}{2r} \frac{n^{r-1}}{r^{r-1}}.$$

But if $e \in G_u \cap G_v$ is arbitrary, then the subgraph $\{E, e \cup \{v\}, e \cup \{u\}\}$ of $\mathcal{H}$ belongs to Σ_r, contrary to $\mathcal{H}$ being Σ_r-free. ∎

By Claim 4.2.8 and Claim 4.2.9 the partition $V(\mathcal{H}) = \bigcup_{i \in [r]} \widehat{V}_i$ defined by

$$\widehat{V}_i = \begin{cases} V_1 \cup \{v\} & \text{if } i = 1, \\ V_i & \text{if } i \in [2, r], \end{cases}$$

is a K_r^r-coloring of $\mathcal{H}$. This completes the proof of Lemma 4.2.7. ∎

We have thereby checked all assumptions of Theorem 4.1.7 and can conclude that for $r \in \{3,4\}$ the family Σ_r is degree-stable with respect to $\mathfrak{T}_r$. In view of Lemma 4.2.6 this implies that Σ_r is degree-stable with respect to $\mathfrak{K}_r^r$ as well.

4.2.4 Proof for hypergraph expansions

Throughout this subsection we fix two integers $\ell \geq r \geq 2$ and an r-graph F with $\ell+1$ vertices satisfying the assumptions of Theorem 4.1.12. Our goal is to conclude from Theorem 4.1.7 that the family $\mathcal{K}_{\ell+1}^F$ is indeed degree-stable with respect to $\mathfrak{K}_\ell^r$.

Since the family $\mathcal{K}_{\ell+1}^F$ is closed under taking homomorphic images, it is blowup-invariant and, clearly, $\mathfrak{K}_\ell^r$ is hereditary. So it remains to show that $\mathcal{K}_{\ell+1}^F$ is symmetrized-stable and vertex-extendable with respect to $\mathfrak{K}_\ell^r$. The fact that all members of $\mathfrak{K}_\ell^r$ are $\mathcal{K}_{\ell+1}^F$-free implies $\pi(\mathcal{K}_{\ell+1}^F) \geq (\ell)_r/\ell^r$ and thus our claim on symmetrized stability follows from the next statement.

Lemma 4.2.10. *There exists some $\epsilon > 0$ such that every symmetrized $\mathcal{K}_{\ell+1}^F$-free r-graph $\mathcal{H}$ with n vertices and $|\mathcal{H}| > \left(\binom{\ell}{r}/\ell^r - \epsilon\right) n^r$ is K_ℓ^r-colorable.*

Proof. We contend that every positive number ϵ satisfing

$$\sup\{\lambda(\mathcal{G})\colon \mathcal{G} \text{ is } F\text{-free but not } K_\ell^r\text{-colorable}\} + \epsilon \leq \binom{\ell}{r}/\ell^r$$

has the desired property. To see this we consider an arbitrary symmetrized $\mathcal{K}_{\ell+1}^F$-free r-graph $\mathcal{H}$ with n vertices and $|\mathcal{H}| > \left(\binom{\ell}{r}/\ell^r - \epsilon\right) n^r$. Since $\mathcal{H}$ is symmetrized, there exists a 2-covered hypergraph $\mathcal{G}$ such that $\mathcal{H}$ is a proper blow-up of $\mathcal{G}$. Now $|\mathcal{H}| \leq \lambda(\mathcal{G}) n^r$ yields $\binom{\ell}{r}/\ell^r - \epsilon < \lambda(\mathcal{G})$.

On the other hand, since $\mathcal{H}$ is $\mathcal{K}^F_{\ell+1}$-free and $\mathcal{G}$ is 2-covered, $\mathcal{G}$ must be F-free. So our choice of ϵ implies that $\mathcal{G}$ is K^r_ℓ-colorable and, hence, so is $\mathcal{H}$. ∎

The next lemma implies that $\mathcal{K}^F_{\ell+1}$ is vertex-extendable with respect to $\mathfrak{K}^r_\ell$ and thus concludes the proof of Theorem 4.1.12.

Lemma 4.2.11. *There exist $\zeta > 0$ and $N_0 \in \mathbb{N}$ such that every $\mathcal{K}^F_{\ell+1}$-free r-graph $\mathcal{H}$ on $n \geq N_0$ vertices satisfying the minimum degree condition $\delta(\mathcal{H}) > \left(\binom{\ell-1}{r-1}/\ell^{r-1} - \zeta\right) n^{r-1}$ and possessing a vertex v such that $\mathcal{H} - v$ is K^r_ℓ-colorable is K^r_ℓ-colorable itself.*

A slight modification of the proof below shows that this holds for $H^F_{\ell+1}$ instead of the family $\mathcal{K}^F_{\ell+1}$ as well.

Proof. Choose $N_0^{-1} \ll \zeta \ll \ell^{-1}$ appropriately and let $\mathcal{H}$ be a $\mathcal{K}^F_{\ell+1}$-free r-graph on $n \geq N_0$ vertices whose minimum degree is at least $\left(\binom{\ell-1}{r-1}/\ell^{r-1} - \zeta\right) n^{r-1}$. Write $V = V(\mathcal{H})$ and suppose that $\mathcal{H}_v = \mathcal{H} - v$ is K^r_ℓ-colorable for some vertex $v \in V$. Consider a K^r_ℓ-coloring $\bigcup_{i\in[\ell]} V_i = V \setminus \{v\}$ of $\mathcal{H}_v$ and the associated blowup $\widehat{K}^r_\ell = K^r_\ell[V_1, \dots, V_\ell]$ of K^r_ℓ. Sets in $\widehat{K}^r_\ell \setminus \mathcal{H}_v$ are called missing edges of $\mathcal{H}_v$; furthermore, for every $u \in V$ sets in $L_{\widehat{K}^r_\ell}(u) \setminus L_{\mathcal{H}_v}(u)$ are called missing edges of $L_{\mathcal{H}_v}(u)$.

Notice that

$$\delta(\mathcal{H}_v) \geq \delta(\mathcal{H}) - n^{r-2} \geq \left(\binom{\ell-1}{r-1}/\ell^{r-1} - 2\zeta\right) n^{r-1}. \tag{4.5}$$

Due to $|\mathcal{H}| \ge n\delta(\mathcal{H})/r > \left(\binom{\ell}{r}/\ell^r - \zeta\right) n^r$ we have, similarly,

$$|\mathcal{H}_v| \ge |\mathcal{H}| - n^{r-1} \ge \left(\binom{\ell}{r}/\ell^r - 2\zeta\right) n^r. \tag{4.6}$$

Consequently, the number of missing edges of $\mathcal{H}_v$ is at most $2\zeta n^r$. We proceed with the following claim.

Claim 4.2.12. *The following hold.*

(a) We have $|V_i| = \left(1/\ell \pm \zeta^{1/3}\right) n$ for every $i \in [\ell]$.

(b) If $i \in [\ell]$ and $u \in V(\mathcal{H}_v) \setminus V_i$, then $|V_i \setminus N_H(u)| \le \zeta^{1/3} n$.

(c) For every $u \in V(\mathcal{H}_v)$ the number of missing edges of $L_{\mathcal{H}_v}(u)$ is at most $\zeta^{1/3} n^{r-1}$. ∎

The proof of our next claim exploits the fact that F fails to be 2-covered, i.e., that there exist two distinct vertices $u, v \in V(F)$ such that $uv \notin \partial_{r-2} F$. Indeed, if F has an isolated vertex this is clear and if $F \subseteq B(r, \ell+1)$ we can take $u = 1$ as well as $v = r+1$.

Claim 4.2.13. *We have $|E \cap V_i| \le 1$ for all $E \in \mathcal{H}$ and $i \in [\ell]$.*

Proof. Otherwise we may assume, without loss of generality, that for some edge E there exist two distinct vertices $w_1, w_1' \in E \cap V_1$. By Claim 4.2.12 (a) and (b) for every $i \in [2, \ell]$ the set $V_i' = V_i \cap N_{\mathcal{H}}(w_1) \cap N_{\mathcal{H}}(w_1')$ satisfies $|V_i'| > n/2\ell$. Applying Lemma 5.3.23 with $S = \{w_1, w_1'\}$ and $T = [2, \ell]$ we obtain vertices $u_i \in V_i'$ for $i \in [2, \ell]$ such that the set $U = \{u_i \colon i \in [2, \ell]\}$ satisfies $\mathcal{H}[U \cup \{w_1\}] \cong \mathcal{H}[U \cup \{w_1'\}] \cong K_\ell^r$. As F is not 2-covered, it is a subgraph of $H = \mathcal{H}[U \cup \{w_1, w_1'\}]$. Thus $H \cup \{E\}$ is a weak expansion of F, contrary to $\mathcal{H}$ being $\mathcal{K}_{\ell+1}^F$-free. ∎

Essentially it remains to be shown that $N_{\mathcal{H}}(v) \cap V_i = \emptyset$ holds for some $i \in [\ell]$. Preparing ourselves we first show the following weaker result.

Claim 4.2.14. *There is no index $i \in [\ell]$ such that $N_{\mathcal{H}}(v) \cap V_i \neq \emptyset$ and $|N_{\mathcal{H}}(v) \cap V_j| \geq 2\zeta^{1/4r}n$ for all $j \in [\ell] \setminus \{i\}$.*

Proof. By symmetry it suffices to deal with the case $i = 1$. Assume for the sake of contradiction that there exists a vertex $u_1 \in N_{\mathcal{H}}(v) \cap V_1$ and moreover, that $|N_{\mathcal{H}}(v) \cap V_j| \geq 2\zeta^{1/4r}n$ for all $j \in [2, \ell]$. We shall show that, contrary to the hypothesis, $\mathcal{H}$ contains a weak expansion of F.

If F has an isolated vertex, we observe that due to Claim 4.2.12 (b) for every $j \in [2, \ell]$ the set $V_j' = V_j \cap N_{\mathcal{H}}(v) \cap N_{\mathcal{H}}(u_1)$ has at least the size $|V_j'| \geq 2\zeta^{1/4r}n - \zeta^{1/3}n > \zeta^{1/4r}n$. So we can apply Lemma 5.3.23 to $S = \{u_1\}$ and $T = [2, \ell]$, thus obtaining a set $U = \{u_j : j \in [2, \ell]\}$ with $u_j \in V_j'$ for $j \in [2, \ell]$ and $\mathcal{H}[U \cup \{u_1\}] \cong K_\ell^r$. For every $i \in [\ell]$ let $e_i \in \mathcal{H}$ be an edge containing both u_i and v. Since at least one vertex of F is isolated, $H = \mathcal{H}[U \cup \{u_1\}] \cup \{e_j : j \in [\ell]\}$ is the desired weak $(\ell + 1)$-expansion of F.

So it remains to consider the case $F \subseteq B(r, \ell + 1)$. Pick an edge $E \in \mathcal{H}$ containing $\{v, u_1\}$. By Claim 4.2.13 we may assume that E is of the form $\{v, u_1, \dots, u_{r-1}\}$, where $u_j \in V_j$ holds for all $j \in [2, r - 1]$. Claim 4.2.12 (b) tells us that for every $k \in [r, \ell]$ the set

$$V_k' = V_k \cap N_{\mathcal{H}}(v) \cap \left(\bigcap_{j \in [1, r-1]} N_{\mathcal{H}}(u_j) \right)$$

has at least the size $|V_k'| \geq 2\zeta^{1/4r}n - (r - 1)\zeta^{1/3}n > \zeta^{1/4r}n$. For this reason Lemma 5.3.23 applied to $S = \{u_1, \dots, u_{r-1}\}$ and $T = [r, \ell]$ leads to a set $U = \{u_k : k \in [r, \ell]\}$ such that

- $u_k \in V'_k$ for every $k \in [r, \ell]$,
- $\mathcal{H}[U] \cong K^r_{\ell-r+1}$,
- and $L_{\mathcal{H}}(u_j)[U] \cong K^{r-1}_{\ell-r+1}$ for every $j \in [r-1]$.

Next we select for every $k \in [r, \ell]$ an edge $E_i \in \mathcal{H}$ containing both u_k and v. Now

$$H = \mathcal{H}[U \cup \{u_1, \dots, u_{r-1}\}] \cup \{E\} \cup \{E_k \colon k \in [r, \ell]\}$$

is a weak expansion of $B(r, \ell+1)$ and, a fortiori, a weak expansion of F. ∎

Let us now consider the set

$$S = \left\{ i \in [\ell] \colon |N_{\mathcal{H}}(v) \cap V_i| \geq 2\zeta^{1/4r} n \right\}.$$

By Claim 4.2.14 we know, in particular, that $S \neq [\ell]$. Pick an arbitrary $i_\star \in [\ell] \setminus S$. Now Claim 4.2.12 (a) and $|d_{\mathcal{H}}(v)| \geq \left(\binom{\ell-1}{r-1} / \ell^{r-1} - \zeta \right) n^{r-1}$ imply $S = [\ell] \setminus \{i_\star\}$ and a further application of Claim 4.2.14 discloses $N_{\mathcal{H}}(v) \cap V_{i_\star} = \emptyset$. Together with Claim 4.2.13 this shows that the partition $V(\mathcal{H}) = \bigcup_{i \in [\ell]} \widehat{V}_i$ defined by

$$\widehat{V}_i = \begin{cases} V_{i_\star} \cup \{v\} & \text{if } i = i_\star, \\ V_i & \text{if } i \neq i_\star, \end{cases}$$

is a K^r_ℓ-coloring of $\mathcal{H}$. This completes the proof of Lemma 4.2.11. ∎

4.2.5 Expansions of Matchings of size 2.

In this subsection we shall derive Theorem 4.1.14 from Theorem 4.1.7. Again it is easy to see that $\mathcal{K}_{2r}^{M_2^r}$ is blowup-invariant and that the class $\mathfrak{S}^r$ is hereditary. Bene Watts, Norin, and Yepremyan proved in [17] that

$$\sup\{\lambda(\mathcal{G})\colon \mathcal{G} \text{ is } M_2^r\text{-free but not semibipartite}\} < \frac{(1-1/r)^{r-1}}{r!}$$

holds for all $r \geq 4$, where, let us recall, the numerator is the supremum of the edge densities of semibipartite r-graphs. Following the proof of Lemma 4.2.10 one easily deduces from this result that $\mathcal{K}_{2r}^{M_2^r}$ is symmetrized-stable with respect to $\mathfrak{S}^r$. So it only remains to establish vertex-extendibility, i.e., the following lemma.

Lemma 4.2.15. *Let $r \geq 4$ and $F = M_2^r$. There exist $\zeta > 0$ and $N_0 \in \mathbb{N}$ such that every $\mathcal{K}_{2r}^F$-free r-graph $\mathcal{H}$ on $n \geq N_0$ vertices satisfying $\delta(\mathcal{H}) \geq \left(\left(1-\frac{1}{r}\right)^{r-1}/(r-1)! - \zeta\right) n^{r-1}$ and possessing a vertex v for which $\mathcal{H} - v$ is semibipartite is semibipartite itself.*

In order to estimate the sizes of the vertex classes of semibipartite hypergraphs with almost the maximum number of edges we use the following estimate.

Fact 4.2.16. *If $r \geq 2$ and $x \in [0,1]$, then*

$$\frac{x(1-x)^{r-1}}{(r-1)!} + \frac{1}{r!}\left(1-\frac{1}{r}\right)^{r-3}\left(x-\frac{1}{r}\right)^2 \leq \frac{1}{r!}\left(1-\frac{1}{r}\right)^{r-1}.$$

Note that equality holds for $x = 1/r$ and $x = 1$.

Proof. The case $x = 1$ being clear we assume $x \in [0, 1)$ from now on. The standard inductive proof of Bernoulli's inequality also shows $(1 + 2h)(1 + h)^{r-2} \geq 1 + rh$ for every real $h \geq -1$. In particular, for $h = (x - 1/r)/(1 - x)$ we obtain

$$\left(\frac{1 - 1/r}{1 - x}\right)^{r-2} \frac{1 + x - 2/r}{1 - x} \geq \frac{(r - 1)x}{1 - x}\,.$$

Multiplying by $(1 - x)^r$ we deduce

$$\begin{aligned}(r - 1)x(1 - x)^{r-1} &\leq (1 - 1/r)^{r-2}(1 - x)(1 + x - 2/r)\\ &= (1 - 1/r)^{r-2}[(1 - 1/r)^2 - (x - 1/r)^2]\end{aligned}$$

and now it remains to divide by $(r - 1)(r - 1)!$. ∎

Proof of Lemma 4.2.15. Fix some sufficiently small $\zeta \ll r^{-1}$ and then some sufficiently large $N_0 \gg \zeta^{-1}$. Let $\mathcal{H}$ be a $\mathcal{K}^F$-free r-graph on $n \geq N_0$ vertices whose minimum degree is at least $\left(\left(1 - \frac{1}{r}\right)^{r-1}/(r - 1)! - \zeta\right) n^{r-1}$. Set $V = V(\mathcal{H})$ and suppose that for some vertex $v \in V$ the r-graph $\mathcal{H}_v = \mathcal{H} - v$ is semibipartite. Fix a partition $V(\mathcal{H}_v) = V_1 \cup V_2$ such that $|E \cap V_1| = 1$ holds for every $E \in \mathcal{H}_v$ and let $\widehat{\mathcal{S}}$ be the complete semibipartite r-graph on $V(\mathcal{H}_v)$ corresponding to this partition. Sets in $\widehat{\mathcal{S}} \setminus \mathcal{H}_v$ are called missing edges of $\mathcal{H}_v$, and for every $u \in V \setminus \{v\}$ sets in $L_{\widehat{\mathcal{S}}}(u) \setminus L_{\mathcal{H}_v}(u)$ are called missing edges of $L_{\mathcal{H}_v}(u)$.

As usual we have

$$\delta(\mathcal{H}_v) \geq \left(\left(1 - \frac{1}{r}\right)^{r-1} /(r-1)! - 2\zeta\right) n^{r-1} \quad \text{and} \quad |\mathcal{H}_v| \geq \left(\left(1 - \frac{1}{r}\right)^{r-1} /r! - 2\zeta\right) n^r.$$

In particular, the number of missing edges of $\mathcal{H}_v$ is at most $2\zeta n^r$.

Claim 4.2.17. *The following statements hold.*

(a) We have $|V_1| = \left(1/r \pm \zeta^{1/3}\right) n$ *and* $|V_2| = \left((r-1)/r \pm \zeta^{1/3}\right) n$.

(b) For every $u \in V(\mathcal{H}_v)$ *the number of missing edges of* $L_{\mathcal{H}_v}(u)$ *is at most* $\zeta^{1/3} n^{r-1}$.

(c) If $u \in V_1$*, then* $|V_2 \setminus N_{\mathcal{H}_v}(u)| \leq \zeta^{1/3} n$.

(d) If $u \in V_2$*, then* $|N_{\mathcal{H}_v}(u)| \geq \left(1 - \zeta^{1/3}\right) n$.

Proof. Setting $x = |V_1|/n$ we have

$$2\zeta > \frac{(1-1/r)^{r-1}}{r!} - \frac{|\widehat{\mathcal{S}}|}{n^r} > \frac{(1-1/r)^{r-1}}{r!} - \frac{x(1-x)^{r-1}}{(r-1)!}$$

and due to $\zeta \ll r^{-1}$ Fact 4.2.16 leads to $|x - 1/r| \leq O_r(\zeta^{1/2}) \leq \zeta^{1/3}$, which proves (a). Moroever, in $\widehat{\mathcal{S}}$ every vertex has degree $\left(\left(1 - \frac{1}{r}\right)^{r-1} /(r-1)! \pm O_r(\zeta^{1/2})\right) n^{r-1}$ and thus for every $u \in V(\mathcal{H}_v)$ there are at most $O_r(\zeta^{1/2}) n^{r-1}$ missing edges of $L_{\mathcal{H}_v}(u)$, which implies (b). Now for part (c) it suffices to observe that every vertex in $|V_2 \setminus N_{\mathcal{H}_v}(u)|$ belongs to $\Omega_r(n^{r-2})$ missing edges of $L_{\mathcal{H}_v}(u)$ and the argument for (d) is similar. ∎

Since $\mathcal{H}$ contains no weak expansion of M_2^r, there cannot exist two disjoint edges $E, E' \in \mathcal{H}$ such that $E \cup E'$ is 2-covered.

Claim 4.2.18. *If two distinct vertices* $u, w \in V(\mathcal{H})$ *satisfy*

$$|L_{\mathcal{H}}(u)[V_2]|, |L_{\mathcal{H}}(w)[V_2]| \geq \left(\left(1 - \frac{1}{r}\right)^{r-1} /(r-1)! - \zeta^{1/4} \right) n^{r-1},$$

then no edge of $\mathcal{H}$ *contains both of them.*

Proof. Assume contrariwise that some edge $E \in \mathcal{H}$ contains u and w. We shall show that this leads to two disjoint edges E_u, E_w of $\mathcal{H}$ such that $u \in E_u \subseteq V_2 \cup \{u\}$, $w \in E_w \subseteq V_2 \cup \{w\}$, and $E_u \cup E_w$ is 2-covered, which is absurd.

Owing to Claim 4.2.17 (a) and our assumption on the links of u and w we have

$$|V_2 \setminus N_{\mathcal{H}}(u)|, |V_2 \setminus N_{\mathcal{H}}(w)| \leq \zeta^{1/5} n.$$

The latter estimate and our lower bound on $|L_{\mathcal{H}}(u)[V_2]|$ show that there exists an edge $E_u \in \mathcal{H}$ such that $u \in E_u$ and $E_u \setminus \{u\} \subseteq V_2 \cap N_{\mathcal{H}}(w)$. Now Claim 4.2.17 (d) and our upper bound on $|V_2 \setminus N_{\mathcal{H}}(u)|$ imply that the set $V_2' = \bigcap_{x \in E_u} N_{\mathcal{H}}(x) \cap (V_2 \setminus E_u)$ has at least the size $|V_2'| \geq |V_2| - 2\zeta^{1/5} n$. Thus there exists an edge $E_w \in \mathcal{H}_v$ with $w \in E_w \subseteq V_2' \cup \{w\}$. Clearly E_u and E_w are as desired. ∎

By our lower bound on $\delta(\mathcal{H}_v)$ any two distinct vertices $u, w \in V_1$ satisfy the hypothesis of Claim 4.2.18, which has the following consequence.

Claim 4.2.19. *We have* $|E \cap V_1| \leq 1$ *for every* $E \in \mathcal{H}$. ∎

Notice that $d_{\mathcal{H}}(v) \geq \left(\left(1-\frac{1}{r}\right)^{r-1}/(r-1)! - \zeta\right) n^{r-1}$ yields

$$|N_{\mathcal{H}}(v)| \geq (1 - 1/r - O_r(\zeta))\, n \geq 2n/3,$$

whence

$$|N_{\mathcal{H}}(v) \cap V_2| \geq 2n/3 - |V_1| \geq n/3. \tag{4.7}$$

If there exists no edge $E_\star \in \mathcal{H}$ with $v \in E_\star \subseteq V_2 \cup \{v\}$, then $V(\mathcal{H}) = V_1 \cup (V_2 \cup \{v\})$ is a partition exemplifying that $\mathcal{H}$ is semibipartite and we are done. So we may suppose from now on that such an edge $E_\star$ exists. Consider the set $X = \bigcap_{w \in E_\star} N_{\mathcal{H}}(w)$. On the one hand, Claim 4.2.17 (d) and Equation 4.7 imply

$$|X \cap V_2| \geq |N_{\mathcal{H}}(v) \cap V_2| - (r-1)\zeta^{1/3} n \geq n/3 - n/12 = n/4.$$

On the other hand, there cannot exist an edge $E' \subseteq X \setminus E_\star$, for then $\{E_\star, E'\}$ would be a matching in $\mathcal{H}$ such that $E_\star \cup E'$ is 2-covered. Since there are at most $2\zeta n^r$ missing edges, this implies $|X \cap V_1| \leq O_r(\zeta) n \leq \zeta^{1/3} n$. As Claim 4.2.17 (d) yields $|N_{\mathcal{H}}(v) \setminus X| \leq (r-1)\zeta^{1/3} n$, we may conclude

$$|N_{\mathcal{H}}(v) \cap V_1| \leq |N_{\mathcal{H}}(v) \setminus X| + |X \cap V_1| \leq r\zeta^{1/3} n,$$

whence

$$|L_{\mathcal{H}}(v)[V_2]| \geq d_{\mathcal{H}}(v) - |N_{\mathcal{H}}(v) \cap V_1| n^{r-2} \geq \Big((1 - 1/r)^{r-1} / (r-1)! - \zeta^{1/4} \Big) n^{r-1}.$$

Now Claim 4.2.18 discloses $N_{\mathcal{H}}(v) \cap V_1 = \emptyset$. In view of Claim 4.2.19 this shows that the partition $V(\mathcal{H}) = (V_1 \cup \{v\}) \cup V_2$ witnesses the semibipartiteness of $\mathcal{H}$. ∎

4.3 Concluding remarks

• In this chapter we provided a framework for proving the degree-stability of certain classes of graph and hypergraph families, and applied it to the degree-stability of Σ_3, Σ_4, and $\mathcal{K}_\ell^F$ for some combinations of F and ℓ. In fact, one could push our results further and show that T_3, T_4, and H_ℓ^F (for some combinations of F and ℓ) are degree-stable by using the degree-stability results obtained here, applying the Removal lemma to prove the vertex-stability of T_3, T_4, and H_ℓ^F, respectively, and finally applying Theorem 4.1.8.

• Generalizing Theorem 4.1.9 one may attempt to characterize for arbitrary $\ell \geq r \geq 2$ the hypergraph families which are vertex-stable or degree-stable with respect to $\mathfrak{K}_\ell^r$. This problem is presumably very difficult and even partial results in this direction would be interesting.

• A classical example in hypergraph Turán theory suggested by Vera T. Sós is the Fano plane, i.e. the 3-graph on vertex set [7] with edge set

$$\{123, 345, 561, 174, 275, 376, 246\}.$$

The Turán density of the Fano plane was determined by De Caen and Füredi in [50]. Later Keevash and Sudakov [142] and, independently, Füredi and Simonovits [113] proved the degree-stability of the Fano plane and used it to determine the Turán number for large n. The complete determination of its Turán number was obtained only recently by Bellmann and the third author [16]. We do not know whether our method can be used to give another proof of the degree-stability of the Fano plane.

• Recall that by Theorem 4.1.1 every family $\mathcal{F}$ of graphs with $\chi(\mathcal{F}) = \ell + 1$ is edge-stable with respect to the family $\{T(n, \ell) \colon n \in \mathbb{N}\}$ of Turán graphs, which has the property that for every $n \in \mathbb{N}$ it contains a unique n-vertex graph. This state of affairs prompted the second author [193] to define for every nondegenerate family $\mathcal{F}$ of r-graphs the (edge-) stability number $\xi_e(\mathcal{F})$ to be the least number t such that there exists a class of r-graphs $\mathfrak{H}$ with the following properties:

- $\mathcal{F}$ is edge-stable with respect to $\mathfrak{H}$;
- for every $n \in \mathbb{N}$ there are t hypergraphs on n vertices in $\mathfrak{H}$.

For instance, the families studied in this article have stability number 1 and standard conjectures imply that the stability number of K_4^3 is infinite. It was shown recently [169; **?**] that for every $t \in \mathbb{N}$ there exists a family $\mathcal{M}_t$ of triple systems such that $\xi_e(\mathcal{M}_t) = t$.

In analogy with Definition 4.1.2 one can also define a vertex-stability number $\xi_v(\mathcal{F})$ and a degree-stability number $\xi_d(\mathcal{F})$. These satisfy the easy estimates $\xi_e(\mathcal{F}) \leq \xi_v(\mathcal{F}) \leq \xi_d(\mathcal{F})$ and it would be interesting to study how "exotic" these parameters can get.

• Our method can also be used in the context of other combinatorial structures, such as families of edge-weigthed graphs. To give an example, we recall the following result of Erdős, Hajnal, Sós, and Szemerédi [68] from Ramsey-Turán theory: For $r \geq 2$ every K_{2r}-free graph with n vertices and more than $\big(\frac{3r-5}{3r-2}+o(1)\big)n^2/2$ edges contains an independent set of size $o(n)$. Here the constant $\frac{3r-5}{3r-2}$ is optimal and the analogous problem with forbidden cliques of odd order is much easier. The proof of this result involves a certain family $\mathcal{F}_{2r}$ of graphs with weights from $\{0, 1/2, 1\}$ assigned to their edges. The main points of the argument are (i) that $\pi(\mathcal{F}_{2r}) = \frac{3r-5}{3r-2}$ and (ii) that the regularity method establishes a connection between $\mathcal{F}_{2r}$ and K_{2r}. Lüders and Reiher [180] recently obtained the sharper result that for $\delta \ll r^{-1}$ every K_{2r}-free graph with n vertices and more than $\big(\frac{3r-5}{3r-2}+\delta-\delta^2\big)n^2/2$ edges contains an independent set of size δn, where the term $\frac{3r-5}{3r-2}+\delta-\delta^2$ is again optimal. Their proof requires some stability result for the family $\mathcal{F}_{2r}$. In fact, they provide a rather ad-hoc proof of vertex-stability (see Proposition 5.5 in [180]) and returned to the topic in [181] proving degree-stability. A straightforward adaptation of the Ψ-trick to weighted graphs yields an alternative (and shorter) proof of the degree-stability of $\mathcal{F}_{2r}$.

• We would like to emphasize that the strongest general stability result in this article, Theorem 4.2.3, can also be used for giving reasonable quantitative versions of edge stability. For instance, combined with the results in Subsection 4.2.3 it tells us that if $\epsilon > 0$ is sufficiently small, then every Σ_4-free quadruple system $\mathcal{H}$ on a sufficiently large number n of vertices with more than $(1/256-\epsilon)n^4$ edges admits a partition $V(\mathcal{H}) = A \cup B \cup C \cup D \cup Z$ such that $|Z| \leq \epsilon^{1/2} n$ and $\mathcal{H} - Z$ is 4-partite with vertex classes A, B, C, and D. Moreover, all vertices in $V(\mathcal{H}) \setminus Z$

have at least the degree $(1/64 - 2\epsilon^{1/2})n^3$. Now a careful calculation shows $|A|, |B|, |C|, |D| = (1/4 \pm 6\epsilon^{1/2})n$ and the proof of Claim 4.2.8 discloses that the sets A, B, C, and D are independent in $\partial_2 \mathcal{H}$. By the proof of Claim 4.2.9, if some $z \in Z$ satisfies $|L_{\mathcal{H}}(z)[A, B, C]| \geq 4\epsilon^{1/2} n^3$, then z has no neighbours in D. So $\mathcal{H}$ can be made 4-partite by the deletion of at most $17\epsilon n^4$ edges, namely (i) at most ϵn^4 edges with two or more vertices in Z; (ii) at most $4\epsilon n^4$ edges $zabc$ with $z \in Z$, $a \in A$, $b \in B$, $c \in C$, and $|L_{\mathcal{H}}(z)[A, B, C]| \geq 4\epsilon^{1/2} n^3$; (iii) and, similarly, at most $4\epsilon n^4$ edges of each of the three the three types $zabd$, $zacd$, $zbcd$. In particular, the edge stability of Σ_4 with respect to $\mathfrak{K}_4^4$ holds with a linear dependence between the error terms. Taking into account that at most $400\epsilon n$ vertices $v \in V(\mathcal{H})$ can satisfy $d_{\mathcal{H}}(v) \leq n^3/80$ one can show the stronger result that $\mathcal{H}$ can be made K_4^4-colorable by the deletion of $7000\epsilon^{3/2} n^4$ edges, which seems to be a new result.

CHAPTER 5

HYPERGRAPHS WITH MANY EXTREMAL CONFIGURATIONS

Previously published as X. Liu and D. Mubayi. A hypergraph Turán problem with no stability. Combinatorica, pages 130, 2022.

5.1 Introduction

The classical Erdős–Stone theorem with the Erdős–Simonovits stability theorem imply that every graph family $\mathcal{F}$ is stable with respect to the Turán graph $T(n, \chi(\mathcal{F}) - 1)$. However, there are many Turán problems for hypergraphs that (perhaps) do not have the stability property. The example K_4^3 we mentioned in Chapter 2 was shown to have exponentially many extremal constructions in the number of vertices (see Kostochka [151] and Brown [32]). We will prove (Proposition 5.1.1) that these constructions can be used to show that K_4^3 does not have the stability property (assuming Conjecture 2.1.1 is true). For K_ℓ^3 with $\ell \geq 5$, different near-extremal constructions were given by Sidorenko [228], and Keevash and Mubayi [135]. Although we do not provide the details, these also show that K_ℓ^3 does not have stability (assuming Conjecture 2.1.1 is true).

The absence of stability seems to be a fundamental barrier in determining the Turán numbers of some families. Indeed, the Turán numbers of the examples we presented above are not known, even asymptotically, and in fact, no Turán number of a family without the stability property has been determined.

This chapter provides the first such example. In Section 5.1.1, we construct a finite family $\mathcal{M}$ of 3-graphs, prove that $\mathcal{M}$ does not have the stability property, and determine $\pi(\mathcal{M})$, and even $\mathrm{ex}(n, \mathcal{M})$ for infinitely many n (Theorems 5.1.4 and 5.1.7). In Section 5.1.2, we construction a finite family $\mathcal{M}_t$ of 3-graphs for every integer t and prove that $\mathcal{M}_t$ has exactly t different extremal configurations. In Section 5.1.3, we extend this construction to r-graphs for all $r \geq 4$.

In order to state our results formally, we need some definitions. Let $\ell \geq r \geq 2$ and $\mathcal{K}_{\ell+1}^r$ be the collection of all r-graphs F on at most $(\ell+1)+(r-2)\binom{\ell+1}{2}$ vertices such that for some $(\ell+1)$-set S, which will be called the core of F, every pair $\{u, v\} \subset S$ is covered by an edge in F [1].

Suppose that $\mathcal{T}$ is an r-graph on s vertices and $t = (t_1, \ldots, t_s)$ with each t_i a positive integer. Then the blowup $\mathcal{T}(t)$ of $\mathcal{T}$ is obtained from $\mathcal{T}$ by replacing each vertex i by a set of size t_i, and replacing every edge in $\mathcal{T}$ by the corresponding complete r-partite r-graph.

A family $\mathcal{F}$ is t-stable if there exist t near-extremal constructions, and every $\mathcal{F}$-free graph (or hypergraph) of size close to $\mathrm{ex}(n, \mathcal{F})$ is structurally close to one of these near-extremal constructions. The stability number of $\mathcal{F}$, denoted by $\xi(\mathcal{F})$, is the minimum integer t such that $\mathcal{F}$ is t-stable. If there is no such integer t, then we let $\xi(\mathcal{F}) = \infty$.

Although the concept of t-stable families was raised over a decade ago (see [193] and [208]), no example of t-stable families are known for any $t \geq 2$ before this work. However, if we assume that Turán's conjecture is true, then the following result shows that the stability number of K_4^3 is infinity.

Proposition 5.1.1. *If Conjecture 2.1.1 is true, then* $\xi(K_4^3) = \infty$.

---

[1] The original definition of $\mathcal{K}_{\ell+1}^r$ in [191] requires that F has at most $\binom{\ell+1}{2}$ edges. The new definition we used here will make our proofs simpler. Notice that $\mathcal{K}_{\ell+1}^r$ is a finite family in both definitions.

5.1.1 A 2-stable family of 3-graphs

In this section we construct a finite family $\mathcal{M}$ of 3-graphs that is 2-stable. Let us define the two extremal configurations first.

Definition 5.1.2. *Let $|A| = \lfloor n/3 \rfloor$ and $|B| = \lceil 2n/3 \rceil$ with $A \cap B = \emptyset$. Define*

$$\mathcal{G}_n^1 = \left\{ abb' \colon a \in A \text{ and } \{b, b'\} \subset B \right\}.$$

Let $\mathcal{G}_6^2$ be the 3-graph with vertex set $[6]$ whose complement is

$$\overline{\mathcal{G}_6^2} = \{123, 126, 345, 456\}.$$

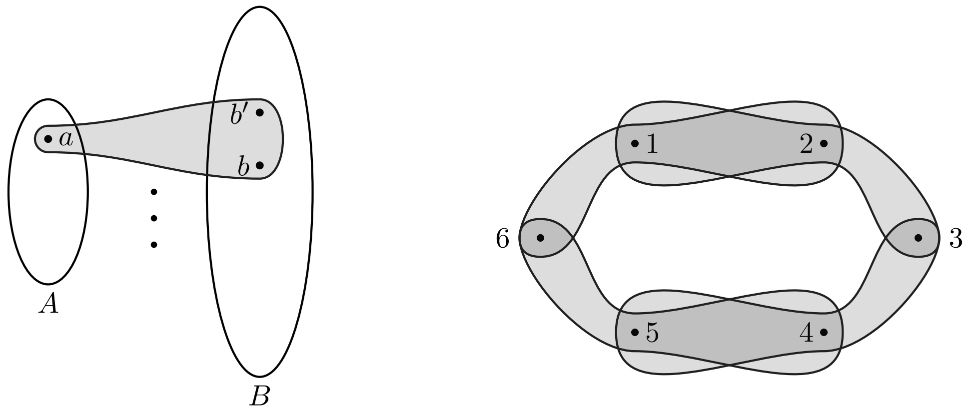
For $n > 6$ let $\mathcal{G}_n^2$ be a 3-graph on n vertices which is a blowup of $\mathcal{G}_6^2$ with the maximum number of edges.

Remarks.

- Notice that $\mathcal{G}_n^1$ is a (unbalanced) blowup of a star.
- Simple calculations show that each part in $\mathcal{G}_n^2$ has size either $\lfloor n/6 \rfloor$ or $\lceil n/6 \rceil$.
- For $i = 1, 2$, let $g_i(n) = |\mathcal{G}_n^i|$. Then $\lim_{n \to \infty} g_i(n)/n^3 = 2/27$.

Definition 5.1.3. *The family $\mathcal{M}$ is the union of the following three finite families.*

(a) M_1 is the set containing the complete 3-graph on five vertices with one edge removed, $M_1 = \left\{ K_5^{3-} \right\}$.



(a) The 3-graph $\mathcal{G}_n^1$.

(b) The complement of $\mathcal{G}_6^2$.

Figure 14. $\mathcal{G}^1$ and $\overline{\mathcal{G}_6^2}$.

(b) M_2 is the collection of all 3-graphs in $\mathcal{K}_7^3$ with a core whose induced subgraph has transversal number at least two.

(c) M_3 is the collection of all 3-graphs $F \in \mathcal{K}_6^3$ such that both $F \not\subset \mathcal{G}_n^1$ and $F \not\subset \mathcal{G}_n^2$ for all positive integers n.

Our first result is about the Turán number of $\mathcal{M}$.

Theorem 5.1.4. *The inequality $ex(n, \mathcal{M}) \leq 2n^3/27$ holds for all positive integers n. Moreover, equality holds whenever n is a multiple of six.*

Note that both $\mathcal{G}_n^1$ and $\mathcal{G}_n^2$ are $\mathcal{M}$-free and $g_1(n) \sim g_2(n) \sim 2n^3/27$. Moreover, it is easy to see that transforming $\mathcal{G}_n^1$ to $\mathcal{G}_n^2$ requires us to delete and add $\Omega(n^3)$ edges. Indeed, $\partial\mathcal{G}_n^1$ contains a clique on $\lfloor 2n/3 \rfloor$ vertices, whereas $\partial\mathcal{G}_n^2$ has clique number six. By Turán's theorem, one must thus delete strictly more that $(1-\pi(K_7))\binom{\lfloor 2n/3\rfloor}{2} = \Omega(n^2)$ edges from $\partial\mathcal{G}_n^2$ to obtain a copy of $\partial\mathcal{G}_n^2$. Since every edge in $\partial\mathcal{G}_n^1$ is covered by $\Omega(n)$ edges in $\mathcal{G}_n^1$, we need to remove at least $\Omega(n^3)$ edges from $\mathcal{G}_n^1$ before getting $\mathcal{G}_n^2$. So this proves that $\mathcal{M}$ does not have the stability property (in the sense of Theorem 4.1.1).

Our next result gives further detail about near-extremal $\mathcal{M}$-free constructions by showing that $\mathcal{M}$ is 2-stable with respect to $\mathcal{G}_n^1$ and $\mathcal{G}_n^2$. More precisely, it shows that $\xi(\mathcal{M}) = 2$.

Definition 5.1.5. *Let $\mathcal{H}$ be a 3-graph. Then $\mathcal{H}$ is called semibipartite if $V(\mathcal{H})$ has a partition $A \cup B$ such that $|E \cap A| = 1$ and $|E \cap B| = 2$ for all $E \in \mathcal{H}$, and $\mathcal{H}$ is called $\mathcal{G}_6^2$-colorable if it is a subgraph of a blowup of $\mathcal{G}_6^2$.*

With some calculations one can get the following observation.

Observation 5.1.6. *Let $\mathcal{H}$ be a 3-graph on n-vertices. If $\mathcal{H}$ is semibipartite, then $|\mathcal{H}| \le g_1(n)$. If $\mathcal{H}$ is $\mathcal{G}_6^2$-colorable, then $|\mathcal{H}| \le g_2(n)$.*

Theorem 5.1.7 (2-stability)**.** *There exist an absolute constant $\epsilon > 0$ such that the following holds for all sufficiently large n. Every $\mathcal{M}$-free 3-graph on n vertices with minimum degree at least $(4/9-\epsilon)\binom{n}{2}$ is either semibipartite or $\mathcal{G}_6^2$-colorable. Consequently, $\xi(\mathcal{M}) = 2$.*

Note that Theorem 5.1.7 is stronger than the requirement in the definition of 2-stability since removing at most δn vertices implies that the number of edges removed is at most δn^3 but not vice versa.

Theorem 5.1.4 together with Theorem 5.1.7 yield the following result.

Theorem 5.1.8. *The set* $\mathrm{proj}\Omega(\mathcal{M}) = [0, 1]$*, and* $g(\mathcal{M}, x) \leq 4/9$ *for all* $x \in [0, 1]$*. Moreover,* $g(\mathcal{M}, x) = 4/9$ *iff* $x \in \{5/6, 8/9\}$*.*

In words, Theorem 5.1.8 says that $\mathcal{M}$-free 3-graphs can have any possible shadow density but the edge density is maximized for exactly two values of the shadow densities (see Figure 15).

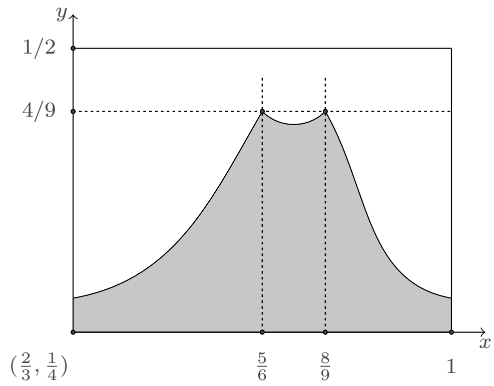


Figure 15. $g(\mathcal{M})$ has exactly two global maxima.

5.1.2 t-stable families of 3-graphs

Our main result in this section is that there exists a finite family $\mathcal{M}_t$ of 3-graphs that is t-stable for every positive integer t.

Theorem 5.1.9. *For every positive integer* t *there exist constants* $0 < n_1 < \cdots < n_t$*,* $0 < \lambda_t < 1/6$*,* t *triple systems* $\mathcal{G}_1, \ldots, \mathcal{G}_t$ *with* $v(\mathcal{G}_i) = n_i$ *for* $i \in [t]$*, and a finite family* $\mathcal{M}_t$ *of triple systems with the following properties.*

(a) *The inequality* $\mathrm{ex}\,(n, \mathcal{M}_t) \leq \lambda_t n^3$ *holds for all positive integers* n*, and moreover, equality holds whenever* n *is a multiple of* n_i *for some* $i \in [t]$*.*

(b) *For every* $\delta > 0$ *there exist* $\epsilon > 0$ *and* N_0 *so that the following holds for all* $n \geq N_0$*. Every* $\mathcal{M}_t$*-free triple system* $\mathcal{H}$ *on* n *vertices with at least* $(\lambda_t - \epsilon)n^3$ *edges can be made* $\mathcal{G}_i$*-colorable for some* $i \in [t]$ *by removing at most* δn *vertices. Moreover,* $\xi(\mathcal{M}_t) = t$*.*

Using Theorem 5.1.9 we obtain the following result.

Theorem 5.1.10. *For every positive integer* t *there exist constants* $0 < n_1 < \cdots < n_t$*,* $0 < \lambda_t < 1/6$*, and a finite family* $\mathcal{M}_t$ *of triple systems such that* $\mathrm{proj}\Omega(\mathcal{M}_t) = [0, 1]$*, and* $g(\mathcal{M}_t, x) \leq 6\lambda_t$ *for all* $x \in [0, 1]$*. Moreover,* $g(\mathcal{M}_t, x) = 6\lambda_t$ *if and only if* $x = 1 - 1/n_i$ *for some* $i \in [t]$*.*

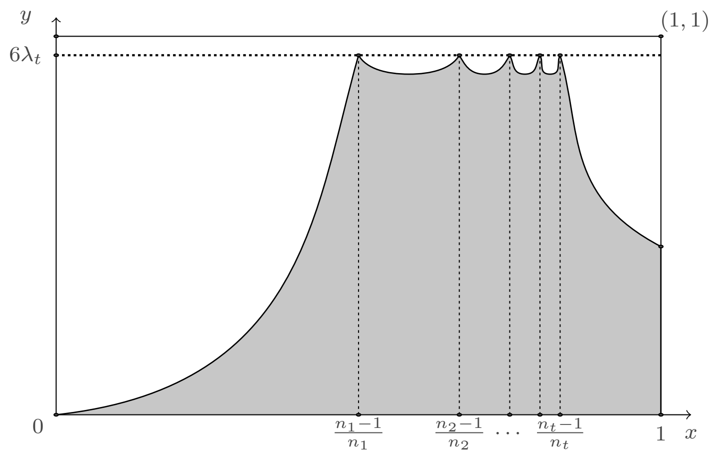


Figure 16. The function $g(\mathcal{M}_t)$ has exactly t global maxima.

Roughly speaking, the connection between these results is as follows. An r-graph is a *star* if there is a vertex v such that all edges contain v, and an r-graph $\mathcal{H}$ is semibipartite if it is $\mathcal{S}$-colorable for some star $\mathcal{S}$. Note that this is the same as saying that $V(\mathcal{H})$ has a partition into two parts A and B such that all edges have exactly one vertex in A and $r-1$ vertices in B. We will see later that our definition of $\mathcal{M}_t$ ensures that every semibipartite 3-graph is $\mathcal{M}_t$-free. By shrinking A, the shadow density of an n-vertex semibipartite 3-graph $\mathcal{H}$ can be made arbitrarily close to 1 as $n \to \infty$, so $\text{proj}\Omega(\mathcal{M}_t) = [0,1]$. The shadows of the triple systems $\mathcal{G}_1, \ldots, \mathcal{G}_t$ from Theorem 5.1.9 are complete graphs and thus their edge densities are the distinct numbers $1-1/n_1, \ldots, 1-1/n_t$. So $g(\mathcal{M}_t, x) = 6\lambda_t$ holds if x is one of those densities and stability allows us to exclude further solutions to this equation.

5.1.3 t-stable families of r-graphs

In this section, we extend the result in the previous section to r-graphs for all $r \geq 4$.

Theorem 5.1.11. *For every $r \geq 4$ and $t \geq 1$ there exist constants $0 < n_1 < \cdots < n_t$, $0 < \lambda_t^{(r)} < 1/r!$, a family $\{\mathcal{G}_1^r, \ldots, \mathcal{G}_t^r\}$ of r-graphs with $v(\mathcal{G}_i^r) = n_i + r - 3$ for $i \in [t]$, and a finite family $\mathcal{M}_t^r$ of r-graphs such that the followings hold.*

(a) The inequality $\text{ex}(n, \mathcal{M}_t^r) \leq \lambda_t^{(r)} n^r$ holds for all positive integers n, and moreover, equality holds whenever n satisfies $r \mid n$ and $rn_i \mid 3n$ for some $i \in [t]$.

(b) For every $\delta > 0$ there exist $\epsilon > 0$ and N_0 so that the following holds for all $n \geq N_0$. Every $\mathcal{M}_t^r$-free r-graph $\mathcal{H}$ on n vertices with at least $(\lambda_t^{(r)} - \epsilon)n^r$ edges is $\mathcal{G}_i^r$-colorable after removing at most δn vertices. Moreover, $\xi(\mathcal{M}_t^r) = t$.

Consequently, we obtain the following result about the feasible region of $\mathcal{M}_t^r$.

Theorem 5.1.12. *For every $r \geq 4$ and $t \geq 1$ there exist constants $0 < n_1 < \cdots < n_t$, $0 < \lambda_t^{(r)} < 1/r!$, and a finite family $\mathcal{M}_t^r$ of r-graphs such that* $\mathrm{proj}\Omega(\mathcal{M}_t^r) = [0, \hat{x}]$ *for some constant $\hat{x} \in (0,1]$, and $g(\mathcal{M}_t^r, x) \leq r!\lambda_t^{(r)}$ for all $x \in [0, \hat{x}]$. Moreover, $g(\mathcal{M}_t^r, x) = r!\lambda_t^{(r)}$ iff $x \in \{x_1, \ldots, x_t\}$, where*

$$x_i = \frac{(r-1)!}{r^{r-1}} \left((r-3) r^r \lambda_t^{(r)} + \frac{9}{2} \left(1 - \frac{1}{n_i} \right) \right) < \hat{x}, \quad \forall i \in [t].$$

Remark. It seems nontrivial to determine the exact value of $\hat{x}$. Here we show a lower bound for $\hat{x}$ which is obtained by optimizing the shadow density of a blow-up $\mathcal{G}_t^r$.

Let V be a set of size n and $V = \bigcup_{i \in [n_t + r - 3]} V_i$ be a partition with $|V_i| = (1 + o(1)) \, (1/r - \delta) \, n$ for $i \in [r-3]$ and $|V_i| = (1 + o(1)) \, (3/r + (r-3)\delta) \, n/n_i$ for $i \in [r-2, n_t + r - 3]$. Let $\mathcal{H}$ be the blow-up of $\mathcal{G}_t^r$ on V. Then

$$\begin{aligned} \lim_{n \to \infty} \frac{|\partial \mathcal{H}|}{n^{r-1}} &= (r-3) \left(\frac{1}{r} - \delta \right)^{r-4} \lambda_t \left(\frac{3}{r} + (r-3)\delta \right)^3 \\ &\quad + \left(\frac{1}{r} - \delta \right)^{r-3} \frac{n_t - 1}{2n_t} \left(\frac{3}{r} + (r-3)\delta \right)^2 . \end{aligned}$$

Letting $A = \frac{27(r-3)\lambda_t}{r^{r-1}}$ and $B = \frac{9}{2r^{r-1}}\left(1 - \frac{1}{n_t}\right)$, we obtain

$$\begin{aligned}
\lim_{n\to\infty} \frac{|\partial\mathcal{H}|}{n^{r-1}} &> A(1-r\delta)^{r-4}\left(1+\frac{r(r-3)}{3}\delta\right)^3 + B(1-r\delta)^{r-3}\left(1+\frac{r(r-3)}{3}\delta\right)^2 \\
&\geq A(1-r(r-4)\delta)\left(1+r(r-3)\delta\right) + B(1-r(r-3)\delta)\left(1+\frac{2r(r-3)}{3}\delta\right) \\
&\geq A\left(1+r\delta-r^4\delta^2\right) + B\left(1-\frac{r(r-3)}{3}\delta - r^4\delta^2\right) \\
&> A+B+rA\delta - \frac{r(r-3)}{3}B\delta - (A+B)r^4\delta^2.
\end{aligned}$$

Since $rA > \frac{r(r-3)}{3}B$, if $\delta > 0$ is sufficiently small, then

$$\hat{x} \geq (r-1)! \lim_{n\to\infty} \frac{|\partial\mathcal{H}|}{n^{r-1}} > (A+B)(r-1)! = x_t.$$

5.2 Proof for the 2-stable family

We prove Theorems 5.1.4, 5.1.7 and 5.1.8 in this section. Let us present some preliminary definitions and results first.

5.2.1 Preliminaries

The next standard lemma gives a relationship between $\lambda(\mathcal{T})$ and the size of a blowup of $\mathcal{T}$.

Lemma 5.2.1. *Let $r \geq 2$ and $\mathcal{T}$ and $\mathcal{H}$ be two r-graphs. Suppose that $\mathcal{H}$ is a blowup of $\mathcal{T}$ with $v(\mathcal{H}) = n$. Then $|\mathcal{H}| \leq \lambda(\mathcal{T})n^r$.*

Proof. Suppose that $|V(\mathcal{T})| = s$ and $\mathcal{H} = \mathcal{T}(t)$ for some $t = (t_1, \ldots, t_s)$. Then

$$|\mathcal{H}| = \sum_{E \in \mathcal{T}} \prod_{i \in E} t_i = n^r \sum_{E \in \mathcal{T}} \prod_{i \in E} \frac{t_i}{n} \leq \lambda(\mathcal{T})n^r,$$

where the last inequality follows from the definition of $\lambda(\mathcal{T})$ and $\sum_{i \in [s]} t_i = n$. ∎

The following lemma shows that $\mathcal{M}$ is blowup-invariant.

Lemma 5.2.2. *A 3-graph $\mathcal{H}$ is $\mathcal{M}$-free if and only if it is $\mathcal{M}$-hom-free.*

Proof. The backward implication is clear. Now suppose conversely that $\mathcal{H}$ fails to be $\mathcal{M}$-hom-free, i.e., that there is a homomorphism $f \colon V(F) \to V(\mathcal{H})$ for some $F \in \mathcal{M}$. If $F \cong K_5^{3-}$, then f is injective due to the fact that K_5^{3-} is 2-covered. However, this implies that $K_5^{3-} \subset \mathcal{H}$, a contradiction. Therefore, $F \in M_2 \cup M_3$. Clearly the restriction of f to the core S of F is injective. So $f(F) \in \mathcal{K}_{|S|}^3 \cap \mathcal{M}$ and in view of $f(F) \subset \mathcal{H}$ it follows that $\mathcal{H}$ fails to be $\mathcal{M}$-free. ∎

Lemma 5.2.3. *Suppose that $\mathcal{T}$ is a 3-graph with at most four vertices. Then $\lambda(\mathcal{T}) \leq 1/16$.*

Proof. Without loss of generality we may assume that $v(\mathcal{T}) = 4$ and $|\mathcal{T}| = 4$, i.e., $\mathcal{T} \cong K_4^3$. It is easy to see that

$$p_{K_4^3}(x) = x_1x_2x_3 + x_1x_2x_4 + x_1x_3x_4 + x_2x_3x_4 \leq 4(1/4)^3 = 1/16.$$

Therefore, $\lambda(\mathcal{T}) \leq 1/16$. ∎

Lemma 5.2.4. *Suppose that $\mathcal{T}$ is a 3-graph on five vertices with at most eight edges. Then $\lambda(\mathcal{T}) < 0.067277$.*

Lemma 5.2.5. $\lambda(\mathcal{G}_6^2) \leq 2/27$.

Proof. Notice that

$$\begin{aligned} p_{\mathcal{G}_6^2}(x_1, \ldots, x_6) &= x_3x_6(x_1 + x_2 + x_4 + x_5) \\ &\quad + (x_1 + x_2)(x_3 + x_6)(x_4 + x_5) + x_1x_2(x_4 + x_5) + x_4x_5(x_1 + x_2). \end{aligned}$$

Set $a = (x_3 + x_6)/2$, $b = (x_1 + x_2)/2$, $c = (x_4 + x_5)/2$, $d = (b + c)/2$. It follows from the AM-GM inequality that

$$\begin{aligned} p_{\mathcal{G}_6^2}(x_1, \ldots, x_6) \leq 2a^2(b + c) + 8abc + 2bc(b + c) &\leq 4a^2d + 8ad^2 + 4d^3 \\ &= 2\left((a + d) \cdot (a + d) \cdot 2d\right) \\ &\leq 2\left(\frac{(a + d) + (a + d) + 2d}{3}\right)^3 = \frac{2}{27}. \end{aligned}$$

∎

Lemma 5.2.6. *Let $\mathcal{T}$ be a 2-covered 3-graph on $k \geq 7$ vertices. Suppose that $\tau(\mathcal{T}[S]) \leq 1$ for all sets $S \subset V(\mathcal{T})$ with $|S| = 7$. Then $\mathcal{T}$ is a star.*

Remark. In fact, a weaker condition that $|S| = 6$ is sufficient for the proof of Lemma 5.2.6.

Proof. Suppose that $\mathcal{T}$ is not a star. Then for every vertex v in $\mathcal{T}$ there exists an edge E_v in $\mathcal{T}$ that does not contain v.

First notice that $\mathcal{T}$ cannot contain two disjoint edges. Therefore, $\mathcal{T}$ is intersecting. Suppose that $\mathcal{T}$ contains two edges $E_1 = \{u, v_1, v_2\}$ and $E_2 = \{u, w_1, w_2\}$, where $\{v_1, v_2\} \cap \{w_1, w_2\} = \emptyset$. Let $E_3 \in \mathcal{T}$ be an edge that does not contain u. Since $\mathcal{T}$ is intersecting, we may assume that $v_1, w_1 \in E_3$. Then, we have $|E_1 \cup E_2 \cup E_3| \leq 6$, and $\tau(\{E_1, E_2, E_3\}) = 2$, a contradiction. Therefore, we may assume that the intersection of every two edges in $\mathcal{T}$ has size two. Let $E_1 = \{u, v, w_1\}$ and $E_2 = \{u, v, w_2\}$ be two edges in $\mathcal{T}$. By assumption there exists an edge $E_3 \in \mathcal{T}$ that does not contain u and, hence, we have $E_3 = \{v, w_1, w_2\}$. Similarly there exists $E_4 \in \mathcal{T}$ that does not contain v and, hence, we have $E_4 = \{u, w_1, w_2\}$. Then, we have $|E_1 \cup E_2 \cup E_3 \cup E_4| = 4$, and $\tau(\{E_1, E_2, E_3, E_4\}) = 2$, a contradiction. ∎

5.2.2 Proof of Theorem 5.1.4

In this section we complete the proof of Theorem 5.1.4.

Proof of Theorem 5.1.4. Let $\mathcal{H}$ be an $\mathcal{M}$-free 3-graph on n vertices. By Theorem 4.1.4, we may assume that $\mathcal{H}$ is symmetrized. Let $T \subset V(\mathcal{H})$ such that T contains exactly one vertex from each equivalent class in $\mathcal{H}$, and let $\mathcal{T} = \mathcal{H}[T]$. Since $\mathcal{H}$ is a blowup of $\mathcal{T}$, by Lemma 5.2.1, it suffices to show that $\lambda(\mathcal{T}) \leq 2/27$. Next, we will consider two cases depending on the size of T: either $|T| \geq 7$ or $|T| \leq 6$.

Case 1: $|T| \geq 7$.

Since $\mathcal{T}$ is 2-covered and it is M_2-free, $\tau(\mathcal{T}[S]) \leq 1$ for all $S \subset T$ with $|S| = 7$, and it follows from Lemma 5.2.6 that $\mathcal{T}$ is a star.

Let us calculate $\lambda(\mathcal{T})$. We may assume that $V(\mathcal{T}) = [s]$ for some integer s and 1 is the center of $\mathcal{T}$. Then,

$$p_{\mathcal{T}}(x) \leq x_1 \left(\sum_{\{i,j\} \subset [s] \setminus \{1\}} x_i x_j \right) \leq \frac{s-2}{2(s-1)} x_1 (1-x_1)^2 < \frac{1}{2} x_1 (1-x_1)^2 \leq \frac{2}{27},$$

which implies that $\lambda(\mathcal{T}) < 2/27$.

Case 2: $|T| \leq 6$.

If $|T| \leq 5$, then Lemmas 5.2.3 and 5.2.4 imply that $\lambda(\mathcal{T}) < 0.67277$. So we may assume that $|T| = 6$.

Since $\mathcal{T}$ is 2-covered, $\mathcal{T} \in \mathcal{K}_6^3$. Since $\mathcal{H}$ does not contain any member in M_3 as a subgraph, either $\mathcal{T} \subset \mathcal{G}_n^1$ or $\mathcal{T} \subset \mathcal{G}_n^2$ for some $n \geq 6$. Due to the fact that $\mathcal{T}$ is 2-covered again, either $\mathcal{T}$ is a star or $\mathcal{T} \subset \mathcal{G}_6^2$. The former case has been handled by Case 1, so we may assume that $\mathcal{T} \subset \mathcal{G}_6^2$, and it follows from Lemma 5.2.5 that $\lambda(\mathcal{T}) \leq \lambda(\mathcal{G}_6^2) \leq 2/27$. ∎

5.2.3 Proof of Theorem 5.1.7

Let $\mathfrak{G}_1$ be the collection of all semibipartite 3-graphs and let $\mathfrak{G}_2$ be the collection of all $\mathcal{G}_6^2$-colorable 3-graphs. According to Theorem 4.1.7, it suffices to prove the vertex-extendability of $\mathcal{M}$ with respects to $\mathfrak{G}_1$ and $\mathfrak{G}_2$.

Lemma 5.2.7. *The family $\mathcal{M}$ is vertex-extendable with respects to $\mathfrak{G}_1$.*

We will use the following stability theorem due to Füredi, Pikhurko, and Simonovits [112] to prove Lemma 5.2.7.

Let $\mathbb{F}_{3,2}$ be the 3-graph with vertex set $[5]$ and edges set $\{123, 124, 125, 345\}$. Füredi, Pikhurko, and Simonovits [112] proved that if n is sufficiently large, then $\mathcal{G}_n^1$ is the unique $\mathbb{F}_{3,2}$-free 3-graph on n vertices with the maximum number of edges. Moreover, they proved the following strong stability result.

Theorem 5.2.8 (Füredi–Pikhurko–Simonovits [112]). *Let $\gamma \le 1/125$ be fixed and $n \ge n_0$. Let $\mathcal{H}$ be an $\mathbb{F}_{3,2}$-free 3-graph on n vertices with $\delta(\mathcal{H}) > (4/9 - \gamma)\binom{n}{2}$. Then $\mathcal{H}$ is semibipartite.*

Proof of Lemma 5.2.7. Let $\mathcal{H}$ be an $(n+1)$-vertex $\mathcal{M}$-free 3-graph with $\delta(\mathcal{H}) \ge \left(\frac{4}{9} - \epsilon\right)\binom{n}{2}$. Let $v \in V(\mathcal{H})$ such that $\mathcal{H}' = \mathcal{H} - v$ is semibipartite. We need to prove that $\mathcal{H}$ is also semibipartite.

Let $A \cup B = V(\mathcal{H}) \setminus \{v\}$ be a partition such that every edge in $\mathcal{H}'$ has exactly one vertex in A.

Claim 5.2.9. *We have $||A| - n/3| < \epsilon^{1/2} n$ and $||B| - 2n/3| < \epsilon^{1/2} n$.*

Proof. Let $\beta = |B|$. Since $\mathcal{H}'$ is semibipartite,

$$|\mathcal{H}'| \leq (n-\beta)\binom{\beta}{2}.$$

On the other hand, it is easy to see from the minimum degree assumption that $|\mathcal{H}'| \geq (4/9 - \epsilon)\binom{n}{3}$. Therefore,

$$(4/9-\epsilon)\binom{\tilde{n}}{3} \leq (n-\beta)\binom{\beta}{2},$$

which implies that $(2/3 - \epsilon^{1/2})n < \beta < (2/3 + \epsilon^{1/2})n$. ∎

Claim 5.2.10. *We have* $|N(v) \cap B| \geq \left(1/3 - 2\epsilon^{1/2}\right) n$.

Proof. By assumption, we have

$$\binom{|N(v)|}{2} \geq d(v) \geq \left(\frac{4}{9} - \epsilon\right)\binom{n}{2},$$

which implies that $|N(v)| \geq (2/3 - \epsilon)n$. By Claim 5.2.9, $|A| \leq (1/3 + \epsilon^{1/2})\tilde{n}$, and hence

$$|N(v) \cap B| \geq (2/3 - \epsilon)\, n - \left(1/3 + \epsilon^{1/2}\right) n > \left(1/3 - 2\epsilon^{1/2}\right) n.$$

∎

Claim 5.2.11. *For every vertex* $w \in V(\mathcal{H}) \setminus \{v\}$ *we have* $|N(w) \cap B| \geq |B| - 4\epsilon^{1/2}n$.

Proof. **Case 1:** $w \in A$.

Let $Z_w = B \setminus N(w)$ and $z_w = |Z_w|$. If follows from our assumption $d(v) \geq \left(\frac{4}{9} - \epsilon\right)\binom{n}{2}$ and Claim 5.2.9 that

$$\frac{1}{2}\left(\frac{2}{3} + \epsilon^{1/2} - z_w\right)^2 n^2 \geq \binom{|B \setminus Z_w|}{2} \geq \left(\frac{4}{9} - \epsilon\right)\binom{n}{2}.$$

It follows that $z_w \leq 2\epsilon^{1/2}n$.

Case 2: $w \in B$.

Let $Z_w = V(\mathcal{H}) \setminus N(w)$ and $z_w = |Z_w|$. If follows from our assumption $d(v) \geq \left(\frac{4}{9} - \epsilon\right)\binom{n}{2}$ and Claim 5.2.9 that

$$\left(\frac{1}{3} + \epsilon^{1/2}\right)\left(\frac{2}{3} + \epsilon^{1/2} - z_w\right) \geq |A \setminus Z_w||B \setminus Z_w| \geq \left(\frac{4}{9} - \epsilon\right)\binom{n}{2}.$$

It follows that $z_w \leq 4\epsilon^{1/2}n$. ∎

We may assume that $\mathcal{H}$ contains a copy of $\mathbb{F}_{3,2}$, since otherwise by Theorem 5.2.8 we are done. Let $S \subset V(\mathcal{H})$ be a set of size 5 such that $\mathbb{F}_{3,2} \subset \mathcal{H}$. Observe that $v \in S$. Let $\{w_1, w_2, w_3, w_4\} = S \setminus \{v\}$. Define $B' = B \cap N(v) \cap \left(\bigcap_{j\in[4]} N(w_j)\right)$. Then Claims 5.2.10 and 5.2.11 imply that $|B'| \geq \left(1/3 - \epsilon^{1/2}\right)n - 4 \times 4\epsilon^{1/2}n > n/6$. Fix a vertex $u \in A$ (it is possible that $u \in \{w_1, w_2, w_3, w_4\}$). that there exists an edge $w_5w_6 \in L(u)[B']$. Let $E \subset \mathcal{H}$ be a set of edges of size at most 10 that covers all pairs in $\{v, w_1, w_2, w_3, w_4\} \times \{w_5, w_6\}$, and let $F = \mathcal{H}[\{v, w_1, w_2, w_3, w_4\}] \cup \{uw_5w_6\} \cup E$. Then it is easy to see that F is a member in M_2 (since $\mathbb{F}_{3,2} \subset \mathcal{H}[\{v, w_1, w_2, w_3, w_4\}]$ has transversal number at least two), a contradiction. ∎

Lemma 5.2.12. *The family $\mathcal{M}$ is vertex-extendable with respects to $\mathfrak{G}_2$.*

Recall the following lemma.

Lemma 5.2.13. *Fix a real $\eta \in (0,1)$ and integers $m, n \geq 1$. Let $\mathcal{G}$ be a 3-graph with vertex set $[m]$ and let $\mathcal{H}$ be a further 3-graph with $v(\mathcal{H}) = n$. Consider a vertex partition $V(\mathcal{H}) = \bigcup_{i\in[m]} V_i$ and the associated blowup $\widehat{\mathcal{G}} = \mathcal{G}[V_1, \ldots, V_m]$ of $\mathcal{G}$. If two sets $T \subseteq [m]$ and $S \subseteq V$ (we allow S to contain vertices from V_i for $i \in T$) have the properties*

(a) $|V_j| \geq (|S|+1)|T|\eta^{1/3}n + |S|$ *for all* $j \in T$,

(b) $|\mathcal{H}[V_{j_1}, V_{j_2}, V_{j_3}]| \geq |\widehat{\mathcal{G}}[V_{j_1}, V_{j_2}, V_{j_3}]| - \eta n^3$ *for all* $\{j_1, j_2, j_3\} \in \binom{T}{3}$, *and*

(c) $|L_{\mathcal{H}}(v)[V_{j_1}, V_{j_2}]| \geq |L_{\widehat{\mathcal{G}}}(v)[V_{j_1}, V_{j_2}]| - \eta n^3$ *for all* $v \in S$ *and* $\{j_1, j_2\} \in \binom{T}{2}$.

then there exists a selection of vertices $u_j \in V_j \setminus S$ for all $j \in [T]$ such that $U = \{u_j \colon j \in T\}$ satisfies $\widehat{\mathcal{G}}[U] \subseteq \mathcal{H}[U]$ and $L_{\widehat{\mathcal{G}}}(v)[U] \subseteq L_{\mathcal{H}}(v)[U]$ for all $v \in S$. In particular, if $\mathcal{H} \subseteq \widehat{\mathcal{G}}$, then $\widehat{\mathcal{G}}[U] = \mathcal{H}[U]$ and $L_{\widehat{\mathcal{G}}}(v)[U] = L_{\mathcal{H}}(v)[U]$ for all $v \in S$.

Proof of Lemma 5.2.12. Let $\mathcal{H}$ be an $(n+1)$-vertex $\mathcal{M}$-free 3-graph with $\delta(\mathcal{H}) \geq \left(\frac{4}{9} - \epsilon\right)\binom{n}{2}$. Let $v \in V(\mathcal{H})$ such that $\mathcal{H}' = \mathcal{H} - v$ is $\mathcal{G}_6^2$-colorable. We need to prove that $\mathcal{H}$ is also $\mathcal{G}_6^2$-colorable.

Let $V = V(\mathcal{H})$, $V' = V \setminus \{v\}$, and

$$\mathcal{P} = \{V_1, \ldots, V_6\}.$$

be the set of six parts in $\mathcal{H}'$ such that there is no edge between $V_1V_2V_3$, $V_1V_2V_6$, $V_3V_4V_5$, and $V_4V_5V_6$ (and every edge in $\mathcal{H}$ hits at most one vertex in V_j for every $j \in [6]$). Let $y_i = |V_i|/n$ for $i \in [6]$.

First we will prove several claims about sets in $\mathcal{P}$. Since V_1 is a representative for sets in $\{V_1, V_2, V_4, V_5\}$ and V_3 is a representative for sets in $\{V_3, V_6\}$, we shall only prove the statements for V_1 and V_3, and by symmetry, the statements hold for all sets in $\mathcal{P}$.

Claim 5.2.14. *We have* $||V_i| - n/6| < 5\epsilon^{1/2}n$ *for* $i \in [6]$.

Proof. First, it follows from the minimum degree assumption that $|\mathcal{H}'| \geq (4/9 - \epsilon)\binom{n}{3}$. On the other hand, since $\mathcal{H}'$ is $\mathcal{G}_6^2$-colorable, we have

$$|\mathcal{H}'| \leq p_{\mathcal{G}_6^2}(y)n^3. \tag{5.1}$$

Therefore, we have

$$p_{\mathcal{G}_6^2}(y_1, \ldots, y_6) \geq (4/9 - \epsilon)\binom{n}{3}/n^3 \geq \frac{2}{27} - \epsilon.$$

Let $a = (y_3 + y_6)/2$, $b = (y_1 + y_2)/2$, $c = (y_4 + y_5)/2$, $d = (b + c)/2$ and recall from the proof of Lemma 5.2.5 that

$$
\begin{aligned}
p_{\mathcal{G}_6^2}(y_1, \dots, y_6) &= y_3 y_6 (y_1 + y_2 + y_4 + y_5) \\
&\quad + (y_1 + y_2)(y_3 + y_6)(y_4 + y_5) + y_1 y_2 (y_4 + y_5) + y_4 y_5 (y_1 + y_2) \\
&\leq 2a^2(b + c) + 8abc + 2bc(b + c) \leq 4a^2 d + 8ad^2 + 4d^3 = 2\left((a + d) \cdot (a + d) \cdot 2d\right).
\end{aligned}
$$

Therefore,

$$(a + d) \cdot (a + d) \cdot 2d \geq 1/27 - \epsilon/2, \tag{5.2}$$

and

$$4d(a^2 - y_3 y_6) \leq \epsilon, \quad 4d(d^2 - bc) \leq \epsilon, \quad 2c(b^2 - y_1 y_2) \leq \epsilon, \quad 2b(c^2 - y_4 y_5) \leq \epsilon. \tag{5.3}$$

Now Equation 5.2 and $2a + 4d = 1$ yield

$$
\begin{aligned}
\epsilon/2 \geq 1/27 - (a + d)^2 \cdot 2d = 1/27 - (1 + 2a)^2(1 - 2a)/32 &= (a - 1/6)^2(a/4 + 5/24) \\
&\geq (a - 1/6)^2/8,
\end{aligned}
$$

whence $|a - 1/6| \le 2\epsilon^{1/2}$. By $2|a-1/6| = 4|d-1/6|$ this implies $|d - 1/6| \le \epsilon^{1/2}$. Since ϵ is sufficiently small, it follows that $a, d \ge 1/8$. So the first inequality in Equation 5.3 leads to $(y_3 - y_6) \le 8\epsilon$, whence $|y_3 - y_6| \le 3\epsilon^{1/2}$. By the triangle inequality we obtain

$$2|y_3 - 1/6| \le |y_3 - y_6| + |y_3 + y_6 - 1/3| \le 3\epsilon^{1/2} + 2|a - 1/6| \le 7\epsilon^{1/2},$$

which shows $|y_3 - 1/6| \le 4\epsilon^{1/2}$. Similarly, $|y_6 - 1/6| \le 4\epsilon^{1/2}$. Applying the same reasoning to the other estimates in Equation 5.3 we obtain first $|b - 1/6|, |c - 1/6| \le 3\epsilon^{1/2}$ and then $|y_i - 1/6| \le 5\epsilon^{1/2}$ for every $i \in \{1, 2, 4, 5\}$. ∎

Denote by $\mathcal{G}$ the blowup $\mathcal{G}_6^2[V_1, \ldots, V_6]$ of $\mathcal{G}_6^2$, and notice that $\mathcal{H}' \subset \mathcal{G}$. For $j \in [6]$ fix a vertex $a_j \in V_j$, let $\widetilde{G}_j = L_{\mathcal{G}}(a_j)$, $G_j = \widetilde{G}_j[\{a_1, \ldots, a_6\} \setminus \{a_j\}]$, and notice that $\widetilde{G}_j$ is a graph on $V(\mathcal{H}) \setminus (V_j \cup \{v\})$ and is a blowup of G_j (see Figure 17).

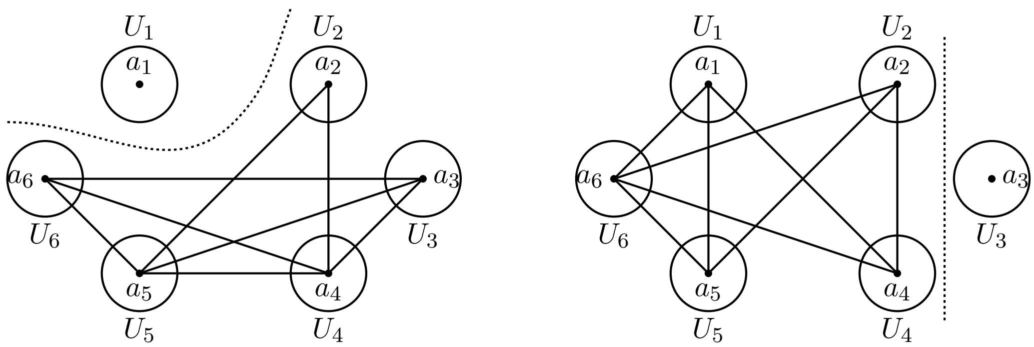
Since $\mathcal{H}$ is $\mathcal{G}_6^2$-colorable, $L_{\mathcal{H}'}(w) \subset \widetilde{G}_j$ for all $j \in [6]$ and $w \in V_j$. For every $j \in [6]$ and every $w \in V_j$ let

$$M(w) = \left\{ w_1 w_2 \in \widetilde{G}_j \colon w_1 w_2 \notin L_{\mathcal{H}'}(w) \right\},$$

and call members in $M(w)$ missing edges of $L_{\mathcal{H}'}(w)$.

Claim 5.2.15. *We have $|M(w)| \le 20\epsilon^{1/2} n^2$ for every $w \in V'$.*

Proof. We shall only prove the case $w \in V_1$, since the arguments for other cases are similar. Fix a vertex $w \in V_1$. Let $\widehat{G}_1$ be the blowup of G_1 obtained by replacing each vertex in $V(G_1)$



(a) The graph G_1 is the 5-vertex graph above, and $\widetilde{G}_1$ is a blowup of G_1.

(b) The graph G_3 is the 5-vertex graph above, and $\widetilde{G}_3$ is a blowup of G_3.

Figure 17. Graphs G_1 and G_3.

with the set in $\mathcal{P}$ that contains it. Since $\mathcal{H}'$ is $\mathcal{G}_6^2$-colorable, $L_{\mathcal{H}'}(w) \subset \widehat{G}_1$. On the other hand, it follows from Claim 5.2.14 that

$$|M(w)| = |\widehat{G}_1| - |L_{\mathcal{H}'}(w)| < 8\left(1/6 + 5\epsilon^{1/2}\right)^2 n^2 - (4/9 - \epsilon)\binom{n}{2} < 20\epsilon^{1/2}n^2.$$

∎

By assumption and Claim 5.2.15, $\mathcal{H}$ and $\mathcal{G}$ satisfy the following statements, which will be useful later when we applying Lemma 5.3.23.

(a) $|\mathcal{H}[A_1, A_2, A_3]| \geq |\mathcal{G}[A_1, A_2, A_3]| - \epsilon n^3$ for every triple $\{A_1, A_2, A_3\} \subset \mathcal{P}$, and

(b) $|L_{\mathcal{H}}(u)[A_1, A_2]| \geq |L_{\mathcal{G}}(u)[A_1, A_2]| - 20\epsilon^{1/2}n^2$ for every $u \in V'$ and every pair $\{A_1, A_2\} \subset \mathcal{P}$ satisfying $u \notin A_1 \cup A_2$.

Claim 5.2.16. *Let $j \in [6]$ and $w \in V_j$. Then $|N(w) \cap (V' \setminus V_j)| > |V' \setminus U_j| - 150\epsilon^{1/2}\tilde{n}$.*

Proof. We shall only prove the case that $j = 1$, since the arguments for other cases are similar. Let $w \in V_1$ and $V'' = V' \setminus V_1$. Since $\delta(G_1) \geq 2$ and $\widetilde{G}_1$ is a blowup of G_1, it follows from Claim 5.2.14 that

$$\delta(\widetilde{G}_1) > 2\left(1/6 - 5\epsilon^{1/2}\right) n \geq \left(1/3 - 10\epsilon^{1/2}\right) n.$$

So it follows from Claim 5.2.15 that the number of vertices in V'' with degree 0 in $L_{\mathcal{H}'}(w)$ is at most

$$\frac{2|M(w)|}{\delta(\widetilde{G}_1)} < \frac{40\epsilon^{1/2}n^2}{\left(1/3 - 10\epsilon^{1/2}\right) n} < 150\epsilon^{1/2}n.$$

∎

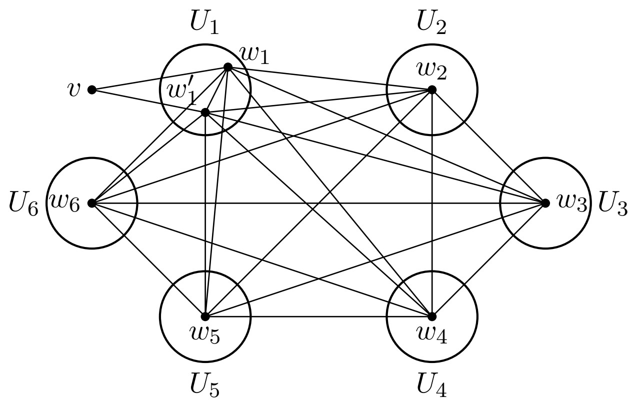


Figure 18. The 3-graph $F = \mathcal{H}_i[\{w_1, w_2, \ldots, w_6\}] \cup \mathcal{H}_i[\{w_1', w_2, \ldots, w_6\}] \cup \{vw_1w_1'\}$ is a member in M_2 with core $\{w_1, w_1', w_2, \ldots, w_6\}$.

Claim 5.2.17. *We have $L(v)[V_i] = \emptyset$ for $i \in [6]$.*

Proof. Suppose to the contrary that there exists an edge $w_1w_1' \in L(v)[V_i]$ for some $i \in [6]$. We shall only prove the case $i = 1$, since the arguments for other cases are similar. It follows from Claim 5.2.16 that

$$|N(w_1) \cap N(w_1') \cap (V' \setminus V_1)| > |V' \setminus V_1| - 300\epsilon^{1/2}n. \tag{5.4}$$

Applying Lemma 5.3.23 with $S = \{w_1, w_1'\}$, $T = [2, 6]$, and $\eta = 20\epsilon^{1/2}$ we obtain $w_j \in V_j$ for $j \in [2, 6]$ (see Figure 18) such that the induced subgraphs of $\mathcal{H}$ on sets $\{w_1, w_2, \dots, w_6\}$ and $\{w_1', w_2, \dots, w_6\}$ are isomorphic to $\mathcal{G}_6^2$. Let $F = \mathcal{H}[\{w_1, w_2, \dots, w_6\}] \cup \mathcal{H}[\{w_1', w_2, \dots, w_6\}] \cup \{vw_1w_1'\}$. Then it is easy to see that $F \in M_2$ with core $\{w_1, w_1', w_2, \dots, w_6\}$ (see Figure 18), a contradiction.

∎

Claim 5.2.18. *There is at most one set $A \in \mathcal{P}$ such that $|N(v) \cap A| < n/48$.*

Proof. Let $V_j' = N(v) \cap V_j$ for $j \in [6]$. By Claim 5.2.17, $L(v)$ is a 6-partite graph (not necessarily complete) on V'. Suppose to the contrary that there are exist two distinct $i, j \in [6]$ such that $|V_j'| \leq n/48$. Then, by Claim 5.2.14,

$$|L(v)| \leq 6\left(1/6 + 5\epsilon^{1/2}\right)^2 n^2 + (n/48)^2 + 8 \times n/48 \times \left(1/6 + 5\epsilon^{1/2}\right) n < \left(2/9 - 10\epsilon^{1/2}\right) n^2,$$

which contradicts the minimum degree assumption. ∎

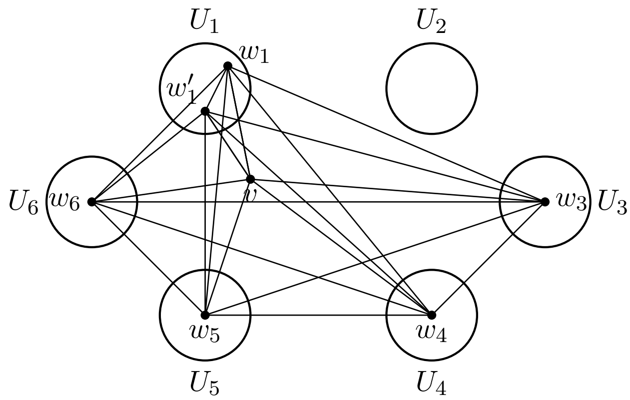


Figure 19. The 3-graph

$F = \mathcal{H}_i[\{w_1, w_3, \ldots, w_6\}] \cup \mathcal{H}_i[\{w_1', w_3, \ldots, w_6\}] \cup \{vw_1w_1'\} \cup \{e_j : j \in [3, 6]\}$ is a member in M_2 with core $\{v, w_1, w_1', w_3, \ldots, w_6\}$. In particular, $\tau(\{w_1w_3w_4, w_1'w_5w_6\}) > 1$.

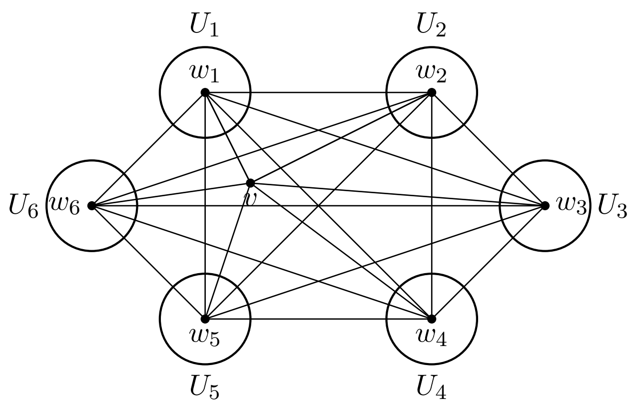


Figure 20. The 3-graph $F = \mathcal{H}_i[\{w_1, \ldots, w_6\}] \cup \{e_j : j \in [6]\}$ is a member in M_2 with core $\{v, w_1, \ldots, w_6\}$. In particular, $\tau(\{w_1w_3w_4, w_2w_5w_6\}) > 1$.

Claim 5.2.19. *There exists $i \in [6]$ such that $N(v) \cap V_i = \emptyset$.*

Proof. Suppose to the contrary that $N(v) \cap V_i \neq \emptyset$ for all $i \in [6]$. By Claim 5.2.18, there are at least five sets $A \in \mathcal{P}$ with $|A \cap N(v)| \geq n/48$. We shall only prove the case that $|N(v) \cap V_i| \geq n/48$ for $i \in [6] \setminus \{1\}$, since the arguments for other cases are similar.

Fix a vertex $w_1 \in N(v) \cap V_1$. Let $V_j' = V_j \cap N(v)$ for $i \in [2, 6]$. By assumption, $|V_j'| \geq n/48$ for $j \in [2, 6]$. So applying Lemma 5.3.23 with $T = \{w_1\}$, $S = [2, 6]$, and $\eta = 20\epsilon^{1/2}$ we obtain $w_j \in V_j'$ for $j \in [2, 6]$ (see Figure 20) such that the induced subgraph of $\mathcal{H}$ on $\{w_1, \ldots, w_6\}$ is isomorphic to $\mathcal{G}_6^2$. For $j \in [6]$ let $e_j \in \mathcal{H}$ be an edge containing v and w_j. Define $F = \mathcal{H}[\{w_1, \ldots, w_6\}] \cup \{e_j \colon j \in [6]\}$. Then it is easy to see that F is a member in M_2 with core $\{v, w_1, \ldots, w_6\}$ (see Figure 20), a contradiction. ∎

Our next step is to show that $\mathcal{H}$ is $\mathcal{G}_6^2$-colorable with the sets of parts $\widetilde{\mathcal{P}}$, where $\widetilde{\mathcal{P}}$ is obtained from $\mathcal{P}$ by replacing A with $A \cup \{v\}$ and the set A is guaranteed by Claim 5.2.19. We shall only prove the case $A = V_1$, since the arguments for other cases are similar.

Let

$$B_v = \left\{ww' \in L(v) \colon ww' \notin \widetilde{G}_1\right\}, \quad \text{and} \quad M_v = \left\{ww' \in \widetilde{G}_1 \colon ww' \notin L(v)\right\}.$$

Members in B_v are called bad edges of $L(v)$ and members in M_v are called missing edges of $L(v)$.

Claim 5.2.20. *We have $|B_v| < 300\epsilon^{1/12}n^2$.*

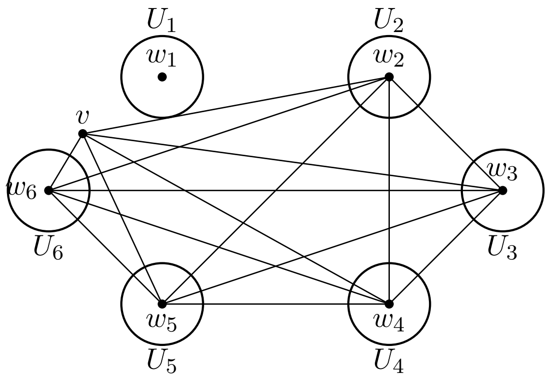


Figure 21. The 3-graph $F = \mathcal{H}_i[\{w_1, \ldots, w_6\}] \cup \{e_j : j \in \{4, 5, 6\}\} \cup \{vw_2w_3\}$ is a member in M_3 with core $\{v, w_1, \ldots, w_5\}$.

Proof. Suppose to the contrary that $|B_v| \geq 300\epsilon^{1/12}\tilde{n}^2$. Notice that every edge in B_v must have one vertex in V_2 and the other vertex in $V_3 \cup V_6$. By symmetry and the Pigeonhole principle, we may assume that at least $|B_v|/2$ edges in B_v have one vertex in V_2 and the other vertex in V_3. Then Claim 5.2.14 and an easy averaging argument show that there exists a vertex $w_2 \in V_2$ such that

$$|N_{B_v}(w_2) \cap V_3| \geq \frac{|B_v|/2}{|V_2|} > \frac{300\epsilon^{1/12}n^2/2}{n/5} > 600\epsilon^{1/12}n.$$

Let $V_3' = N_{B_v}(w_2) \cap V_3$, and $V_j' = N(v) \cap V_j$ for $j \in \{4, 5, 6\}$. Since $|V_3'| \geq 600\epsilon^{1/12}n$ and $|V_j'| \geq n/13$ for $j \in \{4, 5, 6\}$, applying Lemma 5.3.23 with $T = \{w_2\}$, $S = \{1, 3, 4, 5, 6\}$, and $\eta = 20\epsilon^{1/2}$ we obtain $w_1 \in V_1$ and $w_j \in V_j'$ for $j \in \{3, 4, 5, 6\}$ (see Figure 21) such that the induced subgraph of $\mathcal{H}$ on $\{w_1, \ldots, w_6\}$ is a copy of $\mathcal{G}_6^2$. For $j \in \{4, 5, 6\}$ let $e_j \in \mathcal{H}$ be an edge containing v and w_j. Let $F = \mathcal{H}[\{w_1, \ldots, w_6\}] \cup \{e_j : j \in \{4, 5, 6\}\} \cup \{vw_2w_3\}$. It is easy to see that F is a member in $\mathcal{K}_6^3$ with core $\{v, w_2, \dots, w_6\}$ (see Figure 21). So, by assumption, either $F \subset \mathcal{G}_m^1$ or $F \subset \mathcal{G}_m^2$ for any integer m. It is easy to see that the former case cannot hold since the induced subgraph of F on the set $\{w_1, \dots, w_6\}$ is a copy of $\mathcal{G}_6^2$ and $\mathcal{G}_6^2 \not\subset \mathcal{G}_m^1$ for any integer m. So, $F \subset \mathcal{G}_m^2$ for some integer m. In other words, there exists a map $\psi\colon V(F) \to V(\mathcal{G}_6^2)$ such that $\psi(e) \in \mathcal{G}_6^2$ for all $e \in F$. Notice that both $\{w_1, \dots, w_6\}$ and $\{v, w_2, \dots, w_6\}$ are 2-covered in F, so the restrictions of ψ on $\{w_1, \dots, w_6\}$ and $\{v, w_2, \dots, w_6\}$ are both injective (similar to the proof of Lemma 5.2.2), and moreover, $\psi(v) = \psi(w_1)$. Let $w = \psi(v) = \psi(w_1)$. Notice that the induced subgraph of $L_F(w_1)$ on $\{w_2, \dots, w_3\}$ has size 8 and $w_2w_3 \in L_F(v) \setminus L_F(w_1)$. Since ψ preserves edges, the degree of w in $\mathcal{G}_6^2$ should be at least $8+1=9$. However, this contradicts the fact that the maximum degree of $\mathcal{G}_6^2$ is 8. ∎

A consequence of Claim 5.2.20 is that the size of M_v satisfies

$$\begin{aligned}|M_v| = |\widetilde{G}_1 \setminus L(v)| &= |\widetilde{G}_1| - |\widetilde{G}_1 \cap L(v)| \\ &= |\widetilde{G}_1| - (|L(v)| - |B_v|) \\ &< 8\left(1/6 + 5\epsilon^{1/2}\right)^2 n^2 - \left((2/9 - \epsilon)\, n^2 - |B_v|\right) < 400\epsilon^{1/12} n^2.\end{aligned}$$

Claim 5.2.21. *We have $B_v = \emptyset$. In other words, $L(v) \subset \widetilde{G}_1$.*

Proof. Suppose to the contrary that there exists an edge $u_2u_3 \in B_v$ and by symmetry we may assume that $u_2 \in V_2$ and $u_3 \in V_3$. For $j \in \{4,5,6\}$ let $V_j' = V_j \cap N(v) \cap N(u_1) \cap N(u_2)$ and notice that due to $|M_v| \le 400\epsilon^{1/12} n^2$ and Claim 5.2.14 we have $|V_j'| \ge |V_j|/2 > n/20$. Applying

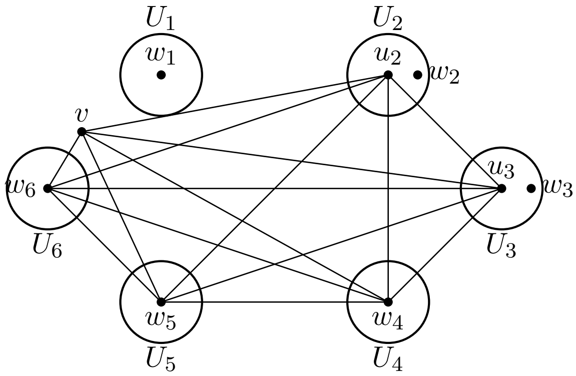


Figure 22. The 3-graph $F = \mathcal{H}_i[\{v, u_2, u_3, w_1, \ldots, w_6\}] \cup \{vu_2u_3\} \cup \{e_{u_3w_4}\}$ is a member in M_3 with core $\{v, u_2, u_3, w_4, w_5, w_6\}$.

Lemma 5.3.23 with $T = \{u_2, u_3\}$, $S = [6]$, and $\eta = 400\epsilon^{1/36}$ we obtain $w_j \in V_j'$ for $j \in [6]$ (see Figure 22) such that

(a) $\mathcal{H}[\{w_1, \ldots, w_6\}] \cong \mathcal{G}_6^2$,

(b) $L(v)[\{w_2, \ldots, w_6\}] = L_{\widehat{\mathcal{G}}}(w_1)[\{w_2, \ldots, w_6\}]$,

(c) $L(u_2)[\{w_1, w_3, \ldots, w_6\}] = L_{\widehat{\mathcal{G}}}(u_2)[\{w_1, w_3, \ldots, w_6\}]$, and

(d) $L(u_3)[\{w_1, w_2, w_4, w_5, w_6\}] = L_{\widehat{\mathcal{G}}}(u_3)[\{w_1, w_2, w_4, w_5, w_6\}]$.

Let $e_{u_3w_4} \in \mathcal{H}$ be an edge containing u_3 and w_4. Let $F = \mathcal{H}[\{v, u_2, u_3, w_1, \ldots, w_6\}] \cup \{vu_2u_3\} \cup \{e_{u_3w_4}\}$. Then it is easy to see that F is a member in $\mathcal{K}_6^3$ with core $\{v, u_2, u_3, w_4, w_5, w_6\}$ (see Figure 22). Similar to the proof of Claim 5.2.20, $F \subset \mathcal{G}_m^2$ for some integer m. In other words, there exists a map $\psi\colon V(F) \to V(\mathcal{G}_6^2)$ such that $\psi(e) \in \mathcal{G}_6^2$ for all $e \in F$. Notice that both $\{w_1, \ldots, w_6\}$ and $\{v, u_2, u_3, w_4, w_5, w_6\}$ are 2-covered in F, so the restrictions of ψ on sets $\{w_1, \ldots, w_6\}$ and $\{v, u_2, u_3, w_4, w_5, w_6\}$ are both injective (similar to the proof of Lemma 5.2.2), and moreover, $\psi(v) = \psi(w_1)$ (due to (b), v is adjacent to all vertices in $\{w_2, \ldots, w_6\}$, so $\psi(v)$ is distinct from $\{\psi(w_2), \ldots, \psi(w_6)\}$), $\psi(u_2) = \psi(w_2)$ (due to (c) and a similar reason), and $\psi(u_3) = \psi(w_3)$ (due to (d) and a similar reason). Let $w = \psi(v) = \psi(w_1)$. Notice that the induced subgraph of $L_F(w_1)$ on $\{w_2, \ldots, w_6\}$ has size 8 and $u_2u_3 \in L_F(v) \setminus L_F(w_1)$. Since ψ preserves edges, the degree of w in $\mathcal{G}_6^2$ should be at least $8 + 1 = 9$. However, this contradicts the fact that the maximum degree of $\mathcal{G}_6^2$ is 8. ∎

Define

$$\widehat{V}_j = \begin{cases} V_1 \cup \{v\}, & \text{if } j = 1, \\ V_j, & \text{otherwise.} \end{cases}$$

By Claim 5.2.21, $L(v) \subset \widetilde{G}_1$. Therefore, $\mathcal{H}$ is $\mathcal{G}_6^2$-colorable with set of parts $\{\widehat{V}_1, \ldots, \widehat{V}_6\}$. This completes the proof of Lemma 5.2.12. ∎

5.2.4 Proof of Theorem 5.1.8

Theorem 5.1.7 gives the following lemma.

Lemma 5.2.22. *Let $\epsilon > 0$ be sufficiently small and n (related to ϵ) be sufficiently large. Suppose that $\mathcal{H}$ is an $\mathcal{M}$-free 3-graph with n vertices and at least $2n^3/27 - \epsilon n^3$ edges. Then,*

$$\text{either } \left| |\partial\mathcal{H}| - \frac{5}{12}n^2 \right| < 100\epsilon^{1/4}n^2 \text{ or } \left| |\partial\mathcal{H}| - \frac{4}{9}n^2 \right| < 100\epsilon^{1/4}n^2.$$

Proof. Let $\epsilon > 0$ be sufficiently small and n (related to ϵ) be sufficiently large. Let $\mathcal{H}$ be an $\mathcal{M}$-free 3-graph with n vertices and at least $2n^3/27 - \epsilon n^3$ edges. By Theorem 5.1.7, there exists $W \subset V(\mathcal{H})$ with $|W| > n - 3\epsilon^{1/2}n$ such that $\delta(\mathcal{H}[W]) \geq 2n^2/9 - 20\epsilon^{1/2}n^2$ and $\mathcal{H}[W]$ is either semibipartite or $\mathcal{G}_6^2$-colorable. Let $Z = V(\mathcal{H}) \setminus W$, $\tilde{n} = |W|$, and $\widetilde{\mathcal{H}} = \mathcal{H}[W]$. Then,

$$|\widetilde{\mathcal{H}}| = \frac{1}{3}\sum_{w\in W} d_{\widetilde{\mathcal{H}}}(w) > \frac{1}{3}\left(n - 3\epsilon^{1/2}n\right)\left(\frac{2}{9}n^2 - 20\epsilon^{1/2}n^2\right) > \frac{2}{27}n^3 - 20\epsilon^{1/2}n^3. \tag{5.5}$$

Suppose that $\mathcal{H}[W]$ is semibipartite and let L and R denote the two parts of $\mathcal{H}[W]$ such that every $E \in \mathcal{H}[W]$ satisfies $|A \cap E| = 1$ and $|B \cap E| = 2$. Note from Claim 5.2.9 that

$$|L| = \frac{|W|}{3} \pm 4\epsilon^{1/4}|W| = \frac{n}{3} \pm 8\epsilon^{1/4}n \tag{5.6}$$

and

$$|L| = \frac{2|W|}{3} \pm 4\epsilon^{1/4}|W| = \frac{2n}{3} \pm 8\epsilon^{1/4}n. \tag{5.7}$$

First we prove the lower bound for $|\partial\mathcal{H}|$. Let

$$(\partial\widetilde{\mathcal{H}})[R] = \left\{uv \in \partial\widetilde{\mathcal{H}} : \{u,v\} \subset R\right\},$$

and

$$(\partial\widetilde{\mathcal{H}})[L,R] = \left\{uv \in \partial\widetilde{\mathcal{H}} : u \in L, v \in R\right\}.$$

Since $\widetilde{\mathcal{H}}$ is semibipartite,

$$|L||(\partial\widetilde{\mathcal{H}})[R]| \geq \sum_{v\in L} d_{\widetilde{H}}(v) = |\widetilde{H}|$$

and

$$|R||(\partial\widetilde{\mathcal{H}})[L,R]| \geq \sum_{u\in R} d_{\widetilde{H}}(u) = 2|\widetilde{H}|.$$

Together with Equation 5.5, Equation 5.6, and Equation 5.7, we obtain

$$\begin{aligned}|\partial\widetilde{\mathcal{H}}| = |(\partial\widetilde{\mathcal{H}})[R]| + |(\partial\widetilde{\mathcal{H}})[L,R]| &\geq \frac{|\widetilde{\mathcal{H}}|}{|L|} + \frac{2|\widetilde{\mathcal{H}}|}{|R|} \\ &> \frac{2n^3/27 - 20\epsilon^{1/2}n^3}{(1/3+8\epsilon^{1/4})n} + \frac{2(2n^3/27 - 20\epsilon^{1/2}n^3)}{(2/3+8\epsilon^{1/4})n} \\ &> \frac{4}{9}n^2 - 100\epsilon^{1/2}n^2.\end{aligned}$$

Therefore, $|\partial\mathcal{H}| \geq |\partial\widetilde{\mathcal{H}}| > 4n^2/9 - 100\epsilon^{1/2}n^2$.

Next, we prove the upper bound for $|\partial\mathcal{H}|$. Let $v \in Z$ and suppose that there exists an edge $w_1w_2 \in L_{\mathcal{H}}(v)[L]$. Since $L_{\mathcal{H}[W]}(w_1)$ and $L_{\mathcal{H}[W]}(w_2)$ are graphs on R and $|L_{\mathcal{H}[W]}(w_1)|, |L_{\mathcal{H}[W]}(w_2)| \geq 2n^2/9 - 20\epsilon^{1/2}n^2$, the edge density of $L_{\mathcal{H}[W]}(w_1) \cap L_{\mathcal{H}[W]}(w_2)$ is close to 1. So, by Turán's theorem, there exists a copy of K_5 in $L_{\mathcal{H}[W]}(w_1) \cap L_{\mathcal{H}[W]}(w_2)$. We may assume that the vertex set of this K_5 is $\{w_3, w_4, w_5, w_6, w_7\}$ (see Figure 23). Let $F = \mathcal{H}[\{w_1, \ldots, w_7\}] \cup \{uw_1w_2\}$. Then

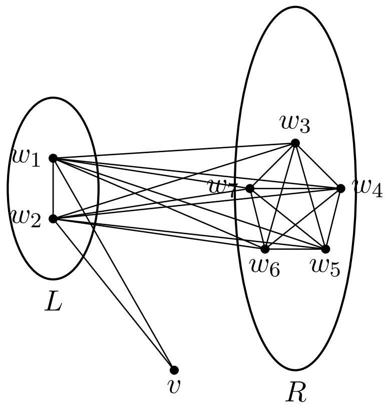


Figure 23. The 3-graph $F = \mathcal{H}[\{w_1, \ldots, w_7\}] \cup \{uw_1w_2\}$ is a member in M_2 with core $\{w_1, \ldots, w_7\}$. In particular, $\tau\{w_1w_3w_4, w_2w_5w_6\} > 1$.

it is easy to see that F is a member in M_2 with core $\{w_1, \ldots, w_7\}$, a contradiction. Therefore, $L_{\mathcal{H}}(v)[L] = \emptyset$ for all $v \in Z$, and it follows that

$$\begin{aligned}|\partial\mathcal{H}| &\leq \binom{|W|}{2} - \binom{|L|}{2} + |Z||W| + \binom{|Z|}{2} \\ &\leq \binom{n}{2} - \binom{|L|}{2} < \frac{1}{2}n^2 - \frac{1}{2}\left(\frac{1}{3} - 8\epsilon^{1/4}\right)n^2 < \frac{4}{9}n^2 + 100\epsilon^{1/4}n^2.\end{aligned}$$

Therefore, if $\widetilde{\mathcal{H}}$ is semibipartite, then

$$\frac{4}{9}n^2 - 100\epsilon^{1/2}n^2 < |\partial\mathcal{H}| < \frac{4}{9}n^2 + 100\epsilon^{1/4}n^2.$$

Suppose that $\widetilde{\mathcal{H}}$ is $\mathcal{G}_6^2$-colorable and let $\mathcal{P} := \{A_1, A_2, A_3, A_4, B_1, B_2\}$ be the set of six parts of $\widetilde{\mathcal{H}}$ such that there is no edge between $A_1A_2B_1$, $A_1A_2B_2$, $A_3A_4B_1$, and $A_3A_4B_2$. Notice from Claim 5.2.14 that

$$|S| = \frac{|W|}{6} \pm 10\epsilon^{1/2}|W| = \frac{n}{6} \pm 10\epsilon^{1/4}n, \text{ for all } S \in \mathcal{P}. \tag{5.8}$$

First, we prove the lower bound for $|\partial\mathcal{H}|$. Since $\widetilde{\mathcal{H}}$ is $\mathcal{G}_6^2$-colorable, $\partial\widetilde{\mathcal{H}}$ is a 6-partite graph the set of six parts $\mathcal{P}$. Let $\mathcal{G}$ denote the blow up of $\mathcal{G}_6^2$ with the set of six parts $\mathcal{P}$ such that there is no edge between $A_1A_2B_1$, $A_1A_2B_2$, $A_3A_4B_1$, and $A_3A_4B_2$. Notice that for every $e \in \partial\mathcal{G} \setminus \partial\widetilde{\mathcal{H}}$ there are at least $2(|W|/6 - 10\epsilon^{1/2}|W|)$ sets $E \in \mathcal{G} \setminus \widetilde{\mathcal{H}}$ such that $e \in E$. Therefore,

$$|\partial\mathcal{G} \setminus \partial\widetilde{\mathcal{H}}| \leq \frac{3|\mathcal{G} \setminus \widetilde{\mathcal{H}}|}{2(|W|/6 - 10\epsilon^{1/2}|W|)} \overset{\text{Equation 5.5, Equation 5.8}}{<} \frac{3 \times 20\epsilon^{1/2}n^2}{2(n/6 - 10\epsilon^{1/4}n)} < 400\epsilon^{1/2}n^2,$$

and it follows that

$$|\partial\widetilde{\mathcal{H}}| > |\partial\mathcal{G}| - 400\epsilon^{1/2}n^2 \overset{\text{Equation 5.8}}{>} \binom{6}{2}\left(\frac{n}{6} - 10\epsilon^{1/4}n\right)^2 - 400\epsilon^{1/2}n^2 > \frac{5}{12}n^2 - 100\epsilon^{1/4}n^2$$

Therefore, $|\partial\mathcal{H}| \geq |\partial\widetilde{\mathcal{H}}| > 5n^2/12 - 100\epsilon^{1/4}n^2$.

Next, we prove the upper bound for $|\partial\mathcal{H}|$. Let $v \in Z$ and suppose that $L_{\mathcal{H}}(v)[S] \neq \emptyset$ for some $S \in \mathcal{P}$. Then Claim 5.2.17[1] implies that $\mathcal{H}$ contains a copy of a 3-graph in M_2, a contradiction. Therefore, $L_{\mathcal{H}}(v)[S] = \emptyset$ for all $S \in \mathcal{P}$, and it follows that

$$|\partial\mathcal{H}| \leq \frac{5}{12}|W|^2 + |Z||W| + \binom{|Z|}{2} < \frac{5}{12}n^2 + 100\epsilon^{1/4}n^2.$$

Therefore, if $\widetilde{\mathcal{H}}$ is $\mathcal{G}^2$-colorable, then

$$\frac{5}{12}n^2 - 100\epsilon^{1/4}n^2 < |\partial\mathcal{H}| < \frac{5}{12}n^2 + 100\epsilon^{1/4}n^2.$$

∎

Now we are ready to prove Theorem 5.1.8.

Proof of Theorem 5.1.8. Let

$$\mathcal{S}_n = \{A \subset [n] : 1 \in A\}.$$

Since $\mathcal{S}_n$ is $\mathcal{M}$-free and $|\partial\mathcal{S}_n| = \binom{n}{2}$, it follows that $\text{proj}\Omega(\mathcal{M}) = [0,1]$. On the other hand, it follows from Theorem 5.1.4 that $g(\mathcal{M}, x) \leq 4/9$ for all $x \in [0,1]$ and $g(\mathcal{M}, 5/6) = g(\mathcal{M}, 8/9) = 4/9$.

---

[1] Even though Claim 5.2.17 was proved only for vertices in W, in fact, its proof does not require v to have a large degree. So it also holds for vertices in Z.

Now suppose that $(\mathcal{H}_k)_{k=1}^{\infty}$ is a sequence of $\mathcal{M}$-free 3-graphs with $\lim_{k\to\infty} v(\mathcal{H}_k) = \infty$, $\lim_{k\to\infty} d(\partial\mathcal{H}_k) = x_0$, and $\lim_{k\to\infty} d(\mathcal{H}_k) = 4/9$. For any sufficiently small $\epsilon > 0$ and sufficiently n_0, there exists k_0 such that $v(\mathcal{H}_k) \geq n_0$ and $|\mathcal{H}_k| > 2(v(\mathcal{H}_k))^3/27 - \epsilon(v(\mathcal{H}_k))^3$ for all $k \geq k_0$. Therefore, by Lemma 5.2.22, for every $k \geq k_0$ either

$$\frac{8}{9} - 200\epsilon^{1/4} < \frac{|\partial\mathcal{H}_k|}{\binom{v(\mathcal{H}_k)}{2}} < \frac{8}{9} + 200\epsilon^{1/4}$$

or

$$\frac{5}{6} - 200\epsilon^{1/4} < \frac{|\partial\mathcal{H}_k|}{\binom{v(\mathcal{H}_k)}{2}} < \frac{5}{6} + 200\epsilon^{1/4}.$$

Letting $\epsilon \to 0$ we obtain either $x_0 = 8/9$ or $x_0 = 5/6$, and this completes the proof. ∎

5.3 Proof for t-stable families

In this section we prove Theorems 5.1.9 and 5.1.10.

5.3.1 Preliminaries

In this section we present some definitions related to the Lagrangian of a hypergraph, introduced by Frankl and Rödl in [105], and prove a result (Proposition 5.4.2 below) about certain almost complete triple systems.

For a pair of vertices $u, v \in V(\mathcal{G})$ the neighborhood of $\{u, v\}$ is

$$N_{\mathcal{G}}(u, v) = \{w \in V(\mathcal{G}) \setminus \{u, v\} \colon \exists A \in \mathcal{G} \text{ such that } \{u, v, w\} \subseteq A\},$$

and $d_{\mathcal{G}}(u, v) = |N_{\mathcal{H}}(u, v)|$ is called the codegree of $\{u, v\}$. Denote by $\delta_2(\mathcal{G}), \Delta_2(\mathcal{G})$ the minimum and maximum codegree of $\mathcal{G}$, respectively.

For a hypergraph $\mathcal{G}$ the maximum number of edges in a blow-up of $\mathcal{G}$ is related to $\lambda(\mathcal{G})$ (e.g. see Frankl and Füredi [102] or Section 3 in Keevash's survey [135]).

Lemma 5.3.1 ([102; 135])**.** *Let $r \geq 2$ and let $\mathcal{G}$, $\mathcal{H}$ be two r-graphs. If $\mathcal{H}$ is a blow up of $\mathcal{G}$, then $|\mathcal{H}| \leq \lambda(\mathcal{G}) v(\mathcal{H})^r$.*

Given a 3-graph $\mathcal{G}$, by plugging $(1/n, \dots, 1/n)$ into $L_{\mathcal{G}}$ one immediately obtains the lower bound $\lambda(L_{\mathcal{G}}) \geq |\mathcal{G}|/n^3$. It is well known that for cliques $\mathcal{H} = K_n^3$ this holds with equality and, moreover, that $(1/n, \dots, 1/n)$ is the only point in the simplex Δ_{n-1}, where $L_{\mathcal{H}}$ attains this maximum value.

The main result of this section, Proposition 5.4.2 below, exhibits a class of almost complete 3-graphs having the same properties. This will allow us later to construct for every given positive integer t a family $\{\mathcal{G}_1, \dots, \mathcal{G}_t\}$ of 3-graphs and a rational number λ_t close to $1/6$ such that $\lambda(\mathcal{G}_i) = |\mathcal{G}_i|/v(\mathcal{G}_i)^3 = \lambda_t$ holds for all $i \in [t]$. The extremal configurations for our hypergraph Turán problem are then going to be balanced blow-ups of $\mathcal{G}_1, \dots, \mathcal{G}_t$. As we can accomplish $v(\mathcal{G}_1) < \dots < v(\mathcal{G}_t)$, this is relevant to Theorem 5.1.10.

Let us observe that every hypergraph $\mathcal{G}$ satisfying $\lambda(\mathcal{G}) = |\mathcal{G}|/v(\mathcal{G})^3$ needs to be regular in the sense that all vertices have the same degree. In the converse direction, regular hypergraphs can still have much larger Lagrangians than $|\mathcal{G}|/v(\mathcal{G})^3$. For instance, the Lagrangian of the Fano plane is $1/27$ but not $1/49$. To avoid such situations we utilize a design theoretic construction.

For the purposes of this article, by an (n, k)-design we shall mean a k-graph $\mathcal{D}$ on n vertices such that every pair of vertices is covered by a unique edge. With every such design $\mathcal{D}$ we associate the 3-graph

$$H(\mathcal{D}) = \bigcup_{E \in \mathcal{D}} \binom{E}{3}.$$

on $V(\mathcal{D})$. Note that

$$|H(\mathcal{D})| = \binom{k}{3} \frac{\binom{n}{2}}{\binom{k}{2}} = \frac{k-2}{6} n(n-1).$$

It will turn out that for $n \geq 18k$ every 3-graph of the form $\mathcal{G} = K_n^3 \setminus H(\mathcal{D})$, where $\mathcal{D}$ is an (n, k)-design on $[n]$, has the property $\lambda(\mathcal{G}) = |\mathcal{G}|/v(\mathcal{G})^3$. In order to increase our control over the resulting value of $\lambda(\mathcal{G})$ Proposition 5.4.2 allows the extra flexibility to subtract a very sparse regular 3-graph from $\mathcal{G}$. Moreover, for reasons related to stability we state slightly more than just the actual value of the Lagrangian.

Proposition 5.3.2. *Suppose that* $n \geq 18k + 3^7 s^3$*,* $\mathcal{D}$ *is an* (n, k)*-design on* $[n]$*, and* $\mathcal{S}$ *is an* s*-regular* 3*-graph on* $[n]$*. If* $\mathcal{S} \cap H(\mathcal{D}) = \emptyset$ *and* $\mathcal{G} = K_n^3 \setminus (H(\mathcal{D}) \cup \mathcal{S})$*, then*

$$L_{\mathcal{G}}(x_1, \ldots, x_n) + \frac{1}{9}\sum_{i=1}^{n}\left(x_i - \frac{1}{n}\right)^2 \leq \frac{|\mathcal{G}|}{n^3} = \frac{1}{6}\left(1 - \frac{k+1}{n} + \frac{k-2s}{n^2}\right) \tag{5.9}$$

holds for all $(x_1, \ldots, x_n) \in \Delta_{n-1}$ *and, consequently,*

$$\lambda(\mathcal{G}) = \frac{1}{6}\left(1 - \frac{k+1}{n} + \frac{k-2s}{n^2}\right). \tag{5.10}$$

We start with a simple observation that will come in handily later.

Fact 5.3.3. *Let* $\mathcal{G}$ *be a* 3*-graph with vertex set* $[n]$ *and let* $\alpha \geq 0$ *be a real number. If the real numbers* $\alpha_1, \ldots, \alpha_n \in [-1, \alpha]$ *sum up to zero, then*

$$L_{\mathcal{G}}(\alpha_1, \ldots, \alpha_n) \leq (\alpha n)^3.$$

Proof. Define $P = \{i \in [n] \colon \alpha_i > 0\}$ to be the set of vertices of $\mathcal{G}$ with positive weight. Let us decompose $L_{\mathcal{G}}(\alpha_1, \ldots, \alpha_n) = S_0 + S_1 + S_2 + S_3$ such that for $m \in \{0, 1, 2, 3\}$ the sum S_m consists of all terms $\alpha_i\alpha_j\alpha_k$ contributing to $L_{\mathcal{G}}$ and satisfying $|P \cap \{i, j, k\}| = m$.

As the sums S_0 and S_2 possess no positive terms, we have $S_0, S_2 \le 0$. Moreover, S_3 has no more than $\binom{|P|}{3} \le n^3/6$ summands each of which amounts to at most α^3, wherefore S_3 is at most $(\alpha n)^3/6$. Thus to conclude the argument it is more than enough to show $S_1 \le (\alpha n)^3/2$.

Writing $W = \sum_{i\in P} \alpha_i$ we have $\sum_{i\in[n]\setminus P} \alpha_i = -W$ and

$$S_1 \le \sum_{i\in P} \alpha_i \cdot \sum_{jk\in\binom{[n]\setminus P}{2}} \alpha_j\alpha_k \le W \cdot (W^2/2) = W^3/2,$$

which by $|W| \le \alpha|P| \le \alpha n$ completes the proof. ∎

Proof of Proposition 5.4.2. Since the left side of Equation 5.28 is continuous in $(x_1, \dots, x_n)$ and Δ_{n-1} is compact, there exists a point $\xi = (\xi_1, \dots, \xi_n) \in \Delta_{n-1}$ such that

$$\omega = L_{\mathcal{G}}(\xi_1, \dots, \xi_n) + \frac{1}{9}\sum_{i=1}^{n}\left(\xi_i - \frac{1}{n}\right)^2 - \frac{1}{6}\left(1 - \frac{k+1}{n} + \frac{k-2s}{n^2}\right) \tag{5.11}$$

is maximum. Assume for the sake of contradiction that $\omega > 0$.

Claim 5.3.4. *There exists an index $i(\star) \in [n]$ such that $\xi_{i(\star)} > \frac{1}{n} + \frac{9s}{n^2}$.*

Proof. Define $\alpha_1, \dots, \alpha_n \in [-1, n-1]$ by $\xi_i = (1+\alpha_i)/n$ for every $i \in [n]$ and observe that

$$\begin{aligned}\omega n^3 &= L_{\mathcal{G}}(1+\alpha_1, \dots, 1+\alpha_n) + \frac{n}{9}\sum_{i=1}^{n}\alpha_i^2 - |\mathcal{G}| \\ &= \sum_{i=1}^{n} d_{\mathcal{G}}(i)\alpha_i + \sum_{1\le i<j\le n} d_{\mathcal{G}}(i,j)\alpha_i\alpha_j + L_{\mathcal{G}}(\alpha_1, \dots, \alpha_n) + \frac{n}{9}\sum_{i=1}^{n}\alpha_i^2\,.\end{aligned}$$

Since all vertices of $\mathcal{G}$ have the same degree and $\sum_{i=1}^{n} \alpha_i = n\big(\sum_{i=1}^{n} \xi_i - 1\big) = 0$, the first sum on the right side vanishes. Moreover, all pairs of vertices have codegree $n-k$ in $K_n^3 \setminus H(\mathcal{D})$ and thus we obtain

$$\omega n^3 = \left(\frac{n}{9} - \frac{n-k}{2}\right) \sum_{i=1}^{n} \alpha_i^2 - \sum_{1 \le i < j \le n} d_{\mathcal{S}}(i,j)\alpha_i\alpha_j + L_{\mathcal{G}}(\alpha_1, \dots, \alpha_n)\,. \tag{5.12}$$

First case: We have $\xi_1, \dots, \xi_n > 0$.

Collecting the quadratic and cubic terms in Equation 5.12 separately we put

$$Q = \left(\frac{n}{9} - \frac{n-k}{2}\right) \sum_{i=1}^{n} \alpha_i^2 - \sum_{1 \le i < j \le n} d_{\mathcal{S}}(i,j)\alpha_i\alpha_j \quad \text{and} \quad K = L_{\mathcal{G}}(\alpha_1, \dots, \alpha_n),$$

so that

$$\omega n^3 = Q + K.$$

Now for every real number C sufficiently close to 1 the point $(\xi_1', \dots, \xi_n')$ defined by $\xi_i' = (1 + C\alpha_i)/n$ belongs to Δ_{n-1} and the maximal choice of ω reveals

$$L_{\mathcal{G}}(\xi_1', \dots, \xi_n') + \frac{1}{9} \sum_{i=1}^{n} \left(\xi_i' - \frac{1}{n}\right)^2 - \frac{1}{6}\left(1 - \frac{k+1}{n} + \frac{k-2s}{n^2}\right) \le \omega.$$

Multiplying by n^3 and repeating the above calculation we obtain $QC^2 + KC^3 \leq Q + K$ and thus

$$0 \leq (1-C)[(1+C)Q + (1+C+C^2)K] \tag{5.13}$$

whenever $|C-1|$ is sufficiently small. Letting C tend to 1 from above and below we obtain $2Q + 3K = 0$. Substituting this back into Equation 5.13 we learn

$$0 \leq (1-C)[(C-1)Q + (C^2+C-2)K] = -(1-C)^2[Q + (C+2)K].$$

Thus $Q + 3K \leq (1-C)K$ holds whenever $|C-1|$ is sufficiently small, which is only possible if $Q + 3K \leq 0$. Together with $Q + K = \omega n^3 > 0$ this yields $K < 0$ and $\omega n^3 < Q - K = (-1)^2 Q + (-1)^3 Q$. So the maximality of ω tells us that for $C = -1$ we have $(\xi_1', \ldots, \xi_n') \notin \Delta_{n-1}$. In other words, there is some $i(\star) \in [n]$ such that

$$\xi_{i(\star)} \geq \frac{2}{n} > \frac{1}{n} + \frac{9s}{n^3},$$

as desired.

Second case: There exists some $j(\star) \in [n]$ satisfying $\xi_{j(\star)} = 0$.

Now $\alpha_{j(\star)} = -1$ and, consequently,

$$\sum_{i=1}^{n} \alpha_i^2 \geq 1. \tag{5.14}$$

Next we observe that the hypothesis that $\mathcal{S}$ be s-regular yields

$$-\sum_{1\le i<j\le n} d_{\mathcal{S}}(i,j)\alpha_i\alpha_j \le \sum_{1\le i<j\le n} d_{\mathcal{S}}(i,j)\frac{\alpha_i^2+\alpha_j^2}{2} = s\sum_{i=1}^{n}\alpha_i^2.$$

Combined with Equation 5.12 and the positivity of ω this shows

$$\left(\frac{n-k}{2}-\frac{n}{9}-s\right)\sum_{i=1}^{n}\alpha_i^2 < L_{\mathcal{G}}(\alpha_1,\ldots,\alpha_n). \tag{5.15}$$

Due to $n \ge 18k + 3^7 s^3$ we have

$$\frac{n-k}{2}-\frac{n}{9}-s > \left(\frac{1}{2}-\frac{1}{36}-\frac{1}{9}-\frac{1}{36}\right)n = \frac{n}{3} > (9s)^3$$

and together with Equation 5.14, Equation 5.15 this establishes

$$(9s)^3 < L_{\mathcal{G}}(\alpha_1,\ldots,\alpha_n).$$

In view of Fact 5.3.3 we deduce that $\alpha_{i(\star)} > 9s/n$ holds for some $i(\star)\in[n]$ and now

$$\xi_{i(\star)} = \frac{1+\alpha_{i(\star)}}{n} > \frac{1}{n}+\frac{9s}{n^2}$$

follows. Thereby Claim 5.3.4 is proved. ∎

Now for every $i \in [n]$ we set

$$D_i = \frac{\partial L_{\mathcal{G}}(x_1, \ldots, x_n)}{\partial x_i}\bigg|_{(\xi_1,\ldots,\xi_n)} = \sum_{jk \in L_i} \xi_j \xi_k,$$

where L_i denotes the link graph of i in $\mathcal{G}$. Owing to the maximality of ω in Equation 5.11 the Lagrange multiplier method leads to the existence of a real number M such that

$$D_i + \frac{2}{9}\left(\xi_i - \frac{1}{n}\right) = M$$

holds for every vertex $i \in [n]$ with $\xi_i > 0$. Notice that

$$M = M\sum_{j=1}^{n} \xi_j = \sum_{j=1}^{n} \xi_j \left(D_j + \frac{2}{9}\left(\xi_j - \frac{1}{n}\right)\right) = 3L_{\mathcal{G}}(\xi_1, \ldots, \xi_n) + \frac{2}{9}\sum_{j=1}^{n}\left(\xi_j - \frac{1}{n}\right)^2$$
$$\overset{\text{Equation 5.11}}{>} \frac{1}{2}\left(1 - \frac{k+1}{n} + \frac{k-2s}{n^2}\right) - \frac{1}{9}\sum_{j=1}^{n}\left(\xi_j - \frac{1}{n}\right)^2.$$

Altogether, this proves

$$D_i + \frac{2}{9}\left(\xi_i - \frac{1}{n}\right) + \frac{1}{9}\sum_{j=1}^{n}\left(\xi_j - \frac{1}{n}\right)^2 > \frac{1}{2}\left(1 - \frac{k+1}{n} + \frac{k-2s}{n^2}\right)$$

for every vertex $i \in [n]$ satisfying $\xi_i > 0$.

By our design theoretic construction, the link in $K_n^3 \setminus H(\mathcal{D})$ of every vertex $i \in [n]$ is a q-partite Turán graph with vertex classes of size $k-1$, where $q = (n-1)/(k-1)$ is an integer. Consequently, there exist real numbers $\beta_1, \ldots, \beta_q$ such that $\xi_i + (\beta_1 + \cdots + \beta_q) = 1$ and

$$D_i \leq \sum_{1 \leq v < w \leq q} \beta_v \beta_w \leq \frac{q-1}{2q}(\beta_1 + \cdots + \beta_q)^2 = \frac{n-k}{2(n-1)}(1-\xi_i)^2.$$

Summarizing, we have

$$\frac{2}{9}\left(\xi_i - \frac{1}{n}\right) + \frac{1}{9}\sum_{j=1}^{n}\left(\xi_j - \frac{1}{n}\right)^2 > \frac{n-k}{2(n-1)}\left(\left(1-\frac{1}{n}\right)^2 - (1-\xi_i)^2\right) - \frac{s}{n^2} \tag{5.16}$$

for every vertex of positive weight. For the rest of the argument we fix a vertex $i(\star) \in [n]$ such that $\xi_{i(\star)}$ is maximal. Let us add the trivial estimate

$$\frac{1}{9}\sum_{j=1}^{n} \xi_j(\xi_{i(\star)} - \xi_j) \geq 0$$

to the case $i = i(\star)$ of Equation 5.16. Because of

$$\sum_{j=1}^{n}\left(\xi_j - \frac{1}{n}\right)^2 + \sum_{j=1}^{n}\xi_j(\xi_{i(\star)} - \xi_j) = \sum_{j=1}^{n}\xi_j\left(\xi_{i(\star)} - \frac{1}{n}\right) - \frac{1}{n}\sum_{j=1}^{n}\left(\xi_j - \frac{1}{n}\right) \tag{5.17}$$

$$= \xi_{i(\star)} - \frac{1}{n} \tag{5.18}$$

this yields

$$\frac{1}{3}\left(\xi_{i(\star)}-\frac{1}{n}\right) > \frac{n-k}{2(n-1)}\left(\xi_{i(\star)}-\frac{1}{n}\right)\left(2-\frac{1}{n}-\xi_{i(\star)}\right)-\frac{s}{n^2}$$
$$\geq \frac{n-k}{2n}\left(\xi_{i(\star)}-\frac{1}{n}\right)-\frac{s}{n^2} \geq \frac{4}{9}\left(\xi_{i(\star)}-\frac{1}{n}\right)-\frac{s}{n^2},$$

whence

$$\xi_{i(\star)} < \frac{1}{n}+\frac{9s}{n^2}.$$

Owing to the maximal choice of $\xi_{i(\star)}$ this contradicts Claim 5.3.4. ∎

5.3.2 Constructions and Turán numbers

Given a positive integer t we define in this section the triple systems $\mathcal{G}_1, \ldots, \mathcal{G}_t$ and the forbidden family $\mathcal{M}_t$ appearing in Theorem 5.1.9. For every $i \in [t]$ there will be three integers n_i, k_i, s_i such that $\mathcal{G}_i = K^3_{n_i} \setminus (H(\mathcal{D}_i) \cup \mathcal{S}_i)$ holds for some (n_i, k_i)-design $\mathcal{D}_i$ on $[n_i]$ and some s_i-regular triple system $\mathcal{S}_i$ on $[n_i]$ that is disjoint to $H(\mathcal{D}_i)$. As we shall have $n_i \gg k_i, s_i$, Proposition 5.4.2 will imply

$$\lambda(\mathcal{G}_i) = \frac{1}{6}\left(1-\frac{k_i+1}{n_i}+\frac{k_i-2s_i}{n_i^2}\right).$$

Part of our goal is that balanced blow-ups of $\mathcal{G}_1, \ldots, \mathcal{G}_t$ should be extremal $\mathcal{M}_t$-free triple systems and for this reason we need to ensure $\lambda(\mathcal{G}_i) = \cdots = \lambda(\mathcal{G}_t)$. We shall achieve this by letting $k_i = 2s_i$ for $i \in [t]$, and by guaranteeing

$$\frac{k_1+1}{n_1} = \cdots = \frac{k_t+1}{n_t}. \tag{5.19}$$

The details of this construction are given in Subsection 5.3.2.1 and the exact Turán numbers of our families $\mathcal{M}_t$ are determined in Subsection 5.3.2.2.

5.3.2.1 The extremal configurations and forbidden family

First, we need the following theorem about the existence of designs due to Wilson [244; 245; 247].

Theorem 5.3.5 (Wilson [244; 245; 247])**.** *For every integer $k \geq 2$ there exists a threshold $n_0(k)$ such that for every integer $n \geq n_0(k)$ satisfying the divisibility conditions $(k-1) \mid (n-1)$ and $(k-1)k \mid (n-1)n$ there exists an (n,k)-design.*

Our next lemma deals with the arithmetic properties the numbers $k_1, \ldots, k_t$ and $n_1, \ldots, n_t$ entering the construction of $\mathcal{G}_1, \ldots, \mathcal{G}_t$ need to satisfy. Apart from Equation 5.19 and the divisibility conditions in Theorem 5.3.5 we will require that $n_1, \ldots, n_t$ are divisible by 3 so that $(k_i/2)$-regular triple systems on n_i vertices exist. Thus the case $q = 3$ of the following lemma is exactly what we need.

Lemma 5.3.6. *Given positive integers t and q there exist t even integers $3 < k_1 < \cdots < k_t$ such that for every constant $C > 0$ there exist t integers $n_1 < \cdots < n_t$ with the following properties.*

(a) We have $q \mid n_i$, $(k_i - 1) \mid (n_i - 1)$, and $k_i(k_i-1) \mid n_i(n_i-1)$ for all $i \in [t]$.

(b) Moreover,

$$Q = \frac{n_1}{k_1 + 1} = \cdots = \frac{n_t}{k_t + 1}$$

is an integer with $Q \geq C$*.*

Proof. Starting with an arbitrary positive multiple s_1 of q we recursively define integers $1 \leq s_1 < \cdots < s_t$ by setting $s_{i+1} = \prod_{j \leq i} s_j(2s_j - 1) + 1$ for every $i \in [t-1]$. Now whenever $1 \leq i < j \leq t$ we have $s_j \equiv 1 \pmod{s_i(2s_i - 1)}$ and, consequently,

$$s_j(2s_j - 1) \equiv 1 \pmod{s_i(2s_i - 1)}.$$

In particular, the numbers

$$s_1(2s_1 - 1), \ldots, s_t(2s_t - 1)$$

are pairwise coprime and by the Chinese remainder theorem there exists an even integer $Q \geq C$ such that $Q/2 \equiv s_i^2 \pmod{s_i(2s_i - 1)}$ holds for all $i \in [t]$. Multiplying these congruences by 2 and setting $k_i = 2s_i$ we obtain

$$Q \equiv k_i^2/2 \pmod{k_i(k_i - 1)}. \tag{5.20}$$

Now it is plain that the numbers $n_i = Q(k_i + 1)$ satisfy (b). Moreover, the case $i = 1$ of Equation 5.20 yields $q \mid k_1 \mid Q$ and, therefore, $n_1, \ldots, n_t$ are divisible by q. Finally, multiplying Equation 5.20 by $k_i + 1$ we learn

$$n_i \equiv k_i(k_i + 1)(k_i/2) \equiv 2k_i(k_i/2) \equiv k_i^2 \equiv k_i \pmod{k_i(k_i - 1)},$$

for which reason $k_i \mid n_i$ and $(k_i - 1) \mid (n_i - 1)$. So altogether (a) holds as well. ∎

Given two r-graphs $\mathcal{H}_1$ and $\mathcal{H}_2$ with the same number of vertices a packing of $\mathcal{H}_1$ and $\mathcal{H}_2$ is a bijection $\phi\colon V(\mathcal{H}_1) \to V(\mathcal{H}_2)$ such that $\phi(E) \notin \mathcal{H}_2$ for all $E \in \mathcal{H}_1$. In order to proceed with our construction of the triple systems $\mathcal{G}_1, \ldots, \mathcal{G}_t$ we need to argue that, under natural assumptions, if $\mathcal{D}_i$ denotes an (n_i, k_i)-design, then there is an s_i-regular 3-graph $\mathcal{S}_i \subseteq K_n^3 \setminus H(\mathcal{D}_i)$, where $s_i = k_i/2$. Provided that $3 \mid n_i$ and $s_i \le \binom{n-1}{2}$ the existence of some s_i-regular 3-graph $\mathcal{S}_i \subseteq K_n^3$ is a well known fact that follows, e.g., from Baranyai's factorisation theorem [14]. For making $\mathcal{S}_i$ and $H(\mathcal{D}_i)$ disjoint we use a packing argument based on the following result of Lu and Székely.

Theorem 5.3.7 (Lu–Székely [179]). *Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be two r-graphs on n vertices. If*

$$\Delta(\mathcal{H}_1)|\mathcal{H}_2| + \Delta(\mathcal{H}_2)|\mathcal{H}_1| < \frac{1}{er}\binom{n}{r},$$

then there is a packing of $\mathcal{H}_1$ and $\mathcal{H}_2$.

In fact, we only require the following consequence.

Corollary 5.3.8. *Suppose* $3 \mid n$ *and that* $\mathcal{D}$ *is an* (n,k)*-design on* $[n]$*. If* $s < \frac{n-2}{6e(k-2)}$*, then there exists an* s*-regular* 3*-graph* $\mathcal{S}$ *on* $[n]$ *such that* $\mathcal{S} \cap H(\mathcal{D}) = \emptyset$*.*

Proof. By $3 \mid n$ and $s \le \binom{n-1}{2}$ there is an s-regular 3-graph $\mathcal{S}'$ on n vertices. Since

$$\begin{aligned}\Delta(\mathcal{S}')|H(\mathcal{D})| + \Delta(H(\mathcal{D}))|\mathcal{S}'| &= s\frac{k-2}{6}n(n-1) + \frac{k-2}{2}(n-1)\frac{sn}{3} \\ &= s\frac{k-2}{3}n(n-1) < \frac{n-2}{6e(k-2)}\frac{k-2}{3}n(n-1) \\ &= \frac{1}{3e}\binom{n}{3},\end{aligned}$$

Theorem 9.1.15 yields a packing $\phi\colon V(\mathcal{S}') \to [n]$ of $\mathcal{S}'$ and $H(\mathcal{D})$. It is clear that $\mathcal{S} = \phi(\mathcal{S}')$ satisfies the requirements of Corollary 5.3.8. ∎

Now we are ready to present the definition of $\mathcal{G}_1, \ldots, \mathcal{G}_t$.

Definition 5.3.9. *Given a positive integer* t *perform the following steps.*

- *Apply Lemma 5.3.6 with* $q = 3$*, thus getting some even integers* $3 < k_1 < \cdots < k_t$*.*
- *Take an integer* $C \ge \max\{n_0(k_1), \ldots, n_0(k_t), 2k_t^3, 3^8\}$*, where the thresholds* $n_0(k_i)$ *are given by Theorem 5.3.5.*
- *Now Lemma 5.3.6 applied to* C *and* $k_1, \ldots, k_t$ *yields integers* $C < n_1 < \cdots < n_k$ *such that, in particular,*

$$Q = \frac{n_1}{k_1+1} = \cdots = \frac{n_t}{k_t+1}$$

is an integer with $Q \ge C$*.*

Now, for every $i \in [t]$

- *let $\mathcal{D}_i$ be an (n_i, k_i)-design on $[n_i]$ (as obtained by Theorem 5.3.5)*
- *let $\mathcal{S}_i$ be a $(k_i/2)$-regular 3-graph on $[n_i]$ such that $\mathcal{S}_i \cap H(\mathcal{D}_i) = \emptyset$ (as obtained by Corollary 5.3.8).*
- *and, finally, define*

$$\mathcal{G}_i = K^3_{n_i} \setminus (H(\mathcal{D}_i) \cup \mathcal{S}_i) \,.$$

By Proposition 5.4.2 we have

$$\lambda(\mathcal{G}_i) = \frac{1}{6}\left(1 - \frac{k_i+1}{n_i} + \frac{k_i - 2k_i/2}{n_i^2}\right) = \frac{1}{6}\left(1 - \frac{1}{Q}\right).$$

for every $i \in [t]$, so some rational λ_t satisfies

$$\lambda_t = \lambda(\mathcal{G}_1) = \cdots = \lambda(\mathcal{G}_t) \in [5/32, 1/6). \tag{5.21}$$

In the remainder of this subsection we introduce the family $\mathcal{M}_t$. For an r-graph $\mathcal{H}$ and a set $S \subseteq V(\mathcal{H})$ we say that S is 2-covered in $\mathcal{H}$ if for every pair of vertices in S there is an edge in $\mathcal{H}$ containing it. If this holds for $S = V(\mathcal{H})$ then $\mathcal{H}$ itself is said to be 2-covered.

For all integers $\ell > r \geq 2$ we let $\mathcal{K}^r_\ell$ denote the family of r-graphs F with at most $\binom{\ell}{2}$ edges that contain a 2-covered set S of ℓ vertices called a core of F. The family $\mathcal{K}^r_\ell$ was first introduced by the second author [191] in order to extend Turán's theorem to hypergraphs. It also plays a key rôle in in the construction of the family $\mathcal{M}$ with two extremal configurations in [169]. In the present work, we also need the larger family $\widehat{\mathcal{K}}^r_\ell$ defined to consist of all r-graphs F with at most $\binom{\ell}{r}$ edges that contain a 2-covered set S of ℓ vertices, which is again called a core of F.

Let us recall that the transversal number of a hypergraph $\mathcal{H}$ is the nonnegative integer

$$\tau(\mathcal{H}) = \min\left\{|S|\colon S \subseteq V(\mathcal{H}) \text{ and } S \cap E \neq \emptyset \text{ for all } E \in \mathcal{H}\right\}.$$

Note that if $\mathcal{H}$ is empty, then we can take $S = \emptyset$, whence $\tau(\mathcal{H}) = 0$ holds in this case. After these preparations, the family $\mathcal{M}_t$ is defined as follows.

Definition 5.3.10. *For every positive integer t the family $\mathcal{M}_t$ consists of all 3-graphs $F \in \bigcup_{\ell \le n_t} \widehat{\mathcal{K}}^3_\ell$ which do not occur as a subgraph in any blow-up of $\mathcal{G}_1, \dots, \mathcal{G}_t$ and which have a core S such that $\tau(F[S]) \ge 2$.*

We conclude this subsection with a simple sufficient condition for 3-graphs $F \in \mathcal{K}^3_{n_t+1}$ guaranteeing that they are in $\mathcal{M}_t$ (see Lemma 5.3.13 below). For this purpose we require the following observation analysing the extent to which $\tau(\mathcal{H}) \ge 2$ is a "local" property of a hypergraph $\mathcal{H}$.

Fact 5.3.11. *If $r \ge 2$ and $\mathcal{H}$ denotes an r-graph with $\tau(\mathcal{H}) \ge 2$, then there is a subgraph $\mathcal{H}' \subseteq \mathcal{H}$ with at most $r+1$ edges satisfying $\tau(\mathcal{H}') \ge 2$.*

Proof. Pick two distinct edges $E', E'' \in \mathcal{H}$ and write $E' \cap E'' = \{v_1, \ldots, v_m\}$, where $0 \le m \le r-1$. For every $i \in [m]$ the assumption that $\{v_i\}$ fails to cover $\mathcal{H}$ yields an edge $E_i \in \mathcal{H}$ such that $v_i \notin E_i$. Now $\mathcal{H}' = \{E', E'', E_1, \ldots, E_m\}$ has the desired properties. ∎

Notice that the example $\mathcal{H} = K^r_{r+1}$ shows that the bound $|\mathcal{H}'| \le r+1$ is optimal.

Lemma 5.3.12. *Suppose that F is a 3-graph and that $S \subseteq V(F)$ is a 2-covered set in F. If $\tau(F[S]) \ge 2$, then F contains a subgraph F' such that $F' \in \mathcal{K}^3_{|S|}$ and $\tau(F'[S]) \ge 2$. Moreover, if $12 \le s \le |S|$, then F has a subgraph $F'' \in \mathcal{K}^3_s$ possessing a core S'' such that $\tau(F''[S'']) \ge 2$.*

Proof. The case $r = 3$ of Fact 5.3.11 yields a subgraph $\mathcal{G}$ of $F[S]$ with at most four edges such that $\tau(\mathcal{G}) \ge 2$. Notice that $|\mathcal{G}| \ge 2$ and $|\partial\mathcal{G}| \ge 5$. Since S is 2-covered in F, we can choose for every pair $uw \in \binom{S}{2} \setminus \partial\mathcal{G}$ an edge $e_{uw} \in F$ containing u and w. Now

$$F' = \left\{ e_{uw} \colon uw \in \tbinom{S}{2} \setminus \partial\mathcal{G} \right\} \cup \mathcal{G}$$

has the properties that S is 2-covered in F' and $\tau(F'[S]) \ge 2$. Together with

$$|F'| \le \binom{\ell}{2} - |\partial\mathcal{G}| + |\mathcal{G}| \le \binom{\ell}{2} - 5 + 4 < \binom{\ell}{2}$$

this proves $F' \in \mathcal{K}^3_{|S|}$. Moreover, if any $s \in [12, |S|]$ is given, we can take a set S'' of size s with $V(\mathcal{G}) \subseteq S'' \subseteq S$ and apply the first part of the lemma to S'' rather than S. ∎

Lemma 5.3.13. *If S denotes a core of $F \in \mathcal{K}^3_{n_t+1}$ and $\tau(F[S]) \ge 2$, then $F \in \mathcal{M}_t$.*

Proof. By the previous lemma and $n_t \geq 12$ there exists a set $S'' \subseteq S$ such that $|S''| = n_t$ and $\tau(F[S'']) \geq 2$. Since $|F| \leq \binom{n_t+1}{2} \leq \binom{n_t}{3}$, we can regard F as a member of $\widehat{\mathcal{K}}^3_{n_t}$ with core S'' and it remains to prove that F cannot be $\mathcal{G}_i$-colorable for any $i \in [t]$. This is due to the fact that the shadows of blow-ups of $\mathcal{G}_i$ are complete n_i-partite graphs, while S induces a K_{n_t+1} in ∂F. ∎

5.3.2.2 Turán numbers of $\mathcal{M}_t$

Having now introduced the main protagonists $\mathcal{G}_1, \ldots, \mathcal{G}_t$ and $\mathcal{M}_t$ we shall determine the extremal numbers $\mathrm{ex}(n, \mathcal{M}_t)$ in this subsection. More precisely, setting

$$\mathfrak{M}(n) = \max\{|\mathcal{G}| \colon \mathcal{G} \text{ is } \mathcal{G}_i\text{-colorable for some } i \in [t] \text{ and } v(\mathcal{G}) = n\}$$

for every positive integer n we shall prove the following result.

Theorem 5.3.14. *The equality* $\mathrm{ex}(n, \mathcal{M}_t) = \mathfrak{M}(n)$ *holds for all positive integers* n.

Notice that in view of Lemma 5.3.1 and Equation 5.21 this implies $\mathrm{ex}(n, \mathcal{M}_t) \leq \lambda_t n^3$ for every positive integer n. Moreover, whenever n is a multiple of n_i for some $i \in [t]$, the balanced blow-up of $\mathcal{G}_i$ with factor n/n_i exemplifies that this holds with equality. For these reasons, Theorem 5.3.14 is stronger than Theorem 5.1.9 (a). Let us start with the lower bound on $\mathrm{ex}(n, \mathcal{M}_t)$.

Fact 5.3.15. *We have* $\mathrm{ex}(n, \mathcal{M}_t) \geq \mathfrak{M}(n)$ *for every positive integer* n.

Proof. This is an immediate consequence of the fact that by Definition 5.3.10 for every $i \in [t]$ all blow-ups of $\mathcal{G}_i$ are $\mathcal{M}_t$-free. ∎

Our proof for the upper bound uses the Zykov symmetrization method [252]. The applicability of this technique in the current situation hinges on the fact that if a hypergraph $\mathcal{H}$ is $\mathcal{M}_t$-free, then there is no homomorphism from a member of $\mathcal{M}_t$ to $\mathcal{H}$ (see Proposition 5.3.16 below). Let us recall that given two r-graphs F and $\mathcal{H}$ a map $\phi\colon V(F) \longrightarrow V(\mathcal{H})$ is said to be a homomorphism if ϕ preserves edges, i.e., if $\phi(E) \in \mathcal{H}$ holds for all $E \in F$. Further, $\mathcal{H}$ is F-hom-free if there is no homomorphism from F to $\mathcal{H}$ or, in other words, if F fails to be $\mathcal{H}$-colourable. For a family $\mathcal{F}$ of r-graphs, we say that $\mathcal{H}$ is $\mathcal{F}$-hom-free if it is F-hom-free for every $F \in \mathcal{F}$.

Proposition 5.3.16. *A 3-graph $\mathcal{H}$ is $\mathcal{M}_t$-hom-free if and only if it is $\mathcal{M}_t$-free.*

Proof. Notice that the forward implication is clear. Now suppose conversely that $\mathcal{H}$ fails to be $\mathcal{M}_t$-hom-free, i.e., that there is a homomorphism $\phi\colon V(F) \longrightarrow V(\mathcal{H})$ for some $F \in \mathcal{M}_t$. Clearly the restriction of ϕ to a core S of F is injective. So $\phi(F) \in \widehat{\mathcal{K}}^3_{|S|} \cap \mathcal{M}_t$ and in view of $\phi(F) \subseteq \mathcal{H}$ it follows that $\mathcal{H}$ fails to be $\mathcal{M}_t$-free. ∎

As an immediate consequence of Definition 5.3.10, semibipartite triple systems are $\mathcal{M}_t$-free. We analyze the semibipartite case as follows.

Lemma 5.3.17. *If $\mathcal{H}$ denotes a semibipartite triple system on n vertices, then*

$$|\mathcal{H}| \leq \min\{2n^3/27, \mathfrak{M}(n)\}.$$

Proof. Fix a partition $V(\mathcal{H}) = A \cup B$ such that $|E \cap A| = 1$ holds for every $E \in \mathcal{H}$. Now the AM-GM inequality yields

$$|\mathcal{H}| \le |A|\binom{|B|}{2} \le \frac{2|A| \cdot |B| \cdot |B|}{4} \le \frac{1}{4}\left(\frac{2|A| + |B| + |B|}{3}\right)^3 = \frac{2n^3}{27}$$

and it remains to show $|\mathcal{H}| \le \mathfrak{M}(n)$. If n is large this is an immediate consequence of $\mathfrak{M}(n) = (\lambda_t - o(1))n^3$ and $\lambda_t \ge 5/32 > 2/27$, but for a complete proof addressing all values of n we need to argue more carefully.

To this end we consider a random map $\phi\colon [n] \longrightarrow [n_1]$ together with the random blow-up $\widehat{\mathcal{G}}$ of $\mathcal{G}_1$ determined by ϕ. Explicitly $\widehat{\mathcal{G}}$ has vertex set $[n]$ and a triple ijk forms an edge of $\widehat{\mathcal{G}}$ if and only if $\phi(i)\phi(j)\phi(k) \in \mathcal{G}_1$. Now every potential edge of $\widehat{\mathcal{G}}$ is present with probability $\frac{6|\mathcal{G}_i|}{n_1^3} = 6\lambda_t$ and thus the expectation of $|\widehat{\mathcal{G}}|$ is $6\lambda_t\binom{n}{3}$. So by averaging we obtain

$$\mathfrak{M}(n) \ge 6\lambda_t\binom{n}{3} \ge \frac{15}{16}\binom{n}{3}, \tag{5.22}$$

which for $n \ge 5$ implies the desired estimate $\mathfrak{M}(n) \ge 2n^3/27$. Moreover, Equation 5.22 yields $\mathfrak{M}(4) \ge 3$, which still suffices for the case $n = 4$ of our lemma. Finally, the case $n \le 3$ is trivial. ∎

The central notion in arguments based on Zykov symmetrization is the following: Given an r-graph $\mathcal{H}$, two non-adjacent vertices $u, v \in V(\mathcal{H})$ are said to be equivalent if $L_{\mathcal{H}}(u) = L_{\mathcal{H}}(v)$. Evidently, equivalence is an equivalence relation. Since any two equivalent vertices have the same degree and the same link, we can write $d_{\mathcal{H}}(C)$ and $L_{\mathcal{H}}(C)$ for the common degree and the common link of all vertices in an equivalence class C, respectively.

Lemma 5.3.18. *Let $\mathcal{H}$ be an $\mathcal{M}_t$-free 3-graph with equivalence classes $C_1, \dots, C_m$. If for all distinct $k, \ell \in [m]$ the shadow $\partial\mathcal{H}$ induces a complete bipartite graph between C_k and C_ℓ, then $\mathcal{H}$ is either semibipartite or $\mathcal{G}_i$-colourable for some $i \in [t]$.*

Proof. Let $T \subseteq V(\mathcal{H})$ be a set containing exactly one vertex from each equivalence class of $\mathcal{H}$, and let $\mathcal{T}$ be the subgraph of $\mathcal{H}$ induced by T. By assumption, $\mathcal{T}$ is 2-covered, $|T| = m$, and $\mathcal{H}$ is a blow-up of $\mathcal{T}$. If $\tau(\mathcal{T}) < 2$, then $\mathcal{T}$ is a star and $\mathcal{H}$ is semibipartite. So we may assume $\tau(\mathcal{T}) \geq 2$ from now on.

Since $\mathcal{T}$ is 2-covered and $|\mathcal{T}| \leq \binom{m}{3}$ we have $\mathcal{T} \in \widehat{\mathcal{K}}^3_m$. So if $m \leq n_t$, then in view of Definition 5.3.10 and $\mathcal{T} \notin \mathcal{M}_t$ there exists an index $i \in [t]$ such that $\mathcal{T}$ is $\mathcal{G}_i$-colorable. As $\mathcal{H}$ is a blow-up of $\mathcal{T}$, it follows that $\mathcal{H}$ is $\mathcal{G}_i$-colorable as well.

Now assume for the sake of contradiction that $m > n_t$. Since $n_t \geq 12$, Lemma 5.3.12 leads to a subgraph $\mathcal{T}'' \in \mathcal{K}^3_{n_t+1}$ of $\mathcal{T}$ having a core S'' such that $\tau(\mathcal{T}''[S'']) \geq 2$. By Lemma 5.3.13 this contradicts $\mathcal{H}$ being $\mathcal{M}_t$-free. ∎

Now we are ready to establish the main result of this subsection.

Proof of Theorem 5.3.14. Fix some positive integer n. By Fact 5.3.15 it suffices to establish the upper bound $\mathrm{ex}(n, \mathcal{M}_t) \leq \mathfrak{M}(n)$. Arguing indirectly we choose an $\mathcal{M}_t$-free triple system $\mathcal{H}$ on n vertices with more than $\mathfrak{M}(n)$ edges such that the number m of equivalence classes of $\mathcal{H}$ is minimal. Let $C_1, \dots, C_m$ be the equivalence classes of $\mathcal{H}$.

By Lemma 5.3.17 we know that $\mathcal{H}$ is not semibipartite and the definition of $\mathfrak{M}(n)$ implies that $\mathcal{H}$ fails to be $\mathcal{G}_i$-colorable for every $i \in [t]$. For these reasons, Lemma 5.3.18 tells us that ∂H is not the complete m-partite graph with vertex classes $C_1, \dots, C_m$. Without loss of generality we may assume that at least one possible edge between C_1 and C_2 is missing in ∂H. Due to the definition of equivalence there are actually no edges between C_1 and C_2 in ∂H. By symmetry we may suppose further that $d_{\mathcal{H}}(C_1) \le d_{\mathcal{H}}(C_2)$.

Now let $\mathcal{H}'$ be the unique 3-graph satisfying $V(\mathcal{H}') = V(\mathcal{H})$, $\mathcal{H}' - C_1 = \mathcal{H} - C_1$, and $L_{\mathcal{H}'}(v) = L_{\mathcal{H}}(w)$ for all $v \in C_1$ and $w \in C_2$. Observe that $\{C_1 \cup C_2, C_3, \dots, C_m\}$ refines the partition of $V(\mathcal{H}')$ into the equivalence classes of $\mathcal{H}'$ and

$$|\mathcal{H}'| = |\mathcal{H}| + |C_1|\big(d_{\mathcal{H}}(C_2) - d_{\mathcal{H}}(C_1)\big) \ge |\mathcal{H}| > \mathfrak{M}(n).$$

So our minimal choice of m implies that $\mathcal{H}'$ cannot be $\mathcal{M}_t$-free. As there exists a homomorphism from $\mathcal{H}'$ to $\mathcal{H}$, it follows that $\mathcal{H}$ fails to be $\mathcal{M}_t$-hom-free. But owing to Proposition 5.3.16 this contradicts $\mathcal{H}$ being $\mathcal{M}_t$-free. ∎

5.3.3 Stability

In this section we prove most of Theorem 5.1.9 (b) – only the proof of $\xi(\mathcal{M}_t) = t$ is postponed to Section 5.3.4. Our goal is to show that after deleting a small number of low-degree vertices an "almost extremal" $\mathcal{M}_t$-free 3-graph becomes $\mathcal{G}_i$-colorable for some $i \in [t]$. More precisely, we aim for the following result.

Theorem 5.3.19. *If $\epsilon > 0$ is sufficiently small, n is sufficiently large, and $\mathcal{H}$ is an $\mathcal{M}_t$-free 3-graph on n vertices with $|\mathcal{H}| \geq (\lambda_t - \epsilon)n^3$, then the set*

$$Z = \left\{u \in V(\mathcal{H}) \colon d_{\mathcal{H}}(u) \leq (3\lambda_t - 2\epsilon^{1/2})n^2\right\}$$

has size at most $\epsilon^{1/2}n$ and the 3-graph $\mathcal{H} - Z$ is $\mathcal{G}_i$-colorable for some $i \in [t]$.

As the proof of this result will occupy the entire section, we would like to start with a quick overview. The argument is somewhat similar in spirit to [208; 30; 169] and ultimately it is based on the Zykov symmetrization method [252]. There are certain kinds of complications that often arise when one uses this strategy in order to establish stability results and we overcome several of these common difficulties by introducing the so-called Ψ-trick in Subsection 5.3.3.1. By means of this trick, the problem to prove Theorem 5.3.19 gets reduced to an apparently much simpler task: If a triple system $\mathcal{H}$ with n vertices and minimum degree $(3\lambda_t - o(1))n^2$ can be made $\mathcal{G}_i$-colorable by deleting a single vertex, then, actually, $\mathcal{H}$ itself is $\mathcal{G}_i$-colorable (see Lemma 5.3.21). The Ψ-trick can also be used to reprove some known stability results with improved control over the dependence of the constants (see [171]).

The proof of Lemma 5.3.21 is still quite long. We will collect some auxiliary results in Subsection 5.3.3.2 and defer the main part of the argument to Subsection 5.3.3.3

5.3.3.1 General preliminaries.

This subsection reduces the task of proving Theorem 5.3.19 to the apparently much simpler task of verifying Lemma 5.3.21 below. There are only few "special properties" of $\mathcal{M}_t$ we are going to utilize in the course of this reduction and we refer to [171] for a more systematic treatment.

Throughout this subsection we use the following notation: For every 3-graph $\mathcal{H}$ on n vertices and every $\epsilon > 0$ we set

$$Z_\epsilon(\mathcal{H}) = \left\{u \in V(\mathcal{H})\colon d_{\mathcal{H}}(u) \le (3\lambda_t - 2\epsilon^{1/2})n^2\right\}.$$

Lemma 5.3.20. *If $\epsilon \in (0,1)$, $n \ge \epsilon^{-1/2}$ and $\mathcal{H}$ is an $\mathcal{M}_t$-free 3-graph on n vertices with at least $(\lambda - \epsilon)n^3$ edges, then*

(a) the set $Z_\epsilon(\mathcal{H})$ has at most the size $\epsilon^{1/2}n$

(b) and the subgraph $\mathcal{H}' = \mathcal{H} - Z_\epsilon(\mathcal{H})$ of $\mathcal{H}$ satisfies $\delta(\mathcal{H}') \ge (3\lambda_t - 3\epsilon^{1/2})n^2$ as well as $|\mathcal{H}'| \ge (\lambda_t - 2\epsilon^{1/2})n^3$.

Proof. Set $Z = Z_\epsilon(\mathcal{H})$. Assuming that part (a) fails we can take a set $X \subseteq Z$ of size $\frac{2}{3}\epsilon^{1/2}n \le |X| \le 2\epsilon^{1/2}n$. The definition of Z leads to

$$\begin{aligned}
|\mathcal{H} - X| &\ge (\lambda_t - \epsilon)n^3 - |X|(3\lambda_t - 2\epsilon^{1/2})n^2 \\
&\ge (\lambda_t - \epsilon)n^3 - |X|(3\lambda_t - 2\epsilon^{1/2})n^2 - \tfrac{3}{4}n(|X| - \tfrac{2}{3}\epsilon^{1/2}n)(2\epsilon^{1/2}n - |X|) \\
&= \lambda_t(n - |X|)^3 + 3(1/4 - \lambda_t)n|X|^2 + \lambda_t|X|^3 > \lambda_t(n - |X|)^3,
\end{aligned}$$

where we used $\lambda_t < 1/6 < 1/4$ in the last step. However, by Theorem 5.1.9 (a) this contradicts the fact that $\mathcal{H} - X$ is $\mathcal{M}_t$-free.

Now we prove part (b). For every $u \in V(\mathcal{H}')$ the definition of Z and (a) yield

$$d_{\mathcal{H}'}(u) \geq d_{\mathcal{H}}(u) - |Z|n \geq (3\lambda_t - 2\epsilon^{1/2})n^2 - \epsilon^{1/2}n^2 = (3\lambda_t - 3\epsilon^{1/2})n^2.$$

Similarly, we have

$$|\mathcal{H}'| \geq |\mathcal{H}| - |Z|n^2 \geq (\lambda_t - \epsilon)n^3 - \epsilon^{1/2}n^3 > (\lambda_t - 2\epsilon^{1/2})n^3. \blacksquare$$

The following lemma will be shown to imply Theorem 5.3.19.

Lemma 5.3.21. *There exist constants $\zeta \in (0,1)$ and $N_0 \in \mathbb{N}$ such that the following holds for all $n \geq N_0$. Let $\mathcal{H}$ be an $\mathcal{M}_t$-free 3-graph on n vertices with at least $(\lambda_t - \zeta)n^3$ edges and $\delta(\mathcal{H}) > (3\lambda_t - \zeta)n^2$. If there exists a vertex $v \in V(\mathcal{H})$ such that $\mathcal{H} - v$ is $\mathcal{G}_i$-colorable for some $i \in [t]$, then $\mathcal{H}$ itself is $\mathcal{G}_i$-colorable as well.*

We postpone the proof of this result to Subsection 5.3.3.3. The deduction of Theorem 5.3.19 from Lemma 5.3.21 factorises through the following statement.

Lemma 5.3.22. *There exists $\epsilon \in (0, 1/16)$ such that the following holds for every sufficiently large integer n. Let $\mathcal{H}$ denote an $\mathcal{M}_t$-free 3-graph with n vertices and at least $(\lambda_t - \epsilon)n^3$ edges. If $Q \subseteq V(\mathcal{H})$ has size $|Q| \leq 2\epsilon^{1/2}n$ and $\mathcal{H} - Q$ is $\mathcal{G}_i$-colourable for some $i \in [t]$, then $\mathcal{H} - Z_\epsilon(\mathcal{H})$ is $\mathcal{G}_i$-colourable as well.*

Proof of Lemma 5.3.22 using Lemma 5.3.21. We show that $\epsilon = \zeta^2/25$ has the desired property, where ζ denotes the constant provided by Lemma 5.3.21. Given a sufficiently large 3-graph $\mathcal{H}$

and a set Q as described in the statement of Lemma 5.3.22 we set $Q' = Q \setminus Z_\epsilon(\mathcal{H})$ and $V' = V(\mathcal{H}) \setminus (Z_\epsilon(\mathcal{H}) \cup Q)$.

By our assumption, there is an index $i(\star) \in [t]$ such that $\mathcal{H}[V']$ is $\mathcal{G}_{i(\star)}$-colorable. Choose a set $S \subseteq Q'$ of maximum size such that $\mathcal{H}[V' \cup S]$ is still $\mathcal{G}_{i(\star)}$-colorable. If $S = Q'$ we are done, so suppose for the sake of contradiction that there exists a vertex $v \in Q' \setminus S$.

Due to the maximality of S the triple system $\mathcal{H}' = \mathcal{H}[V' \cup S \cup \{v\}]$ is not $\mathcal{G}_{i(\star)}$-colorable. On the other hand, Lemma 5.3.20 (a) and $|Q| \le 2\epsilon^{1/2}n$ entail

$$\delta(\mathcal{H}') > (3\lambda_t - 2\epsilon^{1/2})n^2 - |Z(\mathcal{H}) \cup Q|n > (3\lambda_t - 5\epsilon^{1/2})n^2$$

$$\text{and} \qquad |\mathcal{H}'| > (\lambda_t - \epsilon)n^3 - |Z(\mathcal{H}) \cup Q|n^2 > (\lambda_t - 4\epsilon^{1/2})n^3.$$

So by Lemma 5.3.21 and $\zeta = 5\epsilon^{1/2}$ the $\mathcal{G}_{i(\star)}$-colorability of $\mathcal{H}' - v = \mathcal{H}[V' \cup S]$ implies that $\mathcal{H}'$ itself is $\mathcal{G}_{i(\star)}$-colorable as well. This contradiction completes the proof of Lemma 5.3.22. ∎

It remains to deduce Theorem 5.3.19. The argument involves the following invariant of 3-graphs: Given a 3-graph $\mathcal{H}$ with equivalence classes $C_1, \dots, C_m$ we set $\Psi(\mathcal{H}) = \sum_{i=1}^m |C_i|^2$.

Proof of Theorem 5.3.19 using Lemma 5.3.22. Let ϵ be the constant delivered by Lemma 5.3.22 and fix a sufficiently large natural number n. Assuming that the conclusion of Theorem 5.3.19 fails for our values of ϵ and n we pick a counterexample $\mathcal{H}$ such that the pair $(|\mathcal{H}|, \Psi(\mathcal{H}))$ is lexicographically maximal. Let $C_1, \dots, C_m$ be the equivalence classes of $\mathcal{H}$.

Recall that Lemma 5.3.20 (a) tells us $|Z_\epsilon(\mathcal{H})| \leq \epsilon^{1/2}n$. Since $\mathcal{H}$ is a counterexample, it cannot be $\mathcal{G}_i$-colorable for any $i \in [t]$. Moreover, Equation 5.21 yields

$$|\mathcal{H}| > (\lambda_t - \epsilon)n^3 \geq (5/32 - 1/16)n^3 = 3n^3/32 > 2n^3/27$$

and thus $\mathcal{H}$ cannot be semibipartite. So by Lemma 5.3.18 there exist two equivalence classes, say C_1 and C_2, such that $\partial\mathcal{H}$ possesses no edges from C_1 to C_2. We may assume that $(d_\mathcal{H}(C_1), |C_1|) \leq_{\text{lex}} (d_\mathcal{H}(C_2), |C_2|)$, where $\leq_{\text{lex}}$ indicates the lexicographic ordering on $\mathbb{N}^2$.

Pick arbitrary vertices $v_1 \in C_1$ and $v_2 \in C_2$ and symmetrize only them. That is, we let $\mathcal{H}'$ be the 3-graph with $V(\mathcal{H}') = V(\mathcal{H})$, $\mathcal{H}' - v_1 = \mathcal{H} - v_1$ and $L_{\mathcal{H}'}(v_1) = L_\mathcal{H}(v_2)$. Clearly, if $d_\mathcal{H}(v_1) < d_\mathcal{H}(v_2)$, then $|\mathcal{H}'| > |\mathcal{H}|$. Moreover, if $d_\mathcal{H}(v_1) = d_\mathcal{H}(v_2)$, then $|\mathcal{H}'| = |\mathcal{H}|$, $|C_1| \leq |C_2|$, and

$$\Psi(\mathcal{H}') - \Psi(\mathcal{H}) \geq (|C_1| - 1)^2 + (|C_2| + 1)^2 - |C_1|^2 - |C_2|^2 = 2(|C_2| - |C_1| + 1) \geq 2.$$

In both cases $(|\mathcal{H}'|, \Psi(\mathcal{H}'))$ is lexicographically larger than $(|\mathcal{H}|, \Psi(\mathcal{H}))$ and our choice of $\mathcal{H}$ implies that $\mathcal{H}' - Z_\epsilon(\mathcal{H}')$ is $\mathcal{G}_i$-colourable for some $i \in [t]$. By Lemma 5.3.20 (a) the set $Q = Z_\epsilon(\mathcal{H}') \cup \{v_1\}$ has size $|Q| \leq \epsilon^{1/2}n + 1 < 2\epsilon^{1/2}n$. Since the hypergraph $\mathcal{H} - Q = \mathcal{H}' - Q$ is $\mathcal{G}_i$-colourable, Lemma 5.3.22 implies that $\mathcal{H} - Z(\mathcal{H})$ is $\mathcal{G}_i$-colourable too. This contradiction to the choice of $\mathcal{H}$ establishes Theorem 5.3.19. ∎

5.3.3.2 Transversals

Roughly speaking, the hypergraph $\mathcal{H} - v$ appearing in Lemma 5.3.21 arises from an almost balanced blow-up of $\mathcal{G}_i$ by deleting a small number of edges. When we randomly select one vertex from each partition class of $\mathcal{H} - v$ it is thus very likely that the resulting transversal induces a copy of $\mathcal{G}_i$. In the proof of Lemma 5.3.21 there are several places where we argue similarly in situations where some vertices from the transversals have been selected in advance. The precise statement we shall use in these cases is Lemma 5.3.23 below.

Consider a 3-graph with $V(\mathcal{G}) = [m]$ and pairwise disjoint sets $V_1, \dots, V_m$. The blow-up $\mathcal{G}[V_1, \dots, V_m]$ of $\mathcal{G}$ is obtained from $\mathcal{G}$ by replacing each vertex $j \in [m]$ with the set V_j and each edge $\{j_1, j_2, j_3\} \in \mathcal{G}$ with the complete 3-partite 3-graph with vertex classes V_{j_1}, V_{j_2}, and V_{j_3}. For a 3-graph $\mathcal{H}$ we say that a partition $V(H) = \bigcup_{j \in [m]} V_j$ is a $\mathcal{G}$-coloring of $\mathcal{H}$ if $\mathcal{H} \subseteq \mathcal{G}[V_1, \dots, V_m]$.

Lemma 5.3.23. *Fix a real $\eta \in (0,1)$ and integers $m, n \geq 1$. Let $\mathcal{G}$ be a 3-graph with vertex set $[m]$ and let $\mathcal{H}$ be a further 3-graph with $v(\mathcal{H}) = n$. Consider a vertex partition $V(\mathcal{H}) = \bigcup_{i \in [m]} V_i$ and the associated blow-up $\widehat{\mathcal{G}} = \mathcal{G}[V_1, \dots, V_m]$ of $\mathcal{G}$. If two sets $T \subseteq [m]$ and $S \subseteq \bigcup_{j \notin T} V_j$ have the properties*

(a) $|V_j| \geq (|S| + 1)|T|\eta^{1/3} n$ *for all* $j \in T$,

(b) $|\mathcal{H}[V_{j_1}, V_{j_2}, V_{j_3}]| \geq |\widehat{\mathcal{G}}[V_{j_1}, V_{j_2}, V_{j_3}]| - \eta n^3$ *for all* $\{j_1, j_2, j_3\} \in \binom{T}{3}$,

(c) and $|L_{\mathcal{H}}(v)[V_{j_1}, V_{j_2}]| \geq |L_{\widehat{\mathcal{G}}}(v)[V_{j_1}, V_{j_2}]| - \eta n^2$ *for all* $v \in S$ *and* $\{j_1, j_2\} \in \binom{T}{2}$,

then there exists a selection of vertices $u_j \in V_j$ for all $j \in [T]$ such that $U = \{u_j \colon j \in T\}$ satisfies $\widehat{\mathcal{G}}[U] \subseteq \mathcal{H}[U]$ and $L_{\widehat{\mathcal{G}}}(v)[U] \subseteq L_{\mathcal{H}}(v)[U]$ for all $v \in S$. In particular, if $\mathcal{H} \subseteq \widehat{\mathcal{G}}$, then $\widehat{\mathcal{G}}[U] = \mathcal{H}[U]$ and $L_{\widehat{\mathcal{G}}}(v)[U] = L_{\mathcal{H}}(v)[U]$ for all $v \in S$.

Proof. Choose for $j \in T$ the vertices $u_j \in V_j$ independently and uniformly at random and let $U = \{u_j \colon j \in T\}$ be the random transversal consisting of these vertices. By (a) and (b) we have

$$\mathbb{P}\left(\{u_{j_1}, u_{j_2}, u_{j_3}\} \notin \mathcal{H}\right) = 1 - \frac{|\mathcal{H}[V_{j_1}, V_{j_2}, V_{j_3}]|}{|V_{j_1}||V_{j_2}||V_{j_3}|} \leq \frac{\eta n^3}{|V_{j_1}||V_{j_2}||V_{j_3}|} \leq \frac{1}{(|S|+1)^3|T|^3}$$

for all edges $\{j_1, j_2, j_3\} \in \mathcal{G}$. Similarly (a) and (c) lead to

$$\begin{aligned}\mathbb{P}\Big(\{u_{j_1}, u_{j_2}\} \notin L_{\mathcal{H}}(v) \mid \{u_{j_1}, u_{j_2}\} \in L_{\widehat{\mathcal{G}}}(v)\Big) &= 1 - \frac{|L_{\mathcal{H}}(v)[V_{j_1}, V_{j_2}]|}{|V_{j_1}||V_{j_2}|} \leq \frac{\eta n^2}{|V_{j_1}||V_{j_2}|} \\ &\leq \frac{\eta^{1/3}}{(|S|+1)^2|T|^2}\end{aligned}$$

for all $v \in S$ and all distinct $j_1, j_2 \in [m]$. Therefore, the union bound reveals

$$\begin{aligned} &\mathbb{P}\Big(\widehat{\mathcal{G}}[U] \not\subseteq \mathcal{H}[U]\Big) \leq \binom{|T|}{3}\frac{1}{(|S|+1)^3|T|^3} < \frac{1}{6} \\ \text{and} \qquad &\mathbb{P}\Big(L_{\widehat{\mathcal{G}}}(v) \not\subseteq L_{\mathcal{H}}(v)\Big) \leq \binom{|T|}{2}\frac{\eta^{1/3}}{(|S|+1)^2|T|^2} < \frac{1}{2(|S|+1)} \qquad \text{for every } v \in S.\end{aligned}$$

Altogether, the probability that U fails to have the desired properties is at most

$$\frac{1}{6}+\frac{|S|}{2(|S|+1)}<\frac{2}{3}.$$

So the probability that U has these properties is positive. ∎

In practice the sets U obtained by means of Lemma 5.3.23 will be 2-covered and thus they will be cores of some subgraphs $F \in \widehat{\mathcal{K}}^3_{|U|}$ of $\mathcal{H}$. In such situations F will be $\mathcal{M}_t$-free and in order to exploit this fact we need to know that for $i \neq j$ the triple system $\mathcal{G}_i$ is in some sense far from being $\mathcal{G}_j$-colorable (see Lemma 5.3.25 below). The verification of this statement requires that we take a closer look into Construction 5.3.9 and the observation that follows summarizes everything we need in the sequel.

Observation 5.3.24. *The triple systems $\mathcal{G}_1, \ldots, \mathcal{G}_t$ have the following properties.*

(a) For $i \in [t]$ and $v \in \mathcal{G}_i$ the clique number $\omega\left(L_{\mathcal{G}_i}(v)\right)$ of the link graph $L_{\mathcal{G}_i}(v)$ satisfies

$$\frac{n_i-1}{k_i-1}-\frac{k_i}{2}\le \omega\left(L_{\mathcal{G}_i}(v)\right)\le \frac{n_i-1}{k_i-1}.$$

(b) We have

$$\frac{n_i-1}{k_i-1}-\frac{n_{i+1}-1}{k_{i+1}-1}>\frac{Q}{k_i^2}$$

for every $i \in [t-1]$, where

$$Q=\frac{n_1}{k_1+1}=\cdots=\frac{n_t}{k_t+1}\ge 2k_t^3\ge 16.$$

(c) For $i \in [t]$ the 3-graph $\mathcal{G}_i$ is regular with degree $3\lambda_t n_i^2$ and

$$7n_i/8 \le n_i - 3k_i/2 \le \delta_2(\mathcal{G}_i) \le \Delta_2(\mathcal{G}_i) \le n_i - k_i.$$

Proof. Part (a) follows from the fact that due to $\mathcal{G}_i = (K^3_{n_i} \setminus H(\mathcal{D}_i)) \setminus \mathcal{S}_i$ the link $L_{\mathcal{G}_i}(v)$ arises from an $((n_i - 1)/(k_i - 1))$-partite Turán graph by the deletion of $k_i/2$ edges. The proof of part (c) is similar. For part (b) it suffices to calculate

$$\begin{aligned}\frac{n_i - 1}{k_i - 1} - \frac{n_{i+1} - 1}{k_{i+1} - 1} &= \frac{Q(k_i + 1) - 1}{k_i - 1} - \frac{Q(k_{i+1} + 1) - 1}{k_{i+1} - 1} \\ &= (2Q - 1)\left(\frac{1}{k_i - 1} - \frac{1}{k_{i+1} - 1}\right) \\ &\ge Q\left(\frac{1}{k_i - 1} - \frac{1}{k_i}\right) > \frac{Q}{k_i^2}.\end{aligned}$$

∎

As indicated earlier, this has the following consequence.

Lemma 5.3.25. *If $i \in [t]$ and the triple system $\mathcal{G}_i'$ arises from $\mathcal{G}_i$ by the deletion of at most $Q/(2k_i^2)$ vertices, then $\mathcal{G}_i'$ fails to be $\mathcal{G}_j$-colorable for every $j \in [t] \setminus \{i\}$.*

Proof. Suppose first that $j \in [i - 1]$. Due to

$$\delta_2(\mathcal{G}_i') \ge \delta_2(\mathcal{G}_i) - \frac{Q}{2k_i^2} \ge 1$$

we know that $\mathcal{G}_i'$ is 2-covered. Together with

$$v(\mathcal{G}_i') = n_i - Q/(2k_i^2) > Q(k_i+1) - Q \geq Q(k_j+1) = n_j$$

it follows that $\mathcal{G}_i'$ is indeed not $\mathcal{G}_j$-colorable.

If $j \in (i,t]$ we take an arbitrary vertex $v \in V(\mathcal{G}_i')$. The parts (a) and (b) of Observation 5.3.24 yield

$$\omega\left(L_{\mathcal{G}_i'}(v)\right) \geq \omega\left(L_{\mathcal{G}_i}(v)\right) - \frac{Q}{2k_i^2} \geq \frac{n_i-1}{k_i-1} - \frac{k_i}{2} - \frac{Q}{2k_i^2} > \frac{n_j-1}{k_j-1}.$$

On the other hand, by Observation 5.3.24 (a) again, any $\mathcal{G}_j$-coloring of $\mathcal{G}_i'$ would show that

$$\omega\left(L_{\mathcal{G}_i'}(v)\right) \leq \omega\left(L_{\mathcal{G}_j}(v)\right) \leq \frac{n_j-1}{k_j-1}.$$

∎

On most occasions the following corollary of Lemma 5.3.25 will suffice.

Corollary 5.3.26. *If $i \in [t]$, the 3-graph $\mathcal{H}$ is $\mathcal{M}_t$-free and $U \subseteq V(\mathcal{H})$ denotes a 2-covered set of size n_i+1, then $\mathcal{H}[U]$ is $\mathcal{G}_i$-free.*

Proof. Assume for the sake of contradiction that $\mathcal{H}[U]$ has a subgraph isomorphic to $\mathcal{G}_i$. If $i < t$ we can take a subgraph $F \in \widehat{\mathcal{K}}^3_{n_i+1}$ of $\mathcal{H}$ with $F[U] = \mathcal{H}[U]$ having U as a core. As $\mathcal{H}[U]$ contains a copy of $\mathcal{G}_i$, we have $\tau(F[U]) \geq 2$. Now $F \notin \mathcal{M}_t$ implies that F is $\mathcal{G}_j$-colorable for some $j \in [t]$.

In particular, $\mathcal{G}_i$ is $\mathcal{G}_j$-colorable and by Lemma 5.3.25 this leads to $i = j$. In other words, F is $\mathcal{G}_i$-colorable, contrary to the fact that ∂F contains a copy of K_{n_i+1}.

It remains to discuss the case $i = t$. Now Lemma 5.3.12 yields a subgraph F' of F which belongs to $\mathcal{K}^3_{n_t+1}$, and whose induced subgraph on its core has covering number at least 2. By Lemma 5.3.13 this contradicts $\mathcal{H}$ being $\mathcal{M}_t$-free. ∎

5.3.3.3 Proof of the main lemma

This entire subsection is devoted to the proof of Lemma 5.3.21. Select constants ζ and N_0 fitting into the hierarchy

$$N_0^{-1} \ll \zeta \ll n_t^{-1}.$$

Consider an $\mathcal{M}_t$-free 3-graph $\mathcal{H}$ on $n \geq N_0$ vertices satisfying $|\mathcal{H}| \geq (\lambda_t - \zeta)n^3$ and $\delta(\mathcal{H}) \geq (3\lambda - \zeta)n^2$ such that for some $v \in V(\mathcal{H})$ and $i \in [t]$ the 3-graph $\mathcal{H}_v = \mathcal{H} \setminus \{v\}$ is $\mathcal{G}_i$-colorable. Set $V = V(\mathcal{H})$ and fix a partition $\bigcup_{i \in [n_i]} V_i = V \setminus \{v\}$ exemplifying the $\mathcal{G}_i$-colorability of $\mathcal{H}_v$. We divide the argument that follows into three main parts each of which consists of several claims.

Part I. Analysis of $\mathcal{H}_v$. The three claims that follow only deal with $\mathcal{H}_v$ but say nothing about v and its link.

Claim 5.3.27. *We have $|V_j| = n/n_i \pm 5\zeta^{1/2}n$ for every $j \in [n_i]$.*

Proof. Set $x_j = |V_j|/(n-1)$ for every $j \in [n_i]$. By Proposition 5.4.2 (and the proof of Lemma 5.3.1) we obtain

$$|\mathcal{H}_v| = L_{\mathcal{G}_i}(x_1,\dots,x_{n_i})(n-1)^3 \le \left(\lambda_t - \frac{1}{9}\sum_{j\in[n_i]}\left(x_j - \frac{1}{n_i}\right)^2\right) n^3.$$

Combined with

$$|\mathcal{H}_v| \ge (\lambda_t - \zeta)n^3 - d_{\mathcal{H}}(v) > (\lambda_t - 2\zeta)n^3$$

this leads to $\frac{1}{9}\sum_{j\in[n_i]}(x_j - 1/n_i)^2 \le 2\zeta$, whence $x_j = 1/n_i \pm (18\zeta)^{1/2}$ and

$$\big||V_j| - n/n_i\big| \le (n-1)\big|x_j - 1/n_i\big| + 1/n_i \le (18\zeta)^{1/2}n + 1/n_i \le 5\zeta^{1/2}n.$$

∎

Recall that the sets $V_1,\dots,V_{n_i}$ have been chosen in such a way that $\mathcal{H}_v$ is a subgraph of the blow-up $\widehat{\mathcal{G}}_i = \mathcal{G}_i[V_1,\dots,V_{n_i}]$ of $\mathcal{G}_i$. Our next objective is to compare the links of an arbitrary vertex $u \in V\setminus\{v\}$ in $\mathcal{H}_v$ and in $\widehat{\mathcal{G}}$. As a consequence of $\mathcal{H}_v \subseteq \widehat{\mathcal{G}}_i$ we know $L_{\mathcal{H}_v}(u) \subseteq L_{\widehat{\mathcal{G}}_i}(u)$ and $|L_{\mathcal{H}_v}(u)| \le |L_{\widehat{\mathcal{G}}_i}(u)|$. Members of $L_{\widehat{\mathcal{G}}_i}(u)\setminus L_{\mathcal{H}_v}(u)$ are referred to as the missing pairs of u. By Lemma 5.3.1 the global number of missing edges can be bounded from above by

$$\big|\widehat{\mathcal{G}}_i \setminus \mathcal{H}_v\big| \le \lambda_t(n-1)^3 - (\lambda_t - \zeta)n^3 + d_{\mathcal{H}}(v) \le 2\zeta n^3. \tag{5.23}$$

Locally we obtain the following.

Claim 5.3.28. *Every $u \in V \setminus \{v\}$ satisfies $|L_{\widehat{\mathcal{G}}_i}(u)| < (3\lambda_t + 6n_i\zeta^{1/2})n^2$. Moreover the number of missing pairs of u is bounded by $|L_{\widehat{\mathcal{G}}_i}(u) \setminus L_{\mathcal{H}_v}(u)| < 7\zeta^{1/2}n_i n^2$.*

Proof. Since $\mathcal{G}_i$ is $(3\lambda_t n_i^2)$-regular, Claim 5.3.27 yields

$$\big|L_{\widehat{\mathcal{G}}_i}(u)\big| \le 3\lambda_t n_i^2 \left(\frac{n}{n_i} + 5\zeta^{1/2}n\right)^2 = 3\lambda_t n^2 \left(1 + 5\zeta^{1/2}n_i\right)^2 < (3\lambda_t + 6\zeta^{1/2}n_i)n^2,$$

where we used $\lambda_t < 1/6$ and our hierarchy $\zeta \ll n_i^{-1}$. Owing to the minimum degree condition $\delta(\mathcal{H}) \ge (3\lambda_t - \zeta)n^2$ this entails the upper bound

$$\big|L_{\widehat{\mathcal{G}}_i}(u) \setminus L_{\mathcal{H}_v}(u)\big| \le \left(3\lambda_t n^2 + 6\zeta^{1/2}n_i n^2\right) - \left(3\lambda_t n^2 - \zeta n^2 - n\right) < 7\zeta^{1/2}n_i n^2$$

on the number of missing pairs of u. ∎

It can now be shown that in $\mathcal{H}_v$ all neighborhoods have roughly the expected size $\frac{n_i-1}{n_i}n$, but for our concerns it suffices to establish a lower bound.

Claim 5.3.29. *We have $|N_{\mathcal{H}_v}(u)| \ge \frac{n_i-1}{n_i}n - 17\zeta^{1/2}n_i n$ for every $u \in V \setminus \{v\}$.*

Proof. Let $j \in [n_i]$ be the index satisfying $u \in V_j$. Since every vertex in $V \setminus (V_j \cup N_{\mathcal{H}_v}(u) \cup \{v\})$ belongs to at least $\delta_2(\mathcal{G}) \cdot \min\{|V_\ell| \colon \ell \in [n_i]\}$ missing pairs of u, and every missing pair is counted at most twice in this manner, Claim 5.3.28 yields

$$\big|V \setminus (V_j \cup N_{\mathcal{H}_v}(u) \cup \{v\})\big| \cdot \delta_2(\mathcal{G}) \cdot \min\big\{|V_\ell| \colon \ell \in [n_i]\big\} < 14\zeta^{1/2}n_i n^2.$$

So by Observation 5.3.24 (c) and Claim 5.3.27 the assumption $|N_{\mathcal{H}_v}(u)| < \frac{n_i-1}{n_i}n - 17\zeta^{1/2}n_i n$ would yield the contradiction

$$\Big(17\zeta^{1/2}n_i n - 5\zeta^{1/2}n - 1\Big) \cdot \frac{7n_i}{8} \cdot \left(\frac{n}{n_i} - 5\zeta^{1/2}n\right) < 14\zeta^{1/2}n_i n^2.$$

Thereby Claim 5.3.29 is proved. ∎

Part II. Choice of a vertex class for v. Our strategy for showing that $\mathcal{H}$ is $\mathcal{G}_i$-colorable is to adjoin v to one the partition classes $V_1, \dots, V_{n_t}$. In fact, there is only one of these classes v fits into. Before finding this class we show a statement that has to hold if our plan is sound.

Claim 5.3.30. *We have $L_{\mathcal{H}}(v) \cap \binom{V_j}{2} = \emptyset$ for every $j \in [n_i]$.*

Proof. Without loss of generality we may assume that $j = 1$. Let $u_0, u_1 \in V_1$ be two distinct vertices. By Lemma 5.3.23 applied to $S = \{u_0, u_1\}$ and $T = [2, n_i]$ there exist vertices $u_j \in V_j$ for $j \in [2, n_i]$ such that the subgraphs of $\mathcal{H}$ induced by $\{u_0, u_2, \dots, u_{n_i}\}$ and $\{u_1, u_2, \dots, u_{n_i}\}$ are isomorphic to $\mathcal{G}_i$. Now Corollary 5.3.26 informs us that the set $U = \{u_0, u_1, \dots, u_{n_t}\}$ cannot be 2-covered, for which reason $u_0u_1 \notin \partial\mathcal{H}$. So, in particular, we have $u_0u_1 \notin L_{\mathcal{H}}(v)$. ∎

Claim 5.3.31. *There exists $j \in [n_i]$ such that $|N_{\mathcal{H}}(v) \cap V_j| < \zeta^{1/7}n$.*

Proof. Suppose for the sake of contradiction that the sets $W_j = N_{\mathcal{H}}(v) \cap V_j$ satisfy $|W_j| \geq \zeta^{1/7}n$ for every $j \in [n_i]$. Applying Lemma 5.3.23 to W_j here in place of V_j there and to $S = \emptyset$, $T = [n_i]$ we obtain vertices $u_j \in V_j$ for all $j \in [n_i]$ such that the set $U = \{u_1, \dots, u_{n_i}\}$ induces a copy of $\mathcal{G}_i$ in $\mathcal{H}$. But now the 2-covered set $U \cup \{v\}$ contradicts Corollary 5.3.26. ∎

It will turn out later that the index j delivered by Claim 5.3.31 is unique. Without loss of generality we may assume that

$$|N_{\mathcal{H}}(v) \cap V_1| < \zeta^{1/7} n. \tag{5.24}$$

Part III. The link of v. It remains to show that $L_{\mathcal{H}}(v) \subseteq L_{\widehat{G}_i}(V_1)$. To this end we define

$$N_v(u) = \Big\{ j \in [n_i] \colon |N_{\mathcal{H}}(u,v) \cap V_j| \geq \zeta^{1/7} n \Big\}$$

for every $u \in N_{\mathcal{H}}(v)$. The upper bound on $\Delta_2(\mathcal{G}_i)$ in Observation 5.3.24 (c) transfers to these sets as follows.

Claim 5.3.32. *We have $|N_v(u)| \leq n_i - k_i$ for every $u \in N_{\mathcal{H}}(v)$.*

Proof. Assume for the sake of contradiction that there is a set $N_\star \subseteq N_v(u)$ such that $|N_\star| = n_i - k_i + 1 < n_i - 2$. As in the proof of Claim 5.3.31 there exist vertices $u_j \in N_{\mathcal{H}}(u,v) \cap V_j$ for $j \in N_\star$ such that $\mathcal{G}_i[N_\star]$ is isomorphic to $\mathcal{H}[U]$, where $U = \{u_j \colon j \in N_\star\}$.

Now we consider the 3-graph $F = \mathcal{H}[U \cup \{u,v\}]$. Clearly $U \cup \{u,v\}$ is 2-covered in F and $\tau(F) \geq \tau(\mathcal{G}_i[N_\star]) \geq 2$. So $F \notin \mathcal{M}_t$ tells us that F is $\mathcal{G}_s$-colorable for some $s \in [t]$.

On the other hand by Lemma 5.3.25 and $|U| \geq n_i - k_i + 2 > n_i - Q/(2k_i^2)$ the subgraph $F[U]$ of F cannot be $\mathcal{G}_s$-colorable for any $s \in [t] \setminus \{i\}$.

Summarizing this discussion, F is $\mathcal{G}_i$-colorable. As F is also 2-covered, F is actually isomorphic to a subgraph of $\mathcal{G}_i$ and, consequently, $n_i - k_i < |N_\star| = d_F(u,v) \leq \Delta_2(\mathcal{G}_i)$, contrary to Observation 5.3.24 (c). ∎

Claim 5.3.33. *We have* $|N_{\mathcal{H}}(v) \cap V_j| \geq \zeta^{1/7} n$ *for every* $j \in [2, n_i]$.

Proof. The minimum degree condition imposed on $\mathcal{H}$ and $6\lambda_t = 1 - \frac{k_i+1}{n_i}$ yield

$$\left(1 - \frac{k_i + 1}{n_i} - 2\zeta\right) n^2 = 2(3\lambda_t - \zeta)n^2 \leq 2d_{\mathcal{H}}(v) \leq \Delta\big(L_{\mathcal{H}}(v)\big)|N_{\mathcal{H}}(v)|.$$

Claim 5.3.32 allows us to bound the first factor on the right side from above by

$$\Delta\big(L_{\mathcal{H}}(v)\big) \leq (n_i - k_i)\left(\frac{n}{n_i} + 5\zeta^{1/2} n\right) + k_i \zeta^{1/7} n < \frac{n_i - k_i}{n_i} n + 2k_i \zeta^{1/7} n.$$

Altogether we obtain

$$\frac{n_i - (k_i + 1) - 2n_i\zeta}{(n_i - k_i) + 2k_i n_i \zeta^{1/7}} \leq \frac{|N_{\mathcal{H}}(v)|}{n},$$

which due to

$$\frac{n_i - (k_i + 1)}{n_i - k_i} = 1 - \frac{1}{n_i - k_i} > 1 - \frac{5/4}{n_i}$$

and $\zeta \ll n_i^{-1}$ implies

$$\left(1 - \frac{3/2}{n_i}\right) n \leq |N_{\mathcal{H}}(v)|.$$

On the other hand, setting $I = \big\{j \in [2, n_i] \colon |N_{\mathcal{H}}(v) \cap V_j| \geq \zeta^{1/7} n\big\}$ Claim 5.3.27 and Equation 5.24 lead to

$$|N_{\mathcal{H}}(v)| \leq |I| \left(\frac{1}{n_i} + 5\zeta^{1/2}\right) n + \zeta^{1/7} n_i n.$$

Combining both estimates we arrive at $|I| > n_i - 7/4$, whence $I = [2, n_i]$. ∎

Claim 5.3.34. *We have $N_{\mathcal{H}}(v) \cap V_1 = \emptyset$.*

Proof. Suppose that there exists $u_1 \in N_{\mathcal{H}}(v) \cap V_1$. Owing Claim 5.3.33 we can apply Lemma 5.3.23 with $S = \{u_1\}$ and $T = [2, n_i]$ in order to obtain vertices $u_j \in N_{\mathcal{H}}(v) \cap V_j$ for $j \in [2, n_i]$ such that $\mathcal{H}$ induces a copy of $\mathcal{G}_i$ on $U = \{u_1, \dots, u_{n_i}\}$. Since $U \cup \{v\}$ is 2-covered, this contradicts Corollary 5.3.26. ∎

Let us recall that $L_{\widehat{\mathcal{G}}_i}(V_1)$ denotes the common $\widehat{\mathcal{G}}_i$-link of all vertices in V_1.

Claim 5.3.35. *We have $L_{\mathcal{H}}(v) \subseteq L_{\widehat{\mathcal{G}}_i}(V_1)$.*

Proof. Due to the Claims 5.3.30 and 5.3.34 we know that $L_{\mathcal{H}}(v)$ is an $(n_i - 1)$-partite graph with vertex classes $V_2, \dots, V_{n_i}$. So if Claim 5.3.35 fails we may assume without loss of generality $123 \notin \mathcal{G}_i$ and that there exists a pair $u_2u_3 \in L_{\mathcal{H}}(v)$ with $u_2 \in V_2$, $u_3 \in V_3$.

Since $|V_1| > n/(2n_i)$ and $|N_{\mathcal{H}}(v) \cap V_j| \geq \zeta^{1/7}n$ for $j \in [4, n_i]$, Lemma 5.3.23 applied to $S = \{u_2, u_3\}$ and $T = \{1, 4, \dots, n_i\}$ delivers vertices $u_1 \in V_1$ and $u_j \in N_{\mathcal{H}}(v) \cap V_j$ for $j \in [4, n_i]$ such that the set $U' = \{u_1, u_4, \dots, u_{n_i}\}$ satisfies

$$\mathcal{H}[U'] = \widehat{\mathcal{G}}_i[U'] \quad \text{and} \quad L_{\mathcal{H}}(u_\ell)[U'] = L_{\widehat{\mathcal{G}}_i}(u_\ell)[U'] \quad \text{for} \quad \ell = 2, 3. \tag{5.25}$$

Consider the set $U = \{u_1, \dots, u_{n_i}\}$. Because of Equation 5.25 and $123 \notin \mathcal{G}_i$ the map $i \longmapsto u_i$ is an embedding of $L_{\mathcal{G}_i}(1)$ into $\mathcal{H}$ and for this reason we have

$$d_{\mathcal{H}[U]}(u_1) \geq d_{\mathcal{G}_i}(1). \tag{5.26}$$

Next we choose for every $j \in [4, n_i]$ an edge $e_j \in \mathcal{H}$ such that $u_j, v \in e_j$ and observe that U is 2-covered in the 3-graph

$$F = \{vu_2u_3\} \cup \{e_j \colon 4 \leq j \leq n_i\} \cup \mathcal{H}[U].$$

Moreover, $|F| \leq |\mathcal{G}_i| + n_i - 2 < \binom{n_i}{3}$ implies $F \in \widehat{\mathcal{K}}^3_{n_i}$. Since $F[U'] = \mathcal{H}[U']$ is isomorphic to $\mathcal{G}_i - \{2, 3\}$, Lemma 5.3.25 tells us that F cannot be $\mathcal{G}_j$-colorable for any $j \in [t] \setminus \{i\}$. But on the other hand we have $\tau(F[U]) \geq 2$ and $F \notin \mathcal{M}_t$, so altogether F is $\mathcal{G}_i$-colorable.

Fix a homomorphism $\phi\colon V(F) \longrightarrow V(\mathcal{G}_i)$ from F to $\mathcal{G}_i$. Since U and $U_v = U \cup \{v\} \setminus \{u_1\}$ are 2-covered subsets of F whose size is $n_i = v(\mathcal{G}_i)$, the map ϕ has to be bijective on U and U_v, which is only possible if $\phi(v) = \phi(u_1)$. Now ϕ embeds the link $L_{F[U]}(u_1)$ into the link $L_{\mathcal{G}_i}(\phi(u_1))$. Moreover, $vu_2u_3 \in F$ implies that $\phi(u_2)\phi(u_3)$ belongs to the link $L_{\mathcal{G}_i}(\phi(u_1))$ as well and by $123 \notin \mathcal{G}_i$ this edge is not in the image $\phi(L_{F[U]}(u_1))$. Altogether this proves $d_{F[U]}(u_1) + 1 \leq d_{\mathcal{G}_i}(\phi(u_1))$, which in view of $F[U] = \mathcal{H}[U]$ and Equation 5.26 contradicts the regularity of $\mathcal{G}_i$. ∎

By Claim 5.3.35 the partition $\bigcup_{j \in [n_i]} \widehat{V}_j$, where

$$\widehat{V}_j = \begin{cases} V_1 \cup \{v\} & \text{if } j = 1 \\ V_j & \text{if } 2 \leq j \leq n_i \end{cases}$$

is a $\mathcal{G}_i$-coloring of $\mathcal{H}$. This completes the proof of Lemma 5.3.21.

5.3.4 Feasible region of $\mathcal{M}_t$ and $\xi(\mathcal{M}_t)$

We prove Theorem 5.1.10 and that $\xi(\mathcal{M}_t) = t$ in this section. First, let us show a simple lemma.

Lemma 5.3.36. *Suppose that $\mathcal{H}$ is an n-vertex $\mathcal{G}_i$-colorable 3-graph for some $i \in [t]$. If $|\mathcal{H}| \geq (\lambda_t - \epsilon)n^3$, then $|\partial\mathcal{H}| \geq \left(\frac{n_i-1}{2n_i} - 3\epsilon^{1/2}n_i\right)n^2$.*

Proof. Let $V(\mathcal{H}) = \bigcup_{j\in[n_i]} V_j$ be a $\mathcal{G}_i$-coloring of $\mathcal{H}$. Now by Proposition 5.4.2, $|V_j| = (1/n_i \pm 3\epsilon^{1/2})n$ for all $j \in [n_i]$. Call a pair $\{u, v\}$ with $u \in V_j, v \in V_k$ and $j \neq k$ missing if $uv \notin \partial\mathcal{H}$, and let M denote the set of all missing pairs. Since $\delta_2(\mathcal{G}_i) \geq 7n_i/8$, we obtain

$$|M| \cdot \frac{7n_i}{8} \cdot \left(\frac{1}{n_i} - 3\epsilon^{1/2}\right) n \leq 3\epsilon n^3,$$

which yields $|M| < 4\epsilon n^2$. Therefore,

$$|\partial\mathcal{H}| > \binom{n_i}{2} \times \left(\frac{1}{n_i} - 3\epsilon^{1/2}\right)^2 n^2 - |M| > \frac{n_i - 1}{2n_i} n^2 - 3\epsilon^{1/2} n_i n^2.$$

∎

We remark that the stronger conclusion $|\partial\mathcal{H}| \geq \left(\frac{n_i-1}{2n_i} - 5\epsilon n_i\right)n^2$ could be shown by arguing more carefully, but this is immaterial to what follows.

Proof of Theorem 5.1.10. Recall from Section 5.3.2 that semibipartite 3-graphs are $\mathcal{M}_t$-free. This yields $\text{proj}\Omega(\mathcal{M}_t) = [0, 1]$, as for every $x \in [0, 1]$ there exists a good sequence of semibipartite 3-graphs such that the edge densities of their shadows converges to x.

Theorem 5.1.9 (a) implies that $g(\mathcal{M}_t, x) \leq 6\lambda_t$ for all $x \in [0,1]$. Furthermore for every $i \in [t]$ the sequence of balanced blow-ups of $\mathcal{G}_i$ shows the equality $g(\mathcal{M}_t, 1-1/n_i) = 6\lambda_t$. So, in order to finish the proof it suffices to show that if some $x \in [0,1]$ satisfies $g(\mathcal{M}_t, x) = 6\lambda_t$, then there is an index $i \in [t]$ such that $x = 1 - 1/n_i$.

Fix such an $x \in [0,1]$ and let $(\mathcal{H}_n)_{n=1}^{\infty}$ be a good sequence of $\mathcal{M}_t$-free 3-graphs realizing $(x, 6\lambda_t)$. Consider an arbitrary $\delta > 0$ and let $\epsilon > 0, N_0$ be the constants guaranteed by Theorem 5.1.9 (b). Without loss of generality we may assume $\epsilon \leq \delta$. By our choice of $(\mathcal{H}_n)_{n=1}^{\infty}$ there exists $n_0 \in \mathbb{N}$ such that

$$d(\mathcal{H}_n) = 6\lambda_t \pm \epsilon \quad \text{and} \quad d(\partial\mathcal{H}_n) = x \pm \epsilon$$

hold for all $n \geq n_0$. By Theorem 5.1.9 (b), for every $n \geq \max\{n_0, N_0\}$ the 3-graph $\mathcal{H}_n$ is $\mathcal{G}_i$-colorable for some $i = i(n) \in [t]$ after removing at most $\delta v(\mathcal{H}_n)$ vertices. Therefore,

$$|\partial\mathcal{H}_n| \leq \left(\frac{n_i - 1}{2n_i} + \delta\right) v(\mathcal{H}_n)^2,$$

and, on the other hand, by Lemma 5.3.36,

$$|\partial\mathcal{H}_n| > \left(\frac{n_i - 1}{2n_i} - 3\epsilon^{1/2} n_i\right)(1-\delta)^2 v(\mathcal{H}_n)^2 > \frac{n_i - 1}{2n_i} v(\mathcal{H}_n)^2 - \left(3\epsilon^{1/2} n_i + 2\delta\right) v(\mathcal{H}_n)^2.$$

Summarizing and taking $\epsilon \leq \delta$ into account we arrive at

$$\frac{n_i - 1}{n_i} - \left(6\delta^{1/2} n_t + 4\delta\right) < d(\partial\mathcal{H}_n) \leq \frac{n_i - 1}{n_i} + 2\delta, \tag{5.27}$$

where, let us recall, $i = i(n)$ might depend on n. So what (Equation 5.27) means is that if we set

$$I_i(\delta) = \left[\frac{n_i - 1}{n_i} - 6\delta^{1/2} n_i - 4\delta, \frac{n_i - 1}{n_i} + 2\delta\right]$$

for every $i \in [t]$, then

$$d(\partial\mathcal{H}_n) \in I_1(\delta) \cup \cdots \cup I_t(\delta)$$

holds for every $n \geq n_0$. As the set on the right side is closed we obtain

$$x \in I_1(\delta) \cup \cdots \cup I_t(\delta)$$

in the limit $n \to \infty$. Since $\delta > 0$ was arbitrary,

$$x \in \bigcap_{\delta > 0} \big(I_1(\delta) \cup \cdots \cup I_t(\delta)\big) = \big\{1 - 1/n_i \colon i \in [t]\big\}$$

follows. ∎

Recall that we already proved that $\mathcal{M}_t$ is t-stable, which, by definition, shows that $\xi(\mathcal{M}_t) \leq t$. Therefore, in order to prove $\xi(\mathcal{M}_t) = t$ it suffices to show that $\xi(\mathcal{M}_t) \geq t$, and this is an easy consequence of the following proposition and Theorem 5.1.10.

Proposition 5.3.37. *Let $\mathcal{F}$ be a family of r-graphs and let M be the set of global maxima of $g(\mathcal{F})$. If M is finite, then $|M| \le \xi(\mathcal{F})$.*

Proof. If $\mathcal{F}$ is degenerate, then $g(\mathcal{F})$ is the constant function whose value is always 0 and M is infinite. So we may assume that the Turán density $y = \pi(\mathcal{F})$ is positive. Let us write $M = \{(x_i, y) \colon i \in [m]\}$ such that $x_1 < \cdots < x_m$ and $m = |M|$. For every $i \in [m]$ we select a good sequence $(\mathcal{H}_i(n))_{n=1}^{\infty}$ of $\mathcal{F}$-free r-graphs realizing (x_i, y). Without loss of generality we have $v(\mathcal{H}_i(n)) = n$ for every positive integer n. Now suppose for the sake of contradiction that $t = \xi(\mathcal{F})$ is smaller than m.

Claim 5.3.38. *For every $\delta > 0$ there are distinct $i, j \in [m]$ and $n > 1/\delta$ such that*

$$d_1(\mathcal{H}_i(n), \mathcal{H}_j(n)) \le \delta n^r \quad \text{and} \quad \min\{|\mathcal{H}_i(n)|, |\mathcal{H}_j(n)|\} \ge (y - \delta)\binom{n}{r}.$$

Proof. By the definition of $\xi(\mathcal{F}) = t$ there are $n_0 \in \mathbb{N}$ and $\epsilon > 0$ such that for every $n \ge n_0$ there exists a family $\{\mathcal{G}_1(n), \ldots, \mathcal{G}_t(n)\}$ of r-graphs on n vertices such that for every $\mathcal{F}$-free r-graph $\mathcal{H}$ with $v(\mathcal{H}) = n$ and $|\mathcal{H}| \ge (y - \epsilon)\binom{n}{r}$ there is some $s \in [t]$ such that $d_1(\mathcal{H}, \mathcal{G}_s(n)) \le (\delta/2)n^r$ As usual, we may suppose that $\epsilon \le \delta$.

Now choose $n \ge n_0, \delta^{-1}$ such that for every $i \in [m]$ we have $d(\mathcal{H}_i(n)) \ge y - \epsilon$. Stability allows us to select for every $i \in [m]$ an index $s(i) \in [t]$ such that $d_1(\mathcal{H}_i(n), \mathcal{G}_s(n)) \le \delta n^r$. By

$t < m$ the map $i \longmapsto s(i)$ cannot be injective, i.e., there are distinct $i, j \in [m]$ and $s \in [t]$ such that $s(i) = s(j) = s$. Now the triangle inequality yields

$$d_1(\mathcal{H}_i(n), \mathcal{H}_j(n)) \leq d_1(\mathcal{H}_i(n), \mathcal{G}_s(n)) + d_1(\mathcal{G}_s(n), \mathcal{H}_j(n)) \leq \delta n^r,$$

as desired. ∎

Notice that, as stated, Claim 5.3.38 allows i and j to depend on δ. However, a quick thought reveals that there actually have to be two indices $i < j$ that work for every $\delta > 0$. Now we intend to contradict the finiteness of M by proving $[x_i, x_j] \times \{y\} \subseteq M$.

To this end, let $x \in [x_i, x_j]$ and a large integer N be given. It suffices to construct an $\mathcal{F}$-free r-graph $\mathcal{H}$ satisfying $v(\mathcal{H}) > N$, $d(\partial\mathcal{H}) = x \pm 1/N$ and $d(\mathcal{H}) = y \pm 1/N$. By Claim 5.3.38 applied to $\delta \ll N^{-1}$ there is some $n > N$ such that $d_1(\mathcal{H}_i(n), \mathcal{H}_j(n)) \leq \delta n^r$ and $\min\{|\mathcal{H}_i(n)|, |\mathcal{H}_j(n)|\} \geq (y - \delta)\binom{n}{r}$. Assume without loss of generality that

$$|\mathcal{H}_i(n) \triangle \mathcal{H}_j(n)| \leq \delta n^r.$$

Now consider the following process transforming $\mathcal{H}_i(n)$ into $\mathcal{H}_i(n)$: Start with $\mathcal{H}_i(n)$ and remove edges one by one until $\mathcal{H}_i(n) \cap \mathcal{H}_j(n)$ is reached. Then, keep adding edges one by one until you arrive at $\mathcal{H}_j(n)$. Every r-graph occurring along the way is $\mathcal{F}$-free. Moreover, since deleting or adding an edge can affect the size of the shadow by at most r, in every step of the process the shadow density changer by at most $r/\binom{n}{r-1}$. Thus at some moment we pass an r-graph

$\mathcal{H}$ such that $|d(\partial\mathcal{H}) - x| \le r/\binom{n}{r-1} \le \delta$. Finally, $d(\mathcal{H}) \ge d(\mathcal{H}_i(n) \cap \mathcal{H}_j(n)) \ge d(\mathcal{H}_i(n)) - |\mathcal{H}_i(n) \triangle \mathcal{H}_j(n)|/\binom{n}{r} \ge y - O(\delta)$ completes the proof that $\mathcal{H}$ has all desired properties. ∎

5.3.5 Concluding remarks

For every positive integer t we constructed a family of 3-graphs $\{\mathcal{G}_1, \dots, \mathcal{G}_t\}$ that have the same Lagrangian λ_t, and we showed that there is a family $\mathcal{M}_t$ of 3-graphs whose extremal configurations are balanced blow-ups of $\mathcal{G}_1, \dots, \mathcal{G}_t$, and whose stability number is $\xi(\mathcal{M}_t) = t$. Notice that our choice of λ_t is very close to $1/6$, which is the supremum of the Lagrangians of all 3-graphs. It would be interesting to find for every integer $t \ge 2$ the minimum value (if it exists) of $\lambda = \lambda(t)$ so that there exists a t-stable family $\mathcal{F}_t$ with $\pi(\mathcal{F}_t) = 6\lambda$. A result of Erdős [58] implies that there are no Turán densities in the interval $(0, 2/9)$. This motivates the following question.

Problem 5.3.39. *Does there exist a family $\mathcal{F}$ of triple systems with $\pi(\mathcal{F}) = 2/9$ but $\xi(\mathcal{F}) \ne 1$?*

For a family $\mathcal{F}$ of r-graphs let $M(\mathcal{F}) = \{x \in \mathrm{proj}\Omega(\mathcal{F})\colon g(\mathcal{F})(x) = \pi(\mathcal{F}))$ be the set of abscissae of the global maxima of its feasible region function. As we have shown here, $|M(\mathcal{F})|$ can be every finite cardinal except zero. In would be interesting to know whether $M(\mathcal{F})$ can be infinite and, in case the answer is affirmative, there immediately arise further questions.

Problem 5.3.40. *For $r \ge 3$ does there exist a non-degenerate family $\mathcal{F}$ of r-graphs so that $g(\mathcal{F})$ has infinitely many global maxima? If so, can the set $M(\mathcal{F})$ be uncountable? Can it even contain a non-trivial interval?*

Notice that if the last question on intervals has a negative answer, then in Proposition 5.3.37 the assumption that M should be finite can be omitted. In fact, it is somewhat bizarre that we do not know the following.

Problem 5.3.41. *Let $\mathcal{F}$ be a non-degenerate family of r-graphs such that $M(\mathcal{F})$ is infinite. Can it nevertheless happen that $\mathcal{F}$ has finite stability number?*

5.4 Proof for t-stable families of r-graphs

In this section we prove Theorems 5.1.11 and 5.1.12.

5.4.1 Preliminaries

Denote by $M(\mathcal{F})$ the set of global maxima of $g(\mathcal{F})$ and $m(\mathcal{F}) = |M(\mathcal{F})|$.

Proposition 5.4.1. *Let $n_0 > 0, c > 0$ be constants and $\mathcal{F}$ be a family of r-graphs with $\xi_v(\mathcal{F}) = t$. Suppose that $\mathcal{F}$ is vertex-t-stable respects to $\mathcal{G}_1(n), \ldots, \mathcal{G}_t(n)$, and $\mathcal{G}_i(n)$ satisfies $\delta_{r-1}(\mathcal{G}_i(n)) \geq cn$ for every $n \geq n_0$ and $i \in [t]$. Then $m(\mathcal{F}) \leq t$.*

Proof. By assumption we have $\lim_{n\to\infty} d\left(\mathcal{G}_i(n)\right) = \pi(\mathcal{F})$ for $i \in [t]$. Also we may assume that $\lim_{n\to\infty} d\left(\partial\mathcal{G}_i(n)\right) = x_i$ for $i \in [t]$ since otherwise we can take a convergence subsequence. Without loss of generality we may assume that $x_1 \leq \cdots \leq x_t$.

We claim that $M(\mathcal{F}) = \{(x_i, \pi(\mathcal{F})) : i \in [t]\}$. Indeed, suppose this is not true and there exists $x_0 \in \mathrm{proj}\Omega(\mathcal{F})$ with $x_0 \neq x_i$ for $i \in [t]$ such that $(x_0, \pi(\mathcal{F})) \in M(\mathcal{F})$. Then let $(\mathcal{H}(n))_{n=1}^{\infty}$ be an $\mathcal{F}$-free good sequence that realizes $(x_0, \pi(\mathcal{F}))$. Without loss of generality we may assume that $v(\mathcal{H}(n)) = n$ for $n \geq 1$. For every $\delta > 0$ there exist $n(\delta)$ such that for every $n \geq n(\delta)$ there exists a set $Z_n \subset V(\mathcal{H})$ of size at most δn so that the r-graph $\mathcal{H}'(n) = \mathcal{H}(n) - Z_n$ is a subgraph of $\mathcal{G}_i(n)$ for some $i \in [t]$. So,

$$|\partial\mathcal{H}(n)| \leq |\partial\mathcal{H}'(n)| + |Z_n|\binom{n}{r-2} \leq |\partial\mathcal{G}_i(n)| + |Z_n|\binom{n}{r-2} \leq |\partial\mathcal{G}_i(n)| + \delta n\binom{n}{r-2},$$

and by assumption

$$|\partial\mathcal{H}(n)| \geq |\partial\mathcal{H}'(n)| \geq |\partial\mathcal{G}_i(n)| - \frac{r\left(|\mathcal{G}_i(n)| - |\mathcal{H}'(n)|\right)}{cn} \geq |\partial\mathcal{G}_i(n)| - \frac{r\left(\delta n^r + o(n^r)\right)}{cn}$$
$$\geq |\partial\mathcal{G}_i(n)| - r\left(\delta n^{r-1} + o(n^{r-1})\right).$$

Letting $\delta \to 0$ we obtain $x_0 = \lim_{n\to\infty} d\left(\partial\mathcal{H}(n)\right) = \lim_{n\to\infty} d\left(\partial\mathcal{G}_i(n)\right) = x_i$, a contradiction. ∎

For an r-graph $\mathcal{H}$ and a set $S \subset V(\mathcal{H})$ of size $s < r$ the neighborhood of S is

$$N_{\mathcal{H}}(S) = \left\{u \in V(\mathcal{H}) \setminus S \colon \exists A \in \mathcal{H} \text{ such that } \{u\} \cup S \subset A\right\},$$

and the link of S is an $(r-s)$-graph on $N_{\mathcal{H}}(S)$ which is defined as

$$L_{\mathcal{H}}(S) = \left\{A \in \partial_{r-s}\mathcal{H} \colon A \cup S \in \mathcal{H}\right\}.$$

The degree of S is $d_{\mathcal{H}}(S) = |L_{\mathcal{H}}(S)|$, and denote by $\delta_s(\mathcal{H}), \Delta_s(\mathcal{H})$ the minimum and maximum s-degree of $\mathcal{H}$, respectively. In other words,

$$\delta_s(\mathcal{H}) = \min\left\{d_{\mathcal{H}}(S) \colon S \in \binom{V(\mathcal{H})}{s}\right\} \quad \text{and} \quad \Delta_s(\mathcal{H}) = \max\left\{d_{\mathcal{H}}(S) \colon S \in \binom{V(\mathcal{H})}{s}\right\}.$$

Recall that a k-graph $\mathcal{D}$ is an (n,k)-design if it has n vertices and every pair of vertices is covered by a unique edge. With every such design $\mathcal{D}$ we associate the 3-graph $H(\mathcal{D})$ on $V(\mathcal{D})$, where

$$H(\mathcal{D}) = \bigcup_{E\in\mathcal{D}} \binom{E}{3}.$$

The following result determines the Lagrangian of a 3-graph obtained from a complete 3-graph by removing edges in $H(\mathcal{D})$ and a sparse regular 3-graph that is disjoint from $H(\mathcal{D})$.

Proposition 5.4.2. *Suppose that $n \geq 18k + 3^7 s^3$, $\mathcal{D}$ is an (n,k)-design on $[n]$, and $\mathcal{S}$ is an s-regular 3-graph on $[n]$. If $\mathcal{S} \cap H(\mathcal{D}) = \emptyset$ and $\mathcal{G} = K_n^3 \setminus (H(\mathcal{D}) \cup \mathcal{S})$, then*

$$L_{\mathcal{G}}(x_1,\ldots,x_n) + \frac{1}{9}\sum_{i=1}^{n}\left(x_i - \frac{1}{n}\right)^2 \leq \frac{|\mathcal{G}|}{n^3} = \frac{1}{6}\left(1 - \frac{k+1}{n} + \frac{k-2s}{n^2}\right) \tag{5.28}$$

holds for all $(x_1,\ldots,x_n) \in \Delta_{n-1}$.

For every 3-graph $\mathcal{G}$ we define an r-graph $\mathcal{G}^r$ be with vertex set $V(\mathcal{G}) \cup C$, where C (called the center of $\mathcal{G}^r$) is a set of size $r-3$ disjoint from $V(\mathcal{G})$, and

$$\mathcal{G}^r = \{C \cup E : E \in \mathcal{G}\}.$$

Lemma 5.4.3. *Let $\mathcal{G}$ be a 3-graph on n vertices that satisfies the conditions in Proposition 5.4.2. Then the followings hold.*

(a) $\lambda(\mathcal{G}^r) = 27\lambda(\mathcal{G})/r^r$, and

(b) for every $\epsilon \geq 0$ *and* $(x_1, \ldots, x_{\hat{n}}) \in \Delta_{\hat{n}}$*, where* $\hat{n} = n+r-3$*, if* $L_{\mathcal{G}^r}(x_1, \ldots, x_{\hat{n}}) \geq \lambda(\mathcal{G}^r) - \epsilon$*, then*

$$x_j = \begin{cases} 1/r \pm r^{r/2-1}\epsilon^{1/2}, & j \in [r-3], \\ 3/rn \pm 2r^{r/2+1}\epsilon^{1/2}, & j \in [r-2, \hat{n}]. \end{cases}$$

Proof. Let $\hat{r} = r-3$, $\lambda = \lambda(\mathcal{G})$, and $\lambda^{(r)} = \lambda(\mathcal{G}^r)$. Let $(x_1, \ldots, x_{\hat{n}}) \in \Delta_{\hat{n}-1}$ and $x = \sum_{j \in [\hat{r}]} x_j$. Then

$$L_{\mathcal{G}^r}(x_1, \ldots, x_{\hat{n}}) = L_{\mathcal{G}}(x_{\hat{r}+1}, \ldots, x_{\hat{n}}) \prod_{j \in [\hat{r}]} x_j \leq \lambda(1-x)^3 \left(\frac{x}{r-3}\right)^{r-3} \leq \frac{27\lambda}{r^r},$$

and equality holds if

$$x_1 = \cdots = x_{\hat{r}} = 1/r \quad \text{and} \quad x_{\hat{r}+1} = \cdots = x_{\hat{n}} = 3/rn.$$

Therefore, $\lambda(\mathcal{G}^r) = 27\lambda/r^r$.

Now we prove (b). Suppose for the contrary that there exists $j \in [\hat{r}]$ such that $|x_j - 1/r| > \epsilon_1$, where $\epsilon_1 = r^{r/2-1}\epsilon^{1/2}$, and without loss of generality we may assume that $j = 1$. Then

$$\begin{aligned} L_{\mathcal{G}^r}(x_1, \ldots, x_{\hat{n}}) = L_{\mathcal{G}}(x_{\hat{r}+1}, \ldots, x_{\hat{n}}) \prod_{j \in [\hat{r}]} x_j &\leq \lambda(1-x)^3 x_1 \prod_{2 \leq j \leq \hat{r}} x_j \\ &\leq \lambda(1-x)^3 x_1 \frac{(x-x_1)^{r-4}}{(r-4)^{r-4}} \end{aligned}$$

Viewing $(1-x)^3(x-x_1)^{r-4}$ as a function in x we obtain

$$(1-x)^3(x-x_1)^{r-4} \leq \frac{27(r-4)^{r-4}}{(r-1)^{r-1}}(1-x_1)^{r-1}.$$

Therefore,

$$\begin{aligned}
L_{\mathcal{G}^r}(x_1,\dots,x_{\hat{n}}) &\leq \frac{27\lambda}{(r-1)^{r-1}} x_1(1-x_1)^{r-1} \\
&< \frac{27\lambda}{(r-1)^{r-1}}\left(\frac{1}{r}-\epsilon_1\right)\left(\frac{r-1}{r}+\epsilon_1\right)^{r-1} \\
&< \frac{27\lambda}{(r-1)^{r-1}}\frac{1}{r}(1-r\epsilon_1)\times\left(\frac{r-1}{r}\right)^{r-1}(1+r\epsilon_1) \\
&< \frac{27\lambda}{r^r}(1-r^2\epsilon_1^2) = \lambda^{(r)}-\epsilon,
\end{aligned}$$

a contradiction.

Now, suppose that $|x_j - 3/rn| > \epsilon_2$ for some $j \in [\hat{r}+1,\hat{n}]$, where $\epsilon_2 = 2r^{r/2+1}\epsilon^{1/2}$. Without loss of generality we may assume that $j = \hat{r}+1$. Then, by Equation 5.28,

$$\begin{aligned}
L_{\mathcal{G}^r}(x_1,\dots,x_{\hat{n}}) &= L_{\mathcal{G}}(x_{\hat{r}+1},\dots,x_{\hat{n}})\prod_{j\in[\hat{r}]} x_j \\
&\leq \frac{x^{r-3}}{(r-3)^{r-3}}(1-x)^3\left(\lambda-\frac{1}{9}\sum_{j\in[\hat{r}+1,\hat{n}]}\left(\frac{x_j}{1-x}-\frac{1}{n}\right)^2\right) \\
&\leq \frac{27}{r^r}\left(\lambda-\frac{1}{9}\left(\frac{x_1}{1-x}-\frac{1}{n}\right)^2\right) = \lambda^{(r)}-\frac{3}{r^r}\left(\frac{x_1}{1-x}-\frac{1}{n}\right)^2.
\end{aligned}$$

It follows from $1 - x = 1 - \sum_{j\in[\hat{r}]} x_j = 3/r \pm r\epsilon_1$ and $|x_1 - 3/rn| > \epsilon_2$ that

$$\left|\frac{x_1}{1-x} - \frac{1}{n}\right| > \epsilon_2 - r^2\epsilon_1 > \frac{\epsilon_2}{2}.$$

Therefore,

$$L_{\mathcal{G}^r}(x_1, \ldots, x_{\hat{n}}) \leq \lambda^{(r)} - \frac{3}{r^r}\left(\frac{\epsilon_2}{2}\right)^2 < \lambda^{(r)} - \epsilon,$$

a contradiction. ∎

Let $r \geq c \geq 0$ be integers. An r-graph $\mathcal{G}$ is called a c-star if there exists a set C (called the center of $\mathcal{G}$) in $V(\mathcal{G})$ of size c such that $C \subset E$ for all $E \in \mathcal{G}$. Notice that the r-graph $\mathcal{G}^r$ defined above is an $(r-3)$-star, and is not an $(r-2)$-star if $\mathcal{G}$ contains two disjoint egdes.

Lemma 5.4.4. *Suppose that the r-graph $\mathcal{G}$ is a c-star. Then $\lambda(\mathcal{G}) \leq \frac{(r-c)^{r-c}}{(r-c)!}\frac{1}{r^r}$. In particular, if $c = r-2$, then $\lambda(\mathcal{G}) \leq 2/r^r$.*

Proof. Let us assume that the number of vertices of $\mathcal{G}$ is $\hat{n} = n + c$ and $\mathcal{G}$ is a maximal c-star, i.e. all sets in $\mathcal{G}$ contain the set $[c]$. Let $x = x_1 + \cdots + x_c$. Then

$$\begin{aligned} L_{\mathcal{G}}(x_1, \ldots, x_{\hat{n}}) &= \prod_{i\in[c]} x_i \sum_{E\in\binom{[c+1,\hat{n}]}{r-c}} \prod_{i\in E} x_i \\ &\leq \left(\frac{\sum_{i\in[c]} x_i}{c}\right)^c \binom{n}{r-c}\frac{1}{n^{r-c}}\left(\sum_{i\in[c+1,\hat{n}]} x_i\right)^{r-c} \\ &< \frac{x^c(1-x)^{r-c}}{(r-c)!c^c} \leq \frac{1}{(r-c)!c^c}\frac{(r-c)^{r-c}c^c}{r^r} = \frac{(r-c)^{r-c}}{(r-c)!}\frac{1}{r^r}. \end{aligned}$$

This completes the proof of Lemma 5.4.4 ∎

5.4.2 Extremal configurations and the forbidden family

5.4.2.1 Definition

Recall that $3 < k_1 < \cdots < k_t$, $n_1 < \cdots < n_t$ are integers satisfying $3 \mid n_i$, $(k_i - 1) \mid (n_i - 1)$, and $(k_i - 1)k_i \mid (n_i - 1)n_i$ for $i \in [t]$, and

$$\frac{n_1}{k_1 + 1} = \cdots = \frac{n_t}{k_t + 1} = Q,$$

where Q is a sufficiently large constant.

For $i \in [t]$ the 3-graph $\mathcal{G}_i$ has n_i vertices and

$$\mathcal{G}_i = \binom{V(\mathcal{G}_i)}{3} \setminus (H(\mathcal{D}_i) \cup \mathcal{S}_i)\,,$$

where $\mathcal{D}_i$ is an (n_i, k_i)-design on $V(\mathcal{G}_i)$, and $\mathcal{S}_i$ is a $k_i/2$-regular 3-graph on $V(\mathcal{G}_i)$ such that $\mathcal{S}_i \cap H(\mathcal{D}_i) = \emptyset$. Also, recall that

$$\lambda_t = \lambda(\mathcal{G}_1) = \cdots = \lambda(\mathcal{G}_t) = \frac{1}{6}\left(1 - \frac{k_i + 1}{n_i}\right) = \frac{1}{6}\left(1 - \frac{1}{Q}\right).$$

The family $\mathcal{M}_t$ is the union of the following two families:

(a) $\mathcal{M}_{t,1}$ contains every $F \in \bigcup_{\ell \le n_t} \widehat{\mathcal{K}}_\ell^3$ whose induced subgraph on its core has transversal number at least two, and F does not occur as a subgraph in any blow-up of $\mathcal{G}_1, \ldots, \mathcal{G}_t$.

(b) $\mathcal{M}_{t,2}$ contains every 3-graph in $\mathcal{K}^3_{n_t+1}$ whose induced subgraph on its core has transversal number at least two.

Our extremal configurations for r-graphs are balanced blow-ups of $\mathcal{G}^r_1, \ldots, \mathcal{G}^r_t$, and the forbidden family $\mathcal{M}^r_t$ is the union of the following two families:

(a) $\mathcal{M}^r_{t,1}$ contains every $F \in \bigcup_{\ell \leq \hat{n}_t} \widehat{\mathcal{K}}^r_\ell$ that does not occur as a subgraph in any blow-up of $\mathcal{G}^r_i$ for $i \in [t]$, and the induced subgraph of F on its core is nonempty and is not an $(\hat{r}+1)$-star.

(b) $\mathcal{M}^r_{t,2}$ contains every $F \in \mathcal{K}^r_{\hat{n}_t+1}$ whose induced subgraph on its core is nonempty and is not an $(\hat{r}+1)$-star.

Let us identify the vertex set of $\mathcal{G}^r_i$ with $[\hat{n}_i]$, where the set $[\hat{r}]$ is the center of $\mathcal{G}^r_i$. It follows from Lemma 5.4.3 (a) that,

$$\lambda(\mathcal{G}^r_1) = \cdots = \lambda(\mathcal{G}^r_t) = \lambda^{(r)}_t = 27\lambda_t/r^r.$$

Notice that a 3-graph $\mathcal{G}$ that is nonempty and is not a 1-star iff $\tau(\mathcal{G}) \geq 2$, so $\mathcal{M}^3_t = \mathcal{M}_t$. Also, note that our definition of $\mathcal{M}^r_t$ ensures that every $(\hat{r}+1)$-star is $\mathcal{M}^r_t$-free.

5.4.3 Turán number of $\mathcal{M}^r_t$

We determine the Turán number of $\mathcal{M}^r_t$ in this section.

Lemma 5.4.5. *An r-graph $\mathcal{H}$ is $\mathcal{M}^r_t$-hom-free iff it is $\mathcal{M}^r_t$-free, i.e. $\mathcal{M}^r_t$ is blowup-invariant.*

Proof. It is clear that $\mathcal{M}_t^r$-hom-free implies $\mathcal{M}_t^r$-free. So, it suffices to show that $\mathcal{M}_t^r$-free implies $\mathcal{M}_t^r$-hom-free.

Suppose that $\mathcal{H}$ is $\mathcal{M}_t^r$-free. In order to show that $\mathcal{H}$ is $\mathcal{M}_t^r$-hom-free it suffices to show that every blow-up of $\mathcal{H}$ is $\mathcal{M}_t^r$-free. Since every blow-up of $\mathcal{H}$ can be obtained from $\mathcal{H}$ by duplicating vertices one by one, it suffices to show that duplicating a vertex of $\mathcal{H}$ keeps the $\mathcal{M}_t^r$-free.

Suppose for the contrary that there exists $v \in V(\mathcal{H})$ so that the r-graph $\mathcal{H}'$ obtained from $\mathcal{H}$ by duplicating v is not $\mathcal{M}_t^r$-free. Let F be a member in $\mathcal{M}_t^r$ that occurs as a subgraph in $\mathcal{H}'$. Then $V(F)$ must contain both of v and the clone $\hat{v}$ of v. Note that $\{v, \hat{v}\}$ is not contained in any edge of $\mathcal{H}'$. Let C denote the core of F and without loss of generality we may assume that $\hat{v} \notin C$ since otherwise we may replace $\hat{v}$ by v. Let $F' = F - \hat{v}$ and note that $F' \subset \mathcal{H}$. Since $L_F(v) = L_F(\hat{v})$, C is 2-covered in F' and $F[C] = F'[C]$.

Let $\ell = |C|$, and note that by definition, $F[C]$ is nonempty and is not an $(r'+1)$-star. If $\ell = n_t + r' + 1$, then by definition, $F' \in \mathcal{M}_{t,2}^r$, which contradicts the assumption that $\mathcal{H}$ is $\mathcal{M}_t^r$-free. So, $\ell \leq n_t + r$ and hence, $F \in \widehat{K}_\ell^r \cap \mathcal{M}_{t,1}^r$. By definition, F does not occur as a subgraph in any blow-up of $\mathcal{G}_i^r$ for $i \in [t]$. So, F' does not occur as a subgraph in any blow-up of $\mathcal{G}_i^r$ for $i \in [t]$ since F can be obtained from F' by duplicating v. So, by the definition of $\mathcal{M}_{t,1}^r$, $F' \in \widehat{K}_\ell^r \cap \mathcal{M}_{t,1}^r$, which contradicts the assumption that $\mathcal{H}$ is $\mathcal{M}_t^r$-free. ∎

We also need the following two lemmas about c-stars.

Lemma 5.4.6. *Every r-graph that is not a c-star contains a subgraph of size at most $r+2-c$ that is not a c-star.*

Proof. Let $\mathcal{H}$ be an r-graph that is not a c-star. Then for every set $S \subset V(\mathcal{H})$ of size at least c there exists an edge $E(S) \in \mathcal{H}$ such that $S \not\subset E(S)$.

Fix $E \in \mathcal{H}$ and assume that $E = \{v_1, \ldots, v_r\}$. Since $|E| > c$, there exists $E_1 \in \mathcal{H}$ such that $E \not\subset E_1$, i.e. $E \neq E_1$. Let $S_1 = E \cap E_1$, and without loss of generality we may assume that $S_1 = \{v_{i_1}, v_{i_1+1}, \ldots, v_r\}$, where $i_1 \geq 2$. If $|S_1| \leq c - 1$, then $\mathcal{H} = \{E, E_1\}$ is not a c-star, and we are done. So we may assume that $|S_1| \geq c$, Then there exists $E_2 \in \mathcal{H}$ such that $S_1 \not\subset E_2$. Let $S_2 = S_1 \cap E_2$, and without loss of generality we may assume that $S_2 = \{v_{i_2}, v_{i_2+1}, \ldots, v_r\}$, where $i_2 \geq i_1 + 1$. Keep doing this until $|S_j| \leq c - 1$ for some integer $j \leq r + 2 - c$. Then the subgraph $\mathcal{H} = \{E, E_1, \ldots, E_j\}$ is not a c-star. ∎

Let $\mathfrak{H}(n)$ be the family of all n-vertex r-graphs that are $\mathcal{G}_i^r$-colorable for some $i \in [t]$ or are blowups of some $(\hat{r}+1)$-star. The following criteria is useful in determining the structure of an $\mathcal{M}_t^r$-free r-graph.

Lemma 5.4.7. *Suppose that $\mathcal{H}$ is a symmetrized $\mathcal{F}$-free r-graph on n vertices. Then either $\mathcal{H}$ is $\mathcal{G}_i^r$-colourable for some $i \in [t]$, or $\mathcal{H}$ is a blow-up of an $(\hat{r}+1)$-star. In particular, if, in addition, $|\mathcal{H}| > 2n^r/r^r$, then $\mathcal{H}$ is $\mathcal{G}_i^r$-colourable for some $i \in [t]$. In other words, $\mathcal{M}_t^r$ is symmetrized-stable respects to $\mathfrak{H}(n)$.*

Proof. Let $T \subset V(\mathcal{H})$ be a set that contains exactly one vertex from each equivalent class of $\mathcal{H}$, and let $\mathcal{T}$ denote the induced subgraph of $\mathcal{H}$ on T. Then by assumption, $\mathcal{T}$ is 2-covered and $\mathcal{H}$ is a blow-up of $\mathcal{T}$.

If $\mathcal{T}$ is an $(\hat{r}+1)$-star, then $\mathcal{H}$ is a blow-up of an $(\hat{r}+1)$-star, and we are done. So, we may assume that $\mathcal{T}$ is not an $(\hat{r}+1)$-star.

Let $\ell = |T|$. If $\ell \geq \hat{n}_t + 1$, then by Lemma 5.4.6, there exists $F \subset \mathcal{T}$ of size at most four such that F is not an $(\hat{r}+1)$-star. Let $C \subset T$ be a set of size $\hat{n}_t + 1$ with $V(F) \subset C$. By greedily choosing a set in $\mathcal{T} \setminus F$ to cover each pair of vertices in $\binom{C}{2} \setminus \partial_{r-2} F$, it is easy to see that $\mathcal{T}$ contains a subgraph $\mathcal{T}'$ of size at most $\binom{\hat{n}_t+1}{2}$ such that C is 2-covered in $\mathcal{T}'$. By definition, $\mathcal{T}' \in \mathcal{M}_{t,2}^r$, a contradiction. So, $\ell \leq \hat{n}_t$. Then $\mathcal{T} \in \widehat{\mathcal{K}}_\ell^r$. Since $\mathcal{T} \notin \mathcal{M}_{t,1}^r$, $\mathcal{T}$ must occur as a subgraph in some blow-up of $\mathcal{G}_i^r$ for some $i \in [t]$. Therefore, $\mathcal{H}$ is $\mathcal{G}_i^r$-colourable.

Lemma 5.4.4 implies that an n-vertex $(\hat{r}+1)$-star has at most $2n^r/r^r$ edges, so if $|\mathcal{H}| > 2n^r/r^r$, then $\mathcal{H}$ can only be $\mathcal{G}_i^r$-colourable for some $i \in [t]$. ∎

Now we are ready to prove the first part of Theorem 5.1.11.

Proof of Theorem 5.1.11 (a). It follows from Theorem 4.1.4, Lemma 5.4.5, and Lemma 5.4.7 that $\mathrm{ex}(n, \mathcal{M}_t^r) = \max\{|\mathcal{H}| : \mathcal{H} \in \mathfrak{H}(n)\}$. By Lemma 5.4.4, every n-vertex r-graph that is a blowup of an $(\hat{r}+1)$-star has size at most $2n^r/r^r$, which is less than $\mathfrak{M}^{(r)}(n)$. Therefore, $\mathrm{ex}(n, \mathfrak{M}^{(r)}(n)) = \mathfrak{M}^{(r)}(n)$. ∎

5.4.4 Stability of $\mathcal{M}_t^r$

Recall that for every $\mathcal{M}_t^r$-free r-graph $\mathcal{H}$ on n vertices with $|\mathcal{H}| \geq (\lambda_t^{(r)} - \epsilon)n^r$ the set $Z(\mathcal{H})$ is defined as

$$Z(\mathcal{H}) = \left\{ u \in V(\mathcal{H}) \colon d_{\mathcal{H}}(u) \leq (r\lambda_t^{(r)} - 2\epsilon^{1/2})n^{r-1} \right\}$$

It follows from Fact 4.2.1 that $|Z(\mathcal{H})| \leq \epsilon^{1/2}n$.

The main result in the section is the following theorem.

Theorem 5.4.8. *If $\epsilon > 0$ is sufficiently small, n is sufficiently large, and $\mathcal{H}$ is an $\mathcal{M}_t^r$-free r-graph on n vertices with $|\mathcal{H}| \geq (\lambda_t^{(r)} - \epsilon)n^r$, then the hypergraph $\mathcal{H} - Z(\mathcal{H})$ is $\mathcal{G}_i^r$-colourable for some $i \in [t]$.*

The most technical part in the proof of Theorem 5.4.8 is to show the following lemma.

Lemma 5.4.9. *There exist $\zeta > 0$ and N_0 such that the following holds for all $n \geq N_0$. Suppose that $\mathcal{H}$ is an $\mathcal{M}_t^r$-free r-graph on n vertices with $\delta(\mathcal{H}) \geq (r\lambda_t^{(r)} - \zeta)n^{r-1}$. If there exists $v \in V(\mathcal{H})$ such that $\mathcal{H} - v$ is $\mathcal{G}_i^r$-colorable for some $i \in [t]$, then $\mathcal{H}$ is also $\mathcal{G}_i^r$-colorable, i.e. $\mathcal{H}$ is $\mathcal{G}_i^r$-extendable for $i \in [t]$.*

Theorem 5.4.8 is just an easy corollary of Theorem 4.1.7 and Lemma 5.4.9.

5.4.4.1 Preliminaries

Lemma 5.4.10. *Let F be an r-graph and S be a 2-covered set in F. Suppose that there exists a set $C \subset S$ of size $\hat{r}$ and two edges $e_1, e_2 \in F[S]$ with $e_1 \cap e_2 = S$. If $\psi : V(F) \to V(\mathcal{G}_i^r)$ satisfies $\psi(e) \in \mathcal{G}_i^r$ for all $e \in F$, then $\psi(C) = [\hat{r}]$.*

Proof. Since S is 2-covered in F, the induced map ψ_C of ψ on S is a bijection. In particular, $|\psi(C)| = \hat{r}$ and $\psi(e_1) \cap \psi(e_2) = \psi(C)$. So, $|\psi(e_1) \cap \psi(e_2)| = \hat{r}$, and it follows from the property of $\mathcal{G}_i^r$ that $\psi(e_1) \cap \psi(e_2) = [\hat{r}]$. So, $\psi(C) = [\hat{r}]$. ∎

Lemma 5.4.11. *Let $i \in [t]$. Suppose that $W \subset V(\mathcal{G}_i^r)$ is a set of size at least $\hat{n}_i - Q/2k_i^2$ and W contains the center of $\mathcal{G}_i^r$. Then $\mathcal{G}_i^r[W]$ does not occur as a subgraph in any blow-up of $\mathcal{G}_j^r$ for $j \in [t] \setminus \{i\}$.*

Proof. Without loss of generality we may assume that $W = [m]$ for some integer $m \geq \hat{n}_i - Q/2k_i^2$. Let $\mathcal{G} = \mathcal{G}_i^r[W]$, $W' = W \setminus [\hat{r}]$, and $\mathcal{G}'$ be the link of $[\hat{r}]$ in $\mathcal{G}$. It is clear that $\mathcal{G}'$ is a copy of the induced subgraph of $\mathcal{G}_i$ on some vertex set of size $|W'| \geq n_i - Q/2k_i^2$. Since $\delta_2(\mathcal{G}_i) \geq n_i - 3k_i/2 > Q/2k_i^2$, W' is 2-covered in $\mathcal{G}'$. So, $\mathcal{G}$ is 2-covered. It is clear that $\mathcal{G}$ contains two edges E_1, E_2 with $E_1 \cap E_2 = [\hat{r}]$, so if there exists an embedding $\psi\colon W \to [\hat{n}_j]$ of $\mathcal{G}$ into some $\mathcal{G}_j^r$ with $j \in [t] \setminus \{i\}$, then by Lemma 5.4.10, $\psi([\hat{r}]) = [\hat{r}]$. In other words, the induced map $\psi_{W'}$ of ψ on W is an embedding of $\mathcal{G}'$ into $\mathcal{G}_j$, which by a result in [173], is impossible. ∎

The following lemma shows that the family $\mathcal{M}_t^r$ has a nice inductive property on r.

Lemma 5.4.12. *Suppose that $\mathcal{H}$ is an $\mathcal{M}_t^r$-free r-graph. Then for every set $S \in V(\mathcal{H})$ of size $s \leq r-3$ the link $L_{\mathcal{H}}(S)$ is $\mathcal{M}_t^{r-s}$-free.*

Proof. Suppose for the contrary that there exists a set $S \in V(\mathcal{H})$ of size $s \leq \hat{r}$ so that $L_{\mathcal{H}}(S)$ is $\mathcal{M}_t^{r-s}$-free. Let F be a subgraph of $L_{\mathcal{H}}(S)$ that is also a member in $\mathcal{M}_t^{r-s}$. By definition $F \in \bigcup_{\ell \leq n_t + r' - s} \widehat{\mathcal{K}}_\ell^{r-s} \cup \mathcal{K}_{n_t+r'+1-s}^{r-s}$. Let C denote the core of F and suppose that $\ell = |C|$. Let $\hat{F} = F + S$ and $\hat{C} = C \cup S$. By assumption, $F[C] \neq \emptyset$ and it is not an $(r'+1-s)$-star, so $\hat{F}[\hat{C}] \neq \emptyset$ and it is not an $(r'+1)$-star.

If $\ell = n_t + r' + 1 - s$, then $|\hat{C}| = n_t + r' + 1$ and hence, $\hat{F} \in \mathcal{K}_{n_t+r'+1-s}^{r-s}$. Since $\mathcal{H}$ is $\mathcal{M}_{t,2}^r$-free, by the definition of $\mathcal{M}_{t,2}^r$, $\hat{F}[\hat{C}]$ must be either an $(r'+1)$-star or empty, a contradiction.

So, $\ell \le n_t + r' + 1 - s$ and hence, $F \in \widehat{\mathcal{K}}_\ell^{r-s}$. Then $\hat{F} \in \widehat{K}_{\ell+s}^r$ with core $\hat{C}$. Since $\mathcal{H}$ is $\mathcal{M}_{t,1}^r$-free, by the definition of $\mathcal{M}_{t,1}^r$, $\hat{F}$ must be $\mathcal{G}_i^r$-colorable for some $i \in [t]$. Let $\phi : V(\hat{F}) \to V(\mathcal{G}_i^r) = \{u_1, \ldots, u_{r'}, v_1, \ldots, v_{n_i}\}$ be a $\mathcal{G}_i^r$-coloring. Since $\hat{C}$ is 2-covered in $\hat{F}$, ϕ induced a bijection $\phi_{\hat{C}}$ between $\hat{C}$ and $\phi(\hat{C})$. It follows from $\hat{F}[\hat{C}] \neq \emptyset$ that $\{u_1, \ldots, u_{r'}\} \subset \phi(\hat{C})$. Let $w_j = \phi_{\hat{S}}^{-1}(u_j)$ for $j \in [r']$. Then $\{w_1, \ldots, w_{r'}\} \subset e$ for all $e \in \hat{F}[\hat{C}]$. If $S \not\subset \{w_1, \ldots, w_{r'}\}$, then the set $\{w_1, \ldots, w_{r'}\} \setminus S$ has size at least $r' + 1 - s$ and is contained in all $e' \in F[S]$. This implies that $F[C]$ is an $(r' + 1 - s)$-star, a contradiction. So, $S \in \{w_1, \ldots, w_{r'}\}$, that is, $\phi(S) \in \{u_1, \ldots, u_{r'}\}$. Also, since $N_{\hat{F}}(S) = V(\hat{F}) \setminus S$, there is no vertex $w \in V(\hat{F}) \setminus S$ with $\phi(w) \subset \phi(S)$. So, the induced map of ϕ on $V(F) = V(\hat{F}) \setminus S$ is a $\mathcal{G}_i^{r-s}$-coloring of F. However, this contradicts the assumption that $F \in \widehat{K}_\ell^{r-s} \cap \mathcal{M}_t^{r-s}$. ∎

The following lemma shows that if $\mathcal{H}$ is an almost complete subgraph of some blow-up of $\mathcal{G}$, then it contains every subgraph of $\mathcal{G}$.

Lemma 5.4.13. *Fix $r \ge 3, s \ge 0, m \ge t \ge 1, \zeta > 0$. Let $\mathcal{H}$ be an r-graph and $V(\mathcal{H}) = \bigcup_{i \in [m]} V_i$ be a partition. Let $\mathcal{G}$ be an r-graph on m vertices and $\widehat{\mathcal{G}} = \mathcal{G}[V_1, \ldots, V_m]$. Let $T \subset [m]$ be a set of size t and $S \subset V(\mathcal{H})$ be a set of size s. Suppose that*

(a) $|V_j| \ge (s+1)t\zeta^{1/r} n$ for all $j \in T$, and

(b) $|\mathcal{H}[V_{j_1}, \ldots, V_{j_r}]| \ge |\widehat{\mathcal{G}}[V_{j_1}, \ldots, V_{j_r}]| - \zeta n^r$ for all $\{j_1, \ldots, j_r\} \in \binom{T}{r}$, and

(c) $|L_{\mathcal{H}}(v)[V_{j_1}, \ldots, V_{j_{r-1}}]| \ge |L_{\widehat{\mathcal{G}}}(v)[V_{j_1}, \ldots, V_{j_{r-1}}]| - \zeta n^{r-1}$ for all $v \in S$ and $\{j_1, \ldots, j_{r-1}\} \in \binom{T}{r-1}$.

Then there exist $u_j \in V_j$ for all $j \in [T]$ so that the set $U = \{u_j : j \in T\}$ satisfies

$$\widehat{\mathcal{G}}[U] \subset \mathcal{H}[U], \quad \text{and} \quad L_{\widehat{\mathcal{G}}}(v)[U] \subset L_{\mathcal{H}}(v)[U], \quad \forall v \in S.$$

In particular, if $\mathcal{H} \subset \widehat{\mathcal{G}}$,

$$\widehat{\mathcal{G}}[U] = \mathcal{H}[U] \quad \text{and} \quad L_{\widehat{\mathcal{G}}}(v)[U] = L_{\mathcal{H}}(v)[U] \quad \forall v \in S.$$

Proof. Notice that $\widehat{\mathcal{G}}$ is the blow-up of $\mathcal{G}$ on $\bigcup_{i\in[m]} V_i$, so the size of $\widehat{\mathcal{G}}[V_{j_1}, \ldots, V_{j_r}]$ is either 0 or $\prod_{\ell\in[r]} |V_{j_\ell}|$ for all $\{j_1, \ldots, j_r\} \in \binom{[m]}{r}$, and the size of $L_{\widehat{\mathcal{G}}}(v)[V_{j_1}, \ldots, V_{j_{r-1}}]$ is either 0 or $\prod_{\ell\in[r-1]} |V_{j_\ell}|$ for all $v \in V(\widehat{\mathcal{G}})$ and $\{j_1, \ldots, j_{r-1}\} \in \binom{[m]}{r-1}$.

Choose a vertex u_j from V_j for $j \in T$ uniformly at random, and let $U = \{u_j : j \in T\}$. For distinct $j_1, \ldots, j_r \in [m]$ let $P_{j_1,\ldots,j_r}$ denote the probability of $u_{j_1} \ldots u_{j_r} \in \mathcal{H}$ under the condition that $u_{j_1} \ldots u_{j_r} \in \widehat{\mathcal{G}}$ Then by (a) and (b)

$$P_{j_1,\ldots,j_r} = \frac{|\mathcal{H}[V_{j_1}, \ldots, V_{j_r}]|}{|V_{j_1}| \cdots |V_{j_r}|} > 1 - \frac{\zeta n^r}{|V_{j_1}| \cdots |V_{j_r}|} > 1 - \frac{1}{(s+1)^r t^r}.$$

For $v \in S$ and distinct $j_1, \ldots, j_{r-1} \in [m]$ let $P_{j_1,\ldots,j_{r-1}}(v)$ denote the probability of $u_{j_1} \ldots u_{j_{r-1}} \in L_{\mathcal{H}}(v)$ under the condition that $u_{j_1} \ldots u_{j_{r-1}} \in L_{\widehat{\mathcal{G}}}(v)$. Then by (a) and (c)

$$P_{j_1,\ldots,j_{r-1}}(v) = \frac{|L_{\mathcal{H}}(v)[V_{j_1}, \ldots, V_{j_{r-1}}]|}{|V_{j_1}| \cdots |V_{j_{r-1}}|} > 1 - \frac{\zeta n^{r-1}}{|V_{j_1}| \cdots |V_{j_{r-1}}|} > 1 - \frac{\zeta^{1/r}}{(s+1)^{r-1} t(r-1)}.$$

Let E_U denote the event that $\widehat{\mathcal{G}}[U] \subset \mathcal{H}[U]$. Then by the union bound

$$P\left(E_U\right) > 1 - \binom{t}{r}\frac{1}{(s+1)^r t^r} > 1 - \frac{1}{r!(s+1)^r}.$$

For $v \in S$ let E_v denote the event that $L_{\widehat{\mathcal{G}}}(v) \subset L_{\mathcal{H}}(v)$. Then by the union bound

$$P\left(E_v\right) > 1 - \binom{t}{r-1}\frac{\zeta^{1/r}}{(s+1)^(r-1)t^(r-1)} > 1 - \frac{\zeta^{1/r}}{(r-1)!(s+1)^(r-1)}.$$

So, by the union bound again,

$$P\left(E_U \wedge \left(\bigwedge_{v\in S} E_v\right)\right) > 1 - \frac{1}{r!(s+1)^r} - s\frac{\zeta^{1/r}}{(r-1)!(s+1)^(r-1)} > 0$$

Therefore, there exist $u_j \in V_j$ for $j \in [T]$ so that the set $U = \{u_j : j \in T\}$ satisfies $\widehat{\mathcal{G}}[U] \subset \mathcal{H}[U]$ and $L_{\widehat{\mathcal{G}}}(v)[U] \subset L_{\mathcal{H}}(v)[U]$ for all $v \in S$. ∎

5.4.4.2 Proof of Lemma 5.4.9

Proof of Lemma 5.4.9. We proceed by induction on r, and the base case is $r = 3$, which is already proved in Section 5.3. So, let us assume that Lemma 5.4.9 holds for all $r' < r$, and $r \geq 4$.

Let $V = V(\mathcal{H})$, $\mathcal{H}_v = \mathcal{H} - v$, and suppose that $V \setminus \{v\} = \bigcup_{j\in[\hat{n}]} V_j$ is a $\mathcal{G}_i^r$-coloring of $\mathcal{H}_v$. Let $\widehat{\mathcal{G}} = \mathcal{G}_i^r[V_1, \ldots, V_{\hat{n}_i}]$ and note that $\mathcal{H}_v \subset \widehat{\mathcal{G}}$. For every $u \in V \setminus \{v\}$ let $\operatorname{cap}(u) = |L_{\widehat{\mathcal{G}}}(u)|$

denote the capacity of u and note that $|L_{\mathcal{H}_v}(u)| \le \mathrm{cap}(u)$. Call members in $L_{\widehat{\mathcal{G}}}(u) \setminus L_{\mathcal{H}_v}(u)$ the missing edges of $L_{\mathcal{H}_v}(u)$, and call members in $\widehat{\mathcal{G}} \setminus \mathcal{H}_v$ the missing edges of $\mathcal{H}_v$. Note that

$$|\mathcal{H}_v| \ge (\lambda_t^{(r)} - \zeta)n^r - n^{r-1} > (\lambda_t^{(r)} - 2\zeta)n^r,$$

so the number of missing edges of $\mathcal{H}_v$ is at most $2\zeta n^r$.

Claim 5.4.14. *For every $u \in V(\mathcal{H})$ we have*

$$|N_{\mathcal{H}}(u)| \ge \left(\frac{r-1}{r} - r^r\zeta\right) n.$$

Proof. Suppose for the contrary that there exists $u \in V(\mathcal{H})$ with $|N_{\mathcal{H}}(u)| < \left(\frac{r-1}{r} - r^r\zeta\right) n$. By Lemma 5.4.12, $L_{\mathcal{H}}(u)$ is an $\mathcal{M}_t^{r-1}$-free $(r-1)$-graph (with vertex set $N_{\mathcal{H}}(u)$), so by Theorem 5.1.11 (a),

$$\begin{aligned}
|L_{\mathcal{H}}(u)| &\le \lambda_t^{(r-1)} \left(\frac{r-1}{r} - r^r\zeta\right)^{r-1} n^{r-1} \\
&= \left(1 - \frac{r}{r-1} r^r\zeta\right)^{r-1} \lambda_t^{(r-1)} \left(\frac{r-1}{r}\right)^{r-1} n^{r-1} \\
&\le (1 - r^r\zeta)\, r\lambda_t^{(r)} n^{r-1} = \left(r\lambda_t^{(r)} - r^{r+1}\lambda_t^{(r)}\zeta\right) n^{r-1} < \left(r\lambda_t^{(r)} - 2\zeta\right) n^{r-1},
\end{aligned}$$

which contradicts our assumption that $\delta(\mathcal{H}) \ge (r\lambda_t^r - \zeta)n^{r-1}$. ∎

Let $\zeta_1 = 4r^{r/2+1}\zeta^{1/2}$. The following claim is an easy corollary of Lemma 5.4.3 and the fact that $|\mathcal{H}_v| > (\lambda_t^{(r)} - 2\zeta)n^r$.

Claim 5.4.15.

$$|V_j| = \frac{n}{r} \pm \zeta_1 n, \ \forall j \in [\hat{r}], \quad \text{and} \quad |V_j| = \frac{3n}{rn_i} \pm \zeta_1 n, \ \forall j \in [\hat{r}+1, \hat{n}_i].$$

Claim 5.4.16. $\mathrm{cap}(u) \le (1 + 3rn_i\zeta_1)\, r\lambda_t^{(r)} n^{r-1}$ *for all* $u \in V \setminus \{v\}$.

Proof. By Claim 5.4.15, for every $j \in [\hat{r}]$ and $u \in V_j$ we have

$$\begin{aligned}
\mathrm{cap}(u) &\le \left(\frac{n}{r} + \zeta_1 n\right)^{r-4} \times \lambda_t n_i^3 \left(\frac{3n}{rn_i} + \zeta_1 n\right)^3 \\
&= (1 + r\zeta_1)^{r-4} \left(1 + \frac{rn_i}{3}\zeta_1\right)^3 r\lambda_t^{(r)} n^{r-1} \\
&< \left(1 + 2r^2\zeta_1\right)(1 + 2rn_i\zeta_1)\, r\lambda_t^{(r)} n^{r-1} < (1 + 3rn_i\zeta_1)\, r\lambda_t^{(r)} n^{r-1}.
\end{aligned}$$

Similarly, for every $j \in [\hat{r}+1, \hat{n}_i]$ and $u \in V_j$ we have

$$\begin{aligned}
\mathrm{cap}(u) &\le \left(\frac{n}{r} + \zeta_1 n\right)^{r-3} \times 3\lambda_t n_i^2 \left(\frac{3n}{rn_i} + \zeta_1 n\right)^2 \\
&= (1 + r\zeta_1)^{r-3} \left(1 + \frac{rn_i}{3}\zeta_1\right)^2 r\lambda_t^{(r)} n^{r-1} \\
&< \left(1 + 2r^2\zeta_1\right)(1 + 2rn_i\zeta_1)\, r\lambda_t^{(r)} n^{r-1} < (1 + 3rn_i\zeta_1)\, r\lambda_t^{(r)} n^{r-1}.
\end{aligned}$$

∎

Since $L_{\mathcal{H}_v}(u) > \delta(\mathcal{H}) - n^{r-2} > (r\lambda_t^{(r)} - 2\zeta)n^{r-1}$, by Claim 5.4.16, the number of missing edges of $L_{\mathcal{H}_v}$ is at most

$$2\zeta n^{r-1} + 3r^2\lambda_t^{(r)} n_i \zeta_1 n^{r-1} < \zeta_2 n^{r-1},$$

where $\zeta_2 = 4r^2\lambda_t^{(r)} n_i \zeta_1$.

For every $j \in [\hat{n}_i]$ and $u \in V_j$ the vertices in $\bigcup_{\ell \in [\hat{n}_i] \setminus \{j\}} V_\ell \setminus N_{\mathcal{H}_v}(u)$ are called the missing neighbors of u.

Claim 5.4.17. *For every $u \in V \setminus \{v\}$ the number of missing neighbors of u is at most $r^{r-2}\zeta_2 n$ for all $u \in V \setminus \{v\}$.*

Proof. If $u \in V_j$ for some $j \in [\hat{r}]$, then by Claims 5.4.15 and 5.4.14, the number of missing neighbors of u is at most

$$\begin{aligned} n - |U_i| - \left(\frac{r-1}{r} - r^r\zeta\right) n \leq n - \left(\frac{n}{r} - \zeta_1 n\right) - \left(\frac{r-1}{r} - r^r\zeta\right) n &= (\zeta_1 + r^r\zeta)n \\ &< r^{r-2}\zeta_2 n. \end{aligned}$$

Now suppose that $u \in V_j$ for some $j \in [\hat{r}+1, \hat{n}_i]$. By Claim 5.4.15, the degree of a vertex $w \in V \setminus V_j$ in $L_{\widehat{\mathcal{G}}}(u)$ satisfies

$$\begin{aligned} d_{L_{\widehat{\mathcal{G}}(u)}}(w) > \min\{ &\left(\frac{n}{r} - \zeta_1 n\right)^{r-4} 3\lambda_t n_i^2 \left(\frac{3n}{rn_i} - \zeta_1 n\right)^2, \\ &\left(\frac{n}{r} - \zeta_1 n\right)^{r-3} \delta_2(\mathcal{G}_i) \left(\frac{3n}{rn_i} - \zeta_1 n\right) \} > \frac{n^{r-2}}{r^{r-2}}. \end{aligned}$$

Since the number of missing edges in $L_{\mathcal{H}_v}(u)$ is at most $\zeta_2 n^{r-1}$, the number of vertices with degree zero in $L_{\mathcal{H}_v}(u)$ is at most

$$\frac{\zeta_2 n^{r-1}}{n^{r-2}/r^{r-2}} < r^{r-2}\zeta_2 n.$$

∎

The following claim shows that every set in $L_{\mathcal{H}}(v)$ contains at most one vertex from each set V_j for $j \in [\hat{n}_i]$.

Claim 5.4.18. *$|E \cap V_j| \leq 1$ for every $j \in [\hat{n}_i]$ and every $E \in \mathcal{H}$ that contains v.*

Proof. Suppose for the contrary that $|E \cap V_j| \geq 2$ for some $j \in [\hat{n}_i]$ and some $E \in \mathcal{H}$ that contains v, and without loss of generality we may assume that $j = 1$ and $u_1, u_1' \in E \cap V_1$. Applying Lemma 5.4.13 with $T = [\hat{n}_i] \setminus \{1\}$ and $S = \{u_1, u_1'\}$ we obtain $u_j \in V_j$ for $j \in [\hat{n}_i] \setminus \{1\}$ such that sets $U = \{u_1, u_2, \dots, u_{\hat{n}_i}\}$ and $U' = \{u_1', u_2, \dots, u_{\hat{n}_i}\}$ satisfy

$$\mathcal{H}_v[U] = \widehat{\mathcal{G}}[U] \quad \text{and} \quad \mathcal{H}_v[U'] = \widehat{\mathcal{G}}[U'].$$

Let $\mathcal{G} = \mathcal{H}_v[U]$, $\mathcal{G}' = \mathcal{H}_v[U']$, and

$$F = \mathcal{G} \cup \mathcal{G}' \cup \{E\}.$$

It is clear that $F \in \widehat{\mathcal{K}}^r_{\hat{n}_i+1}$ with core $C = \{u_1, u_1', u_2, \dots, u_{\hat{n}_i}\}$.

If $i = t$, then using Lemma 5.4.6, it is easy to see that F contains a subgraph $F \in \mathcal{K}^r_{\hat{n}_t+1}$ with core C, and $F'[C]$ is nonempty and not an $(\hat{r}+1)$-star. In other words, $F' \in \mathcal{M}^r_{t,2}$, a contradiction.

Suppose that $i < t$. Since $\mathcal{G} \subset F$ and $\mathcal{G}$ is a copy of $\mathcal{G}_i^r$, F does not occur as a subgraph in any blow-up of $\mathcal{G}_j^r$ for $j \in [t] \setminus \{i\}$. On the other hand, since $\mathcal{G}_i^r$ is $\mathcal{K}_{n_i+1}^r$-free, F does not occur as a subgraph in any blow-up of $\mathcal{G}_i^r$. So, $F \in \widehat{\mathcal{K}}_{\hat{n}_i+1}^r \cap \mathcal{M}_{t,1}^r$, a contradiction. ∎

To summarize, $\mathcal{H}$ and $\mathcal{H}_v$ have the following properties.

(a) $|\mathcal{H}_v[V_{j_1}, \dots, V_{j_r}]| > |\widehat{\mathcal{G}}[V_{j_1}, \dots, V_{j_r}]| - 2\zeta n^{r-1}$ for all distinct $j_1, \dots, j_r \in [\hat{n}_i]$, and

(b) $|L_{\mathcal{H}_v}(u)[V_{j_1}, \dots, V_{j_{r-1}}]| > |L_{\widehat{\mathcal{G}}}(u)[V_{j_1}, \dots, V_{j_{r-1}}]| - \zeta_2 n^{r-1}$ for all distinct $j_1, \dots, j_{r-1} \in [\hat{n}_i]$, and

(c) $|N_{\mathcal{H}_v}(u) \cap V_j| > |N_{\widehat{\mathcal{G}}}(u) \cap V_j| - r^{r-2}\zeta_2 n$ for all $j \in [\hat{n}_i]$ such that $u \notin V_j$, and

(d) every set in $L_{\mathcal{H}}(v)$ has at most one vertex in each set V_j for $j \in [\hat{n}_i]$.

Define

$$S_g = \left\{ j \in [\hat{n}_i] \colon |N_{\mathcal{H}}(v) \cap V_j| \geq n_i^2 \zeta_2^{1/r} n \right\},$$

and let $s_g = |S_g|$.

Claim 5.4.19. *There exists $j \in [\hat{n}_i]$ such that $|N_{\mathcal{H}}(v) \cap V_j| < n_i^2 \zeta_2^{1/r} n$. In other words, $s_g \leq \hat{n}_i - 1$.*

Proof. Let $V_j' = N_{\mathcal{H}}(v) \cap V_j$ and suppose for the contrary that $|V_j'| \geq n_i^2 \zeta_2^{1/r} n$ for all $j \in [\hat{n}_i]$. Applying Lemma 5.4.13 with $T = [\hat{n}_i]$ and $S = \emptyset$, we obtain $u_j \in V_j$ for $j \in [\hat{n}_i]$ such that the set $U = \{u_1, \ldots, u_{\hat{n}_i}\}$ satisfies $\mathcal{H}_v[U] = \widehat{\mathcal{G}}[U]$. For every $u_j \in U$ let $e_j \in \mathcal{H}$ be an edge that contains v and u_j. Define $\mathcal{G} = \mathcal{H}_v[U]$ and

$$F = \mathcal{G} \cup \{e_j \colon j \in [\hat{n}_i]\}\,.$$

It is clear that $F \in \widehat{\mathcal{K}}^r_{\hat{n}_i+1}$ with core $C = U \cup \{v\}$. Then similar to the proof of Claim 5.4.18, one can show that either F contains a subgraph F' with $F' \in \mathcal{M}^r_{t,2}$ (the case $i = t$) or $F \in \widehat{\mathcal{K}}^r_{\hat{n}_i+1} \cap \mathcal{M}^r_{t,1}$ (the case $i < t$), both contradict the assumption that $\mathcal{H}$ is $\mathcal{M}^r_t$-free. ∎

The next claim shows that in order to finish the proof of Lemma 5.4.9 it suffices to prove that $s_g = \hat{n}_i - 1$.

Claim 5.4.20. *Suppose that $s_g = \hat{n}_i - 1$ and $\{i_0\} = [\hat{n}_i] \setminus S_g$. Then $N_{\mathcal{H}}(v) \cap V_{i_0} = \emptyset$ and $L_{\mathcal{H}}(v) \subset L_{\widehat{\mathcal{G}}}(u)$ for $u \in V_{i_0}$. In particular, $V = \bigcup_{j\in[\hat{n}_i]} \widehat{V}_j$ with $\widehat{V}_{i_0} = V_{i_0} \cup \{v\}$ and $\widehat{V}_j = V_j$ for $j \in [\hat{n}_i] \setminus \{i_0\}$ is a $\mathcal{G}^r_i$-coloring of $\mathcal{H}$.*

Proof. Let $V'_j = V_j \cap N_{\mathcal{H}}(v)$ for $j \in [\hat{n}_i]$ and without loss of generality we may assume that $i_0 = 1$. Suppose for the contrary that there exists a vertex $u_1 \in V'_1$. Since $|V'_j| \geq n_i^2 \zeta_2^{1/r} n$ for $j \in [\hat{n}_i] \setminus \{1\}$, apply Lemma 5.4.13 with $T = [\hat{n}_i] \setminus \{1\}$ and $S = \{u_1\}$, we obtain $u_j \in V'_j$ for $j \in [\hat{n}_i] \setminus \{1\}$ such that the set $U = \{u_1, \ldots, u_{\hat{n}_i}\}$ satisfies $\mathcal{H}_v[U] = \widehat{\mathcal{G}}[U]$. For every $u_j \in U$ let $e_j \in \mathcal{H}$ be an edge that contains v and u_j. Define $\mathcal{G} = \mathcal{H}_v[U]$ and

$$F = \mathcal{G} \cup \{e_j \colon j \in [\hat{n}_i]\}\,.$$

It is clear that $F \in \widehat{\mathcal{K}}^r_{\hat{n}_i+1}$ with core $C = U \cup \{v\}$. Then similar to the proof of Claim 5.4.18, one can show that either F contains a subgraph F' with $F' \in \mathcal{M}^r_{t,2}$ (the case $i = t$) or $F \in \widehat{\mathcal{K}}^r_{\hat{n}_i+1} \cap \mathcal{M}^r_{t,1}$ (the case $i < t$), both contradict the assumption that $\mathcal{H}$ is $\mathcal{M}^r_t$-free. Therefore, $V'_{i_0} = \emptyset$.

Now let us prove the second part. Suppose for the contrary that there exists $e_v \in L_{\mathcal{H}}(v) \setminus L_{\widehat{\mathcal{G}}}(u)$ for $u \in V_1$, and without loss of generality we may assume that $e_v \cap V_j = \{\hat{v}_j\}$ for $2 \leq j \leq a$ and $\hat{r}+1 \leq j \leq \hat{r}+b$, where $a, b \in \mathbb{N}$ satisfy $a + b = r$.

Define

$$V'_j = \begin{cases} V_j \cap \left(\bigcap_{u \in e_v} N_{\mathcal{H}}(u)\right), & j = 1, \\ U_j \cap N_{\mathcal{H}}(v) \cap \left(\bigcap_{u \in e_v \setminus \{v_j\}} N_{\mathcal{H}}(u)\right), & j \in [2, a] \cup [\hat{r}+1, \hat{r}+b], \\ U_j \cap N_{\mathcal{H}}(v) \cap \left(\bigcap_{u \in e_v} N_{\mathcal{H}}(u)\right), & \text{otherwise.} \end{cases}$$

Applying Lemma 5.4.13 with $T = [\hat{n}_i]$ and $S = \emptyset$, we obtain $\hat{v}_j \in V'_j$ for all $j \in [\hat{u}]$ such that the set $U = \{\hat{v}_j \colon j \in [\hat{u}]\}$ satisfies

$$\mathcal{H}[U] = \widehat{\mathcal{G}}[U] =: \mathcal{G}.$$

Let $E_v = e_v \cup \{v\}$. For every pair $(w, \hat{w}) \in E_v \times U$ with $\{w, \hat{w}\} \not\in V_j$ for all $j \in [\hat{n}_i]$ there exists $e_{w,\hat{w}} \in \mathcal{H}$ such that $\{w, \hat{w}\} \subset e_{w_1,w_2}$. Define

$$F = \{E_v\} \cup \{e_{w_1,w_2} \colon (w, \hat{w}) \in E_v \times U, \{w, \hat{w}\} \not\in V_j, \forall j \in [\hat{n}_i]\} \cup \mathcal{G}.$$

By Lemma 5.4.11, F does not occur as a subgraph in any blow-up of $\mathcal{G}_j^r$ for all $j \in [t] \setminus \{i\}$, so F is a subgraph in some blow-up of $\mathcal{G}_i^r$. In other words, there exists $\psi\colon V(F) \to V(\mathcal{G}_i^r)$ such that $\psi(e) \in \mathcal{G}_i^r$ for all $e \in F$. Let

$$U_v = E_v \cup U \setminus (\{\hat{v}_j \colon j \in [a] \cup [\hat{r}+1, \hat{r}+b]\}).$$

Since U and U_v are 2-covered in F and $|U| = |U_v| = |V(\mathcal{G}_i^r)|$, the induced maps ψ_U and ψ_{U_v} of ψ on U and U_v are bijections. Moreover, $\psi(v_1) = \psi(v)$ and $\psi(\hat{v}_j) = \psi(v_j)$ for $j \in \cup[\hat{r}+1, \hat{r}+b]$. It follows from $E_v \in F$ that $\psi(E_v) \in \mathcal{G}_i^r$. So,

$$|\mathcal{G}_i^r| \geq 1 + |F[U]| = 1 + |\mathcal{G}| = 1 + |\mathcal{G}_i^r|,$$

a contradiction. ∎

Claim 5.4.21. *If $|N_{\mathcal{H}}(v) \cap V_j| < n_i^2 \zeta_2^{1/r} n$ for some $j \in [\hat{r}]$, then $s_g = \hat{n}_i - 1$.*

Proof. Suppose for the contrary that $s_g < \hat{n}_i - 2$. Then by Claim 5.4.15,

$$\begin{aligned} |N_{\mathcal{H}}(v)| &\leq (\hat{r}-1)\left(\frac{n}{r} + \zeta_1 n\right) + (n_i - 1)\left(\frac{3n}{rn_i} + \zeta_1 n\right) \\ &= \frac{r-1}{r}n - \frac{3n}{rn_i} + \hat{n}_i \zeta_1 n < \frac{r-1}{r}n - r^r \zeta n, \end{aligned}$$

which contradicts Claim 5.4.14. ∎

Claims 5.4.20 and 5.4.21 imply that if $|N_{\mathcal{H}}(v) \cap V_j| < n_i^2 \zeta_2^{1/r} n$ for some $j \in [\hat{r}]$, then we are done. So we may assume that

$$|N_{\mathcal{H}}(v) \cap V_j| \geq n_i^2 \zeta_2^{1/r} n, \quad \forall j \in [\hat{r}],$$

and $|N_{\mathcal{H}}(v) \cap V_j| < n_i^2 \zeta_2^{1/r} n$ for some $j \in [\hat{r}+1, \hat{n}_i]$.

Let us first show a trivial lower bound for s_g.

Claim 5.4.22. $s_g \geq 2n_i/3$.

Proof. Suppose for the contrary that $s_g < 2n_i/3$. Then by Claim 5.4.15,

$$\begin{aligned} |N_{\mathcal{H}}(v)| &\leq \hat{r}\left(\frac{n}{r} + \zeta_1 n\right) + \left(\frac{2n_i}{3} - 1\right)\left(\frac{3n}{rn_i} + \zeta_1 n\right) \\ &= \frac{r-1}{r} n - \frac{3n}{rn_i} + \hat{n}_i \zeta_1 n < \frac{r-1}{r} n - r^r \zeta n, \end{aligned}$$

which contradicts Claim 5.4.14. ∎

Recall that Claim 5.4.18 implies that every set $E \in \mathcal{H}$ containing v has at most one vertex from each V_j for $j \in [\hat{n}_i]$. Our next claim shows that E has exactly one vertex in V_j for every $j \in [\hat{r}]$.

Claim 5.4.23. *Every set $E \in \mathcal{H}$ containing v has exactly one vertex in V_j for $j \in [\hat{r}]$.*

Proof. Suppose for the contrary that there exists a set $E_v \in \mathcal{H}$ containing v such that E_v has $a < \hat{r}$ vertices in $\bigcup_{j\in[\hat{r}]} V_j$. Without loss of generality we may assume that $E_v \cap V_j = \{v_j\}$ for $j \in [a] \cup [\hat{r}+1, \hat{r}+b]$, where $b = r-1-a$. Let

$$V_j' = \begin{cases} V_j \cap \left(\bigcap_{u\in E_v\setminus\{v_j\}} N_{\mathcal{H}}(u)\right), & j \in [a] \cup [\hat{r}+1, \hat{r}+b], \\ V_j \cap \left(\bigcap_{u\in E_v} N_{\mathcal{H}}(u)\right), & \text{otherwise.} \end{cases}$$

Then Claim 5.4.17 implies that

$$|V_j'| \geq n_i^2 \zeta_2^{1/r} n - (r-1) \times r^{r-2}\zeta_2 n > n_i(n_i-1)\zeta_2^{1/r} n, \quad \forall j \in S_g.$$

Applying Lemma 5.4.13 with $T = S_g \setminus ([a] \cup [\hat{r}+1, \hat{r}+b])$ and $S = \emptyset$ we obtain $\hat{v}_j \in V_j'$ for all $j \in S_g$ such that the set $U = \{\hat{v}_j \colon j \in S_g\}$ satisfies $\mathcal{H}[U] = \widehat{\mathcal{G}}[U]$. Let $\widehat{G} = \widehat{\mathcal{G}}[U]$ and

$$F = \widehat{G} \cup \{E_v\} \cup \{e_{w,\hat{w}} : (w, \hat{w}) \in E_v \times U, \{w, \hat{w}\} \not\subseteq V_j, \forall j \in [a] \cup [\hat{r}+1, \hat{n}_i]\}.$$

It is clear that U is 2-covered in F and $F[U] = \widehat{G}$ is not an $(\hat{r}+1)$-star. So, F must occur as a subgraph in some blow-up of $\mathcal{G}_j^r$ for $j \in [t]$. In other words, there exists $\psi \colon V(F) \to V(\mathcal{G}_j^r)$ such that $\psi(e) \in \mathcal{G}_j^r$ for all $e \in F$. Let $w_\ell = \psi(\hat{v}_\ell)$ for $\ell \in [\hat{r}]$. Then by Lemma 5.4.10, $w_1, \ldots, w_{\hat{r}}$ are distinct and the set $\{w_1, \ldots, w_{\hat{r}}\}$ is the center of $\mathcal{G}_j^r$. Since $v, v_{\hat{r}+1}, \ldots, v_{\hat{r}+b}$ are adjacent to all vertices in U, we must have

$$\{\psi(v), \psi(v_{\hat{r}+1}), \ldots, \psi(v_{\hat{r}+b})\} \subset V(\mathcal{G}_j^r) \setminus \psi(U).$$

In particular, $|\psi(E_v) \cap \psi(E)| \leq a < \hat{r}$ for all $E \in F$, a contradiction. ∎

For every $e \in \prod_{j\in[\hat{r}]} V_j$ denote by $d(e)$ the degree of e in $L_{\mathcal{H}}(v)$, and denote by G_e the link of $e \cup \{v\}$ in $\mathcal{H}$. Since $\sum_e d(e) = |L_{\mathcal{H}}(v)| \geq (r\lambda_t^{(r)} - \zeta)n^{r-1}$, there exists a set e with

$$d(e) \geq \frac{(r\lambda_t^{(r)} - \zeta)n^{r-1}}{\left(\frac{n}{r} + \zeta_1 n\right)^{r-3}} > (r^{r-2}\lambda_t^{(r)} - 2r^r\zeta_1)n^2 > (3\lambda_t - 2r^r\zeta_1)\left(\frac{3n}{r}\right)^2.$$

Fix such a set e_v. Then we have the following claim.

Claim 5.4.24. *The set*

$$N(u) = \left\{j \in [\hat{r}+1, \hat{n}_i] \colon |N_{G_e} \cap V_j| \geq n_i^2 \zeta_2^{1/r} n\right\}$$

has size at most $n_i - k_i$.

Proof. Let $m = |N(u)|$ and suppose for the contrary that $m > n_i - k_i$. We may assume that $m = n_i - k_i + 1$ since otherwise we can take an $(n_i - k_i + 1)$-subset of $N(u)$. Let

$$V_j' = \begin{cases} V_j \cap N_{\mathcal{H}}(v) \cap N_{\mathcal{H}}(u), & j \in [\hat{r}], \\ \\ V_j \cap N_{G_{e_v}}(u), & j \in N(u). \end{cases}$$

Then by Claim 5.4.17, $|V_j'| \geq n_i^2 \zeta_2^{1/r} n - r^{r-2}\zeta_2 n > n_i(n_i - 1)\zeta_2^{1/r} n$ for $j \in [\hat{r}] \cup N(u)$. Applying Lemma 5.4.13 with $T = [\hat{r}] \cup N(u)$ and $S = \emptyset$, we obtain $u_j \in V_j$ for $j \in T$ such that the set U satisfies $\mathcal{H}[U] = \widehat{\mathcal{G}}[U]$. Let $\mathcal{G} = \mathcal{H}[U]$, $e_j, e_j' \in \mathcal{H}$ such that $\{v, u_j\} \subset e_j$ and $\{u, u_j\} \subset e_j'$ for $j \in [\hat{r}]$, and

$$F = \mathcal{G} \cup \{e_j \colon j \in [\hat{r}]\} \cup \left\{e_j' : j \in [\hat{r}]\right\} \cup \{e_v \cup \{v, u, u_j\} \colon j \in N(u)\}.$$

It is clear that $F \in \widehat{\mathcal{K}}_m^r$ with core U and $F[U] = \mathcal{G}$ is nonempty and not an $(\hat{r}+1)$-star. Since $|U| = \hat{r} + m > \hat{n}_i - Q/2k_i^2$ and U contains the center, by Lemma 5.4.11, F does not occur as a subgraph in any blow-up of $\mathcal{G}_j^r$ for $j \in [t] \setminus \{i\}$. So, F is a subgraph in some blow-up of $\mathcal{G}_i^r$. In other words, there exists $\psi : V(F) \to V(\mathcal{G}_i^r)$ such that $\psi(e) \in \mathcal{G}_i^r$ for all $e \in F$.

Let $w_j = u_j$ for $j \in [\hat{r}]$. It is clear that $\mathcal{G}$ contains two edges E_1, E_2 with $E_1 \cap E_2 = \{u_1, \ldots, u_{\hat{r}}\}$, so by Lemma 5.4.10, the set $C = \{\psi(u_1), \ldots, \psi(u_{\hat{r}})\} = \{w_1, \ldots, w_{\hat{r}}\}$ is the center of $\mathcal{G}_i^r$. Since the set $\{v, u\} \cup U$ is 2-covered in F, $\psi(v), \psi(u) \notin C$ and $\psi(u_j) \notin C$ for all $j \in N(u)$. Therefore, $\psi(e) = C$ and hence $\psi\left(L_F(C)\right)$ is a subgraph of $\mathcal{G}_i$. However, the pair $\{u, v\}$ has degree at least $n_i - k_i + 1$ in $L_F(C)$ contradicting the fact that $\Delta_2(\mathcal{G}_i) = n_i - k_i$. ∎

Claim 5.4.25. $s_g = \hat{n}_i - 1$.

Proof. Suppose for the contrary that $s_g < \hat{n}_i - 1$. Then

$$
\begin{aligned}
d(e_v) &= \frac{1}{2}\left(\sum_{u \in V_j : j \in S_g} d_{G_e}(u) + \sum_{u \in V_j : j \notin S_g} d_{G_e}(u)\right) \\
&\le \frac{1}{2}\left((n_i - 2)\left(\frac{3n}{rn_i} + \zeta_1 n\right) + 2 \times n_i^2 \zeta_2^{1/r} n\right)\left((n_i - k_i)\left(\frac{3n}{rn_i} + \zeta_1 n\right) + k_i \times n_i^2 \zeta_2^{1/r} n\right) \\
&< \frac{1}{2}\left((n_i - 2)\frac{3n}{rn_i} + 3n_i^2 \zeta_2^{1/r} n\right)\left((n_i - k_i)\frac{3n}{rn_i} + 2k_i n_i^2 \zeta_2^{1/r} n\right) \\
&< \frac{(n_i - 1)(n_i - k_i) - k_i}{2n_i^2}\left(\frac{3n}{r}\right)^2 - \frac{n_i - 2k_i}{2n_i^2}\left(\frac{3n}{r}\right)^2 + 3k_i n_i^2 \zeta_2^{1/r} n^2 \\
&< (3\lambda_t - 2r^r \zeta_1)\left(\frac{3n}{r}\right)^2,
\end{aligned}
$$

a contradiction. Here we used the fact that $\lambda_t = \frac{1}{6}\left(1 - \frac{k_i + 1}{n_i}\right) = \frac{(n_i - 1)(n_i - k_i) - k_i}{6n_i^2}$. ∎

Claim 5.4.25 completes the proof of Lemma 5.4.9. ∎

5.4.5 Feasible region of $\mathcal{M}_t^r$ and $\xi_v(\mathcal{M}_t^r)$

For $i \in [t]$ let $\mathcal{H}_i(n)$ be the balanced blowup of $\mathcal{G}_i^r$ on n vertices, and let

$$
x_i = \lim_{n \to \infty} d\left(\partial \mathcal{H}_i(n)\right) = \frac{(r-1)!}{r^{r-1}}\left((r-3)r^r \lambda_t^{(r)} + \frac{9}{2}\left(1 - \frac{1}{n_i}\right)\right).
$$

By Theorem 5.1.11 (a), $\{(x_i, \pi(\mathcal{M}_t^r)) \colon i \in [t]\} \subset M(\mathcal{M}_t^r)$. In particular, $|M(\mathcal{M}_t^r)| \ge t$.

It follows from Theorem 5.1.11 that $\mathcal{M}_t^r$ is vertex-t-stable respects to $\mathcal{H}_1(n), \ldots, \mathcal{H}_t(n)$. On the other hand, it is clear that there exists a constant $c > 0$ such that every n-vertex r-graph $\mathcal{H}$ that is the balanced blowup of $\mathcal{G}_i^r$ for some $i \in [t]$ satisfies $\delta_{r-1}(\mathcal{H}) \ge cn$. Therefore, by

Proposition 5.4.1, $|M(\mathcal{M}_t^r)| \leq \xi_v(\mathcal{M}_t^r) \leq t$. Combined with the conclusion above we obtain $|M(\mathcal{M}_t^r)| = \xi_v(\mathcal{M}_t^r) = t$ and $M(\mathcal{M}_t^r) = \{(x_i, \pi(\mathcal{M}_t^r)) : i \in [t]\}$.

CHAPTER 6

EXTREMAL SET THEORY

Previously published as X. Liu. d-cluster-free sets with a given matching number. European J. Combin., 82:103000, 19, 2019; X. Liu. Structural results for conditionally intersecting families and some applications. Electron. J. Combin., 27(2):Paper No. 2.33, 13, 2020; and X. Liu and D. Mubayi. Tight bounds for Katonas shadow intersection theorem. European J. Combin., 97:Paper No. 103391, 17, 2021.

6.1 d-cluster-free sets with a given matching number

6.1.1 Introduction

Recall that a d-cluster of k-sets is a collection of d different k-sets $A_1, ..., A_d$ such that

$$|A_1 \cup \cdots \cup A_d| \leq 2k, \text{ and } |A_1 \cap \cdots \cap A_d| = 0.$$

A family $\mathcal{F} \subset \binom{[n]}{k}$ is d-cluster-free if it does not contain d-clusters. Note that a family is intersecting if and only if it is 2-cluster-free. The celebrated Erdős–Ko–Rado theorem [69] states that if $n \geq 2k$ and $\mathcal{F} \subset \binom{[n]}{k}$ is an intersecting family, then $|\mathcal{F}| \leq \binom{n-1}{k-1}$. When $n > 2k$, equality holds only if $\mathcal{F}$ is a star. In [88], Frankl showed that this theorem still holds for $n \geq dk/(d-1)$ when the intersecting condition is replaced by the d-wise intersecting condition, i.e. every d sets of $\mathcal{F}$ have nonempty intersection.

Theorem 6.1.1 (Frankl [88]). *Let $k \geq d \geq 3$ be fixed and $n \geq dk/(d-1)$. If $\mathcal{F} \subset \binom{[n]}{k}$ is a d-wise intersecting family, then $|\mathcal{F}| \leq \binom{n-1}{k-1}$, with equality only if $\mathcal{F}$ is a star.*

Later, Frankl and Füredi [97] relaxed the intersection condition and proved that for every $n \geq k^2 + 3k$, if $\mathcal{F} \subset \binom{[n]}{k}$ is 3-cluster-free, then $|\mathcal{F}| \leq \binom{n-1}{k-1}$. Moreover, they conjectured that the lower bound for n can be improved to $3k/2$. In [190], Mubayi settled their conjecture, and posed the following more general conjecture.

Conjecture 6.1.2 (Mubayi [190]). *Let $k \geq d \geq 3$ and $n \geq dk/(d-1)$. Suppose that $\mathcal{F} \subset \binom{[n]}{k}$ is d-cluster-free. Then $|\mathcal{F}| \leq \binom{n-1}{k-1}$, with equality only if $\mathcal{F}$ is a* star.

In [197], Mubayi proved Conjecture 6.1.2 for the case $d = 4$ with n sufficiently large. Later, Mubayi and Ramadurai [197], and independently, Füredi and Özkahya [111] proved this conjecture for sufficiently large n. Chen, Liu and Wang [37] proved this conjecture for the case $d = k$. In [185], Mammoliti and Britz showed that this conjecture is true for stable families, i.e. families that are invariant respect to shifting. Very recently, Currier [48] completely solved Conjecture 6.1.2 by proving the following stronger result.

Theorem 6.1.3 (Currier [48]). *Let $2 \le d \le k \le n/2$. Furthermore, suppose $\mathcal{F}^* \subset \mathcal{F} \subset \binom{[n]}{k}$ have the property that any d-cluster in $\mathcal{F}$ is contained entirely in $\mathcal{F}^*$. Then*

$$|\mathcal{F}^*| + \frac{n}{k}|\mathcal{F} - \mathcal{F}^*| \le \binom{n}{k}.$$

Furthermore, excepting the case where both $d = 2$ and $n = 2k$, equality implies one of the following:

1. *$\mathcal{F}^* = \emptyset$ and $\mathcal{F}$ is a maximum-sized star.*
2. *$\mathcal{F} = \mathcal{F}^* = \binom{[n]}{k}$.*

Note that the theorem above indeed implies Conjecture 1.2 since $\mathcal{F}$ is d-cluster-free if and only if $\mathcal{F}^* = \emptyset$, and the case $dk/(d-1) \le n < 2k$ has been settled by Theorem 1.1.

In this paper, we focus on a conjecture raised by Mammoliti and Britz. In [185], they sharpened Conjecture 6.1.2 further by distinguishing the two conditions given by Theorem 6.1.1 and Conjecture 6.1.2, and considered families that are d-cluster-free but that are not d-wise intersecting. In particular, they posed the following conjecture.

Conjecture 6.1.4 (Mammoliti–Britz [185])**.** *For $k \geq d \geq 3$ and sufficiently large n every family $\mathcal{F} \subset \binom{[n]}{k}$ that is d-cluster-free but that is not intersecting has size at most $\binom{n-k-1}{k-1} + 1$, and equality holds only if $\mathcal{F}$ is the disjoint union of a k-set and a star.*

Let $f(n, k, d, \nu)$ denote the maximum size of a d-cluster-free family $\mathcal{F} \subset \binom{[n]}{k}$ with a matching number at least $\nu+1$. Note that by definition $f(n, k, d, 0)$ is the maximum size of a d-cluster-free k-uniform family, and $f(n, k, d, 1)$ is the maximum size of a k-uniform family that is d-cluster-free but not intersecting. Conjecture 1.4 says that $f(n, k, d, 1) \leq \binom{n-k-1}{k-1}+1$ holds for sufficiently large n.

In this section, we mainly consider the function $f(n, k, d, \nu)$ for ν fixed and n sufficiently large. Let g, h be two functions of n. Then $f = o(g)$ means that $\lim_{n\to\infty} f/g = 0$. A lower bound and an upper bound for $f(n, k, d, \nu)$ will be given in the remaining part. The lower bound is given by some constructions, and it is related to the Turán functions on hypergraphs. On the other hand, the proof of the upper bound is based on a stability theorem proved by Mubayi in [192]. So, before stating our results formally, first let us give some definitions.

An r-uniform family is also called an r-graph. We use the term r-graph to emphasize that multiple edges are not allowed in such a hypergraph, and use the term r-multigraph to emphasize that multiple edges are allowed in such a hypergraph. Suppose that $\mathcal{G}$ is an r-multigraph and $E \in \mathcal{G}$ is an edge with multiplicity ℓ, then E is counted ℓ times in $e(\mathcal{G})$. Intuitively, one can view E as a set with ℓ different colors $c_1, ..., c_\ell$, and use (E, c_i) to represent the edge E with color c_i. Pairs $(E, c_i), (E, c_j)$ are considered as different edges in $\mathcal{G}$ if $c_i \neq c_j$.

Definition 6.1.5. *Let $\mathcal{H}_v^e$ to be the collection of all r-multigraphs on v vertices with e edges. Let H_v^e be the collection of r-graphs in $\mathcal{H}_v^e$. An r-multigraph $\mathcal{G}$ is $\mathcal{H}_v^e$-free if it does not contain any element in $\mathcal{H}_v^e$ as a subgraph. An r-graph G is H_v^e-free if it does not contain any element in H_v^e as a subgraph.*

Let $\mathrm{EX}^r(n, \mathcal{H}_v^e)$ denote the maximum number of edges in an n-vertex $\mathcal{H}_v^e$-free r-multigraph. Let $\mathrm{ex}^r(n, H_v^e)$ denote the maximum number of edges in an n-vertex H_v^e-free r-graph. Sometimes we omit the superscript r if there is no cause of any ambiguity.

Let n, r, t, λ be integers and $n \geq r \geq t \geq 0$, $\lambda \geq 1$. A t-(n, r, λ)-design is an r-graph $\mathcal{G}$ on $[n]$ such that for every t-subset T of $[n]$ there are exactly λ members of $\mathcal{G}$ containing T. The existence of certain designs was established by Keevash [136].

For $\ell \geq 1$ and $r \geq 2$ a tight ℓ-path P_ℓ^r is an r-graph with edge set $\{v_i v_{i+1} ... v_{i+r-1} : 1 \leq i \leq \ell\}$. Let $\mathrm{ex}\,(n, P_\ell^r)$ denote the maximum number of edges in an n-vertex P_ℓ^r-free r-graph. Notice that an r-graph $\mathcal{G}$ on $[n]$ is P_2^r-free if and only if every $(r-1)$-subset of $[n]$ is contained in at most one edge in $\mathcal{G}$. Therefore, we have $\mathrm{ex}\,(n, P_2^r) \leq \frac{1}{r}\binom{n}{r-1}$. On the other hand, by results in [136], for infinitely many n, an $(r-1)$-$(n, r, 1)$-design exists and, hence, we know that $\mathrm{ex}\,(n, P_2^r) \geq \frac{1}{r}\binom{n}{r-1}$ holds for infinitely many n.

Now we are ready to state our results formally.

Theorem 6.1.6. *There exist two constants c_1, c_2 that are only related to k, ν and satisfying*

$$\max\left\{(k-1)ex\left(\nu, P_2^3\right), 2(k-1)\left\lfloor\frac{\nu}{2}\right\rfloor\right\} \leq c_1 \leq c_2 \leq \frac{k}{3}\binom{\nu}{2} + (k-1)\nu$$

such that

$$f(n,k,3,\nu) \geq \binom{n-k\nu-1}{k-1} + \left\lfloor \frac{\nu}{2} \right\rfloor \binom{n-k\nu-1}{k-3} + c_1 \binom{n-k\nu-1}{k-4} + \nu \text{ holds for all } n,$$

and

$$f(n,k,3,\nu) \leq \binom{n-k\nu-1}{k-1} + \left\lfloor \frac{\nu}{2} \right\rfloor \binom{n-k\nu-1}{k-3} + (c_2 + o(1)) \binom{n-k\nu-1}{k-4} + M_3$$

holds for sufficiently large n, where M_3 is a constant only related to k, v, and $M_3 \leq f(k\nu, k, 3, \nu-1)$.

Theorem 6.1.7. *There exist two constants $c_1', c_2' \geq k \left\lfloor \frac{\nu^2}{4} \right\rfloor$ such that*

$$f(n,k,4,\nu) \geq \binom{n-k\nu-1}{k-1} + c_1' \binom{n-k\nu-1}{k-3} \text{ holds for all } n,$$

and

$$f(n,k,4,\nu) \leq \binom{n-k\nu-1}{k-1} + c_2' \binom{n-k\nu-1}{k-3} \text{ holds for sufficiently large } n.$$

In particular, if $\nu = 1$, then

$$f(n,k,4,1) \geq \binom{n-k-1}{k-1} + \mathrm{ex}(n-k-1, P_2^{k-2}) + 1 \text{ holds for all } n.$$

Theorem 6.1.8. *Suppose that $d \geq 5$. Then*

$$f(n,k,d,\nu) \geq \binom{n-k\nu-1}{k-1} + \nu \mathrm{EX}^{k-2}\left(n-k\nu-1, \mathcal{H}_{k-1}^{d-2}\right) + \nu \text{ holds for all } n,$$

and

$$f(n,k,d,\nu) \leq \binom{n-k\nu-1}{k-1} + (\nu+o(1))\mathrm{EX}^{k-2}\left(n-k\nu-1, \mathcal{H}_{k-1}^{d-2}\right)$$

holds for sufficiently large n.

For the special case $\nu = 1$, we have the following result.

Theorem 6.1.9. *For sufficiently large n, we have*

$$f(n,k,3,1) = \binom{n-k-1}{k-1} + 1,$$

with equality only for the disjoint union of a k-set and a star.

Theorem 6.1.9 shows that Conjecture 6.1.4 is true for $d = 3$. However, Theorems 6.1.7 and 6.1.7 imply that Conjecture 6.1.4 is false for $d \geq 4$.

Note that in [185] Mammoliti and Britz also asked for the maximum size of a k-uniform family that is d-cluster-free but that is not d-wise intersecting. Let $g(n,k,d,t)$ denote the maximum size of a k-uniform family $\mathcal{F}$ on $[n]$ that is d-cluster-free but not t-wise intersecting. i.e. for all distinct sets $A_1, ..., A_d \in \mathcal{F}$, we have $A_1 \cap ... \cap A_d \neq \emptyset$ whenever $|A_1 \cup ... \cup A_d| \leq 2k$, but there exist t sets $A'_1, ..., A'_t \in \mathcal{F}$ such that $A'_1 \cap ... \cap A'_t = \emptyset$. Later, it will be shown that a family $\mathcal{F}$ that is d-cluster-free but not t-wise intersecting and of large size is actually not intersecting. Therefore, we have the following result.

Theorem 6.1.10. *The equation $g(n,k,d,t) = f(n,k,d,1)$ holds for sufficiently large n.*

6.1.2 Preliminaries

For r-graphs, it is well known that the Turán density $\pi\left(H_v^e\right) = \lim_{n\to\infty} ex\left(n, H_v^e\right)/\binom{n}{r}$ exists, and have the Supersaturation Lemma. A similar result is also true for r-multigraphs.

Lemma 6.1.11. *The limit $\lim_{n\to\infty} \mathrm{EX}^r\left(n, \mathcal{H}_v^e\right)/\binom{n}{r}$ exists.*

Proof. Let $\mathcal{G}$ be an n-vertex $\mathcal{H}_v^e$-free r-multigraph with $\mathrm{EX}(n, \mathcal{H}_v^e)$ edges. Choose an $(n-1)$-subset S of $V(\mathcal{G})$ uniformly at random. For every edge $E \in \mathcal{G}$, the probability that E is contained in S is $(n-r)/n$. So, the expected number of edges in S is $((n-r)/n)\,\mathrm{EX}(n, \mathcal{H}_v^e)$. Therefore, there exists a set S of size $n-1$ with at least $((n-r)/n)\,\mathrm{EX}(n, \mathcal{H}_v^e)$ edges in $\mathcal{G}[S]$. Since $\mathcal{G}[S]$ is also $\mathcal{H}_v^e$-free, we therefore have that $((n-r)/n)\,\mathrm{EX}(n, \mathcal{H}_v^e) \le \mathrm{EX}(n-1, \mathcal{H}_v^e)$. It follows that

$$\frac{\mathrm{EX}(n, \mathcal{H}_v^e)}{\binom{n}{r}} \le \frac{\mathrm{EX}(n-1, \mathcal{H}_v^e)}{\binom{n-1}{r}}.$$

So $\mathrm{EX}(n, \mathcal{H}_v^e)/\binom{n}{r}$ is non-increasing respect to n, and this implies the existence of the limit $\lim_{n\to\infty} \mathrm{EX}^r(n, \mathcal{H}_v^e)/\binom{n}{r}$. ∎

Define the Turán density $\Pi\left(\mathcal{H}_v^e\right)$ of $\mathcal{H}_v^e$ as $\Pi\left(\mathcal{H}_v^e\right) = \lim_{n\to\infty} \mathrm{EX}^r\left(n, \mathcal{H}_v^e\right)/\binom{n}{r}$. Notice that in the proof of Lemma 6.1.11, we showed that $\mathrm{EX}^r\left(n, \mathcal{H}_v^e\right)/\binom{n}{r}$ is non-increasing respect to n. Therefore, we have $\mathrm{EX}^r\left(n, \mathcal{H}_v^e\right) \le \left(\mathrm{EX}^r\left(v, \mathcal{H}_v^e\right)/\binom{v}{r}\right)\binom{n}{r} < \left(e/\binom{v}{r}\right)\binom{n}{r}$. On the other hand, since every H_v^e-free r-graph is also an $\mathcal{H}_v^e$-free r-multigraph, we have $\mathrm{EX}^r(n, \mathcal{H}_v^e) \ge \mathrm{ex}^r(n, H_v^e)$.

Lemma 6.1.12 (Supersaturation)**.** *For any $\mathcal{H}_v^e$ and any $a > 0$, there exist $b > 0$ and n_0 such that any r-multigraph $\mathcal{G}$ on $n > n_0$ vertices with at least $(\Pi(\mathcal{H}_v^e) + a)\binom{n}{r}$ edges contains at least $b\binom{n}{v}$ copies of elements in $\mathcal{H}_v^e$. Moreover, we have $b \geq (a/2)/\binom{M}{v}$, where M is the smallest integer satisfying both $M \geq \max\{r, v\}$ and $\mathrm{EX}(M, \mathcal{H}_v^e) \leq (\Pi(\mathcal{H}_v^e) + a/2)\binom{M}{r}$.*

Let $\mathcal{F} \subset \binom{[n]}{k}$ and $x \in [n]$, define $\mathcal{F}(x) = \{F \in \mathcal{F} \colon x \in F\}$ and $\mathcal{F}(\bar{x}) = \{F \in \mathcal{F} \colon x \notin F\}$. The following stability theorem for d-cluster-free families is an important tool in our proofs.

Theorem 6.1.13 (Stability, [192])**.** *Fix $2 \leq d \leq k$. For every $\delta > 0$, there exists $\epsilon > 0$ and n_0 such that the following holds for all $n > n_0$. Suppose that $\mathcal{F} \subset \binom{[n]}{k}$ is a d-cluster-free family. If $|\mathcal{F}| \geq (1-\epsilon)\binom{n-1}{k-1}$, then there exists a vertex $x \in [n]$ such that $|\mathcal{F}(\bar{x})| < \delta\binom{n-1}{k-1}$.*

Now let $\mathcal{F}$ be a d-cluster-free family with a matching number at least $\nu+1$ and of size exactly $f(n,k,d,\nu)$. In order to apply Theorem 6.1.13 to $\mathcal{F}$, we need a lower bound for $f(n,k,d,\nu)$. So, let us give a simple construction of a d-cluster-free family $\mathcal{S}_\nu$ with a matching number exactly $\nu+1$.

Fix a vertex $y \in [n]$, and choose ν disjoint sets $C_1, ..., C_\nu$ from $\binom{[n]-y}{k}$. Let $J = \bigcup_{i=1}^{\nu} C_i$ and $W = [n] - y - J$. Let

$$\mathcal{S}_\nu = \left\{\{y\} \cup A : A \in \binom{W}{k-1}\right\} \cup \{C_1, ..., C_\nu\}.$$

Note that the size of $\mathcal{S}_\nu$ is $\binom{n-k\nu-1}{k-1} + \nu$. Therefore, we have $f(n,k,d,\nu) \geq \binom{n-k\nu-1}{k-1} + \nu$.

For fixed ν and k we have $\lim_{n\to\infty}\binom{n-k\nu-1}{k-1}/\binom{n-1}{k-1} = 1$. Choose $\delta > 0$ to be sufficiently small, which will be determined later in the proof of Lemma 6.1.16, and let ϵ, n_0 be given by

Theorem 6.1.13. Let n be sufficiently large so that $n > n_0$ and $\binom{n-k\nu-1}{k-1} > (1-\epsilon)\binom{n-1}{k-1}$. By Theorem 6.1.13, there exists a vertex $x \in [n]$ such that $|\mathcal{F}(\bar{x})| < \delta\binom{n-1}{k-1}$. Since $\mathcal{F}$ contains at least $\nu+1$ pairwise disjoint sets, we know that $\mathcal{F}(\bar{x})$ contains at least ν pairwise disjoint sets. So we can choose ν pairwise disjoint sets $B_1, ..., B_\nu$ from $\mathcal{F}(\bar{x})$. Let $I = \bigcup_{i=1}^{\nu} B_i$ and $U = [n] - x - I$. Let $m = |\mathcal{F}(\bar{x})|$ and note that $m < \delta\binom{n-1}{k-1}$. Actually, the following lemmas will show that if $m \geq c\binom{n-1}{k-2}$ holds for some absolute constant $c > 0$, then there exists a d-cluster in $\mathcal{F}$, which contradicts our assumption.

Lemma 6.1.14 ([197])**.** *Fix $2 \leq d \leq k$, $1 \leq p \leq k$, and $k < u_1 \leq n/2$ with n sufficiently large. Suppose that $[n]$ has a partition $U_1 \cup U_2$, $u_1 = |U_1|$, $u_2 = |U_2|$ and $\mathcal{F}$ is a collection of k-sets of $[n]$ such that $|F \cap U_1| = p$ for every $F \in \mathcal{F}$. If $\mathcal{F}$ contains no d-cluster, then $|\mathcal{F}| \leq k u_1^{p-1} u_2^{k-p}$.*

The original form of the next lemma is Claim 1 in [192]. Note that it is assumed in the proof of Claim 1 that the size of $\mathcal{F}$ is at least $\binom{n-1}{k-1}$. However, in our proof, we can only assume that $|\mathcal{F}| \geq \binom{n-k\nu-1}{k-1} + \nu$. So we add an extra assumption that $m \geq c\binom{n-1}{k-2}$ holds for some constant $c > 0$ in the next lemma, and the conclusion is also sightly different from that in Claim 1.

Lemma 6.1.15. *Suppose that $m \geq c\binom{n-1}{k-2}$ holds for some constant $c > 0$. Then, there are pairwise disjoint $(k-2)$-sets $S_1, S_2, S_3 \subset [n] - x$ such that for each i*

$$d_{\mathcal{F}(x)}(S_i) := |\{y \in [n] \colon \{x, y\} \cup S_i \in \mathcal{F}\}| \geq n - k + 1 - \frac{\left(k^2/c + 2k\right) m}{\binom{n-1}{k-2}}.$$

The proof of the next lemma appeared in [197] as a part of the proof of its main theorem. For completeness, we state it formally as a lemma and include its proof in Appendix A.

Lemma 6.1.16 ([197])**.** *Suppose that $m \geq c\binom{n-1}{k-2}$ holds for some constant $c > 0$. Then, there is a d-cluster in $\mathcal{F}$.*

Before presenting our proofs, we would like to remaind the reader that in the proof of the upper bound for $f(n,k,d,\nu)$, we always assume that n is sufficiently large. Our constructions are obtained from $\mathcal{S}_\nu$ by adding some extra k-sets. We will continue using the notations y, J, W and $C_1, ..., C_\nu$ in the lower bound parts, and continue using the notations x, I, U and $B_1, ..., B_\nu$ in the upper bound parts.

6.1.3 Proofs of Theorems 6.1.6 and 6.1.9

The proof of Theorem 6.1.6 is consisting of two parts. In the first part, we present two constructions to give two lower bounds for $f(n,k,3,\nu)$. In the second part, we prove the upper bound for $f(n,k,3,\nu)$.

6.1.3.1 Lower Bound

Recall that the family S_ν is the disjoint union of a star and ν pairwise disjoint k-sets $C_1, ..., C_\nu$.

First construction for $d = 3$.

Choose one vertex v_i from each set C_i. For every $\ell \in \{1, \ldots, \lfloor \nu/2 \rfloor\}$ let $P_\ell = C_{2\ell-1} \cup C_{2\ell}$. For every $i \in \{2, \ldots, k-1\}$ define

$$\mathcal{G}_i = \left\{ A \in \bigcup_{\ell=1}^{\lfloor \nu/2 \rfloor} \binom{P_\ell}{i} : \{v_{2\ell-1}, v_{2\ell}\} \subset A \text{ for some } \ell \right\}.$$

Let

$$\mathcal{L}_1 = \mathcal{S}_\nu \cup \left(\bigcup_{i=2}^{k-1} \left\{ \{y\} \cup A \cup B \colon B \in \binom{W}{k-1-i} \text{ and } A \in \mathcal{G}_i \right\} \right).$$

Note that the size of $\mathcal{G}_i$ is $\lfloor \frac{\nu}{2} \rfloor \binom{2k-2}{i-2}$ for all $i \in \{2, ..., k-1\}$. Therefore, we have

$$|\mathcal{L}_1| = \binom{n-k\nu-1}{k-1} + \sum_{i=2}^{k-1} \left\lfloor \frac{\nu}{2} \right\rfloor \binom{2k-2}{i-2} \binom{n-k\nu-1}{k-1-i} + \nu.$$

Since $\mathcal{L}_1$ is a 3-cluster-free family with $\nu(\mathcal{L}_1) = \nu + 1$, we therefore have that

$$f(n,k,3,\nu) \geq \binom{n-k\nu-1}{k-1} + \sum_{i=2}^{k-1} \left\lfloor \frac{\nu}{2} \right\rfloor \binom{2k-2}{i-2} \binom{n-k\nu-1}{k-1-i} + \nu.$$

Second construction for $d = 3$.

Suppose that $C_i = \{c_1^i, \ldots, c_k^i\}$ for $1 \leq i \leq \nu$. Then let $V_j = \{c_j^1, \ldots, c_j^\nu\}$ for every $j \in [k]$. Let G_1 be the graph on V_1 with edge set $\left\{\{c_1^{2i-1}, c_1^{2i}\} \colon 1 \leq i \leq \lfloor \nu/2 \rfloor \right\}$. For every $j \in \{2, \ldots, k\}$ let $\mathcal{G}_j$ be a P_2^3-free 3-graph on V_j with exactly $ex(\nu, P_2^3)$ edges. Let

$$\mathcal{L}_2' = \mathcal{S}_\nu \cup \left\{ \{y\} \cup A \cup B \colon B \in \binom{W}{k-3} \text{ and } A \in E(G_1) \right\}.$$

Then let

$$\mathcal{L}_2 = \mathcal{L}_2' \cup \left(\bigcup_{j=2}^{k} \left\{ \{y\} \cup A \cup B \colon B \in \binom{W}{k-4} \text{ and } A \in \mathcal{G}_j \right\} \right).$$

It is easy to see that

$$|\mathcal{L}_2| = \binom{n-k\nu-1}{k-1} + \left\lfloor \frac{\nu}{2} \right\rfloor \binom{n-k\nu-1}{k-3} + (k-1)\mathrm{ex}(\nu, P_2^3)\binom{n-k\nu-1}{k-4} + \nu.$$

Since $\mathcal{L}_2$ is a 3-cluster-free family with $\nu(\mathcal{L}_2) = \nu + 1$, we therefore have that

$$f(n,k,3,\nu) \geq \binom{n-k\nu-1}{k-1} + \left\lfloor \frac{\nu}{2} \right\rfloor \binom{n-k\nu-1}{k-3} + (k-1)\mathrm{ex}(\nu, P_2^3)\binom{n-k\nu-1}{k-4} + \nu.$$

6.1.3.2 Upper Bound

First we claim that $|F \cap F'| \leq k-2$ holds for every $F \in \mathcal{F}(x)$ and every $F' \in \mathcal{F}(\bar{x})$. Indeed, suppose that there exists an edge $F \in \mathcal{F}(x)$ and an edge $F' \in \mathcal{F}(\bar{x})$ such that $|F \cap F'| = k-1$. Then for every set $S \in \binom{[n]-x-F'}{k-1}$ we have $\{x\} \cup S \notin \mathcal{F}$, since otherwise $\{x\} \cup S, F$ and F' would form a 3-cluster, a contradiction. So in this case we would have

$$|\mathcal{F}| \leq \binom{n-1}{k-1} - \binom{n-k-1}{k-1} + \delta\binom{n-1}{k-1} < \binom{n-k\nu-1}{k-1},$$

and this contradicts our assumption that $\mathcal{F}$ is of size $f(n,k,3,\nu)$.

Let M_3 be the maximum possible number of sets in $\mathcal{F}$ that are completely contained in I, and it is easy to see that $M_3 \leq f(k\nu, k, 3, \nu-1)$. For every subset C of U that of size at most $k-2$ let

$$\mathcal{F}'(C) = \{F - x - C \colon F \in \mathcal{F}(x) \text{ and } F \cap U = C\},$$

and let $\mathcal{F}(C) = \mathcal{F}'(C) - \bigcup_{i=1}^{\nu} \binom{B_i}{k-1-|C|}$. For every $j \in \{0, \ldots, k-1\}$ let

$$\mathcal{F}_j = \{F \in \mathcal{F}(x) \colon |F \cap I| = j\}.$$

Intuitively, one can view $\mathcal{F}'(C)$ as the collection of neighbors of C in I, and view $|\mathcal{F}'(C)|$ as the degree of C in $\mathcal{F}(x)$. Our goal is to give an upper bound for $|\mathcal{F}(x)|$, and this is done by giving an upper bound for each $|\mathcal{F}'(C)|$.

Lemma 6.1.17. *Let $C \in \binom{U}{k-3}$. Then $|\mathcal{F}(C)| \leq \left\lfloor \frac{\nu}{2} \right\rfloor$.*

Proof. Let $C \in \binom{U}{k-3}$ and let G denote the graph $\mathcal{F}(C)$. Note that G is a graph on I. By the definition of $\mathcal{F}(C)$, we know that B_i is an independent set in G for $1 \leq i \leq \nu$.

For every pair $\{i, j\} \subset \{1, \ldots, \nu\}$ let $E(B_i, B_j)$ denote the collection of edges in G that have one endpoint in B_i and the other endpoint in B_j, and let $e(B_i, B_j)$ denote the size of $E(B_i, B_j)$. First, we claim that $e(B_i, B_j) \leq 1$ for every pair $\{i, j\} \subset \{1, \ldots, \nu\}$. Indeed, suppose that there are two edges $e_1, e_2 \in E(B_i, B_j)$ for some pair $\{i, j\}$. Assume that $e_1 = \{b_1^i, b_1^j\}, e_2 = \{b_2^i, b_2^j\}$ and $b_1^i, b_2^i \in B_i, b_1^j, b_2^j \in B_j$. We may assume that $b_1^i \neq b_2^i$, otherwise we consider b_1^j and b_2^j instead. However, the three sets $B_i, \{x, b_1^i, b_1^j\} \cup C$ and $\{x, b_2^i, b_2^j\} \cup C$ form a 3-cluster, a contradiction. Therefore, we have $e(B_i, B_j) \leq 1$. Next, we show that for every $i \in \{1, \ldots, \nu\}$ there is at most one edge that has nonempty intersection with B_i. Indeed, suppose there are two edges e_1, e_2 such that $e_1 \cap B_i \neq \emptyset$ and $e_2 \cap B_i \neq \emptyset$. Assume that $e_1 = \{b_1^i, b_1^j\}, e_2 = \{b_2^i, b_2^k\}$ and $b_1^i, b_2^i \in B_i, b_1^j \in B_j, b_2^k \in B_k$. By the argument above, we know that $j \neq k$. However, if $b_1^i \neq b_2^i$, then $B_i, \{x, b_1^i, b_1^j\} \cup C$ and $\{x, b_2^i, b_2^k\} \cup C$ form a 3-cluster, a contradiction. If $b_1^i = b_2^i$, then $B_j, \{x, b_1^i, b_1^j\} \cup C$ and $\{x, b_2^i, b_2^k\} \cup C$ form a 3-cluster, a contradiction. So every B_i has nonempty intersection with at most one edge of G. Therefore, we have $|\mathcal{F}(C)| = e(G) \leq \lfloor \nu/2 \rfloor$.

∎

Assume that $k \geq 4$. Let $C \in \binom{U}{k-4}$ and view $\mathcal{F}(C)$ as a 3-graph on I. By the definition of $\mathcal{F}(C)$, every $E \in \mathcal{F}(C)$ has nonempty intersection with at least two sets in $\{B_1, \ldots, B_\nu\}$. We call E a long edge if E has nonempty intersection with three sets in $\{B_1, \ldots, B_\nu\}$, otherwise we call E a short edge. Let $\mathcal{L}_c$ be the collection of all long edges in $\mathcal{F}(C)$ and let $\mathcal{S}_c$ be the collection of all short edges in $\mathcal{F}(C)$. For every $i \in [\nu]$ let G_i be the graph on B_i with edge set $\partial \mathcal{S}_c \cap \binom{B_i}{2}$. For every pair $\{i, j\} \subset [\nu]$ let $G_{i,j}$ be the bipartite graph on $B_i \cup B_j$ with edge set $\partial \mathcal{L}_c \cap \binom{B_i \cup B_j}{2}$.

Claim 6.1.18. *The matching number of G_i is at most one for every $i \in [\nu]$.*

Proof. Suppose there are two vertex disjoint edges e_1, e_2 in $E(G_i)$ for some $i \in [\nu]$. By the definition of $E(G_i)$, there exist two sets $S_1, S_2 \in \mathcal{S}_c$ such that $S_1 \cap B_i = e_1$ and $S_2 \cap B_i = e_2$. However, the three sets $B_i, \{x\} \cup C \cup S_1$ and $\{x\} \cup C \cup S_2$ form a 3-cluster in $\mathcal{F}$, a contradiction. Therefore, the matching number of G_i is at most one. ∎

Claim 6.1.19. *For every $i \in [\nu]$ and every $e \in E(G_i)$ there is exactly one set $S \in \mathcal{S}_c$ such that $S \cap B_i = e$.*

Proof. Suppose that there exist two vertices $v_1 \in B_j$ and $v_2 \in B_k$ for some j, k such that $S_1 = \{v_1\} \cup e$ and $S_2 = \{v_2\} \cup e$ are both contained in $\mathcal{S}_c$. Here $j \neq i$ and $k \neq i$ but j, k might be the same. However, the three sets $B_j, \{x\} \cup C \cup S_1$ and $\{x\} \cup C \cup S_2$ form a 3-cluster in $\mathcal{F}$, a contradiction. Therefore, there is exactly one set $S \in \mathcal{S}_c$ such that $S \cap B_i = e$. ∎

Claim 6.1.18 implies that the size of $E(G_i)$ is at most $k-1$ for every $i \in [\nu]$. Combining Claim 6.1.18 with Claim 6.1.19, we obtain that $|\mathcal{S}_c| = \sum_{i=1}^{\nu} |E(G_i)| \leq (k-1)v$. Next, we will give an upper bound for $|\mathcal{L}_c|$.

Claim 6.1.20. *For every pair $\{i, j\} \subset [\nu]$ every vertex in $G_{i,j}$ has degree at most* 1.

Proof. Suppose that there exist two edges $e_1, e_2 \in E(G_{i,j})$ for some pair $\{i, j\} \subset [\nu]$ such that $e_1 \cap e_2 \neq \emptyset$. Without loss of generality, we may assume that the common endpoint of e_1, e_2 lies in B_i. By the definition of $G_{i,j}$, there exist two sets $S_1, S_2 \in \mathcal{L}_c$ such that $e_1 \subset S_1$ and $e_2 \subset S_2$. However, the three sets $B_j, \{x\} \cup C \cup S_1$ and $\{x\} \cup C \cup S_2$ form a 3-cluster in $\mathcal{F}$, a contradiction. Therefore, every vertex in $G_{i,j}$ has degree at most 1. ∎

Claim 6.1.21. *For every $e \in E(G_{i,j})$ there is exactly one set $S \in \mathcal{L}_c$ containing e.*

Proof. Suppose there exist two vertices $v_1 \in B_k$ and $v_2 \in B_\ell$ such that $S_1 = \{v_1\} \cup e$ and $S_2 = \{v_2\} \cup e$ are both contained in $\mathcal{L}_c$. Here $k, \ell \notin \{i, j\}$ but k, ℓ might be the same. However, $B_k, \{x\} \cup C \cup S_1$ and $\{x\} \cup C \cup S_2$ form a 3-cluster in $\mathcal{F}$, a contradiction. Therefore, there is exactly one set in $\mathcal{L}_c$ that contains e. ∎

Claim 6.1.20 implies that $|E(G_{i,j})| \leq k$ for every pair $\{i, j\} \subset [\nu]$. Combining Claim 6.1.20 with Claim 6.1.21, we obtain that $|\mathcal{L}_c| = \frac{1}{3} \sum_{1 \leq i < j \leq \nu} |E(G_{i,j})| \leq \frac{k}{3}\binom{\nu}{2}$. Since $|\mathcal{F}(C)| = |\mathcal{S}_c| + |\mathcal{L}_c|$, we therefore obtain the following lemma.

Lemma 6.1.22. *Suppose that $k \geq 4$ and $C \in \binom{U}{k-4}$. Then $|\mathcal{F}(C)| \leq \frac{k}{3}\binom{\nu}{2} + (k-1)v$.*

Now we are ready to prove the upper bound for Theorem 6.1.6.

Proof of Theorem 6.1.6. **Case 1:** the family $\mathcal{F}(\bar{x})$ is completely contained in $\binom{I}{k}$.

For every $j \in [k-2]$ define $\mathcal{B}_j = \bigcup_{i=1}^{\nu} \binom{B_i}{j}$, and let

$$\mathcal{G}_j = \left\{ A \in \binom{U}{j} : \exists B \in \mathcal{B}_{k-1-j} \text{ such that } \{x\} \cup A \cup B \in \mathcal{F} \right\}.$$

Let $S \in \binom{U}{k-1}$, we say that S is *bad* if it contains an edge $E \in \mathcal{G}_j$ for some $j \in [k-2]$. Note that if S is bad, then $\{x\} \cup S \notin \mathcal{F}$, since otherwise there would be a set B contained in B_i for some i such that $F = \{x\} \cup E \cup B$ is contained in $\mathcal{F}$. However, the three sets B_i, F and $\{x\} \cup S$ form a 3-cluster in $\mathcal{F}$, a contradiction.

For every $j \in [k-2]$ let g_j denote the size of $\mathcal{G}_j$. Let β denote the number of bad sets in $\binom{U}{k-1}$. Let $E \in \mathcal{G}_j$. Then for every $A \in \binom{U-E}{k-1-j}$, we know that $A \cup E$ is a bad set in $\binom{U}{k-1}$. Therefore, we have $\beta \geq \frac{1}{2^{2k}} \sum_{i=1}^{k-2} g_i \binom{|U|-i}{k-1-i}$.

For every $j \in [k-1]$, we have $|\mathcal{F}_j| \leq \binom{|I|}{j}\binom{|U|}{k-1-j}$. Therefore, we obtain $\sum_{j=4}^{k-1} |\mathcal{F}_j| = o(1)\binom{|U|}{k-4}$. Let $c' = \frac{k}{3}\binom{\nu}{2} + (k-1)\nu$, by Lemmas 6.1.17 and 6.1.22, we have

$$\begin{aligned} |\mathcal{F}| &= \sum_{i=0}^{k-1} |\mathcal{F}_i| + |\mathcal{F}(\bar{x})| \\ &\leq \binom{|U|}{k-1} - \beta + 2^k \nu \sum_{i=1}^{k-2} g_i + \left\lfloor \frac{\nu}{2} \right\rfloor \binom{|U|}{k-3} + (c' + o(1))\binom{|U|}{k-4} + M_3. \end{aligned}$$

For every $j \in [k-2]$, we have $\frac{1}{2^{2k}}\binom{|U|-i}{k-1-j} > 2^k\nu$. Therefore, we have $-\beta + 2^k\nu\sum_{i=1}^{k-2} g_i \leq 0$ and, hence, we obtain

$$\begin{aligned}|\mathcal{F}| &\leq \binom{|U|}{k-1} + \left\lfloor\frac{\nu}{2}\right\rfloor\binom{|U|}{k-3} + (c' + o(1))\binom{|U|}{k-4} + M_3\\ &= \binom{n-k\nu-1}{k-1} + \left\lfloor\frac{\nu}{2}\right\rfloor\binom{n-k\nu-1}{k-3} + (c' + o(1))\binom{n-k\nu-1}{k-4} + M_3.\end{aligned}$$

Case 2: the family $\mathcal{F}(\bar{x})$ is not completely contained in $\binom{I}{k}$.

Then there exists a set $B_{\nu+1} \in \mathcal{F}(\bar{x})$ such that $B_{\nu+1} - I \neq \emptyset$. Now let $I' = I \cup B_{\nu+1}$ and $U' = [n] - x - I'$. Let

$$\mathcal{G} = \left\{E \in \binom{U'}{k-2} : \exists b \in I' \text{ such that } \{x, b\} \cup E \in \mathcal{F}\right\}.$$

Let $S \in \binom{U'}{k-1}$, we say that S is *bad* if it contains an edge $E \in \mathcal{G}$. Note that if S is bad, then $\{x\} \cup S \notin \mathcal{F}$, since otherwise there would be a vertex b contained in B_i for some i such that $\{x, b\} \cup E \in \mathcal{F}$. However, the three sets $B_i, \{x, b\} \cup E$ and $\{x\} \cup S$ form a 3-cluster in $\mathcal{F}$, a contradiction.

Let g denote the size of $\mathcal{G}$ and let β denote the number of bad sets in $\binom{U'}{k-1}$. Let $E \in \mathcal{G}$. Then for every $v \in U' - E$, we know that $\{v\} \cup E$ is a bad set in $\binom{U'}{k-1}$. So we have $\beta \geq \frac{|U'|-k+2}{k-1} g$.

Let $\mathcal{F}_{\geq 2} = \{F \in \mathcal{F}(x) \colon |F \cap I'| \geq 2\}$, and note that $|\mathcal{F}_{\geq 2}| \leq \sum_{i=2}^{k-1} \binom{|I'|}{i}\binom{|U'|}{k-1-i} < \frac{1}{2}\binom{|U'|}{k-2}$. Therefore, we have

$$|\mathcal{F}| = |\mathcal{F}(x)| + |\mathcal{F}(\bar{x})| \leq \binom{|U'|}{k-1} - \beta + g|I'| + \frac{1}{2}\binom{|U'|}{k-2} + m$$
$$\leq \binom{|U'|}{k-1} - \left(\frac{|U'|-k+2}{k-1} - |I'|\right) g + \frac{1}{2}\binom{|U'|}{k-2} + m.$$

Since $\frac{|U'|-k+2}{k-1} > |I'|$ and $|U'| \leq n - k\nu - 2$, we therefore have that

$$|\mathcal{F}| \leq \binom{n-k\nu-2}{k-1} + \frac{1}{2}\binom{n-k\nu-2}{k-2} + m = \binom{n-k\nu-1}{k-1} - \frac{1}{2}\binom{n-k\nu-2}{k-2} + m.$$

By the assumption that $|\mathcal{F}| = f(n,k,3,\nu)$, we obtain $m \geq \frac{1}{2}\binom{n-k\nu-2}{k-2} \geq \frac{1}{4}\binom{n-1}{k-2}$. However, Lemma 6.1.16 implies that $\mathcal{F}$ contains a 3-cluster, a contradiction. Therefore, Case 2 is impossible and, hence, we obtain

$$f(n,k,3,\nu) \leq \binom{n-k\nu-1}{k-1} + \left\lfloor\frac{\nu}{2}\right\rfloor\binom{n-k\nu-1}{k-3} + (c' + o(1))\binom{n-k\nu-1}{k-4} + M_3.$$

∎

6.1.3.3 Proof of Theorem 6.1.9

Proof of Theorem 6.1.9. Let C be a subset of U that of size at most $k-2$. Since $\nu = 1$, every set in $\mathcal{F}'(C)$ is contained in B_1 and, hence, we have $\mathcal{F}(C) = \emptyset$. Note that in the argument above, we already showed that Case 2 is impossible. Therefore, it suffices to only consider Case 1 and, hence, we obtain

$$f(n,k,3,1) = |\mathcal{F}| \leq \binom{|U|}{k-1} - \beta + 2^k \sum_{i=1}^{k-2} g_i + 1 \leq \binom{n-k-1}{k-1} + 1,$$

and equality holds only if $g_i = 0$ holds for every $i \in [k-2]$, i.e., $\mathcal{F}$ is the disjoint union of a k-set and a star. ∎

6.1.4 Proof of Theorem 6.1.7

6.1.4.1 Lower Bound

A construction for $\nu = 1$.

Let $k' = k-2$ and $n' = n - k\nu - 1$ for short. Let $\mathcal{G}$ be a $P_2^{k'}$-free k'-graph on W with exactly $ex(n', P_2^{k'})$ edges. Let $v \in J$ be fixed and define

$$\mathcal{L}_3 = \mathcal{S}_\nu \cup \{\{y,v\} \cup A : A \in \mathcal{G}\}.$$

It is easy to see that

$$|\mathcal{L}_3| = \binom{n-k-1}{k-1} + \mathrm{ex}(n', P_2^{k'}) + 1.$$

Since $\mathcal{L}_3$ is 4-cluster-free and $\nu(\mathcal{L}_3) = 2$, we therefore have

$$f(n,k,4,1) \geq \binom{n-k-1}{k-1} + \mathrm{ex}(n', P_2^{k'}) + 1.$$

A construction for $\nu \geq 2$.

Let $\mathcal{C}_\ell = \{C_1, \ldots, C_{\lfloor \nu/2 \rfloor}\}$ and $\mathcal{C}_r = \{C_{\lfloor \nu/2 \rfloor + 1}, \ldots, C_\nu\}$. For every pair (C_i, C_j) with $C_i \in \mathcal{C}_\ell$ and $C_j \in \mathcal{C}_r$ add k vertex disjoint edges between C_i and C_j, and let G denote the resulting graph. Note that the number of edges in G is $k \left\lfloor \nu^2/4 \right\rfloor$. Let

$$\mathcal{L}_4 = \mathcal{S}_\nu \cup \left\{ \{y\} \cup e \cup B \colon B \in \binom{W}{k-3} \text{ and } e \in E(G) \right\}.$$

Then, it is easy to see that

$$|\mathcal{L}_4| = \binom{n - k\nu - 1}{k-1} + k \left\lfloor \frac{\nu^2}{4} \right\rfloor \binom{n - k\nu - 1}{k-3} + \nu.$$

Since $\mathcal{L}_4$ is 4-cluster-free and $\nu(\mathcal{L}_4) = \nu + 1$, we therefore have that

$$f(n, k, 4, \nu) \geq \binom{n - k\nu - 1}{k-1} + k \left\lfloor \frac{\nu^2}{4} \right\rfloor \binom{n - k\nu - 1}{k-3} + \nu.$$

6.1.5 Upper Bound

Let M_4 be the maximum possible number of sets in $\mathcal{F}$ that are completely contained in I, and it is easy to see that $M_4 \leq f(k\nu, k, 4, \nu - 1)$.

Proof of Theorem 6.1.7. **Case 1:** the family $\mathcal{F}(\bar{x})$ is completely contained in $\binom{I}{k}$.

For every $i \in [\nu]$ let

$$\mathcal{G}_i = \left\{ A \in \binom{U}{k-2} : \exists b \in B_i \text{ such that } \{x, b\} \cup A \in \mathcal{F} \right\},$$

and let g_i denote the size of $\mathcal{G}_i$. Without loss of generality, we may assume that $g_1 \geq \cdots \geq g_\nu$. Let $\mathcal{P}_2$ be the collection of all tight 2-paths in $\mathcal{G}_1$. Then we have

$$|\mathcal{P}_2| = \sum_{E \in \binom{U}{k-3}} \binom{d_{\mathcal{G}_1}(E)}{2} \geq \binom{|U|}{k-3} \binom{\sum d_{\mathcal{G}_1}(E)/\binom{|U|}{k-3}}{2} = \frac{(k-2)g_1}{2} \left(\frac{(k-2)g_1}{\binom{|U|}{k-3}} - 1 \right).$$

Let $S \in \binom{U}{k-1}$, we say that S is bad if it contains at least two sets $E_1, E_2 \in \mathcal{G}_i$ for some i. Note that if S is bad, then $\{x\} \cup S \notin \mathcal{F}$, since otherwise there would be two vertices $b_1, b_2 \in B_i$ such that $\{x, b_1\} \cup E_1, \{x, b_2\} \cup E_2$ are both contained in $\mathcal{F}$. However, the four sets $B_i, \{x\} \cup S, \{x, b_1\} \cup E_1$ and $\{x, b_2\} \cup E_2$ form a 4-cluster in $\mathcal{F}$, a contradiction.

Let β denote the number of bad sets. Since every tight 2-path in $\mathcal{G}_1$ forms a bad set, we have $\beta \geq \frac{g_1}{k-1} \left((k-2)g_1/\binom{|U|}{k-3} - 1 \right)$.

Let $\mathcal{F}_{\geq 2} = \{F \in \mathcal{F}(x) \colon |F \cap I| \geq 2\}$. Then there exists a constant c such that $|\mathcal{F}_{\geq 2}| \leq \sum_{i=2}^{k-1} \binom{|I|}{i} \binom{|U|}{k-1-i} \leq c\binom{|U|}{k-3}$.

For every $E \in \mathcal{G}_i$ there are at most two vertices b_1, b_2 in B_i such that $\{x, b_1\} \cup E, \{x, b_2\} \cup E \in \mathcal{F}$. Indeed, suppose there are three vertices $b_1, b_2, b_3 \in B_i$ such that $\{x, b_1\} \cup E, \{x, b_2\} \cup E, \{x, b_3\} \cup E$ are all contained in $\mathcal{F}$. Then the four sets $\{x, b_1\} \cup E, \{x, b_2\} \cup E, \{x, b_3\} \cup E$ and B_i would form a 4-cluster in $\mathcal{F}$, a contradiction. Therefore, we have

$$\begin{aligned} |\mathcal{F}| = |\mathcal{F}(x)| + |\mathcal{F}(\bar{x})| &\leq \binom{|U|}{k-1} - \beta + \sum_{i=1}^{\nu} 2g_i + c\binom{|U|}{k-3} + M_4 \\ &\leq \binom{n-k\nu-1}{k-1} + 2\nu g_1 - \frac{g_1}{k-1} \left(\frac{(k-2)g_1}{\binom{|U|}{k-3}} - 1 \right) + c\binom{n-k\nu-1}{k-3} + M_4. \end{aligned}$$

Viewing g_1 as a variable to obtain that $2\nu g_1 - \frac{g_1}{k-1}\left((k-2)g_1/\binom{|U|}{k-3} - 1\right) \leq \frac{(2\nu(k-1)+1)^2}{4(k-1)(k-2)}\binom{|U|}{k-3}$. Since $\frac{(2\nu(k-1)+1)^2}{4(k-1)(k-2)}$ is a constant only related to k and ν, we obtain that

$$|\mathcal{F}| \leq \binom{n-k\nu-1}{k-1} + c_2'\binom{n-k\nu-1}{k-3} + M_4,$$

where c_2' is a constant only related to k and ν.

Case 2: the family $\mathcal{F}(\bar{x})$ is not completely contained in $\binom{I}{k}$.

Then there exists a set $B_{\nu+1} \in \mathcal{F}(\bar{x})$ such that $B_{\nu+1} - I \neq \emptyset$. Now let $I' = I \cup B_{\nu+1}$ and $U' = [n] - x - I'$. For every $i \in [\nu+1]$ let

$$\mathcal{G}_i = \left\{A \in \binom{U'}{k-2} : \exists b \in B_i \text{ such that } \{x, b\} \cup A \in \mathcal{F}\right\},$$

and let g_i denote the size of $\mathcal{G}_i$. We may assume that $g_1 \geq \cdots \geq g_{\nu+1}$. Let $\mathcal{P}_2$ be the collection of all tight 2-paths in $\mathcal{G}_1$. Then $|\mathcal{P}_2| \geq \frac{(k-2)g_1}{2}\left((k-2)g_1/\binom{|U'|}{k-3} - 1\right)$.

Let $S \in \binom{U'}{k-1}$, we say that S is bad if S contains two edges E_1, E_2 in $\mathcal{G}_i$ for some i. Note that if S is bad, then $\{x\} \cup S \notin \mathcal{F}$. Let β denote the number of bad sets in $\binom{U'}{k-1}$. Since every tight 2-path in $\mathcal{G}_1$ forms a bad set, we have $\beta \geq \frac{g_1}{k-1}\left((k-2)g_1/\binom{|U'|}{k-3} - 1\right)$.

Let $\mathcal{F}_{\geq 2} = \{F \in \mathcal{F}(x) \colon |F \cap I'| \geq 2\}$ and note that there exists a constant c such that $|\mathcal{F}_{\geq 2}| \leq \sum_{i=2}^{k-1} \binom{|I'|}{i}\binom{|U'|}{k-1-i} \leq c\binom{|U'|}{k-3}$. Therefore, we have

$$\begin{aligned}
|\mathcal{F}| &\leq \binom{|U'|}{k-1} - \beta + \sum_{i=1}^{\nu+1} 2g_i + c\binom{|U'|}{k-3} + m \\
&\leq \binom{|U'|}{k-1} + 2(\nu+1)g_1 - \frac{g_1}{k-1}\left(\frac{(k-2)g_1}{\binom{|U'|}{k-3}} - 1\right) + c\binom{|U'|}{k-3} + m.
\end{aligned}$$

Since $2(\nu+1)g_1 - \frac{g_1}{k-1}\left((k-2)g_1/\binom{|U'|}{k-3} - 1\right) \leq \frac{(2(\nu+1)(k-1)+1)^2}{4(k-1)(k-2)}\binom{|U'|}{k-3}$, there exists a constant c' such that

$$|\mathcal{F}| \leq \binom{|U'|}{k-1} + c'\binom{|U'|}{k-3} + m \leq \binom{n-k\nu-1}{k-1} - \binom{n-k\nu-2}{k-2} + c'\binom{n-k\nu-2}{k-3} + m.$$

By the assumption that $|\mathcal{F}| = f(n,k,4,\nu)$, we have $m > \binom{n-k\nu-2}{k-2} - c'\binom{n-k\nu-2}{k-3} \geq \frac{1}{2}\binom{n-1}{k-2}$. However, Lemma 6.1.16 implies that $\mathcal{F}$ contains a 4-cluster, a contradiction. Therefore, Case 2 is impossible and, hence, there exists a constant c_2 such that

$$f(n,k,4,\nu) \leq \binom{n-k\nu-1}{k-1} + c_2\binom{n-k\nu-1}{k-3}.$$

∎

6.1.6 Proof of Theorem 6.1.8

Let $k' = k-2$ and $n' = n - k\nu - 1$ for short.

6.1.6.1 Lower Bound

Let $\mathcal{G}$ be an n'-vertex $\mathcal{H}_{k-1}^{d-2}$-free k'-multigraph on W with exactly $EX^{k'}\left(n', \mathcal{H}_{k-1}^{d-2}\right)$ edges. Let $E \in \mathcal{G}$ be an edge of multiplicity ℓ. For every $i \in [\nu]$ choose ℓ distinct vertices $c_1^i, \ldots, c_\ell^i$ from C_i and add $\{y, c_1^i\} \cup E, \ldots, \{y, c_l^i\} \cup E$ into $\mathcal{S}_\nu$. Let $\mathcal{L}_5$ denote the resulting family. It is easy to see that

$$|\mathcal{L}_5| = \binom{n-k\nu-1}{k-1} + \nu \mathrm{EX}^{k'}\left(n', \mathcal{H}_{k-1}^{d-2}\right) + \nu.$$

When $d \geq 5$, every k'-graph in H_{k-1}^{d-2} is nondegenerate, i.e. the Turán density $\pi\left(H_{k-1}^{d-2}\right)$ of H_{k-1}^{d-2} is not 0 (we refer the reader to [135] for more details), we therefore have that $\Pi\left(\mathcal{H}_{k-1}^{d-2}\right) \geq \pi\left(H_{k-1}^{d-2}\right) \geq \frac{(k-2)!}{(k-2)^{k-2}}$. Since $\mathcal{L}_5$ is a d-cluster-free family with $\nu(\mathcal{L}_5) = \nu + 1$, we therefore have that

$$f(n,k,d,\nu) \geq \binom{n-k\nu-1}{k-1} + \nu EX^{k'}\left(n', \mathcal{H}_{k-1}^{d-2}\right) + \nu.$$

6.1.6.2 Upper Bound

Let M_d be the maximum possible number of sets in $\mathcal{F}$ that are completely contained in I, and it is easy to see that $M_d \leq f(k\nu, k, d, \nu - 1)$.

Proof of Theorem 6.1.8. **Case 1:** the family $\mathcal{F}(\bar{x})$ is completely contained in $\binom{I}{k}$.

For every $i \in [\nu]$ define the k'-multigraph $\mathcal{G}_i$ on U as

$$\mathcal{G}_i = \left\{ E \in \binom{U}{k-2} : \exists b \in B_i \text{ such that } \{x, b\} \cup E \in \mathcal{F} \right\}.$$

Let $E \in \mathcal{G}_i$. Then the multiplicity of E is the number of vertices b in B_i such that $\{x, b\} \cup E \in \mathcal{F}$. For every $i \in [\nu]$ let g_i denote the number of edges in $\mathcal{G}_i$. Without loss of generality, we may assume that $g_1 \geq \cdots \geq g_\nu$.

Let $S \in \binom{U}{k-1}$, we say that S is bad if $\mathcal{G}_i[S] \in \mathcal{H}_{k-1}^{d-2}$ holds for some i. Note that if S is bad, then $\{x\} \cup S \notin \mathcal{F}$, since otherwise there would be $d-2$ edges $E_1, \ldots, E_{d-2}$ in $\mathcal{G}_i$ for some i such that they are all contained in S. By the definition of $\mathcal{G}_i$, there exist $d-2$ vertices $b_1, \ldots, b_{d-2} \in B_i$ such that $\{x, b_1\} \cup E_1, \ldots, \{x, b_{d-2}\} \cup E_{d-2}$ are all contained in $\mathcal{F}$. However, the d sets $B_i, \{x\} \cup S, \{x, b_1\} \cup E_1, \ldots, \{x, b_{d-2}\} \cup E_{d-2}$ form a d-cluster in $\mathcal{F}$, a contradiction.

Let $\mathcal{F}_{\geq 2} = \{F \in \mathcal{F}(x) \colon |F \cap I| \geq 2\}$. Then there exists a constant c such that $|\mathcal{F}_{\geq 2}| \leq \sum_{i=2}^{k-1} \binom{|I|}{i}\binom{|U|}{k-1-i} \leq c\binom{|U|}{k-3}$. Therefore, we have

$$|\mathcal{F}| \leq \binom{|U|}{k-1} - \beta + \sum_{i=1}^{\nu} g_i + c\binom{|U|}{k-3} + M_d.$$

If $g_1 \leq (1+o(1))\mathrm{EX}^{k'}\left(n', \mathcal{H}_{k-1}^{d-2}\right)$, then we are done. Therefore, we may assume that $g_1 = (1+a)\mathrm{EX}^{k-2}\left(n', \mathcal{H}_{k-1}^{d-2}\right)$ with $a \geq 2\sigma$ holds for some absolute constant $\sigma > 0$. By Lemma 6.1.12, the graph $\mathcal{G}_1$ contains at least $\frac{a/2}{\binom{N}{k-1}}\binom{|U|}{k-1}$ copies of elements in $\mathcal{H}_{k-1}^{d-2}$, where N is the smallest integer satisfying both $\mathrm{EX}^{k'}\left(N, \mathcal{H}_{k-1}^{d-2}\right) \leq (1+\sigma)\Pi\left(\mathcal{H}_{k-1}^{d-2}\right)\binom{N}{k-2}$ and $N \geq k-1$.

Let β denote the number of bad sets. Since every copy of element in $\mathcal{H}_{k-1}^{d-2}$ forms a bad set in $\binom{U}{k-1}$, we therefore have

$$\beta \geq \frac{1}{(k-1)^{d-2}} \frac{a}{2\binom{N}{k-1}} \binom{|U|}{k-1} =: ac'\binom{|U|}{k-1},$$

where $c' = \frac{1}{2(k-1)^{d-2}\binom{N}{k-1}} > 0$ is a constant. Therefore, the size of $\mathcal{F}$ satisfies

$$|\mathcal{F}| \le \binom{|U|}{k-1} - ac'\binom{|U|}{k-1} + \nu(1+a)\mathrm{EX}^{k'}\left(n', \mathcal{H}_{k-1}^{d-2}\right) + c\binom{|U|}{k-3} + M_d.$$

Since $\mathrm{EX}^{k'}\left(n', \mathcal{H}_{k-1}^{d-2}\right) \le \frac{d-2}{k-1}\binom{|U|}{k-2}$, we obtain that $c'\binom{|U|}{k-1} > \nu\mathrm{EX}^{k'}\left(n', \mathcal{H}_{k-1}^{d-2}\right)$ and, hence, we have

$$|\mathcal{F}| \le \binom{|U|}{k-1} + \nu(1+o(1))\mathrm{EX}^{k'}\left(n', \mathcal{H}_{k-1}^{d-2}\right).$$

Case 2: the family $\mathcal{F}(\bar{x})$ is not completely contained in $\binom{I}{k}$.

Then there exists a set $B_{\nu+1} \in \mathcal{F}(\bar{x})$ such that $B_{\nu+1} - I \ne \emptyset$. Now let $I' = I \cup B_{\nu+1}$ and let $U' = [n] - x - I$. For every $i \in [\nu+1]$ define the k'-multigraph $\mathcal{G}_i$ on U' as

$$\mathcal{G}_i = \left\{E \in \binom{U'}{k-2} \colon \exists b \in B_i \text{ such that } \{x,b\} \cup E \in \mathcal{F}\right\}.$$

Let $E \in \mathcal{G}_i$. Then the multiplicity of E is the number of vertices b in B_i such that $\{x,b\} \cup E \in \mathcal{F}$. For every $i \in [\nu+1]$ let g_i denote the number of edges in $\mathcal{G}_i$. Without loss of generality, we may assume that $g_1 \ge \cdots \ge g_{\nu+1}$. Let $S \in \binom{U'}{k-1}$, we say that S is bad if $\mathcal{G}_i[S] \in \mathcal{H}_{k-1}^{d-2}$ holds for some i. Note that if S is bad, then $\{x\} \cup S \notin \mathcal{F}$. Let β denote the number of bad sets.

Let $\mathcal{F}_{\ge 2} = \{F \in \mathcal{F}(x) \colon |F \cap I'| \ge 2\}$. Then there exists a constant c such that

$$|\mathcal{F}_{\ge 2}| \le \sum_{i=2}^{k-1} \binom{|I'|}{i}\binom{|U'|}{k-1-i} \le c\binom{|U'|}{k-3}.$$

Therefore, we have

$$\begin{aligned}
|\mathcal{F}| &\le \binom{|U'|}{k-1} - \beta + \sum_{i=1}^{\nu+1} g_i + c\binom{|U'|}{k-3} + m \\
&\le \binom{n-k\nu-2}{k-1} - \beta + \sum_{i=1}^{\nu+1} g_i + c\binom{n-k\nu-2}{k-3} + m \\
&= \binom{n-k\nu-1}{k-1} - \binom{n-k\nu-2}{k-2} - \beta + \sum_{i=1}^{\nu+1} g_i + c\binom{n-k\nu-2}{k-3} + m.
\end{aligned}$$

If $g_1 \le (1+o(1))\mathrm{EX}^{k-2}\left(|U'|, \mathcal{H}_{k-1}^{d-2}\right) \le (1+o(1))\frac{d-2}{k-1}\binom{n-k\nu-2}{k-2}$, then

$$|\mathcal{F}| \le \binom{n-k\nu-1}{k-1} + \nu\mathrm{EX}^{k-2}\left(n', \mathcal{H}_{k-1}^{d-2}\right) + m - \frac{k-d+1}{2(k-1)}\binom{n-1}{k-2}.$$

By the assumption that $|\mathcal{F}| = f(n,k,d,\nu)$, we have $m > \frac{k-d+1}{4(k-1)}\binom{n-1}{k-2}$. However, Lemma 6.1.16 implies that $\mathcal{F}$ contains a d-cluster, a contradiction. Therefore, we may assume that $g_1 = (1+a)\mathrm{EX}^{k'}\left(|U'|, \mathcal{H}_{k-1}^{d-2}\right)$ with $a \ge 2\sigma$ holds for some absolute constant $\sigma > 0$. Lemma 6.1.12 implies that $\mathcal{G}_1$ contains at least $\frac{a/2}{\binom{N}{k-1}}\binom{|U'|}{k-1}$ copies of elements in $\mathcal{H}_{k-1}^{d-2}$. Therefore, we have

$$\beta \ge \frac{1}{(k-1)^{d-2}}\frac{a}{2\binom{N}{k-1}}\binom{|U'|}{k-1} =: ac'\binom{|U'|}{k-1},$$

where $c' = \frac{1}{2(k-1)^{d-2}\binom{N}{k-1}} > 0$ is a constant. So the size of $\mathcal{F}$ satisfies

$$|\mathcal{F}| \le \binom{|U'|}{k-1} - ac'\binom{|U'|}{k-1} + (\nu+1)(1+a)\mathrm{EX}^{k'}(|U'|, \mathcal{H}_{k-1}^{d-2}) + c\binom{|U'|}{k-3} + m.$$

Since $c'\binom{|U|}{k-1} > (\nu+1)\mathrm{EX}^{k'}\left(n', \mathcal{H}_{k-1}^{d-2}\right)$, we therefore have that

$$\begin{aligned}
|\mathcal{F}| &\leq \binom{|U'|}{k-1} + (\nu+1)(1+o(1))\mathrm{EX}^{k'}\left(|U'|, \mathcal{H}_{k-1}^{d-2}\right) + m \\
&\leq \binom{n-k\nu-1}{k-1} - \binom{n-k\nu-2}{k-2} + (\nu+1)(1+o(1))EX^{k'}\left(|U'|, \mathcal{H}_{k-1}^{d-2}\right) + m \\
&\leq \binom{n-k\nu-1}{k-1} + \nu\mathrm{EX}^{k'}\left(|U'|, \mathcal{H}_{k-1}^{d-2}\right) + m - \frac{k-d+1}{2(k-1)}\binom{n-1}{k-2}.
\end{aligned}$$

By the assumption that $|\mathcal{F}| = f(n,k,d,\nu)$, we have $m > \frac{k-d+1}{4(k-1)}\binom{n-1}{k-2}$. However, Lemma6.1.16 implies that $\mathcal{F}$ contains a d-cluster, a contradiction. Therefore, we have

$$f(n,k,d,\nu) \leq \binom{n-k\nu-1}{k-1} + \nu(1+o(1))\mathrm{EX}^{k-2}\left(n-k\nu-1, \mathcal{H}_{k-1}^{d-2}\right).$$

∎

6.1.7 Proof of Theorem 6.1.10

Proof of Theorem 6.1.10. Let $\mathcal{K} \subset \binom{[n]}{k}$ be a family that is d-cluster-free but not t-wise intersecting and of size $g(n,k,d,t)$. Notice that a family that is not intersecting is also not t-wise intersecting. Therefore, we have $g(n,k,d,t) \geq f(n,k,d,1) > \binom{n-k-1}{k-1}$.

Now choose $\delta' > 0$ to be sufficiently small such that $\delta' < 2\binom{n-k-1}{k-1}/\binom{n-1}{k-1} - 1$ holds for sufficiently large n, and let ϵ', n_0' be given by Theorem 6.1.13. Let n be sufficiently large such that $n > n_0'$ and $\binom{n-k-1}{k-1} > (1-\epsilon')\binom{n-1}{k-1}$. By Theorem 6.1.13, there exists $z \in [n]$ such that $|\mathcal{K}(\bar{z})| < \delta'\binom{n-1}{k-1}$.

Notice that $\mathcal{K}(\bar{z})$ is nonempty, since otherwise every set in $\mathcal{K}$ would contain z, and this contradicts our assumption that $\mathcal{K}$ is not t-wise intersecting. So, let D be a set in $\mathcal{K}(\bar{z})$ and consider the family $\mathcal{K}(z)$. We claim that there exists a set $E \in \mathcal{K}(z)$ that is disjoint from D. Indeed, suppose that every set in $\mathcal{K}(z)$ has nonempty intersection with D. Then the size of $\mathcal{K}(z)$ is at most $\binom{n-1}{k-1} - \binom{n-k-1}{k-1}$, and, hence, we have

$$|\mathcal{K}| \leq \binom{n-1}{k-1} - \binom{n-k-1}{k-1} + \delta'\binom{n-1}{k-1} < \binom{n-k-1}{k-1},$$

a contradiction. Therefore, there exists a set $E \in \mathcal{K}(z)$ that is disjoint from D. However, this implies that $\mathcal{K}$ is not intersecting and, hence, we have $g(n,k,d,t) \leq f(n,k,d,1)$. Therefore, the equation $g(n,k,d,t) = f(n,k,d,1)$ holds for sufficiently large n. ∎

6.1.8 Concluding Remarks

In Section 6.1.3 we give two constructions for the lower bounds for $f(n,k,3,\nu)$. The first construction shows that

$$f(n,k,3,\nu) \geq \binom{n-k\nu-1}{k-1} + \sum_{i=2}^{k-1} \left\lfloor \frac{\nu}{2} \right\rfloor \binom{2k-2}{i-2}\binom{n-k\nu-1}{k-1-i} + \nu,$$

while the second construction shows that

$$f(n,k,3,\nu) \geq \binom{n-k\nu-1}{k-1} + \left\lfloor \frac{\nu}{2} \right\rfloor \binom{n-k\nu-1}{k-3} + (k-1)\mathrm{ex}(\nu, P_2^3)\binom{n-k\nu-1}{k-4} + \nu.$$

Since $\mathrm{ex}(\nu, P_2^3) \geq \binom{\nu}{2}/3$ holds for infinitely many ν, the second construction is better than the first one for large ν. However, when ν is small, say smaller than 7, then the first construction is better. So determining the extremal families for $f(n, k, 3, \nu)$ seems very complicated in general.

6.2 Conditionally interesting families

6.2.1 Introduction

Recall that $\mathcal{F} \subset \binom{[n]}{k}$ is (d, s)-conditionally intersecting if it does not contain d sets with union of size at most s and empty intersection. In particular, a family $\mathcal{F}$ is $(d, 2k)$-conditionally intersecting if it does not contain d-clusters, and a k-uniform family is $(2, 2k)$-conditionally intersecting if and only if it is intersecting. We use $h(n, k, d, s)$ to denote the maximum size of a (d, s)-conditionally intersecting family $\mathcal{F}$.

In this section, we consider the structure of conditionally intersecting families, which is motivated by a structural theorem for $(3, 6)$-conditionally intersecting family proved by Frankl [95].

Definition 6.2.1. *Let $\mathcal{H} \subset 2^{[n]}$, and let $H \in \mathcal{H}$. A subset $G \subset H$ is called unique if there is no other set in $\mathcal{H}$ containing G.*

The following result of Bollobás [23] gives an upper bound for the size of a family in which every set has a unique subset.

Theorem 6.2.2 (Bollobás [23])**.** *Suppose that for every member H of the family $\mathcal{H} \subset 2^{[n]}$ the set $G(H) \subset H$ is a unique subset. Then*

$$\sum_{H \in \mathcal{H}} \frac{1}{\binom{n-|H-G(H)|}{|G(H)|}} \leq 1.$$

Frankl [95] proved the following structural result for $(3, 6)$-conditionally intersecting families.

Theorem 6.2.3 (Frankl [95])**.** *Suppose that* $\mathcal{F} \subset \binom{[n]}{3}$ *is a* $(3,6)$*-conditionally intersecting family. Then* $\mathcal{F}$ *can be partitioned into two families* $\mathcal{H}$ *and* $\mathcal{B}$*, and the ground set* $[n]$ *can be partitioned into two disjoint subsets* Y *and* Z *such that the following statements hold.*

(a) $\mathcal{H} \subset \binom{Y}{3}$ *and every set* $H \in \mathcal{H}$ *contains a unique* 2*-subset.*

(b) $\mathcal{B} \subset \binom{Z}{3}$ *and* $\mathcal{B}$ *is the vertex disjoint union of* $|Z|/4$ *copies of complete* 3*-graphs on* 4 *vertices.*

First, let us show how to use Theorem 6.2.3 to get an upper bound for $|\mathcal{F}|$. Let $\mathcal{F} \subset \binom{[n]}{3}$ be a $(3,6)$-conditionally intersecting family, and let $Y, Z, \mathcal{B}$ and $\mathcal{H}$ be given by Theorem 6.2.3. Since every set in $\mathcal{H}$ contains a unique 2-subset, it follows from Theorem 6.2.2 that $|\mathcal{H}| \le \binom{|Y|-1}{2}$. On the other hand, it is easy to see that $|\mathcal{B}| = |Z|$. Therefore,

$$|\mathcal{F}| = |\mathcal{H}| + |\mathcal{F}| \le \binom{|Y|-1}{2} + |Z| \le \binom{n-1}{2},$$

and equality holds only if $Z = \emptyset$.

In [95], Frankl also asked for a structural result for a $(3,2k)$-conditionally intersecting family $\mathcal{F} \subset \binom{[n]}{k}$ which can imply the $\binom{n-1}{k-1}$ bound for $|\mathcal{F}|$. Here we consider a more general question, namely the structures of $(d, 2k+d-3)$-conditionally intersecting families for all $k \ge d \ge 3$, and we obtain the following result.

Let $\mathcal{L}_k$ denote the collection of all k-graphs on at most $2k$ vertices.

Theorem 6.2.4. *Let $k \geq d \geq 3$ be fixed. Suppose that $\mathcal{F} \subset \binom{[n]}{k}$ is a $(d, 2k+d-3)$-conditionally intersecting family. Then $\mathcal{F}$ can be partitioned into three families $\mathcal{H}$, $\mathcal{B}$ and $\mathcal{S}$, and the ground set $[n]$ can be partitioned into two subsets Y and Z such that the following statements hold.*

(a) $\mathcal{H} \subset \binom{Y}{k}$ *and every set $H \in \mathcal{H}$ contains a unique $(k-1)$-subset.*

(b) *Z has a partition $V_1 \cup \cdots \cup V_t$ with each V_i of size at most $2k$ such that $\mathcal{B} \subset \bigcup_{i=1}^{t} \binom{V_i}{k}$, i.e., the family $\mathcal{B}$ is the vertex disjoint union of copies of k-graphs in $\mathcal{L}_k$*

(c) $\mathcal{S} \subset \binom{[n]}{k} - \binom{Y}{k}$*, and for every set $S \in \mathcal{S}$ and every $V_i \subset Z$ the size of $S \cap V_i$ is either 0 or at least d.*

Note that the constraint on $|S \cap V_i|$ in (c) for $S \in \mathcal{S}$ and $V_i \subset Z$ implies that the family $\mathcal{S}$ is actually very sparse. Therefore, the term $|\mathcal{S}|$ contributes very little to $|\mathcal{F}|$.

Our next result gives a structure for $(k, 2k)$-intersecting families for all $k \geq 3$.

Theorem 6.2.5. *Let $k \geq 3$ be fixed. Suppose that $\mathcal{F} \subset \binom{[n]}{k}$ is a $(k, 2k)$-conditionally intersecting family. Then $\mathcal{F}$ can be partitioned into two families $\mathcal{H}$ and $\mathcal{B}$, and the ground set $[n]$ can be partitioned into two subsets Y and Z such that the following statements hold.*

(a) $\mathcal{H} \subset \binom{Y}{k}$ *and every set $H \in \mathcal{H}$ contains a unique $(k-1)$-subset.*

(b) $\mathcal{B} \subset \binom{Z}{k}$ *and $\mathcal{B}$ is the vertex disjoint union of $\frac{|Z|}{k+1}$ copies of complete k-graphs on $(k+1)$ vertices.*

Applying the structural results above we are able to give some new proofs to the following theorems.

Theorem 6.2.6. *Let $k \geq d \geq 3$ be fixed and $n \geq 3k^5$. Suppose that $\mathcal{F} \subset \binom{[n]}{k}$ is a $(d, 2k+d-3)$-conditionally intersecting family. Then $|\mathcal{F}| \leq \binom{n-1}{k-1}$.*

Note that Theorem 6.2.6 is true for every $n \geq 3k/2$ according to the result in [190], but in our proof we need the assumption that $n \geq 3k^5$ to keep the calculations simple.

Theorem 6.2.7. *Let $k \geq 3$ be fixed and $n \geq k^2/(k-1)$. Suppose that $\mathcal{F} \subset \binom{[n]}{k}$ is a $(k, 2k)$-conditionally intersecting family. Then $|\mathcal{F}| \leq \binom{n-1}{k-1}$.*

Theorem 6.2.8. *Let $k \geq 3$ be fixed and $n \geq 3k\binom{2k}{k}$. Let $\mathcal{F} \subset \binom{[n]}{k}$ be a family that is $(3, 2k)$-conditionally intersecting but not intersecting. Then $|\mathcal{F}| \leq \binom{n-k-1}{k-1} + 1$.*

Note that Theorem 6.2.8 shows that Mammoliti and Britz's conjecture is true for $d = 3$, and the proof here is completely different from the proof in the previous section.

6.2.2 Structural Results

Let $\mathcal{F}$ be a k-uniform family on $[n]$ and $B \in \mathcal{F}$. We say B is bad if it does not contain any unique $(k-1)$-subset. Suppose that $B = \{b_1, \ldots, b_k\}$ is a bad set in $\mathcal{F}$, then there exist k distinct sets $C_1, \ldots, C_k$ in $\mathcal{F}$ such that $B \cap C_i = B - \{b_i\}$ for all $i \in [k]$. Let $V_B = B \cup C_1 \cdots \cup C_k$ and $H_B = \{B, C_1, \ldots, C_k\}$. First let us prove Theorem 6.2.5.

Proof of Theorem 6.2.5. Suppose that $\mathcal{F}$ is a $(k, 2k)$-conditionally intersecting family, and suppose that $B = \{b_1, \ldots, b_k\}$ is a bad set in $\mathcal{F}$. Let $C_1, \cdots, C_k, V_B, H_B$ be defined as above. Since $|V_B| \leq 2k$, by assumption we have $C_1 \cap \cdots \cap C_k \neq \emptyset$. It follows that $|V_B| = k+1$ and, hence, the family H_B is a complete k-graph on V_B. Let b_{k+1} denote the vertex in $V_B - B$, and let $F \in \mathcal{F} - H_B$. Then we claim that $F \cap V_B = \emptyset$. Indeed, suppose that $F \cap V_B \neq \emptyset$. We may assume that $F \cap V_B = \{b_1, \ldots, b_\ell\}$ for some $\ell \in [k-1]$. Now, rename the edges in H_B as $B_i = V_B - b_i$ for all $i \in [k+1]$. Since $|F \cup B_1 \cup \cdots \cup B_{k-1}| \leq 2k$ and $F \cap B_1 \cap \cdots \cap B_{k-1} = \emptyset$, the k sets $F, B_1, \ldots, B_{k-1}$ form a k-cluster in $\mathcal{F}$, a contradiction. Therefore, $F \cap V_B = \emptyset$. To finish the proof we just let $\mathcal{B}$ be the collection of all bad sets in $\mathcal{F}$, and let $\mathcal{H} = \mathcal{F} - \mathcal{B}$. ∎

Before proving Theorem 6.2.4 let us present a useful lemma. Let $s = 2k + d - 3$.

Lemma 6.2.9. *Suppose that $\mathcal{F}$ is a (d, s)-conditionally intersecting family and B is a bad set in $\mathcal{F}$. Then for every $F \in \mathcal{F}$ either $|F \cap V_B| = 0$ or $|F \cap V_B| \geq d$.*

Proof. Let B is a bad set in $\mathcal{F}$ and let V_B be the set as we defined before. Suppose that $F \in \mathcal{F}$ has nonempty intersection with V_B. It suffices to show that $|F \cap V_B| \geq d$. For contradiction, suppose that $|F \cap B| = x$, $|F \cap (V_B - B)| = y$ and $x + y \leq d - 1$. Suppose that $F \cap B = \{b_{m_1}, \ldots, b_{m_x}\}$ and $F \cap (V_B - B) = \{c_{n_1}, \ldots, c_{n_y}\}$.

If $x = d - 1$, then $y = 0$ and, hence, the d sets $F, C_{m_1}, \ldots, C_{m_{d-1}}$ satisfy $|F \cup C_{m_1} \cup \cdots \cup C_{m_{d-1}}| \leq 2k$ and $F \cap C_{m_1} \cap \cdots \cap C_{m_{d-1}} = \emptyset$, a contradiction. If $x = d - 2$, then the d sets $F, B, C_{m_1}, \ldots, C_{m_{d-2}}$ satisfy $|F \cup B \cup C_{m_1} \cup \cdots \cup C_{m_{d-2}}| \leq 2k$ and $F \cap B \cap C_{m_1} \cap \cdots \cap C_{m_{d-2}} = \emptyset$, a contradiction. Therefore, we may assume that $x \leq d - 3$. Let $p = d - (x + 2)$. Choose p sets $C_{q_1}, \ldots, C_{q_p}$ from $\{C_1, \ldots, C_k\} - \{C_{m_1}, \ldots, C_{m_x}\}$. Then the d sets $F, B, C_{m_1}, \ldots, C_{m_x}, C_{q_1}, \ldots, C_{q_p}$ satisfy $|F \cup B \cup C_{m_1} \cup \cdots \cup C_{m_x} \cup C_{q_1} \cup \cdots \cup C_{q_p}| \leq 2k + p$ and $F \cap B \cap C_{m_1} \cap \cdots \cap C_{m_x} \cap C_{q_1} \cap \cdots \cap C_{q_p} = \emptyset$. By assumption we have $2k + p \geq s$ and, hence, $x = 0$ and $y \geq 1$.

Let $p' = d - (y+2)$, and choose p' sets $C_{q_1}, \dots, C_{q_{p'}}$ from $\{C_1, \dots, C_k\} - \{C_{n_1}, \dots, C_{n_y}\}$. Then the d sets $F, B, C_{n_1}, \dots, C_{n_y}, C_{q_1}, \dots, C_{q_{p'}}$ satisfy $|F \cup B \cup C_{n_1} \cup \cdots \cup C_{n_y} \cup C_{q_1} \cup \cdots \cup C_{q_{p'}}| \leq 2k + p' \leq s$ and $F \cap B \cap C_{n_1} \cap \cdots \cap C_{n_y} \cap C_{q_1} \cap \cdots \cap C_{q_{p'}} = \emptyset$, a contradiction. Therefore, we have $|F \cap V_b| \geq d$. ∎

Now we are ready to prove Theorem 6.2.4.

Proof of Theorem 6.2.4. Let $\mathcal{F}$ be a (d,s)-conditionally intersecting family. Choose a collection of bad sets $\{B_1, \dots, B_t\}$ for some t from $\mathcal{F}$ such that the sets $V_{B_1}, \dots, V_{B_t}$ are pairwise disjoint, and any other bad set in $\mathcal{F}$ has nonempty intersection with V_{B_i} for some $i \in [t]$. Note that this can be done by greedy choosing each B_i from $\mathcal{F}$ such that B_i is disjoint from $\bigcup_{j<i} V_{B_j}$, and by Lemma 6.2.9 the set V_{B_i} is also disjoint from $\bigcup_{j<i} V_{B_j}$.

Now let $V_i = V_{B_i}$ and $H_i = H_{B_i}$ for $i \in [t]$. Let $Z = \bigcup_{i\in[t]} V_i$ and $Y = [n] - Z$. Let $\mathcal{B} = \bigcup_{i\in[t]} H_i$, $\mathcal{H} = \mathcal{F} \cap \binom{Y}{k}$ and $\mathcal{S} = \mathcal{F} - \mathcal{B} - \mathcal{H}$. Suppose that $S \in \mathcal{S}$. Then by Lemma 6.2.9, either $|S \cap V_i| = 0$ or $|S \cap V_i| \geq d$ for every $i \in [t]$, and this completes the proof of Theorem 6.2.4. ∎

6.2.3 Applications

In this section we show some applications of Theorems 6.2.4 and 6.2.5 by giving new proofs to Theorems 6.2.6, 6.2.7, and 6.2.8. First let us prove Theorem 6.2.7.

Proof of Theorem 6.2.7. Suppose that $\mathcal{F}$ is a $(k, 2k)$-conditionally intersecting family on $[n]$. Let $Y, Z, \mathcal{B}$ and $\mathcal{H}$ be given by Theorem 6.2.5. By Theorem 6.2.2, $\mathcal{H} \leq \binom{|Y|-1}{k-1}$. On the other hand, it is easy to see that $|\mathcal{B}| = (k+1) \times |Z|/(k+1) = |Z|$. Therefore, $|\mathcal{F}| = |\mathcal{H}| + |\mathcal{B}| \leq \binom{|Y|-1}{k-1} + |Z| \leq \binom{n-1}{k-1}$, and equality holds only if $Z = \emptyset$. ∎

Now we apply Theorem 6.2.4 to prove Theorem 6.2.6.

Proof of Theorem 6.2.6. Let $\mathcal{F}$ be a $(d, 2k+d-3)$-conditionally intersecting family on $n \geq 3k^5$ vertices. Let $Y, Z, \mathcal{B}, \mathcal{H}$ and $\mathcal{S}$ be given by Theorem 6.2.4. Let $v_i = |V_i|$ for $i \in [t]$. Let $Y_0 = Y$ and $Y_i = Y_{i-1} \cup V_i$ for $i \in [t]$ and let $y_i = |Y_i|$ for $0 \leq i \leq t$. Define $\mathcal{H}_i = \mathcal{F} \cap \binom{Y_i}{k}$ and let $h_i = |\mathcal{H}_i|$. By Lemma 6.2.9, every set $H \in \mathcal{H}_i$ is either disjoint from V_i or has an intersection of size at least d with V_i. Therefore, $|\mathcal{H}_i| \leq |\mathcal{H}_{i-1}| + \sum_{\ell=d}^{k} \binom{v_i}{\ell}\binom{y_{i-1}}{k-\ell}$. Inductively, we obtain

$$|\mathcal{F}| \leq |\mathcal{H}| + \sum_{i=0}^{t-1}\sum_{\ell=d}^{k} \binom{v_{i+1}}{\ell}\binom{y_i}{k-\ell} \leq \binom{y_0-1}{k-1} + \sum_{i=0}^{t-1}\sum_{\ell=d}^{k} \binom{2k}{\ell}\binom{n-k-1}{k-\ell}.$$

Since $\binom{2k}{\ell}\binom{n-k-1}{k-\ell} \geq \binom{2k}{\ell+1}\binom{n-k-1}{k-\ell-1}$, we obtain

$$\begin{aligned}
|\mathcal{F}| &\leq \binom{y_0-1}{k-1} + \sum_{i=0}^{t-1}(k-d)\binom{2k}{d}\binom{n-k-1}{k-d} \\
&\leq \binom{y_0-1}{k-1} + (k-d)\binom{2k}{d}\binom{n-k-1}{k-d}\frac{n-y_0}{k+1} \\
&\leq \binom{y_0-1}{k-1} + \binom{2k}{3}\binom{n-k-1}{k-3}(n-y_0).
\end{aligned}$$

Now let $\delta = \left(2\binom{2k}{3}\right)^{-1}$. If $n - y_0 \leq \delta n$, then

$$|\mathcal{F}| < \binom{n-1}{k-1} - k\binom{n-k-1}{k-2} + \frac{n}{2}\binom{n-k-1}{k-3} < \binom{n-1}{k-1},$$

and we are done. Therefore, we may assume that $y_0 \le (1-\delta)\, n$. Then

$$|\mathcal{F}| \le \left(1 - \frac{1}{4\binom{2k}{3}}\right)\binom{n-1}{k-1} + \binom{n-k-1}{k-3}\frac{n}{2} \le \binom{n-1}{k-1},$$

and this completes the proof of Theorem 6.2.6. ∎

The remaining part of this section is devoted to prove Theorem 6.2.8. We will use the following lemma in our proof.

Lemma 6.2.10. *Suppose that $\mathcal{H} \subset \binom{[n]}{k}$, and every set $H \in \mathcal{H}$ has a unique $(k-1)$-subset $G(H) \subset H$. Then*

$$|\mathcal{H}| \le \frac{n-k+1}{n}|\partial\mathcal{H}|.$$

Proof. Consider a weight function $\omega(G,H)$ for all pairs $G \subset H \in \mathcal{F}$ with $|G| = k-1$. For every $G \in \partial\mathcal{H}$ and every $H \in \mathcal{H}$ assign weight 1 to (G,H) if $G = G(H)$ and $(n-k+1)^{-1}$ if $G \ne G(H)$. Then an easy double counting gives

$$\left(1 + \frac{k-1}{n-k+1}\right)|\mathcal{H}| = \sum_{(G,H)} \omega(G,H) \le |\partial\mathcal{H}|,$$

which implies $|\mathcal{H}| \le (n-k+1)|\partial\mathcal{H}|/n$. ∎

Definition 6.2.11. *Let $\mathcal{F} \subset \binom{[n]}{k}$ and $S \subset [n]$. Then $\mathcal{F}$ is a full star on S if it is the collection of all k-subsets of S that contain a fixed vertex v, and $\mathcal{F}$ is a star if it is a subfamily of some full star on S. In either case, we call v the core of $\mathcal{F}$.*

Now we prove Theorem 6.2.8.

Proof of Theorem 6.2.8. Let $n \geq 3k\binom{2k}{k}$ and let $\mathcal{F}$ be a family on $[n]$ such that $\mathcal{F}$ is $(3, 2k)$-conditionally intersecting but not intersecting. Suppose that $B \in \mathcal{F}$ is a bad set. Let V_B, H_B be as defined at the beginning of this section and let $\mathcal{F}' = \mathcal{F} \cap \binom{[n]-V_B}{k}$. Since $\mathcal{F}'$ is also $(3, 2k)$-intersecting, by result in [190], $|\mathcal{F}'| \leq \binom{n-|V_B|-1}{k-1}| \leq \binom{n-k-2}{k-1}$. Then by Lemma 6.2.9,

$$
\begin{aligned}
|\mathcal{F}| &\leq |\mathcal{F}'| + \sum_{i=3}^{k}\binom{2k}{i}\binom{n-k-1}{k-i} \\
&\leq \binom{n-k-2}{k-1} + k\binom{2k}{3}\binom{n-k-1}{k-3} \\
&= \binom{n-k-1}{k-1} - \left(\binom{n-k-2}{k-2} - k\binom{2k}{3}\binom{n-k-1}{k-3}\right) < \binom{n-k-1}{k-1} + 1,
\end{aligned}
$$

and we are done. So we may assume that every $F \in \mathcal{F}$ has a unique $(k-1)$-subset $G(F)$.

Since $\mathcal{F}$ is not intersecting, there exist two disjoint sets A, B in $\mathcal{F}$. Assume that $A = \{a_1, \ldots, a_k\}$ and $B = \{b_1, \ldots, b_k\}$. Let $I = \{a_1, \ldots, a_k, b_1, \ldots, b_k\}$ and let $U = [n] - I$. For every set $C \subset U$ of size at most $k-1$ define the family $\mathcal{F}(C)$ on I as follows:

$$\mathcal{F}(C) = \{F - C \colon F \in \mathcal{F} \text{ and } F \cap U = C\}.$$

For every $i \in \{0, 1, \ldots, k\}$ let

$$\mathcal{F}_i = \{F \in \mathcal{F} \colon |F \cap I| = i\}.$$

First notice that $\mathcal{F}_k = \{A, B\}$, since any extra edge in $\mathcal{F}_k$ together with A, B would form a 3-cluster in $\mathcal{F}$. Next, we will prove

$$\sum_{i=0}^{\ell} |\mathcal{F}_i| \leq \sum_{i=1}^{\ell} \binom{n-2k}{k-i}\binom{k-1}{i-1}. \tag{6.1}$$

for all $\ell \in [k]$. Suppose that Equation 6.1 is true, then by letting $\ell = k$ we obtain

$$|\mathcal{F}| = \sum_{i=0}^{k} |\mathcal{F}_i| \leq \sum_{i=1}^{k-1} \binom{n-2k}{k-i}\binom{k-1}{i-1} + 2 = \binom{n-k-1}{k-1} + 1,$$

and this will complete the proof of Theorem 6.2.8. One could compare Equation 6.1 with a similar inequality in [190], which is

$$|\mathcal{F}| \leq \sum_{\ell=1}^{k} \binom{n-tk}{k-\ell}\binom{tk-1}{\ell-1} = \binom{n-1}{k-1}, \tag{6.2}$$

where t is the maximum number of pairwise disjoint sets in $\mathcal{F}$. For the case $t = 2$, the summand in Equation 6.2 is $\binom{n-2k}{k-\ell}\binom{2k-1}{\ell-1}$, but the summand in Equation 6.1 is $\binom{n-2k}{k-\ell}\binom{k-1}{\ell-1}$, which is smaller when $\ell \geq 2$.

Claim 6.2.12. *Let $F \in \mathcal{F}_1$. Then the set $F \cap U$ is a unique $(k-1)$-subset of F in $\mathcal{F}$.*

Proof. Without lose of generality, we may assume that $F = \{a_1, f_1, \ldots, f_{k-1}\}$, where $f_1, \ldots, f_{k-1}$ are contained in U. Suppose that there is another edge $F' \in \mathcal{F}$ containing $\{f_1, \ldots, f_{k-1}\}$. Then the three sets A, F, F' form a 3-cluster in $\mathcal{F}$, a contradiction. Therefore, $F \cap U = \{f_1, \ldots, f_{k-1}\}$ is a unique $(k-1)$-subset of F in $\mathcal{F}$. ∎

Now we prove Equation 6.1 for $\ell = 1$. Let us consider the family $\mathcal{F}_0 \cup \mathcal{F}_1$. Define

$$\mathcal{M} = \left\{ G \in \binom{U}{k-1} : \exists F \in \mathcal{F}_0 \cup \mathcal{F}_1 \text{ such that } G \subset F \right\}.$$

By assumption, every set $F \in \mathcal{F}_0 \cup \mathcal{F}_1$ has a unique $(k-1)$-subset $G(F)$, and by Claim 6.2.12, we may assume that $G(F) \subset U$. Let $\mathcal{G} = \{G(F) : F \in \mathcal{F}_1\}$. For every set $F_1 \in \mathcal{F}_1$, the set $G(F_1)$ cannot be contained in $\partial \mathcal{F}_0$, since otherwise one could easily find a 3-cluster. Therefore, $\mathcal{G}$ and $\partial \mathcal{F}_0$ are disjoint. Since $|\mathcal{G}| = |\mathcal{F}_1|$, by Lemma 6.2.10, we have

$$\frac{|U|}{|U| - k + 1} |\mathcal{F}_0| + |\mathcal{F}_1| \leq |\mathcal{M}| \leq \binom{n-2k}{k-1},$$

and hence $|\mathcal{F}_0| + |\mathcal{F}_1| \leq \binom{n-2k}{k-1}$.

To prove Equation 6.1 for $\ell \geq 2$, we need to give an upper bound for $|\mathcal{F}_i|$ for every $2 \leq i \leq k-1$. Since $|\mathcal{F}_i| = \sum_{C \in \binom{U}{k-i}} |\mathcal{F}(C)|$, it suffices to give an upper bound for $|\mathcal{F}(C)|$ for every $C \in \binom{U}{k-i}$. Unfortunately, the inequality $|\mathcal{F}(C)| \leq \binom{k-1}{i-1}$ is not true in general. So, in our proof, we will build a relationship between $\mathcal{F}_i$ and $\bigcup_{j<i} \mathcal{F}_j$ and then use this relation to prove Equation 6.1.

The basic idea in our proof is showing that if $|\mathcal{F}(C)|$ is bigger than its expected value $\binom{k-1}{k-|C|-1}$, then there must be many sets D containing C such that the size of $\mathcal{F}(D)$ is smaller than its expected value $\binom{k-1}{k-|D|-1}$.

Let $C \subset U$ be a set of size at most $k-2$. We say C is perfect if the family $\mathcal{F}(C)$ is a full star on either A or B. Let $D \subset U$ be a set of size $k-1$. We say D is perfect if there exists a set F in $\mathcal{F}$ that contains D.

For every $i \in [k-1]$ let $\mathcal{P}_i$ be the collection of all perfect sets in $\binom{U}{k-i}$, and let $\mathcal{N}_i$ be the collection of non-perfect sets in $\binom{U}{k-i}$. Let $p_i = |\mathcal{P}_i|$ and $n_i = |\mathcal{N}_i|$ for $i \in [k-1]$ and notice that $p_i + n_i = \binom{|U|}{k-i}$.

For every $i \in \{2, \dots, k-1\}$ let $\mathcal{P}'_i$ denote the collection of all sets $C \in \binom{U}{k-i}$ such that C is contained in a perfect set in $\binom{U}{k-i+1}$, and let $\mathcal{N}'_i$ denote the collection all of sets $D \in \binom{U}{k-i}$ such that D is not contained in any perfect set in $\binom{U}{k-i+1}$. Let $p'_i = |\mathcal{P}'_i|$ and $n'_i = |\mathcal{N}'_i|$ for $i \in \{2, \dots, k-1\}$. Let $\mathcal{G}_i = \mathcal{N}_i \cap \mathcal{P}'_i$ and $\mathcal{B}_i = \mathcal{N}_i \cap \mathcal{N}'_i$, and let $g_i = |\mathcal{G}_i|$ and $b_i = |\mathcal{B}_i|$ for $i \in \{2, \dots, k-1\}$. Let $\mathcal{G}_1 = \mathcal{N}_1$, and let $g_1 = n_1$, $b_1 = 0$. Note that by definition, $b_i + g_i = n_i$ and $n'_i \geq b_i$ for $i \in [k-1]$.

By the definition of perfect sets, $|\mathcal{F}(C)| = \binom{k-1}{i-1}$ for all $C \in \mathcal{P}_i$. Later we will show that $|\mathcal{F}(C)| < \binom{k-1}{i-1}$ for all $C \in \mathcal{G}_i$. For every $C \in \mathcal{B}_i$ it could be true that $|\mathcal{F}(C)| > \binom{k-1}{i-1}$. However, for every $C \in \mathcal{B}_i$ there are either many sets in $\mathcal{G}_{i-1}$ containing C, which means that there are many sets $D \in \binom{U}{k-i+1}$ with $|\mathcal{F}(D)|$ smaller than its expected value, or there are many sets in $\mathcal{B}_{i-1}$, in which case we turn to consider sets in $\binom{U}{k-i+2}$ and repeat this argument until we end up with many sets P in $\binom{U}{k-1}$ with $|\mathcal{F}(P)|$ smaller than its expected value.

The next claim gives a relation between n_i and b_{i+1}.

Claim 6.2.13. *For every $i \in [k-2]$ we have*

$$n_i \geq \frac{n-3k}{k} b_{i+1}.$$

Proof. Let $C \in \mathcal{N}'_{i+1}$, and let $u \in U - C$. By definition $C \cup \{u\}$ is a non-perfect set in $\binom{U}{k-i}$. Therefore, we have $(k-i)n_i \geq n'_{i+1}(n-3k+i+1) \geq b_{i+1}(n-3k)$. It follows that $n_i \geq (n-3k)b_{i+1}/k$. ∎

Claim 6.2.14. *The following statement holds for all $\ell \geq (k+1)/2$. Suppose that $C \subset U$ is a perfect set of size ℓ, and $\mathcal{F}(C)$ is a full star on A (or on B) with core v. Then for every $(\ell-1)$-subset C' of C the family $\mathcal{F}(C')$ is a star on A (or on B) with core v.*

Proof. Let $C \subset U$ such that $\mathcal{F}(C)$ is a full star on A with core $v \in A$. Without loss of generality we may assume that $v = a_1$. Let $E' \in \mathcal{F}(C')$. If $E' \subset B$, then choose a set E from $\mathcal{F}(C)$, and the three sets $E \cup C, E' \cup C', B$ form a 3-cluster in $\mathcal{F}$, a contradiction. If $E' \cap A \neq \emptyset$ and $E' \cap B \neq \emptyset$, then let $x = |E' \cap A|$ and $y = |E' \cap B|$. Since $x + y = k - \ell + 1$, we have $x \leq k - \ell$ and $y \leq k - \ell$. If $a_1 \notin E' \cap A$, then by the assumption that $\ell \geq (k+1)/2$ and $\mathcal{F}(C)$ is a full star, there exists a set $E \in \mathcal{F}(C)$ such that $(E' \cap A) \cap E = \emptyset$. So the three sets $E' \cup C', E \cup C, A$ form a 3-cluster in $\mathcal{F}$, a contradiction. If $a_1 \in E' \cap A$, then by assumption there exists a set $E \in \mathcal{F}(C)$ such that $E' \cap A \subset E$. However, the three sets $E \cup C, E' \cup C', B$ form a 3-cluster in $\mathcal{F}$, a contradiction. Therefore, every set in $\mathcal{F}(C')$ is completely contained in A.

Next, we show that every set $E' \in \mathcal{F}(C')$ contains a_1. Suppose there exists a set $E' \in \mathcal{F}(C')$ such that $a_1 \notin E'$. By assumption we have $k - \ell + 1 + k - \ell \leq k$, so there exists a set $E \in \mathcal{F}(C)$

such that $E \cap E' = \emptyset$. However, the three sets $E' \cup C', E \cup C, A$ form a 3-cluster in $\mathcal{F}$, a contradiction. Therefore, the family $\mathcal{F}(C')$ is a star on A with core a_1. ∎

For every $i \in [k-1]$ let $w_i = \binom{k-1}{i-1}\binom{n-2k}{k-i}$ and $k_i = \binom{2k}{i} - \binom{k-1}{i-1} + 1$. Our next claim gives an upper bound for $|\mathcal{F}_i|$ for $2 \le i \le (k+1)/2$.

Claim 6.2.15. *For every i satisfying $2 \le i \le (k+1)/2$ we have*

$$|\mathcal{F}_i| \le w_i + k_i b_i - n_i.$$

Proof. Let us give an upper bound for $|\mathcal{F}(C)|$ for every $C \in \binom{U}{k-i}$. First notice that by definition $|\mathcal{F}(C)| = \binom{k-1}{i-1}$ for all $C \in \mathcal{P}_i$. By Claim 6.2.14, $|\mathcal{F}(C)| \le \binom{k-1}{i-1} - 1$ for all $C \in \mathcal{G}_i$. On the other hand, it is trivially true that $|\mathcal{F}(C)| \le \binom{2k}{i}$ for all $C \in \mathcal{B}_i$. Therefore,

$$\begin{aligned}
|\mathcal{F}_i| &= \sum_{C \in \mathcal{P}_i} |\mathcal{F}(C)| + \sum_{C \in \mathcal{G}_i} |\mathcal{F}(C)| + \sum_{C \in \mathcal{B}_i} |\mathcal{F}(C)| \\
&\le \binom{k-1}{i-1} p_i + \left(\binom{k-1}{i-1} - 1\right) g_i + \binom{2k}{i} b_i \\
&= \binom{k-1}{i-1}\binom{n-2k}{k-i} + \left(\binom{2k}{i} - \binom{k-1}{i-1} + 1\right) b_i - n_i = w_i + k_i b_i - n_i.
\end{aligned}$$

Here we used that fact that $b_i + g_i = n_i$ and $n_i + p_i = \binom{n-2k}{k-i}$. ∎

Recall that Claim 6.2.13 says that $n_i \ge (n-3k) b_{i+1}/k$. Since $n \ge 3k\binom{2k}{k}$ and $k_{i+1} < \binom{2k}{k}$, we have $n_i/2 \ge k_{i+1} b_{i+1}$. Combining this inequality with Claim 6.2.15 we obtain the following claim.

Claim 6.2.16. *For every ℓ satisfying $1 \leq \ell \leq (k+1)/2$ we have*

$$\sum_{i=0}^{\ell} |\mathcal{F}_i| \leq \sum_{i=1}^{\ell} w_i - \sum_{i=1}^{\ell} \frac{n_i}{2}.$$

Proof. The case $\ell = 1$ follows from the inequality that

$$|\mathcal{F}_0| + |\mathcal{F}_1| \leq |\mathcal{M}| = \binom{n-2k}{k-1} - n_1.$$

For $\ell \geq 2$ by Claim 6.2.15 we obtain

$$\sum_{i=0}^{\ell} |\mathcal{F}_i| \leq \sum_{i=1}^{\ell} (w_i + k_i b_i - n_i) = \sum_{i=1}^{\ell} w_i - \sum_{i=1}^{\ell-1} (n_i - k_{i+1} b_{i+1}) - n_\ell \leq \sum_{i=1}^{\ell} w_i - \sum_{i=1}^{\ell} \frac{n_i}{2}.$$

∎

The next step is to extend Claim 6.2.16 to all $\ell > (k+1)/2$.

Claim 6.2.17. *Let $C \subset U$ be a set of size $\ell \geq 2$. Suppose that $\mathcal{F}(C)$ is a full-star on A (or on B) with core v and there exists a perfect set $P \in \binom{U}{k-1}$ containing C. Then, for every $(\ell-1)$-subset $C' \subset C$ the family $\mathcal{F}(C')$ is a star on A (or on B) with core v.*

Proof. Let $C \subset U$ be a set of size ℓ such that $\mathcal{F}(C)$ is a full-star on A with core v. Without loss of generality we may assume that $v = a_1$. Let $P \in \binom{U}{k-1}$ be a perfect set containing C. By the definition of perfect set there exists a set $F \in \mathcal{F}$ containing P. Suppose that $F = P \cup \{u\}$, and we want to show that $u = a_1$. Suppose that $u \notin A$. Then for every $E \in \mathcal{F}(C)$ the three sets $A, F, E \cup C$ form a 3-cluster in $\mathcal{F}$, a contradiction. Therefore, $u \in A$.

Now suppose for the contrary that $u \neq a_1$. Then by assumption there exists a set $E \in \mathcal{F}(C)$ not containing u and, hence, the three sets $A, F, E \cup C$ form a 3-cluster in $\mathcal{F}$, a contradiction. Therefore, $u = a_1$.

Let $C' \subset C$ be a set of size $\ell - 1$ and $E' \in \mathcal{F}(C')$. If $E' \subset B$, then for every $E \in \mathcal{F}(C)$ the three sets $E \cup C, E' \cup C', B$ form a 3-cluster in $\mathcal{F}$, a contradiction. If $E' \cap A \neq \emptyset$ and $E' \cap B \neq \emptyset$, then let $x = |E' \cap A|$ and $y = |E' \cap B|$. Since $x + y = k - \ell + 1$, we have $x \leq k - \ell$ and $y \leq k - \ell$. If $x \leq k - \ell - 1$, then by assumption there exists a set $E \in \mathcal{F}(C)$ containing $E' \cap A$. However, the three sets $E \cup C, E' \cup C', B$ form a 3-cluster in $\mathcal{F}$, a contradiction. Therefore, we may assume that $x = k - \ell$. If $a_1 \in E' \cap A$, then there exists a set $E \in \mathcal{F}(C)$ such that $E' \cap A = E$. However, the three sets $E \cup C, E' \cup C', B$ form a 3-cluster in $\mathcal{F}$, a contradiction. If $a_1 \notin E' \cap A$, then the three sets $A, F, E' \cup C'$ form a 3-cluster in $\mathcal{F}$, a contradiction. Therefore, every set in $\mathcal{F}(C')$ is completely contained in A.

Suppose that there is a set $E' \in \mathcal{F}(C')$ not containing a_1, then the three sets $A, F, E' \cup C'$ would form a 3-cluster in $\mathcal{F}$, a contradiction. Therefore, every set in $\mathcal{F}(C')$ contains a_1, and this complete the proof of Claim 6.2.17. ∎

Let $c = \lfloor (k+1)/2 \rfloor$ and let $m = \lfloor k/2 \rfloor$, and notice that $m + c = k$. The next claim shows that Equation 6.1 holds for $\ell = c + 1$.

Claim 6.2.18. *We have*

$$\sum_{i=0}^{c+1} |\mathcal{F}_i| \leq \sum_{i=1}^{c+1} w_i - \sum_{i=1}^{c+1} \frac{n_i}{4}.$$

Proof. Similar to the proof of Claim 6.2.15, for every $C \in \mathcal{P}_{c+1}$ we have $|\mathcal{F}(C)| = \binom{k-1}{c}$, and for every $C \in \mathcal{B}_{c+1}$ we have $|\mathcal{F}(C)| \leq \binom{2k}{c+1}$.

For every perfect set $D \in \binom{U}{m}$ we say that D is a good container if D itself is contained in a perfect $(k-1)$-set, otherwise we say that D is a bad container. Let $\mathcal{S}$ be the collection of all sets in $\mathcal{G}_{c+1}$ that are contained in a good container. Let $\mathcal{T}$ be the collection of all sets in $\mathcal{G}_{c+1}$ that are not contained in any good container. Let $s = |\mathcal{S}|$ and $t = |\mathcal{T}|$. Since every bad container in $\binom{U}{m}$ has m subsets of size $m-1$, the number of bad containers in $\binom{U}{m}$ is at least t/m.

Let $D \in \binom{U}{m}$ be a bad container. Then for every $E \in \binom{U-D}{k-m-1}$ the set $D \cup E$ is non-perfect in $\binom{U}{k-1}$. Therefore, $n_1 \geq \binom{n-2k-m}{c-1} t / \left(m\binom{k-1}{m}\right)$. By definition, every set $C \in \mathcal{G}_{c+1}$ is contained in a perfect set $D \in \binom{U}{m}$. If $C \in \mathcal{S}$, then by Claim 6.2.17, $|\mathcal{F}(C)| \leq \binom{k-1}{c} - 1$. If $C \in \mathcal{T}$, then it is trivially true that $|\mathcal{F}(C)| \leq \binom{k}{c+1}$. Therefore,

$$\begin{aligned}
|\mathcal{F}_{c+1}| &= \sum_{C \in \mathcal{P}_{c+1}} |\mathcal{F}(C)| + \sum_{C \in \mathcal{B}_{c+1}} |\mathcal{F}(C)| + \sum_{C \in \mathcal{S}} |\mathcal{F}(C)| + \sum_{C \in \mathcal{T}} |\mathcal{F}(C)| \\
&\leq \binom{k-1}{c} p_{c+1} + \binom{2k}{c+1} b_{c+1} + \left(\binom{k-1}{c} - 1\right) s + \binom{2k}{c+1} t \\
&= w_{c+1} + k_{c+1} b_{c+1} + k_{c+1} t - n_{c+1}.
\end{aligned}$$

Here we used the fact that $s+t=g_{c+1}$, $g_{c+1}+b_{c+1}=n_{c+1}$ and $n_{c+1}+p_{c+1}=\binom{n-2k}{k-c-1}$. Combining the inequality above with Claim 6.2.15, we obtain

$$\sum_{i=0}^{c+1}|\mathcal{F}_i| \le \sum_{i=1}^{c+1}(w_i+k_ib_i-n_i)+k_{c+1}t.$$

Since $n_1/4 \ge k_{c+1}t$ and $n_i/2 \ge k_{i+1}b_{i+1}$,

$$\sum_{i=0}^{c+1}|\mathcal{F}_i| \le \sum_{i=1}^{c+1}w_i - \sum_{i=1}^{c+1}\frac{n_i}{4}.$$

∎

Claim 6.2.19. *Every set $C \subset U$ of size at most $k-c$ is contained in a perfect $(k-1)$-set.*

Proof. Let $C \subset U$ be a set of size $\ell \le k-c$. Suppose that C is not contained in any perfect $(k-1)$-set. Then for every $S \in \binom{U-C}{k-\ell-1}$ the set $C \cup S$ is non-perfect and of size $k-1$. Therefore, we have $n_1 \ge \binom{n-2k-\ell}{k-\ell-1}/\binom{k-1}{\ell} \ge \binom{n-2k-\ell}{c-1}/\binom{k-1}{\ell}$. On the other hand, we have $\sum_{i=c+2}^{k-1}|\mathcal{F}_i| \le \sum_{i=c+2}^{k-1}\binom{2k}{i}\binom{n-2k}{k-i}$. Since $n \ge 3k\binom{2k}{k}$, $n_1/4 > \sum_{i=c+2}^{k-1}|\mathcal{F}_i|$. Therefore, by Claim 6.2.18,

$$\sum_{i=0}^{k-1}|\mathcal{F}_i| = \sum_{i=1}^{c+1}|\mathcal{F}_i| + \sum_{i=c+2}^{k-1}|\mathcal{F}_i| \le \sum_{i=1}^{c+1}w_i - \sum_{i=1}^{c+1}\frac{n_i}{4} + \sum_{i=c+2}^{k-1}\binom{2k}{i}\binom{n-2k}{k-c-2} < \sum_{i=1}^{k-1}w_i,$$

and we are done. So we may assume that C is contained in a perfect $(k-1)$-set. ∎

Claim 6.2.20. *The inequality $|\mathcal{F}_i| \le w_i + t_ib_i - n_i$ holds for all $i \ge c+1$.*

Proof. By Claim 6.2.19, every set $C \subset U$ of size at most $k-c$ is contained in a perfect $(k-1)$-set. Therefore, by Claim 6.2.17,

$$\begin{aligned}|\mathcal{F}_i| &= \sum_{C\in\mathcal{P}_i} |\mathcal{F}(C)| + \sum_{C\in\mathcal{G}_i} |\mathcal{F}(C)| + \sum_{C\in\mathcal{B}_i} |\mathcal{F}(C)| \\ &\le \binom{k-1}{i-1} p_i + \left(\binom{k-1}{i-1} - 1\right) g_i + \binom{2k}{i} b_i.\end{aligned}$$

∎

By Claims 6.2.13, 6.2.15, and 6.2.20,

$$\begin{aligned}\sum_{i=0}^{k-1} |\mathcal{F}_i| &\le \sum_{i=1}^{k-1} (w_i + t_i b_i - n_i) = \sum_{i=1}^{k-1} w_i - \sum_{i=1}^{k-2} (n_i - t_{i+1} b_{i+1}) - n_{k-1} \\ &\le \sum_{i=1}^{k-1} w_i - \sum_{i=1}^{k-1} \frac{n_i}{2},\end{aligned}$$

which proves Equation 6.1, and equality holds if and only if C is perfect for every $C \in \binom{U}{i}$ and for every $i \in [k-1]$, which implies that $\mathcal{F}$ is the disjoint union of a k-set and a full star. ∎

6.3 Katona's intersecting shadow theorem

6.3.1 Introduction

The seminal Kruskal–Katona theorem gives a tight upper bound for $|\mathcal{H}|$ as a function of $|\partial\mathcal{H}|$. In order to state its precise form, we need the following definition.

The colex order on $\binom{[n]}{k}$ is defined as follows:

$$A \prec B \text{ iff } \max\{(A \setminus B) \cup (B \setminus A)\} \in B.$$

Write $L_m\mathcal{H}$ to denote the set of the first m elements of $\mathcal{H} \subset \binom{[n]}{k}$ in the colex order. When $\mathcal{H} = \binom{[n]}{k}$, we abuse notation by simply writing $L_m\binom{[n]}{k}$.

The Kruskal–Katona theorem states that the families in $\binom{[n]}{k}$ with a fixed number of sets and minimum shadow size are initial elements of the colex order.

Theorem 6.3.1 (Kruskal–Katona [132; 154]). *For $n \geq k > \ell \geq 1$ and $\mathcal{H} \subset \binom{[n]}{k}$ with $|\mathcal{H}| = m$,*

$$|\partial_\ell \mathcal{H}| \geq \left|\partial_\ell L_m \binom{[n]}{k}\right|.$$

6.3.1.1 Katona's shadow intersection theorem

The Kruskal–Katona theorem was extended to many families with additional properties. One such central result is due to Katona [131] who proved the following theorem for t-intersecting families, which are families in which every two sets have at least t common elements.

Theorem 6.3.2 (Katona [131]). *Let $n \geq k > t \geq \ell \geq 1$. If $\mathcal{H} \subset \binom{[n]}{k}$ is t-intersecting, then*

$$|\partial_\ell \mathcal{H}| \geq \frac{\binom{2k-t}{k-\ell}}{\binom{2k-t}{k}} |\mathcal{H}|.$$

The only case of equality in Theorem 6.3.2 is when $n = 2k - t$ and $\mathcal{H} \cong \binom{[2k-t]}{k}$ (see [1]).

Theorem 6.3.2 is a foundational result in extremal set theory with many applications. Its first application was to prove a conjecture of Erdős-Ko-Rado on the maximum size of a t-intersecting family in $2^{[n]}$. It was used to obtain short new proofs for several classical results. For example, Frankl–Füredi [103] used it to give a short proof for the Erdős–Ko–Rado theorem, and Frankl–Tokushige [106] used it to obtain a short proof for the Hilton–Milner theorem. It also has many applications to Sperner families and other types of intersection problems [28; 42; 98; 100; 116; 185; 224; 250].

This paper is concerned with improving the bounds in Theorem 6.3.2 and related results about shadows of families with certain properties. In many cases the bounds we prove are best possible.

Our first result improves Theorem 6.3.2 for intersecting families (the case $t = 1$) and applies to all $n > 2k$. It is convenient to define the family

$$\mathrm{EKR}(n, k) = \left\{ A \in \binom{[n]}{k} : 1 \in A \right\}.$$

Theorem 6.3.3. *Let $n > 2k \geq 6$ and $1 \leq \ell < k$. Suppose that $\mathcal{H} \subset \binom{[n]}{k}$ is intersecting and*

$$|\mathcal{H}| = m > m(n,k) = \begin{cases} 3n-8, & \text{if } k=3, \\ \binom{n-1}{k-1} - \binom{n-k}{k-1} + \binom{n-k-2}{k-3} + 2, & \text{if } k \geq 4. \end{cases}$$

Then $|\partial_\ell \mathcal{H}| \geq |\partial_\ell L_m \mathrm{EKR}(n,k)|$. In particular, if for some $x \in \mathbb{R}$

$$|\mathcal{H}| = \binom{x-1}{k-1} > m(n,k) \tag{6.3}$$

then $|\partial_\ell \mathcal{H}| \geq \binom{x}{k-\ell}$.

Remarks.

- For $k = 3$ and $m = 3n - 8$, the inequality $|\partial \mathcal{H}| < |\partial L_m \mathrm{EKR}(n,k)|$ is possible (see Fact 6.3.27 with $t = 1$), so Theorem 6.3.3 is best possible in this sense. In fact, when $k = 3$ one can compute the sharp lower bound for $|\partial \mathcal{H}|$ for all intersecting families $\mathcal{H}$ using our proof method but we do not carry out all these details.
- For fixed $k > 3$ and $n \to \infty$, we will lower the value of $m(n,k)$ from $(k-1+o(1))\binom{n}{k-2}$ to $(3+o(1))\binom{n}{k-2}$ in Theorem 6.3.10 and the constant 3 will be shown to be tight.

Ahlswede, Aydinian, and Khachatrian [1] considered large t-intersecting families on $\mathbb{N}$. Let $\binom{\mathbb{N}}{k}$ denote the collection of all k-subsets of $\mathbb{N}$ and let

$$\mathrm{EM}(\mathbb{N},k,s,t) = \left\{ A \in \binom{\mathbb{N}}{k} : |A \cap [s]| \geq t \right\}.$$

Theorem 6.3.4 (Ahlswede–Aydinian–Khachatrian [1]). *Let $\mathcal{H} \subset \binom{\mathbb{N}}{k}$ be a t-intersecting family.*

- *For $1 \le \ell \le t < k$, there exists $m_1(k,t,\ell) \in \mathbb{N}$ such that if $|\mathcal{H}| = m \ge m_1(k,t,\ell)$, then $|\partial_\ell \mathcal{H}| \ge |\partial_\ell L_m \mathrm{EM}(\mathbb{N}, k, 2k-2-t, k-1)|$.*
- *For $1 \le t < \ell < k$, there exists $m_2(k,t,\ell) \in \mathbb{N}$ such that if $|\mathcal{H}| = m \ge m_2(k,t,\ell)$, then $|\partial_\ell \mathcal{H}| \ge |\partial_\ell L_m \mathrm{EM}(\mathbb{N}, k, t, t)|$.*

Let

$$\mathrm{EM}(n,k,s,t) = \left\{ A \subset \binom{[n]}{k} : |A \cap [s]| \ge t \right\},$$

and set $\mathrm{EM}(n,k,s,t) = \emptyset$ if $t > \min\{k,s\}$, and $\mathrm{EM}(n,k,s,t) = \binom{[n]}{k}$ if $t \le 0$. Notice that if $t \le k \le n \le s$, then $\mathrm{EM}(n,k,s,t)$ is a complete k-graph on n vertices.

Notice that for every $m \le \binom{n-t}{k-t}$ we have $L_m \mathrm{EM}(n,k,t,t) = L_m \mathrm{EM}(\mathbb{N},k,t,t)$. Therefore, Theorem 6.3.4 implies the following result.

Corollary 6.3.5. *Let $1 \le t < \ell < k$ and $\mathcal{H} \subset \binom{[n]}{k}$ be a t-intersecting family with $|\mathcal{H}| = m > m_2(k,t,\ell)$. Then $|\partial_\ell \mathcal{H}| \ge |\partial_\ell L_m \mathrm{EM}(n,k,t,t)|$.*

However, for the case $\ell \le t$ we show that the smallest possible size of the ℓ-th shadow of large t-intersecting families on $[n]$ is different than the formula in Theorem 6.3.4. Let

$$\mathrm{AK}(n,k,t) = \left\{ A \in \binom{[n]}{k} : [t] \subset A \text{ and } [t+1,k+1] \cap A \ne \emptyset \right\} \cup \left(\bigcup_{i \in [t]} \{[k+1] \setminus \{i\}\} \right).$$

Notice that $\mathrm{AK}(n,k,t)$ and $\mathrm{EM}(n,k,t+2,t+1)$ are both t-intersecting,

$$|\mathrm{AK}(n,k,t)| \sim (k-t+1)\binom{n}{k-t-1},$$

$$|\mathrm{EM}(n,k,t+2,t+1)| \sim (t+2)\binom{n}{k-t-1}.$$

Our next result is a finite version of Theorem 6.3.4.

Theorem 6.3.6. *Let $t \geq 1, k \geq 3, 1 \leq \ell < k$, and $n > (t+1)(k-t+1)$. Suppose that $\mathcal{H} \subset \binom{[n]}{k}$ is t-intersecting and*

$$|\mathcal{H}| = m > m(n,k,t) = \begin{cases} \max\left\{|\mathrm{AK}(n,k,t)|, |\mathrm{EM}(n,k,t+2,t+1)|\right\}, & \text{if } t < \frac{k-1}{2}, \\ |\mathrm{EM}(n,k,t+2,t+1)|, & \text{if } t \geq \frac{k-1}{2}. \end{cases}$$

Then $|\partial_\ell \mathcal{H}| \geq |\partial_\ell L_m \mathrm{EM}(n,k,t,t)|$. In particular, if

$$|\mathcal{H}| = \binom{x-t}{k-t} > m(n,k,t) \tag{6.4}$$

for some $x \in \mathbb{R}$. Then $|\partial_\ell \mathcal{H}| \geq \sum_{i=t-\ell}^{k-\ell} \binom{t}{i}\binom{x-t}{k-\ell-i}$. For $1 \leq \ell \leq t$ the value of $m(n,k,t)$ is tight for $t \geq \frac{k-1}{2}$ and is tight up to a constant multiplicative factor independent of n for $t < \frac{k-1}{2}$.

Remarks.

- Theorem 6.3.6 implies that for a t-intersecting family $\mathcal{H} \subset \binom{[n]}{k}$ with large size,

$$\frac{|\partial_\ell \mathcal{H}|}{|\mathcal{H}|} > \binom{t}{\ell} \geq \frac{\binom{2k-t}{k-\ell}}{\binom{2k-t}{k}}$$

  for $1 \leq \ell \leq t$ with equality in the second inequality iff $\ell = t$. Hence our bound is better than that in Theorem 6.3.2 (as expected since our bound is best possible).

- For $t < \frac{k-1}{2}$ we will show in the last section that the lower bound for $|\mathcal{H}|$ in Theorem 6.3.6 can be improved slightly.

6.3.1.2 Frankl's theorem

The matching number of $\mathcal{H}$, denoted by $\nu(\mathcal{H})$, is the maximum number of pairwise disjoint edges in $\mathcal{H}$. Notice that $\nu(\mathrm{EM}(n,k,s,1)) \leq s$ with equality iff $n \geq ks$ and

$$|\mathrm{EM}(n,k,s,1)| = \binom{n}{k} - \binom{n-s}{k} \sim s\binom{n}{k-1} \qquad (n \to \infty).$$

The Erdős matching conjecture [59] says that for all $n \geq (s+1)k-1$, if $\mathcal{H} \subset \binom{[n]}{k}$ and $\nu(\mathcal{H}) \leq s$, then

$$|\mathcal{H}| \leq \max\left\{\binom{(s+1)k-1}{k}, \binom{n}{k} - \binom{n-s}{k}\right\}. \tag{6.5}$$

When $s = 1$, Equation 6.5 follows from the Erdős–Ko–Rado theorem [69].

Theorem 6.3.7 (Erdős–Ko–Rado [69]). *Let $k \geq 2$ and $n \geq 2k$, $\mathcal{H} \subset \binom{[n]}{k}$ be an intersecting family. Then $\mathcal{H} \leq \binom{n-1}{k-1}$ and when $n > 2k$ equality holds iff $\mathcal{H} \cong \mathrm{EKR}(n,k)$.*

The Erdős matching conjecture is still open and the current record on this conjecture is due to Frankl [93].

Theorem 6.3.8 (Frankl [93]). *Let $k \geq 2$ and $n \geq (2s+1)k - s$, $\mathcal{H} \subset \binom{[n]}{k}$ and $\nu(\mathcal{H}) \leq s$. Then $\mathcal{H} \leq \binom{n}{k} - \binom{n-s}{k}$ with equality iff $\mathcal{H} \cong \mathrm{EM}(n, k, s, 1)$.*

If we take $t = 1$ in Theorem 6.3.2, then every intersecting family $\mathcal{H} \subset \binom{[n]}{k}$ satisfies $|\partial\mathcal{H}| \geq |\mathcal{H}|$. Frankl generalized this as follows.

Theorem 6.3.9 (Frankl [92; 93]). *Let $n \geq k \geq 2$ and $\mathcal{H} \subset \binom{[n]}{k}$. If $\nu(\mathcal{H}) = s \geq 1$, then*

$$|\partial\mathcal{H}| \geq \frac{|\mathcal{H}|}{s}$$

with equality iff $\mathcal{H} \cong \binom{[(s+1)k-1]}{k}$.

Theorem 6.3.9 is a crucial tool in the proof of Theorem 6.3.8 and any improvement in Theorem 6.3.9 for small values of n could lead to a corresponding improvement in Theorem 6.3.8. Our final result provides such an improvement (for large n) that is sharp if $|\mathcal{H}|$ is large.

Theorem 6.3.10. *For every $k \geq 3$ and every $s \geq 1$ there exists $c = c(k, s)$ such that the following holds as $n \to \infty$. Suppose that $\mathcal{H} \subset \binom{[n]}{k}$ satisfies $\nu(\mathcal{H}) \leq s$ and $|\mathcal{H}| = m \geq (1 + o(1))c\binom{n}{k-2}$. Then*

$$|\partial\mathcal{H}| \geq |\partial L_m \mathrm{EM}(n, k, s, 1)|.$$

In particular, if $|\mathcal{H}| = \binom{x}{k} - \binom{x-s}{k} \geq (1 + o(1))c\binom{n}{k-2}$ for some $x \in \mathbb{R}$, then $|\partial\mathcal{H}| \geq \binom{x}{k-1}$.

The families $\mathcal{H}$ above can have size as large as $\Theta(n^{k-1})$, so the point of Theorem 6.3.10 is that it applies to $\mathcal{H}$ that are quite a bit smaller. In fact, as we will show below, the order of magnitude n^{k-2} is best possible for such a result and even the constant c is sharp in many cases.

Our proof of Theorem 6.3.10 gives

$$c(k,s) = \begin{cases} 3 & \text{if } s = 1 \\ \binom{2s+1}{2} & \text{if } k = 3 \\ \binom{s+1}{2}k & \text{if } k \geq 4, s \geq 2. \end{cases} \tag{6.6}$$

On the other hand, the following construction shows that the lower bound for $c(k,s)$ (given by Equation 6.6) in Theorem 6.3.10 is tight for (k,s) if $s = 1$ or $k = 3$, and is tight up to a constant factor for all other (s,k).

Let $\mathcal{G} = EM(n, k, 2s{+}1, 2)$ and $m = |\mathcal{G}| \sim \binom{2s+1}{2}\binom{n}{k-2}$ and let $x \in \mathbb{R}$ such that $\binom{x}{k} - \binom{x-s}{k} = m$. Since $\binom{x}{k} - \binom{x-s}{k} \sim s\binom{x}{k-1}$, $x = \Theta(n^{\frac{k-2}{k-1}})$. Notice that

$$\begin{aligned} s|\partial\mathcal{G}| - m &= s\sum_{i=1}^{k-1}\binom{2s+1}{i}\binom{n-2s-1}{k-1-i} - \sum_{i=2}^{k}\binom{2s+1}{i}\binom{n-2s-1}{k-i} \\ &= \Theta(n^{k-3}), \end{aligned}$$

and

$$\begin{aligned} s|\partial L_m \mathrm{EM}(n,k,s,1)| - m &\geq s\binom{x}{k-1} - \left(\binom{x}{k} - \binom{x-s}{k}\right) \\ &= \Theta(x^{k-2}) = \Theta(n^{\frac{(k-2)^2}{k-1}}). \end{aligned}$$

Since $\frac{(k-2)^2}{k-1} > k-3$, $|\partial L_m \mathrm{EM}(n,k,s,1)| > |\partial \mathrm{EM}(n,k,2s+1,2)|$ for sufficiently large n. Therefore, we obtain the following result.

Fact 6.3.11. *For every $k \geq 3$ and sufficiently large n there exists $\mathcal{G} \subset \binom{[n]}{k}$ with $\nu(\mathcal{G}) = s$ and $|\mathcal{G}| = (1+o_n(1))\binom{2s+1}{2}\binom{n}{k-2}$ such that $|\partial \mathcal{G}| < |\partial L_{|\mathcal{G}|} \mathrm{EM}(n,k,s,1)|$.*

It would be interesting to determine the minimum value of $c(k,s)$ such that the conclusion in Theorem 6.3.10 holds for all $|\mathcal{H}| > c(k,s)\binom{n}{k-2}$ and sufficiently large n.

6.3.2 Proofs

6.3.2.1 Extension of the k-cascade representation

In this section, we prove an extension of the well-known k-cascade representation of a number. The k-cascade representation plays an important role in the Kruskal–Katona theorem and the extension that we prove plays an analogous role for our theorems. As a convention, let $\binom{a}{b} = 0$ if $b < 0$ or $a < b$, and let $\binom{a}{0} = 1$ for all $a \geq 0$.

For an r-graph $\mathcal{H}$ and a vertex set S that is disjoint from $V(\mathcal{H})$ define

$$\mathcal{H} + S = \{A \cup S : A \in \mathcal{H}\}.$$

For every $i \in \mathbb{N}$ let $\widehat{i} = i + 1$.

Lemma 6.3.12. *Let $n \geq k \geq t \geq 0$ and $s \geq t \geq 0$. Then the following hold.*

(a) $|\mathrm{EM}(n,k,s,t)| = \binom{n}{k} - \sum_{j=0}^{t-1} \binom{s}{j}\binom{n-s}{k-j}$.

(b) For every $m \geq 1$ there exist integers $a_k > a_{k-1} > \cdots > a_h \geq h \geq \max\{t, 1\}$ such that

$$L_m \mathrm{EM}(n,k,s,t) = \mathrm{EM}(a_k,k,s,t) \cup \bigcup_{i=h}^{k-1} \left(\mathrm{EM}(a_i,i,s,t) + \{\widehat{a}_{i+1}, \ldots, \widehat{a}_k\}\right)$$

Proof. (a) is clear. So let us consider (b).

First, it follows from the definition that the colex order of $\mathrm{EM}(n',k,s,t)$ is the initial segment of the colex on $\mathrm{EM}(n,k,s,t)$ for all $n' < n$. Let $\mathcal{F} = L_m\mathrm{EM}(n,k,s,t)$. Without loss of generality we may assume that $\mathcal{F} \neq \mathrm{EM}(n',k,s,t)$ for all $n' \leq n$ since otherwise we can let $h = k$ and $a_k = n'$ and we are done. So there exists a_k such that $\mathrm{EM}(a_k,k,s,t) \subset \mathcal{F} \subset \mathrm{EM}(a_k+1,k,s,t)$ and hence every set in $\mathcal{F}\backslash\mathrm{EM}(a_k,k,s,t)$ contains a_k+1. Therefore, $\mathcal{F} = \mathrm{EM}(a_k,k,s,t) \cup (\mathcal{F}_k + \{\widehat{a}_k\})$ for some $\mathcal{F}_k \subset \mathrm{EM}(a_k,k-1,s,t)$.

Let $m' = |\mathcal{F}_k|$. Then it follows from the definition of colex order that $\mathcal{F}_k = L_{m'}\mathrm{EM}(a_k,k-1,s,t)$. So we can repeat the argument above to show that there exists a_{k-1} such that $\mathcal{F}_k = \mathrm{EM}(a_{k-1},k-1,s,t) \cup (\mathcal{F}_{k-1} + \{\widehat{a}_{k-1}\})$. This means that

$$\mathcal{F} = \mathrm{EM}(a_k,k,s,t) \cup (\mathrm{EM}(a_{k-1},k-1,s,t) + \{\widehat{a}_k\}) \cup (\mathcal{F}_{k-1} + \{\widehat{a}_k, \widehat{a}_{k-1}\}).$$

Inductively, one will get a decomposition of $\mathcal{F}$ as in Lemma 6.3.12. ∎

Lemma 6.3.13. *For every integers $m \geq 1$, there exists a unique representation of m in the form*

$$m = \sum_{i=h}^{k} \binom{a_i}{i} - \sum_{j=0}^{t-1} \binom{s}{j} \sum_{i=h}^{k} \binom{a_i - s}{i - j},$$

where $a_k > \cdots > a_h \geq h \geq \max\{t, 1\}$ are integers.

Proof. If $t = 0$, then this is just the k-cascade representation of m. So we may assume that $t \geq 1$. Let $n \in \mathbb{N}$ be sufficiently large such that $m \leq |\mathrm{EM}(n, k, s, t)|$. Then the existence of such a representation follows from Lemma 6.3.12 since

$$\begin{aligned} m = |L_m EM(n, k, s, t)| = \sum_{i=h}^{k} |EM(a_i, i, s, t)| &= \sum_{i=h}^{k} \left(\binom{a_i}{i} - \sum_{j=0}^{t-1} \binom{s}{j} \binom{a_i - s}{i - j} \right) \\ &= \sum_{i=h}^{k} \binom{a_i}{i} - \sum_{j=0}^{t-1} \binom{s}{j} \sum_{i=h}^{k} \binom{a_i - s}{i - j}. \end{aligned}$$

Next, we prove the uniqueness of such representation of m. Suppose that there exists $a_k > \cdots > a_h \geq h \geq t$ and $b_k > \cdots > b_{h'} \geq h' \geq t$ such that

$$\sum_{i=h}^{k} \binom{a_i}{i} - \sum_{j=0}^{t-1} \binom{s}{j} \sum_{i=h}^{k} \binom{a_i - s}{i - j} = m = \sum_{i=h'}^{k} \binom{b_i}{i} - \sum_{j=0}^{t-1} \binom{s}{j} \sum_{i=h'}^{k} \binom{b_i - s}{i - j}. \tag{6.7}$$

Without loss of generality we may assume that $a_k \neq b_k$ since otherwise we can consider $m' = m - \left(\binom{a_k}{i} - \sum_{j=0}^{t-1}\binom{s}{j}\binom{a_k-s}{i-j}\right)$ instead. Let

$$\mathcal{F}_a = EM(a_k,k,s,t) \cup \bigcup_{i=h}^{k-1} \left(EM(a_i,i,s,t) + \{\widehat{a}_{i+1},\ldots,\widehat{a}_k\}\right)$$

and

$$\mathcal{F}_b = EM(b_k,k,s,t) \cup \bigcup_{i=h'}^{k-1} \left(EM(b_i,i,s,t) + \{\widehat{b}_{i+1},\ldots,\widehat{b}_k\}\right).$$

Then

$$|\mathcal{F}_a| = \sum_{i=h}^{k}\binom{a_i}{i} - \sum_{j=0}^{t-1}\binom{s}{j}\sum_{i=h}^{k}\binom{a_i-s}{i-j},$$

and

$$|\mathcal{F}_b| = \sum_{i=h'}^{k}\binom{b_i}{i} - \sum_{j=0}^{t-1}\binom{s}{j}\sum_{i=h'}^{k}\binom{b_i-s}{i-j}.$$

Without loss of generality we may assume that $a_k \geq b_k + 1$. However, notice that in this case $\mathcal{F}_b$ is a proper subset of $\mathcal{F}_a$, since every set of $\mathcal{F}_b$ has maximum element at most $b_k + 1 \leq a_k$. This contradicts Equation 6.7. ∎

6.3.2.2 Shifting

For every $A \in \mathcal{H}$ and $1 \le i < j \le n$ define

$$S_{ij}(A) = \begin{cases} (A \setminus \{j\}) \cup \{i\}, & \text{if } j \in A, i \notin A, \text{ and } (A \setminus \{j\}) \cup \{i\} \notin \mathcal{H}, \\ A, & \text{otherwise.} \end{cases}$$

Let $S_{ij}(\mathcal{H}) = \{S_{ij}(A) \colon A \in \mathcal{H}\}$ and call $\mathcal{H}$ shifted if $\mathcal{H} = S_{ij}(\mathcal{H})$ for all $1 \le i < j \le n$.

Fact 6.3.14 (see [91])**.** *The following statements hold for all* $\mathcal{H} \subset \binom{[n]}{k}$ *and all* $1 \le i < j \le n$ *and all* $1 \le t, \ell \le k-1$.

- $|\mathcal{H}| = |S_{ij}(\mathcal{H})|$.
- $\partial_\ell S_{ij}(\mathcal{H}) \subset S_{ij}(\partial_\ell \mathcal{H})$ *and in particular,* $|S_{ij}(\partial_\ell \mathcal{H})| \ge |\partial_\ell S_{ij}(\mathcal{H})|$
- $\nu(S_{ij}(\mathcal{H})) \le \nu(\mathcal{H})$.
- *If* $\mathcal{H}$ *is* t*-intersecting, then* $S_{ij}(\mathcal{H})$ *is also* t*-intersecting.*

6.3.2.3 Main Lemma

Fact 6.3.14 shows that it suffices to consider shifted families in all proofs in this paper. The main technical statement in this work is Lemma 6.3.16 below which is a generalization of the Kruskal-Katona theorem. For two families $\mathcal{H}_1$ and $\mathcal{H}_2$ we write $\mathcal{H}_1 \subset \mathcal{H}_2$ if $\mathcal{H}_1$ is isomorphic to a subgraph of $\mathcal{H}_2$.

Given a family $\mathcal{H}$, let $\mathcal{H}(1) = \{A \setminus \{1\} \colon 1 \in A \in \mathcal{H}\}$ and $\mathcal{H}(\bar{1}) = \{A \in \mathcal{H} \colon 1 \notin A\}$. It is easy to see that if $\mathcal{H}$ is shifted, then $\partial \mathcal{H}(\bar{1}) \subset \mathcal{H}(1)$ and hence $|\partial \mathcal{H}| = |\mathcal{H}(1)| + |\partial \mathcal{H}(1)|$.

Lemma 6.3.15. *Let $n \geq k \geq t \geq 0$ and $s \geq t \geq 0$. Suppose that*

$$m = \sum_{i=h}^{k} \binom{a_i}{i} - \sum_{j=0}^{t-1} \binom{s}{j} \sum_{i=h}^{k} \binom{a_i - s}{i - j}$$

for integers $a_k > \cdots > a_h \geq \max\{t, 1\}$. Then for $1 \leq \ell \leq k-1$

$$|\partial_\ell L_m \mathrm{EM}(n,k,s,t)| = \sum_{i=h}^{k} \binom{a_i}{i-\ell} - \sum_{j=0}^{t-1-\ell} \binom{s}{j} \sum_{i=h}^{k} \binom{a_i - s}{i - \ell - j}.$$

Proof. Fix $1 \leq \ell \leq k-1$. By Lemma 6.3.12,

$$L_m \mathrm{EM}(n,k,s,t) = \mathrm{EM}(a_k,k,s,t) \cup \bigcup_{i=h}^{k-1} \left(\mathrm{EM}(a_i,i,s,t) + \{\widehat{a}_{i+1}, \ldots, \widehat{a}_k\}\right).$$

Notice that

$$\partial \mathrm{EM}(a_k,k,s,t) = \mathrm{EM}(a_k,k-1,s,t-1),$$

and for every $h \leq i \leq k-1$ we have

$$\partial \left(\mathrm{EM}(a_i,i,s,t) + \{\widehat{a}_{i+1}, \ldots, \widehat{a}_k\}\right) = \left(\mathrm{EM}(a_i,i-1,s,t-1) + \{\widehat{a}_{i+1}, \ldots, \widehat{a}_k\}\right) \cup \bigcup_{j=i+1}^{k} \left(\mathrm{EM}(a_i,i,s,t) + \{\widehat{a}_{i+1}, \ldots, \widehat{a}_k\} \setminus \{\widehat{a}_j\}\right).$$

On the other hand, for all $h \le i < j \le k-2$ since $a_j > a_i$,

$$\mathrm{EM}(a_i, i, s, t) + \{\widehat{a}_{i+1}, \dots, \widehat{a}_k\} \setminus \{\widehat{a}_j\} \subset \mathrm{EM}(a_j, j-1, s, t-1) + \{\widehat{a}_{j+1}, \dots, \widehat{a}_k\}.$$

For all $h \le i \le k-1$ since $a_k > a_i$,

$$\mathrm{EM}(a_i, i, s, t) + \{\widehat{a}_{i+1}, \dots, \widehat{a}_k\} \setminus \{\widehat{a}_k\} \subset \mathrm{EM}(a_k, k-1, s, t-1).$$

Therefore,

$$\partial L_m \mathrm{EM}(n, k, s, t) = \bigcup_{i=h}^{k} \left(\mathrm{EM}(a_i, i-1, s, t-1) + \{\widehat{a}_{i+1}, \dots, \widehat{a}_k\}\right),$$

and inductively we obtain for all $1 \le \ell \le k-1$

$$\partial_\ell L_m \mathrm{EM}(n, k, s, t) = \bigcup_{i=h}^{k} \left(\mathrm{EM}(a_i, i-\ell, s, t-\ell) + \{\widehat{a}_{i+1}, \dots, \widehat{a}_k\}\right).$$

Therefore,

$$\begin{aligned}
|\partial_\ell L_m \mathrm{EM}(n, k, s, t)| &= \sum_{i=h}^{k} |\mathrm{EM}(a_i, i-\ell, s, t-\ell)| \\
&= \sum_{i=h}^{k} \left(\binom{a_i}{i-\ell} - \sum_{j=0}^{t-1-\ell} \binom{s}{j} \binom{a_i}{i-\ell-j} \right) \\
&= \sum_{i=h}^{k} \binom{a_i}{i-\ell} - \sum_{j=0}^{t-1-\ell} \binom{s}{j} \sum_{i=h}^{k} \binom{a_i - s}{i-\ell-j}.
\end{aligned}$$

This completes the proof of Lemma 6.3.15. ∎

Lemma 6.3.16. *Let $s \geq t \geq 0$. If $\mathcal{H} \subset \mathrm{EM}(n,k,s,t)$ and $|\mathcal{H}| = m$, then*

$$|\partial \mathcal{H}| \geq |\partial L_m \mathrm{EM}(n,k,s,t)|.$$

Proof. By Lemma 6.3.13, there exists $a_k > \cdots > a_h \geq \max\{t, 1\}$ such that

$$m = \sum_{i=h}^{k} \binom{a_i}{i} - \sum_{j=0}^{t-1} \binom{s}{j} \sum_{i=h}^{k} \binom{a_i - s}{i - j}.$$

Then, by Lemma 6.3.15 it suffices to show that

$$|\partial \mathcal{H}| \geq \sum_{i=h}^{k} \binom{a_i}{i-1} - \sum_{j=0}^{t-2} \binom{s}{j} \sum_{i=h}^{k} \binom{a_i - s}{i - 1 - j}.$$

We prove this statement by induction on k, s, t. When $s = 0$ or $k = 1$ the statement is trivially true. When $t = 0$ the statement follows from the Kruskal–Katona theorem. So we may assume that $s \geq t \geq 1$ and $k \geq 2$.

Claim 6.3.17. $|\mathcal{H}(1)| \geq \sum_{i=h}^{k} \binom{a_i - 1}{i-1} - \sum_{j=0}^{t-2} \binom{s-1}{j} \sum_{i=h}^{k} \binom{a_i - s}{i-1-j}$.

Proof. Suppose not. Then

$$
\begin{aligned}
|\mathcal{H}(\bar{1})| &= |\mathcal{H}| - |\mathcal{H}(1)| \\
&> \sum_{i=h}^{k} \binom{a_i}{i} - \sum_{j=0}^{t-1} \binom{s}{j} \sum_{i=h}^{k} \binom{a_i - s}{i - j} - \left(\sum_{i=h}^{k} \binom{a_i - 1}{i - 1} - \sum_{j=0}^{t-2} \binom{s-1}{j} \sum_{i=h}^{k} \binom{a_i - s}{i - 1 - j} \right) \\
&= \sum_{i=h}^{k} \binom{a_i}{i} - \sum_{i=h}^{k} \binom{a_i - 1}{i - 1} - \sum_{j=0}^{t-1} \left(\binom{s}{j} - \binom{s-1}{j-1} \right) \sum_{i=h}^{k} \binom{a_i - s}{i - j} \\
&= \sum_{i=h}^{k} \binom{a_i - 1}{i} - \sum_{j=0}^{t-1} \binom{s-1}{j} \sum_{i=h}^{k} \binom{a_i - s}{i - j}.
\end{aligned}
$$

Since $\mathcal{H}(\bar{1}) \subset \mathrm{EM}(n, k, s-1, t)$, by the induction hypothesis

$$
|\partial\mathcal{H}(\bar{1})| > \sum_{i=h}^{k} \binom{a_i - 1}{i - 1} - \sum_{j=0}^{t-2} \binom{s-1}{j} \sum_{i=h}^{k} \binom{a_i - s}{i - 1 - j} > |\mathcal{H}(1)|,
$$

which contradicts the assumption that $\mathcal{H}$ is shifted. ∎

Since $\mathcal{H}(1) \subset EM(n, k-1, s-1, t-1)$, by the induction hypothesis and Claim 6.3.17,

$$
\begin{aligned}
|\partial\mathcal{H}| &\geq |\mathcal{H}(1)| + |\partial\mathcal{H}(1)| \\
&\geq \sum_{i=h}^{k} \binom{a_i - 1}{i - 1} - \sum_{j=0}^{t-2} \binom{s-1}{j} \sum_{i=h}^{k} \binom{a_i - s}{i - 1 - j} \\
&\quad + \sum_{i=h}^{k} \binom{a_i - 1}{i - 2} - \sum_{j=0}^{t-3} \binom{s-1}{j} \sum_{i=h}^{k} \binom{a_i - s}{i - 2 - j} \\
&= \sum_{i=h}^{k} \binom{a_i}{i - 1} - \sum_{j=0}^{t-2} \left(\binom{s-1}{j} + \binom{s-1}{j-1} \right) \sum_{i=h}^{k} \binom{a_i - s}{i - 1 - j} \\
&= \sum_{i=h}^{k} \binom{a_i}{i - 1} - \sum_{j=0}^{t-2} \binom{s}{j} \sum_{i=h}^{k} \binom{a_i - s}{i - 1 - j}.
\end{aligned}
$$

This completes the proof of Lemma 6.3.16. ∎

Corollary 6.3.18. *Let* $s \geq t \geq 0$ *and* $1 \leq \ell \leq k-1$. *Suppose that* $\mathcal{H} \subset \mathrm{EM}(n,k,s,t)$ *and* $|\mathcal{H}| = m$. *Then* $|\partial_\ell \mathcal{H}| \geq |\partial_\ell L_m \mathrm{EM}(n,k,s,t)|$.

Proof. Similar to the proof of Lemma 6.3.16, it suffices to show that if for some integers $a_k > \cdots > a_h \geq h \geq \max\{t, 1\}$

$$|\mathcal{H}| = \sum_{i=h}^{k} \binom{a_i}{i} - \sum_{j=0}^{t-1} \binom{s}{j} \sum_{i=h}^{k} \binom{a_i - s}{i-j},$$

then

$$|\partial_\ell \mathcal{H}| \geq \sum_{i=h}^{k} \binom{a_i}{i-\ell} - \sum_{j=0}^{t-1-\ell} \binom{s}{j} \sum_{i=h}^{k} \binom{a_i - s}{i-\ell-j}.$$

We proceed by induction on ℓ. When $\ell = 1$, this is Lemma 6.3.16. So we may assume that $\ell \geq 2$. By the induction hypothesis

$$|\partial_{\ell-1} \mathcal{H}| \geq \sum_{i=h}^{k} \binom{a_i}{i-\ell+1} - \sum_{j=0}^{t-\ell} \binom{s}{j} \sum_{i=h}^{k} \binom{a_i - s}{i-\ell+1-j}.$$

Since $\partial_{\ell-1}\mathcal{H} \subset \mathrm{EM}(n,k,s,t-\ell+1)$, by Lemma 6.3.16,

$$|\partial_\ell \mathcal{H}| = |\partial \partial_{\ell-1} \mathcal{H}| \geq \sum_{i=h}^{k} \binom{a_i}{i-\ell} - \sum_{j=0}^{t-\ell-1} \binom{s}{j} \sum_{i=h}^{k} \binom{a_i - s}{i-\ell-j}.$$

This completes the proof of Corollary 6.3.18. ∎

The same induction argument as above gives the following technically simpler version of Corollary 6.3.18.

Lemma 6.3.19 (Simplified version of Lemma 6.3.16)**.** *Let $s \geq t \geq 0$ and $1 \leq \ell \leq k-1$. Suppose that $\mathcal{H} \subset \mathrm{EM}(n,k,s,t)$ and $|\mathcal{H}| = \binom{x}{k} - \sum_{j=0}^{t-1}\binom{s}{j}\binom{x-s}{k-j}$ for some $x \in \mathbb{R}$. Then $|\partial_\ell \mathcal{H}| \geq \binom{x}{k-\ell} - \sum_{j=0}^{t-1-\ell}\binom{s}{j}\binom{x-s}{k-\ell-j}$.*

Let

$$\mathrm{HM}(n,k,s,t) = \left\{A \in \binom{[n]}{k} : |A \cap [s-1]| \geq 1\right\} \cup \left\{A \in \binom{[n]}{k} : s \in A \text{ and } |A \cap [s+1,s+t]| \geq 1\right\}.$$

Note that there is no constraint on the relation between s and t for $\mathrm{HM}(n,k,s,t)$.

Similar to Lemmas 6.3.12, 6.3.13, and 6.3.15 we have the following result for $\mathrm{HM}(n,k,s,t)$.

Lemma 6.3.20. *Let $n \geq k$. Then the following hold.*

(a) $|\mathrm{HM}(n,k,s,t)| = \binom{n}{k} - \binom{n-s}{k} - \binom{n-s-t}{k-1}$ *and* $|\partial \mathrm{HM}(n,k,s,t)| = \binom{n}{k-1}$.

(b) For every $m \leq |\mathrm{HM}(n,k,s,t)|$ there exist integers $a_k > \cdots > a_h \geq h \geq 1$ such that

$$L_m \mathrm{HM}(n,k,s,t) = \mathrm{HM}(a_k,k,s,t) \cup \bigcup_{i=h}^{k-1} \left(\mathrm{HM}(a_i,i,s,t) + \{\widehat{a}_{i+1},\ldots,\widehat{a}_k\}\right).$$

(c) *For every* $m \geq 1$ *there exists a unique sequence of integers* $a_k > \cdots > a_h \geq h \geq 1$ *such that*

$$m = \sum_{i=h}^{k} \binom{a_i}{i} - \sum_{i=h}^{k} \binom{a_i - s}{i} - \sum_{i=h}^{k} \binom{a_i - s - t}{i-1}.$$

(d) *If* m *is given by the equation above, then*

$$|\partial L_m \mathrm{HM}(n,k,s,t)| = \sum_{i=h}^{k} \binom{a_i}{i-1}.$$

Lemma 6.3.21. *If* $\mathcal{H} \subset \mathrm{HM}(n,k,s,t)$ *and* $|\mathcal{H}| = m$, *then* $|\partial\mathcal{H}| \geq |\partial L_m \mathrm{HM}(n,k,s,t)|$. *In particular, if* $|\mathcal{H}| = \sum_{i=h}^{k} \binom{a_i}{i} - \sum_{i=h}^{k} \binom{a_i-s}{i} - \sum_{i=h}^{k} \binom{a_i-s-t}{i-1}$ *for some integers* $a_k > \cdots > a_h \geq h \geq 1$, *then* $|\partial\mathcal{H}| \geq \sum_{i=h}^{k} \binom{a_i}{i-1}$.

Proof. Let $a_k > \cdots > a_h \geq h \geq 1$ be integers such that

$$m = \sum_{i=h}^{k} \binom{a_i}{i} - \sum_{i=h}^{k} \binom{a_i - s}{i} - \sum_{i=h}^{k} \binom{a_i - s - t}{i-1}.$$

Then by Lemma 6.3.20, it suffices to show

$$|\partial\mathcal{H}| \geq \sum_{i=h}^{k} \binom{a_i}{i-1} = |\partial L_m \mathrm{HM}(n,k,s,t)|.$$

We proceed by induction on s and t. When $s = 0$, this is trivially true. When $t = 0$, we have $\mathrm{HM}(n, k, s, 0) = \mathrm{EM}(n, k, s, 1)$, so the conclusion follows from Lemma 6.3.16. So we may assume that $s \geq 1$ and $t \geq 1$.

Claim 6.3.22. $|\mathcal{H}(1)| \geq \sum_{i=h}^{k} \binom{a_i-1}{i-1}$

Proof. Suppose not. Then

$$|\mathcal{H}(\bar{1})| = |\mathcal{H}| - |\mathcal{H}(1)| > \sum_{i=h}^{k} \binom{a_i - 1}{i} - \sum_{i=h}^{k} \binom{a_i - s}{i} - \sum_{i=h}^{k} \binom{a_i - s - t}{i - 1}.$$

Since $\mathcal{H}(\bar{1}) \subset HM(n, k, s-1, t)$, by the induction hypothesis $|\partial\mathcal{H}(\bar{1})| > \sum_{i=h}^{k} \binom{a_i-1}{i-1} > |\mathcal{H}(1)|$, which contradicts the assumption that $\mathcal{H}$ is shifted. ∎

Now, by Claim 6.3.22 and the Kruskal-Katona theorem,

$$|\partial\mathcal{H}| \geq |\mathcal{H}(1)| + |\partial\mathcal{H}(1)| \geq \sum_{i=h}^{k} \binom{a_i - 1}{i - 1} + \sum_{i=h}^{k} \binom{a_i - 1}{i - 2} = \sum_{i=h}^{k} \binom{a_i}{i - 1}.$$

This completes the proof of Lemma 6.3.21. ∎

Similarly, the same induction argument as above gives the following technically simpler version of Lemma 6.3.21.

Lemma 6.3.23 (Simplified version of Lemma 6.3.21)**.** *Suppose that $\mathcal{H} \subset \mathrm{HM}(n, k, s, t)$ and $|\mathcal{H}| = \binom{x}{k} - \binom{x-s}{k} - \binom{x-s-t}{k-1}$ for some $x \in \mathbb{R}$. Then $|\partial\mathcal{H}| \geq \binom{x}{k-1}$.*

6.3.2.4 Proof of Theorem 6.3.3

The proof of Theorem 6.3.3 uses the following structural theorem for intersecting families. For $1 \le t \le k-1$ let

$$\begin{aligned}\mathrm{HM}(n,k,t) = &\{A \in \mathrm{EKR}(n,k)\colon A \cap [2,k] \neq \emptyset \text{ or } [k+1,k+t] \subset A\} \cup \\ &\bigcup_{i=1}^{t} \{\{2,\ldots,k,k+i\}\}.\end{aligned}$$

Note that $|\mathrm{HM}(n,k,2)| = \binom{n-1}{k-1} - \binom{n-k}{k-1} + \binom{n-k-2}{k-3} + 2$ and $\mathrm{HM}(n,k,1)$ is the extremal configuration in the Hilton–Milner theorem on nontrivial intersecting families.

Theorem 6.3.24 (Han–Kohayakawa [119])**.** *Let $k \ge 3$ and $n > 2k$ and let $\mathcal{H}$ be an n-vertex intersecting k-graph. If $\mathcal{H} \not\subset \mathrm{EKR}(n,k)$ and $\mathcal{H} \not\subset \mathrm{HM}(n,k,1)$ and for $k = 3$ $\mathcal{H} \not\subset \mathrm{EM}(n,3,3,2)$ as well, then $|\mathcal{H}| \le |\mathrm{HM}(n,k,2)|$.*

Proof of Theorem 6.3.3. By the assumption on the size of $\mathcal{H}$ and Theorem 6.3.24, for $k \ge 4$ either $\mathcal{H} \subset \mathrm{EKR}(n,k)$ or $\mathcal{H} \subset \mathrm{HM}(n,k,1)$, and for $k = 3$ we have $\mathcal{H} \subset \mathrm{EKR}(n,3)$.

Suppose that $k = 3$. Since $\mathcal{H} \subset \mathrm{EKR}(n,3) = \mathrm{EM}(n,3,1,1)$, by Corollary 6.3.18, $|\partial_\ell \mathcal{H}| \ge |\partial_\ell L_m \mathrm{EKR}(n,3)|$ for $1 \le \ell \le 2$.

Now suppose that $k \ge 4$. Let $a_k > \cdots > a_h \ge h \ge 1$ be integers such that $|\mathcal{H}| = \sum_{i=h}^{k} \binom{a_i}{i} - \sum_{i=h}^{k} \binom{a_i-1}{i}$. If $\mathcal{H} \subset \mathrm{EKR}(n,k) = \mathrm{EM}(n,k,1,1)$, then by Corollary 6.3.18, $|\partial_\ell \mathcal{H}| \ge |\partial_\ell L_m \mathrm{EM}(n,k,1,1)| = |\partial_\ell L_m \mathrm{EKR}(n,k)|$ and we are done. So we may assume that $\mathcal{H} \subset \mathrm{HM}(n,k,1)$ and we are going to show that $|\partial_\ell \mathcal{H}| > \sum_{i=h}^{k} \binom{a_i}{i-\ell} = |\partial_\ell L_m \mathrm{EKR}(n,k)|$ in this case.

Suppose that $|\partial_\ell \mathcal{H}| \leq \sum_{i=h}^{k} \binom{a_i}{i-\ell}$. Let $\mathcal{H}' = \mathcal{H} \setminus \{\{2, \dots, k+1\}\}$ and note that $|\partial_\ell \mathcal{H}'| \leq |\partial_\ell \mathcal{H}| \leq \sum_{i=h}^{k} \binom{a_i}{i-\ell}$. Applying the the contrapositive of the Kruskal-Katona theorem to $\mathcal{H}'$ we obtain $|\partial \mathcal{H}'| \leq \sum_{i=h}^{k} \binom{a_i}{i-1}$. On the other hand, since $\mathcal{H} \subset \mathrm{HM}(n,k,1)$, $\mathcal{H}' \subset \mathrm{HM}(n,k,1) \setminus \{\{2, \dots, k+1\}\} = \mathrm{HM}(n,k,1,k)$. So applying the contrapositive of Lemma 6.3.21 to $\mathcal{H}'$ we obtain

$$
\begin{aligned}
|\mathcal{H}| \leq |\mathcal{H}'| + 1 &\leq \sum_{i=h}^{k} \binom{a_i}{i} - \sum_{i=h}^{k} \binom{a_i - 1}{i} - \sum_{i=h}^{k} \binom{a_i - 1 - k}{i-1} + 1 \\
&= \sum_{i=h}^{k} \binom{a_i - 1}{i-1} - \sum_{i=h}^{k} \binom{a_i - 1 - k}{i-1} + 1.
\end{aligned}
$$

Claim 6.3.25. *$a_k \geq 2k$ and if $a_k = 2k$ then $a_{k-1} = 2k-1$.*

Proof. First, suppose that $a_k \leq 2k-1$. Then

$$
\begin{aligned}
|\mathcal{H}| &\leq \sum_{i=h}^{k} \binom{a_i - 1}{i-1} - \sum_{i=h}^{k} \binom{a_i - 1 - k}{i-1} + 1 \\
&\leq \sum_{i=1}^{k} \binom{k+i-2}{i-1} - \sum_{i=1}^{k} \binom{i-2}{i-1} + 1 \\
&= \binom{2k-1}{k-1} + 1 < \binom{2k}{k-1} - \binom{k}{k-1} - \binom{k-1}{k-2} + 3,
\end{aligned}
$$

which contradicts the assumption that $|\mathcal{H}| > |\mathrm{HM}(n,k,2)|$ and $n > 2k$. Therefore, $a_k \geq 2k$.

Now suppose that $a_k = 2k$ and $a_{k-1} \leq 2k-2$. Then,

$$\begin{aligned}
|\mathcal{H}| &\leq \sum_{i=h}^{k}\binom{a_i-1}{i-1} - \sum_{i=h}^{k}\binom{a_i-1-k}{i-1} + 1 \\
&\leq \binom{2k-1}{k-1} + \sum_{i=1}^{k-1}\binom{k+i-2}{i-1} - \sum_{i=1}^{k}\binom{i-2}{i-1} \\
&= \binom{2k-1}{k-1} + \binom{2k-2}{k-2} < \binom{2k}{k-1} - \binom{k}{k-1} - \binom{k-1}{k-2} + 3,
\end{aligned}$$

where the strict inequality uses $k \geq 4$ and $n > 2k$. This contradicts the assumption that $|\mathcal{H}| > |\mathrm{HM}(n,k,2)|$. ∎

Claim 6.3.25 implies that $\sum_{i=h}^{k}\binom{a_i-1-k}{i-1} - 1 > 0$. Therefore,

$$|\mathcal{H}| \leq \sum_{i=h}^{k}\binom{a_i-1}{i-1} - \sum_{i=h}^{k}\binom{a_i-1-k}{i-1} + 1 < \sum_{i=h}^{k}\binom{a_i-1}{i-1},$$

contradicts the assumption that $|\mathcal{H}| = \sum_{i=h}^{k}\binom{a_i}{i} - \sum_{i=h}^{k}\binom{a_i-1}{i} = \sum_{i=h}^{k}\binom{a_i-1}{i-1}$. Therefore, if $\mathcal{H} \subset \mathrm{HM}(n,k,1)$, then $|\partial_\ell \mathcal{H}| > \sum_{i=h}^{k}\binom{a_i}{i-\ell}$, and this completes the proof of Theorem 6.3.3. ∎

6.3.2.5 Proof of Theorem 6.3.6

In this section we prove Theorem 6.3.6. We need the following theorem for t-intersecting families.

Theorem 6.3.26 (Ahlswede–Khachatrian [2])**.** *Let $t \geq 1, k \geq 3$, and $n > (t+1)(k-t+1)$. Suppose that $\mathcal{H} \subset \binom{[n]}{k}$ is a t-intersecting family and*

$$|\mathcal{H}| = m > \begin{cases} \max\{|\mathrm{AK}(n,k,t)|, |\mathrm{EM}(n,k,t+2,t+1)|\}, & \text{if } t < \frac{k-1}{2}, \\ |\mathrm{EM}(n,k,t+2,t+1)|, & \text{if } t \geq \frac{k-1}{2}. \end{cases}$$

Then $\mathcal{H} \subset \mathrm{EM}(n,k,t,t)$.

Proof of Theorem 6.3.6. Suppose $\mathcal{H}$ is given as in Theorem 6.3.6. By Theorem 6.3.26, $\mathcal{H} \subset \mathrm{EM}(n,k,t,t)$ and by Corollary 6.3.18, we have $|\partial_\ell \mathcal{H}| \geq |\partial_\ell L_m \mathrm{EM}(n,k,t,t)|$.

We now show that the value of $m(n,k,t)$ in the theorem cannot be reduced for $t \geq \frac{k-1}{2}$ and is tight up to a constant multiplicative factor for $t < \frac{k-1}{2}$. Indeed, we would just take $\mathcal{H} = \mathrm{EM}(n,k,t+2,t+1)$ and hence it suffices to prove the following.

Fact 6.3.27. *Let n be sufficiently large and $m = |\mathrm{EM}(n,k,t+2,t+1)|$. Then*

$$|\partial_\ell \mathrm{EM}(n,k,t+2,t+1)| < |\partial_\ell L_m \mathrm{EM}(n,k,t,t)| \quad \text{for all} \quad 1 \leq \ell \leq t.$$

In particular, for $1 \leq \ell \leq t$ the lower bound $m(n,k,t)$ for $|\mathcal{H}|$ in Theorem 6.3.6 cannot be reduced to be less than $|\mathrm{EM}(n,k,t+2,t+1)| \sim (t+2)\binom{n}{k-t-1}$.

Note that when $t < \frac{k-1}{2}$ Fact 6.3.27 implies that the constant multiplicative factor is at most $|\mathrm{AK}(n,k,t)|/|\mathrm{EM}(n,k,t+2,t+1)| \sim \frac{k-t+1}{t+2}$ which is independent of n.

Let $x \in \mathbb{R}$ such that $\binom{x-t}{k-t} = |\mathrm{EM}(n,k,t+2,t+1)| = (t+2)\binom{n-t-2}{k-t-1} + \binom{n-t-2}{k-t-2}$, then $x = \Theta(n^{\frac{k-t-1}{k-t}})$. Applying Lemma 6.3.15 to $\mathrm{EM}(n,k,t,t)$, we obtain

$$\begin{aligned}
|\partial_\ell L_m \mathrm{EM}(n,k,t,t)| &\geq \sum_{i=t-\ell}^{k-\ell} \binom{t}{i}\binom{x-t}{k-\ell-i} \\
&= \binom{t}{t-\ell}\binom{x-t}{k-t} + (1+o(1))\binom{t}{t-\ell+1}\binom{x-t}{k-t-1} \\
&= (t+2)\binom{t}{t-\ell}\binom{n}{k-t-1} + \Theta(n^{\frac{(k-t-1)^2}{k-t}}),
\end{aligned}$$

and

$$\begin{aligned}
|\partial_\ell \mathrm{EM}(n,k,t+2,t+1)| &= \sum_{i=t+1-\ell}^{k-\ell} \binom{t+2}{i}\binom{n-t+2}{k-\ell-i} \\
&= \binom{t+2}{t+1-\ell}\binom{n}{k-t-1} + \Theta(n^{k-t-2}).
\end{aligned}$$

If $\ell < t$, then

$$|\partial_\ell L_m \mathrm{EM}(n,k,t,t)| \sim (t+2)\binom{t}{t-\ell}\binom{n}{k-t-1}$$

and

$$|\partial_\ell \mathrm{EM}(n,k,t+2,t+1)| \sim \binom{t+2}{t+1-\ell}\binom{n}{k-t-1}.$$

Since $\binom{t+2}{t+1-\ell} < (t+2)\binom{t}{t-\ell}$ for $\ell < t$,

$$|\partial_\ell \mathrm{EM}(n,k,t+2,t+1)| < |\partial_\ell L_m \mathrm{EM}(n,k,t,t)|$$

for large n.

If $\ell = t$, then

$$|\partial_\ell L_m \mathrm{EM}(n,k,t,t)| \sim (t+2)\binom{n}{k-t-1} + \Theta(n^{\frac{(k-t-1)^2}{k-t}})$$

and

$$|\partial_\ell \mathrm{EM}(n,k,t+2,t+1)| \sim (t+2)\binom{n}{k-t-1} + \Theta(n^{k-t-2}).$$

Since $\frac{(k-t-1)^2}{k-t} > k-t-2$,

$$|\partial_{t+1}\mathrm{EM}(n,k,t+2,t+1)| < |\partial_{t+1} L_m \mathrm{EM}(n,k,t,t)|$$

for large n. Consequently, Fact 6.3.27 holds and the proof is complete. ∎

6.3.2.6 Proof of Theorem 6.3.10

Before proving Theorem 6.3.10 we need some structure theorems for a family with large size and a given matching number.

Definition 6.3.28. *Let $n \geq sk+1$, $k \geq 3$, and $s \geq 1$. Let $v_0, \ldots, v_{s-1} \in [n]$ be distinct vertices, $T_1, \ldots, T_s \subset [n]$ be pairwise disjoint k-sets, and $v_i \in T_i$ for all $1 \leq i \leq s-1$, and $v_0 \notin T_i$ for all $1 \leq i \leq s$. Let*

$$\mathrm{PF}(n,k,s) = \{T_1, \ldots, T_s\} \cup \left\{ A \in \binom{[n]}{k} : \exists 0 \leq i \leq s-1 \text{ such that } x_i \in A \text{ and } |A \cap \bigcup_{j=i+1}^{s} T_i| \geq 1 \right\}.$$

Notice that $|\mathrm{PF}(n,k,s)| \sim k\binom{s+1}{2}\binom{n}{k-2}$.

Theorem 6.3.29 (Kostochka–Mubayi [148])**.** *For every $k \geq 3$, $s \geq t \geq 2$, there exists n_0 such that the following holds for all $n \geq n_0$. Suppose that $\mathcal{H} \subset \binom{[n]}{k}$ satisfies $\nu(\mathcal{H}) = s$ and*

$$|\mathcal{H}| > \begin{cases} |\mathrm{EM}(n,3,s-t,1)| + |\mathrm{EM}(n-s+t,3,2s+1,2)| & \text{if } k = 3, \\ |\mathrm{EM}(n,k,s-t,1)| + |\mathrm{PF}(n-s+t,k,t)| & \text{if } k \geq 4. \end{cases}$$

Then there exists $X \subset [n]$ with $|X| = s-t+1$ such that $\nu(H - X) = t-1$. The bound on $|\mathcal{H}|$ is tight. In particular, if

$$|\mathcal{H}| > \begin{cases} |\mathrm{EM}(n,3,2s+1,2)| & \text{for } k = 3, \\ |\mathrm{PF}(n,k,s)| & \text{for } k \geq 4, \end{cases}$$

then there exists $v \in [n]$ such that $\nu(H - v) = s-1$.

Note that the proof of Theorem 6.3.29 was not included in [148], but one can easily prove it using results in [96] (Theorems 4.1 and 4.2 in [96]).

We also need the following structure theorems for intersecting families. Let

- $H_0^3(n) = \left\{A \in \binom{[n]}{3} \colon |A \cap [3]| \geq 2\right\}$.
- $H_1^3(n) = \left\{A \in \binom{[n]}{3} \colon 1 \in A \text{ and } |A \cap \{2,3,4\}| \geq 1\right\} \cup \{234\}$.
- $H_2^3(n) = \left\{A \in \binom{[n]}{3} \colon 1 \in A \text{ and } |A \cap \{2,3\}| \geq 1\right\} \cup \{234, 235, 145\}$.
- $H_3^3(n) = \left\{A \in \binom{[n]}{3} \colon \{1,2\} \in A\right\} \cup \{134, 135, 145, 234, 235, 245\}$.
- $H_4^3(n) = \left\{A \in \binom{[n]}{3} \colon \{1,2\} \in A\right\} \cup \{134, 156, 235, 236, 245, 246\}$.
- $H_5^3(n) = \left\{A \in \binom{[n]}{3} \colon \{1,2\} \in A\right\} \cup \{134, 156, 136, 235, 236, 246\}$.
- For $k \geq 4$ and $0 \leq i \leq 5$, $H_i^k(n) = \left\{A \in \binom{[n]}{k} \colon \exists B \in H_i^3(n) \text{ such that } B \subset A\right\}$.

Fact 6.3.30. *The following holds for all $n \geq k \geq 3$.*

- $|H_0^k(n)| = 3\binom{n-3}{k-2} + \binom{n-3}{k-3} < 3\binom{n}{k-2} - 2\binom{n}{k-3}$.
- $|H_1^k(n)| = 3\binom{n-4}{k-2} + 4\binom{n-4}{k-3} + \binom{n-4}{k-4} < 3\binom{n}{k-2} - 2\binom{n}{k-3}$.
- $\max\{|H_i^k(n)| \colon 2 \leq i \leq 5\} \leq 2\binom{n}{k-2}$.

Definition 6.3.31. *Let $n \geq 2k$ and $k \geq 3$. Let $Y = [2, k+1], Z = [k+2, 2k]$. The n-vertex k-graph $\mathrm{PF}(n,k)$ consists of all k-subsets of $[n]$ containing a member of the family*

$$G = \{A \colon 1 \in A \text{ and } |A \cap Y| = 1 \text{ and } |A \cap Z| = 1\} \cup$$
$$\{Y, \{1, k, k+1\}, Z \cup \{k\}, Z \cup \{k+1\}\}.$$

Note that $|\mathrm{PF}(n,k)| = O(n^{k-3})$.

Theorem 6.3.32 (Kostochka–Mubayi [148]). *Let $k \geq 4$ be fixed and n be sufficiently large. Then there is $C > 0$ such that for every intersecting n-vertex k-graph $\mathcal{H}$ with $|\mathcal{H}| > |\mathrm{PF}(n,k)| = O(n^{k-3})$, one can remove from $\mathcal{H}$ at most Cn^{k-4} edges so that the resulting k-graph $\mathcal{H}'$ is contained in one of $H_0^k(n), \ldots, H_5^k(n), \mathrm{EKR}(n,k)$.*

For intersecting 3-graphs there is a stronger result.

Theorem 6.3.33 (Kostochka–Mubayi [148]). *Let $\mathcal{H}$ be an intersecting 3-graph and $n = |V(\mathcal{H})| \geq 6$. If $\tau(\mathcal{H}) \leq 2$, then $\mathcal{H}$ is contained in one of $\mathrm{EKR}(n,3), H_1^3(n), \ldots, H_5^3(n)$.*

The following result shows that the size of an intersecting 3-graph $\mathcal{H}$ with $\tau(\mathcal{H}) \geq 3$ is bounded by a constant.

Theorem 6.3.34 (Frankl [89]). *Let $k \geq 3$ and n be sufficiently large. Then every intersecting n-vertex k-graph $\mathcal{H}$ with $\tau(\mathcal{H}) \geq 3$ satisfies $|\mathcal{H}| \leq |\mathrm{PF}(n,k)|$. Moreover, if $k \geq 4$, then equality holds only if $\mathcal{H} \cong \mathrm{PF}(n,k)$.*

Now we are ready to prove Theorem 6.3.10.

Proof of Theorem 6.3.10. Let n be sufficiently large and $c = c(k,s)$ be given by Equation 6.6. We may assume that $\mathcal{H}$ is shifted and $\nu(\mathcal{H}) = s$. For every $v \in [n]$ let $d_{\mathcal{H}}(v) = |\{A \in \mathcal{H} \colon v \in A\}|$, and let $\Delta = \max\{d_{\mathcal{H}}(v) : v \in [n]\}$. Suppose that $m = \sum_{i=h}^{k} \binom{a_i}{i} - \sum_{i=h}^{k} \binom{a_i - s}{i}$ for some integers $a_k > \cdots > a_h \geq h \geq 1$. Then by Lemma 6.3.15, it suffice to show that

$$|\partial \mathcal{H}| \geq \sum_{i=h}^{k} \binom{a_i}{i-1} = |\partial L_m \mathrm{EM}(n,k,s,1)|.$$

Note that $a_k \to \infty$ as $n \to \infty$ and so $m \sim s\binom{a_k-1}{k-1}$.

Claim 6.3.35. $\Delta \leq \sum_{i=h}^{k} \binom{a_i-1}{i-1} = (1+o(1))\frac{m}{s}$.

Proof. Suppose that there exists $v \in [n]$ with $d_{\mathcal{H}}(v) > \sum_{i=h}^{k} \binom{a_i-1}{i-1}$. Let $\mathcal{H}(v) = \{A \setminus \{v\} \colon v \in A \in \mathcal{H}\}$. Then by the Kruskal–Katona theorem,

$$|\partial\mathcal{H}| \geq |\mathcal{H}(v)| + |\partial\mathcal{H}(v)| > \sum_{i=h}^{k} \binom{a_i-1}{i-1} + \sum_{i=h}^{k} \binom{a_i-1}{i-2} = \sum_{i=h}^{k} \binom{a_i}{i-1},$$

and we are done. ∎

We are going to use Theorem 6.3.29 and Claim 6.3.35 to define a sequence of distinct vertices $v_1, \ldots, v_{s-1}$ and a sequence of k-graphs $\mathcal{H}_1, \ldots, \mathcal{H}_{s-1}$ such that $\nu(\mathcal{H}_i) = s - i$ and $|\mathcal{H}_i| > (1-o(1))\frac{s-i}{s}m$ for all $1 \leq i \leq s-1$. Since $\mathcal{H}$ is shifted, we may assume that $v_i = i$ for $1 \leq i \leq s-1$.

First, by the assumption on the size of $\mathcal{H}$ and Theorem 6.3.29, there exists $v_1 \in [n]$ such that $\mathcal{H}_1 := \mathcal{H} - v_1$ satisfies $\nu(\mathcal{H}_1) = s-1$. By Claim 6.3.35, $d_{\mathcal{H}}(v_1) < (1+o(1))m/s$, so $|\mathcal{H}_1| > (1-o(1))\frac{s-1}{s}m$.

Now suppose that we have defined $\mathcal{H}_i$ for some $1 \leq i \leq s-2$ such that $\nu(\mathcal{H}_i) = s-i$ and $|\mathcal{H}_i| > (1-o(1))\frac{s-i}{s}m$. Since

$$|\mathcal{H}_i| > (1-o(1))\frac{s-i}{s}m \geq \frac{s-i}{s}c\binom{n}{k-2} \geq \begin{cases} |\mathrm{EM}(n,3,2(s-i)+1,2)|, & \text{for } k=3, \\ |\mathrm{PF}(n,k,s-i)|, & \text{for } k \geq 4, \end{cases}$$

by Theorem 6.3.29, there exists $v_{i+1} \in [n]$ such that $\mathcal{H}_{i+1} := \mathcal{H}_i - v_{i+1}$ satisfies $\nu(\mathcal{H}_{i+1}) = s - i - 1$. By Claim 6.3.35, $|\mathcal{H}_{i+1}| > (1 - o(1))\frac{s-i-1}{s}m$.

Note that $\mathcal{H}_{s-1}$ satisfies $\nu(\mathcal{H}_{s-1}) = 1$ and

$$|\mathcal{H}_{s-1}| > (1 - o(1))\frac{1}{s}m \geq \frac{1}{s}c\binom{n}{k-2} \geq 3\binom{n}{k-2} > |\mathrm{PF}(n,k)|.$$

If $k = 3$, then by Theorem 6.3.34, $\tau(\mathcal{H}_{s-1}) \leq 2$. Therefore, by Theorem 6.3.33, $\mathcal{H}$ is contained in one of $\mathrm{EKR}(n,3), H_1^3(n), \ldots, H_5^3(n)$. Since $|\mathcal{H}_{s-1}| > 3n$ and by Fact 6.3.30, $\max_{0\leq i\leq 5}\{|H_i^3|\} \leq 3n - 8$, we must have $\mathcal{H} \subset \mathrm{EKR}(n,3) = \mathrm{EM}(n,3,1,1)$. Note that $\mathcal{H}_{s-1}$ is obtained from $\mathcal{H}$ by removing $s-1$ vertices, so $\mathcal{H} \subset \mathrm{EM}(n,3,s,1)$. Therefore, by Lemma 6.3.16, $|\partial\mathcal{H}| \geq |\partial L_m \mathrm{EM}(n,3,s,1)|$ and we are done.

Now we may assume that $k \geq 4$. Then, by Theorem 6.3.32, one can remove at most Cn^{k-4} edges from $\mathcal{H}_{s-1}$ such that the resulting k-graph $\mathcal{H}'$ is contained in one of $H_0^k(n), \ldots, H_5^k(n)$, $\mathrm{EKR}(n,k)$. Note that $|\mathcal{H}'| \geq |\mathcal{H}_{s-1}| - Cn^{k-4} > 3\binom{n}{k-2} - \binom{n}{k-3}$ and by Fact 6.3.30, $\max_{0\leq i\leq 5}\{|H_i^k|\} < 3\binom{n}{k-2} - 2\binom{n}{k-3}$, so $\mathcal{H}' \subset \mathrm{EKR}(n,k) = \mathrm{EM}(n,k,1,1)$. Here we need n to be sufficient large so that $Cn^{k-4} < \binom{n}{k-3}$.

Note that $\mathcal{H}_{s-1}$ is obtained from $\mathcal{H}$ by removing $s-1$ vertices. If $\mathcal{H}_{s-1} \subset \mathrm{EM}(n,k,1,1)$, then $\mathcal{H} \subset \mathrm{EM}(n,k,s,1)$ and by Lemma 6.3.16 we are done. So we may assume that $\mathcal{H}_{s-1} \not\subset \mathrm{EM}(n,k,1,1)$, i.e. $\mathcal{H}_{s-1} \setminus \mathcal{H}' \neq \emptyset$. Let $A \in \mathcal{H}_{s-1} \setminus \mathcal{H}'$ and since $\mathcal{H}_{s-1}$ is shifted, we may assume that $A = \{s+1, \ldots, s+k\}$. Since $\mathcal{H}_{s-1}$ is intersecting, every edge in $\mathcal{H}'$ must have nonempty intersecting with A. So $\mathcal{H}' \subset \mathrm{HM}(n,k,1,k)$. This implies that one can remove at most Cn^{k-4} edges from $\mathcal{H}$ such that the resulting k-graph $\mathcal{H}''$ satisfies $\mathcal{H}'' \subset \mathrm{HM}(n,k,s,k)$.

Let $y \in \mathbb{R}$ satisfy $|\mathcal{H}''| = \binom{y}{k} - \binom{y-s}{k} - \binom{y-s-k}{k-1}$. Then by Lemma 6.3.23, $|\partial\mathcal{H}| \geq |\partial\mathcal{H}''| \geq \binom{y}{k-1}$. Let $x \in \mathbb{R}, a_k, \dots, a_h \in \mathbb{N}$ such that $a_k > \cdots > a_h \geq h \geq 1$ and $|\mathcal{H}| = \binom{x}{k} - \binom{x-s}{k} = \sum_{i=h}^{k}\binom{a_i}{i} - \sum_{i=h}^{k}\binom{a_i-s}{i}$. It is easy to see that $x \leq a_k + 1$.

Claim 6.3.36. $y > x + 1$.

Proof. Suppose not. Then

$$\begin{aligned}&\binom{x+1}{k} - \binom{x+1-s}{k} - \binom{x+1-s-k}{k-1}\\ \geq&\binom{y}{k} - \binom{y-s}{k} - \binom{y-s-k}{k-1} \geq \binom{x}{k} - \binom{x-s}{k} - Cn^{k-4}.\end{aligned}$$

Since $|\mathcal{H}| \geq c\binom{n}{k-2}$, $x = \Omega(n^{\frac{k-2}{k-1}})$. Therefore,

$$\begin{aligned}&\binom{x+1}{k} - \binom{x+1-s}{k} - \binom{x+1-s-k}{k-1}\\ =&\binom{x}{k} - \binom{x-s}{k} + \binom{x}{k-1} - \binom{x-s}{k-1} - \binom{x+1-s-k}{k-1}\\ <&\binom{x}{k} - \binom{x-s}{k} - \frac{1}{2}\binom{x-s-k}{k-1}\\ <&\binom{x}{k} - \binom{x-s}{k} - Cn^{k-4},\end{aligned}$$

a contradiction. This completes the proof of Claim 6.3.36. ∎

By Claim 6.3.36,

$$|\partial\mathcal{H}| \geq \binom{y}{k-1} > \binom{x+1}{k-1} \geq \sum_{i=h}^{k} \binom{a_i}{i-1}.$$

This completes the proof of Theorem 6.3.10. ∎

6.3.3 Concluding Remarks

Let $\mathcal{H} \subset \binom{[n]}{3}$ be an intersecting family with $|\mathcal{H}| \geq \mathrm{PF}(n,3) = 10$. Then Theorems 6.3.33 and 6.3.34 completely determine the structure of $\mathcal{H}$. One can use this structural result to determine the minimum size of $|\partial_\ell \mathcal{H}|$ completely for $1 \leq \ell \leq 2$. However, the calculation is very complicated and tedious, so we omit it here.

As we mentioned before, for $1 \leq \ell \leq t$ the lower bound for $|\mathcal{H}|$ in Theorem 6.3.6 above is tight for $t \geq \frac{k-1}{2}$ and can be improved for $t < \frac{k-1}{2}$. Indeed, one can use the Δ-system method (see [148]) to prove the following result.

Theorem 6.3.37. *Let $t \geq 1$, $k \geq 3$, $\epsilon > 0$, and n be sufficiently large. Suppose that $\mathcal{H} \subset \binom{[n]}{k}$ is t-intersecting. If $t < \frac{k-1}{2}$ and $|\mathcal{H}| > (k-t+\epsilon)\binom{n}{k-t-1}$, then either $\mathcal{H} \subset \mathrm{AK}(n,k,t)$ or $\mathcal{H} \subset \mathrm{EM}(n,k,t,t)$. If $t \geq \frac{k-1}{2}$ and $|\mathcal{H}| > (t+1+\epsilon)\binom{n}{k-t-1}$, then $\mathcal{H}$ is contained in one of $\mathrm{AK}(n,k,t), \mathrm{EM}(n,k,t+2,t+1), \mathrm{EM}(n,k,t,t)$.*

One can easily use Corollary 6.3.18 to show that for $1 \leq \ell \leq t$, $|\partial_\ell L_m \mathrm{AK}(n,k,t)| > |\partial_\ell L_m \mathrm{EM}(n,k,t,t)|$ for sufficiently large n and m. Therefore, by Theorem 6.3.37, we obtain the following result.

Fact 6.3.38. *Let $k \geq 3$, $t < \frac{k-1}{2}$, $1 \leq \ell \leq t$, $\epsilon > 0$, and n be sufficiently large. Then every t-intersecting family $\mathcal{H} \subset \binom{[n]}{k}$ with $|\mathcal{H}| = m > (k - t + \epsilon)\binom{n}{k-t-1}$ satisfies $|\partial_\ell \mathcal{H}| \geq |\partial_\ell L_m \mathrm{EM}(n, k, t, t)|$.*

6.4 Hypergraphs without non-trivial subgraphs

6.4.1 Introduction

For $d \geq 2$ a hypergraph $\mathcal{F}$ is d-wise-intersecting if $\bigcap_{i\in[d]} E_i \neq \emptyset$ for all $E_1, \ldots, E_d \in \mathcal{F}$, and $\mathcal{F}$ is non-trivial d-wise-intersecting if it is d-wise-intersecting but $\bigcap_{E\in\mathcal{F}} E = \emptyset$. If $d = 2$, then we simply call $\mathcal{F}$ intersecting and non-trivial intersecting, respectively.

A d-simplex is a collection of $d+1$ sets $\{A_1, \ldots, A_{d+1}\}$ such that $\bigcap_{i\neq j} A_i \neq \emptyset$ for all $j \in [d+1]$, but $\bigcap_{i\in[d+1]} A_i = \emptyset$. The Chvátal Simplex Conjecture [41] states that for every $k \geq d+1 \geq 3$ and $n \geq (d+1)k/d$ if a hypegraph $\mathcal{H} \subset \binom{[n]}{k}$ does not contain a d-simplex as a subgraph, then $|\mathcal{H}| \leq \binom{n-1}{k-1}$, with equality only if $\mathcal{H}$ is a star, i.e. all sets in $\mathcal{H}$ contain a fixed vertex. The case $k = d+1$ was proved by Chvátal [41]. Mubayi and Verstraëte [198] proved the conjecture for all $k \geq 3$ and $d = 2$. Recently, Currier [47] proved this conjecture for all $k \geq d+1 \geq 3$ and $n \geq 2k$. The Chvátal Simplex Conjecture is still open in general for $n < 2k$ and $3 \leq d \leq k-2$, and we refer the reader to [19; 46; 101; 88; 90; 140; 145; 157] and their references for more results related to this conjecture.

It is easy to see that the family of all d-simplexes is the same as the family of all non-trivial d-wise-intersecting hypergraphs of size $d+1$, and if a hypergraph is d-wise-intersecting, then it is also d'-wise-intersecting for all $2 \leq d' \leq d$.

In the proof for the Chvátal Simplex Conjecture for $d = 2$ Mubayi and Verstraëte actually proved the following stronger result.

Theorem 6.4.1 (Mubayi–Verstraëte [198]). *Let $k \geq d+1 \geq 3$ and $n \geq (d+1)k/d$. Suppose that $\mathcal{H} \subset \binom{[n]}{k}$ contains no non-trivial intersecting subgraph of size $d+1$. Then $|\mathcal{H}| \leq \binom{n-1}{k-1}$, with equality only if $\mathcal{H}$ is a star.*

Mubayi and Verstraëte also remarked that their proof of Theorem 6.4.1 actually works for d slightly greater than k as well, and they posed the following conjecture.

Conjecture 6.4.2 (Mubayi–Verstraëte [198]). *Let $d \geq k \geq 4$ and n be sufficiently large. Suppose that $\mathcal{H} \subset \binom{[n]}{k}$ contains no non-trivial intersecting subgraph of size $d+1$. Then $|\mathcal{H}| \leq \binom{n-1}{k-1}$, with equality only if $\mathcal{H}$ is a star.*

Let $m \geq 2$. Recall that a Steiner $(n, 3, m-1)$-system is a 3-graph $\mathcal{S}$ on n vertices such that every pair of vertices in $V(\mathcal{S})$ is contained in exactly $m-1$ edges of $\mathcal{S}$. It follows from Keevash's result [138] that if n is a multiple of 3 and sufficiently large, then there exists a Steiner $(n, 3, m-1)$-system.

Notice that a Steiner $(n, 3, m-1)$-system has size $\frac{m-1}{3}\binom{n}{2}$, which is greater than $\binom{n-1}{2}$ when $m \geq 4$. It was observed by Mubayi and Verstraëte [198] that a Steiner $(n, 3, m-1)$-system does not contain non-trivial intersecting subgraph of size $3m+1$. Therefore, they made the following conjecture for 3-graphs.

Conjecture 6.4.3 (Mubayi–Verstraëte [198]). *Let $m \geq 4$ and n be sufficiently large. Suppose that $\mathcal{H} \subset \binom{[n]}{3}$ contains no non-trivial intersecting family of size $3m+1$. Then $|\mathcal{H}| \leq \frac{m-1}{3}\binom{n}{2}$, with equality holds iff $\mathcal{H}$ is a Steiner $(n, 3, m-1)$-system.*

In this section, we confirm Conjecture 6.4.2 by proving a stronger statement (Theorem 6.4.6), and disprove Conjecture 6.4.3 by showing a construction with more than $\frac{m-1}{3}\binom{n}{2}$ edges and contains no non-trivial intersecting subgraph of size $3m+1$.

Let $s \geq 2$. A family $\mathcal{D} = \{D_1, \ldots, D_s\}$ is a Δ-system (or a sunflower) if $D_i \cap D_j = C$ for all $\{i, j\} \subset [s]$. The set C is called the center of $\mathcal{D}$.

Definition 6.4.4. *Let $k, d \geq p \geq 2$, and $\vec{a} = (a_1, \ldots, a_p)$, $\vec{b} = (b_1, \ldots, b_p)$ be two sequences of positive integers with $\sum_{i=1}^{p} a_i = k$.*

(1) An $\vec{a}$-partition of a k-set E is a partition $E = \bigcup_{i\in[p]} A_i$ such that $|A_i| = a_i$ for $i \in [p]$.

(2) A semi-$(\vec{a}, \vec{b})$-Δ-system is a collection of sets $\{E_0, E_1^1, \ldots, E_1^{b_1}, \ldots, E_p^1, \ldots, E_p^{b_p}\}$ such that for some $\vec{a}$-partition of $E_0 = \bigcup_{i\in[p]} A_i$, the family $\{E_0, E_i^1, \ldots, E_i^{b_i}\}$ is a Δ-system with center $E_0 \setminus A_i$ for all $i \in [p]$. The set E_0 is called the host of this semi-$(\vec{a}, \vec{b})$-Δ-system.

(3) An $(\vec{a}, \vec{b})$-Δ-system is a semi-$(\vec{a}, \vec{b})$-Δ-system $\{E_0, E_1^1, \ldots, E_1^{b_1}, \ldots, E_p^1, \ldots, E_p^{b_p}\}$ such that sets $E_1^1 \setminus E_0, \ldots, E_1^{b_1} \setminus E_0, \ldots, E_p^1 \setminus E_0, \ldots, E_p^{b_p} \setminus E_0$ are pairwise disjoint.

(4) An $(\vec{a}, d)$-Δ-system is a $(\vec{a}, \vec{b})$-Δ-system for some $\vec{b}$ such that $\sum_{i=1}^{p} b_i = d$.

From the definitions one can easily obtain the following observation.

Observation 6.4.5. *Let $k, d \geq p \geq 3$ and $\vec{a} = (a_1, \ldots, a_p)$ be a sequence of integers with $\sum_{i=1}^{p} a_i = k$. Then an $(\vec{a}, d)$-Δ-system is a non-trivial $(p-1)$-wise-intersecting hypergraph with $d+1$ edges.*

An $(\vec{a}, d)$-Δ-system in which $d = p$, i.e. $b_1 = \cdots = b_p = 1$ was studied by Füredi and Özkahya in [85]. In this note we employ a machinery (a complicated version of the delta-system method) developed by them and even earlier by Frankl and Füredi [99], to obtain the following tight bound for the size of a hypergraph without $(\vec{a}, d)$-Δ-systems for all $d \geq p \geq 2$.

Theorem 6.4.6. *Let $k > p \geq 2$, $d \geq p$, and $\vec{a} = (a_1, \ldots, a_p)$ be a sequence of positive integers with $\sum_{i=1}^{p} a_i = k$. Suppose that $n \geq n_0(k, d)$ is sufficiently large and $\mathcal{H} \subset \binom{[n]}{k}$ does not contain a $(\vec{a}, d)$-Δ-system as a subgraph. Then $|\mathcal{H}| \leq \binom{n-1}{k-1}$, with equality only if $\mathcal{H}$ is a star.*

Remark. Our proof of Theorem 6.4.6 uses the delta-system method and Theorem 6.4.15 due to Füredi, so our lower bound for $n_0(k, d)$ is at least exponential in k and d. It would be interesting to find the minimum value of $n_0(k, d)$ such that the statement in Theorem 6.4.6 holds for all $n \geq n_0(k, d)$.

The following result is an immediate consequence of Theorem 6.4.6 and Observation 6.4.5.

Theorem 6.4.7. *Let $k > p \geq 3$, $d \geq p$. Suppose that $n \geq n_0(k, d)$ is sufficiently large and $\mathcal{H} \subset \binom{[n]}{k}$ does not contain a non-trivial $(p-1)$-wise-intersecting subgraph of size $d + 1$. Then $|\mathcal{H}| \leq \binom{n-1}{k-1}$, with equality only if $\mathcal{H}$ is a star.*

Note that Conjecture 6.4.2 is a special case of Theorem 6.4.7, i.e. $p = 3$.

We are also able to prove the following stability version of Theorem 6.4.6.

Theorem 6.4.8. *Let $k > p \geq 2$, $d \geq p$, and $\vec{a} = (a_1, \ldots, a_p)$ be a sequence of positive integers with $\sum_{i=1}^{p} a_i = k$. For every $\delta > 0$ there exists $\epsilon > 0$ and $n_0(k, d, \delta)$ such that the following holds for all $n \geq n_0(k, d, \delta)$. Suppose that $\mathcal{H} \subset \binom{[n]}{k}$ does not contain a $(\vec{a}, d)$-Δ-system as a subgraph, and $|\mathcal{H}| \geq (1 - \epsilon)\binom{n-1}{k-1}$. Then there exists a vertex $v \in [n]$ such that v is contained in all but at most δn^{k-1} edges in $\mathcal{H}$.*

For 3-graphs the following result shows that Conjecture 6.4.3 is not true in general.

Theorem 6.4.9. *Let $m \geq 4$, n be a multiple of 3 and sufficiently large. Then there exists a 3-graph $\widehat{\mathcal{S}}$ on n vertices with $\frac{m-1}{3}\binom{n}{2} + \frac{n}{3}$ edges and contains no non-trivial intersecting subgraph of size $3m+1$.*

6.4.2 Constructions

In this section we give a construction to show that Conjecture 6.4.3 is not true in general. We need the following structural theorem of intersecting 3-graphs due to Kostochka and Mubayi [148]. Define

- $\mathrm{EKR}(n) = \left\{A \in \binom{[n]}{3} \colon 1 \in A\right\}$.
- $H_0(n) = \left\{A \in \binom{[n]}{3} \colon |A \cap [3]| \geq 2\right\}$.
- $H_1(n) = \left\{A \in \binom{[n]}{3} \colon 1 \in A \text{ and } |A \cap \{2,3,4\}| \geq 1\right\} \cup \{234\}$.
- $H_2(n) = \left\{A \in \binom{[n]}{3} \colon 1 \in A \text{ and } |A \cap \{2,3\}| \geq 1\right\} \cup \{234, 235, 145\}$.
- $H_3(n) = \left\{A \in \binom{[n]}{3} \colon \{1,2\} \in A\right\} \cup \{134, 135, 145, 234, 235, 245\}$.
- $H_4(n) = \left\{A \in \binom{[n]}{3} \colon \{1,2\} \in A\right\} \cup \{134, 156, 235, 236, 245, 246\}$.
- $H_5(n) = \left\{A \in \binom{[n]}{3} \colon \{1,2\} \in A\right\} \cup \{134, 156, 136, 235, 236, 246\}$.

Theorem 6.4.10 (Kostochka–Mubayi [148])**.** *Every intersecting 3-graph with at least 11 edges is contained in one of* $\mathrm{EKR}(n), H_0(n), H_1(n), \ldots, H_5(n)$.

For a 3-graph $\mathcal{H}$ and $\{u,v\} \subset V(\mathcal{H})$ let $\deg_{\mathcal{H}}(uv)$ denote the number of edges in $\mathcal{H}$ that contain both u and v. Let $\Delta_2(\mathcal{H}) = \max\{\deg_{\mathcal{H}}(uv) \colon \{u,v\} \subset V(\mathcal{H})\}$.

Observation 6.4.11. *Let $\mathcal{H}$ be a 3-graph with e edges. If $\mathcal{H} \subset H_0(n)$, then $\Delta_2(\mathcal{H}) \geq \lceil \frac{e}{3} \rceil$. If $\mathcal{H} \subset H_2(n)$, then $\Delta_2(\mathcal{H}) \geq \lceil \frac{e-3}{2} \rceil$. If $\mathcal{H}$ is contained in $H_3(n)$, $H_4(n)$, or $H_5(n)$, then $\Delta_2(\mathcal{H}) \geq e - 6$.*

Now we define the construction. Let n be a multiple of 3 and sufficiently large. Let $\mathcal{S} \subset \binom{[n]}{3}$ be a Steiner $(n, 3, m-1)$-system. Then the complement of $\mathcal{S}$, which is $\bar{\mathcal{S}} := \binom{[n]}{3} \setminus \mathcal{S}$, satisfies that $d_{\bar{\mathcal{S}}}(uv) = n - m + 1$ for all $\{u, v\} \subset V(\mathcal{S})$. Therefore, by the Rödl–Ruciński–Szemerédi Theorem [221], $\bar{\mathcal{S}}$ contains a matching $\mathcal{M}$ with $n/3$ edges. Let $\widehat{\mathcal{S}} = \mathcal{S} \cup \mathcal{M}$. Then it is easy to see that

$$|\widehat{\mathcal{S}}| = |\mathcal{S}| + |\mathcal{M}| = \frac{m-1}{3}\binom{n}{2} + \frac{n}{3}.$$

The following proposition proves Theorem 6.4.9.

Proposition 6.4.12. *Let $m \geq 4$. Then $\widehat{\mathcal{S}}$ does not contain a non-trivial intersecting subgraph of size $3m+1$.*

Proof. Suppose to the contrary that this is not true. Let $\mathcal{F} \subset \widehat{\mathcal{S}}$ be a non-trivial intersecting subgraph with $3m + 1 \geq 11$ edges. By Theorem 6.4.10, $\mathcal{F}$ is contained in one of $H_0(n), H_1(n), \ldots, H_5(n)$. Notice that $\Delta_2(\mathcal{F}) \leq \Delta_2(\widehat{\mathcal{S}}) = m$. If $\mathcal{F}$ is contained in one of $H_0(n), H_2(n), \ldots, H_5(n)$, then by Observation 6.4.11, $\Delta_2(\mathcal{F}) \geq \min\left\{\lceil \frac{3m+1}{3} \rceil, \lceil \frac{3}{2}m - 1 \rceil, 3m-5\right\} > m$, a contradiction. Therefore, $\mathcal{F} \subset H_1(n)$. Then $\mathcal{F}$ contains four vertices v_0, v_1, v_2, v_3 such that $\deg_{\widehat{\mathcal{S}}}(v_0v_1) + \deg_{\widehat{\mathcal{S}}}(v_0v_2) + \deg_{\widehat{\mathcal{S}}}(v_0v_3) \geq 3m$, which implies $\deg_{\widehat{\mathcal{S}}}(v_0v_1) = \deg_{\widehat{\mathcal{S}}}(v_0v_2) =$

$\deg_{\widehat{\mathcal{S}}}(v_0v_3) = m$. However, this is impossible because the set $\{\{u,v\} \subset V(\mathcal{S})\colon \deg_{\widehat{\mathcal{S}}}(uv) = m\}$ consists of $n/3$ copies of pairwise vertex-disjoint triangles. ∎

6.4.3 Lemmas

In this section we present some preliminary lemmas for the proofs of Theorems 6.4.6 and 6.4.8. Our first lemma shows that a sufficiently large semi-$(\vec{a},\vec{c})$-Δ-system contains an $(\vec{a},\vec{b})$-Δ-system.

Lemma 6.4.13. *Let $k, d \geq p \geq 2$ and $\vec{a} = (a_1, \dots, a_p)$, $\vec{b} = (b_1, \dots, b_p)$, $\vec{c} = (c_1, \dots, c_p)$ be sequences of positive integers with $\sum_{i=1}^{p} a_i = k$. Suppose that $c_i \geq b_i + \sum_{j=1}^{i-1} a_j b_j$ for $i \in [p]$. Then every semi-$(\vec{a},\vec{c})$-Δ-system contains an $(\vec{a},\vec{b})$-Δ-system. In particular, if $c_1 \geq 1$ and $c_i \geq kd$ for $2 \leq i \leq p$, then every semi-$(\vec{a},\vec{c})$-Δ-system contains an $(\vec{a},d)$-Δ-system.*

Proof. Let $\mathcal{F} = \{E_0, E_1^1, \dots, E_1^{c_1}, \dots, E_p^1, \dots, E_p^{c_p}\}$ be a semi-$(\vec{a},\vec{c})$-Δ-system. Our goal is to choose $\{F_i^1, \dots, F_i^{b_i}\} \subset \{E_i^1, \dots, E_i^{c_i}\}$ for all $i \in [p]$ so that sets E_0, $F_1^1, \dots, F_1^{b_1}$, …, $F_p^1, \dots, F_p^{b_p}$ form a $(\vec{a},\vec{b})$-Δ-system.

Since $c_1 \geq b_1$, we can simply let $F_1^j = E_1^j$ for $j \in [b_1]$. Now suppose that we have defined sets $\{F_1^1, \dots, F_1^{b_1}, \dots, F_i^1, \dots, F_i^{b_i}\}$ for some $i \in [p-1]$. We are going to define sets $F_{i+1}^1, \dots, F_{i+1}^{b_{i+1}}$. Note that for every $1 \leq j \leq i$ and $1 \leq \ell \leq b_j$ the set $F_j^\ell \setminus E_0$ can have nonempty intersection with at most a_j sets in $\{E_{i+1}^1, \dots, E_{i+1}^{c_{i+1}}\}$. Since $c_{i+1} \geq b_{i+1} + \sum_{j=1}^{i} a_j b_j$, there exist at least b_{i+1} sets in $\{E_{i+1}^1, \dots, E_{i+1}^{c_{i+1}}\}$ that have empty intersection with all sets in $\{F_1^1 \setminus E_0, \dots, F_1^{b_1} \setminus E_0, \dots, F_i^1 \setminus E_0, \dots, F_i^{b_i} \setminus E_0\}$, and choose any b_i sets from them to form $\{F_{i+1}^1, \dots, F_{i+1}^{b_{i+1}}\}$.

The process terminates when $i = p$, and clearly, sets $E_0, F_1^1, \ldots, F_1^{b_1}, \ldots, F_p^1, \ldots, F_p^{b_p}$ form an $(\vec{a}, \vec{b})$-Δ-system.

Now suppose that $c_1 \geq 1$ and $c_i \geq kd$ for $2 \leq i \leq p$. Let $b_1 = 1$ and $b_i \geq 1$ for $2 \leq i \leq p$ such that $\sum_{i=2}^{p} b_i = d - 1$. Since $c_i \geq kd \geq b_i + \sum_{j=1}^{i-1} a_j b_j$, by the previous argument, $\mathcal{F}$ contains an $(\vec{a}, \vec{b})$-Δ-system, which is an $(\vec{a}, d)$-Δ-system. ∎

For a hypergraph $\mathcal{H}$ and $E \in \mathcal{H}$. The intersection structure of E with respect to $\mathcal{H}$ is

$$\mathcal{I}(E, \mathcal{H}) := \{E \cap E' : E' \in \mathcal{H} \setminus \{E\}\}.$$

A hypergraph $\mathcal{H} \subset \binom{[n]}{k}$ is k-partite if there exists a partition $[n] = V_1 \cup \cdots \cup V_k$ such that $|E \cap V_i| = 1$ for all $i \in [k]$. Suppose that $\mathcal{H}$ is k-partite with k parts $V_1, \ldots, V_k$. Then for every $S \subset [n]$, its projection is $\Pi(S) := \{i \colon S \cap V_i \neq \emptyset\}$. For every family $\mathcal{F} \subset 2^{[n]}$, its projection is $\Pi(\mathcal{F}) := \{\Pi(F) \colon F \in \mathcal{F}\}$.

Definition 6.4.14. *Let $s \geq 2$. A hypergraph $\mathcal{H} \subset \binom{[n]}{k}$ is s-homogeneous if it satisfies the following conditions.*

(1) $\mathcal{H}$ is k-partite.

(2) There exists a family $\mathcal{J} \subset 2^{[k]} \setminus \{[k]\}$ such that $\Pi(\mathcal{I}(E, \mathcal{H})) = \mathcal{J}$ for all $E \in \mathcal{H}$, where $\mathcal{J}$ is called the intersection pattern of $\mathcal{H}$.

(3) $\mathcal{J}$ is closed under intersection, i.e. if $A, B \in \mathcal{J}$, then $A \cap B \in \mathcal{J}$.

(4) For every $E \in \mathcal{H}$ every set in $\mathcal{I}(E, \mathcal{H})$ is the center of a Δ-system $\mathcal{D}$ of size s formed by edges of $\mathcal{H}$ and containing E, i.e. $E \in \mathcal{D} \subset \mathcal{H}$.

A hypergraph $\mathcal{H} \subset \binom{[n]}{k}$ *is homogeneous if it is* s*-homogeneous for some* $s \geq 2$.

Füredi [109] showed that for every $s \geq 2$, every hypergraph contains a large s-homogeneous subgraph.

Theorem 6.4.15 (Füredi [109]). *For every* $s, k \geq 2$*, there exists a constant* $c(k, s) > 0$ *such that every hypergraph* $\mathcal{H} \subset \binom{[n]}{k}$ *contains a* s*-homogeneous subgraph* $\mathcal{H}^*$ *with* $|\mathcal{H}^*| \geq c(k, s)|\mathcal{H}|$.

For a family $\mathcal{J} \subset 2^{[k]} \setminus \{[k]\}$ the rank of $\mathcal{J}$ is

$$r(\mathcal{J}) := \min\{|A| : A \subset [k], A \notin \mathcal{J} \text{ and } \nexists B \in \mathcal{J} \text{ such that } A \subset B\}.$$

It is easy to see from the definition that $r(\mathcal{J}) = k$ iff $\mathcal{J} = 2^{[k]} \setminus \{[k]\}$.

The following lemma gives an upper bound for the size of a homogeneous hypergraph $\mathcal{H}$ in terms of the rank of its intersection pattern and its shadow.

Lemma 6.4.16. *Let* $\mathcal{H} \subset \binom{[n]}{k}$ *be a homogeneous hypergraph with intersection pattern* $\mathcal{J} \subset 2^{[k]} \setminus \{[k]\}$*. Then* $|\mathcal{H}| \leq |\partial_{k-r(\mathcal{J})}\mathcal{H}|$.

Proof. Let $r = r(\mathcal{J})$. By the definition of rank, there exists an r-set $S \subset [k]$ that is not contained in $\mathcal{J}$, and moreover, every $T \subset [k]$ that contains S is also not contained in $\mathcal{J}$. Since $\Pi(\mathcal{I}(E, \mathcal{H})) = \mathcal{J}$ for all $E \in \mathcal{H}$, there exists an r-set in every $E \in \mathcal{H}$ that is not contained in any other edges in $\mathcal{H}$. Therefore, $|\mathcal{H}| \leq |\partial_{k-r}\mathcal{H}|$. ∎

The following lemma shows that if a hypergraph is s-homogeneous for sufficiently large s and does not contain an $(\vec{a}, d)$-Δ-system as a subgraph, then the rank of its intersection pattern is at most $k-1$.

Lemma 6.4.17. *Let $d \geq p \geq 2$, $k > p$, $s \geq kd+1$, and $\vec{a} = (a_1, \ldots, a_p)$ be a sequence of positive integers with $\sum_{i=1}^{p} a_i = k$. Let $\mathcal{H} \subset \binom{[n]}{k}$ be a s-homogeneous hypergraph with intersection pattern $\mathcal{J} \subset 2^{[k]} \setminus \{[k]\}$. If $r(\mathcal{J}) = k$, then $\mathcal{H}$ contains an $(\vec{a}, d)$-Δ-system.*

Proof. Since $r(\mathcal{J}) = k$, $\mathcal{J} = 2^{[k]} \setminus \{[k]\}$. Let $E \in \mathcal{H}$ and let $\bigcup_{i=1}^{p} A_i = E$ be an $\vec{a}$-partition of E. Since $\Pi(\mathcal{I}(E, \mathcal{H})) = \mathcal{J}$, we have $E \setminus A_i \in \mathcal{I}(E, \mathcal{H})$ for all $i \in [p]$. Since $\mathcal{H}$ is s-homogeneous, there exists a Δ-system $\mathcal{D}_i$ of size s with center $E \setminus A_i$ for $i \in [p]$. By assumption, $s \geq kd+1$, therefore, by Lemma 6.4.13, $\mathcal{H}$ contains an $(\vec{a}, d)$-Δ-system. ∎

Lemmas 6.4.17 and 6.4.16, and Theorem 6.4.15 implies that following proposition.

Proposition 6.4.18. *Let $d \geq p \geq 2$, $k > p$, and $\vec{a} = (a_1, \ldots, a_p)$ be a sequence of positive integers with $\sum_{i=1}^{p} a_i = k$. Let $\mathcal{H} \subset \binom{[n]}{k}$ be a hypergraph that contains no $(\vec{a}, d)$-Δ-systems. Then there exists a constant $c(k, d) > 0$ such that $|\partial \mathcal{H}| \geq c(k, d)|\mathcal{H}|$.*

Proof. Let $s = kd+1$ and $\mathcal{H}^*$ be a maximum s-homogeneous subgraph of $\mathcal{H}$ with intersection pattern $\mathcal{J}$. Then by Theorem 6.4.15, $|\mathcal{H}^*| \geq c(k, d)|\mathcal{H}|$ for some constant $c(k, d) > 0$. Since $\mathcal{H}^*$ contains no $(\vec{a}, d)$-Δ-systems, by Lemma 6.4.17, $r(\mathcal{J}) \leq k-1$. So by Lemma 6.4.16, $|\mathcal{H}^*| \leq |\partial \mathcal{H}^*|$. Therefore, $|\partial \mathcal{H}| \geq |\partial \mathcal{H}^*| \geq |\mathcal{H}^*| \geq c(k, d)|\mathcal{H}|$. ∎

The next lemma gives another condition that guarantees a hypergraph to contain an $(\vec{a}, d)$-Δ-system as a subgraph.

Lemma 6.4.19. *Let $d \geq p \geq 2$, $k > p$, $s \geq kd+1$, and $\vec{a} = (a_1, \ldots, a_p)$ be a sequence of positive integers with $\sum_{i=1}^{p} a_i = k$. Let $\mathcal{H} \subset \binom{[n]}{k}$ and $\mathcal{H}^*$ be a s-homogeneous subgraph of $\mathcal{H}$. Let $E_0 \in \mathcal{H}^*$ and $\bigcup_{i \in [p]} A_i = E_0$ be an $\vec{a}$-partition of E_0. Suppose that there exists $i_0 \in [p]$ such that $E \setminus A_i \in \mathcal{I}(E, \mathcal{H}^*)$ for all $i \in [p] \setminus \{i_0\}$, and there exists $E \in \mathcal{H}$ such that $E \cap E_0 = E_0 \setminus A_{i_0}$, Then $\mathcal{H}$ contains an $(\vec{a}, d)$-Δ-system.*

Proof. Without loss of generality, we may assume that $i_0 = 1$. By assumption, $E \setminus A_i$ is the center of Δ-system of size $s \geq kd+1$ in $\mathcal{H}^* \subset \mathcal{H}$ for $2 \leq i \leq k$, and $E_0 \setminus A_1$ is the center of a Δ-system of size 2 in $\mathcal{H}$, i.e. $\{E_0, E\}$. So by Lemma 6.4.13, $\mathcal{H}$ contains an $(\vec{a}, d)$-Δ-system. ∎

Lemma 6.4.20. *Let $d \geq p \geq 2$, $k > p$, $s \geq kd+1$, and $\vec{a} = (a_1, \ldots, a_p)$ be a sequence of positive integers with $\sum_{i=1}^{p} a_i = k$. Let $\mathcal{H} \subset \binom{[n]}{k}$ be a hypergraph that does not contain $(\vec{a}, d)$-Δ-systems. Let $\mathcal{H}^*$ be a s-homogeneous subgraph of $\mathcal{H}$ with intersection pattern $\mathcal{J}$. Suppose that $r(\mathcal{J}) = k-1$ and $\mathcal{J}$ contains exactly $k-1$ $(k-1)$-sets. Let $v \in E \in \mathcal{H}^*$ be the vertex that is contained in all $(k-1)$-sets in $\mathcal{I}(E, \mathcal{H}^*)$. Then $v \in F$ for all $F \in \mathcal{H}$ that satisfies $|F \cap E| \geq k - a_1$*

Proof. Let $E = \{v_1, \ldots, v_k\} \in \mathcal{H}^*$ and suppose that v_1 is contained in all $(k-1)$-sets in $\mathcal{I}(E, \mathcal{H}^*)$. Let $F \in \mathcal{H}$ and suppose that $|E \cap F| = k - t$ for some $1 \leq t \leq a_1$, but $v_1 \notin F$. If $t = a_1$, then let $\bigcup_{i \in [p]} A_i = E$ be an $\vec{a}$-partition such that $A_1 = E \setminus F$. For $2 \leq i \leq p$ since $v_1 \in E \setminus A_i$, $E \setminus A_i \in \mathcal{I}(E, \mathcal{H}^*)$. Therefore, $E \setminus A_i$ is the canter of a Δ-system of size s in $\mathcal{H}^*$ for $2 \leq i \leq p$. Since $E \setminus A_1$ is the center of a Δ-system of size 2, i.e. $\{E, F\}$, by Lemma 6.4.13, $\mathcal{H}$ contains an $(\vec{a}, d)$-Δ-system, a contradiction. So, $t < a_1$.

Let $M \subset E$ such that $E \setminus F \subset M$ and $|M| = k - a_1 + t$. Since $|M| \leq k-1$ and $v_1 \in M$, $M \in \mathcal{I}(E, \mathcal{H}^*)$. Therefore, M is the center of a Δ-system of size s in $\mathcal{H}^*$, which means that there exists $E_1 \in \mathcal{H}^*$ such that $E_1 \cap E = M$ and $(E_1 \setminus E) \cap F = \emptyset$. This implies that $E_1 \cap F = M \setminus (E \setminus F)$ and $|E_1 \cap F| = k - a_1$. Since $\Pi(\mathcal{I}(E_1, \mathcal{H}^*)) = \Pi(\mathcal{I}(E, \mathcal{H}^*))$, applying the same argument as above to E_1 and F we obtain that $\mathcal{H}$ contains an $(\vec{a}, d)$-Δ-system, a contradiction. ∎

For a hypergraph $\mathcal{H}$ and $E \in \mathcal{H}$ the weight of E is

$$\omega_{\mathcal{H}}(E) := \sum_{E' \subset E,\ |E'|=k-1} \frac{1}{\deg_{\mathcal{H}}(E')},$$

where $\deg_{\mathcal{H}}(E')$ is the number of edges in $\mathcal{H}$ containing E'. We have the following identity:

$$\sum_{E \in \mathcal{H}} \omega_{\mathcal{H}}(E) = \sum_{E \in \mathcal{H}} \sum_{E' \subset E,\ |E'|=k-1} \frac{1}{\deg_{\mathcal{H}}(E')} = \sum_{E' \in \partial\mathcal{H}} \sum_{E \in \mathcal{H},\ E' \subset E} \frac{1}{\deg_{\mathcal{H}}(E')} = |\partial \mathcal{H}|. \tag{6.8}$$

The following lemma gives a lower bound for $\omega_{\mathcal{H}}(E)$ regarding the intersection structure of E in a homogeneous subgraph of $\mathcal{H}$.

Lemma 6.4.21. *Let $d \geq p \geq 2$, $k > p$, $s \geq kd + 1$, and $\vec{a} = (a_1, \ldots, a_p)$ be a sequence of positive integers with $\sum_{i=1}^{p} a_i = k$. Let $\mathcal{H} \subset \binom{[n]}{k}$ be a hypergraph that does not contain $(\vec{a}, d)$-Δ-systems. Let $\mathcal{H}^*$ be a s-homogeneous subgraph of $\mathcal{H}$ with intersection pattern $\mathcal{J}$. Suppose that $r(\mathcal{J}) = k - 1$. Then the following hold.*

(1) If $\mathcal{J}$ contains exactly $k-1$ $(k-1)$-sets, then every $E \in \mathcal{H}^*$ contains a $(k-1)$-subset that is not contained in any other edges in $\mathcal{H}$. In particular, $\omega_{\mathcal{H}}(E) \geq 1$ for all $E \in \mathcal{H}^*$.

(2) If $\mathcal{J}$ contains at most $k-2$ $(k-1)$-sets, then $\omega_{\mathcal{H}}(E) \geq \frac{k}{k-1}$ for all $E \in \mathcal{H}^*$.

Proof. We prove (1) first. We may assume that $a_1 \geq \cdots \geq a_k$, and note that $a_1 \geq 2$ since $\sum_{i=1}^{p} a_i = k > p$. Let $E = \{v_1, \ldots, v_k\} \in \mathcal{H}^*$. Since $\Pi(\mathcal{I}(E, \mathcal{H}^*)) = \mathcal{J}$, by assumption, there are exactly $k-1$ $(k-1)$-sets in $\mathcal{I}(E, \mathcal{H}^*)$. Without loss of generality we may assume that $E \setminus \{v_i\} \in \mathcal{I}(E, \mathcal{H}^*)$ for $2 \leq i \leq k$. We claim that $\{v_2, \ldots, v_k\}$ is not contain in any set in $\mathcal{H} \setminus \{E\}$. Indeed, if there exists $E_1 \in \mathcal{H}$ such that $\{v_2, \ldots, v_k\} \subset E_1$, then $|E_1 \cap E| \geq k-1$. So, by Lemma 6.4.20, $v_1 \in E_1$, which implies that $E_1 = E$.

Now we prove (2). Suppose that $\mathcal{J}$ has exactly $k-t$ $(k-1)$-sets for some $2 \leq t \leq k$. Let $E = \{v_1, \ldots, v_k\} \in \mathcal{H}^*$. Without loss of generality, we may assume that $E \setminus \{v_i\} \in \mathcal{I}(E, \mathcal{H}^*)$ for $t+1 \leq i \leq k$.

Claim 6.4.22. *There does not exist a $(t-1)$-set $I \subset [t]$ and $t-1$ distinct vertices $\{u_i : i \in I\}$, such that $(E \setminus \{v_i\}) \cup \{u_i\} \in \mathcal{H}$ for all $i \in I$.*

Proof. Suppose not, and without loss of generality we may assume that $F_i := (E \setminus \{v_i\}) \cup \{u_i\} \in \mathcal{H}$ for all $2 \leq i \leq t$, where $u_2, \ldots, u_t$ are distinct vertices.

By assumption $\mathcal{I}(E, \mathcal{H}^*)$ contains all $(k-1)$-sets that contain $\{v_1, \ldots, v_t\}$. Since $\mathcal{I}(E, \mathcal{H}^*)$ is closed under intersection, $\mathcal{I}(E, \mathcal{H}^*)$ contains all proper subsets of E that contain $\{v_1, \ldots, v_t\}$, i.e. if $A \subset \{v_{t+1}, \ldots, v_k\}$, then $E \setminus A \in \mathcal{I}(E, \mathcal{H}^*)$.

On the other hand, since $r(\mathcal{I}(E,\mathcal{H}^*)) = k-1 \geq k-2$ and $E \setminus \{v_i\}, E \setminus \{v_j\} \notin \mathcal{I}(E,\mathcal{H}^*)$ for $i,j \in [t]$, we have $E \setminus \{v_i, v_j\} \in \mathcal{I}(E,\mathcal{H}^*)$ for all $\{i,j\} \subset [t]$. This together with the previous argument and the property that $\mathcal{I}(E,\mathcal{H}^*)$ is closed under intersection imply that if $|A \cap \{v_1,\dots,v_t\}| \geq 2$, then $E \setminus A \in \mathcal{I}(E,\mathcal{H}^*)$.

Let $i_0 \in [p]$ such that $\sum_{i=1}^{i_0-1} a_i < t \leq \sum_{i=1}^{i_0} a_i$, and let $\ell = t - \sum_{i=1}^{i_0-1} a_i$. Recall that $a_1 \geq \cdots \geq a_p \geq 1$ and $a_1 \geq 2$. Suppose that $\ell \geq 2$. Then there exists an $\vec{a}$-partition $E = \bigcup_{i\in[p]} A_i$ such that $A_1,\dots,A_{i_0-1} \subset \{v_1,\dots,v_t\}$, $|A_{i_0} \cap \{v_1,\dots,v_t\}| \geq \ell \geq 2$, and $A_{i_0+1},\dots,A_p \subset \{v_{t+1},\dots,v_k\}$. Since $a_1 \geq \cdots \geq a_{i_0-1} \geq a_{i_0} \geq 2$, by the argument above, $E \setminus A_i \in \mathcal{I}(E,\mathcal{H}^*)$ for all $i \in [p]$. Therefore, $E \setminus A_i$ is the center of a Δ-system of size $s \geq kd+1$ in $\mathcal{H}^*$ for $i \in [p]$, so by Lemma 6.4.13, $\mathcal{H}$ contains an $(\vec{a},d)$-Δ-system, a contradiction. Therefore, $\ell = 1$.

Suppose that $a_{i_0} = 1$. Then let $E = \bigcup_{i\in[p]} A_i$ be an $\vec{a}$-partition such that $\bigcup_{i\in[i_0]} A_i = \{v_1,\dots,v_t\}$ and $v_1 \in A_1$. Since $A_1 \subset \{v_1,\dots,v_t\}$ and $|A_1| \geq 2$, $E \setminus A_1 \in \mathcal{I}(E,\mathcal{H}^*)$. So $E \setminus A_1$ is the center of a Δ-system of size s. Without loss of generality we may assume that $a_2 = \cdots = a_{i_0} = 1$ since other cases can be proved similarly. For $i_0+1 \leq i \leq p$ since $A_i \subset \{v_{t+1},\dots,v_k\}$, $E \setminus A_i \in \mathcal{I}(E,\mathcal{H}^*)$. So $E \setminus A_i$ is is the center of a Δ-system of size s for $i_0+1 \leq i \leq p$. Notice that by assumption for every $2 \leq i \leq i_0$ there exists $F_{j_i} \in \mathcal{H}$ such that $F_{j_i} \cap E = E \setminus \{v_{j_i}\}$ for $2 \leq j_i \leq t$, and moreover, $\{F_{j_i} \setminus E : 2 \leq i \leq i_0\}$ are distinct. Therefore, by a similar argument as in the proof of Lemma 6.4.13, $\mathcal{H}$ contains an $(\vec{a},d)$-Δ-system, a contradiction. Therefore, $a_{i_0} \geq 2$.

Suppose that $a_1 \geq 3$. Then let $E = \bigcup_{i\in[p]} A_i$ be an $\vec{a}$-partition such that $A_2 \cup \cdots \cup A_{i_0-1} \subset \{v_1,\dots,v_t\}$, $v_{t+1} \in A_1$ and $A_1 \setminus \{v_{t+1}\} \subset \{v_1,\dots,v_t\}$, and $\{v_1,\dots,v_t\} \setminus \left(\bigcup_{i\in[i_0-1]} A_i\right) \subset A_{i_0}$.

Then $|A_i \cap \{v_1, \ldots, v_t\}| \geq 2$ for all $i \in [i_0]$ and $A_j \subset \{v_{t+1}, \ldots, v_k\}$ for all $i_0 + 1 \leq j \leq p$. Therefore, $E \setminus A_i$ is the center of a Δ-system of size $s \geq kd + 1$ in $\mathcal{H}^*$ for $i \in [p]$, so by Lemma 6.4.13, $\mathcal{H}$ contains an $(\vec{a}, d)$-Δ-system, a contradiction. Therefore, $a_1 = 2$.

Suppose that $a_p = 1$. Then let $E = \bigcup_{i \in [p]} A_i$ be an $\vec{a}$-partition such that $A_1 \cup \cdots \cup A_{i_0 - 1} = \{v_1, \ldots, v_{t-1}\}$ and $A_p = \{v_t\}$. Then $E \setminus A_i$ is the center of a Δ-system of size $s \geq kd + 1$ in $\mathcal{H}^*$ for $i \in [p-1]$ and $E \setminus A_p$ is the center of a Δ-system of size 2 in $\mathcal{H}$, i.e. $\{E, F_t\}$. Therefore, by Lemma 6.4.13, $\mathcal{H}$ contains an $(\vec{a}, d)$-Δ-system, a contradiction. Therefore, $a_p = 2$.

Now we have $a_1 = \cdots = a_p = 2$ and $\ell = 1$. Then t is odd, $t \geq 3$, and k is even, $k > t$. Since $E \setminus \{v_k\} \in \mathcal{I}(E, \mathcal{H}^*)$, $E \setminus \{v_k\}$ is the center of a Δ-system of size s in $\mathcal{H}^*$. So there exists $F_k := (E \setminus \{v_k\}) \cup \{u_k\} \in \mathcal{H}^*$ such that $u_k \notin \{u_2, \ldots, u_t\}$. Let $F_k = \bigcup_{i \in [p]} A_i$ be an $\vec{a}$-partition such that $A_1 = \{v_2, u_k\}$, $A_2 = \{v_1, v_3\}$, and $A_i = \{v_{2i-2}, v_{2i-1}\}$ for $3 \leq i \leq p$. Then for every $i \in [p] \setminus \{1\}$, either $A_i \subset \{v_1, \ldots, v_t\}$ or $A_i \subset \{v_{t+1}, \ldots, v_{k-1}\}$. Since $\Pi(\mathcal{I}(F_k, \mathcal{H}^*)) = \Pi(\mathcal{I}(E, \mathcal{H}^*))$, $F_k \setminus A_i$ is the center of a Δ-system of size $s \geq kd + 1$ in $\mathcal{H}^*$ for $i \in [p] \setminus \{1\}$. Since $V \setminus A_1$ is the center of a Δ-system of size 2 in $\mathcal{H}$, i.e. $\{F_k, F_2\}$. Therefore, by Lemma 6.4.13, $\mathcal{H}$ contains an $(\vec{a}, d)$-Δ-system, a contradiction. This completes the proof of Claim 6.4.22. ∎

Define a bipartite graph G with two parts $L = \{v_1, \ldots, v_t\}$ and $R = [n] \setminus E$, and for every $v_i \in L$ and $u \in R$, $v_i u$ is an edge in G iff $(E \setminus v_i) \cup \{u\} \in \mathcal{H}$. Claim 6.4.22 implies that there are at most $t - 2$ pairwise disjoint edges in G. Therefore, by the König-Hall theorem, G contains a vertex cover S with $|S| \le t-2$. Let $\ell = |L \setminus S| \ge 2$. Then $|S \cap R| \le \ell - 2$. For every $v \in L \setminus S$ since $N_G(v) \subset S \cap R$, we obtain

$$\deg_{\mathcal{H}}(E \setminus \{v\}) = \deg_G(v) + 1 \le \ell - 1,$$

which implies that

$$\omega_{\mathcal{H}}(E) = \sum_{E' \subset E,\ |E'|=k-1} \frac{1}{\deg_{\mathcal{H}}(E')} > \sum_{v \in L \setminus S} \frac{1}{\deg_{\mathcal{H}}(E \setminus \{v\})} \ge \frac{\ell}{\ell - 1} \ge \frac{k}{k-1}.$$

This completes the proof of Lemma 6.4.21. ∎

6.4.4 Proofs

In this section we prove Theorems 6.4.6 and 6.4.8. First, let us prove Theorem 6.4.8.

Proof of Theorem 6.4.8. Let $k > p \ge 2$, $d \ge p$, and $\vec{a} = (a_1, \dots, a_p)$ be a sequence of integers such that $a_1 \ge \cdots \ge a_p \ge 1$ and $\sum_{i \in [p]} a_i = k$. Let $\epsilon > 0$ and n be sufficiently large. Let $\mathcal{H} \subset \binom{[n]}{k}$ be a hypergraph that contains no $(\vec{a}, d)$-Δ-systems and $|\mathcal{H}| \ge (1-\epsilon)\binom{n-1}{k-1}$.

Let $s = kd + 1$ and let $\mathcal{H}_1$ be a maximum s-homogeneous subgraph of $\mathcal{H}$. Suppose now we have defined $\mathcal{H}_1, \dots, \mathcal{H}_i$ for some $i \ge 1$. Let $\mathcal{H}_{i+1}$ be the maximum s-homogeneous subgraph of $\mathcal{H} \setminus \left(\bigcup_{j=1}^{i} \mathcal{H}_j\right)$. This process terminates if $\mathcal{H} \setminus \left(\bigcup_{j=1}^{m} \mathcal{H}_j\right) = \emptyset$ or the intersection pattern of

$\mathcal{H}_{m+1}$ has rank at most $k-2$ for some $m \geq 1$. Let $\mathcal{J}_i$ denote the intersection pattern of $\mathcal{H}_i$ for $i \in [m]$, and note that by definition and Lemma 6.4.17, $r(\mathcal{J}_i) = k-1$ for $i \in [m]$. Let

$$\widehat{\mathcal{H}}_1 = \bigcup_i \{\mathcal{H}_i \colon i \in [m] \text{ and } \mathcal{J}_i \text{ contains exactly } k-1 \ (k-1)\text{-sets}\},$$

$$\widehat{\mathcal{H}}_2 = \bigcup_i \{\mathcal{H}_i \colon i \in [m] \text{ and } \mathcal{J}_i \text{ contains at most } k-2 \ (k-1)\text{-sets}\},$$

$$\widehat{\mathcal{H}}_3 = \mathcal{H} \setminus \left(\widehat{\mathcal{H}}_1 \cup \widehat{\mathcal{H}}_2\right) = \mathcal{H} \setminus \left(\bigcup_{i \in [m]} \mathcal{H}_i\right).$$

Our first step is to show that the sizes of $\widehat{\mathcal{H}}_2$ and $\widehat{\mathcal{H}}_3$ are small.

Claim 6.4.23. $|\widehat{\mathcal{H}}_2| + |\widehat{\mathcal{H}}_3| < 3\epsilon k\binom{n-1}{k-1}$.

Proof. First we show that $\widehat{\mathcal{H}}_3 = O(n^{k-2})$. We may assume that $\widehat{\mathcal{H}}_3 \neq \emptyset$. Recall that $\mathcal{H}_{m+1}$ is a maximum s-homogeneous subgraph of $\widehat{\mathcal{H}}_3$ with intersection pattern $\mathcal{J}_{m+1}$. By Theorem 6.4.15, there exists a constant $c(k,s) > 0$ such that $|\mathcal{H}_{m+1}| \geq c(k,s)|\widehat{\mathcal{H}}_3|$. By definition, $r(\mathcal{J}_{m+1}) \leq k-2$, so by Lemma 6.4.16, $|\mathcal{H}_3| \leq |\partial_2 \mathcal{H}_3| \leq \binom{n}{k-2}$. Therefore, $|\widehat{\mathcal{H}}_3| \leq \frac{1}{c(k,s)}\binom{n}{k-2}$.

Next we show that $\widehat{\mathcal{H}}_2 = O(n^{k-2})$. By Lemma 6.4.19 and Equation 6.8,

$$|\partial \mathcal{H}| = \sum_{E \in \mathcal{H}} \omega_{\mathcal{H}}(E) = \sum_{E \in \widehat{\mathcal{H}}_1} \omega_{\mathcal{H}}(E) + \sum_{E \in \widehat{\mathcal{H}}_2} \omega_{\mathcal{H}}(E) \geq |\widehat{\mathcal{H}}_1| + \frac{k}{k-1}|\widehat{\mathcal{H}}_2|.$$

Therefore, $|\widehat{\mathcal{H}}_1| + \frac{k}{k-1}|\widehat{\mathcal{H}}_2| \le \binom{n}{k-1}$, which implies that

$$\begin{aligned}|\widehat{\mathcal{H}}_2| &= (k-1)\left(|\widehat{\mathcal{H}}_1| + \frac{k}{k-1}|\widehat{\mathcal{H}}_2| + |\widehat{\mathcal{H}}_3| - |\mathcal{H}|\right)\\ &\le (k-1)\left(\binom{n}{k-1} + |\widehat{\mathcal{H}}_3| - (1-\epsilon)\binom{n-1}{k-1}\right) < 2\epsilon k\binom{n-1}{k-1}.\end{aligned}$$

This completes the proof of Claim 6.4.23. ∎

Note that the proof of Claim 6.4.23 also shows that

$$|\mathcal{H}| \le |\widehat{\mathcal{H}}_1| + \frac{k}{k-1}|\widehat{\mathcal{H}}_2| + |\widehat{\mathcal{H}}_3| \le \binom{n}{k-1} + O(n^{k-2}). \tag{6.9}$$

Claim 6.4.23 implies that

$$|\widehat{\mathcal{H}}_1| = |\mathcal{H}| - \left(|\widehat{\mathcal{H}}_2| + |\widehat{\mathcal{H}}_3|\right) > (1-4\epsilon k)\binom{n-1}{k-1}. \tag{6.10}$$

By definition, for every $E \in \widehat{\mathcal{H}}_1$ there exists a unique s-homogeneous hypergraph $\mathcal{H}_i$ for some i such that $E \in \mathcal{H}_i$, moreover, $r(\mathcal{J}_i) = k-1$ and $\mathcal{J}_i$ contains exactly $k-1$ $(k-1)$-sets. Therefore, $\mathcal{I}(E, \mathcal{H}_i)$ contains a unique vertex $c \in E$ such that every $(k-1)$-subset of E that contains c is contained in $\mathcal{I}(E, \mathcal{H}_i)$. Let $c(E)$ denote this unique vertex c for every $E \in \widehat{\mathcal{H}}_1$. Define $\mathcal{G}_i = \left\{E \in \widehat{\mathcal{H}}_1 \colon c(E) = i\right\}$ for $i \in [n]$, and notice that $\bigcup_{i\in[n]} \mathcal{G}_i = \widehat{\mathcal{H}}_1$ is a partition. Let $\mathcal{G}_i(i) = \{E \setminus \{i\} \colon E \in \mathcal{G}_i\}$ for $i \in [n]$. From the proof of Lemma 6.4.21 (1), for every $i \in [n]$ and every $E \in \mathcal{G}_i$ the set $E \setminus \{i\}$ is not contained in any set in $\mathcal{H} \setminus \{E\}$. Therefore, $\mathcal{G}_i(i) \cap \mathcal{G}_j(j) = \emptyset$ for all $\{i, j\} \subset [n]$.

Claim 6.4.24. *$\partial\mathcal{G}_i(i) \cap \partial\mathcal{G}_j(j) = \emptyset$ for all $\{i,j\} \subset [n]$.*

Proof. Suppose not. Without loss of generality we may assume that there exists $A \in \partial\mathcal{G}_1(1) \cap \partial\mathcal{G}_2(2)$. Then there exists $E_1 \in \mathcal{G}_1$ and $E_2 \in \mathcal{G}_2$ such that $E_1 = \{1,u\} \cup A$ and $E_2 = \{2,v\} \cup A$ for some $u, v \in [n]$. Since $\mathcal{G}_1$ is s-homogeneous and $|E_2 \cap E_1| \geq k-2 \geq k-a_1$, by Lemma 6.4.20, $1 \in E_2$. Similarly, we obtain $2 \in E_1$. Therefore, $E_1 = E_2 = \{1,2\} \cup A$, which implies that $\{1,2\} \cup A \in \mathcal{G}_1 \cap \mathcal{G}_2$, a contradiction. ∎

Let $x_i \in \mathbb{R}$ such that $|\mathcal{G}_i| = |\mathcal{G}_i(i)| = \binom{x_i}{k-1}$ for $i \in [n]$. Without loss of generality we may assume that $x_1 \geq \cdots \geq x_n \geq 0$. By the Kruskal-Katona theorem (e.g. see [175]),

$$|\mathcal{G}_i(i)| \leq \frac{\binom{x_i}{k-1}}{\binom{x_i}{k-2}} |\partial\mathcal{G}_i(i)| = \frac{x_i - k + 2}{k-1} |\partial\mathcal{G}_i(i)|,$$

for $i \in [n]$. Therefore by Equation 6.10 and Claim 6.4.24,

$$\begin{aligned}(1-4\epsilon k)\binom{n-1}{k-1} < |\widehat{\mathcal{H}}_1| = \sum_{i\in[n]} |\mathcal{G}_i| = \sum_{i\in[n]} |\mathcal{G}_i(i)| &\leq \sum_{i\in\mathcal{H}} \frac{x_i - k + 2}{k-1} |\partial\mathcal{G}_i(i)| \\ &\leq \frac{x_1 - k + 2}{k-1} \sum_{i\in\mathcal{H}} |\partial\mathcal{G}_i(i)| \leq \frac{x_1 - k + 2}{k-1} \binom{n}{k-2},\end{aligned}$$

which implies that

$$x_1 \geq (k-1)\frac{(1-4\epsilon k)\binom{n-1}{k-1}}{\binom{n}{k-2}} + k - 2 > (1-5\epsilon k)n.$$

Therefore,

$$|\mathcal{G}_1| = \binom{x_1}{k-1} > \binom{(1-5\epsilon k)n}{k-1} > (1-5\epsilon k^2)\binom{n-1}{k-1},$$

which together with Equation 6.9 implies that all but at most $5\epsilon k^2 n^{k-1}$ edges in $\mathcal{H}$ contain the vertex 1. ∎

Now we prove Theorem 6.4.6.

Proof of Theorem 6.4.6. Let $d \geq p \geq 2$, $k > p$, $s = kd + 1$, and $\vec{a} = (a_1, \ldots, a_p)$ be a sequence of positive integers such that $a_1 \geq \cdots \geq a_p$ and $\sum_{i \in [p]} a_i = k$. Let $n \geq n_0(k, d)$ be sufficiently large. Let $\mathcal{H} \subset \binom{[n]}{k}$ be a hypergraph that contains no $(\vec{a}, d)$-Δ-systems and $|\mathcal{H}| = \binom{n-1}{k-1}$. It suffices to show that all edges in $\mathcal{H}$ contain a fixed vertex.

From the proof of Theorem 6.4.8 we know that $\mathcal{H}$ contains a subgraph $\mathcal{G}_1$ such that all edges in $\mathcal{G}_1$ contains a fixed vertex (we may assume that this vertex is 1), moreover, $\mathcal{G}_1$ consists of pairwise edge-disjoint s-homogeneous hypergraphs whose intersection patterns have rank $k-1$ and contain all $(k-1)$-subsets of $[k]$ that contain 1.

Define

$$\mathcal{B}_0 = \{E \in \mathcal{H} \colon 1 \notin E\},$$

$$\mathcal{B}_1 = \{E \in \mathcal{H} \colon 1 \in E \text{ and } |E \cap B| \geq k - a_1 \text{ for some } B \in \mathcal{B}_0\},$$

$$\mathcal{G} = \{E \in \mathcal{H} \setminus \mathcal{B}_1 \colon 1 \in E,\ \forall S \subset E \text{ with } 1 \in S \text{ is the center of a } \Delta\text{-system in } \mathcal{H} \text{ of size } s\},$$

$$\mathcal{B}_2 = \{E \in \mathcal{H} \colon 1 \in E\} \setminus (\mathcal{B}_1 \cup \mathcal{G}).$$

Note that $\mathcal{G}_1 \subset \mathcal{G}$. Let

$$\mathcal{B}_1(1) = \{E \setminus 1 \colon E \in \mathcal{B}_1\}, \quad \mathcal{G}(1) = \{E \setminus 1 \colon E \in \mathcal{G}\}, \quad \text{and} \quad \mathcal{B}_2(1) = \{E \setminus 1 \colon E \in \mathcal{B}_2\}.$$

Let $\mathcal{B}_1^*(1), \mathcal{B}_2^*(1)$ be maximum s-homogeneous subgraphs of $\mathcal{B}_1(1), \mathcal{B}_2(1)$, respectively. Then by Theorem 6.4.15, $|\mathcal{B}_i^*(1)| \geq c(k,s)|\mathcal{B}_i(1)|$ for some constant $c(k,s) > 0$ and $i = 1, 2$. Recall that for every $E \in \partial\mathcal{G}(1)$, $\deg_{\mathcal{G}(1)}(E)$ is the number of edges in $\mathcal{G}(1)$ that contain E. Since $\sum_{E \in \partial\mathcal{G}(1)} \deg_{\mathcal{G}(1)}(E) = (k-1)|\mathcal{G}(1)|$ and $\deg_{\mathcal{G}(1)}(E) \leq n - k + 1$, we have

$$|\partial\mathcal{G}(1)| \geq \frac{k-1}{n-k+1}|\mathcal{G}(1)|. \tag{6.11}$$

Claim 6.4.25. $|\mathcal{G}| + 4|\mathcal{B}_0| \leq \binom{n-1}{k-1}$.

Proof. Notice that by definition $|E \cap B| \leq k - a_1 - 1 \leq k - 3$ for all $E \in \mathcal{G}(1)$ and $B \in \mathcal{B}_0$. Therefore, $\partial\mathcal{G}(1) \cap \partial_2\mathcal{B}_0 = \emptyset$, and hence $|\partial\mathcal{G}(1)| + |\partial_2\mathcal{B}_0| \leq \binom{n-1}{k-2}$. Let $x \in \mathbb{R}$ such that $|\partial\mathcal{B}_0| = \binom{x}{k-1}$, then by the Kruskal-Katona theorem and Proposition 6.4.18,

$$|\partial_2\mathcal{B}_0| \geq \frac{k-1}{x-k+1}|\partial\mathcal{B}_0| \geq \frac{k-1}{x-k+1}c(k,s)|\mathcal{B}_0|.$$

Therefore, together with Equation 6.11 we obtain

$$\frac{k-1}{n-k+1}|\mathcal{G}(1)| + \frac{k-1}{x-k+1}c(k,s)|\mathcal{B}_0| \leq \binom{n-1}{k-2},$$

which implies $|\mathcal{G}|+c(k,s)\frac{n-k+1}{x-k+1}|\mathcal{B}_0| \leq \binom{n-1}{k-1}$. By Theorem 6.4.8, $\binom{x}{k-1} = |\partial\mathcal{B}_0| \leq k|\mathcal{B}_0| \leq \delta n^{k-1}$ for all sufficiently small $\delta > 0$ (as long as n is sufficiently large), so $x < \delta' n$ for some sufficiently small $\delta' > 0$ (depending on δ). Choosing $\delta' \ll c(k,s)$ we obtain $c(k,s)\frac{n-k+1}{\delta' n-k+1} > 4$, this completes the proof of Claim 6.4.25. ∎

Claim 6.4.26. *Every $E \in \mathcal{B}_1^*(1)$ has a $(k-2)$-subset that is not contain in any other set in $\mathcal{B}_1^*(1) \cup \mathcal{G}'$.*

Proof. Suppose not. Let $E = \{v_1, \ldots, v_{k-1}\} \in \mathcal{B}_1^*(1)$ such that $E \setminus \{v_i\}$ is contained in some set in $\mathcal{B}_1^*(1) \cup \mathcal{G}(1)$ for $1 \leq i \leq k-1$. Without loss of generality we may assume that $E \setminus \{v_i\} \in \mathcal{I}(E, \mathcal{G}(1))$ for $1 \leq i \leq \ell$, and $E \setminus \{v_i\} \in \mathcal{I}(E, \mathcal{B}_1^*(1))$ for $\ell+1 \leq i \leq k-1$.

Let $\mathcal{J}_{\mathcal{B}_1^*(1)}$ be the intersection pattern of $\mathcal{B}_1^*(1)$. Let $\mathcal{B}_1^* = \{E \cup \{1\} \colon E \in \mathcal{B}_1^*(1)\}$, and note that $\mathcal{B}_1^*$ is also s-homogeneous with intersection pattern $\mathcal{J}_{\mathcal{B}_1^*} := \{A \cup \{1\} : A \in \mathcal{J}_{\mathcal{B}_1^*(1)}\}$. Let $\widehat{E} = E \cup \{1\} \in \mathcal{B}_1^*$.

If $\ell = 0$, then $\mathcal{J}_{\mathcal{B}_1^*(1)} = 2^{[k-1]} \setminus \{[k-1]\}$, and hence $r(\Pi(\mathcal{I}(\widehat{E}, \mathcal{B}_1^*))) = k-1$ and $\Pi(\mathcal{I}(\widehat{E}, \mathcal{B}_1^*))$ contains all $(k-1)$-subsets of $\widehat{E}$ that contain 1. By definition there exists $B \in \mathcal{B}_0$ such that $|B \cap \widehat{E}| \geq k - a_1$. However, by Lemma 6.4.20, $1 \in B$, a contradiction. Therefore, $\ell \geq 1$.

Let $E_i \in \mathcal{G}$ such that $E_i \cap \widehat{E} = \widehat{E} \setminus \{v_i\}$ for $1 \leq i \leq \ell$. Let $B \in \mathcal{B}_0$ such that $|B \cap \widehat{E}| \geq k - a_1$ and suppose that $|B \cap \widehat{E}| = k - t$ for some $1 \leq t \leq a_1$. Then for $1 \leq i \leq \ell$ we have $|B \cap E_i| \geq k - t - 1$. However, by the definition of $\mathcal{G}$, $|B \cap E_i| \leq k - a_1 - 1$ for $1 \leq i \leq \ell$. Therefore, $|B \cap \widehat{E}| = k - a_1$ and $v_i \in B$ for all $1 \leq i \leq \ell$. Let $\bigcup_{i \in [p]} A_i = \widehat{E}$ be an $\vec{a}$-partition such that $A_1 = \widehat{E} \setminus B$. Note that for $2 \leq i \leq p$, either $1 \in \widehat{E} \setminus A_i \subset E_{j_i}$ for some $1 \leq j_i \leq \ell$, which by the definition of $\mathcal{G}$, is the center of some Δ-system of size s in $\mathcal{H}$, or $\{1, v_1, \ldots, v_\ell\} \subset \widehat{E} \setminus A_i$, which implies that $\widehat{E} \setminus A_i \in \mathcal{I}(\widehat{E}, \mathcal{B}_1^*)$ and hence is the center of some Δ-system of size s in $\mathcal{B}_1^*$. Note that $E \setminus A_1$ is the center of a Δ-system of size 2, i.e. $\{\widehat{E}, B\}$. Therefore, by Lemma 6.4.13, $\mathcal{H}$ contains an $(\vec{a}, d)$-Δ-system, a contradiction. ∎

By Claim 6.4.26, we obtain $|\partial \mathcal{G}(1)| + |\mathcal{B}_1^*(1)| \leq \binom{n-1}{k-2}$, which implies $|\mathcal{G}| + c(k,s)\frac{n-k+1}{k-1}|\mathcal{B}_1| \leq \binom{n-1}{k-1}$. Note that $c(k,s)\frac{n-k+1}{k-1} \gg 1$, so

$$|\mathcal{G}| + 4|\mathcal{B}_1| \leq \binom{n-1}{k-1}.$$

Claim 6.4.27. *Every $E \in \mathcal{B}_2^*(1)$ has a $(k-2)$-subset that is not contain in any other set in $\mathcal{B}_2^*(1) \cup \mathcal{G}'$.*

Proof. Suppose not. Let $E = \{v_1, \ldots, v_{k-1}\} \in \mathcal{B}_2^*(1)$ such that $E \setminus \{v_i\}$ is contained in some set in $\mathcal{B}_2^*(1) \cup \mathcal{G}(1)$ for $1 \leq i \leq k-1$. Without loss of generality we may assume that $E \setminus \{v_i\} \in \mathcal{I}(E, \mathcal{G}(1))$ for $1 \leq i \leq \ell$, and $E \setminus \{v_i\} \in \mathcal{I}(E, \mathcal{B}_2^*(1))$ for $\ell + 1 \leq i \leq k-1$.

Let $\mathcal{J}_{\mathcal{B}_2^*(1)}$ be the intersection pattern of $\mathcal{B}_2^*(1)$. Let $\mathcal{B}_2^* = \{E \cup \{1\} \colon E \in \mathcal{B}_2^*(1)\}$, and note that $\mathcal{B}_2^*$ is also s-homogeneous with intersection pattern $\mathcal{J}_{\mathcal{B}_2^*} := \{A \cup \{1\} \colon A \in \mathcal{J}_{\mathcal{B}_2^*(1)}\}$. Let $\widehat{E} = E \cup \{1\} \in \mathcal{B}_2^*$.

If $\ell = 0$, then $\mathcal{J}_{\mathcal{B}_2^*(1)} = 2^{[k-1]} \setminus \{[k-1]\}$, and hence $r(\Pi(\mathcal{I}(\widehat{E}, \mathcal{B}_2^*))) = k-1$ and $\Pi(\mathcal{I}(\widehat{E}, \mathcal{B}_2^*))$ contains all $(k-1)$-subsets of $\widehat{E}$ that contain 1. Since $\mathcal{I}(\widehat{E}, \mathcal{B}_2^*)$ is closed under intersection, all proper subsets of $\widehat{E}$ that contain 1 is contained in $\mathcal{I}(\widehat{E}, \mathcal{B}_2^*)$, which by definition, implies that $\widehat{E} \in \mathcal{G}$, a contradiction. Therefore, $\ell \geq 1$.

Let $E_i \in \mathcal{G}$ such that $E_i \cap \widehat{E} = \widehat{E} \setminus \{v_i\}$ for $1 \leq i \leq \ell$. For every proper subset $S \subset \widehat{E}$ with $1 \in S$, if $v_i \notin S$ for some $1 \leq i \leq \ell$, then $S \subset E_i$, which, by the definition of $\mathcal{G}$, means that S is the center of some Δ-system of size s in $\mathcal{H}$. If $\{v_1, \ldots, v_\ell\} \subset S$, then $S \in \mathcal{I}(\widehat{E}, \mathcal{B}_2^*)$ and hence S is the center of some Δ-system of size s in $\mathcal{B}_2^*$. Therefore, every proper subset $S \subset \widehat{E}$ with $1 \in S$ is the center of some Δ-system of size s in $\mathcal{H}$, which by definition, implies that $\widehat{E} \in \mathcal{G}$, a contradiction. ∎

Similarly, we obtain

$$|\mathcal{G}| + 4|\mathcal{B}_2| \leq \binom{n-1}{k-1}.$$

Therefore, by the assumption that $|\mathcal{H}| = \binom{n-1}{k-1}$ we obtain

$$\begin{aligned} 3\binom{n-1}{k-1} &\leq 3|\mathcal{H}| + |\mathcal{B}_0| + |\mathcal{B}_1| + |\mathcal{B}_2| \\ &= |\mathcal{G}| + 4|\mathcal{B}_0| + |\mathcal{G}| + 4|\mathcal{B}_1| + |\mathcal{G}| + 4|\mathcal{B}_2| \leq 3\binom{n-1}{k-1}, \end{aligned}$$

which implies that $|\mathcal{G}| = \binom{n-1}{k-1}$ and $\mathcal{B}_0 = \mathcal{B}_1 = \mathcal{B}_2 = \emptyset$. This completes the proof of Theorem 6.4.6. ∎

CHAPTER 7

EXTENSION OF THE TURÁN THEOREM

Previously published as X. Liu and D. Mubayi. On a generalized Erdős–Rademacher problem. J. Graph Theory, 2020; and X. Liu and J. Ma. Sparse halves in K_4 -free graphs. J. Graph Theory, 99(1):525, 2022.

7.1 Sparse halves in K_4-free graphs

7.1.1 Introduction

Generalizing Turán's theorem, Erdős [61] initialized the study of the following problem: Given a constant $0 \leq \alpha \leq 1$, what is the minimum value $\beta = \beta(\alpha, r)$ such that every n-vertex K_r-free graph contains a vertex set of size $\lfloor \alpha n \rfloor$ which spans at most βn^2 edges? This is often referred as the local density problem.

The case $\alpha = 1/2$ is of special interest. Erdős [63] offered $250 for the first solution to the following long-standing conjecture on triangle-free graphs.

Conjecture 7.1.1 (Erdős [61])**.** *Every triangle-free graph on n vertices contains a vertex set of size* $\lfloor n/2 \rfloor$ *that spans at most* $n^2/50$ *edges.*

Both of the balanced blow-ups of the 5-cycle and the Petersen graph show that the bound $n^2/50$ would be best possible if this conjecture is true. Despite extensive research [153; 143; 203; 15], Conjecture 7.1.1 is still open.

A similar question also has been asked for K_4-free graphs. Chung and Graham [40], and Erdős, Faudree, Rousseau and Schelp [65] posted the following conjecture.

Conjecture 7.1.2 (Chung et al. [40], Erdős et al. [65])**.** *Every* K_4*-free graph on n vertices contains a vertex set of size* $\lfloor n/2 \rfloor$ *that spans at most* $n^2/18$ *edges.*

The Turán graph $T_3(n)$ shows that the bound $n^2/18$ in Conjecture 7.1.2 would be best possible if it is true. A closely related conjecture of Erdős (see [64]), which was proved by Sudakov [237], states that every K_4-free graphs on n vertices can be made bipartite by deleting at most $n^2/9$ edges. An interesting interplay between these problems for regular graphs was observed by Krivelevich [153], where he pointed out that a bound in the local density problem can imply a bound (doubled) in the problem of making a graph bipartite; also see [237] for an illustration.

The main result of this section is to confirm Conjecture 7.1.2 for all regular graphs. We prove it in the following form, which also characterizes the unique extremal graph.

Theorem 7.1.3. *Let G be a K_4-free regular graph on n vertices. If every vertex set of size $\lfloor n/2 \rfloor$ in G spans at least $n^2/18$ edges, then n is divisible by 6 and $G \cong T_3(n)$.*

We would like to remark that our proof of Theorem 7.1.3 actually shows that Conjecture 7.1.2 holds for all almost regular graphs, i.e. graphs whose difference of maximum degree and minimum degree is bounded by ϵn for some absolute constant $\epsilon > 0$.[1] [2]

As a corollary, Theorem 7.1.3 implies the following slightly stronger version of Sudakov's theorem in the case of regular graphs.

Corollary 7.1.4. *Let $n \in \mathbb{N}$ be even. Then every regular K_4-free graph on n vertices can be made bipartite by removing at most $n^2/9$ edges such that each part has size exactly $n/2$.*

For odd $n \in \mathbb{N}$, one could easily obtain a similar result as in Corollary 7.1.4.

We now introduce a crucial tool in our proof of Theorem 7.1.3, which also can be viewed as a strengthening of the local density problem. Erdős, Faudree, Rousseau and Schelp conjectured

---

[1] Our calculations indicate that ϵ can be chosen as $\epsilon = 1/500$.

[2] Conjecture 7.1.2 has been completely resolved by Reiher [218] recently.

in [65] that for every $\alpha \in [17/30, 1]$, every triangle-free graph on n vertices contains a vertex set of size $\lfloor \alpha n \rfloor$ that spans at most $(2\alpha - 1)n^2/4$ edges. This was confirmed by Krivelevich [153] for all $\alpha \in [3/5, 1]$. The coming result shows that the bound $(2\alpha - 1)n^2/4$ can be improved in the range where α is relatively large.

Theorem 7.1.5. *Let $\alpha, c \in [0, 1]$ satisfy $\alpha + c \geq 1$. Then the following hold:*

(1). Every n-vertex triangle-free graph with cn^2 edges contains a vertex set of size $\lfloor \alpha n \rfloor$ that spans at most $(2\alpha - 1)cn^2$ edges.

(2). Assume that $\alpha n \in \mathbb{N}$ and G is an n-vertex triangle-free graph. If every vertex set of size αn in G spans at least $(2\alpha - 1)cn^2$ edges, then G is regular, and vice versa.

Note that by Mantel's theorem [186], we have $(2\alpha - 1)cn^2 \leq (2\alpha - 1)n^2/4$.

7.1.2 Preliminaries

Recall that for two disjoint vertex sets $S, T \subset V(G)$, we let $G[S, T]$ be the induced bipartite subgraph of G with two parts S, T and let $e_G(S, T)$ be the number of edges in $G[S, T]$. If it is clear from the context we omit the subscript G. We also omit floors and ceilings when they are not essential in our proofs.

The following propositions can be found in the literature (e.g. [143]). For completeness we include their proofs here.

Proposition 7.1.6. *Let $0 \leq \alpha \leq 1$. Then every n-vertex graph G with e edges contains a vertex set of size αn that spans at most $\alpha^2 e$ edges.*

Proof. Choose $S \subset V(G)$ with $|S| = \alpha n$ uniformly at random. Then for every edge e, the probability that e is contained in S is $\frac{\alpha n}{n} \cdot \frac{\alpha n - 1}{n-1} \leq \alpha^2$. So, the expected value of $e(S)$ is at most $\alpha^2 e$. Hence there exists a vertex set of size αn in G that spans at most $\alpha^2 e$ edges. ∎

Proposition 7.1.7. *Let G be an n-vertex graph with e edges. Let $A \cup B = V(G)$ be a partition with $|A| = \alpha n \leq n/2$. Then there exists $S \subset B$ with $|S| = (1/2 - \alpha)\, n$ such that*

$$
\begin{aligned}
e(A \cup S) &\leq e(A) + \frac{1/2 - \alpha}{1 - \alpha} e(A, B) + \left(\frac{1/2 - \alpha}{1 - \alpha} \right)^2 e(B) \\
&= e(G) - \frac{1}{2(1 - \alpha)} e(A, B) - \frac{3/2 - 2\alpha}{2(1 - \alpha)^2} e(B).
\end{aligned}
$$

Proof. Choose $S \subset B$ with $|S| = (1/2 - \alpha)\, n$ uniformly at random. Then, for every $e \in E(G[A, B])$ the probability that e is contained in $A \cup S$ is $\frac{1/2-\alpha}{1-\alpha}$. Similar to the proof of Proposition 7.1.6, for every $e' \in E(G[B])$ the probability that e' is contained in S is at most $\left(\frac{1/2-\alpha}{1-\alpha}\right)^2$. So, the expected value of $e(A \cup S)$ is at most $e(A) + \frac{1/2-\alpha}{1-\alpha} e(A, B) + \left(\frac{1/2-\alpha}{1-\alpha}\right)^2 e(B)$. Therefore, there exists $S \subset B$ with $|S| = (1/2 - \alpha)\, n$ such that the desired inequality holds. ∎

7.1.3 Local densities in triangle-free graphs

In this section we prove Theorem 7.1.5. First we show the following proposition for the "vice versa" part of Theorem 7.1.5 (2).

Proposition 7.1.8. *Let $\alpha, c \in [0, 1]$, $n \in \mathbb{N}$ such that $\alpha n \in \mathbb{N}$. Suppose that G is a triangle-free regular graph on n vertices with cn^2 edges. Then every $S \subseteq V(G)$ with $|S| = \alpha n$ spans at least $(2\alpha - 1)cn^2$ edges.*

Proof. Let $S \subset V(G)$ be a set with size αn let $T = V(G) \setminus S$. Since G is regular, every vertex has degree $2cn$, which shows that

$$2e(S) + e(S,T) = \sum_{v \in S} d(v) = 2\alpha cn^2 \quad \text{and} \quad e(S,T) \leq \sum_{v \in T} d(v) = 2(1-\alpha)cn^2.$$

Therefore,

$$e(S) = \frac{1}{2}\left(2e(S) + e(S,T) - e(S,T)\right) \geq \frac{1}{2}(2\alpha cn^2 - 2(1-\alpha)cn^2) = (2\alpha - 1)cn^2,$$

which completes the proof of Proposition 7.1.8. ∎

Now we prove Theorem 7.1.5. The core of the proof is a probabilistic argument. For convenience we will assume $\alpha n \in \mathbb{N}$ in the coming presentation, while the proof for the case $\alpha n \notin \mathbb{N}$ holds analogously.

Proof of Theorem 7.1.5. Let $\alpha + c \geq 1$ and G be an n-vertex triangle-free graph with cn^2 edges. Our goal is to find a subset $S \subseteq V(G)$ with $|S| = \alpha n$ that spans at most $(2\alpha - 1)cn^2$ edges. It is clear that we may assume $\alpha < 1$. We divide the proof into two cases by considering the value of $\delta(G)$.

First suppose that $\delta(G) \geq (1-\alpha)n$.[1] Suppose for the contrary that every subset of size αn spans more than $(2\alpha - 1)cn^2$ edges. For every $v \in V(G)$, let $B_v = N(v)$ and $A_v = V(G) \setminus B_v$. Since G is triangle-free, B_v is an independent set and hence $e(A_v) + e(A_v, B_v) = cn^2$. Let

---

[1] We point out that this case holds even without requiring $\alpha + c \geq 1$.

$d_v = d(v)/n$. By a similar argument as in Proposition 7.1.7, there exists $S \subseteq B_v$ with $|S| = (\alpha + d_v - 1)n$ such that

$$e(A_v \cup S) \leq e(A_v) + \frac{\alpha + d_v - 1}{d_v} e(A_v, B_v).$$

Since $|A_v \cup S| = \alpha n$, by assumption, we have

$$e(A_v) + \frac{\alpha + d_v - 1}{d_v} e(A_v, B_v) \geq e(A_v \cup S) > (2\alpha - 1)cn^2,$$

which together with $e(A_v) + e(A_v, B_v) = cn^2$ gives

$$cn^2 - \frac{1-\alpha}{d_v} e(A_v, B_v) > (2\alpha - 1)cn^2.$$

Therefore,

$$\sum_{v \in V(G)} \left(cn^2 - \frac{1-\alpha}{d_v} e(A_v, B_v) \right) d_v > \sum_{v \in V(G)} (2\alpha - 1)cn^2 d_v,$$

which implies

$$(1-\alpha) \sum_{v \in V(G)} e(A_v, B_v) < 2(1-\alpha)cn^2 \sum_{v \in V(G)} d_v.$$

Since $\sum_{v \in V(G)} d_v = 2cn$ and $\alpha < 1$, this gives

$$\sum_{v \in V(G)} e(A_v, B_v) < 4c^2 n^3.$$

On the other hand, since B_v is independent for each v, by the Cauchy–Schwarz inequality

$$\sum_{v\in V(G)} e(A_v,B_v) = \sum_{v\in V(G)}\sum_{u\in N(v)} d(u) = \sum_{u\in V(G)} (d(u))^2 \geq \frac{1}{n}\left(\sum_{u\in V(G)} d(u)\right)^2 = 4c^2n^3,$$

which is a contradiction. Therefore, if $\delta(G) \geq (1-\alpha)n$, then there exists a vertex set of size αn that spans at most $(2\alpha-1)cn^2$ edges. Note that if every vertex set of size αn spans at least $(2\alpha-1)cn^2$ edges, then by the above arguments, we see that $d(v)$ must be the same for all $v \in V(G)$, that is, G is regular.

Now suppose that $\delta(G) < (1-\alpha)n$, where $\alpha + c \geq 1$. Choose $v \in V(G)$ such that $d(v) = \delta(G) < (1-\alpha)n$ and remove v from G. We iteratively remove a vertex with the minimum degree in the remaining graph until there is no vertex left or the remaining graph G' satisfies $\delta(G') \geq (1-\alpha)n$. Let A denote the set of vertices we removed in this process and let $k = |A|/n$. If $|A| = n$, then $e(G) < (1-\alpha)n^2 \leq cn^2$, a contradiction. So $|A| < n$, which implies that $G' \neq \emptyset$. Since $\delta(G') \geq (1-\alpha)n$, we have $|V(G')| > (1-\alpha)n$. Therefore, $k = |A|/n = (n-|V(G')|)/n < \alpha$. Let $B = V(G)\setminus A$ and let $G' = G[B]$. Also let $\tilde{n} = (1-k)n$ and $\tilde{\alpha} = \frac{\alpha-k}{1-k}$. Since $\delta(G') \geq (1-\alpha)n = (1-\tilde{\alpha})\tilde{n}$, by the previous case, there exists $S \subseteq B$ with

$|S| = \tilde{\alpha}\tilde{n}$ such that $e(S) \le (2\tilde{\alpha}-1)e(B)$. Now we obtain a desired subset $A \cup S$ in G with size $|A \cup S| = kn + \tilde{\alpha}\tilde{n} = \alpha n$ and

$$\begin{aligned} e(A \cup S) &= e(A) + e(A,S) + e(S) \le e(A) + e(A,B) + (2\tilde{\alpha}-1)e(B) \\ &= (2\tilde{\alpha}-1)\left(e(A) + e(A,B) + e(B)\right) + 2(1-\tilde{\alpha})\left(e(A) + e(A,B)\right) \\ &< (2\tilde{\alpha}-1)cn^2 + 2(1-\tilde{\alpha})k(1-\alpha)n^2 \le (2\alpha-1)cn^2, \end{aligned}$$

where the second last inequality is strict since $e(A) + e(A,B) < |A|(1-\alpha)n = k(1-\alpha)n^2$, and the last inequality follows from

$$(2\tilde{\alpha}-1)c + 2(1-\tilde{\alpha})(1-\alpha)k - (2\alpha-1)c = \frac{2k(1-\alpha)(\alpha+c-1)}{k-1} \le 0.$$

Therefore in case of $\delta(G) < (1-\alpha)n$, there always exists a subset of size αn spanning strictly less than $(2\alpha-1)cn^2$ edges. Together with Proposition 7.1.8, we have finished the proofs of Theorem 7.1.5 for both (1) and (2). ∎

7.1.4 Sparse halves

In this section we prove Theorem 7.1.3. Let G be a K_4-free graph on n vertices. For a vertex set $S \subset V(G)$ with $|S| = \lfloor n/2 \rfloor$, we call it a sparse half of G if $e(S) \le n^2/18$.

We will consider three cases regarding the edge density of G and use quite different techniques in each case. If G is sparse, then we will use some probabilistic arguments to show that it contains a sparse half. If G is dense, then a result of Lyle [182] gives a nice structure on G

and this enables us to find a sparse half. The most intricate case is when the edge density of G is intermediate. In this case, assuming G does not contain a sparse half, we will first find three large disjoint independent sets in G (by using Theorem 7.1.5), and then building on these sets, use probabilistic arguments (in a complicated way) to derive a contradiction. Finally, we infer Theorem 7.1.3 from these cases in Section 7.1.4.4.

In the rest of this section we will state our results without assuming the parities of integers n. However for convenience, in the proofs we will always view n as even in order to avoid the floors (while the same arguments also work for odd n). For Theorem 7.1.3, we will see in Section 4.4 that it suffices to only consider when n is divisible by 6.

7.1.4.1 Sparse range

In this section we will prove the following for graphs with few edges.

Theorem 7.1.9. *Suppose that G is a K_4-free graph on n vertices with at most $0.26n^2$ edges. Then G contains a vertex set of size $\lfloor n/2 \rfloor$ that spans strictly less than $n^2/18$ edges.*

We need the following two lemmas from [237] which are proved by probabilistic arguments. Let $t(G)$ denote the number of triangles in G.

Lemma 7.1.10 (Sudakov [237])**.** *Every graph G on n vertices contains a bipartite subgraph G' such that*

$$e(G') \geq \frac{1}{n} \sum_{v \in V(G)} (d(v))^2 - \frac{2}{n} \sum_{v \in V(G)} e\left(N(v)\right) \geq \frac{4\left(e(G)\right)^2}{n^2} - \frac{6t(G)}{n}.$$

Lemma 7.1.11 (Sudakov [237])**.** *Every K_4-free graph on n vertices contains a bipartite subgraph G' such that*

$$e(G') \geq \frac{e(G)}{2} + \frac{1}{n} \sum_{v \in V(G)} \left(\frac{4\,(e(N(v)))^2}{(d(v))^2} - \frac{e(N(v))}{2} \right).$$

The next lemma shows that if a K_4-free graph G contains a large enough bipartite subgraph, then it contains a sparse half.

Lemma 7.1.12. *Let G be a K_4-free graph on n vertices with cn^2 edges. Suppose that there is a partition $A \cup B = V(G)$ such that $e(A, B) > 9c^2n^2/4$. Then G contains a vertex set of size $\lfloor n/2 \rfloor$ that spans strictly less than $n^2/18$ edges.*

Proof. Suppose for the contrary that every vertex set of size $n/2$ in G spans at least $n^2/18$ edges. Assume that $a := |A|/n \leq 1/2$. Applying Proposition 7.1.6 to $G[B]$, we obtain a vertex set $S \subset B$ with $|S| = n/2$ such that $e(S) \leq \left(\frac{1/2}{1-a}\right)^2 e(B)$. By assumption we have $\left(\frac{1/2}{1-a}\right)^2 e(B) \geq n^2/18$, which implies

$$e(B) \geq \frac{2(1-a)^2}{9} n^2.$$

Now applying Proposition 7.1.7 to $A \cup B$, we see that there exists $T \subset V(G)$ with $|T| = n/2$ such that $A \subset T$ and $e(T) \leq cn^2 - \frac{1}{2(1-a)} e(A, B) - \frac{3/2-2a}{2(1-a)^2} e(B)$. By assumption, we have

$$\frac{n^2}{18} \leq cn^2 - \frac{1}{2(1-a)} e(A, B) - \frac{3/2 - 2a}{2(1-a)^2} e(B) \leq cn^2 - \frac{1}{2(1-a)} e(A, B) - \frac{3/2 - 2a}{9} n^2,$$

which implies

$$e(A,B) \le 2(1-a)\left(c - \frac{3/2-2a}{9} - \frac{1}{18}\right)n^2 = \left(2(1-a)c - \frac{4}{9}(1-a)^2\right)n^2 \le \frac{9}{4}c^2n^2,$$

a contradiction. Here the last inequality follows from $\left(\frac{2}{3}(1-a) - \frac{3}{2}c\right)^2 n^2 \ge 0$. ∎

Now we are ready to prove Theorem 7.1.9.

Proof of Theorem 7.1.9. Let $c = e(G)/n^2$ and let $\lambda = 8/13$. By Lemmas 7.1.10 and 7.1.11, there exists a partition $A \cup B = V(G)$ such that

$$\begin{aligned} e(A,B) &\ge (1-\lambda)\left(\frac{1}{n}\sum_{v\in V(G)} (d(v))^2 - \frac{2}{n}\sum_{v\in V(G)} e\left(N(v)\right)\right) \\ &\quad + \lambda\left(\frac{e(G)}{2} + \frac{1}{n}\sum_{v\in V(G)}\left(\frac{4\left(e(N(v))\right)^2}{(d(v))^2} - \frac{e(N(v))}{2}\right)\right) \\ &= \frac{\lambda}{2}e(G) + \frac{4\lambda}{n}\sum_{v\in V(G)} (d(v))^2\left(\left(\frac{e\left(N(v)\right)}{(d(v))^2}\right)^2 - \frac{2-3\lambda/2}{4\lambda}\frac{e\left(N(v)\right)}{(d(v))^2} + \frac{1-\lambda}{4\lambda}\right). \end{aligned}$$

Since

$$x^2 - \frac{2-3\lambda/2}{4\lambda}x + \frac{1-\lambda}{4\lambda} \ge \frac{88\lambda - 73\lambda^2 - 16}{256\lambda^2},$$

we obtain

$$\begin{aligned} e(A,B) &\geq \frac{\lambda}{2}e(G) + \frac{88\lambda - 73\lambda^2 - 16}{64\lambda} \sum_{v\in V(G)} \frac{d(v)^2}{n} \\ &\geq \frac{\lambda}{2}e(G) + \frac{88\lambda - 73\lambda^2 - 16}{64\lambda} \sum_{v\in V(G)} \left(\frac{\sum_{v\in V(G)} d(v)}{n}\right)^2 \\ &= \left(\frac{\lambda}{2}c + \frac{88\lambda - 73\lambda^2 - 16}{16\lambda}c^2\right)n^2 = \left(\frac{4}{13}c + \frac{111}{104}c^2\right)n^2. \end{aligned}$$

Since $\frac{4}{13}c + \frac{111}{104}c^2 > \frac{9}{4}c^2$ holds for all $c \in (0, \frac{32}{123})$ and $\frac{32}{123} > 0.26$, we derive that $e(A,B) > \frac{9}{4}c^2n^2$ whenever $c \leq 0.26$. Therefore, by Lemma 7.1.12, G contains a vertex set of size $n/2$ that spans strictly less than $n^2/18$ edges. ∎

7.1.4.2 Dense range

In this section we prove the following for graphs with high minimum degree.

Theorem 7.1.13. *Suppose that G is a K_4-free graph on n vertices with $\delta(G) \geq 0.59n$. Then G contains a vertex set of size $\lfloor n/2 \rfloor$ that spans at most $n^2/18$ edges. Moreover, if every vertex set of size $\lfloor n/2 \rfloor$ in G spans at least $n^2/18$ edges, then $G \cong T_3(n)$.*

To show this, we need a structural result on dense K_4-free graphs. A K_r-free graph G is maximal if adding any new edge to G will result in a copy of K_r. Let G_1 and G_2 be two vertex disjoint graphs. The join of G_1 and G_2, denoted by $G_1 \vee G_2$, is a graph with $V(G_1 \vee G_2) = V(G_1) \cup V(G_2)$ and

$$E(G_1 \vee G_2) = E(G_1) \cup E(G_2) \cup \{uv\colon u \in V(G_1), v \in V(G_2)\}.$$

Theorem 7.1.14 (Lyle [182]). *Let G be a maximal K_4-free on n vertices with $\delta(G) \geq 4n/7$. Then either G contains an independent set of size at least $4\delta(G) - 2n$ or G is the join of an independent set and a triangle-free graph.*

Our next lemma shows that if a K_4-free graph G contains a large induced triangle-free graph, then G contains a sparse half.

Lemma 7.1.15. *Let G be a K_4-free graph on n vertices. Suppose that G contains an induced triangle-free subgraph Γ with at least $2n/3$ vertices. Then G contains a vertex set of size $n/2$ that spans at most $n^2/18$ edges. Moreover, if $|V(\Gamma)| > 2n/3$, then G contains a vertex set of size $n/2$ which spans strictly less than $n^2/18$ edges.*

Proof. Let $A \subset V(G)$ such that $\Gamma = G[A]$ and let $x = |A|/n$. We may assume that $x \leq 5/6$ since otherwise we could choose $A' \subset A$ with $|A'| = 5n/6$ and consider $G[A']$ instead. Let $\alpha = 1/(2x)$. Then $\alpha \geq 3/5$. By a result of Krivelevich on triangle-free graphs [153], there exists $T \subset A$ with $|T| = \alpha|A| = n/2$ such that

$$e(T) \leq \frac{2 \times \frac{1}{2x} - 1}{4}|A|^2 = \frac{(1-x)x}{4}n^2 \leq \frac{n^2}{18},$$

where in the last inequality we used the assumption that $x \geq 2/3$. Notice that if $x > 2/3$, then the inequality above is strict. This proves the lemma. ∎

We also need the following slightly stronger version of Krivelevich's theorem on local densities of triangle-free graphs. A proof is included in the appendix, which follows from a detailed analysis of Krivelevich's proof in [153] as well as the proof of Erdős et al. in [65].

Theorem 7.1.16 (Krivelevich [153])**.** *Let $3/5 < \alpha \leq 1$, $n \in \mathbb{N}$ and $\alpha n \in \mathbb{N}$. Let G be a triangle-free graph on n vertices. If every vertex set of size αn in G spans at least $\frac{2\alpha-1}{4}n^2$ edges, then $G \cong T_2(n)$.*

Now we are ready to prove Theorem 7.1.13.

Proof of Theorem 7.1.13. It is clear that to prove Theorem 7.1.13, it suffices to consider maximal K_4-free graphs. Let G be a maximal K_4-free graph on n vertices with $\delta(G) \geq 0.59n > 4n/7$. Then by Theorem 7.1.14, either G is the join of an independent set and a triangle-free graph or G contains an independent set of size at least $4\delta(G) - 2n$.

First, suppose that the former case occurs, that is, G is the join of an independent set I and a triangle-free graph Γ. Let $\alpha = |V(\Gamma)|/n$. So $|I| = (1-\alpha)n$. We may assume that $\alpha > 1/2$ since otherwise we can simply choose a subset of I with size $n/2$ which spans none of edges. On the other hand, if $\alpha > 2/3$, then by Lemma 7.1.15, we are done. So we may assume that $1/2 < \alpha \leq 2/3$.

Let $c = e(\Gamma)/(\alpha n)^2$. If $c < 2/9$, then by Proposition 7.1.6, there exists $S \subset V(\Gamma)$ with $|S| = n/2$ such that

$$e(S) \leq \left(\frac{1/2}{\alpha}\right)^2 c(\alpha n)^2 = \frac{1}{4}cn^2 < \frac{n^2}{18}.$$

So we may assume that $c \geq 2/9$. Since Γ has αn vertices and at least $2\alpha^2 n^2/9$ edges, there exists some $v \in V(\Gamma)$ such that $d_\Gamma(v) \geq 4\alpha n/9 \geq (\alpha - 1/2)n$, where the last inequality holds as $\alpha \leq 2/3$. Let $T \subset N_\Gamma(v)$ be any subset with $|T| = (\alpha - 1/2)n$. Since Γ is triangle-free, T is an independent set. Therefore, $I \cup T$ has size $n/2$ and satisfies

$$e(I \cup T) \leq (1-\alpha)\left(\alpha - \frac{1}{2}\right)n^2 \leq \frac{n^2}{18},$$

where the last inequality uses the assumption that $\alpha \leq 2/3$. Notice that if $\alpha < 2/3$, then the inequality above is strict.

Now we may assume that G contains an independent set A whose size is at least $4\delta(G) - 2n \geq 9n/25$. We may just take A such that $|A| = 9n/25$. Let $B = V(G) \setminus A$. By Proposition 7.1.7, there exists $U \subset B$ with $|U| = 7n/50$ such that

$$\begin{aligned} e(A \cup U) &\leq \frac{7/50}{16/25}e(A,B) + \left(\frac{7/50}{16/25}\right)^2 e(B) = \frac{175}{1024}e(A,B) + \frac{49}{1024}(e(A,B) + e(B)) \\ &\leq \frac{175}{1024}\left(\frac{9}{25}n \times \frac{16}{25}n\right) + \frac{49}{1024} \times \frac{n^2}{3} = \frac{4249}{76800}n^2 < \frac{n^2}{18}. \end{aligned}$$

Therefore, $A \cup U$ is a sparse half with $e(A \cup U) < n^2/18$.

From the arguments above, one could see that if every vertex set of size $n/2$ in G spans at least $n^2/18$ edges, then G must be the join of a triangle-free graph Γ and an independent set I with $|V(\Gamma)| = 2n/3$. Since every vertex set of size $n/2 = \frac{3}{4}|V(\Gamma)|$ in Γ spans at least $n^2/18 = (2 \cdot 3/4 - 1)|V(\Gamma)|^2/4$ edges, by Theorem 7.1.16 we have $\Gamma \cong T_2(2n/3)$, which implies $G \cong T_3(n)$. This finishes the proof of Theorem 7.1.13. ∎

7.1.4.3 Intermediate range

In this section we will prove the following result for regular graphs.

Theorem 7.1.17. *Every K_4-free regular graph G on n vertices with $e(G) \leq 0.297n^2$ contains a vertex set of size $\lfloor n/2 \rfloor$ that spans strictly less than $n^2/18$ edges.*

We would like to remind the reader that the assumption that G is regular in Theorem 7.1.17 can be replaced by $\Delta(G) - \delta(G) \leq \epsilon n$ for some absolute (but small) constant $\epsilon > 0$. However, in order to keep the proof simple we shall only consider regular graphs.

The proof ideas are as follows. First, under the assumption that all subsets of size $n/2$ span at least $n^2/18$ edges, we show that G must contain many triangles. Then we show that there exists a partition $V(G) = V_1 \cup V_2 \cup V_3 \cup V_4$ such that V_1, V_2, V_3 are independent sets and $|V_1| + |V_2| + |V_3|$ is relatively large. Finally, utilizing this partition, we employ some ad hoc probabilistic arguments to find a sparse half and thus reach a contradiction.

Recall that $t(G)$ denotes the number of triangles in G.

Lemma 7.1.18. *Let G be an n-vertex K_4-free graph with cn^2 edges and $n/2 \leq \delta(G) \leq \Delta(G) \leq 9n/14$. Suppose that every vertex set of size $\lfloor n/2 \rfloor$ in G spans at least $n^2/18$ edges. Then we have $t(G) \geq \frac{c}{27(1-2c)}n^3$.*

Proof. For every $v \in V(G)$ let $\alpha_v = \frac{n}{2d(v)}$ and $c_v = e(N(v))/ (d(v))^2$. First notice that $c_v \geq 2/9$ for all $v \in V(G)$, since otherwise by Proposition 7.1.6, there would be a set $S \subset N(v)$ with $|S| = n/2$ such that $e(S) \leq \alpha_v^2 \cdot e(N(v)) < n^2/18$, a contradiction.

Fix $v \in V(G)$. Since $d(v) \leq 9n/14$, we see $\alpha_v \geq 7/9 \geq 1 - c_v$. By Theorem 7.1.5 and our assumption, there exists a vertex set $T \subset N(v)$ with $|T| = n/2$ such that $\frac{n^2}{18} \leq e(T) \leq (2\alpha_v - 1)e(N(v))$. This implies that for all $v \in V(G)$,

$$e(N(v)) \geq \frac{n^2}{18}\frac{1}{2\alpha_v - 1} = \frac{n^2}{18}\frac{d(v)}{n - d(v)}.$$

Summing over all $v \in V(G)$, we obtain that $t(G) = \frac{1}{3}\sum_{v\in V(G)} e(N(v))$ is at least

$$\frac{n^2}{54}\sum_{v\in V(G)}\frac{d(v)}{n-d(v)} \geq \frac{n^2}{54}\frac{\sum_{v\in V(G)} d(v)}{\sum_{v\in V(G)}(n-d(v))} = \frac{n^2}{54}\frac{2e(G)}{n^2 - 2e(G)} = \frac{c}{27(1-2c)}n^3.$$

Here we used Jensen's inequality and the fact that $\frac{x}{n-x}$ is convex for $x \in (0, n)$. ∎

We also need the following lemma in [237]. For distinct $u, v \in V(G)$, let $N(uv)$ denote the set of common neighbors of u and v and let $d(uv) = |N(uv)|$.

Lemma 7.1.19 (Sudakov [237])**.** *Every graph G with e edges and m triangles contains a triangle uvw such that $d(uv) + d(vw) + d(wu) \geq \frac{9m}{e}$.*

Notice that if G is K_4-free, then $N(uv)$ is independent for all $uv \in E(G)$ and $N(uv) \cap N(vw) = \emptyset$ for all triangles uvw in G. The following lemma is an immediate consequence of Lemmas 7.1.18 and 7.1.19.

Lemma 7.1.20. *Let G be an n-vertex K_4-free graph with cn^2 edges and $n/2 \le \delta(G) \le \Delta(G) \le 9n/14$. Suppose every vertex set of size $\lfloor n/2 \rfloor$ in G spans at least $n^2/18$ edges. Then there exist three disjoint independent sets V_1, V_2, V_3 in G such that*

$$|V_1| + |V_2| + |V_3| \ge \frac{n}{3(1-2c)}.$$

Now we are ready to prove Theorem 7.1.17.

Proof of Theorem 7.1.17. Let G be a K_4-free regular graph on n vertices with cn^2 edges, where $c \le 0.297$. Suppose that every vertex set of size $n/2$ in G spans at least $n^2/18$ edges. By Theorem 7.1.9, we may assume that $c \in [1/4, 0.297]$. So every vertex has degree $2cn$ with $n/2 \le 2cn \le 0.594n < 9n/14$. Then by Lemma 7.1.20, there exist three disjoint independent sets V_1, V_2, V_3 in G such that

$$|V_1| + |V_2| + |V_3| = g(c)n, \quad \text{where} \quad g(c) = \frac{1}{3(1-2c)}.$$

Let $V_4 = V(G) \setminus \left(\bigcup_{i=1}^{3} V_i \right)$ and let $x_i = |V_i|/n$ for $i \in [4]$. Without loss of generality we may assume that $1/2 > x_1 \ge x_2 \ge x_3$. Let $e_{ij} = e(V_i, V_j)$ for all $\{i, j\} \subset [4]$ (so $e_{ij} = e_{ji}$) and let $e_4 = e(V_4)$. We will consider four cases depending on the values of x_1, x_2 and x_3.

Case 1: $x_1 + x_2 \ge x_1 + x_3 \ge x_2 + x_3 \ge \frac{1}{2}$.

Now we choose different $n/2$ vertices from G according to Table 1. For example, the second row in Table 1 means to choose all vertices of V_1 and choose a set $S \subset V_2$ with $|S| = (1/2 - x_1)\, n$

V_1	V_2	V_3	V_4
x_1	$1/2 - x_1$	0	0
x_1	0	$1/2 - x_1$	0
0	x_2	$1/2 - x_2$	0
$1/2 - x_2$	x_2	0	0
$1/2 - x_3$	0	x_3	0
0	$1/2 - x_3$	x_3	0

TABLE I

DIFFERENT SCHEMES FOR CHOOSING $N/2$ VERTICES FROM G.

uniformly at random. Then the expected value of $e(V_1 \cup S)$ is $\frac{1/2-x_1}{x_2}e_{12}$. So there exists $S \subset V_2$ with $|S| = (1/2 - x_1)\,n$ such that $e(V_1 \cup S) \leq \frac{1/2-x_1}{x_2}e_{12}$. By assumption, we have

$$\frac{1/2 - x_1}{x_2}e_{12} \geq \frac{n^2}{18} \quad \Rightarrow \quad e_{12} \geq \frac{1}{9}\frac{x_2}{1-2x_1}n^2.$$

Similarly, one can get from Table 1 that for all $(i,j) \in [3] \times [3]$ with $i \neq j$,

$$\frac{1/2 - x_i}{x_j}e_{ij} \geq \frac{n^2}{18} \quad \Rightarrow \quad e_{ij} \geq \frac{1}{9}\frac{x_j}{1-2x_i}n^2.$$

Adding them up, we obtain that $e_{12} + e_{13} + e_{23} = \frac{1}{2}\sum_{i \neq j} e_{ij}$ is at least

$$\frac{1}{18}\left(\frac{x_2 + x_3}{1 - 2x_1} + \frac{x_1 + x_3}{1 - 2x_2} + \frac{x_2 + x_1}{1 - 2x_3}\right)n^2 = \frac{1}{18}\left(\frac{g(c) - x_1}{1 - 2x_1} + \frac{g(c) - x_2}{1 - 2x_2} + \frac{g(c) - x_3}{1 - 2x_3}\right)n^2.$$

Since $\frac{g(c)-x}{1-2x}$ is convex, by Jensen's inequality we see that

$$e_{12} + e_{13} + e_{23} \geq \frac{1}{6} \cdot \frac{g(c) - (x_1 + x_2 + x_3)/3}{1 - 2(x_1 + x_2 + x_3)/3} n^2 = \frac{g(c)}{3(3 - 2g(c))} n^2. \tag{7.1}$$

On the other hand, since G is regular,[1] we have

$$e_{14} + e_{24} + e_{34} + 2e_4 = \sum_{v \in V_4} d(v) = 2cn \times |V_4| = 2c(1 - g(c))n^2. \tag{7.2}$$

Since $G[V_4]$ is K_4-free, by Turán's theorem we get

$$e_4 \leq \frac{1}{3}|V_4|^2 = \frac{(1 - g(c))^2}{3} n^2. \tag{7.3}$$

Therefore, it follows from Equation 7.1, Equation 7.2 and Equation 7.1 that (recall that V_1, V_2, V_3 are independent)

$$cn^2 + \frac{(1 - g(c))^2}{3} n^2 \geq e(G) + e_4 \geq \frac{g(c)}{3(3 - 2g(c))} n^2 + 2c(1 - g(c))n^2,$$

---

[1] We point out that throughout the proof of Theorem 7.1.17, this is the only place where we need the restriction that G is regular.

which is a contradiction because

$$h(c) := \frac{g(c)}{3(3-2g(c))} + 2c\,(1-g(c)) - \left(c + \frac{(1-g(c))^2}{3}\right)$$

is decreasing in c for $c \in [1/4, 0.297]$ and $h(0.297) > 0$ (see [183]). This proves Case 1.

Case 2: $x_2 + x_3 \le x_1 + x_3 \le x_1 + x_2 < \frac{1}{2}$.

Note that this case can exist only when $g(c) < 3/4$, which implies $c < 5/18$.

V_1	V_2	V_3	V_4
x_1	x_2	$1/2 - x_1 - x_2$	0
x_1	$1/2 - x_1 - x_3$	x_3	0
$1/2 - x_2 - x_3$	x_2	x_3	0
x_1	0	0	$1/2 - x_1$
0	x_2	0	$1/2 - x_2$
0	0	x_3	$1/2 - x_3$

TABLE II

DIFFERENT SCHEMES FOR CHOOSING $N/2$ VERTICES FROM G.

Now we choose $n/2$ vertices according to Table 2. Then similar to Case 1, we obtain that for every $k \in [3]$ and $\{i, j\} = [3]\backslash\{k\}$,

$$e_{ij} + \frac{1/2 - x_i - x_j}{x_k}(e_{ik} + e_{jk}) \geq \frac{n^2}{18}, \tag{7.4}$$

and for all $i \in [3]$

$$\frac{1/2 - x_i}{x_4} e_{i4} + \left(\frac{1/2 - x_i}{x_4}\right)^2 e_4 \geq \frac{n^2}{18}. \tag{7.5}$$

By simplifying the linear combination of

$$\sum_{k \in [3]} \left(\frac{x_k}{x_4} \times \text{Equation 7.4}\right) + \sum_{i \in [3]} \left(\frac{x_4}{1/2 - x_i} \times \text{Equation 7.5}\right),$$

we can derive that

$$\begin{aligned} e(G) + \frac{e_4}{2x_4} \geq \frac{x_1 + x_2 + x_3}{18x_4} n^2 + \frac{x_4}{18}\left(\frac{1}{1/2 - x_1} + \frac{1}{1/2 - x_2} + \frac{1}{1/2 - x_3}\right) n^2 \\ \geq \frac{x_1 + x_2 + x_3}{18x_4} n^2 + \frac{x_4}{18} \times \frac{3}{1/2 - (x_1 + x_2 + x_3)/3} n^2 = \frac{1 - x_4}{18x_4} n^2 + \frac{x_4}{3 - 2(1 - x_4)} n^2. \end{aligned}$$

Since $e_4 \leq |V_4|^2/3 = x_4^2 n^2/3$ and $x_4 = 1 - g(c)$, the inequality above implies

$$c + \frac{1 - g(c)}{6} \geq \frac{1 - (1 - g(c))}{18(1 - g(c))} + \frac{1 - g(c)}{3 - 2(1 - (1 - g(c)))} = \frac{g(c)}{18(1 - g(c))} + \frac{1 - g(c)}{3 - 2g(c)},$$

which is a contradiction because

$$k(c) := c + \frac{1-g(c)}{6} - \left(\frac{g(c)}{18(1-g(c))} + \frac{1-g(c)}{3-2g(c)} \right)$$

is strictly smaller than 0 for $c \in [1/4, 5/18)$ (see [183]). This proves Case 2.

Case 3: $x_2 + x_3 < \frac{1}{2} \leq x_1 + x_3 \leq x_1 + x_2$.

V_1	V_2	V_3	V_4
x_1	$1/2 - x_1$	0	0
x_1	0	$1/2 - x_1$	0
$1/2 - x_2 - x_3$	x_2	x_3	0
x_1	0	0	$1/2 - x_1$
$1/2 - x_2 - x_4$	x_2	0	x_4
$1/2 - x_3 - x_4$	0	x_3	x_4

TABLE III

DIFFERENT SCHEMES FOR CHOOSING $N/2$ VERTICES FROM G.

We choose $n/2$ vertices according to Table 3. Similar as above, we can obtain that

$$\frac{1/2-x_1}{x_i}e_{1i} \geq \frac{n^2}{18} \quad \text{for} \quad i \in \{2,3\}, \tag{7a}$$

$$e_{23} + \frac{1/2-x_2-x_3}{x_1}(e_{12}+e_{13}) \geq \frac{n^2}{18}, \tag{7b}$$

$$\frac{1/2-x_1}{x_4}e_{14} + \left(\frac{1/2-x_1}{x_4}\right)^2 e_4 \geq \frac{n^2}{18}, \quad \text{and} \tag{7c}$$

$$e_{j4} + e_4 + \frac{1/2-x_j-x_4}{x_1}(e_{1j}+e_{14}) \geq \frac{n^2}{18} \quad \text{for} \quad j \in \{2,3\}. \tag{7d}$$

By simplifying the linear combination of

$$\sum_{i=2,3}\left(\frac{x_i^2}{(1/2-x_1)x_1} \times (7a)\right) + (7b) + \frac{x_4^2}{(1/2-x_1)x_1} \times (7c) + \sum_{j=2,3}(7d),$$

we derive that

$$e(G) + \frac{e_4}{2x_1} \geq \left(\frac{1}{6} + \frac{x_2^2+x_3^2+x_4^2}{9x_1(1-2x_1)}\right)n^2.$$

Since $e_4 \leq |V_4|^2/3 = x_4^2 n^2/3$, the inequality above implies that

$$c + \frac{x_4^2}{6x_1} \geq \frac{1}{6} + \frac{x_2^2+x_3^2+x_4^2}{9x_1(1-2x_1)}.$$

This is a contradiction due to the following claim whose proof can be found in the appendix.

Claim 7.1.21. *Under the conditions of Case 3, we have*

$$\frac{1}{6}+\frac{x_2^2+x_3^2+x_4^2}{9x_1(1-2x_1)}-\frac{x_4^2}{6x_1}-c>0.$$

This contradiction completes the proof of Case 3.

Case 4: $x_2+x_3 \le x_1+x_3 < \frac{1}{2} \le x_1+x_2$.

V_1	V_2	V_3	V_4
x_1	$1/2-x_1$	0	0
x_1	0	$1/2-x_1$	0
x_1	$1/2-x_1-x_3$	x_3	0
$1/2-x_2-x_3$	x_2	x_3	0
x_1	0	0	$1/2-x_1$
0	x_2	0	$1/2-x_2$
$1/2-x_3-x_4$	0	x_3	x_4
0	$1/2-x_3-x_4$	x_3	x_4

TABLE IV

DIFFERENT SCHEMES FOR CHOOSING $N/2$ VERTICES FROM G.

Choosing $n/2$ vertices according to Table 4, we obtain that

$$\frac{1/2-x_i}{x_{3-i}}e_{i,3-i} \geq \frac{n^2}{18} \quad \Rightarrow \quad e_{i,3-i} \geq \frac{x_{3-i}}{1/2-x_i}\frac{n^2}{18} \quad \text{for each } i \in [2], \tag{7e}$$

$$e_{j3} + \frac{1/2-x_j-x_3}{x_{3-j}}(e_{12}+e_{3-j,3}) \geq \frac{n^2}{18} \quad \text{for each } j \in [2], \tag{7f}$$

$$\frac{1/2-x_k}{x_4}e_{k4} + \left(\frac{1/2-x_k}{x_4}\right)^2 e_4 \geq \frac{n^2}{18} \quad \text{for each } k \in [2], \quad \text{and} \tag{7g}$$

$$e_{34} + e_4 + \frac{1/2-x_3-x_4}{x_\ell}(e_{\ell 3}+e_{\ell 4}) \geq \frac{n^2}{18} \quad \text{for each } \ell \in [2]. \tag{7h}$$

By simplifying the linear combination of

$$\frac{1}{2}\left(1+\frac{1}{1-2x_3}-\frac{1}{x_1+x_2}\right)\sum_{i\in[2]}(7e) + \frac{1}{1-2x_3}\sum_{j\in[2]}\left(\frac{x_{3-j}}{x_1+x_2}\times(7f)\right)$$
$$+\frac{1}{2(x_1+x_2)}\sum_{k\in[2]}\left(\frac{x_4}{1/2-x_k}\times(7g)\right) + \frac{x_1+x_2}{1/2-x_3-x_4}\sum_{\ell\in[2]}\left(\frac{x_\ell}{1/2-x_3-x_4}\times(7h)\right),$$

it yields that

$$e(G) + \frac{1-x_1-x_2}{2(x_1+x_2)x_4}e_4 \geq \left(\frac{1}{2}\left(1+\frac{1}{1-2x_3}-\frac{1}{x_1+x_2}\right)\left(\frac{x_2}{1/2-x_1}+\frac{x_1}{1/2-x2}\right)\right.$$
$$\left.+\frac{1}{1-2x_3}+\frac{1}{2(x_1+x_2)}\left(\frac{x_4}{1/2-x_1}+\frac{x_4}{1/2-x_2}\right)+1\right)\frac{n^2}{18}.$$

Since $e_4 \leq |V_4|^2/3 = x_4^2n^2/3$, the inequality above implies that

$$c + \frac{(1-x_1-x_2)x_4}{6(x_1+x_2)} \geq \frac{1}{18}\left(\frac{1}{2}\left(1+\frac{1}{1-2x_3}-\frac{1}{x_1+x_2}\right)\left(\frac{x_2}{1/2-x_1}+\frac{x_1}{1/2-x2}\right)\right.$$
$$\left.+\frac{1}{1-2x_3}+\frac{1}{2(x_1+x_2)}\left(\frac{x_4}{1/2-x_1}+\frac{x_4}{1/2-x_2}\right)+1\right).$$

Again, this is a contradiction because of the following claim, whose proof is included in the appendix.

Claim 7.1.22. *Under the conditions of Case 4, we have*

$$c + \frac{(1-x_1-x_2)x_4}{6(x_1+x_2)} < \frac{1}{18}\left(\frac{1}{2}\left(1+\frac{1}{1-2x_3}-\frac{1}{x_1+x_2}\right)\left(\frac{x_2}{1/2-x_1}+\frac{x_1}{1/2-x2}\right)\right.$$
$$\left.+\frac{1}{1-2x_3}+\frac{1}{2(x_1+x_2)}\left(\frac{x_4}{1/2-x_1}+\frac{x_4}{1/2-x_2}\right)+1\right).$$

This completes the proof of Theorem 7.1.17. ∎

7.1.4.4 Proof of Theorem 7.1.3

Let G be a K_4-free regular graph on n vertices such that every vertex set of size $\lfloor n/2 \rfloor$ in G spans at least $n^2/18$ edges. Our goal is to show that n is divisible by 6 and $G \cong T_3(n)$

First we show that it suffices to consider the case that n is divisible by 6. Assume that we have proved for all n that are divisible by 6, and now consider the case that n is not divisible by 6. Let H be the blow-up of G obtained by replacing every vertex $i \in V(G)$ by a set V_i of size 6 and replacing every edge $ij \in E(G)$ by a complete bipartite graph with parts V_i and V_j. Then H contains $N := 6n$ vertices and is K_4-free and regular, hence by our assumption, if let $S \subset V(H)$ be a subset of size $N/2 = 3n$ spanning the minimum number of edges, then we have $e(S) \le N^2/18$. We may assume that S either contains V_i or is disjoint from V_i for all but at most one i, since if there are two indices i, j satisfying $1 \le |S \cap V_\ell| \le 5$ for $\ell \in \{i, j\}$, then we could increase one of the intersections and decreasing the other until $|S \cap V_\ell| \in \{0, 6\}$

for some ℓ, without increasing $e(S)$. So, S contains $\lfloor n/2 \rfloor$ sets V_i. By our assumption on G, $e(S) \geq 36 \lceil n^2/18 \rceil > N^2/18$, a contradiction.

Now we assume that n is divisible by 6. Let $e(G) = cn^2$ for some $c \in (0, 1/3]$. Then every vertex in G has degree $2cn$. If $c \leq 0.26$, then by Theorem 7.1.9, there exists a vertex set of size $n/2$ that spans strictly less than $n^2/18$ edges, a contradiction. If $c \geq 0.295$, then by Theorem 7.1.13, we can derive that $G \cong T_3(n)$. So it remains to consider $0.26 < c < 0.295$. In this case, by Theorem 7.1.17, G contains a vertex set of size $n/2$ that spans strictly less than $n^2/18$ edges, again a contradiction. We have completed the proof of Theorem 7.1.3. ∎

7.1.5 Concluding remarks

Another problem that is closely related to the Sparse halves problem is making a graph bipartite. A famous conjecture of Erdős [61] states that every triangle-free graph on n vertices can be made bipartite by deleting at most $n^2/25$ edges. This is still open, with the extremal graphs to be the balanced blow-ups of the 5-cycle. Following from Krivelevich's observation [153], we see that for regular graphs, Conjecture 7.1.1 would imply the above conjecture of Erdős. So it seems interesting (but perhaps still difficult) to attack Conjecture 7.1.1 for regular graphs.

For analogous problems on K_r-free graphs and other related problems, we direct interested readers to [40; 64; 65; 237].

7.2 A generalize Erdős–Rademacher problem

7.2.1 Introduction

A classical result of Mantel [186] states that every graph on n vertices with $\lfloor n^2/4 \rfloor + 1$ edges contains at least one copy of K_3. Rademacher showed that there are actually at least $\lfloor n/2 \rfloor$ copies of K_3 in such graphs. Later, Erdős [56; **?**] proved that if $t \leq cn$ for some small constant $c > 0$, then every graph on n vertices with $\lfloor n^2/4 \rfloor + t$ edges contains at least $t \lfloor n/2 \rfloor$ copies of K_3. Erdős also conjectured that the same conclusion holds for all $t < n/2$. Later, Lovász and Simonovits [178] proved Erdős' conjecture and they also proved a similar result for K_k with $k \geq 4$. In [194], Mubayi extended their results by proving tight bounds on the number of copies of color critical graphs in a graph with a prescribed number of vertices and edges.

For a fixed graph F let $N_F(G)$ denote the number of copies of F in G. The F-covering number $\tau_F(G)$ of G is the minimum size of $S \subset V(G)$ such that every copy of F in G has at least one vertex in S. If $F = K_k$, then we simply use $N_k(G)$ and $\tau_k(G)$ to denote $N_{K_k}(G)$ and $\tau_{K_k}(G)$, respectively.

The classical Erdős–Rademacher problem is to determine the minimum value of $N_F(G)$ for graphs G with fixed number of vertices and edges. Very recently, Xiao and Katona [249] posed a generalized Erdős–Rademacher problem by putting constraints on $\tau_F(G)$. More precisely, they asked for the minimum value of $N_F(G)$ for graphs G with a fixed number of vertices and edges and a fixed F-covering number. In particular, they proved that every graph G on n vertices with $\lfloor n^2/4 \rfloor + 1$ edges and $\tau_3(G) = 2$ must contain at least $n - 2$ copies of K_3, which is substantially greater than the bound guaranteed by Rademacher's result. This phenomenon motivatedp them to pose the following conjectures for the general case.

Conjecture 7.2.1 (Xiao–Katona [249]). *Let $s > t \geq 1$ be fixed integers and let $n \geq n_0 = n_0(s,t)$ be sufficiently large. Then every graph G on n vertices with $\lfloor n^2/4 \rfloor + t$ edges and $\tau_3(G) \geq s$ contains at least $(s-1)\lfloor n/2 \rfloor + \lceil n/2 \rceil - 2(s-t)$ copies of K_3.*

Let V be a set of size n. Then a partition $V = V_1 \cup \cdots \cup V_{k-1}$ is called balanced if $\lceil n/(k-1) \rceil \geq |V_i| \geq \lfloor n/(k-1) \rfloor$ for all $i \in [k-1]$. For $k \geq 2$ define $t_k(n) = \prod_{1 \leq i < j \leq k} |V_i||V_j|$, where $V_1 \cup \cdots \cup V_k = [n]$ is a balanced partition.

Conjecture 7.2.2 (Xiao–Katona [249]). *Let $s > t \geq 1, k \geq 4$ be fixed integers. Then every graph G on n vertices with $t_{k-1}(n)+1$ edges and $\tau_k(G) \geq 2$ contains at least $(|V_1| + |V_2| - 2) \prod_{i=3}^{k-1} |V_i|$ copies of K_k, where $V_1 \cup \cdots \cup V_{k-1}$ is a balanced partition of $[n]$ with $|V_1| \geq \cdots \geq |V_{k-1}|$.*

Xiao and Katona claimed that there is a common generalization of Conjectures 7.2.1 and 7.2.2 without writing it explicitly. They also observed that the case $s \leq t$ of these questions is a consequence of the previously mentioned results of Rademacher, Erdős [56; **?**] and Lovász–Simonovits [178]. Indeed, it follows from a result of Lovász and Simonovits [178] that the graph obtained from the balanced complete $(k-1)$-partite n-vertex graph by adding t pairwise vertex-disjoint edges into a largest part minimizes the number of copies of K_k among all n-vertex graph with $t_{k-1}(n)+t$ edges. Moreover, this graph clearly has K_k-covering number $t \geq s$. It therefore suffices to consider only the case $s > t$ for these questions.

We show that Conjecture 7.2.1 is not true in general and give the correct bound on the number of copies of K_3 for all s, t and sufficiently large n. On the other hand, we prove Conjecture 7.2.2 for sufficiently large n and we also prove several generalizations of Conjecture 7.2.2 for graphs G with $t_{k-1}(n) + t$ edges and $\tau_k(G) \geq s$. Our method also gives a bound, which is tight up to a smaller order error term, for the number of color critical graphs F in a graph with a fixed number of vertices and edges and a fixed F-covering number.

7.2.1.1 Triangles

To motivate the following definitions let us look at a simple construction first. Suppose that n is an even integer and $s - t$ is a square. Then the graph G obtained from the complete bipartite graph with part sizes $n/2 + (s-t)^{1/2}$ and $n/2 - (s-t)^{1/2}$ by adding s pairwise vertex-disjoint edges to the larger part satisfies $\tau_3(G) = s$ and $e(G) = n^2/4 + t$. Moreover, $N_3(G) = s\left(n/2 - (s-t)^{1/2}\right)$, which is smaller than the bound in Conjecture 7.2.1 for all $t \geq 2$.

Now let us present the definitions we need in this section. Let $\mathbb{N} = \{0, 1, \ldots\}$ be the set of nonnegative integers. For $s > t \geq 1$ and $n \in \mathbb{N}$ let $e(n) = n^2 - 4t_2(n) = n^2 - 4\lfloor n^2/4 \rfloor \in \{0, 1\}$ and

$$M_{s,t} = M_{s,t}(n) = \left\{ m \in \mathbb{N} \colon\ (4s - 4t - 4m + e(n))^{1/2} \in \mathbb{N} \right\}.$$

Note that $M_{s,t} \neq \emptyset$ since $s - t \in M_{s,t}$. Let

$$m_{s,t} = m_{s,t}(n) = \min M_{s,t},$$

and let

$$R_3(n,s,t) = (4s - 4t - 4m_{s,t} + e(n))^{1/2} \in \mathbb{N}.$$

Define

$$n_{s,t}^{+} = \frac{1}{2}\left(n + R_3(n,s,t)\right) \quad \text{and} \quad n_{s,t}^{-} = \frac{1}{2}\left(n - R_3(n,s,t)\right).$$

Let $B_{s,t}(n)$ be the complete bipartite graph on n vertices with two parts V_1 and V_2 such that $|V_1| = n_{s,t}^{+}$ and $|V_2| = n_{s,t}^{-}$.

Let $\mathcal{BM}_{s,t}(n)$ consist of all graphs obtained from $B_{s,t}(n)$ as follows: take distinct vertices $u_1, \ldots, u_s, v_1, \ldots, v_s$ in V_1, add the edges $u_1v_1, \ldots, u_sv_s$ and remove $m_{s,t}$ distinct edges $e_1, \ldots, e_{m_{s,t}}$ such that every e_i has one endpoint in $\{u_1, \ldots, u_s, v_1, \ldots, v_s\}$ and the other endpoint in V_2, and there is no triangle with three edges in $\{e_1, \ldots, e_{m_{s,t}}, u_1v_1, \ldots, u_sv_s\}$ (see Figure 24 (a) and (b)).

Let $\mathcal{BS}_{s,t}(n)$ consists of all graphs obtained from $B_{s,t}(n)$ as follows: take distinct vertices $u'_1, \ldots, u'_{s-1}, v'_1, \ldots, v'_{s-1}$ in V_1 and distinct vertices u'_s, v'_s in V_2, add the edges $u'_1v'_1, \ldots, u'_sv'_s$ and remove $m_{s,t}$ distinct edges $e'_1, \ldots, e'_{m_{s,t}}$ such that every e'_i has one endpoint in the set

$$\{u'_1, \ldots, u'_{s-1}, v'_1, \ldots, v'_{s-1}\}$$

and the other endpoint in $\{u'_s, v'_s\}$ and there is no triangle with three edges in the set (see Figure 24 (c) and (d))

$$\{e'_1, \ldots, e'_{m_{s,t}}, u'_1v'_1, \ldots, u'_sv'_s\}.$$

.

We abuse notation by letting $BM_{s,t}(n)$ and $BS_{s,t}(n)$ denote a generic member of $\mathcal{BM}_{s,t}(n)$ and $\mathcal{BS}_{s,t}(n)$ respectively.

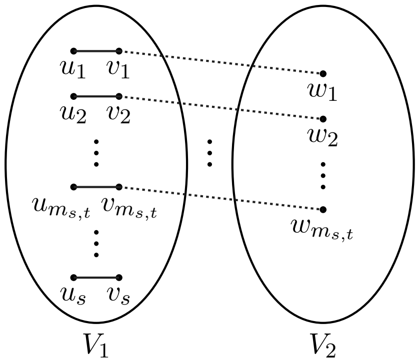
Remark. To compare with the original Erdős–Rademacher problem, i.e. without the $\tau_3(G) \geq s$ constraint, recall that the extremal graphs for that problem are graphs obtained from the balanced complete bipartite graph by adding a triangle-free graph with t edges into the larger part.

Fact 7.2.3. *The following statements hold.*

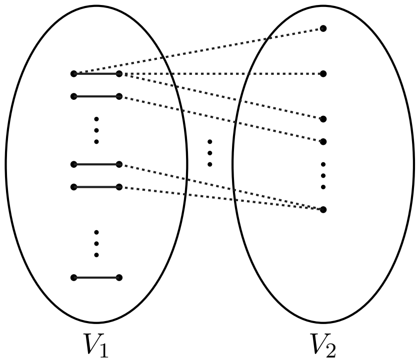
- $e\left(BM_{s,t}(n)\right) = e\left(BS_{s,t}(n)\right) = t_2(n) + t.$
- $\tau_3\left(BM_{s,t}(n)\right) = \tau_3\left(BS_{s,t}(n)\right) = s.$
- $N_3\left(BM_{s,t}(n)\right) = s \cdot n^-_{s,t} - m_{s,t}.$
- $N_3\left(BS_{s,t}(n)\right) = (s-1)n^-_{s,t} + n^+_{s,t} - 2m_{s,t} = s \cdot n^-_{s,t} - m_{s,t} + (n^+_{s,t} - n^-_{s,t} - m_{s,t}).$

By Lemma 7.2.20, if for some $p \in \mathbb{N}$

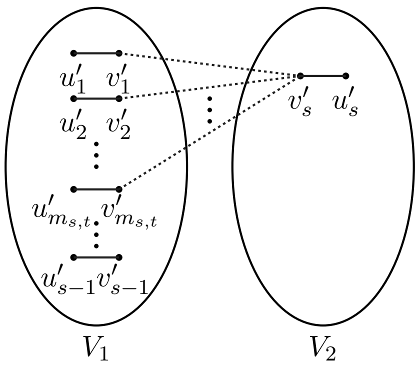
$$s - t = \begin{cases} p^2 - 1, & \text{if } n \text{ is even,} \\ p(p+1) - 1, & \text{if } n \text{ is odd,} \end{cases}$$



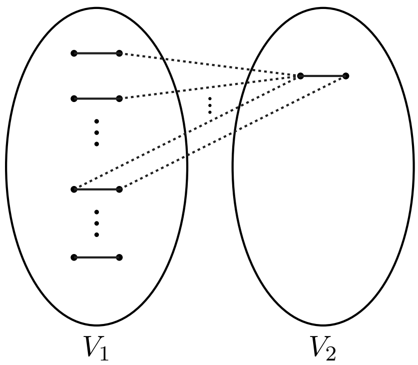
(a) $BM_{s,t}(n)$.



(b) $BM_{s,t}(n)$.



(c) $BS_{s,t}(n)$.



(d) $BS_{s,t}(n)$.

Figure 24. Several examples of graphs in $\mathcal{BM}_{s,t}(n)$ and $\mathcal{BS}_{s,t}(n)$.

then $N_3\left(BM_{s,t}(n)\right) = N_3\left(BS_{s,t}(n)\right) = s \cdot n^-_{s,t} - m_{s,t}$.

Our first result shows that $BM_{s,t}(n)$ (and also $BS_{s,t}(n)$ for some special values of s,t) contains the least number of copies of K_3 among all n-vertex graphs with $t_2(n)+t$ edges and K_3-covering number at least s.

Theorem 7.2.4. *Let $s > t \geq 1$. Then there exists $n_0 = n_0(s,t)$ such that the following holds for all $n \geq n_0$. Let G be a graph on n vertices with $t_2(n)+t$ edges. If $\tau_3(G) = s$, then*

$$N_3(G) \geq s \cdot n^-_{s,t} - m_{s,t}$$

Moreover, equality holds only if $G \cong BM_{s,t}(n)$ or $G \cong BS_{s,t}(n)$ except when n is even and $(s,t) \in \{(2,1),(3,1),(4,1)\}$, or n is odd and $(s,t) \in \{(3,2),(4,1),(5,1),(6,1)\}$. For these exceptional cases there are other examples showing that the bound is best possible.

Note that Theorem 7.2.4 shows that Conjecture 7.2.1 is not true in general. For example, let n be even, $(s-t)^{1/2} \in \mathbb{N}$ and $s-t > 4$. Then

$$N_3\left(BM_{s,t}(n)\right) = s \cdot n^-_{s,t} - m_{s,t} = s \cdot n^-_{s,t} = \frac{sn}{2} - (s-t)^{1/2}s,$$

which is strictly less that $sn/2 - 2(s-t)$.

Let $V_1 \cup \cdots \cup V_{k-1}$ be a partition of $[n]$ with $|V_1| \geq \cdots \geq |V_{k-1}|$. Let $K[V_1,\ldots,V_{k-1}]$ be the complete $(k-1)$-partite graph on $[n]$ with parts $V_1,\ldots,V_{k-1}$. If $V_1 \cup \cdots \cup V_{k-1}$ is a balanced partition, then $K[V_1,\ldots,V_{k-1}]$ is called the Turán graph $T_{k-1}(n)$. Notice that

$t_{k-1}(n) = |T_{k-1}(n)|$. The celebrated Turán theorem [242] states that the maximum number of edges of an n-vertex K_k-free graph is uniquely achieved by $T_{k-1}(n)$.

For $s > m \geq 0$ and $\vec{x} = (x_1, \ldots, x_{k-1}) \in \mathbb{N}^{k-1}$ with $\sum_{i=1}^{k-1} x_i = n$ let $V_1 \cup \cdots \cup V_{k-1}$ be a partition of $[n]$ with $|V_i| = x_i$ for $i \in [k-1]$. Let $\mathcal{KM}_{m,s}(\vec{x})$ consist of all graphs that are obtained from $K[V_1, \ldots, V_{k-1}]$ as follows: take distinct vertices $u_1, \ldots, u_s$, $v_1, \ldots, v_s$ in V_1, add the edges $u_1v_1, \ldots, u_sv_s$ and remove m distinct edges $e_1, \ldots, e_m$ such that every e_i contains one vertex from $\{u_1, \ldots, u_s, v_1, \ldots, v_s\}$ and one vertex from V_{k-1} and there is no triangle with edges in $\{e_1, \ldots, e_m, u_1v_1, \ldots, u_sv_s\}$. We abuse notation by letting $KM_{s,t}(\vec{x})$ denote a generic member in $\mathcal{KM}_{m,s}(\vec{x})$. It is easy to see that

$$e\left(KM_{m,s}(\vec{x})\right) = \sum_{1 \leq i < j < k} x_i x_j + s - m \quad \text{and} \quad N_k\left(KM_{m,s}(\vec{x})\right) = s \prod_{i=2}^{k-1} x_i - m \prod_{i=2}^{k-2} x_i.$$

Let us now consider some special cases of $KM_{m,s}(\vec{x})$ in more detail.

For $n \in \mathbb{N}$, write

$$n = q_{n,k}(k-1) + r_{n,k} \qquad \text{where} \qquad 0 \leq r_{n,k} < k-1.$$

Writing $r = r_{n,k}$ and $q = q_{n,k}$, let $\vec{y}_r \in \mathbb{N}^{k-1}$ be defined as follows:

$$\vec{y}_r = \begin{cases} (q+1, q, \ldots, q, q-1) & \text{if } r = 0 \\ (q+1, q, \ldots, q) & \text{if } r = 1 \\ (q+2, \underbrace{q+1, \ldots, q+1}_{r-2 \text{ times}}, \underbrace{q, \ldots, q}_{k-r \text{ times}}) & \text{if } r \geq 2. \end{cases}$$

Define

$$N_k(n,s) = \begin{cases} s \cdot q^{k-3}\,(q-1) & \text{if } r = 0, \\ s \cdot q^{k-2} - q^{k-3} & \text{if } r = 1, \\ s \cdot (q+1)^{r-2}\, q^{k-r} & \text{if } r \geq 2. \end{cases}$$

Observe that

$$e\left(KM_{0,s}(\vec{y}_r)\right) = e\left(KM_{1,s}(\vec{y}_1)\right) = t_{k-1}(n) + s - 1 \quad \text{for } r \neq 1$$

and

$$N_k\left(KM_{m,s}(\vec{y}_r)\right) = N_k(n,s)$$

for $m = 0, r \neq 1$ and $m = 1, r = 1$.

Our next result shows that the constructions defined above contain the least number of copies of K_k in an n-vertex graph G with $t_{k-1}(n)+s-1$ edges and $\tau_k(G)=s$.

Theorem 7.2.5. *Let $k\geq 4$ and $s\geq 2$ be fixed integers. Then there exists $n_1=n_1(k,s)$ such that the following holds for all $n\geq n_1$. Let G be a graph on n vertices with $t_{k-1}(n)+s-1$ edges. If $\tau_k(G)=s$, then $N_k(G)\geq N_k(n,s)$. Moreover, for $s\geq 3$ equality holds iff $G\cong KM_{0,s}(\vec{y}_{r_{n,k}})$ if $r_{n,k}\neq 1$ and $G\cong KM_{1,s}(\vec{y}_1)$ if $r_{n,k}=1$.*

For $s\geq 2$, the following construction which was defined in [249] also achieves the bound $N_k(n,2)$. Let $V_1\cup\cdots\cup V_{k-1}$ be a balanced partition of $[n]$ with $|V_1|\geq\cdots\geq|V_{k-1}|$. Let $T_{k-1}^{\sqsubset}$ be obtained from $K[V_1,\ldots,V_{k-1}]$ as follows: take two distinct vertices $u_1,v_1\in V_1$ and two distinct vertices $u_2,v_2\in V_2$, add edges u_1v_1,u_2v_2 and remove the edge v_1v_2. One can easily check that $N_k(T_{k-1}^{\sqsubset})=(|V_1|+|V_2|-2)\prod_{i=3}^{k-1}|V_i|=N_k(n,s)$. Therefore, Theorem 7.2.5 shows that Conjecture 7.2.2 is true for large n.

7.2.1.2 k-cliques for large s

Recall that for given n and k, $q_{n,k}=\lfloor n/(k-1)\rfloor$ and $r_{n,k}=n-(k-1)q_{n,k}$. Given $s>t\geq 1$ and $k\geq 3$, let

$$R_k(n,s,t)=\left(\frac{2(k-1)(s-t)+(k-1-r_{n,k})r_{n,k}}{k-2}\right)^{1/2}.$$

We note that while $R_k(n,s,t)$ depends on n it is bounded from above by a function of only k,s,t. Let

$$n_{k,s,t}^{+}=\frac{n+(k-2)R_k(n,s,t)}{k-1}\qquad\text{and}\qquad n_{k,s,t}^{-}=\frac{n-R_k(n,s,t)}{k-1}.$$

Suppose that $n^-_{k,s,t} \in \mathbb{N}$. Then let $V_1 \cup \cdots \cup V_{k-1}$ be a partition of $[n]$ with $|V_1| = n^+_{k,s,t}$ and $|V_i| = n^-_{k,s,t}$ for $2 \leq i \leq k-1$. Let $KM(n,k,s,t)$ be obtained from $K[V_1, \ldots, V_{k-1}]$ by taking distinct vertices $u_1, \ldots, u_s, v_1, \ldots, v_s$ in V_1 and then adding $u_1v_1, \ldots, u_sv_s$. Using Lemma 7.2.11 one can easily check that

$$e\left(KM(n,k,s,t)\right) = t_{k-1}(n) + t \quad \text{ and } \quad N_k\left(KM(n,k,s,t)\right) = s \cdot (n^-_{k,s,t})^{k-2}.$$

The following result shows that if s is large, then $KM(n,k,s,t)$ minimizes the number of copies of K_k among all n-vertex graphs G with $t_{k-1}(n) + t$ edges and $\tau_k(G) = s$.

Theorem 7.2.6. *Let $s > t \geq 1$ and $k \geq 4$ be fixed integers. There exists $n_2 = n_2(k,s,t)$ such that the following holds for all $n \geq n_2$ and $s > 2R_k(n,s,t)$. If G is a graph on n vertices with $t_{k-1}(n) + t$ edges and $\tau_k(G) = s$, then*

$$N_k(G) \geq s \cdot (n^-_{k,s,t})^{k-2}.$$

Moreover, if $n^-_{k,s,t} \in \mathbb{N}$, then equality holds iff $G \cong KM(n,k,s,t)$.

Note that we are not able to determine the exact minimum value of $N_k(G)$ for small s because, similar to the situation in Theorem 7.2.4, when s is small there could be many constructions that achieve the minimum value of $N_k(G)$. On the other hand, for the case $n^-_{k,s,t} \notin \mathbb{N}$ our bound might be not tight and actually, we think there might be a better bound for $N_k(G)$ in this case.

Let $R_k(s,t) = (2(k-1)(s-t)/(k-2))^{1/2}$. If $r_{n,k} = 0$, then $R_k(n,s,t) = R_k(s,t)$. Since $k \geq 4$ and $t \geq 1$, $s > 2R_k(s,t)$ holds for all $s \geq 11$. Therefore, Theorem 7.2.6 gives the following corollary.

Corollary 7.2.7. *Let $s > t \geq 1$ and $k \geq 4$ be fixed integers. Suppose that $s \geq 11$. Then there exists $n_3 = n_3(k,s,t)$ such that the following holds for all $n \geq n_3$ and $n \equiv 0 \mod k-1$. If G is a graph on n vertices with $t_{k-1}(n) + t$ edges and $\tau_k(G) = s$, then $N_k(G) \geq s \cdot (n^{-}_{k,s,t})^{k-2}$. Moreover, if $n^{-}_{k,s,t} \in \mathbb{N}$, then equality holds iff $G \cong KM(n,k,s,t)$.*

After this work was done we found that similar results as in Theorems 7.2.4, 7.2.5, and 7.2.6 were recently proved by Balogh and Clemen [12].

7.2.1.3 Color-critical graphs

Given a graph G let $\chi(G)$ denote the chromatic number of G. Let H be a subgraph of G. Then the graph $G - H$ is obtained from G by removing all edges that are contained in H. In particular, if $e \in E(G)$, then $G - e$ is obtained from G by removing e.

Definition 7.2.8. *Let $k \geq 3$. A graph F is k-critical if $\chi(F) = k$ and there exists $e \in E(F)$ such that $\chi(F - e) < k$.*

Let $k \geq 3$ and let F be a k-critical graph. Let $c(n,F)$ denote the minimum number of copies of F in the graph obtained from $T_{k-1}(n)$ by adding one edge. The number $c(n,F)$ can be calculated using a formula in [194] and in particular there exists a constant $\alpha_F > 0$ depending only on F such that $c(n,F) = \alpha_F n^{f-2} + \Theta(n^{f-3})$.

The second author proved [194] that for any k-critical graph F there exists a constant $\delta = \delta_F > 0$ such that for every $1 \le t \le \delta n$ every n-vertex graph G with $t_{k-1}(n) + t$ edges contains at least $t \cdot c(n, F)$ copies of F. We prove the analogous theorem for $\tau_F(G) = s$.

Theorem 7.2.9. *Let $s > t \ge 1$ and $k \ge 3$ be fixed integers. Let F be a fixed k-critical graph on f vertices. Then there exists constants $C = C(F, s, t)$ and $n_4 = n_4(F, s, t)$ such that the following holds for all $n \ge n_4$. If G is a graph on n vertices with $t_{k-1}(n) + t$ edges and $\tau_k(G) = s$, then $N_F(G) \ge s \cdot c(n, F) - Cn^{f-3}$.*

Remark. For graphs that are not color critical it remains open in general to determine even their Turán numbers exactly. Therefore, one could expect that a Erdős-Rademacher-type result (or result as Theorem 7.2.9) for these graphs can be very hard in general.

This bound is tight up to an error term since the graph obtained from $T_{k-1}(n)$ by adding s pairwise disjoint edges into one part of $T_{k-1}(n)$ contains at most $s \cdot c(n, F) + C'n^{f-3}$ copies of F for some constant $C' > 0$.

7.2.2 Proofs

7.2.2.1 Lemmas

In this section we prove several lemmas that will be used in our proofs.

Definition 7.2.10. *Let $k \ge 3$ and let F be a k-critical graph. Let $c(x_1, \ldots, x_{k-1}, F)$ be the number of copies of F in the graph obtained from the complete $(k-1)$-partite graph with parts of sizes $x_1, \ldots, x_{k-1}$ by adding one edge to the part of size x_1.*

The following explicit expression for $t_{k-1}(n)$ is very useful in our calculations.

Lemma 7.2.11 (e.g. see [178]). *Let* $k \geq 3$ *and suppose that* $n \equiv r \mod (k-1)$ *for some* $0 \leq r \leq k-2$. *Then*

$$t_{k-1}(n) = \frac{(k-2)}{2(k-1)}n^2 - \frac{(k-1-r)r}{2(k-1)}.$$

The following lemma gives a relation between $c(x_1, \ldots, x_{k-1}, F)$ and $c(n, F)$.

Lemma 7.2.12 ([194]). *Let* $k \geq 3$ *and* F *be a* k*-critical graph. Then there exists a constant* $\gamma_F > 0$ *depending only on* F *such that the following holds for all sufficiently large* n*. If* $\sum_{i=1}^{k-1} x_i = n$ *and* $\lfloor n/(k-1) \rfloor - d \leq x_i \leq \lceil n/(k-1) \rceil + d$ *for all* $i \in [k-1]$*and* $d \leq \frac{n}{3(k-1)}$, *then*

$$c(x_1, \ldots, x_{k-1}, F) \geq c(n, F) - \gamma_F d n^{f-3}.$$

The following lemma, which can be found in several places (e.g. see [194]), gives a bound on the size of each part for a $(k-1)$-partite graph whose number of edges is close to $t_{k-1}(n)$.

Lemma 7.2.13 (e.g. see [194]). *Suppose that* $k \geq 3$ *is fixed,* n *is sufficiently large,* $d < n$ *and* $\sum_{i=1}^{k-1} x_i = n$. *If*

$$\sum_{1 \leq i < j \leq k-1} x_i x_j \geq t_{k-1}(n) - d,$$

then $\lfloor n/(k-1) \rfloor - d \leq x_i \leq \lceil n/(k-1) \rceil + d$ *for all* $i \in [k-1]$.

The following two results will be key in our proofs.

Theorem 7.2.14 (Graph removal lemma, e.g. see [66]). *Let F be a graph with f vertices. Suppose that G is a graph on n vertices with $N_F(G) = o(n^f)$. Then one can remove $o(n^2)$ edges from G such that the resulting graph is F-free.*

Theorem 7.2.15 (Erdős–Simonovits stability theorem [233]). *Let $k \geq 3$ and F be a k-critical graph. Suppose that G is an F-free graph on n vertices with $t_{k-1}(n) - o(n^2)$ edges. Then G can be made $(k-1)$-partite by removing $o(n^2)$ edges.*

Now we use the results above to obtain a rough structure of a graph with a fixed number of vertices and edges and a fixed F-covering number that contains not many copies of F.

Given a graph G and $v \in V(G)$ we use $N_G(v)$ to denote the neighbors of v in G and let $d_G(v) = |N_G(v)|$. For a partition $V_1 \cup \cdots \cup V_{k-1}$ of $V(G)$ we use $G[V_1, \ldots, V_{k-1}]$ to denote the induced $(k-1)$-partite subgraph of G on $V_1 \cup \cdots \cup V_{k-1}$. We use $B_G(V_1, \ldots, V_{k-1})$ to denote the set of edges in G that are contained inside V_i for some $i \in [k-1]$, i.e. $B_G(V_1, \ldots, V_{k-1}) = G - G[V_1, \ldots, V_{k-1}]$. We use $M_G(V_1, \ldots, V_{k-1})$ to denote the set of pairs which intersect two parts that are not edges in G, i.e. $M_G(V_1, \ldots, V_{k-1}) = K[V_1, \ldots, V_{k-1}] - G[V_1, \ldots, V_{k-1}]$. If it is clear from the context we will use B and M to represent $B_G(V_1, \ldots, V_{k-1})$ and $M_G(V_1, \ldots, V_{k-1})$, respectively.

For a k-critical graph F a potential copy of F in G (with respect to the partition $V(G) = V_1 \cup \cdots \cup V_{k-1}$) is a copy of F in $G \cup M$ that uses exactly one edge of B (so every other edge is between parts).

Lemma 7.2.16. *Let $s \geq 1, f \geq k \geq 3$ be fixed integers and F be a fixed k-critical graph on f vertices. Then the following holds for sufficiently large n. If G is a graph on n vertices with* *at least* $t_{k-1}(n)+1$ *edges and* $N_F(G) \le (s+1/2)\cdot c(n,F)$*, then* G *contains a* $(k-1)$*-partite subgraph* H *such that* $e(H) \ge e(G) - s$*.*

Proof. Let $\delta_1, \delta_2, \delta_3, \delta_4, \epsilon, \epsilon_1, \epsilon_2$ be constants such that

$$0 < \delta_1 \ll \delta_2 \ll \delta_3 \ll \delta_4 \ll \epsilon_2 \ll \epsilon_1 \ll \epsilon \ll s^{-1}.$$

Let n be sufficiently large and in particular $n \gg s/\epsilon_2$.

Since $N_F(G) \le (s+1/2)\cdot c(n,F) < 2s\alpha_F n^{f-2} = o(n^f)$, by the Graph removal lemma, we can remove at most $\delta_1 n^2$ edges from G such that the resulting graph G_1 is F-free. Since $e(G_1) \ge e(G) - \delta_1 n^2 > t_{k-1}(n) - \delta_1 n^2$, by the Erdős-Simonovits stability theorem, G_1 contains a $(k-1)$-partite subgraph G_2 such that $e(G_2) \ge t_{k-1}(n) - \delta_2 n^2$.

Now let H be a $(k-1)$-partite subgraph of G with the maximum number of edges. Then by the previous argument, $e(H) \ge e(G_2) \ge t_{k-1}(n) - \delta_2 n^2$. Let $V_1 \cup \cdots \cup V_{k-1}$ be a partition of $V(G)$ such that $H = G[V_1, \ldots, V_{k-1}]$ and let $x_i = |V_i|$ for $i \in [k-1]$. An easy calculation shows that $|x_i - n/(k-1)| \le \delta_3 n$ for all $i \in [k-1]$.

Suppose that $|H| = t_{k-1}(n) - \ell$ for some $\ell \ge 0$. Then $|M| \le \ell$ and $|B| \ge \ell + 1$. For every $e \in B$ let $F(e)$ denote the number of copies of F in G containing the unique edge e from B. Let

$$B_1 = \{e \in B \colon F(e) > (1-\epsilon)c(n,F)\}$$

and $B_2 = B \setminus B_1$.

Claim 7.2.17. $|B_1| \geq (1-\epsilon)|B|$.

Proof. Suppose that $|B_2| \geq \epsilon|B|$. Let $e \in B_2$ and without loss of generality we may assume that $e \subset V_1$. Then by Lemma 7.2.12 the number of potential copies of F containing e is

$$c(x_1, \ldots, x_{k-1}, F) \geq c(n, F) - \gamma_F(\delta_3 n)n^{f-3} > (1-\delta_4)c(n, F).$$

At least $\epsilon \cdot c(n, F)/2$ of these potential copies of F have a pair from M, since otherwise

$$F(e) \geq (1-\delta_4)c(n, F) - \frac{\epsilon}{2}c(n, F) > (1-\epsilon)c(n, F),$$

a contradiction. Now suppose that at least $\epsilon \cdot c(n, F)/4$ of these potential copies of F have a pair from M that does not intersect e. For every $e' \in M$ with $e \cap e' = \emptyset$ the number of potential copies of F in G that contains both e and e' is at most n^{f-4}. On the other hand, every potential copy of F contains at most f^2 pairs from M. Therefore,

$$\frac{\epsilon}{4}c(n, F) \geq |M| f^2 n^{f-4},$$

which implies that

$$\delta_2 n^2 \geq |M| \geq \frac{\frac{\epsilon}{4}c(n, F)}{f^2 n^{f-4}} > \frac{\epsilon \alpha_F}{8f^2} n^2,$$

a contradiction. Here we used $|M| \le t_{k-1}(n) - e(H) \le \delta_2 n^2$. Therefore, we may assume that at least $\epsilon \cdot c(n,F)/4$ of these potential copies of F have a pair from M which has nonempty intersection with e. Similarly, since every $e'' \in M$ with $e'' \cap e \neq \emptyset$ is contained in at most n^{f-3} members in $F(e)$ and every potential copy of F contains at most f^2 pairs from M, the number of pairs from M that has nonempty intersection with e is at least

$$\frac{\frac{\epsilon}{4}c(n,F)}{f^2 n^{k-3}} \ge \frac{\epsilon \alpha_F}{8f^2} n.$$

Therefore, there exists $x \in e$ such that $d_M(x) \ge \frac{\epsilon\alpha_F}{16f^2}n$.

Let $A = \left\{ v \in V(G) \colon d_M(v) \ge \frac{\epsilon\alpha_F}{16f^2}n \right\}$. Since every $e \in B$ contains a vertex in A,

$$\sum_{v \in A} d_{B_2}(v) \ge |B_2| \ge \epsilon|B| \ge \epsilon|M| \ge \frac{\epsilon}{2}\sum_{v\in A} d_M(v) \ge \frac{\epsilon^2\alpha_F}{32f^2}n|A|.$$

Therefore, there exists $v \in A$ such that $d_{B_2}(v) \ge \frac{\epsilon^2\alpha_F}{32f^2}n$ and without loss of generality we may assume that $v \in V_1$. Let $V_i' = N_G(v) \cap V_i$ for $i \in [k-1]$. Then by the maximality of H we have $|V_i'| \ge |V_1'| \ge \frac{\epsilon^2\alpha_F}{32f^2}n$ for all $2 \le i \le k-1$. Let $u \in V_1'$. Then by Lemma 7.2.12, the number of potential copies of F containing uv in the complete $(k-1)$-partite graph $K[V_1', \ldots, V_{k-1}']$ is at least

$$c\left(|V_1'|, \ldots, |V_{k-1}'|, F\right) \ge \frac{1}{2}\alpha_F \left(\frac{\epsilon^2\alpha_F}{32f^2}n\right)^{k-2} \ge \epsilon_1 n^{k-2}.$$

Summing over all $u \in V_1'$, there are at least

$$\frac{\epsilon^2 \alpha_F}{32 f^2} n \times \epsilon_1 n^{f-2} \geq \epsilon_2 n^{f-1} \geq 3s \cdot c(n, F)$$

potential copies of F containing v. By the assumption that $N_F(G) \leq (s + 1/2) \cdot c(n, F)$, at least half of these potential copies of F must contain a pair from M, and this pair cannot be incident with v, since v is adjacent to all vertices in $\bigcup_{i=1}^{k-1} V_i'$. Since the number of potential copies of F that contain both v and a pair from M that is disjoint from v is at most n^{f-3} and each potential copy of F contains at most f^2 pairs from M, we obtain

$$\delta_2 n^2 \geq |M| \geq \frac{\epsilon_2 n^{f-1}/2}{f^2 n^{f-3}} \geq \frac{\epsilon_2}{2f^2} n^2,$$

a contradiction. ∎

Claim 7.2.18. *$|B| \leq s$.*

Proof. Suppose that $|B| \geq s + 1$. Then by Claim 7.2.17,

$$\begin{aligned} N_F(G) \geq \sum_{e \in B_1} F(e) \geq \sum_{e \in B_1} (1 - \epsilon) c(n, F) &\geq (1 - \epsilon)^2 |B| c(n, F) \\ &\geq (1 - \epsilon)^2 (s + 1) c(n, F) > (s + 1/2) \cdot c(n, F), \end{aligned}$$

a contradiction. ∎

Therefore, by Claim 7.2.18, $e(H) = e(G) - |B| \geq e(G) - s$. This completes the proof of Lemma 7.2.16. ∎

Now we use Lemma 7.2.16 to obtain a fine structure for graphs with a fixed F-covering number and not many copies of F.

Lemma 7.2.19. *Let $f \geq k \geq 3, s > t \geq 1$ be fixed integers and F be a fixed k-critical graph on f vertices. Then the following holds for sufficiently large n. Let G be a graph on n vertices with $t_{k-1}(n) + t$ edges. If $\tau_F(G) = s$ and $N_F(G) \leq (s + 1/2) \cdot c(n, F)$, then there exists a partition $V(G) = V_1 \cup \cdots \cup V_{k-1}$ such that $G - G[V_1, \ldots, V_{k-1}]$ is a matching with s edges.*

Proof. Let H be a $(k-1)$-partite subgraph of G with the maximum number of edges and let $B = G - H$. Since $N_F(G) \leq (s + 1/2) \cdot c(n, F)$, by Lemma 7.2.16, $|B| \leq s$. So it suffices to show that $|B| \geq s$ and B is a matching.

Recall that $\tau(B) = \min \{|S| \colon S \subset V(G), e \cap S \neq \emptyset \text{ for all } e \in B\}$. Since every copy of F in G must contain at least one edge in B, $\tau_F(G) \leq \tau(B)$. Therefore, $\tau(B) \geq s$. Since $|B| \leq s$, the only possibility is that B is a matching of size s. ∎

7.2.2.2 Proof of Theorem 7.2.4

In this section we prove Theorem 7.2.4. Recall that for $s > t \geq 1$ and $n \in \mathbb{N}$

$$n^+_{s,t} = \frac{1}{2}\left(n + R_3(n,s,t)\right) \quad \text{and} \quad n^-_{s,t} = \frac{1}{2}\left(n - R_3(n,s,t)\right),$$

where $R_3(n,s,t) = \left(4s - 4t - 4m_{s,t} + n^2 - 4t_2(n)\right)^{1/2}$ and

$$m_{s,t} = \min\left\{m \in \mathbb{N}\colon\ \left(4s - 4t - 4m + n^2 - 4t_2(n)\right)^{1/2} \in \mathbb{N}\right\}.$$

We will use the following lemma in our proof.

Lemma 7.2.20. *Let $s > t \geq 1$ and $n \in \mathbb{N}$. Then*

$$n^+_{s,t} - n^-_{s,t} - m_{s,t} = \begin{cases} 0 & \textit{if } n \textit{ is even and } s - t = p^2 - 1 \textit{ for some } p \in \mathbb{N}, \\ 0 & \textit{if } n \textit{ is odd and } s - t = p(p+1) - 1 \textit{ for some } p \in \mathbb{N}, \\ > 0 & \textit{otherwise.} \end{cases}$$

Proof. First, notice that $n^+_{s,t} - n^-_{s,t} - m_{s,t} = \left(4s - 4t - 4m_{s,t} + n^2 - 4t_2(n)\right)^{1/2} - m_{s,t}$.

If n is even, then $n^2 - 4t_2(n) = 0$. Let $p \in \mathbb{N}$ be the largest integer such that $s - t = p^2 + q$ for some $q \in \mathbb{N}$. Note that $q \leq 2p$ since otherwise we would have $p^2 + q \geq (p+1)^2$, a contradiction. Then $m_{s,t} = q$ and hence

$$\left(4s - 4t - 4m_{s,t} + n^2 - 4t_2(n)\right)^{1/2} - m_{s,t} = 2p - m_{s,t} \geq 0$$

and equality holds iff $q = 2p$.

If n is odd, then $n^2-4t_2(n)=1$. Let $p\in\mathbb{N}$ be the largest integer such that $s-t=p(p+1)+q$ for some $q\in\mathbb{N}$. Note that $q\le 2p+1$ since otherwise we would have $p(p+1)+q\ge(p+1)(p+2)$, a contradiction. Then $m_{s,t}=q$ and hence

$$\left(4s-4t-4m_{s,t}+n^2-4t_2(n)\right)^{1/2}-m_{s,t}=2p+1-m_{s,t}\ge 0$$

and equality holds iff $q=2p+1$. ∎

Now we are ready to prove Theorem 7.2.4.

Proof of Theorem 7.2.4. Let $s>t\ge 1$ be fixed and let n be sufficiently large. Let G be a graph on n vertices with $t_2(n)+t$ edges and $\tau_3(G)=s$. Since $s\cdot n^-_{s,t}-m_{s,t}<(s+1/2)\cdot c(n,K_3)$, we may assume that $N_3(G)\le(s+1/2)\cdot c(n,K_3)$. So, by Lemma 7.2.19, there exists a partition $V(G)=V_1\cup V_2$ such that $B:=G-G[V_1,V_2]$ is a matching of size s.

Let $x=|V_1|$ and $y=|V_2|$ and note that $x+y=n$. Without loss of generality we may assume that $x\ge y$. Let $H=G[V_1,V_2]$, $M=K[V_1,V_2]-H$, and $m=|M|$. Since $G-B=H=K[V_1,V_2]-M$, we obtain $t_2(n)+t-s=xy-m=(n-y)y-m$. Therefore, $m\in M_{s,t}$ and

$$y=\frac{1}{2}\left(n-\left(4s-4t-4m+n^2-4t_2(n)\right)^{1/2}\right).$$

Let $s_i=|B\cap\binom{V_i}{2}|$ for $i=1,2$ and note that $s_1+s_2=s$. It is easy to see that the number of potential copies of K_3 is s_1y+s_2x. We will consider two cases: either $s_i=s$ for some $i\in\{1,2\}$ or $s_1\ge 1$ and $s_2\ge 1$.

Case 1: $s_i = s$ for some $i \in \{1,2\}$.

We may assume that $s_2 = 0$ and the case $s_1 = 0$ can be solved using a similar argument. Notice that for every $e \in M$ there is at most one potential copy of K_3 containing e. Therefore,

$$N_3(G) \geq sy - m = \frac{sn}{2} - \frac{s}{2}\left(4s - 4t - 4m + n^2 - 4t_2(n)\right)^{1/2} - m =: f(m).$$

Then

$$\frac{\mathrm{d}f(m)}{\mathrm{d}m} = \frac{s}{\left(4s - 4t - 4m + n^2 - 4t_2(n)\right)^{1/2}} - 1.$$

First let us assume that $s \geq 3$. Then

$$s^2 \geq 4s - 4 + 1 \geq 4s - 4t - 4m + n^2 - 4t_2(n).$$

Therefore, $\frac{\mathrm{d}f(m)}{\mathrm{d}m} > 0$ for all $m > 0$, which implies that $f(m)$ is increasing in m. Therefore, for $s \geq 3$

$$N_3(G) \geq \frac{sn}{2} - \frac{s}{2}\left(4s - 4t - 4m_{s,t} + n^2 - 4t_2(n)\right)^{1/2} - m_{s,t} = s \cdot n^-_{s,t} - m_{s,t}.$$

For the case $s = 2$, one could easily check that the minimum of $f(m)$ is uniquely attained at $m = m_{s,t}$. Therefore, if $s_i = s$ for some $i \in \{1,2\}$, then $N_3(G) \geq s \cdot n^-_{s,t} - m_{s,t}$ for all $s > t \geq 1$.

If $N_3(G) = s\cdot n^-_{s,t} - m_{s,t}$, then the argument above shows that we must have $|V_1| = n - n^-_{s,t} = n^+_{s,t}$ and $|V_2| = n^-_{s,t}$, all edges in B are contained in V_1, all pairs in M must be contained in one potential copy of K_3, and no two pairs in the same potential copy. Therefore, $G \cong BM_{s,t}(n)$.

Case 2: $s_1 \geq 1$ and $s_2 \geq 1$.

Notice that for every $e \in M$ there are at most two potential copies of K_3 containing e. Since $x \geq y$, this gives

$$\begin{aligned} N_3(G) \geq s_1 y + s_2 x - 2m &\geq (s-1)y + x - 2m \\ &= (s-2)y + n - 2m \\ &= \frac{sn}{2} - \frac{s-2}{2}\left(4s - 4t - 4m + n^2 - 4t_2(n)\right)^{1/2} - 2m =: g(m). \end{aligned}$$

Let us first assume that $s \geq 20$. Since

$$\frac{\mathrm{d}g(m)}{\mathrm{d}m} = \frac{s-2}{\left(4s - 4t - 4m + n^2 - 4t_2(n)\right)^{1/2}} - 2.$$

and

$$s - 2 > 2\left(4s - 4 + 1\right)^{1/2} \geq 2\left(4s - 4t - 4m + n^2 - 4t_2(n)\right)^{1/2},$$

$\frac{\mathrm{d}g(m)}{\mathrm{d}m} > 0$ for $m > 0$. Therefore, by Lemma 7.2.20,

$$N_3(G) \geq g(m_{s,t}) = f(m_{s,t}) + \left(4s - 4t - 4m_{s,t} + n^2 - 4t_2(n)\right)^{1/2} - m_{s,t} \geq f(m_{s,t}),$$

and equality holds iff for some $p \in \mathbb{N}$

$$s - t = \begin{cases} p^2 - 1, & \text{if } n \equiv 0 \mod 2, \\ p(p+1) - 1, & \text{if } n \equiv 1 \mod 2. \end{cases} \qquad (\star)$$

For $s \leq 19$ a computer-aided calculation shows that $f(m_{s,t}) \leq \min_m\{g(m)\}$ always holds[1]. Moreover, the minimum of $g(m)$ is uniquely achieved at $m = m_{s,t}$ except for when $(s,t) \in \{(2,1),(3,1),(4,1)\}$ and n even, or $(s,t) \in \{(3,2),(4,1),(5,1),(6,1)\}$ and n odd.

If $N_3(G) = s \cdot n^-_{s,t} - m_{s,t}$, then the argument above shows that $(\star)$ holds, $|V_1| = n - n^-_{s,t} = n^+_{s,t}$ and $|V_2| = n^-_{s,t}$, exactly one edge $e \in B$ is contained in V_2, all other edges in B are contained in V_1, and all pairs in M must be contained in two potential copies of K_3. Therefore, $G \cong BS_{s,t}(n)$.

For $(s,t) \in \{(2,1),(3,1),(4,1)\}$ and n even, or $(s,t) \in \{(3,2),(4,1),(5,1),(6,1)\}$ and n odd, our bound $s \cdot n^-_{s,t} - m_{s,t}$ in Theorem 7.2.4 is also tight, but there are more constructions that achieve this bound. One could easily recover all these constructions using our calculation file. ∎

7.2.2.3 Proof of Theorem 7.2.5

In this section we prove Theorem 7.2.5. Recall that for $n,k \in \mathbb{N}$, $q_{n,k} = \lfloor n/(k-1) \rfloor$ and $r_{n,k} = n - (k-1)q_{n,k}$.

Proof of Theorem 7.2.5. Let $s \geq 2, k \geq 4$ be fixed integers and n be sufficiently large. Let $q = q_{n,k}$ and $r = r_{n,k}$. Let G be a graph on n vertices with $t_{k-1}(n) + s - 1$ edges and

---

[1] A simple Mathematica worksheet verifying this fact can be found at the web page `http://homepages.math.uic.edu/~mubayi/papers/ErdosRademacher.pdf`.

$\tau_k(G) = s$. Notice that $N_k(n,s) = (1+o(1))s\left(\frac{n}{k-1}\right)^{k-2}$ while $c(n,K_k) = (1+o(1))\left(\frac{n}{k-1}\right)^{k-2}$, so $N_k(n,s) < (s+1/2)\cdot c(n,K_k)$. Therefore, we may assume that $N_k(G) \le (s+1/2)\cdot c(n,K_k)$. So by Lemma 7.2.19, there exists a partition $V(G) = V_1\cup\cdots\cup V_{k-1}$ such that $B := G - G[V_1,\ldots,V_{k-1}]$ is a matching of size s.

Let $x_i = |V_i|$ for $i \in [k-1]$ and without loss of generality we may assume that $x_1 \ge \cdots \ge x_{k-1}$. Let $H = G[V_1,\ldots,V_{k-1}]$, $M = K[V_1,\ldots,V_{k-1}] - H$, and $m = |M|$. Since $t_{k-1}(n) - 1 = |H| = |K[V_1,\ldots,V_{k-1}]| - m$, we obtain $m \in \{0,1\}$ and

$$\sum_{1\le i<j\le k-1} x_i x_j = t_{k-1}(n) - 1 + m.$$

Suppose that $m = 1$. Then $\sum_{1\le i<j\le k-1} x_i x_j = t_{k-1}(n)$, so $x_1 = \cdots = x_r = q+1$ and $x_{r+1} = \cdots = x_{k-1} = q$.

Let $s_i = |B \cap \binom{V_i}{2}|$ for $i \in [k-1]$ and $S = \{i \in [k-1] : s_i \ge 1\}$.

Case 1: $|S| = 1$.

Let $i_0 \in [k-1]$ such that $s_{i_0} = s$. Then there are $s\cdot\prod_{i\ne i_0} x_i$ potential copies of K_k. Let $uv \in M$. If uv has empty intersection with all edges in B, then there are at most $s\cdot n^{k-4} = o(n^{k-3})$ potential copies of K_k containing uv. If uv has nonempty intersection with some $e \in B$, then every potential copy of K_k that contains uv must contain e as well. So in this case there are at most $\left(\prod_{i\notin\{i_0\}} x_i\right)/x_{k-1}$ potential copies of K_k containing uv. Therefore,

$$\begin{aligned}
N_k(G) &\geq s \cdot \prod_{i\notin\{i_0\}} x_i - \frac{1}{x_{k-1}} \prod_{i\notin\{i_0\}} x_i \\
&\geq \left(s - \frac{1}{x_{k-1}}\right) \prod_{i=2}^{k-1} x_i = \begin{cases} \left(s - \frac{1}{q}\right) q^{k-2}, & \text{if } r \leq 1, \\ \left(s - \frac{1}{q}\right) (q+1)^{r-1} q^{k-r-1}, & \text{if } 2 \leq r \leq k-2, \end{cases} \\
&\geq N_k(n, s),
\end{aligned}$$

and equality holds only if $r = 1$.

Case 2: $|S| \geq 2$.

The number of potential copies of K_k is $\sum_{i=1}^{k-1}\left(s_i \cdot \prod_{j\neq i} x_j\right)$. Suppose that the pair in M has nonempty intersection with V_{i_0} and V_{i_1} for some $i_0, i_1 \in [k-1]$. If $s_{i_0} = 0$, then there are at most $\left(\prod_{i\neq i_0} x_i\right)/x_{k-1}$ potential copies of K_k containing the pair in M. If both $s_{i_0} \geq 1$ and $s_{i_1} \geq 1$, then there are most $2\prod_{i\neq i_0,i_1} x_i$ potential copies of K_k containing the pair in M. Therefore,

$$\begin{aligned}
N_k(G) \geq \sum_{i=1}^{k-1}\left(s_i \cdot \prod_{j\neq i} x_j\right) - 2\prod_{i\neq i_0,i_1} x_i &= \left(\sum_{i=1}^{k-1} \frac{s_i}{x_i} - \frac{2}{x_{i_0}x_{i_1}}\right) \prod_{j=1}^{k-1} x_j \\
&\geq \left(\frac{s-2}{x_1} + \frac{1}{x_{i_0}} + \frac{1}{x_{i_1}} - \frac{2}{x_{i_0}x_{i_1}}\right) \prod_{j=1}^{k-1} x_j
\end{aligned}$$

Since

$$\frac{1}{x_{i_0}} + \frac{1}{x_{i_1}} - \frac{2}{x_{i_0}x_{i_1}} = \frac{1}{2} - 2\left(\frac{1}{2} - \frac{1}{x_{i_0}}\right)\left(\frac{1}{2} - \frac{1}{x_{i_1}}\right)$$

is decreasing in x_{i_0} and x_{i_1},

$$\frac{1}{x_{i_0}} + \frac{1}{x_{i_1}} - \frac{2}{x_{i_0}x_{i_1}} \geq \frac{1}{x_1} + \frac{1}{x_2} - \frac{2}{x_1x_2}.$$

Therefore,

$$N_k(G) \geq \left(\frac{s-1}{x_1} + \frac{1}{x_2} - \frac{2}{x_1x_2}\right)\prod_{j=1}^{k-1} x_j = \begin{cases} \left(s - \frac{2}{q}\right) q^{k-2}, & \text{if } r = 0, \\ \left(s - \frac{1}{q}\right) q^{k-2}, & \text{if } r = 1, \\ \left(s - \frac{2}{q+1}\right)(q+1)^{r-1}q^{k-r-1}, & \text{if } 2 \leq r \leq k-2. \end{cases}$$

$$\geq N_k(n, s).$$

Note that if $s \geq 3$, then the first inequality above is strict since there are copies of K_k in G containing at least two edges in B.

Now we may assume that $m = 0$. Then every $e \in B$ is contained in at least $\prod_{i=2}^{k-1} x_i$ copies of K_k and hence

$$N_k(G) \geq s \cdot \prod_{i=2}^{k-1} x_i.$$

So we just need to find the minimum of $\prod_{i=2}^{k-1} x_i$ subject to the constraint that $\prod_{i=1}^{k-1} x_i = t_{k-1}(n) - 1$.

If $r = 0$, then $x_1 = q+1, x_2 = \cdots = x_{k-2} = q$, and $x_{k-1} = q-1$. Therefore, $\prod_{i=2}^{k-1} x_i = q^{k-3}(q-1)$.

If $r = 1$, then $x_1 = x_2 = q+1, x_3 = \cdots = x_{k-2} = q$, and $x_{k-1} = q-1$. Therefore, $\prod_{i=2}^{k-1} x_i = q^{k-4}(q+1)(q-1)$.

If $r \geq 2$, then

$$\begin{aligned} \text{either} \quad & x_1 = \cdots = x_{r+1} = q+1, x_{r+2} = \cdots = x_{k-2} = q, x_{k-1} = q-1 \\ \text{or} \quad & x_1 = q+2, x_2 = \cdots = x_{r-1} = q+1, x_r = \cdots = x_{k-1} = q. \end{aligned}$$

The later one gives a smaller $\prod_{i=2}^{k-1} x_i$, which is $(q+1)^{r-2}p^{k-r}$.

Therefore, for the case $m = 0$

$$\begin{aligned} N_k(G) \geq & \begin{cases} s \cdot q^{k-3}(q-1), & \text{if } r = 0, \\ s \cdot q^{k-4}(q+1)(q-1), & \text{if } r = 1, \\ s \cdot (q+1)^{r-2}q^{k-r}, & \text{if } 2 \leq r \leq k-2. \end{cases} \\ \geq & N_k(n,s), \end{aligned}$$

and equality only if $r \neq 1$. ∎

7.2.2.4 Proof of Theorem 7.2.6

In this section we prove Theorem 7.2.6. Recall that for $n, k \in \mathbb{N}$, $q_{n,k} = \lfloor n/(k-1) \rfloor$ and $r_{n,k} = n - (k-1)q_{n,k}$. For $s > t \geq 1, k \geq 3$,

$$R_k(n,s,t) = \left(\frac{2(k-1)(s-t)}{k-2} + \frac{(k-1-r_{n,k})\, r_{n,k}}{k-2} \right)^{1/2},$$

$n^+_{k,s,t} = \frac{n+(k-2)R_k(n,s,t)}{k-1}$, and $n^-_{k,s,t} = \frac{n-R_k(n,s,t)}{k-1}$.

Proof of Theorem 7.2.6. Let $k \geq 4$, $s > t \geq 2$ be fixed integers and n be sufficiently large. Suppose that $s > 2R_k(n,s,t)$. Let $q = q_{n,k}$, $r = r_{n,k}$, and $R = R_k(n,s,t)$. Let G be a graph on n vertices with $t_{k-1}(n) + t$ edges and $\tau_k(G) = s$. Since $s \cdot \left(n^-_{k,s,t} \right) < (s + 1/2) \cdot c(n, K_k)$, we may assume that $N_k(G) \leq (s+1/2) \cdot c(n, K_k)$. So by Lemma 7.2.19, there exists a partition $V(G) = V_1 \cup \cdots \cup V_{k-1}$ such that $B := G - G[V_1, \ldots, V_{k-1}]$ is a matching of size s.

Let $x_i = |V_i|$ for $i \in [k-1]$ and without loss of generality we may assume that $x_1 \geq \cdots \geq x_{k-1}$. Let $H = G[V_1, \ldots, V_{k-1}]$, $M = K[V_1, \ldots, V_{k-1}] - H$, and $m = |M|$. Since $t_{k-1}(n) + t - s = |H| = |K[V_1, \ldots, V_{k-1}]| - m$,

$$\sum_{1 \leq i < j \leq k-1} x_i x_j = t_{k-1}(n) + t - s + m,$$

which is equivalent to

$$\sum_{i=1}^{k-1} x_i^2 = n^2 - 2t_{k-1}(n) + 2s - 2t - 2m.$$

Let $s_i = |B \cap \binom{V_i}{2}|$ for $i \in [k-1]$ and $S = \{i \in [k-1] : s_i \geq 1\}$.

Case 1: $|S| = 1$.

Without loss of generality we may assume that $s_1 = s$ since the other cases can be solved using a similar argument. Notice that there are $s \cdot \prod_{i=2}^{k-1} x_i$ potential copies of K_k, and for every $e \in M$ there are at most $\prod_{i=2}^{k-2} x_i$ potential copies of K_k containing e. Therefore,

$$N_k(G) \geq s \cdot \prod_{i=2}^{k-1} x_i - m \cdot \prod_{i=2}^{k-2} x_i = \left(s - \frac{m}{x_{k-1}}\right) \cdot \prod_{i=2}^{k-1} x_i.$$

Fix $0 \leq m \leq s - t$. Let $\mathbb{R}_{\geq 0}$ be the collection of all nonnegative real numbers. Define

$$C_m(\mathbb{N}) = \left\{(x_1, \ldots, x_{k-1}) \in \mathbb{N}^{k-1} \colon \sum_{i=1}^{k-1} x_i = n, \sum_{i=1}^{k-1} x_i^2 = n^2 - 2t_{k-1}(n) + 2s - 2t - 2m\right\},$$

and

$$C_m(\mathbb{R}) = \left\{(x_1, \ldots, x_{k-1}) \in \mathbb{R}_{\geq 0}^{k-1} \colon \sum_{i=1}^{k-1} x_i = n, \sum_{i=1}^{k-1} x_i^2 = n^2 - 2t_{k-1}(n) + 2s - 2t - 2m\right\}.$$

Note that $C_m(\mathbb{N}) \subset C_m(\mathbb{R})$. In order to get a lower bound for $N_k(G)$ we need to solve the following optimization problem.

$$\mathrm{OPT}_m^{\mathrm{A}} : \begin{cases} \text{Minimize} & \left(s - \frac{m}{x_{k-1}}\right) \cdot \prod_{i=2}^{k-1} x_i \\ \text{subject to} & (x_1, \ldots, x_{k-1}) \in C_m(\mathbb{N}). \end{cases}$$

However, it is not easy to get an optimal solution for $\mathrm{OPT}_m^{\mathrm{A}}$. So we are going to consider the following two auxiliary optimization problems. Let

$$\mathrm{OPT}_m^{\mathrm{B}}: \begin{cases} \text{Minimize} & \left(s - \frac{m}{x_{k-1}}\right) \cdot \prod_{i=2}^{k-1} x_i \\ \text{subject to} & (x_1, \ldots, x_{k-1}) \in C_m(\mathbb{R}), \end{cases}$$

and

$$\mathrm{OPT}_m^{\mathrm{C}}: \begin{cases} \text{Minimize} & \prod_{i=2}^{k-1} x_i \\ \text{subject to} & (x_1, \ldots, x_{k-1}) \in C_m(\mathbb{R}). \end{cases}$$

Let $\mathrm{opt}_m^{\mathrm{a}}$, $\mathrm{opt}_m^{\mathrm{b}}$, and $\mathrm{opt}_m^{\mathrm{c}}$ denote the optimal value of the optimization problems $\mathrm{OPT}_m^{\mathrm{A}}$, $\mathrm{OPT}_m^{\mathrm{B}}$, $\mathrm{OPT}_m^{\mathrm{C}}$, respectively. It is easy to see that $\mathrm{opt}_m^{\mathrm{a}} \geq \mathrm{opt}_m^{\mathrm{b}}$. Moreover, if $\mathrm{OPT}_m^{\mathrm{B}}$ has an optimal solution $x_1, \ldots, x_{k-1}$ such that $x_i \in \mathbb{N}$, then $\mathrm{opt}_m^{\mathrm{a}} = \mathrm{opt}_m^{\mathrm{b}}$. Our goal is to find $\mathrm{opt}_m^{\mathrm{b}}$ and it will be a lower bound for $N_k(G)$.

Claim 7.2.21. *There exists a constant $C' > 0$ such that*

$$\left(s - \frac{k-1}{n} m\right) \cdot \mathrm{opt}_m^{\mathrm{c}} - C' n^{k-4} < \mathrm{opt}_{\mathrm{m}}^{\mathrm{b}} \leq \left(s - \frac{k-1}{n} m\right) \cdot \mathrm{opt}_m^{\mathrm{c}} + C' n^{k-4}.$$

Proof. We abuse notation by assuming that $x_1, \ldots, x_{k-1}$ is an optimal solution of $\mathrm{OPT}^{\mathrm{B}}_m$. Since $\sum_{1\le i<j\le k-1} x_i x_j = t_{k-1}(n) + t - s + m > t_{k-1}(n) - s$, by Lemma 7.2.13, $n/(k-1) - s \le x_i \le n/(k-1) + s$ for all $i \in [k-1]$. Therefore,

$$\begin{aligned}
\mathrm{opt}^{\mathrm{b}}_m = \left(s - \frac{m}{x_{k-1}}\right) \cdot \prod_{i=2}^{k-1} x_i &\ge \left(s - \frac{m}{n/(k-1) - s}\right) \cdot \prod_{i=2}^{k-1} x_i \\
&= \left(s - \frac{(k-1)m}{n}\right) \cdot \prod_{i=2}^{k-1} x_i - \frac{(k-1)^2 sm}{n(n-ks+s)} \cdot \prod_{i=2}^{k-1} x_i \\
&> \left(s - \frac{(k-1)m}{n}\right) \cdot \prod_{i=2}^{k-1} x_i - C' n^{k-4} \\
&\ge \left(s - \frac{(k-1)m}{n}\right) \cdot \mathrm{opt}^{\mathrm{c}}_m - C' n^{k-4},
\end{aligned}$$

where C' is a constant depending only on k, s, m.

Now let $x'_1, \ldots, x'_{k-1}$ be an optimal solution of $\mathrm{OPT}^{\mathrm{C}}_m$. Then similarly we have

$$\begin{aligned}
\left(s - \frac{(k-1)m}{n}\right) \cdot \mathrm{opt}^{\mathrm{c}}_m &= \left(s - \frac{m}{n/(k-1)}\right) \cdot \prod_{i=2}^{k-1} x'_i \\
&\ge \left(s - \frac{m}{x'_{k-1} - s}\right) \cdot \prod_{i=2}^{k-1} x'_i \\
&= \left(s - \frac{m}{x'_{k-1}}\right) \cdot \prod_{i=2}^{k-1} x'_i - \frac{sm}{x'_{k-1}(x'_{k-1}+s)} \prod_{i=2}^{k-1} x'_i \ge \mathrm{opt}^{\mathrm{b}}_m - C' n^{k-4}.
\end{aligned}$$

∎

Claim 7.2.21 shows that $\mathrm{opt}^{\mathrm{b}}_m = \left(s - \frac{k-1}{n} m\right) \cdot \mathrm{opt}^{\mathrm{c}}_m \pm C' n^{k-4}$. So we could view $\left(s - \frac{k-1}{n} m\right) \cdot \mathrm{opt}^{\mathrm{c}}_m$ as a "trajectory" (in other words, the "expected" value) for $\mathrm{opt}^{\mathrm{b}}_m$, and this will be useful later for us to show that $\mathrm{opt}^{\mathrm{b}}_{m-1} \le \mathrm{opt}^{\mathrm{b}}_m$.

Let us solve the optimization problem $\mathrm{OPT}_m^{\mathrm{C}}$ first. We use the Lagrangian multiplier method. Let

$$\mathcal{L}(\vec{x}, \lambda, \mu) = \prod_{i=2}^{k-1} x_i + \lambda \left(\sum_{i=1}^{k-1} x_i - n \right) + \mu \left(\sum_{i=1}^{k-1} x_i^2 - \big(n^2 - 2t_{k-1}(n) + 2s - 2t - 2m\big) \right).$$

Again, we abuse notation here by assuming that $(x_1, \ldots, x_{k-1}) \in C_m(\mathbb{R})$ is an optimal solution of $\mathrm{OPT}_m^{\mathrm{C}}$. Then by the Lagrangian multiplier method,

$$\begin{cases} \frac{\partial \mathcal{L}}{\partial x_1} & = \lambda + 2\mu x_1 = 0 \Rightarrow x_1 = -\frac{\lambda}{2\mu}, \\ \frac{\partial \mathcal{L}}{\partial x_j} & = \frac{\prod_{i=2}^{k-1} x_i}{x_j} + \lambda + 2\mu x_j = 0, \\ \frac{\partial \mathcal{L}}{\partial \lambda} & = \sum_{i=1}^{k-1} x_i - n = 0, \\ \frac{\partial \mathcal{L}}{\partial \mu} & = \sum_{i=1}^{k-1} x_i^2 - \big(n^2 - 2t_{k-1}(n) + 2s - 2t - 2m\big) = 0. \end{cases}$$

Note that $x_1 \neq 0$ so it is an interior point and hence we can apply the Lagrangian multiplier here.

Let $\pi = \prod_{i=2}^{k-1} x_i$. Note that the equation

$$\frac{\pi}{x} + \lambda + 2\mu x = 0$$

has only two solutions

$$x' = \frac{-\lambda + \sqrt{\lambda^2 - 8\mu\pi}}{4\mu} \quad \text{and} \quad x'' = \frac{-\lambda - \sqrt{\lambda^2 - 8\mu\pi}}{4\mu}.$$

Therefore, referring to $\partial\mathcal{L}/\partial x_j$, for every $2 \le j \le k-1$ either $x_j = x'$ or $x_j = x''$.

Before we state the next claim let us recall from the beginning of Case 1 that $s_1 = s$.

Claim 7.2.22. $x_1 \ge x_2 = \cdots = x_{k-1}$.

Proof. First we show that $x_1 \ge x_i$ for all $2 \le i \le k-1$. Suppose to the contrary that there exists some $i \in [k-1] \setminus \{1\}$ such that $x_i > x_1$, and without loss of generality we may assume that $x_2 > x_1$. Then let $x_i' = x_i$ for $3 \le i \le k-1$, $x_1' = x_2$, and $x_2' = x_1$. It is clear that $(x_1', \ldots, x_{k-1}') \in C_m(\mathbb{R})$, but $\prod_{i=2}^{k-1} x_i' < \prod_{i=2}^{k-1} x_i$, which contradicts our assumption that $(x_1, \ldots, x_{k-1})$ is an optimal solution of $\mathrm{OPT}_m^{\mathrm{C}}$. Therefore, $x_1 \ge x_i$ for all $2 \le i \le k-1$.

Now we show that $x_2 = \cdots = x_{k-1}$. Suppose that $x_{i_1} \ne x_{i_2}$ for some $2 \le i_1 < i_2 \le k-1$. Then $\{x_{i_1}, x_{i_2}\} = \{x', x''\}$, which implies that $x_{i_1} + x_{i_2} = -\lambda/(2\mu) = x_1$. Since $\sum_{1 \le i < j \le k-1} x_i x_j = t_{k-1}(n) + t - s + m > t_{k-1}(n) - s$, by Lemma 7.2.13, $|x_i - n/(k-1)| < s$ for all $i \in [k-1]$. Therefore,

$$x_{i_1} + x_{i_2} > 2 \times \frac{n}{k-1} - 2s > \frac{n}{k-1} + s > x_1,$$

a contradiction. Therefore, $x_2 = \cdots = x_{k-1}$. ∎

By Lemma 7.2.11,

$$n^2 - 2t_{k-1}(n) = \frac{n^2}{k-1} + \frac{(k-1-r)r}{k-1}.$$

Let $x = x_1$, $y = x_2 = \cdots = x_{k-1}$. Since $(x_1, \ldots, x_{k-1}) \in C_m(\mathbb{R})$,

$$\begin{cases} x + (k-2)y = n, \\ x^2 + (k-2)y^2 = \frac{n^2}{k-1} + \frac{(k-1-r)r}{k-1} + 2s - 2t - 2m, \\ x_i \geq 0, \forall i \in [k-1], \end{cases}$$

which implies

$$\begin{cases} x = \frac{n}{k-1} + (k-2)\Delta_m \\ y = \frac{n}{k-1} - \Delta_m, \end{cases}$$

where

$$\Delta_m := \frac{(2(k-1)(k-2)(s-t-m) + (k-2)(k-1-r)r)^{1/2}}{(k-1)(k-2)}.$$

Therefore,

$$\mathrm{opt}^{\mathrm{c}}_m = y^{k-2} = \left(\frac{n}{k-1} - \Delta_m\right)^{k-2}.$$

Now we are going to use $\mathrm{opt}^{\mathrm{c}}_m$ to describe the behavior of $\mathrm{opt}^{\mathrm{b}}_m$.

Claim 7.2.23. *The value* $\mathrm{opt}^{\mathrm{b}}_m$ *is strictly increasing in* m*. In particular,* $\mathrm{opt}^{\mathrm{b}}_0 < \mathrm{opt}^{\mathrm{b}}_m$ *for all* $m > 0$.

Proof. Since

$$\mathrm{opt}^{\mathrm{c}}_m = \left(\frac{n}{k-1} - \Delta_m\right)^{k-2} = \left(\frac{n}{k-1}\right)^{k-2} - (k-2)\Delta_m \left(\frac{n}{k-1}\right)^{k-3} + \Theta(n^{k-4}),$$

by Claim 7.2.21, there exists a constant $C' > 0$ such that

$$\begin{aligned}\mathrm{opt}^{\mathrm{b}}_m &= \left(s - \frac{k-1}{n}m\right) \cdot \mathrm{opt}^{\mathrm{c}}_m \pm C'n^{k-4} \\ &= s\left(\frac{n}{k-1}\right)^{k-2} - (m + s(k-2)\Delta_m)\left(\frac{n}{k-1}\right)^{k-3} \pm C''n^{k-4},\end{aligned}$$

where $C'' > C'$ is a constant depending only on s, k, m. Therefore,

$$\mathrm{opt}^{\mathrm{b}}_{m-1} - \mathrm{opt}^{\mathrm{b}}_m = (1 - s(k-2)\left(\Delta_{m-1} - \Delta_m\right))\left(\frac{n}{k-1}\right)^{k-3} \pm 2C''n^{k-4}.$$

Now view Δ_m as a function of the variable m. Then it is easy to see that Δ_m is concave down, i.e. $\mathrm{d}^2\Delta_m/\mathrm{d}m^2 < 0$ for $0 \le m \le s-t$. Therefore,

$$\begin{aligned}s(k-2)\left(\Delta_{m-1} - \Delta_m\right) &\ge s(k-2)(-1) \cdot \left.\frac{\mathrm{d}\Delta_m}{\mathrm{d}m}\right|_{m=0} \\ &= \frac{s(k-2)}{(2(k-1)(k-2)(s-t) + (k-2)(k-1-r)r)^{1/2}}.\end{aligned}$$

Since

$$s > 2R = 2\frac{(2(k-1)(k-2)(s-t) + (k-2)\left(k-1-r\right)r)^{1/2}}{k-2},$$

we obtain $s(k-2)\left(\Delta_{m-1}-\Delta_m\right)>2$. Therefore,

$$1-s(k-2)\left(\Delta_{m-1}-\Delta_m\right)<-1, \tag{$*$}$$

and hence $\mathrm{opt}^{\mathrm{b}}_{m-1}-\mathrm{opt}^{\mathrm{b}}_m<-\left(n/(k-1)\right)^{k-3}+\Theta(n^{k-4})<0$. ∎

Therefore,

$$N_k(G)\geq \mathrm{opt}^{\mathrm{a}}_m\geq \mathrm{opt}^{\mathrm{b}}_m\geq \mathrm{opt}^{\mathrm{b}}_0=s\cdot \mathrm{opt}^{\mathrm{c}}_0=s\cdot\left(\frac{n}{k-1}-\Delta_0\right)^{k-2}=s\cdot\left(n^{-}_{k,s,t}\right)^{k-2}.$$

Here we used that fact that $\Delta_0=R/(k-1)$.

Case 2: $|S|\geq 2$.

The number of potential copies of K_k is $\sum_{i=1}^{k-1}\left(s_i\cdot\prod_{j\neq i}x_j\right)$. Suppose that $uv\in M$ satisfies $u\in V_{i_0}$ and $v\in V_{i_1}$ for some $i_0,i_1\in[k-1]$. Similar to the proof of Theorem 7.2.6 we may assume that $s_{i_0}\geq 1$ and $s_{i_1}\geq 1$. Then there are at most $\prod_{i\neq i_0,i_1}x_i$ potential copies of K_k containing uv. Therefore,

$$N_k(G)\geq\sum_{i=1}^{k-1}\left(s_i\cdot\prod_{j\neq i}x_j\right)-2\sum_{uv\in M}\prod_{\substack{i\neq i_0,i_1\\ u\in V_{i_0},v\in V_{i_1}}}x_i=\left(\sum_{i=1}^{k-1}\frac{s_i}{x_i}-\sum_{\substack{uv\in M\\ u\in V_{i_0},v\in V_{i_1}}}\frac{2}{x_{i_0}x_{i_1}}\right)\prod_{i=1}^{k-1}x_i.$$

We abuse notation by assuming that $x_{i_0}x_{i_1} = \min\{x_ix_j : \exists uv \in M \text{ such that } u \in V_i, v \in V_j\}$.

Then

$$N_k(G) \geq \left(\sum_{i=1}^{k-1} \frac{s_i}{x_i} - \frac{2m}{x_{i_0}x_{i_1}}\right)\prod_{i=1}^{k-1} x_i = \left(\frac{s-2}{x_1} + \frac{1}{x_{i_0}} + \frac{1}{x_{i_1}} - \frac{2m}{x_{i_0}x_{i_1}}\right)\prod_{i=1}^{k-1} x_i.$$

Since

$$\frac{1}{x_{i_0}} + \frac{1}{x_{i_1}} - \frac{2m}{x_{i_0}x_{i_1}} = \frac{1}{2m} - 2m\left(\frac{1}{2m} - \frac{1}{x_{i_0}}\right)\left(\frac{1}{2m} - \frac{1}{x_{i_1}}\right),$$

is decreasing in x_{i_0} and x_{i_1},

$$\frac{1}{x_{i_0}} + \frac{1}{x_{i_1}} - \frac{2m}{x_{i_0}x_{i_1}} \geq \frac{1}{x_1} + \frac{1}{x_2} - \frac{2m}{x_1x_2}.$$

Therefore,

$$\begin{aligned} N_k(G) &\geq \left(\frac{s-1}{x_1} + \frac{1}{x_2} - \frac{2m}{x_1x_2}\right)\prod_{i=1}^{k-1} x_i \\ &= \left(\frac{s}{x_1} + \frac{x_1 - x_2}{x_1x_2} - \frac{2m}{x_1x_2}\right)\prod_{i=1}^{k-1} x_i = \left(s + \frac{x_1 - x_2}{x_2} - \frac{2m}{x_2}\right)\prod_{i=2}^{k-1} x_i. \end{aligned}$$

Therefore, in order to get a lower bound for $N_k(G)$ we need solve the following optimization problem.

$$\mathrm{OPT}_m^{\mathrm{D}}: \begin{cases} \text{Minimize} & \left(s + \frac{x_1 - x_2}{x_2} - \frac{2m}{x_2}\right) \prod_{i=2}^{k-1} x_i \\ \text{subject to} & (x_1, \ldots, x_{k-1}) \in C_m(\mathbb{N}). \end{cases}$$

Similarly, we are going to consider the following auxiliary optimization problem.

$$\mathrm{OPT}_m^{\mathrm{E}}: \begin{cases} \text{Minimize} & \left(s + \frac{x_1 - x_2}{x_2} - \frac{2m}{x_2}\right) \prod_{i=2}^{k-1} x_i \\ \text{subject to} & (x_1, \ldots, x_{k-1}) \in C_m(\mathbb{R}). \end{cases}$$

Theoretically, one could solve $\mathrm{OPT}_m^{\mathrm{E}}$ exactly using the Lagrange multiplier method. However, the optimal solution of $\mathrm{OPT}_m^{\mathrm{E}}$ is very complicated. So we are going to compare $\mathrm{OPT}_m^{\mathrm{E}}$ with $\mathrm{OPT}_m^{\mathrm{C}}$.

Let $\mathrm{opt}_m^{\mathrm{d}}$ and $\mathrm{opt}_m^{\mathrm{e}}$ denote the optimal values of the optimization problems $\mathrm{OPT}_m^{\mathrm{D}}$ and $\mathrm{OPT}_m^{\mathrm{E}}$, respectively. It is easy to see that $\mathrm{opt}_m^{\mathrm{d}} \geq \mathrm{opt}_m^{\mathrm{e}}$. The following claim is very similar to Claim 7.2.21, and can be proved in a similar fashion so so we omit the proof.

Claim 7.2.24. *There exists a constant $\hat{C} > 0$ such that*

$$\left(s - \frac{2(k-1)m}{n}\right) \cdot \mathrm{opt}_m^{\mathrm{c}} - \hat{C}n^{k-4} < \mathrm{opt}_m^{\mathrm{e}} \leq \left(s - \frac{2(k-1)m}{n}\right) \cdot \mathrm{opt}_m^{\mathrm{c}} + \hat{C}n^{k-4}.$$

Claim 7.2.25. *The value* $\mathrm{opt}^{\mathrm{e}}_m$ *is strictly increasing in* m*. In particular,* $\mathrm{opt}^{\mathrm{e}}_0 < \mathrm{opt}^{\mathrm{e}}_m$ *for all* $m > 0$*.*

Proof. The proof is basically the same as the proof for Claim 7.2.23. The only difference is that $s > 2R$ implies that there exists $\varepsilon > 0$ such that $s(k-2)\left(\Delta_{m-1} - \Delta_m\right) > 2 + \varepsilon$. Therefore, $(*)$ now becomes

$$2 - s(k-2)\left(\Delta_{m-1} - \Delta_m\right) < -\varepsilon,$$

which implies that

$$\begin{aligned} \mathrm{opt}^{\mathrm{b}}_{m-1} - \mathrm{opt}^{\mathrm{b}}_m &= \left(2 - s(k-2)\left(\Delta_{m-1} - \Delta_m\right)\right)\left(\frac{n}{k-1}\right)^{k-3} \pm 2C'n^{k-4} \\ &< -\varepsilon\left(n/(k-1)\right)^{k-3} + \Theta(n^{k-4}) < 0. \end{aligned}$$

∎

Therefore, if $s > 2R$, then

$$N_k(G) \geq \mathrm{opt}^{\mathrm{d}}_m \geq \mathrm{opt}^{\mathrm{e}}_m \geq \mathrm{opt}^{\mathrm{e}}_0 \geq s \cdot \mathrm{opt}^{\mathrm{c}}_0 = s \cdot \left(\frac{n}{k-1} - \Delta_0\right)^{k-2} = s \cdot \left(n^{-}_{k,s,t}\right)^{k-2}.$$

Note that we may assume that $s - t \geq 2$ since the case $s - t = 1$ has been solved by Theorem 7.2.5. Therefore, there exists copies of K_k in G that contains at least two edges in B, which implies that the first inequality above is strict. ∎

7.2.2.5 Proof of Theorem 7.2.9

In this section we prove Theorem 7.2.9. We need the following lemma.

Lemma 7.2.26 ([194]). *Fix $k \geq 3$ and a k-critical graph F with f vertices. Then there are positive constants α_F and β_F such that if n is sufficiently large, then $|c(n, F) - \alpha_F n^{f-2}| < \beta_F n^{f-3}$.*

Now we are ready to prove Theorem 7.2.9.

Proof of Theorem 7.2.9. Let $s > t \geq 1, k \geq 3$ be fixed integers and let F be a k-critical graph on f vertices. Let n be sufficiently large. Let G be a graph on n vertices with $t_{k-1}(n) + t$ edges and $\tau_F(G) = s$. We may assume that $N_F(G) \leq s \cdot c(n, F)$, since otherwise we are done.

By Lemma 7.2.19, there exists a partition $V(G) = V_1 \cup \cdots \cup V_{k-1}$ such that $B := G - G[V_1, \ldots, V_{k-1}]$ is a matching of size s. Let $x_i = |V_i|$ for $i \in [k-1]$ and without loss of generality we may assume that $x_1 \geq \cdots \geq x_{k-1}$. Let $H = G[V_1, \ldots, V_{k-1}]$, $M = K[V_1, \ldots, V_{k-1}] - H$, and $m = |M|$. Since $t_{k-1}(n) - t = |H| = |K[V_1, \ldots, V_{k-1}]| - m$,

$$\sum_{1 \leq i < j \leq k-1} x_i x_j = t_{k-1}(n) + t - s + m.$$

Therefore, by Lemma 7.2.13, $n/(k-1) - s < x_i < n/(k-1) + s$ for all $i \in [k-1]$. Let

$$c_{min} = \min\{c(x_{\sigma(1)}, \ldots, x_{\sigma(k-1)}) \colon \sigma \in S_{k-1}\},$$

where S_{k-1} is the collection of all permutations of $[k-1]$. By Lemma 7.2.12, $c_{min} \geq c(n, F) - \gamma_F s n^{f-3}$ for some constant γ_F. Note the the number of potential copies of K_k is at least $s \cdot c_{min}$. Since every $e \in M$ is contained in at most n^{f-3} potential copies of K_k,

$$N_F(G) \geq s \cdot c_{min} - mn^{f-3} \geq s \cdot c(n, F) - Cn^{f-3}$$

for some constant C. This completes the proof of Theorem 7.2.9. ∎

7.2.3 Concluding remarks

We proved several bounds on the number of copies of K_k (and also for k-critical graphs F) in a graph G on n vertices with $t_{k-1}(n) + t$ edges and $\tau_k(G) = s$. In our proof we need s and t to be fixed. Using the same method we are able to show that the same conclusions as in Theorems 7.2.4, 7.2.5, 7.2.6, and 7.2.9 hold for all $s > t \geq 1$ (for Theorem 7.2.6 we still need $s > 2R_k(n, s, t)$) as long as $s(s-t)^{1/2} < \xi n$ for some small constant $\xi > 0$. In particular, if $s - t < C$ for some constant C, then the conclusions hold for all $s < \xi' n$ for some small constant $\xi' > 0$. The proofs are more involved and tedious, so we chose to omit them here.

CHAPTER 8

THE FEASIBLE REGION OF INDUCED GRAPHS

8.1 The feasible region of induced graphs

8.1.1 Introduction

8.1.1.1 Feasible regions

A quantum graph Q is a formal linear combination of finitely many graphs, i.e., an expression of the form

$$Q = \sum_{i=1}^{m} \lambda_i F_i \,,$$

where m is a nonnegative integer, the numbers $\lambda_1, \ldots, \lambda_m$ are real, and $F_1, \ldots, F_m$ are graphs. We call F_i a constituent of Q if $\lambda_i \neq 0$. Two quantum graphs Q, Q' are equal if they have the same constituents and the same (nonzero) coefficients for each constituent. The complement of Q is $\overline{Q} = \sum_{i=1}^{m} \lambda_i \overline{F}_i$, where $\overline{F}_i$ denotes the complement of F_i for each $i \in [m]$. A quantum graph Q is self-complementary if $Q = \overline{Q}$. Every graph parameter f can be extended linearly to quantum graphs by stipulating $f(Q) = \sum_{i=1}^{m} \lambda_i f(F_i)$. In particular,

$$N(Q, G) = \sum_{i=1}^{m} \lambda_i N(F_i, G) \quad \text{and} \quad \rho(Q, G) = \sum_{i=1}^{m} \lambda_i \rho(F_i, G) \,.$$

The main notion investigated in this article is the following.

Definition 8.1.1 (Feasible region). *Let $Q = \sum_{i=1}^{m} \lambda_i F_i$ be a quantum graph.*

- *A sequence $(G_n)_{n=1}^{\infty}$ of graphs is Q-*good *if $\lim_{n\to\infty} v(G_n) = \infty$, $\lim_{n\to\infty} \rho(G_n)$ exists, and for every $i \in [m]$ the limit $\lim_{n\to\infty} \rho(F_i, G_n)$ exists.*

- *A Q-good sequence of graphs* $(G_n)_{n=1}^{\infty}$ realizes *a point* $(x, y) \in [0,1] \times \mathbb{R}$ *if*

$$\lim_{n\to\infty} \rho(G_n) = x \quad \text{and} \quad \lim_{n\to\infty} \rho(Q, G_n) = y.$$

- *The* feasible region $\Omega_{\text{ind}}(Q)$ *of (induced) Q is the collection of points* $(x, y) \in [0,1] \times \mathbb{R}$ *realized by some Q-good sequence* $(G_n)_{n=1}^{\infty}$.

We commence a systematic study of the feasible region of quantum graphs Q. As we shall see soon, $\Omega_{\text{ind}}(Q)$ is determined by its boundary, so it suffices to consider for every $x \in [0,1]$ the numbers

$$i(Q, x) = \inf\{y \colon (x, y) \in \Omega_{\text{ind}}(Q)\} \qquad \text{and} \qquad I(Q, x) = \sup\{y \colon (x, y) \in \Omega_{\text{ind}}(Q)\}\,.$$

Determining the values of $i(Q, x)$ and $I(Q, x)$ under some constraints is a central topic in extremal combinatorics. For example, the classical Kruskal-Katona theorem [131; 154] implies

$$I(K_r, x) = x^{r/2} \quad \text{for all } r \geq 2 \text{ and } x \in [0,1]\,.$$

Turán's seminal theorem [241] and supersaturation show that for every integer $r \geq 3$,

$$i(K_r, x) > 0 \qquad \Longleftrightarrow \qquad x > (r-2)/(r-1)\,.$$

Determining $i(K_r, x)$ for all $x > (r-2)/(r-1)$ is highly nontrivial and was solved for $r = 3$ by Razborov [216], for $r = 4$ by Nikiforov [201], and for all r by Reiher [217].

Regarding quantum graphs with at least two constituents, a classical result of Goodman [115] says that $i(K_3 + \overline{K}_3, x) \geq 1/4$ and equality holds only for $x = 1/2$. Erdős [57] conjectured that $i(K_r + \overline{K}_r, x) \geq 2^{1-\binom{r}{2}}$ for $r \geq 4$ with equality for $x = 1/2$. This conjecture was disproved by Thomason [238] for all $r \geq 4$, but even for $r = 4$ the minimum value of $i(K_r + \overline{K}_r, x)$ is still unknown.

For a single graph F the function $I(F, x)$ is closely related to the inducibility

$$\operatorname{ind}(F) = \lim_{n\to\infty} \max\{\rho(F, G) \colon v(G) = n\}$$

of F introduced by Pippenger and Golumbic [211]. In fact, $\operatorname{ind}(F) = \max\{I(F, x) \colon x \in [0, 1]\}$, where the maximum exists due to the continuity of $I(F, x)$ (see Theorem 8.1.2 below).

Determining the feasible region $\Omega_{\text{ind}}(F)$ of a single graph F is a special case of the more general problem to determine the graph profile $T(\mathcal{F})$ of a given finite family of graphs $\mathcal{F} = \{F_1, \ldots, F_k\}$. Here $T(\mathcal{F}) \subseteq [0, 1]^k$ is the collection of limit points of $((\rho(F_1, G_i), \ldots, \rho(F_k, G_i)))_{i=1}^{\infty}$ with $v(G_i) \to \infty$. Besides the clique density theorem, very few results are known about graph profiles (see [121; 127; 33; 120]).

Our results are of two flavors.

- We prove some general results about the shape of $\Omega_{\text{ind}}(Q)$. Our main result here is Theorem 8.1.2, which states that $I(Q,x)$ and $i(Q,x)$ are continuous and almost everywhere differentiable.

- We study $\Omega_{\text{ind}}(Q)$ for some specific choices of Q for which $\text{ind}(Q)$ has been investigated by many researchers. We focus on quantum graphs whose constituents are complete multipartite graphs and prove a general upper bound for $I(Q,x)$. Prior to this work, $\Omega_{\text{ind}}(F)$ for a single graph F was determined only when F is a clique or an independent set. Here we extend this to the case $F = K_{1,2}$ and also obtain results for complete bipartite graphs. Furthermore we study $\Omega_{\text{ind}}(K_r^-)$, where K_r^- arises from the clique K_r by the deletion of a single edge. As a consequence of our results, we determine the inducibility $\text{ind}(K_r^-)$, which is new for $r \geq 5$.

8.1.1.2 General results

The following result describes the shape of the feasible region of an arbitrary quantum graph.

Theorem 8.1.2. *For every quantum graph Q we have*

$$\Omega_{\text{ind}}(Q) = \left\{(x,y) \in [0,1] \times \mathbb{R} \colon i(Q,F) \leq y \leq I(Q,F)\right\}.$$

Moreover, the boundary functions $i(Q,x)$ and $I(Q,F)$ are continuous and almost everywhere differentiable.

In contrast to Theorem 8.1.2 Hatami and Norin [120] gave an example of a finite family $\mathcal{F}$ of graphs such that the intersection of the graph profile $T(\mathcal{F})$ with some hyperplane has a nowhere differentiable boundary.

For every quantum graph Q the feasible regions of Q, $-Q$ and $\overline{Q}$ are closely related. Indeed, using the formulae

$$N(F,G) = N(\overline{F},\overline{G}) \quad \text{and} \quad \rho(F,G) = \rho(\overline{F},\overline{G}),$$

which are valid for all graphs F and G, one easily confirms the following observation.

Fact 8.1.3. *Let Q be a quantum graph.*

(a) The feasible regions of Q and $-Q$ are symmetric to each other about the x-axis. Hence, $I(-Q,x) = -i(Q,x)$ and $i(-Q,x) = -I(Q,x)$ hold for all $x \in [0,1]$.

(b) The feasible regions of Q and $\overline{Q}$ are symmetric to each other about the line $x = 1/2$. Thus we have $I(Q,x) = I(\overline{Q}, 1-x)$ and $i(Q,x) = i(\overline{Q}, 1-x)$ for every $x \in [0,1]$. In particular, if Q is self-complementary, then $I(Q,x) = I(Q,1-x)$ and $i(Q,x) = i(Q,1-x)$, i.e. the functions $I(Q,x)$ and $i(Q,x)$ are symmetric around $x = 1/2$. ∎

The next result shows that for most single graphs F the lower boundary function $i(F,x)$ vanishes identically. The only exceptions occur when F is a clique and $i(F,x)$ is given by the clique density theorem (see Theorem 8.1.10), or if F is the complement of a clique and $i(F,x)$ is given by the Kruskal–Katona theorem (and Fact 8.1.3 (b)).

Proposition 8.1.4. *If F denotes a graph which is neither complete nor empty, then $i(F,x)=0$ for all $x\in[0,1]$.*

We proceed with some estimates based on random graphs. Given a quantum graph $Q=\sum_{i=1}^{m}\lambda_i F_i$ we define

$$\mathrm{rand}(Q,x)=\sum_{i\in[m]}\lambda_i\frac{(v(F_i))!}{|\mathrm{Aut}(F_i)|}x^{e(F_i)}(1-x)^{e(\overline{F}_i)}\qquad\text{for every } x\in[0,1]\,,$$

where $\mathrm{Aut}(F_i)$ is the automorphism group of F_i for $i\in[m]$. Equivalently,

$$\mathrm{rand}(Q,x)=\lim_{n\to\infty}\mathbb{E}\,\rho(Q,G(n,x))\,,$$

where $G(n,x)$ denotes the standard binomial random graph. It is well known that the random variables $\rho(G(n,x))$, $\rho(Q,G(n,x))$ are tightly concentrated around their expectations. This shows the following observation.

Fact 8.1.5. *If Q denotes a quantum graph and $x\in[0,1]$, then*

$$I(Q,x)\geq\mathrm{rand}(Q,x)\geq i(Q,x)\,.$$

In particular, for a single graph F the inequality $I(F,x)>0$ holds for all $x\in(0,1)$.

Let $P_{4,1}$ be the 5-vertex graph that is the disjoint union of a path on 4 vertices and an isolated vertex. It was asked in [79] whether the inducibility of some graph is achieved by a random graph and, in particular, whether the inducibility $\mathrm{ind}(P_{4,1})$ is achieved by the Erdős–Rényi random graph $G(n, 3/10)$. Here we pose an easier question of a similar flavor.

Problem 8.1.6. *Do there exist a graph F and some $x \in (0,1)$ such that $I(F,x) = \mathrm{rand}(F,x)$?*

8.1.1.3 Complete multipartite graphs

We now present our results on $I(Q,x)$ for specific quantum graphs Q. Our focus is on quantum graphs whose constituents are complete multipartite graphs (a graph whose edge set is empty is viewed as complete multipartite with only one part). A case of particular interest is $Q = K_r + \overline{K}_r$ for $r \geq 3$. Goodman [115] proved that for every graph G on n vertices $\rho(K_3 + \overline{K}_3, G) \geq 1/4 + o(1)$ and the random graph $G(n, 1/2)$ shows that this bound is tight. Therefore, $i(K_3 + \overline{K}_3, x) \geq 1/4$ and equality holds when $x = 1/2$. Combining Goodman's result [115] with a theorem of Olpp [206] one can determine $\Omega_{\mathrm{ind}}(K_3 + \overline{K}_3)$ completely.

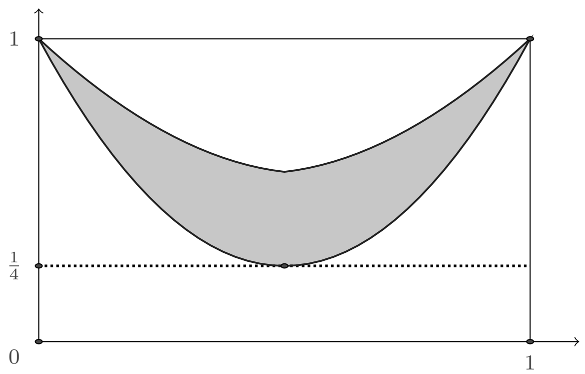


Figure 25. $\Omega_{\mathrm{ind}}(K_3 + \overline{K}_3)$ is the shaded area above.

Theorem 8.1.7 (Goodman [115], Olpp [206]). *For every $x \in [0,1]$ we have*

$$i(K_3 + \overline{K}_3, x) = 1 - 3x + 3x^2 \quad \text{and}$$
$$I(K_3 + \overline{K}_3, x) = 1 - 3\min\left\{x - x^{3/2}, (1-x) - (1-x)^{3/2}\right\}.$$

For $r \geq 4$ determining $\Omega_{\text{ind}}(K_r + \overline{K}_r)$ seems beyond current methods.

Problem 8.1.8. *Determine $\Omega_{\text{ind}}(K_r + \overline{K}_r)$ for $r \geq 4$.*

Another well-studied problem concerns the determination of $\Omega_{\text{ind}}(K_r)$ for $r \geq 3$. We already mentioned that $I(K_r, x) = x^{r/2}$ follows from the Kruskal-Katona theorem [154; 132]. For the lower bound $i(K_r, x)$ we consider (independently of r) the following complete multipartite graphs.

Definition 8.1.9. *For integers $n \geq k \geq 2$ and real $x \in \left(\frac{k-2}{k-1}, \frac{k-1}{k}\right]$ let $H^\star(n,x)$ be the complete k-partite graph on n vertices with parts $V_1, \ldots, V_k$ of sizes $|V_1| = \cdots = |V_{k-1}| = \lfloor \alpha_k n \rfloor$ and $|V_k| = n - (k-1)\lfloor \alpha_k n \rfloor$, where*

$$\alpha_k = \frac{1}{k}\left(1 + \sqrt{1 - \frac{k}{k-1}x}\right).$$

Moreover, $H^\star(n,0)$ and $H^\star(n,1)$ denote the empty and the complete graph on n vertices.

One checks immediately that $\lim_{n\to\infty} \rho(H^\star(n,x)) = x$ holds for every $x \in [0,1]$. Consequently, for every $r \geq 2$ the function $g_r(x) = \lim_{n\to\infty} \rho(K_r, H^\star(n,x))$ is an upper bound on $i(K_r, x)$.

A more explicit description of g_r is as follows. Clearly $g_r(x) = 0$ holds for every $x \le \frac{r-2}{r-1}$ and $g(1) = 1$. If $x \in \big(\frac{r-2}{r-1}, 1\big)$ there exists a unique integer $k \ge r$ such that $x \in \big(\frac{k-2}{k-1}, \frac{k-1}{k}\big]$ and a short calculation reveals

$$g_r(x) = \frac{(k)_r}{k^r}\left(1+\sqrt{1-\frac{k}{k-1}x}\right)^{r-1}\left(1-(r-1)\sqrt{1-\frac{k}{k-1}x}\right),$$

where $(k)_r = k(k-1)\cdots(k-r+1)$. Lovász and Simonovits conjectured in the seventies that this function coincides with $i(K_r, x)$ and the third author proved that this is indeed the case.

Theorem 8.1.10 (Clique density theorem, Reiher [217])**.** *For all integers $r \ge 3$ and real $x \in [0,1]$ we have $i(K_r, x) = g_r(x)$.* ∎

The non-asymptotic problem to determine for given natural numbers n and m the exact minimum number of r-cliques an n-vertex graph with m edges needs to contain is still wide open in general. But for triangles there has recently been spectacular progress by Liu, Pikhurko, and Staden [158].

Easy calculations show that the function $g_r(x)$ is non-differentiable at the critical values $x = 1 - 1/q$, where $q \ge r-1$ denotes an integer. Moreover, $g_r(x)$ is piecewise concave between any two consecutive critical values. An old result of Bollobás [25] (proved long before the clique density theorem) asserts that the piece-wise linear function interpolating between the critical values of $g_r(x)$ is a lower bound on $i(K_r, x)$. Here we extend this result to quantum graphs whose constituents are complete multipartite graphs.

To state this generalization we need the following concepts. For every positive integer $r \geq 2$ and every quantum graph Q we define the complete r-partite feasible region $\Omega_{\mathrm{ind-r}}(Q)$ to be the collection of all points in $[0, (r-1)/r] \times \mathbb{R}$ that can be realized by a Q-good sequence $(G_n)_{n=1}^{\infty}$ of complete r-partite graphs (isolated vertices are not allowed). For $x \in [0, (r-1)/r]$, let

$$i_r(Q, x) = \inf\{y\colon (x, y) \in \Omega_{\mathrm{ind-r}}(Q)\} \quad \text{and} \quad I_r(Q, x) = \sup\{y\colon (x, y) \in \Omega_{\mathrm{ind-r}}(Q)\}\,.$$

Optimizing over r we put

$$m(Q, x) = \inf\left\{i_r(Q, x)\colon r \geq \left\lceil \frac{1}{1-x} \right\rceil\right\} \quad \text{and} \quad M(Q, x) = \sup\left\{I_r(Q, x)\colon r \geq \left\lceil \frac{1}{1-x} \right\rceil\right\}$$

for every quantum graph Q and every real $x \in [0, 1)$ as well as

$$m(Q, 1) = M(Q, 1) = \lim_{n\to\infty} \rho(Q, K_n)\,.$$

Clearly, we have

$$i(Q, x) \leq m(Q, x) \leq M(Q, x) \leq I(Q, x)\,.$$

Next we observe that for every bounded function $f\colon [0, 1] \longrightarrow \mathbb{R}$ there exist a point-wise minimum concave function $\mathrm{cap}(f) \geq f$ and, similarly, a maximum convex function $\mathrm{cup}(f) \leq f$. In fact, $\mathrm{cap}(f)$ is given by

$$\mathrm{cap}(f)(x) = \sup\Big\{\lambda_1 f(x_1) + \cdots + \lambda_n f(x_n)\colon n \geq 1, (\lambda_1, \ldots, \lambda_n) \in \Delta_{n-1}, \text{ and } \sum_{i=1}^{n} \lambda_i x_i = x\Big\}$$

for all $x \in [0,1]$, where

$$\Delta_{n-1} = \big\{(\lambda_1, \dots, \lambda_n) \in [0,1]^n \colon \lambda_1 + \cdots + \lambda_n = 1\big\}$$

denotes the $(n-1)$-dimensional standard simplex. Moreover, replacing the supremum by an infimum one obtains a formula for $\mathrm{cup}(f)(x)$.

Theorem 8.1.11. *Let $Q = \sum_{i=1}^{m} \lambda_i F_i$ be a quantum graph all of whose constituents are complete multipartite graphs.*

(a) If every F_i with $\lambda_i > 0$ is complete, then

$$i(Q,x) \geq \mathrm{cup}\left(m(Q,x)\right) \qquad \textit{for all } x \in [0,1]\,.$$

(b) If every F_i with $\lambda_i < 0$ is complete, then

$$I(Q,x) \leq \mathrm{cap}\left(M(Q,x)\right) \qquad \textit{for all } x \in [0,1]\,.$$

The aforementioned result of Bollobás is the case $Q = K_r$ of Theorem 8.1.11 (a).

8.1.1.4 Almost complete graphs

For every integer $t \geq 3$ we let K_t^- denote the graph obtained from a clique K_t by deleting one edge. As these graphs are neither complete nor empty, Proposition 8.1.4 tells us that the feasible regions $\Omega_{\mathrm{ind}}(K_t^-)$ are completely determined by the functions $I(K_t^-, x)$. For $t = 3$

we have the following exact result showing that the graphs $H^\star(n,x)$ minimizing the triangle density also maximize the induced K_3^--density.

Theorem 8.1.12. *The equality* $I(K_3^-,x) = \frac{3}{2}\,(x - g_3(x))$ *holds for all* $x \in [0,1]$.

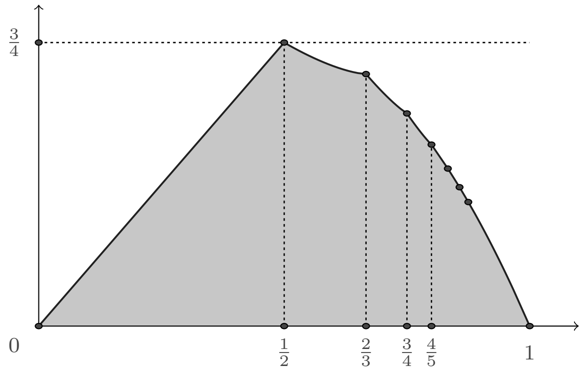


Figure 26. $\Omega_{\mathrm{ind}}(K_3^-)$.

For $t \geq 4$ we show a piecewise linear upper bound on $I(K_t^-,x)$ that yields the correct value of the inducibility $\mathrm{ind}(K_t^-)$. In the statement that follows, we set

$$k(t) = \begin{cases} \lceil (t+1)(3t-8)/6 \rceil & \text{if } t \neq 5, 8, 11, 14, 17, 20 \\ (t-2)(3t+1)/6 & \text{if } t = 5, 8, 11, 14, 17, 20. \end{cases}$$

Theorem 8.1.13. *For all $t \geq 4$ and $x \in [0,1]$ we have $I(K_t^-, x) \leq h_t(x)$, where h_t denotes the piecewise linear function interpolating between $h_t(0) = 0$ and*

$$h_t(1-1/r) = \binom{t}{2} \frac{(r-1)_{t-2}}{r^{t-1}} \qquad \textit{for } r \geq k(t)\,.$$

Furthermore,

$$\operatorname{ind}(K_t^-) = \binom{t}{2} \frac{(q(t)-1)_{t-2}}{q(t)^{t-1}}, \qquad \textit{where } q(t) = \lceil (t-2)(3t+1)/6 \rceil\,. \tag{8.1}$$

For instance, for $t = 4$ we have $q(4) = 5$ and, hence, $\operatorname{ind}(K_4^-) = 72/125$. This was originally proved by Hirst [124], whose computer assisted argument is based on the flag algebra method. Moreover, Theorem 8.1.13 yields the upper bound $I(K_4^-, x) \leq 3x/4$ for $x \in [0, 3/4]$. For small values of x we have the following stronger bound.

Proposition 8.1.14. *If $x \in [0, 1/2]$, then $I(K_4^-, x) \leq 3x^2/2$.*

Finally, we remark that our determination of $\operatorname{ind}(K_t^-)$ in Equation 8.1 implies

$$\lim_{t\to\infty} \operatorname{ind}(K_t^-) = 1/e\,. \tag{8.2}$$

This is closely related to the so-called edge-statistics conjecture of Alon, Hefetz, Krivelevich, and Tyomkyn [7]. Given positive integers k and $\ell \leq \binom{k}{2}$ let the quantum graph $Q_{k,\ell}$ be the sum of all k-vertex graphs with ℓ edges. Alon et al. conjectured $\operatorname{ind}(Q_{k,\ell}) \leq 1/e + o_k(1)$ and proved this for some range of ℓ. Following the work of Kwan, Sudakov, and Tran [156], the edges

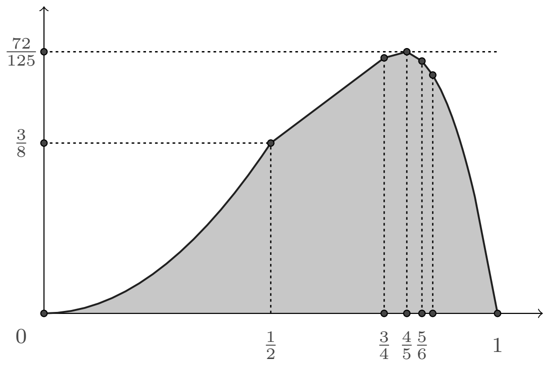


Figure 27. $\Omega_{\text{ind}}(K_4^-)$ is contained in the shaded area above.

statistics conjecture was resolved by Fox and Sauermann [86] and, independently, by Martinsson, Mousset, Noever, and Trujić [187]. Part of the original motivation for the edges statistics conjecture was the observation that for $\ell = 1$ we have $Q_{k,1} = \overline{K_k^-}$ and $\text{rand}(\overline{K_k^-}, 1/\binom{k}{2}) \to 1/e$ as $k \to \infty$. Thus the asymptotic formula Equation 8.2 follows from the results in [86; 187]. However, the exact values $\text{ind}(K_5^-) = 525/1024$, $\text{ind}(K_6^-) = 178200/13^5$, etc. implied by Theorem 8.1.13 are new.

8.1.1.5 Stars

A second case of asymptotic equality in the edge-statistics conjecture occurs for stars. For every positive integer t we denote the star with t edges by S_t. As the case $S_1 = K_2$ is trivial, we may assume $t \geq 2$ in the sequel. A quick calculation shows that the induced S_t-density of a complete bipartite graph the sizes of whose vertex classes have roughly the ratio $1 : t$ is $1/e + o_t(1)$.

A precise formula for the inducibility of stars was discovered by Brown and Sidorenko [31] (see Theorem 8.1.34 below). Here we shall show that for small densities x the values $I(S_t, x)$ of the upper bound function of the feasible region are realized by complete bipartite graphs.

Toward this goal we consider for every real $x \in [0, 1/2]$ a sequence $(B(n, x))_{n=1}^{\infty}$ of complete bipartite graphs with $v(B(n, x)) = n$ for every $n \in \mathbb{N}$ and $\lim_{n\to\infty} \rho(B(n, x)) = x$. The vertex classes of $B(n, x)$ have the sizes αn and $(1 - \alpha)n$ for some $\alpha \in [0, 1/2]$ satisfying $\alpha(1 - \alpha) = x/2 + o(1)$. Since $\rho(S_t, B(n, x)) = (t + 1)\big(\alpha(1 - \alpha)^t + (1 - \alpha)\alpha^t\big) + o_n(1)$ we are lead to the function $s_t\colon [0, 1/2] \longrightarrow \mathbb{R}$ defined by

$$s_t(x) = \lim_{n\to\infty} \rho(S_t, B(n, x)) = \frac{t+1}{2^t} x \left(\big(1 - \sqrt{1-2x}\big)^{t-1} + \big(1 + \sqrt{1-2x}\big)^{t-1}\right). \tag{8.3}$$

As we shall show in Section 8.1.5, there is a unique point $x = x^\star(t) \in [0, 1/2]$ at which $s_t(x)$ attains its maximum. Moreover,

$$x^\star(2) = x^\star(3) = \frac{1}{2} \quad \text{and} \quad \frac{2t}{(t+1)^2} < x^\star(t) < \frac{2}{t+1} \text{ holds for } t \geq 4\,.$$

Using Theorem 8.1.11 we determine $I(S_t, x)$ for $x \in [0, x^\star(t)]$.

Theorem 8.1.15. *If $t \geq 2$ is an integer and $x \in [0, x^\star(t)]$, then $I(S_t, x) = s_t(x)$.*

Notice that for $t = 2$ this tells us $I(K_3^-, x) = 3x/2$ for $x \in [0, 1/2]$, which follows from Theorem 8.1.12 as well. It seems hard to determine $I(S_t, x)$ for $t \geq 3$ and $x \geq x^\star(t)$ (some remarks on this problem are given in Section 8.1.7).

For future reference it is convenient to extend the definitions of this subsection to the trivial case $t = 1$ by setting $x^\star(1) = 1/2$ and $s_1(x) = x$ for every $x \in [0, 1/2]$ (which is one half of the values one would obtain by plugging $t = 1$ into Equation 8.3). It is then still true that we have $I(S_1, x) = s_1(x)$ for every $x \in [0, x^\star(1)]$ and that equality holds for the sequence $(B(n, x))_{n=1}^{\infty}$ of bipartite graphs.

8.1.1.6 Complete bipartite graphs

For positive integers s and t let $K_{s,t}$ denote the complete bipartite graph whose vertex classes are of size s and t. So $K_{1,t} = S_t$ is a star and it turns out that the calculation of $I(K_{s,t}, x)$ reduces to $I(S_{|s-t|+1}, x)$ for $x \in [0, x^\star(|s-t|+1)]$.

Theorem 8.1.16. *Let $t \geq s \geq 2$ be integers. Then for every $x \in [0, 1]$ we have*

$$I(K_{s,t}, x) \leq \frac{1}{2^{s-1}(t-s+2)} \binom{s+t}{s} x^{s-1} I(S_{t-s+1}, x),$$

and equality holds for $x \leq x^\star(t-s+1)$. In particular, for $x \in [0, x^\star(t-s+1)]$,

$$I(K_{s,t}, x) = \begin{cases} \frac{1}{2^t}\binom{2t}{t} x^t & \text{if} \quad t = s, \\ \frac{1}{2^t}\binom{s+t}{s} x^s \left(\left(1 - \sqrt{1-2x}\right)^{t-s} + \left(1 + \sqrt{1-2x}\right)^{t-s} \right) & \text{if} \quad t > s. \end{cases}$$

The remainder of this subsection focuses on the case $s = t = 2$. Observe that $K_{2,2} = C_4$ is a four-cycle. Theorem 8.1.16 yields $I(C_4, x) = 3x^2/2$ for every $x \in [0, 1/2]$, where equality is achieved by the sequence $(B(n, x))_{n=1}^{\infty}$ of bipartite graphs. For $x \geq 1/2$ we believe that $I(C_4, x)$ is related to the constructions for the clique density theorem (see Construction 8.1.9).

Conjecture 8.1.17. *For every real number $x \in [1/2, 1]$ we have*

$$I(C_4, x) = \lim_{n\to\infty} \rho\left(C_4, H^\star(n, x)\right) .$$

This conjecture predicts $I(C_4, 1-1/k) = 3(k-1)/k^3$ for every integer $k \geq 2$ and our next result shows that this is indeed the case.

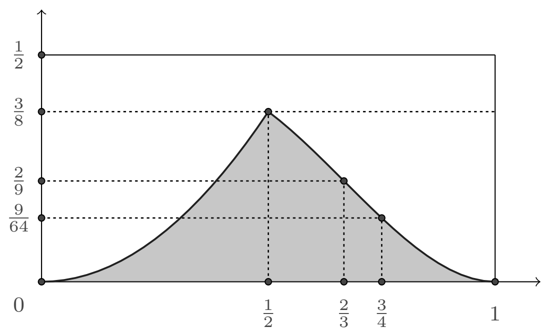


Figure 28. $\Omega_{\text{ind}}(C_4)$ is contained in the shaded area above.

Theorem 8.1.18. *If $x \in [1/2, 1]$, then*

$$I(C_4, x) \leq 3x(1-x)^2 .$$

Moreover, the bound is tight for all $x \in \{(k-1)/k \colon k \in \mathbb{N}$ and $k \geq 2\}$.

8.1.2 Proofs of general results

We prove Theorem 8.1.2 and Proposition 8.1.4 in this section. The following result can be proved using a similar argument as Proposition 1.3 in [167].

Proposition 8.1.19. *For every quantum graph Q the set $\Omega_{\text{ind}}(Q)$ is closed.* ∎

Therefore the definitions of $i(Q, x)$ and $I(Q, x)$ rewrite as

$$i(Q, x) = \min \{y \colon (x, y) \in \Omega_{\text{ind}}(Q)\} \quad \text{and} \quad I(Q, x) = \max \{y \colon (x, y) \in \Omega_{\text{ind}}(Q)\} .$$

Next we show that $\Omega_{\text{ind}}(Q)$ is determined by $i(Q, x)$ and $I(Q, x)$.

Proposition 8.1.20. *Let Q be a quantum graph, $x \in [0, 1]$ and $y_1 < y_2$. If $(x, y_1) \in \Omega_{\text{ind}}(Q)$ and $(x, y_2) \in \Omega_{\text{ind}}(Q)$, then $(x, y) \in \Omega_{\text{ind}}(Q)$ holds for all $y \in [y_1, y_2]$.*

Proof. Fix $y \in [y_1, y_2]$. Let $(G'_n)_{n=1}^{\infty}$ be a Q-good sequence of graphs that realizes (x, y_1), and let $(G''_n)_{n=1}^{\infty}$ be a Q-good sequence of graphs that realizes (x, y_2). Without loss of generality we may assume that $V(G'_n) = V(G''_n) = [n]$ for $n \geq 1$. We shall construct a sequence of graphs $(G_n)_{n=1}^{\infty}$ with $V(G_n) = [n]$ for every $n \geq 1$ that realizes (x, y).

For fixed $n \geq 1$ we consider a finite sequence of graphs $G_n^1, \dots, G_n^{m(n)}$ with common vertex set $[n]$ which interpolates between $G_n^1 = G'_n$ and $G^{m(n)} = G''_n$ in the sense that

- for $1 \leq m < m(n)$ the graph G_n^{m+1} arises from G_n^m by adding or deleting a single edge,
- and $\min\{\rho(G'_n), \rho(G''_n)\} \leq \rho(G_n^m) \leq \max\{\rho(G'_n), \rho(G''_n)\}$ for every $m \in [m(n)]$.

Due to the first bullet we have $\rho(Q, G_n^{m+1}) = \rho(Q, G_n^m) + o(1)$ for every $m \in [m(n) - 1]$. Combined with $\rho(Q, G_n^1) = y_1 + o(1)$ and $\rho(Q, G_n^{m(n)}) = y_2 + o(1)$ this proves that there exists some $k(n) \in [m(n)]$ such that the graph $G_n = G_n^{k(n)}$ satisfies $\rho(Q, G_n) = y + o(1)$. Owing to the second bullet we also have $\rho(G_n) = x + o(1)$. ∎

Towards the continuity of $I(Q, x)$ we now establish the following lemma.

Lemma 8.1.21. *For every quantum graph Q there exist constants $\ell \geq 1$ and $C \geq 0$ such that for all x, x' with $0 < x \leq x' \leq 1$ we have*

$$\frac{I(Q, x')}{(x')^\ell} \leq \frac{I(Q, x)}{x^\ell} + C \cdot \left(\left(\frac{1}{x}\right)^\ell - \left(\frac{1}{x'}\right)^\ell \right) . \tag{8.4}$$

Proof. Fix $0 < x \leq x' \leq 1$, set $\alpha = (x'/x)^{1/2} - 1$, and consider a Q-good sequence $(G'_n)_{n=1}^\infty$ that realizes $(x', I(Q, x'))$. Without loss of generality we may assume $v(G'_n) = n$ for every $n \geq 1$. Let G_n be the graph which is the union of G'_n and a set of $\lfloor \alpha n \rfloor$ isolated vertices. Since

$$\rho\left(G_n\right) = \frac{\rho\left(G'_n\right) \binom{n}{2}}{\binom{n + \lfloor \alpha n \rfloor}{2}} \to \frac{x'}{(1 + \alpha)^2} = x \quad \text{as} \quad n \to \infty \, ,$$

we have

$$I(Q, x) \geq \limsup_{n \to \infty} \rho(Q, G_n) \, . \tag{8.5}$$

To estimate the right side we write $Q = \sum_{i\in P} \lambda_i F_i + \sum_{j\in N} \lambda_j F_j$ with $\lambda_i > 0$ for $i \in P$ and $\lambda_j < 0$ for $j \in N$. Set $\ell_i = v(F_i)$ for every $i \in P \cup N$ and $\ell = \max\{\ell_i/2\colon i \in P \cup N\}$. For every $i \in P$ the fact that G'_n is a subgraph of G_n yields

$$\rho\,(F_i, G_n) \geq \frac{\rho\,(F_i, G'_n)\binom{n}{\ell_i}}{\binom{n+\lfloor \alpha n\rfloor}{\ell_i}} \geq \frac{\rho\,(F_i, G'_n)}{(1+\alpha)^{\ell_i}} = \frac{\rho\,(F_i, G'_n)}{(x'/x)^{\ell_i/2}} \geq \frac{\rho\,(F_i, G'_n)}{(x'/x)^{\ell}}. \tag{8.6}$$

For $j \in N$ we use that every induced copy of F_j in G_n is either already contained in G'_n or involves one of the new isolated vertices, which implies

$$\rho\,(F_j, G_n) \leq \frac{\rho\,(F_j, G'_n)\binom{n}{\ell_j} + \alpha n \cdot \binom{n+\lfloor \alpha n\rfloor}{\ell_j - 1}}{\binom{n+\lfloor \alpha n\rfloor}{\ell_j}} \leq \rho(F_j, G'_n) + \frac{\ell_j \cdot \alpha}{1+\alpha} + o_n(1)\,.$$

Taking into account that

$$\rho(F_j, G'_n) \leq \frac{\rho\,(F_j, G'_n)}{(x'/x)^{\ell}} + \left(1 - \left(\frac{x}{x'}\right)^{\ell}\right)$$

and

$$\frac{\alpha}{1+\alpha} = 1 - \left(\frac{x}{x'}\right)^{1/2} \leq 1 - \left(\frac{x}{x'}\right)^{\ell}$$

we obtain

$$\rho\,(F_j, G_n) \leq \frac{\rho(F_j, G'_n)}{(x'/x)^{\ell}} + (\ell_j + 1)\left(1 - \left(\frac{x}{x'}\right)^{\ell}\right) + o_n(1)\,.$$

Combined with Equation 8.6 this entails

$$\begin{aligned}\rho\left(Q,G_n\right) &= \sum_{i\in P}\lambda_i\rho\left(F_i,G_n\right)+\sum_{j\in N}\lambda_j\rho\left(F_j,G_n\right)\\ &\geq \sum_{i\in P}\lambda_i\frac{\rho\left(F_i,G_n'\right)}{(x'/x)^{\ell}}+\sum_{j\in N}\lambda_j\left(\frac{\rho\left(F_j,G_n'\right)}{(x'/x)^{\ell}}+(\ell_j+1)\left(1-\left(\frac{x}{x'}\right)^{\ell}\right)\right)-o_n(1)\\ &\geq \frac{\rho(Q,G_n')}{(x'/x)^{\ell/2}}-C\cdot\left(1-\left(\frac{x}{x'}\right)^{\ell}\right)-o_n(1)\,,\end{aligned}$$

where $C=\sum_{j\in N}(-\lambda_j)(\ell_j+1)\geq 0$. Now Equation 8.5 reveals

$$I(Q,x)\geq\frac{I(Q,x')}{(x'/x)^{\ell}}-C\cdot\left(1-\left(\frac{x}{x'}\right)^{\ell}\right)$$

and upon multiplying both sides by $x^{-\ell}$ the claim follows. ∎

For later use we record the following consequence.

Corollary 8.1.22. *Given a quantum graph Q and $x'\in[0,1]$, $\epsilon>0$, there exists some $\delta>0$ such that $I(Q,x)>I(Q,x')-\epsilon$ holds for all $x\in[0,x')$ with $|x-x'|<\delta$.*

Now we are ready to prove the main result of Subsection 8.1.1.2.

Proof of Theorem 8.1.2. Given a quantum graph Q the formula

$$\Omega_{\text{ind}}(Q)=\big\{(x,y)\in[0,1]\times\mathbb{R}\colon i(Q,F)\leq y\leq I(Q,F)\big\}$$

follows immediately from Proposition 8.1.20. Now, due to Fact 8.1.3 (a) it suffices to show that $I(Q,x)$ is continuous and almost everywhere differentiable.

Let $\ell \geq 1$, $C \geq 0$ be the constants provided by Lemma 8.1.21. Owing to Equation 8.4 the function $F\colon (0,1] \longrightarrow \mathbb{R}$ defined by $F(x) = (I(Q,x) + C)/x^{\ell}$ is decreasing. It follows that F is almost everywhere differentiable and that for every $x \in (0,1]$ the left-sided limit $\lim_{x \to x_0^-} F(x)$ exists. Consequently, the function $I(Q,x)$ has the same properties.

Let us show next that $I(Q,x)$ is left-continuous. Given an arbitrary $x_0 \in (0,1]$ we already know that the limit $y_0 = \lim_{x \to x_0^-} I(Q,x)$ exists. Proposition 8.1.19 yields $(x_0, y_0) \in \Omega_{\text{ind}}(Q)$, whence $I(Q,x_0) \geq y_0$. But $I(Q,x_0) > y_0$ would contradict Corollary 8.1.22 and thus we have indeed $I(Q,x_0) = y_0$. By Fact 8.1.3 (b) the function $I(Q,x) = I(\overline{Q}, 1-x)$ is right-continuous as well. This concludes the proof. ∎

Proof of Proposition 8.1.4. For every $n \in \mathbb{N}$ and $x \in [0,1]$ we let $H'(n,x)$ denote the n-vertex graph consisting of a clique of order $\lfloor x^{1/2} n \rfloor$ and $n - \lfloor x^{1/2} n \rfloor$ isolated vertices. Moreover, we set $H''(n,x) = \overline{H'(n,1-x)}$. Notice that $\lim_{n\to\infty} \rho(H'(n,x)) = \lim_{n\to\infty} \rho(H''(n,x)) = x$ holds for every $x \in [0,1]$.

Now suppose that F is a graph which is neither complete nor empty. If F has no isolated vertex, then $\rho(F, H'(n,x)) = 0$ holds for all $n \in \mathbb{N}$ and $x \in [0,1]$, which leads to $i(F,x) = 0$. If F has an isolated vertex we get the same conclusion from $\rho(F, H''(n,x)) = 0$. ∎

8.1.3 Proof for complete multipartite graphs

We prove Theorem 8.1.11 in this section. The following result of Schelp and Thomason [225] will be useful in our argument.

Theorem 8.1.23 (Schelp–Thomason [225])**.** *Let $Q = \sum_{i\in[m]} \lambda_i F_i$ be a quantum graph whose constituents are complete multipartite graphs and let $n \in \mathbb{N}$. If every F_i with $\lambda_i < 0$ is complete, then among all n-vertex graphs G maximizing $\rho(Q, G)$ there is a complete multipartite one.*

Definition 8.1.24. *Suppose that $H\colon [0,1] \to \mathbb{R}$ is a concave function and $L\colon [0,1] \to \mathbb{R}$ is a linear function. We say L is a tangent line of H at $x_0 \in [0,1]$ if $L(x) \geq H(x)$ holds for $x \in [0,1]$ with equality for $x = x_0$.*

It is easy to see that for every concave function $F\colon [0,1] \to \mathbb{R}$ and every $x_0 \in (0,1)$ there always exists a (not necessarily unique) tangent line of F at x_0.

Proof of Theorem 8.1.11. By Fact 8.1.3 (a) it suffices to show part (b). Let $Q = \sum_{i\in[m]} \lambda_i F_i$ be a quantum graph whose constituents are complete multipartite graphs such that every F_i with and $\lambda_i < 0$ is complete. For brevity we set $H(x) = \operatorname{cap}(M(Q,x))$ for every $x \in [0,1]$. Clearly

$$H(0) = M(Q,0) = \lim_{n\to\infty} \rho(Q, \overline{K_n}) = I(Q,0)$$

and a similar argument shows $H(1) = I(Q,1)$. So it remains to prove $H(x_0) \geq I(Q,x_0)$ for every $x_0 \in (0,1)$. To this end we choose a tangent line $L(x) = kx + p$ of H at x_0, so that

$$H(x) \leq kx + p \quad \text{for all } x \in [0,1] \quad \text{and} \quad H(x_0) = kx_0 + p\,. \tag{8.7}$$

Now let $(G_n)_{n=1}^{\infty}$ be a sequence of graphs that realizes $(x_0, I(Q, x_0))$. By Theorem 8.1.23 applied to the quantum graph $Q^\star = Q - kK_2$ there exists for every $n \geq 1$ a multipartite n-vertex graph G'_n such that $v(G'_n) = v(G_n)$ and

$$\rho(Q, G_n) - k\rho(G_n) = \rho(Q^\star, G_n) \leq \rho(Q^\star, G'_n) = \rho(Q, G'_n) - k\rho(G'_n)\,. \tag{8.8}$$

By passing to a subsequence of $(G'_n)_{n=1}^{\infty}$ we may assume that the limits $x_1 = \lim_{n\to\infty} \rho(G'_n)$ and $y_1 = \lim_{n\to\infty} \rho(Q, G'_n)$ exist. Due to the definition of $M(Q, x_1)$ and Equation 8.7 we have

$$y_1 \leq M(Q, x_1) \leq H(x_1) \leq kx_1 + p$$

and taking the limit $n \to \infty$ in Equation 8.8 it follows that

$$I(Q, x_0) - kx_0 \leq y_1 - kx_1 \leq p\,.$$

Together with Equation 8.7 this leads to the desired estimate $I(Q, x_0) \leq kx_0 + p = H(x_0)$. ∎

8.1.4 Proofs for almost complete graphs

In this section we prove Theorems 8.1.12 and 8.1.13 as well as Proposition 8.1.14.

8.1.4.1 Cherries

We begin with the proof of Theorem 8.1.12. Consider a graph $G = (V, E)$ with $|V| = n$ vertices. Counting the number of pairs $(\{x, y\}, z) \in E \times V$ with $z \neq x, y$ in two different ways, we obtain

$$(n-2)|E| = N(\overline{K_3^-}, G) + 2N(K_3^-, G) + 3N(K_3, G)\,.$$

Dividing by $2\binom{n}{3}$ and rearranging we deduce

$$\rho(K_3^-, G) = \frac{3}{2}\big(\rho(K_2, G) - \rho(K_3, G)\big) - \frac{1}{2}\rho(\overline{K_3^-}, G)\,.$$

Therefore the clique density theorem yields for every $x \in [0, 1]$ the upper bound $I(K_3^-, x) \leq \frac{3}{2}(x - g_3(x))$. Moreover, for every $x \in [0, 1]$ the sequence of multipartite graphs $(H^\star(n, x))_{n=1}^{\infty}$ is $\overline{K_3^-}$-free and establishes the lower bound $I(K_3^-, x) \geq \frac{3}{2}(x - g_3(x))$.

8.1.4.2 Piecewise linear upper bounds

Roughly speaking we show in this subsection that a concave piecewise linear function is an upper bound on $I(K_t^-, x)$ if it respects the constraints coming from Turán graphs.

Lemma 8.1.25. *Suppose that an integer $s \geq 1$ and real numbers λ, μ have the property that*

$$\frac{1}{r^{s+1}}\binom{r-1}{s} \leq \lambda\frac{r-1}{2r} + \mu \tag{8.9}$$

holds for every positive integer r. If $m \geq 1$ and $(\alpha_1, \ldots, \alpha_m) \in \Delta_{m-1}$, then

$$\sum_{i=1}^{m} \sum_{W \in \binom{[m] \setminus \{i\}}{s}} \alpha_i^2 \prod_{j \in W} \alpha_j \leq \lambda \sum_{\{i,j\} \in \binom{[m]}{2}} \alpha_i \alpha_j + \mu\,.$$

Proof. Assume for the sake of contradiction that this fails and let m denote the least positive integer for which there exists a counterexample. Appealing to a theorem of Weierstraß, we pick a point $(\alpha_1^\star, \ldots, \alpha_m^\star) \in \Delta_{m-1}$ such that the difference

$$\Phi = \sum_{i=1}^{m} \sum_{W \in \binom{[m] \setminus \{i\}}{s}} (\alpha_i^\star)^2 \prod_{j \in W} \alpha_j^\star - \lambda \sum_{\{i,j\} \in \binom{[m]}{2}} \alpha_i^\star \alpha_j^\star$$

is maximal. Due to our indirect assumption we know $\Phi > \mu$. The case $r = m$ of Equation 8.9 reveals that $\alpha_1^\star = \cdots = \alpha_m^\star = 1/m$ is false. Therefore, we have $m \geq 2$ and and for reasons of symmetry we may assume that $\alpha_1^\star < \alpha_2^\star$.

Given two real numbers $\alpha_1, \alpha_2 \geq 0$ satisfying

$$\alpha_1 + \alpha_2 = \alpha_1^\star + \alpha_2^\star$$

we write $\Phi(\alpha_1, \alpha_2)$ for the result of replacing $\alpha_1^\star$, $\alpha_2^\star$ in the above formula for Φ by α_1, α_2. So $\Phi(\alpha_1^\star, \alpha_2^\star) = \Phi$ and there are constants $c_1, \ldots, c_5$ depending only on $\alpha_3^\star, \ldots, \alpha_m^\star$, and λ such that

$$\Phi(\alpha_1, \alpha_2) = c_1 + c_2(\alpha_1 + \alpha_2) + c_3(\alpha_1^2 + \alpha_2^2) + c_4\alpha_1\alpha_2 + c_5(\alpha_1 + \alpha_2)\alpha_1\alpha_2\,.$$

Since $\alpha_1+\alpha_2$ is constant and $\alpha_1^2+\alpha_2^2$, $2\alpha_1\alpha_2$ add up to the constant $(\alpha_1^\star+\alpha_2^\star)^2$, it follows that there are constants c_6, c_7 such that

$$\Phi(\alpha_1,\alpha_2)=c_6\alpha_1\alpha_2+c_7\,.$$

If $c_6\neq 0$ we can find a real number $\xi\neq 0$ such that $|\xi|$ is very small and $\Phi(\alpha_1^\star+\xi,\alpha_2^\star-\xi)>\Phi$ contradicts the maximality of Φ. So $c_6=0$ and $\Phi(\alpha_1,\alpha_2)=c_7=\Phi$ is constant. But now $\Phi(\alpha_1^\star+\alpha_2^\star,0)=\Phi$ contradicts the minimality of m. This completes the proof. ∎

Lemma 8.1.26. *Suppose that* $t\geq 3$ *and that* $f\colon[0,1]\longrightarrow\mathbb{R}$ *is a piecewise linear concave function. If for every positive integer* r *we have*

$$f(1-1/r)\geq\binom{t}{2}\frac{(r-1)\cdots(r-(t-2))}{r^{t-1}}\,, \tag{8.10}$$

then $I(K_t^-,x)\leq f(x)$ *holds for every* $x\in[0,1]$.

Proof. Since f is the pointwise minimum of a family of linear functions, it suffices to deal with the case that $f(x)=\lambda x+\mu$ is itself linear. By Theorem 8.1.11 (b) it is enough to show $M(K_t^-,x)\leq\lambda x+\mu$ for every $x\in[0,1]$. We shall establish the more precise estimate that every complete multipartite graph G on n vertices satisfies

$$N(K_t^-,G)\leq(2\lambda|E(G)|+\mu n^2)n^{t-2}/t!\,. \tag{8.11}$$

Let $a_1, \dots, a_m$ be the sizes of the vertex classes of G and set $\alpha_i = a_i/n$ for every $i \in [m]$. Now $\sum_{i=1}^m \alpha_i = 1$ and

$$N(K_t^-, G) = \sum_{i=1}^m \binom{a_i}{2} \sum_{W \in \binom{[m]\setminus\{i\}}{t-2}} \prod_{j \in W} a_j \le \frac{n^t}{2} \sum_{i=1}^m \alpha_i^2 \sum_{W \in \binom{[m]\setminus\{i\}}{t-2}} \prod_{j \in W} \alpha_j$$

and, therefore, instead of Equation 8.11 it suffices to show

$$\sum_{i=1}^m \alpha_i^2 \sum_{W \in \binom{[m]\setminus\{i\}}{t-2}} \prod_{j \in W} \alpha_j \le \frac{4\lambda}{t!} \sum_{\{i,j\} \in \binom{[m]}{2}} \alpha_i \alpha_j + \frac{2\mu}{t!}\,.$$

By Lemma 8.1.25 applied to $t-2$, $4\lambda/t!$, $2\mu/t!$ here in place of s, λ, μ there this inequality follows from the fact that

$$\frac{1}{r^{t-1}} \binom{r-1}{t-2} \le \frac{4\lambda}{t!} \cdot \frac{r-1}{2r} + \frac{2\mu}{t!} = \frac{2f(1-1/r)}{t!}$$

holds for every $r \ge 1$, which is in turn equivalent to the hypothesis Equation 8.10. ∎

8.1.4.3 Precise calculations

Fix an integer $t \ge 4$. Our next goal is to show that the function h_t introduced in Theorem 8.1.13 satisfies the assumptions of Lemma 8.1.26. To this end we set $A_r = \binom{t}{2} \frac{(r-2)_{t-3}}{r^{t-2}}$ for every integer $r \ge 2$.

Lemma 8.1.27. *Let $t \ge 4$ and $r \ge t-1$ be integers.*

(a) If $r \le (3t^2 - 5t - 4)/6$, then $A_{r-1} < A_r$.

(b) If $r = (3t^2 - 5t - 2)/6$, *then* $A_{r-1} < A_r$ *or* $A_{r-1} > A_r$ *holds depending on whether* $t \leq 20$ *or* $t > 20$.

(c) If $r \geq (3t^2 - 5t)/6$, *then* $A_{r-1} > A_r$.

In particular, there exists a unique integer $k \geq t - 2$ *satisfying* $A_k = \max\{A_r \colon r \geq t - 2\}$, *namely* $k = k(t)$.

Proof. One confirms easily that

$$A_{r-1} < A_r \iff 1 - \frac{t-1}{r} < \left(1 - \frac{1}{r}\right)^{t-2} \left(1 - \frac{2}{r}\right).$$

Due to the approximations

$$\sum_{i=0}^{3} \frac{(-1)^i}{r^i} \binom{t-2}{i} \leq \left(1 - \frac{1}{r}\right)^{t-2} \leq \sum_{i=0}^{4} \frac{(-1)^i}{r^i} \binom{t-2}{i}$$

we obtain the implications

$$\frac{(t+2)(t-2)(t-3)}{6} < r \left(\frac{(t+1)(t-2)}{2} - r \right) \implies A_{r-1} < A_r$$

and

$$\frac{(t+2)(t-2)(t-3)}{6} > r \left(\frac{(t+1)(t-2)}{2} - r \right) + \frac{t+3}{4r} \binom{t-2}{3} \implies A_{r-1} > A_r$$

(see also the proof of Lemma 8.1.29).

So for the proof of part (a) it suffices to observe that $(t-1)/3 \le r \le (3t^2-5t-4)/6$ implies

$$r\left(\frac{(t+1)(t-2)}{2}-r\right) \ge \frac{3t^2-5t-4}{6}\cdot\frac{t-1}{3} \ge \frac{t(t-1)(t-2)}{6} > \frac{(t+2)(t-2)(t-3)}{6}\,.$$

Similarly, if $r \ge (3t^2-5t)/6 > (t-2)(t+3)/6$ we have

$$\begin{aligned} r\left(\frac{(t+1)(t-2)}{2}-r\right)+\frac{t+3}{4r}\binom{t-2}{3} &< \frac{3t^2-5t}{6}\cdot\frac{t-3}{3}+\frac{(t-3)(t-4)}{4} \\ &< \frac{(t+2)(t-2)(t-3)}{6}\,, \end{aligned}$$

which proves part (c).

We proceed with the case $r = (3t^2-5t-2)/6$, which requires $t \equiv 2 \pmod 3$. Direct calculations show $A_{r-1} < A_r$ for $t \in \{5, 8, 11, 14, 17, 20\}$ and $A_{r-1} > A_r$ for $t = 23$. As soon as $t \ge 26$ we have $8(t-8)r > 3(t+3)(t-3)(t-4)$ and hence

$$\begin{aligned} r\left(\frac{(t+1)(t-2)}{2}-r\right)+\frac{t+3}{4r}\binom{t-2}{3} &< \frac{3t^2-5t-2}{6}\cdot\frac{t-2}{3}+\frac{(t-2)(t-8)}{9} \\ &= \frac{(t+2)(t-2)(t-3)}{6}\,, \end{aligned}$$

which concludes the discussion of (b). Finally, (a) – (c) together imply

$$A_{t-2} < A_{t-1} < \cdots < A_{k(t)} \quad \text{and} \quad A_{k(t)} > A_{k(t)+1} > \ldots\,,$$

whence $A_{k(t)} = \max\{A_r \colon r \ge t-2\}$. ∎

Lemma 8.1.28. *We have $I(K_t^-, x) \leq h_t(x)$ for every $x \in [0,1]$.*

Proof. For later use we observe that the number $k = k(t)$ satisfies

$$k \geq \frac{t(t-2)}{2}\,. \tag{8.12}$$

Indeed, if $t \neq 5, 8, 11, 14, 17, 20$, then $k - t(t-2)/2 = \lceil (t-8)/6 \rceil \geq \lceil -2/3 \rceil = 0$ and in the remaining cases we have $k - t(t-2)/2 = (t-2)/6 \geq 0$.

Next we show

$$h_t(1-1/r) \geq \binom{t}{2} \frac{(r-1)\cdots(r-(t-2))}{r^{t-1}}$$

for every positive integer r. The cases $r \leq t-2$ and $r \geq k$ are clear. Now suppose that $t-1 \leq r < k$. Since $1 - 1/r \leq 1 - 1/k$ and $h_t(x) = A_k \cdot x$ for all $x \in [0, 1-1/k]$ we have

$$h_t(1-1/r) = A_k \cdot (1-1/r) \geq A_r \cdot (1-1/r) = \binom{t}{2} \frac{(r-1)\cdots(r-(t-2))}{r^{t-1}}\,,$$

as desired.

According to Lemma 8.1.26 it only remains to show that h_t is concave. Now $A_k > A_{k+1}$ rewrites as

$$\frac{h_t(1-1/k)}{1-1/k} > \frac{h_t(1-1/(k+1))}{1-1/(k+1)}$$

and, therefore, h_t is concave in some sufficiently small neighbourhood around $x = 1 - 1/k$. Define $F\colon \big[0, \frac{1}{k}\big] \longrightarrow \mathbb{R}$ by $F(x) = x\prod_{i=1}^{t-2}(1 - ix)$. Since $h_t(1 - 1/r) = \binom{t}{2}F(1/r)$ holds for every $r \geq k$, it suffices to show that F is concave. If $x \in [0, 1/k]$, then

$$\sum_{i=1}^{t-2} \frac{i}{1 - ix} \leq \frac{1 + \cdots + (t-2)}{1 - (t-2)x} \overset{\text{Equation 8.12}}{\leq} \frac{(t-2)(t-1)}{2(1 - 2/t)} = \frac{(t-1)t}{2} < \frac{2}{x}$$

and thus

$$\frac{F''(x)}{F(x)} = \sum_{1 \leq i < j \leq t-2} \frac{ij}{(1 - ix)(1 - jx)} - \frac{1}{x}\sum_{i=1}^{t-2} \frac{i}{1 - ix} < \frac{1}{2}\Big(\sum_{i=1}^{t-2} \frac{i}{1 - ix}\Big)^2 - \frac{1}{x}\sum_{i=1}^{t-2} \frac{i}{1 - ix} \leq 0\,,$$

which proves that F is indeed concave. ∎

The only part of Theorem 8.1.13 still lacking verification is Equation 8.1. Setting $B_r = \binom{t}{2}\frac{(r-1)_{t-2}}{r^{t-1}}$ for every $r \geq t-2$ and $f = \lceil (t-2)(3t+1)/6 \rceil$ we are to show $B_f = \max\{B_r \colon r \geq t-2\}$. It turns out that this holds in the following slightly stronger form.

Lemma 8.1.29. *We have* $0 = B_{t-2} < B_{t-1} < \cdots < B_f$ *and* $B_f > B_{f+1} > \ldots$ *.*

Proof. First we show $B_{r-1} < B_r$ for every integer $r \in [t-1, f]$. The fact that $(t-2)(3t+1)$ is even yields $f \leq (t-2)(3t+1)/6 + 2/3 = (t-1)(3t-2)/6$, whence

$$\frac{t-1}{3} \leq r \leq \binom{t}{2} - \frac{t-1}{3}\,.$$

For this reason we have

$$r\left(\binom{t}{2}-r\right) \geq \frac{t-1}{3}\left(\binom{t}{2}-\frac{t-1}{3}\right) > \binom{t}{3},$$

which rewrites as

$$1-\frac{t-1}{r} < 1-\frac{t}{r}+\binom{t}{2}\frac{1}{r^2}-\binom{t}{3}\frac{1}{r^3}.$$

As the right side is at most $(1-1/r)^t$, this proves

$$1 < \frac{(r-1)^t}{r^{t-1}(r-(t-1))} = \frac{B_r}{B_{r-1}},$$

as desired.

Next we show $B_{r-1} > B_r$ for every $r \geq f+1$. Due to $r \geq (3t^2-5t+4)/6 > \frac{1}{2}\binom{t}{2}$ we have

$$r\left(\binom{t}{2}-r\right) < \frac{3t^2-5t+4}{6}\cdot\frac{t-2}{3} = \binom{t}{3} - \frac{(t-2)^2}{9}.$$

Moreover, $r \geq t(t-2)/2$ implies

$$\binom{t}{4}\cdot\frac{1}{r} < \frac{(t-1)(t-3)}{12} < \frac{(t-2)^2}{9}.$$

Adding the previous two estimates we obtain

$$r\left(\binom{t}{2}-r\right)+\binom{t}{4}\cdot\frac{1}{r} < \binom{t}{3},$$

which rewrites as

$$1 - \frac{t-1}{r} > 1 - \frac{t}{r} + \binom{t}{2}\frac{1}{r^2} - \binom{t}{3}\frac{1}{r^3} + \binom{t}{4}\frac{1}{r^4}.$$

As the right side is an upper bound on $(1-1/r)^t$ we can conclude

$$1 > \frac{(r-1)^t}{r^{t-1}(r-(t-1))} = \frac{B_r}{B_{r-1}}. \blacksquare$$

8.1.4.4 More on K_4^-

Our last result on $\Omega_{\text{ind}}(K_4^-)$, Proposition 8.1.14, is an immediate consequence of the following result.

Lemma 8.1.30. *Every graph G satisfies $N(K_4^-, G) \le \frac{1}{2}\binom{|E(G)|}{2}$.*

Proof. Notice that an abstract K_4^- has two perfect matchings. Now with every induced copy of K_4^- in G we associate its two perfect matchings, viewed as members of $\binom{E(G)}{2}$. We are thereby considering $2N(K_4^-, G)$ pairs of edges of G. Since every pair $\{e, f\} \in \binom{E(G)}{2}$ can be associated to at most one copy of K_4^- in G (namely the copy induced by $e \cup f$, if it exists), this proves the claim. ∎

8.1.5 Proofs for stars

In this section we prove Theorem 8.1.15. Recall from Section 8.1.1.5 that for every integer $t \ge 3$ and every real $x \in [0, 1/2]$ we defined

$$s_t(x) = \frac{t+1}{2^t} x \left(\left(1 - \sqrt{1-2x}\right)^{t-1} + \left(1 + \sqrt{1-2x}\right)^{t-1} \right).$$

We commence by showing that there is a unique $x^\star(t) \in [0, 1/2]$, where the function s_t attains its maximum. For $t = 3$ we have $s_3(x) = 2x(1-x)$ and, hence, $x^\star(3) = 1/2$ is as desired. The case $t \geq 4$ is addressed by the next lemma.

Lemma 8.1.31. *For $t \geq 4$ there exists a unique real $x^\star(t) \in \left(\frac{2t}{(t+1)^2}, \frac{2}{t+1}\right)$ such that the function s_t is strictly increasing on $[0, x^\star(t)]$ and strictly decreasing on $[x^\star(t), 1/2]$.*

Proof. Define the auxiliary function $h\colon [0,1] \longrightarrow \mathbb{R}$ by $h(y) = 1 - ty + ty^{t-1} - y^t$. Due to $h''(y) = t(t-1)y^{t-3}(t-2-y) > 0$ for $y \in (0,1]$ this function is strictly convex. Together with $h(0) = 1$, $h(1) = 0$, and $h'(1) = t(t-3) > 0$ this shows that there exists a unique $y^\star = (0,1)$ such that $h(y^\star) = 0$, $h(y) > 0$ for $y \in [0, y^\star)$, and $h(y) < 0$ for $y \in (y^\star, 1)$.

Due to

$$\frac{\mathrm{d}}{\mathrm{d}y}\frac{y + y^t}{(1+y)^{t+1}} = \frac{h(y)}{(1+y)^{t+2}}$$

it follows that $\frac{y+y^t}{(1+y)^{t+1}}$ is strictly increasing on $[0, y^\star)$ and strictly decreasing on $(y^\star, 1]$. As $\frac{2y}{(1+y)^2}$ is strictly increasing on $[0,1]$ and

$$s_t\left(\frac{2y}{(1+y)^2}\right) = \frac{(t+1)(y+y^t)}{(1+y)^{t+1}}\,,$$

it follows that s_t has the desired monotonicity properties for $x^\star(t) = \frac{2y^\star}{(1+y^\star)^2}$.

Next, due to $h(1/t) = t^{2-t} - t^{-t} > 0$ we have $y^\star > \frac{1}{t}$ and, consequently, $x^\star(t) > \frac{2t}{(t+1)^2}$. Similarly,

$$h\left(\frac{1}{t-1}\right) < -\frac{1}{t-1} + \frac{t}{(t-1)^{t-1}} \leq \frac{t-(t-1)^2}{(t-1)^3} < 0$$

yields $y^\star < \frac{1}{t-1}$, whence

$$x^\star(t) < \frac{2(t-1)}{t^2} < \frac{2}{t+1}\,. \blacksquare$$

Lemma 8.1.32. *For every integer $t \geq 3$ the function s_t is increasing and concave on $[0, x^\star(t)]$.*

Proof. Our choice of $x^\star(t)$ guarantees that s_t is indeed increasing. So it suffices to show that s_t is concave on the interval $I_t = \left[0, \frac{2}{t+1}\right]$. Since

$$s_t(x) = \frac{t+1}{2^{t-1}} \sum_{0 \leq n \leq (t-1)/2} \binom{t-1}{2n} x(1-2x)^n$$

it suffices to show for every positive integer $n \leq (t-1)/2$ that $x(1-2x)^n$ is concave on I_t. This follows immediately from

$$\frac{\mathrm{d}^2}{\mathrm{d}x^2} x(1-2x)^n = 4n(1-2x)^{n-2}[(n+1)x - 1]\,.$$

$\blacksquare$

Our next step is to show $M(S_t, x) = I_2(S_t, x) = s_t(x)$ for $x \in [0, x^\star(t)]$. To this end we use the following result due to Brown and Sidorenko, which is implicit in the proof of Proposition 2 in [31].

Proposition 8.1.33 (Brown–Sidorenko [31])**.** *Let r, s, t, n be positive integers with $r \geq 3$. For every complete r-partite graph G on n vertices there exists a complete $(r-1)$-partite graph G' on the same vertex set such that $e(G') \leq e(G)$ and $N(K_{s,t}, G') \geq N(K_{s,t}, G)$.*

The proof proceeds by "merging" two smallest vertex classes of G, i.e., if $V_1, \ldots, V_r$ with $|V_1| \leq \cdots \leq |V_r|$ are the vertex classes of G, then one constructs G' so as to have the vertex classes $V_1 \cup V_2, V_3, \ldots, V_r$. Clearly, $r-2$ iterations of this process lead to a complete bipartite graph G'' such that $V(G'') = V(G)$, $e(G'') \leq e(G)$, and $N(K_{s,t}, G') \geq N(K_{s,t}, G)$. This shows that for the determination of the inducibility of $K_{s,t}$ only complete bipartite host graphs are relevant. This establishes the following result on stars.

Theorem 8.1.34 (Brown–Sidorenko [31])**.** *For every integer $t \geq 2$ the inducibility of S_t is given by* $\mathrm{ind}(S_t) = I_2(S_t, x^\star(t))$.

We proceed with another simple consequence of Proposition 8.1.33.

Lemma 8.1.35. *If $r, t \geq 2$ are integers and $x \in [0, x^\star(t)]$, then $I_2(S_t, x) \geq I_r(S_t, x)$.*

Proof. Let $y_2 = I_2(S_t, x)$, $y_r = I_r(S_t, x)$ and consider an S_t-good sequence of complete r-partite graphs $(G_n)_{n=1}^{\infty}$ that realizes (x, y_r). In view of Proposition 8.1.33 there exists a sequence $(G'_n)_{n=1}^{\infty}$ of complete bipartite graphs such that

$$V(G'_n) = V(G_n), \quad e(G'_n) \leq e(G_n), \quad \text{and} \quad N(K_{s,t}, G'_n) \geq N(K_{s,t}, G_n) \tag{8.13}$$

hold for every positive integer n. By passing to a subsequence we may assume that the limits $x' = \lim_{n\to\infty} \rho(G'_n)$ and $y'_2 = \lim_{n\to\infty} \rho(S_t, G'_n)$ exist. Now Equation 8.13 implies

$$x' \leq x \quad \text{and} \quad y'_2 \geq y_r\,, \tag{8.14}$$

and as $(G'_n)_{n=1}^{\infty}$ is an S_t-good sequence of complete bipartite graphs that realizes (x', y'_2) we have $y'_2 \le I_2(S_t, x')$. Since $I_2(S_t, \cdot) = s_t(\cdot)$ is increasing on $[0, x^\star(t)]$, the first estimate in Equation 8.14 entails $I_2(S_t, x') \le I_2(S_t, x)$. So altogether we obtain

$$y_r \le y'_2 \le I_2(S_t, x') \le I_2(S_t, x)\,,$$

which concludes the proof. ∎

Now we are ready to prove Theorem 8.1.15.

Proof of Theorem 8.1.15. The case $t = 2$ already being understood in Theorem 8.1.12 we may assume that $t \ge 3$. It is clear that $I(S_t, x) \ge I_2(S_t, x) = s_t(x)$ holds for $x \in [0, 1/2]$ and thus we just need to show $I(S_t, x) \le s_t(x)$ for $x \in [0, x^\star(t)]$. Define $f\colon [0,1] \longrightarrow \mathbb{R}$ by

$$f(x) = \begin{cases} s_t(x) & \text{for} \quad x \in [0, x^\star(t)] \\ s_t\left(x^\star(t)\right) & \text{for} \quad x \in [x^\star(t), 1]. \end{cases}$$

Lemma 8.1.32 informs us that f is concave. Moreover, we have $f(x) \ge M(S_t, x)$ for all $x \in [0, 1]$. Indeed, if $x \in [0, x^\star(t)]$ this follows from Lemma 8.1.35 and for $x \in [x^\star(t), 1]$ we can appeal to Theorem 8.1.34 instead. Summarizing, $f(x)$ is a concave upper bound on $M(S_t, x)$. Owing to Theorem 8.1.11 this proves $I(S_t, x) \le f(x) = s_t(x)$ for every $x \in [0, x^\star(t)]$. ∎

8.1.6 Proofs for complete bipartite graphs

In this section we prove Theorems 8.1.16 and 8.1.18. The upper bound on $I(K_{s,t}, x)$ stated in Theorem 8.1.16 is an immediate consequence of the following result.

Proposition 8.1.36. *If $t \geq s \geq 2$ are positive integers, then for every graph G we have*

$$N(K_{s,t}, G) \leq \frac{(t-s+1)!}{s!t!} N(S_{t-s+1}, G) \cdot (e(G))^{s-1} .$$

Proof. Notice that for an abstract $K_{s,t}$ the number of ordered partitions $V(K_{s,t}) = U_1 \cup \cdots \cup U_s$ such that U_1 induces a star S_{t-s+1} and each of $U_2, \ldots, U_s$ induces an edge is $\binom{t}{t-s+1}(s-1)!s!$. This is because there are $s\binom{t}{t-s+1}$ possibilities for U_1; moreover, if $i \in [2, s]$ and $U_1, \ldots, U_{i-1}$ are already fixed, then there are $(s-i+1)^2$ possibilities for U_i.

By double counting it follows that $\binom{t}{t-s+1}(s-1)!s!N(K_{s,t}, G)$ is at most the number of s-tuples $(U_1, \ldots, U_s)$ of subsets of G such that $G[U_1] \cong S_{t-s+1}$ and $G[U_i] \cong K_2$ for all $i \in [2, s]$, whence

$$\binom{t}{t-s+1}(s-1)!s!N(K_{s,t}, G) \leq N(S_{t-s+1}, G) \cdot (e(G))^{s-1} .$$

Now it remains to observe $\binom{t}{t-s+1}(s-1)!s! = \frac{s!t!}{(t-s+1)!}$. ∎

We remark that this argument is asymptotically optimal if G is a complete bipartite graph. More precisely, for $x \leq x^\star(t-s+1)$ the sequence $(B(n, x))_{n=1}^{\infty}$ establishes the equality case in Theorem 8.1.16. This observation concludes the proof of Theorem 8.1.16.

In the remainder of this section we show the following explicit version of Theorem 8.1.18.

Theorem 8.1.37. *Every graph G on n vertices with $xn^2/2$ edges satisfies*

$$N(C_4, G) \leq \frac{x(1-x)^2}{8} n^4 + 2n^3 \,.$$

For the proof we need the following well-known result due to Goodman [115], whose short proof we include for the sake of completeness.

Proposition 8.1.38 (Goodman [115])**.** *For every real number $x \in [0,1]$, every positive integer n, and every graph G on n vertices with $xn^2/2$ edges we have*

$$\sum_{v \in V(G)} e(v) \geq \sum_{v \in V(G)} d(v)^2 - xn^3/2 \,,$$

where $e(v) = e(G[N(v)])$ denotes the number of triangles containing the vertex v.

Proof. Counting the number of pairs $(u, \{v, w\}) \in V(G) \times E(G)$ with $v, w \in N(u)$ in two different ways, we obtain

$$\sum_{u \in V(G)} e(u) \geq \sum_{vw \in G} \big(d(v) + d(w) - n\big) = \sum_{v \in V(G)} d(v)^2 - e(G) \cdot n \,.$$

∎

Goodman's formula has the following consequence, which will assist us in the inductive proof of Theorem 8.1.37.

Corollary 8.1.39. *Every graph G with n vertices and $xn^2/2$ edges possesses a vertex v satisfying*

$$e(v) \geq \frac{d(v)^2}{2} + \frac{(1-4x+3x^2)n^2}{4} - \frac{(1-x)^3 n^3}{4(n-d(v))}\,.$$

Proof. The Cauchy–Schwarz inequality implies $\sum_{v \in V(G)} d(v)^2 \geq x^2 n^3$ and because of

$$\sum_{v \in V(G)} (n - d(v)) = (1-x)n^2$$

we also have

$$\sum_{v \in V(G)} \frac{1}{n-d(v)} \geq \frac{1}{1-x}\,.$$

Consequently,

$$\begin{aligned}\sum_{v \in V(G)} \Big(\frac{d(v)^2}{2} + \frac{(1-4x+3x^2)n^2}{4} - \frac{(1-x)^3 n^3}{4(n-d(x))}\Big) &\leq \sum_{v \in V(G)} \frac{d(v)^2}{2} + \frac{(x^2-x)n^3}{2} \\ &\leq \sum_{v \in V(G)} d(v)^2 - xn^3/2\,.\end{aligned}$$

Due to Proposition 8.1.38 the result now follows by averaging. ∎

Proof of Theorem 8.1.37. We argue by induction on n. The base case $n \leq 3$ is clear, for there are no 4-cycles in graphs with less than four vertices. Now suppose $n \geq 4$ and that our claim holds for every graph on $n-1$ vertices.

Given a graph G on n vertices with $xn^2/2$ edges we invoke Corollary 8.1.39 and get a vertex $v \in V(G)$ such that

$$e \geq \frac{d^2}{2} + \frac{(1-4x+3x^2)n^2}{4} - \frac{(1-x)^3 n^3}{4(n-d)}\,, \tag{8.15}$$

where $d = d(v)$ and $e = e(v)$. We contend that

$$N(C_4, G) \leq N(C_4, G-v) + (d^2/2 - e)(n-d)\,, \tag{8.16}$$

or, in other words, that there are at most $(d^2/2 - e)(n-d)$ induced copies of $K_{2,2}$ in G which contain the vertex v. The reason for this is that each such copy contains a pair of non-adjacent members of $N(v)$ and a fourth vertex belonging to $V(G) \setminus N(v)$. Clearly there are at most $d^2/2 - e$ possibilities for such a non-adjacent pair and at most $n-d$ possibilities for the fourth vertex.

Claim 8.1.40. *We have*

$$8N(C_4, G-v) \leq x(1-x)^2(n^4 - 4n^3) + 2(xn-d)(1-4x+3x^2)n^2 + 16n^3\,.$$

Proof. The induction hypothesis yields

$$8N(C_4, G-v) \leq x'(1-x')^2(n-1)^4 + 16(n-1)^3\,, \tag{8.17}$$

where x' is defined by

$$x' = \frac{2|E(G-v)|}{(n-1)^2} = \frac{xn^2 - 2d}{(n-1)^2}\,.$$

The function $h(x) = x(1-x)^2$ has derivatives $h'(x) = 1 - 4x + 3x^2$ and $h''(x) = -4 + 6x$. Therefore we have $\|h'\|_{[0,1]} = 1$ and $\|h''\|_{[0,1]} = 4$, where $\|\cdot\|_{[0,1]}$ denotes the supremum norm with respect to the unit interval. So Taylor's formula and Equation 8.17 imply

$$8N(C_4, G-v) \le x(1-x)^2(n-1)^4 + (1-4x+3x^2)(x'-x)(n-1)^4 \\ + 2(x'-x)^2(n-1)^4 + 16(n-1)^3\,.$$

Here

$$x(1-x)^2(n-1)^4 \le x(1-x)^2(n^4 - 4n^3 + 6n^2) \le x(1-x)^2(n^4 - 4n^3) + n^2$$

and due to

$$x' - x = \frac{(2n-1)x - 2d}{(n-1)^2} \tag{8.18}$$

we have $2(x'-x)^2(n-1)^4 = 2\big|(2n-1)x - 2d\big|^2 \le 8n^2$. For these reasons it suffices to establish

$$(1-4x+3x^2)(x'-x)(n-1)^4 \le 2(xn-d)(1-4x+3x^2)n^2 + 7n^2\,. \tag{8.19}$$

Now the triangle inequality yields

$$\begin{aligned}
&\left|(x'-x)(n-1)^4-2(xn-d)n^2\right| \\
&\qquad \le \left|(x'-x)(n-1)^2-2(xn-d)\right|(n-1)^2+2|xn-d|\big(n^2-(n-1)^2\big) \\
&\qquad \overset{\text{Equation 8.18}}{\le} x(n-1)^2+4n^2 \le 5n^2
\end{aligned}$$

and together with $\|h'\|_{[0,1]}=1$ this proves Equation 8.19. Thereby Claim 8.1.40 is proved. ∎

Now combining Equation 8.15, Equation 8.16, and Claim 8.1.40 we obtain

$$\begin{aligned}
8N(C_4,G) &\le x(1-x)^2(n^4-4n^3)+2(xn-d)(1-4x+3x^2)n^2+16n^3 \\
&\qquad -2(1-4x+3x^2)n^2(n-d)+2(1-x)^3n^3 \\
&= x(1-x)^2n^4+16n^3\,,
\end{aligned}$$

as desired. ∎

8.1.7 Concluding remarks

8.1.7.1 General questions

As the example $Q=K_3+\overline{K}_3$ shows, for a quantum graph Q the function $I(Q,x)$ can have at least two global maxima. We do not know whether this is possible for single graphs F as well.

Problem 8.1.41. *Does there exist a graph F such that the function $I(F,x)$ has at least two global maxima?*

Two questions of a similar flavor are as follows.

Problem 8.1.42. *Does there exist a graph F such that for some nontrivial interval J we have $I(F,x) = \mathrm{ind}(F)$ for all $x \in J$?*

Problem 8.1.43. *Does there exist a graph F such that the function $I(F,x)$ has a nontrivial local maximum (minimum)?*

Recall that for a self-complementary graph F the function $I(F,x)$ is symmetric around $x = 1/2$. One may thus wonder whether some appropriate self-complementary graph F yields an affirmative solution to Problem 8.1.41. This approach leads to the following question.

Problem 8.1.44. *Let F be a self-complementary graph. Is it true that $I(F,x) = \mathrm{ind}(F)$ holds if and only if $x = 1/2$?*

8.1.7.2 Problems for specific graphs

The smallest problem left open by our results on stars in Section 8.1.5 is to determine $I(S_3,x)$ for $x \in [1/2,1]$. In an interesting contrast to the case $S_2 = K_3^-$ one can show that the clique density construction (see Construction 8.1.9) is not extremal for this problem. For $x \in [4\sqrt{2}-5,1]$ the best construction we are aware of is the complement of a clique of order $\lfloor (1-x)^{1/2} n \rfloor$, which leads to the bound

$$I(S_3,x) \geq 4(1-(1-x)^{1/2})(1-x)^{3/2}\,. \tag{8.20}$$

For $x \in [0.91, 0.93]$ we have a complicated argument based on the results in [219] which shows that equality holds in Equation 8.20. In the range $x \in [1/2, 4\sqrt{2} - 5)$ the complement of two disjoint cliques of order $\lfloor ((1 - x)/2)^{1/2} n \rfloor$ shows that $I(S_3, x)$ is strictly larger than the right side of Equation 8.20. We hope to return to this problem in the near future.

Finally, we would like to emphasize Conjecture 8.1.17 again: Is it true that for $x \in [1/2, 1]$ the graphs in Construction 8.1.9 minimizing the triangle density maximize the induced C_4 density?

CHAPTER 9

INDEPENDENT SETS IN SPARSE HYPERGRAPHS

Previously published as T. Bohman, X. Liu, and D. Mubayi. Independent sets in hypergraphs omitting an intersection. Random Structures Algorithms, 2021.

9.1 Independent set in hypergraphs that omit one intersection

9.1.1 Introduction

The seminal Turán Theorem [241] implies that $\alpha(G) \geq n/(d+1)$ for every graph G on n vertices with average degree d. Spencer [234] extended Turán's result to hypergraphs and proved that for all $k \geq 3$ every n-vertex k-graph $\mathcal{H}$ with average degree d satisfies

$$\alpha(\mathcal{H}) \geq c_k \frac{n}{d^{1/(k-1)}} \tag{9.1}$$

for some constant $c_k > 0$.

The bound for $\alpha(\mathcal{H})$ can be improved if we forbid some family $\mathcal{F}$ of hypergraphs in $\mathcal{H}$. For $\ell \geq 2$ a (Berge) cycle of length ℓ in $\mathcal{H}$ is a collection of ℓ edges $E_1, \ldots, E_\ell \in \mathcal{H}$ such that there exists ℓ distinct vertices $v_1, \ldots, v_\ell$ with $v_i \in E_i \cap E_{i+1}$ for $i \in [\ell - 1]$ and $v_\ell \in E_\ell \cap E_1$. A seminal result of Ajtai, Komlós, Pintz, Spencer, and Szemerédi [4] states that for every n-vertex k-graph $\mathcal{H}$ with average degree d that contains no cycles of length $2, 3$, and 4, there exists a constant $c'_k > 0$ such that

$$\alpha(\mathcal{H}) \geq c'_k \frac{n}{d^{1/(k-1)}} (\log d)^{1/(k-1)}. \tag{9.2}$$

Moreover, this is tight apart from c'_k.

Spencer [200] conjectured and Duke, Lefmann, and Rödl [53] proved that the same conclusion holds even if $\mathcal{H}$ just contains no cycles of length 2. Their result was further extended by Rödl and Šiňajová [222] to the larger family of (n, k, ℓ)-systems defined in the following section.

9.1.1.1 (n, k, ℓ)-systems and (n, k, ℓ)-omitting systems

Let $k > \ell \geq 1$. Recall that an n-vertex k-graph $\mathcal{H}$ is an (n, k, ℓ)-system if the intersection of every pair of edges in $\mathcal{H}$ has size less than ℓ, and $\mathcal{H}$ is an (n, k, ℓ)-omitting system if it has no two edges whose intersection has size exactly ℓ. It is clear from the definition that an (n, k, ℓ)-system is an (n, k, ℓ)-omitting system, but not vice versa, since an (n, k, ℓ)-omitting system may have pairwise intersection sizes greater than ℓ.

Define

$$f(n, k, \ell) = \min\{\alpha(\mathcal{H})\colon \mathcal{H} \text{ is an } (n, k, \ell)\text{-system}\}, \quad \text{and}$$
$$g(n, k, \ell) = \min\{\alpha(\mathcal{H})\colon \mathcal{H} \text{ is an } (n, k, \ell)\text{-omitting system}\}.$$

The study of $f(n, k, \ell)$ has a long history (e.g. [222; 150; 78; 239]) and, in particular, Rödl and Šiňajová [222] proved that

$$f(n, k, \ell) = \Theta\left(n^{\frac{k-\ell}{k-1}} (\log n)^{\frac{1}{k-1}}\right) \text{ for all fixed } k > \ell \geq 2. \tag{9.3}$$

It follows that

$$g(n, k, \ell) \leq f(n, k, \ell) = O\left(n^{\frac{k-\ell}{k-1}} (\log n)^{\frac{1}{k-1}}\right). \tag{9.4}$$

One important difference between (n, k, ℓ)-systems and (n, k, ℓ)-omitting systems is their maximum sizes. By definition, every set of ℓ vertices in an (n, k, ℓ)-system is contained in at most one edge, thus every (n,k,ℓ)-system has size at most $\binom{n}{\ell}/\binom{k}{\ell} = O\left(n^{\ell}\right)$. However, this is not true for (n,k,ℓ)-omitting systems. Indeed, the following result of Frankl and Füredi [99] shows that the maximum size of an (n,k,ℓ)-omitting system can be much larger than that of an (n,k,ℓ)-system when $k > 2\ell+1$.

Let $k > \ell \geq 1$ and $\lambda \geq 1$ be integers. Recall that the k-graph $S^k_\lambda(\ell)$ consists of λ edges $E_1,\ldots,E_\lambda$ such that $E_i \cap E_j = S$ for $1 \leq i < j \leq \lambda$ and some fixed set S (called the center) of size ℓ. When $\ell = 1$ we just write S^k_λ, and we will omit the superscript k in $S^k_\lambda(\ell)$ if it is obvious. It is easy to see that an n-vertex k-graph is an (n,k,ℓ)-omitting system iff it is $S_2(\ell)$-free, and is an (n,k,ℓ)-system iff it is $\{S_2(\ell),\ldots,S_2(k-1)\}$-free.

Theorem 9.1.1 (Frankl–Füredi [99]). *Let $k > \ell \geq 1$ and $\lambda > 1$ be fixed integers and $\mathcal{H}$ be an $S_\lambda(\ell)$-free k-graph on n vertices. Then $|\mathcal{H}| = O\left(n^{\max\{\ell,k-\ell-1\}}\right)$. Moreover, the bound is tight up to a constant multiplicative factor.*

Theorem 9.1.1 together with Equation 9.1 imply that for fixed k,ℓ,

$$g(n,k,\ell) = \begin{cases} \Omega\left(n^{\frac{k-\ell}{k-1}}\right) & k \leq 2\ell+1, \\ \Omega\left(n^{\frac{\ell+1}{k-1}}\right) & k > 2\ell+1. \end{cases} \tag{9.5}$$

Notice that for $k \leq 2\ell+1$ the bounds given by Equation 9.4 and Equation 9.5 match except for a factor of $(\log n)^{1/(k-1)}$, but for $k > 2\ell+1$, these two bounds have a gap in the exponent of n.

Our main goal in this paper is to extend the results of Rödl and Šiňajová to the larger class of (n, k, ℓ)-omitting systems and improve the bounds given by Equation 9.4 and Equation 9.5. In other words, the question we focus on is the following:

What is the value of $g(n, k, \ell)$?

Our results for (n, k, ℓ)-omitting systems are divided into two parts. For $k \leq 2\ell + 1$, we believe that the behavior is similar to that of (n, k, ℓ)-systems and prove a nontrivial lower bound for the first open case $\ell = k - 2$. For $k > 2\ell + 1$ we give new lower and upper bounds which show that the minimum independence number of (n, k, ℓ)-omitting systems has a very different behavior than for (n, k, ℓ)-systems.

9.1.1.2 $k \leq 2\ell + 1$

As mentioned above, for this range of ℓ and k, the issue at hand is only the polylogarithmic factor in $g(n, k, \ell)$. It follows from the definition that an $(n, k, k-1)$-omitting system is also an $(n, k, k-1)$-system, thus Rödl and Šiňajová's result implies that

$$g(n, k, k-1) = f(n, k, k-1) = \Theta\left(n^{\frac{1}{k-1}} (\log n)^{\frac{1}{k-1}}\right).$$

So, the first open case in the range of $k \leq 2\ell + 1$ is $\ell = k - 2$, and for this case we prove the following nontrivial lower bound for $g(n, k, k-2)$, which improves Equation 9.5.

Theorem 9.1.2. *Suppose that $k \geq 4$. Then every $(n, k, k-2)$-omitting system has an independent set of size $\Omega\left(n^{2/(k-1)} (\log\log n)^{1/(k-1)}\right)$. In other words,*

$$g(n, k, k-2) = \Omega\left(n^{\frac{2}{k-1}} (\log\log n)^{\frac{1}{k-1}}\right).$$

Unfortunately, our method for proving Theorem 9.1.2 cannot be extended to the entire range of $k \leq 2\ell + 1$, but we make the following conjecture.

Conjecture 9.1.3. *For all fixed integers $k > \ell \geq 2$ that satisfy $k \leq 2\ell + 1$ there exists a function $\omega(n) \to \infty$ as $n \to \infty$ such that $g(n, k, \ell) = \Omega\left(n^{\frac{k-\ell}{k-1}} \omega(n)\right)$.*

Theorem 9.1.2 shows that Conjecture 9.1.3 is true for $\ell = k - 2$. The smallest open case is $k = 5$ and $\ell = 2$.

9.1.1.3 $k > 2\ell + 1$

Recall that in the range of $k > 2\ell + 1$ the bounds given by Equation 9.4 and Equation 9.5 leave a gap in the exponent of n. The following result shows that for a wide range of k and ℓ neither of them gives the correct order of magnitude.

Theorem 9.1.4. *Let $\ell \geq 2$ and $k > 2\ell + 1$ be fixed. Then*

$$\Omega\left(\max\left\{n^{\frac{\ell+1}{3\ell-1}}, n^{\frac{\ell+1}{k-1}}\right\}\right) = g(n, k, \ell) = O\left(n^{\frac{\ell+1}{2\ell}} (\log n)^{\frac{1}{\ell}}\right).$$

Remarks.

(a) The lower bound $n^{\frac{\ell+1}{3\ell-1}}$ can be improved to $n^{\frac{3-\sqrt{5}}{2}+o_\ell(1)} \sim n^{0.38196+o_\ell(1)}$. See the remark in the end of Section 9.1.3 for details.

(b) It is clear that Theorem 9.1.4 improves the bound given by Equation 9.5 for $k > 3\ell$, and it also improves the bound given by Equation 9.4 for $k > 2\ell + 1$ as $\frac{k-\ell}{k-1} - \frac{\ell+1}{2\ell} = \frac{(\ell-1)(k-2\ell-1)}{2\ell(k-1)} > 0$ for $k > 2\ell + 1$.

It would be interesting to determine $g(n, k, \ell)$ for $k > 2\ell + 1$. Here, we are not able to offer a conjecture for the exponent of n.

Problem 9.1.5. *Determine the order of magnitude of $g(n, k, \ell)$ for $k > 2\ell + 1$.*

For the first open case $(k, \ell) = (6, 2)$ Theorem 9.1.4 gives $\Omega\left(n^{3/5}\right) = g(n, 6, 2) = O\left(n^{3/4+o(1)}\right)$. Similar to Remark (a) above the lower bound for $g(n, 6, 2)$ can be improved to $\Omega\left(n^{2/3}\right)$. See the remark in the end of Section 9.1.3 for details.

9.1.1.4 (n, k, ℓ, λ)-systems and (n, k, ℓ, λ)-omitting systems

We consider the following generalization of (n, k, ℓ)-omitting systems and (n, k, ℓ)-systems in this section.

Recall that an n-vertex k-graph $\mathcal{H}$ is an (n, k, ℓ, λ)-system if every set of ℓ vertices is contained in at most λ edges, and $\mathcal{H}$ is an (n, k, ℓ, λ)-omitting system if it does not contain $S_{\lambda+1}(\ell)$ as a subgraph.

Define

$$f(n,k,\ell,\lambda) = \min\{\alpha(\mathcal{H})\colon \mathcal{H} \text{ is an } (n,k,\ell,\lambda)\text{-system}\}, \quad \text{and}$$

$$g(n,k,\ell,\lambda) = \min\{\alpha(\mathcal{H})\colon \mathcal{H} \text{ is an } (n,k,\ell,\lambda)\text{-omitting system}\}.$$

When λ is a fixed constant, the value of $f(n,k,\ell,\lambda)$ is essentially the same as $f(n,k,\ell)$ (e.g. see [222]), i.e. $f(n,k,\ell,\lambda) = \Theta(f(n,k,\ell))$. Similarly, the same conclusions as in Theorems 9.1.2 and 9.1.4 also hold for $g(n,k,\ell,\lambda)$, since Theorem 9.1.1 holds for all $S_\lambda(\ell)$-free hypergraphs and using it one can easily extend the proof for the case $\lambda = 1$ to the case $\lambda > 1$. For the sake of simplicity, we will prove Theorem 9.1.2 only for the case $\lambda = 1$.

When λ is not a constant, even the value of $f(n,k,\ell,\lambda)$ is not known in general. Here is a summary of the known results.

- $\ell = 1$: An $(n,k,1,\lambda)$-system is just a k-graph with maximum degree λ and here complete k-graphs and Equation 9.1 yield

  $$f(n,k,1,\lambda) = \Theta\left(\frac{n}{\lambda^{1/(k-1)}}\right).$$

  On the other hand a result of Loh [174] implies

  $$g(n,k,1,\lambda) = \frac{n}{\lambda+1} \qquad \text{whenever} \qquad (\lambda+1)(k-1) \mid n.$$

  If the divisibility condition fails then we have a small error term above.

- $\ell = k-1$: Kostochka, Mubayi, and Verstraëte [150] proved that

$$f(n,k,k-1,\lambda) = \Theta\left(\left(\frac{n}{\lambda}\right)^{\frac{1}{k-1}}\left(\log\frac{n}{\lambda}\right)^{\frac{1}{k-1}}\right) \quad for \quad 1 \le \lambda \le \frac{n}{(\log n)^{3(k-1)^2}}.$$

- $2 \le \ell \le k-2$: Tian and Liu [239] proved that

$$f(n,k,\ell,\lambda) = \Omega\left(\left(\frac{n}{\lambda}\log\frac{n}{\lambda}\right)^{1/\ell}\right) \quad for \quad k \ge 5,\ \frac{2k+4}{5} < \ell \le k-2,\ \lambda = o\left(n^{\frac{5\ell-2k-4}{3k-9}}\right).$$

  They also gave a construction which implies that

$$f(n,k,\ell,\lambda) = O\left(\left(\frac{n^{k-\ell}}{\lambda}\right)^{\frac{1}{k-1}}\left(\log\frac{n}{\lambda}\right)^{\frac{1}{k-1}}\right) \quad for \quad 2 \le \ell \le k-1,\ \log n \ll \lambda \ll n.$$

Since for every $\lambda > 0$ an (n,k,ℓ,λ)-system has size $O\left(\lambda n^{\ell}\right)$, it follows from Equation 9.1 that

$$f(n,k,\ell,\lambda) = \Omega\left(\left(\frac{n^{k-\ell}}{\lambda}\right)^{\frac{1}{k-1}}\right),$$

which, by Tian and Liu's upper bound, is tight up to a factor of $(\log n)^{1/(k-1)}$ when $\log n \ll \lambda \ll n$.

Using a result of Duke, Lefmann, and Rödl [53] we are able to improve the lower bound for $f(n,k,\ell,\lambda)$ to match the upper bound obtained by Tian and Liu for a wide range of λ.

Theorem 9.1.6. *Let $k > \ell \geq 2$ be fixed. If there exists a constant $\delta > 0$ such that $0 < \lambda < n^{\frac{\ell-1}{k-2}-\delta}$, then*

$$f(n,k,\ell,\lambda) = \Omega\left(\left(\frac{n^{k-\ell}}{\lambda}\right)^{\frac{1}{k-1}} (\log n)^{\frac{1}{k-1}}\right).$$

Remark. It remains open to determine $f(n,k,\ell,\lambda)$ for $\Omega\left(n^{\frac{\ell-1}{k-2}-o(1)}\right) = \lambda = O\left(n^{k-\ell}\right)$.

Since Theorem 9.1.1 does not hold when λ is not a constant, our method of proving Theorems 9.1.2 and 9.1.4 cannot be extended to this case.

9.1.1.5 Applications in Ramsey theory

For a k-graph $\mathcal{F}$ the Ramsey number $r_k(\mathcal{F}, t)$ is the smallest integer n such that every $\mathcal{F}$-free k-graph on n vertices has an independent set of size at least t. Determining the minimum independence number of an $\mathcal{F}$-free k-graph on n vertices is essentially the same as determining the value of $r_k(\mathcal{F}, t)$. So, our results above can be applied to determine the Ramsey number of some hypergraphs.

First, Theorem 9.1.2 and Equation 9.4 imply the following corollary.

Corollary 9.1.7. *Let $k \geq 4$ and $\lambda \geq 2$ be fixed integers. Then*

$$\Omega\left(\frac{t^{(k-1)/2}}{(\log t)^{1/2}}\right) = r_k(S_\lambda(k-2), t) = O\left(\frac{t^{(k-1)/2}}{(\log\log t)^{1/2}}\right).$$

Similarly, Theorem 9.1.4 gives the following corollary.

Corollary 9.1.8. *Let $\ell \geq 2$, $k > 2\ell + 1$, and $\lambda \geq 2$ be fixed integers. Then*

$$\Omega\left(\frac{t^{2\ell/(\ell+1)}}{(\log t)^{2/(\ell+1)}}\right) = r_k(S_\lambda(\ell), t) = O\left(\min\left\{t^{\frac{3\ell-1}{\ell+1}}, t^{\frac{k-1}{\ell+1}}\right\}\right).$$

Remark. According to Remark (a) after Theorem 9.1.4, the upper bound $t^{\frac{3\ell-1}{\ell+1}}$ above can be improved to $t^{\frac{3+\sqrt{5}}{2}+o_\ell(1)} \sim t^{2.61803+o_\ell(1)}$.

The following result about $r_k(S^k_\lambda, t)$ follows from a more general result of Loh [174].

Theorem 9.1.9 (Loh [174])**.** *Let $t \geq k \geq 2$, $t - 1 = q(k-1) + r$ for some $q, r \in \mathbb{N}$ with $0 \leq r \leq k-2$. Then for every $\lambda \geq 2$*

$$\lambda q(k-1) + r + 1 \leq r_k(S^k_\lambda, t) \leq \lambda q(k-1) + \lambda r + 1.$$

In particular, $r_k(S^k_\lambda, t) = \lambda(t-1) + 1$ whenever $(k-1) \mid (t-1)$.

The k-*Fan*, denoted by F^k, is the k-graph consisting of $k+1$ edges $E_1, \ldots, E_k, E$ such that $E_i \cap E_j = v$ for all $1 \leq i < j \leq k$, where $v \notin E$, and $|E_i \cap E| = 1$ for $1 \leq i \leq k$. In other words, F^k is obtained from S^k_k by adding an edge omitting v that intersects each edge of S^k_k. It is easy to see that F^2 is just the triangle K_3. The k-graph F^k was first introduced by Mubayi and Pikhurko [195] in order to extend Mantel's theorem to hypergraphs. Unlike the case $k = 2$, where it is well known that $r_2(K_3, t) = \Theta\left(t^2/\log t\right)$ (e.g. see [5; 146]), the following result shows that $r_k(F^k, t) = \Theta(t^2)$ for all $k \geq 3$.

Theorem 9.1.10. *Suppose that $t \geq k \geq 3$. Then*

$$\left\lfloor \frac{t}{2} \right\rfloor \left\lfloor \frac{t-1}{2(k-2)} \right\rfloor < r_k(F^k, t) \leq t(t-1)+1.$$

As $t \to \infty$, it remains open to determine $\lim r_k(F^k, t)/t^2$.

9.1.2 Proof of Theorem 9.1.2

In this section we prove Theorem 9.1.2. Let us show some preliminary results first.

9.1.2.1 Preliminaries

For a pair of distinct vertices $u, v \in V(\mathcal{H})$ the $(k-1)$-codegree of u and v is the number of $(k-1)$-sets $S \subset V(\mathcal{H})$ such that $S \cup \{u\} \in \mathcal{H}$ and $S \cup \{v\} \in \mathcal{H}$. Denoted by $\Gamma(\mathcal{H})$ the maximum $(k-1)$-codegree of $\mathcal{H}$.

The random greedy independent set algorithm. We begin with $\mathcal{H}(0) = \mathcal{H}, V(0) = V(\mathcal{H})$ and $I(0) = \emptyset$. Given independent set $I(i)$ and hypergraph $\mathcal{H}(i)$ on vertex set $V(i)$, a vertex $v \in V(i)$ is chosen uniformly at random and added to $I(i)$ to form $I(i+1)$. The vertex set $V(i+1)$ is set equal to $V(i)$ less v and all vertices u such that $\{u, v\}$ is an edge in $\mathcal{H}(i)$. The hypergraph $\mathcal{H}(i+1)$ is formed form $\mathcal{H}_i$ by

1. removing v from all edges of size at least three in $\mathcal{H}(i)$ that contain v, and
2. removing every edge that contains a vertex u such that the pair $\{u, v\}$ is an edge of $\mathcal{H}(i)$.

The process terminates when $V(i) = \emptyset$. At this point $I(i)$ is a maximal independent set in $\mathcal{H}$. Let $i_{\max}$ denote the step where the algorithm terminates.

In [18], Bennett and Bohman analyzed the random greedy independent set algorithm using the differential equation method, and they proved that if a k-graph satisfies certain degree and codegree conditions, then the random greedy independent set algorithm produces a large independent set with high probability.

Theorem 9.1.11 (Bennett–Bohman [18])**.** *Let k and $\epsilon > 0$ be fixed. Let $\mathcal{H}$ be a D-regular k-graph on n vertices such that $D > n^{\epsilon}$. If*

$$\Delta_i(\mathcal{H}) < D^{\frac{k-i}{k-1}-\epsilon} \quad for \quad 2 \leq i \leq k-1, \quad and \quad \Gamma(\mathcal{H}) < D^{1-\epsilon},$$

then the random greedy independent set algorithm produces an independent set I in $\mathcal{H}$ of size $\Omega\left((\log n)^{1/(k-1)} \cdot n/D^{1/(k-1)}\right)$ with probability $1 - o(1)$.

The lower bound on independence number in Theorem 9.1.11 can easily be proved by applying a theorem of Duke–Lefmann–Rödl [53] (see Theorem 9.1.13), so the main novelty of Theorem 9.1.11 is the fact that the random greedy independent set algorithm produces an independent set of this size with high probability.

Let $S \subset V(\mathcal{H})$ be a set of bounded size s such that S contains no edge in $\mathcal{H}$. A nice property of the random greedy independent set algorithm is that S is contained in the set $I(i)$ with probability $(1 + o(1)) (i/n)^s$, which is almost the probability that S is contained in a random i-subset of $V(\mathcal{H})$.

Using this property we can easily control the size of the induced subgraph of $\mathcal{G}$ on $I(i)$, where $\mathcal{G}$ is a hypergraph that has the same vertex set with $\mathcal{H}$.

Proposition 9.1.12 (Bennett–Bohman [18]). *Let $\mathcal{H}$ be a hypergraph that satisfies the conditions in Theorem 9.1.11 and $\mathcal{G}$ be a k'-graph on $V(\mathcal{H})$ (i.e. $\mathcal{G}$ and $\mathcal{H}$ are on the same vertex set). If $i \leq i_{\max}$ is fixed, then the expected number of edges of $\mathcal{G}$ contained in $I(i)$ is at most* $(1 + o(1))\,(i/n)^{k'} \cdot |\mathcal{G}|$.

For $2 \leq j \leq k-1$ and two edges E, E' in a k-graph $\mathcal{H}$ we say $\{E, E'\}$ is a $(2,j)$-cycle if $|E \cap E'| = j$. Denote by $C_{\mathcal{H}}(2,j)$ the number of $(2,j)$-cycles in $\mathcal{H}$. Duke, Lefmann, and Rödl [53] proved the following result for hypergraphs with few $(2,j)$-cycles.

Theorem 9.1.13 (Duke–Lefmann–Rödl [53]). *Let $\mathcal{H}$ be a k-graph on n vertices satisfying $\Delta(\mathcal{H}) \leq t^{k-1}$, where $t \gg k$. If $C_{\mathcal{H}}(2,j) \leq nt^{2k-j-1-\epsilon}$ for $2 \leq j \leq k-1$ and some constant $\epsilon > 0$, then $\alpha(\mathcal{H}) \geq c(k,\epsilon)\,(\log t)^{1/(k-1)} \cdot n/t$.*

Recall that a hypergraph is linear if every pair of edges has at most one vertex in common. It is easy to see that $\mathcal{H}$ is linear iff $C_{\mathcal{H}}(2,j) = 0$ for $2 \leq j \leq k-1$. The following easy corollary of Theorem 9.1.13 will be handy for proofs in the next section.

Corollary 9.1.14 (see e.g. [108]). *Suppose that $\mathcal{H}$ is a linear k-graph with n vertices and average degree d. Then $\alpha(\mathcal{H}) = \Omega\left((\log d)^{1/(k-1)} \cdot n/d^{1/(k-1)}\right)$.*

For a (not necessarily uniform) hypergraph $\mathcal{H}$ on n vertices (assuming that $V(\mathcal{H}) = [n]$) and a family $\mathcal{F} = \{\mathcal{G}_1, \ldots, \mathcal{G}_n\}$ of m-vertex k-graphs with $V(\mathcal{G}_1) = \cdots = V(\mathcal{G}_n) = V_{\mathcal{F}}$ the Cartesian product of $\mathcal{H}$ and $\mathcal{F}$, denoted by $\mathcal{H} \square \mathcal{F}$, is a hypergraph on $V(\mathcal{H}) \times V_{\mathcal{F}}$ and

$$\mathcal{H} \square \mathcal{F} = \{(E, v) : E \in \mathcal{H} \text{ and } v \in V_{\mathcal{F}}\} \cup \{(i, F) \colon i \in [n] \text{ and } F \in \mathcal{G}_i\}.$$

Since the hypergraphs we considered here are not necessarily regular, Theorem 9.1.11 cannot be applied directly to our situations. To overcome this issue we use an adaption of a trick used by Shearer in [226], that is, for every nonregular hypergraph $\mathcal{H}$ we take the Cartesian product of $\mathcal{H}$ and a family of linear hypergraphs to get a new hypergraph $\widehat{\mathcal{H}}$ that is regular. Then we apply Theorem 9.1.11 to $\widehat{\mathcal{H}}$ to get a large independent set, and by the Pigeonhole principle, this ensures that $\mathcal{H}$ has a large independent set.

First, we need the following theorem to show the existence of sparse regular linear hypergraphs.

Given two k-graphs $\mathcal{H}_1$ and $\mathcal{H}_2$ with the same number of vertices a packing of $\mathcal{H}_1$ and $\mathcal{H}_2$ is a bijection $\phi : V(\mathcal{H}_1) \to V(\mathcal{H}_2)$ such that $\phi(E) \notin \mathcal{H}_2$ for all $E \in \mathcal{H}_1$.

Theorem 9.1.15 (Lu–Székely [179])**.** *Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be two k-graphs on n vertices. If*

$$\Delta(\mathcal{H}_1)|\mathcal{H}_2| + \Delta(\mathcal{H}_2)|\mathcal{H}_1| < \frac{1}{ek}\binom{n}{k},$$

then there is a packing of $\mathcal{H}_1$ and $\mathcal{H}_2$.

Theorem 9.1.15 enables us to construct sparse regular linear hypergraphs inductively.

Lemma 9.1.16. *For every positive integer n that satisfies $k \mid n$ and every positive integer d that satisfies*

$$d \le \frac{(n-k+2)(n-k+1)}{ek^2(k-1)^2 n} + 1,$$

there exists a d-regular linear k-graph with n vertices.

Proof. We proceed by induction on d and note that the case $d = 1$ is trivial since a perfect matching on n vertices is a 1-regular linear k-graph. Now suppose that $d \geq 2$. By the induction hypothesis, there exists a $(d-1)$-regular linear k-graph on n vertices, and let $\mathcal{H}_{d-1}$ be such a k-graph. Let $\mathcal{H}_1$ be a perfect matching on n vertices. Define the extended k-graph $\widehat{\mathcal{H}}_1$ of $\mathcal{H}_1$ as

$$\widehat{\mathcal{H}}_1 = \left\{ \{u,v\} \cup A \colon \{u,v\} \in \partial_{k-2}\mathcal{H}_1 \text{ and } A \in \binom{V(\mathcal{H}_1) \setminus \{u,v\}}{k-2} \right\}.$$

It is clear from the definition that $\mathcal{H}_1 \subset \widehat{\mathcal{H}}_1$, $|\widehat{\mathcal{H}}_1| < \frac{n}{k}\binom{k}{2}\binom{n}{k-2}$, and $\widehat{\mathcal{H}}_1$ is regular. So,

$$\Delta(\widehat{\mathcal{H}}_1) = \frac{k|\widehat{\mathcal{H}}_1|}{n} < \frac{k}{n}\frac{n}{k}\binom{k}{2}\binom{n}{k-2} = \binom{k}{2}\binom{n}{k-2}.$$

By assumption

$$\begin{aligned} \Delta(\mathcal{H}_{d-1})|\widehat{\mathcal{H}}_1| + \Delta(\widehat{\mathcal{H}}_1)|\mathcal{H}_{d-1}| &< (d-1)\frac{n}{k}\binom{k}{2}\binom{n}{k-2} + \frac{(d-1)n}{k}\binom{k}{2}\binom{n}{k-2} \\ &= 2(d-1)\frac{n}{k}\binom{k}{2}\binom{n}{k-2} \leq \frac{1}{ek}\binom{n}{k}. \end{aligned}$$

Therefore, by Theorem 9.1.15, there exist a bijection $\phi : V(\mathcal{H}_{d-1}) \to V(\mathcal{H}_1)$ such that $|\phi(E) \cap E'| \leq k-1$ for all $E \in \mathcal{H}_{d-1}$ and $E' \in \widehat{\mathcal{H}}_1$, and this implies that $|\phi(E) \cap E''| \leq 1$ for all $E \in \mathcal{H}_{d-1}$ and all $E'' \in \mathcal{H}_1$. Therefore, $\mathcal{H}_1 \cup \phi(\mathcal{H}_{d-1})$ is a d-regular linear k-graph on n vertices. ∎

9.1.2.2 Proofs

First we use Theorem 9.1.11 and Proposition 9.1.12 to prove a result about the common independent set of two hypergraphs on the same vertex set.

Theorem 9.1.17. *Let $k_1, k_2 \geq 2$ be integers, $\epsilon > 0$, $n, D \in \mathbb{N}$, and $d > 0$. Suppose that*

(a) $\mathcal{H}$ is an n-vertex k_1-graph, $\mathcal{G}$ is an n-vertex k_2-graph, and $V(\mathcal{H}) = V(\mathcal{G}) = V$,

(b) $D > n^\epsilon$ and $d\,(\log n/D)^{\frac{k_2-1}{k_1-1}} \gg 1$,

(c) $\mathcal{H}$ satisfies that $\Delta(\mathcal{H}) \leq D$,

$$\Delta_i(\mathcal{H}) < D^{\frac{k_1-i}{k_1-1}-\epsilon} \quad \textit{for} \quad 2 \leq i \leq k_1 - 1, \quad \textit{and} \quad \Gamma(\mathcal{H}) < D^{1-\epsilon},$$

(d) $\mathcal{G}$ satisfies that $d(\mathcal{G}) \leq d$ and

$$C_{\mathcal{G}}(2,i) \ll n\,(D/\log n)^{\frac{2k_2-i-1}{k_1-1}} \quad \textit{for} \quad 2 \leq i \leq k_2 - 1.$$

Then, $\alpha(\mathcal{H} \cup \mathcal{G}) = \Omega\left(\omega \cdot n/d^{1/(k_2-1)}\right)$, where

$$\omega = \omega(n, D, d, k_1, k_2) = \left(\log\left((\log n/D)^{\frac{k_2-1}{k_1-1}}\, d\right)\right)^{1/(k_2-1)}.$$

Remarks.

- Although Theorem 9.1.17 imposes no condition on k_1 and k_2, we will only apply the result in the case $k_2 = k_1 + 1$.

- Spencer's bound Equation 9.1 implies that $\alpha(\mathcal{G}) = \Omega\left(n/d^{1/(k_2-1)}\right)$. Theorem 9.1.17 improves it in two ways: first it improves the bound by a factor of ω, second it is a lower bound for the independence number of $\mathcal{G} \cup \mathcal{H}$. Ajtai, Komlós, Pintz, Spencer, and Szemerédi's result Equation 9.2 implies that the upper bound for ω is $(\log n)^{1/(k_2-1)}$. However, we are not able to show that $\omega = \Omega\left((\log n)^{1/(k_2-1)}\right)$ in general, and it would be interesting to determine the optimal value of ω.

- If $\mathcal{H}$ and $\mathcal{G}$ satisfy conditions (a) and (c) in Theorem 9.1.17 and also satisfy

  (b') $D > n^{\epsilon}$ and $d\left(\log n/D\right)^{\frac{k_2-1}{k_1-1}} \ll 1$,

  then $\alpha\left(\mathcal{H} \cup \mathcal{G}\right) = \Omega\left((\log n)^{1/(k_1-1)} \cdot n/D^{1/(k_1-1)}\right)$. Moreover, if $\mathcal{G} = \emptyset$, then $\alpha(\mathcal{H}) = \Omega\left((\log n)^{1/(k_1-1)} \cdot n/D^{1/(k_1-1)}\right)$ which is the bound in Theorem 9.1.11. The proof is similar to the proof of Theorem 9.1.17.

Proof of Theorem 9.1.17. For $2 \leq i \leq k_2 - 1$ define

$$\mathcal{G}^i = \left\{ S \in \binom{V}{2k_2 - i} : \mathcal{G}[S] \text{ contains a } (2,i)\text{-cycle} \right\}.$$

Fix $m \in \mathbb{N}$ such that $D \ll m = O(n^{k_1})$, and $k_1 \mid m$. Notice that D has a trivial upper bound n^{k_1-1}, so such an integer m exists. For every $v \in V$ let $D_v = D - d_{\mathcal{H}}(v)$. Since $m \gg D$ and

$k_1 \mid m$, by Lemma 9.1.16, there exists a D_v-regular linear k_1-graph $\mathcal{F}(v)$ on $[m]$ for every $v \in V$. Let

$$\mathcal{H}' = \mathcal{H} \cup \mathcal{G} \cup \left(\bigcup_{2 \le i \le k_2 - 1} \mathcal{G}^i \right), \quad \mathcal{F} = \{\mathcal{F}(v) : v \in V\}, \quad \textit{and} \quad \widehat{\mathcal{H}}' = \mathcal{H}' \Box \mathcal{F}.$$

Note that $\widehat{\mathcal{H}}'$ is consisting of

1. the k_1-graph $\widehat{\mathcal{H}} = \mathcal{H} \Box \mathcal{F}$,
2. the k_2-graph $\widehat{\mathcal{G}}$ that is the union of m pairwise vertex-disjoint copies of $\mathcal{G}$, and
3. the $(2k_2 - i)$-graph $\widehat{\mathcal{G}}^i$ that is the union of m pairwise vertex-disjoint copies of $\mathcal{G}^i$ for $2 \le i \le k_2 - 1$.

For every $v \in V(\widehat{\mathcal{H}})$ we have $d_{\widehat{\mathcal{H}}}(v) = d_{\mathcal{H}}(v) + D_v = D$. Moreover,

$$\Delta_i(\widehat{\mathcal{H}}) = \Delta_i(\mathcal{H}) < D^{\frac{k_1 - i}{k_1 - 1} - \epsilon} \quad \textit{for} \quad 2 \le i \le k_1 - 1, \quad \textit{and} \quad \Gamma(\widehat{\mathcal{H}}) = \Gamma(\mathcal{H}) < D^{1-\epsilon},$$

Applying the random greedy independent set algorithm and Theorem 9.1.11 to $\widehat{\mathcal{H}}$, we obtain an independent set $\hat{I}$ of size at least $c\,(\log nm)^{1/(k_1-1)} \cdot nm/D^{1/(k_1-1)}$ for some constant $c > 0$ with probability $1 - o(1)$. Let $p = c\,((\log nm)/D)^{1/(k_1-1)}$ and we may assume that $|\hat{I}| = pnm$ since otherwise we can take the set of the first pnm vertices generated by the random greedy independent set algorithm instead.

Applying Proposition 9.1.12 to $\widehat{\mathcal{G}}, \widehat{\mathcal{G}}^2, \ldots, \widehat{\mathcal{G}}^{k_2-1}$ and by assumption (d) we obtain

$$\mathbb{E}\left[\left|\widehat{\mathcal{G}}[\hat{I}]\right|\right] \leq (1+o(1))p^{k_2}|\widehat{\mathcal{G}}| < 2dnmp^{k_2},$$

and for $2 \leq i \leq k_2 - 1$

$$\mathbb{E}\left[\left|\widehat{\mathcal{G}}^i[\hat{I}]\right|\right] = (1+o(1))p^{2k_2-i} \cdot m \cdot C_{\mathcal{G}}(2,i) = o(pnm).$$

So, by Markov's inequality and the union bound, with probability at least 1/2 both

$$\left|\widehat{\mathcal{G}}[\hat{I}]\right| \leq 10dnmp^{k_2} \quad and \quad \left|\widehat{\mathcal{G}}^i[\hat{I}]\right| = o(pnm) \quad \forall\ 2 \leq i \leq k_2 - 1$$

hold.

Fix a set $\hat{I}$ such that $|\hat{I}| = pnm$ and the events above hold. Then by removing $o(pnm)$ vertices we obtain a subset $\hat{I}' \subset \hat{I}$ such that

$$\left|\widehat{\mathcal{G}}^i[\hat{I}]\right| = 0 \quad for \quad 2 \leq i \leq k_2 - 1.$$

In other words, the k_2-graph $\widehat{\mathcal{G}}[\hat{I}']$ is linear. Since

$$d\left(\widehat{\mathcal{G}}[\hat{I}']\right) \leq \frac{k_2 \cdot 10dnmp^{k_2}}{(1-o(1))pnm} \leq 20k_2dp^{k_2-1},$$

by Corollary 9.1.14, it has an independent set I' of size at least

$$\Omega\left(\frac{pnm}{(20k_2dp^{k_2-1})^{1/(k_2-1)}}\left(\log 20dp^{k_2-1}\right)^{\frac{1}{k_2-1}}\right)=\Omega\left(m\frac{n}{d^{1/(k_2-1)}}\left(\log p^{k_2-1}d\right)^{\frac{1}{k_2-1}}\right)$$
$$=\Omega\left(m\frac{n}{d^{1/(k_2-1)}}\omega\right).$$

Here we used assumption (b) to ensure that $20k_2dp^{k_2-1}\geq 1$.

By the Pigeonhole principle, there exists $j\in[m]$ such that $I=I'\cap(V\times\{j\})$ has size at least $|\hat{I}|/m=\Omega\left(\omega\cdot n/d^{1/(k_2-1)}\right)$, and it is clear that I is independent in both $\mathcal{H}$ and $\mathcal{G}$. ∎

Next we use Theorem 9.1.17 to prove Theorem 9.1.2. The idea is to first decompose an $(n,k,k-2)$-omitting system $\mathcal{H}$ into two parts: $\mathcal{H}_{k-1}\subset\partial\mathcal{H}$ and $\mathcal{H}_k\subset\mathcal{H}$, and then apply Theorem 9.1.17 to $\mathcal{H}_{k-1}$ and $\mathcal{H}_k$ to find a large set $I\subset V$ that is independent in both of them. It will be easy to see that the set I is independent in $\mathcal{H}$.

Proof of Theorem 9.1.2. Let $\mathcal{H}$ be an $(n,k,k-2)$-omitting system and let $V=V(\mathcal{H})$. By Theorem 9.1.1, there exists a constant C_1 such that $|\mathcal{H}|\leq C_1n^{k-2}$. Let $\beta=\beta(k)>0$ be a constant such that $\frac{k}{2(k-1)}<\beta<1$, for example, take $\beta=4/5$. Define

$$\mathcal{H}_{k-1}=\left\{A\in\partial\mathcal{H}\colon d_{\mathcal{H}}(A)\geq\frac{n^{\frac{k-3}{k-1}}}{(\log n)^{\beta}}\right\}\quad\textit{and}\quad\mathcal{H}_k=\left\{E\in\mathcal{H}\colon\binom{E}{k-1}\cap\mathcal{H}_{k-1}=\emptyset\right\}.$$

Let $k_1 = k-1$, $k_2 = k$, $D = n^{k-4+2/(k-1)}(\log n)^\beta$, $d = C_1 n^{k-3}$, and ϵ be a constant such that $0 < \epsilon < 1/(k-1)$. Then $D > n^\epsilon$ and

$$d\left(\frac{\log n}{D}\right)^{\frac{k_2-1}{k_1-1}} = C_1 n^{k-3}\left(\frac{\log n}{n^{k-4+\frac{2}{k-1}}(\log n)^\beta}\right)^{\frac{k-1}{k-2}} = C_1(\log n)^{(1-\beta)\frac{k-1}{k-2}} \gg 1.$$

Therefore, condition (b) in Theorem 9.1.17 is satisfied. Next we show that $\mathcal{H}_{k-1}$ and $\mathcal{H}_k$ satisfy (c) and (d) in Theorem 9.1.17 with our choice of k_1, k_2, D, d, ϵ.

Claim 9.1.18. *The $(k-1)$-graph $\mathcal{H}_{k-1}$ is an $(n, k-1, k-2)$-system with $\Delta_{k-3}(\mathcal{H}_{k-1}) \leq n^{2/(k-1)}(\log n)^\beta$.*

Proof. First we prove that $\mathcal{H}_{k-1}$ is an $(n, k-1, k-2)$-system. Indeed, suppose to the contrary that there exist $e_1, e_2 \in \mathcal{H}_{k-1}$ such that $S = e_1 \cap e_2$ has size $k-2$. By the definition of $\mathcal{H}_{k-1}$, $|N_{\mathcal{H}}(e_i)| \geq n^{\frac{k-3}{k-1}}/(\log n)^\beta > 2k$ for $i = 1, 2$. So there exist $v_1, v_2 \in V \setminus (e_1 \cup e_2)$ such that $E_i = e_i \cup \{v_i\} \in \mathcal{H}$ for $i = 1, 2$. However, $E_1 \cap E_2 = S$ has size $k-2$, contradicting the assumption that $\mathcal{H}$ is an $(n, k, k-2)$-omitting system. Therefore, $\mathcal{H}_{k-1}$ is an $(n, k-1, k-2)$-system.

Now suppose to the contrary that there exists a set $A \subset V$ of size $k-3$ with $d_{\mathcal{H}_{k-1}}(A) = m > n^{2/(k-1)}(\log n)^\beta$. Since $\mathcal{H}_{k-1}$ is an $(n, k-1, k-2)$-system, $L_{\mathcal{H}_{k-1}}(A)$ is a matching consisting of m edges. Suppose that $L_{\mathcal{H}_{k-1}}(A) = \{e_1, \ldots, e_m\}$, and let $B_i = A \cup e_i$ for $1 \leq i \leq m$. Since $B_i \in \mathcal{H}_{k-1}$, by definition, there exists a set $N_i \subset V$ of size at least $n^{\frac{k-3}{k-1}}/(\log n)^\beta$ such that $B_i \cup \{u\} \in \mathcal{H}$ for all $u \in N_i$.

Suppose that there exists $v \in N_i \cap N_j$ for some distinct $i, j \in [m]$. Then the two sets $A \cup e_i \cup \{v\}$ and $A \cup e_j \cup \{v\}$ are edges in $\mathcal{H}$ and have an intersection of size $k-2$, a contradiction. Therefore, $N_i \cap N_j = \emptyset$ for all distinct $i, j \in [m]$. It follows that

$$n = |V| \geq \sum_{i \in [m]} |N_i| \geq m n^{\frac{k-3}{k-1}} / (\log n)^\beta > n^{\frac{2}{k}} (\log n)^\beta n^{\frac{k-3}{k-1}} / (\log n)^\beta > n,$$

a contradiction. Therefore, $\Delta_{k-3}(\mathcal{H}_{k-1}) \leq n^{2/(k-1)} (\log n)^\beta$. ∎

Since $\Delta_{k-3}(\mathcal{H}_{k-1}) \leq n^{2/(k-1)} (\log n)^\beta$, for every set $S \subset V$ of size i with $i \in [k-4]$ the link $L_{\mathcal{H}_{k-1}}(S)$ is an $\left(n, k-1-i, k-3-i, n^{2/(k-1)} (\log n)^\beta\right)$-system. Therefore, for $i \in [k-4]$

$$\Delta_i(\mathcal{H}_{k-1}) \leq n^{\frac{2}{k-1}} (\log n)^\beta \binom{n}{k-3-i} \Big/ \binom{k-1-i}{k-3-i} < n^{k-3-i+\frac{2}{k-1}} (\log n)^\beta.$$

Since

$$\left(k-4+\frac{2}{k-1}\right)\frac{k-1-i}{k-2} - \left(k-3-i+\frac{2}{k-1}\right) = \frac{2(i-1)}{k-1} > \epsilon,$$

we obtain

$$\Delta_i(\mathcal{H}_{k-1}) < n^{k-3-i+\frac{2}{k-1}} (\log n)^\beta < D^{\frac{k-1-i}{k-1-1}-\epsilon} \quad for \quad 2 \leq i \leq k-3.$$

On the other hand, since $\mathcal{H}$ is an $(n, k-1, k-2)$-system, $\Delta_{k-2}(\mathcal{H}_{k-1}) \leq 1 < D^{\frac{k-1-(k-2)}{k-1-1}-\epsilon}$ and $\Gamma(\mathcal{H}_{k-1}) = 0 < D^{1-\epsilon}$. Therefore, $\mathcal{H}_{k-1}$ satisfies condition (c) in Theorem 9.1.17.

Claim 9.1.19. *The k-graph $\mathcal{H}_k$ satisfies $d(\mathcal{H}_k) \leq C_1 k n^{k-2}$,*

$$C_{\mathcal{H}_k}(2,i) = O\left(n^{2k-4-i}\right) \quad for \quad 2 \leq i \leq k-3,$$

$C_{\mathcal{H}_k}(2, k-2) = 0$*, and* $C_{\mathcal{H}_k}(2, k-1) = O\left(n^{k-2+\frac{k-3}{k-1}}/(\log n)^{\beta}\right)$.

Proof. First, it is clear that $C_{\mathcal{H}_k}(2, k-2) = 0$ since there is no pair of edges in $\mathcal{H}_k$ with an intersection of size $k-2$.

Let $2 \leq i \leq k-3$ and $S \subset V$ be a set of size i. Since $\mathcal{H}_k$ is an $(n, k, k-2)$-omitting system, the link $L_{\mathcal{H}_k}(S)$ is an $(n, k-i, k-2-i)$-omitting system. So, by Theorem 9.1.1, $|L_{\mathcal{H}_k}(S)| = O\left(n^{k-2-i}\right)$, which implies that

$$C_{\mathcal{H}_k}(2,i) \leq |\mathcal{H}_k| \cdot \binom{k}{i} \cdot O\left(n^{k-2-i}\right) = O\left(n^{2k-4-i}\right) \quad for \quad 2 \leq i \leq k-3.$$

Now let $S \subset V$ be a set of size $k-1$. By the definition of $\mathcal{H}_k$, $d_{\mathcal{H}_k}(S) \leq n^{2/(k-1)}/(\log n)^{\beta}$. Therefore,

$$C_{\mathcal{H}_k}(2, k-1) \leq |\mathcal{H}_k| \cdot \binom{k}{k-1} \cdot n^{\frac{k-3}{k-1}}/(\log n)^{\beta} = O\left(n^{k-2+\frac{k-3}{k-1}}/(\log n)^{\beta}\right).$$

∎

Since

$$1+\left(k-4+\frac{2}{k-1}\right)\frac{2k-1-i}{k-2}-(2k-4-i)=\frac{2(i-1)}{k-1}>\epsilon,$$

by Claim 9.1.19,

$$C_{\mathcal{H}_k}(2,i)=O\left(n^{2k-4-i}\right)=o\left(n\left(D/\log n\right)^{\frac{2k-i-1}{k-1-1}}\right)\quad for\quad 2\leq i\leq k-3.$$

Moreover, $C_{\mathcal{H}_k}(2,k-2)=0\ll n\left(D/\log n\right)^{\frac{2k-(k-2)-1}{k-1-1}}$, and

$$C_{\mathcal{H}_k}(2,k-1)=O\left(\frac{n^{k-2+\frac{k-3}{k-1}}}{(\log n)^{\beta}}\right)\ll\frac{n^{k-2+\frac{k-3}{k-1}}}{(\log n)^{(1-\beta)\frac{k}{k-2}}}=n\left(\frac{D}{\log n}\right)^{\frac{2k-(k-1)-1}{k-1-1}},$$

where the inequality follows from the assumption that $\beta>\frac{k}{2(k-1)}$. Therefore, $\mathcal{H}_k$ satisfies condition (d) in Theorem 9.1.17.

So, by Theorem 9.1.17, there exists a set $I\subset V$ of size $\Omega\left(\omega\cdot n/n^{\frac{k-3}{k-1}}\right)=\Omega\left(n^{2/(k-1)}\omega\right)$ such that I is independent in both $\mathcal{H}_{k-1}$ and $\mathcal{H}_k$. Here

$$\begin{aligned}\omega=\left(\log\left(((\log n)/D)^{\frac{k_2-1}{k_1-1}}d\right)\right)^{1/(k_2-1)}&=\left(\log\left(\log n\right)^{(1-\beta)\frac{k-1}{k-2}}\right)^{1/(k-1)}\\&=\Omega\left((\log\log n)^{1/(k-1)}\right).\end{aligned}$$

∎

9.1.3 Proof of Theorem 9.1.4

9.1.3.1 Lower bound

We prove the lower bound in Theorem 9.1.4 in this section. The proof idea is similar to that used in the proof of Theorem 9.1.2, that is, we decompose an (n, k, ℓ)-omitting system into many different hypergraphs so that each hypergraph contains the information of a certain subset of edges in the original hypergraph. Then we use a probabilistic argument to show that there exists a large common independent set of these hypergraphs.

Recall that an n-vertex k-graph $\mathcal{H}$ is an (n, k, ℓ, λ)-omitting system iff it is $S_{\lambda+1}(\ell)$-free. While Theorem 9.1.4 as stated provides a lower bound on the independence number of (n, k, ℓ)-omitting systems, the result holds in the more general setting of (n, k, ℓ, λ)-omitting systems. We present the proof in this more general setting.

Let $k \geq k_0 > \ell \geq 1$, $\lambda \geq 2$, and $\mathcal{H}$ be an $S_\lambda(\ell)$-free k-graph. We say $\mathcal{H}$ is (k_0, λ)-indecomposable if

- $k = k_0$, or
- $k > k_0$ and $\mathcal{H}$ is $\{S_{\lambda_1}(k-1), \ldots, S_{\lambda_{k-k_0}}(k_0)\}$-free, where $\lambda_i = (k\lambda)^{2^{i-1}}$ for $i \in [k - k_0]$.

Otherwise, we say $\mathcal{H}$ is (k_0, λ)-decomposable.

Call a family $\mathcal{F}$ of hypergraphs (k_0, λ)-indecomposable if every member in it is (k_0, λ)-indecomposable. Otherwise, we say $\mathcal{F}$ is (k_0, λ)-decomposable.

The decomposition algorithm.

Input: An $S_\lambda(\ell)$-free k-graph $\mathcal{H}$ and a threshold k_0 with $k \geq k_0 > \ell$.

Output: A family $\mathcal{F}$ of $S_\lambda(\ell)$-free (k_0, λ)-indecomposable hypergraphs.

Operation: We start with the family $\mathcal{F} = \{\mathcal{H}\}$. If $\mathcal{F}$ is (k_0, λ)-indecomposable, then we terminate this algorithm. Otherwise, let $\mathcal{G} \in \mathcal{F}$ be a (k_0, λ)-decomposable hypergraph and let k' denote the size of each edge in $\mathcal{G}$. Let $i_0 \in \{1, \ldots, k' - k_0\}$ be the smallest integer such that $\mathcal{G}$ contains a copy of $S_{\lambda_{i_0}}(k' - i_0)$, where $\lambda_{i_0} = (k\lambda)^{2^{i_0-1}}$. Define

$$\mathcal{G}_{k'-i_0} = \left\{ A \in \binom{V(\mathcal{H})}{k' - i_0} : d_{\mathcal{G}}(A) \geq \lambda_{i_0} \right\} \quad \textit{and} \quad \mathcal{G}_{k'} = \left\{ B \in \mathcal{G} \colon \binom{B}{k' - i_0} \cap \mathcal{G}_{k'-i_0} = \emptyset \right\}.$$

Update $\mathcal{F}$ by removing $\mathcal{G}$ and adding $\mathcal{G}_{k'-i_0}$ and $\mathcal{G}_{k'}$. Repeat this operation until $\mathcal{F}$ is (k_0, λ)-indecomposable.

We need the following lemmas to show that the algorithm defined above always terminates. Write $\nu(\mathcal{H})$ for the size of a maximum matching in $\mathcal{H}$.

Lemma 9.1.20. *Let $\mathcal{H}$ be an $\{S_{\lambda_1}(k-1), \ldots, S_{\lambda_{k-1}}(1)\}$-free k-graph with m edges. Then*

$$\nu(\mathcal{H}) \geq \frac{m}{\prod_{i=1}^{k-1}(i+1)\lambda_i}.$$

Proof. For $j \in [k-1]$ let $\Lambda_j = \prod_{i=1}^{j}(i+1)\lambda_i$. We prove this lemma by induction on k. Suppose that $k = 2$. Since $\mathcal{H}$ is $S_{\lambda_1}(1)$-free, $d_{\mathcal{H}}(v) \leq \lambda_1 - 1$ for all $v \in V(\mathcal{H})$. Therefore, by greedily choosing an edge e and removing all edges that have nonempty intersection with e, we obtain at least $m/(2\lambda_1)$ pairwise disjoint edges in $\mathcal{H}$.

Now suppose that $k \geq 3$. We claim that $d_{\mathcal{H}}(v) \leq (\lambda_{k-1} - 1)\Lambda_{k-2}$ for all $v \in V(\mathcal{H})$. Indeed, suppose to the contrary that there exists $v_0 \in V(\mathcal{H})$ with $d_{\mathcal{H}}(v_0) \geq (\lambda_{k-1} - 1)\Lambda_{k-2} + 1$. Since

$\mathcal{H}$ is $\{S_{\lambda_1}(k-1), \dots, S_{\lambda_{k-2}}(2)\}$-free, the link $L_{\mathcal{H}}(v_0)$ is $\{S_{\lambda_1}(k-2), \dots, S_{\lambda_{k-2}}(1)\}$-free. By the induction hypothesis,

$$\nu(L_{\mathcal{H}}(v_0)) \geq \frac{(\lambda_{k-1}-1)\Lambda_{k-2}+1}{\Lambda_{k-2}} > \lambda_{k-1}-1,$$

but this contradicts the assumption that $\mathcal{H}$ is $S_{\lambda_{k-1}}(1)$-free. Therefore, $d_{\mathcal{H}}(v) \leq (\lambda_{k-1}-1)\Lambda_{k-2}$ for all $v \in V(\mathcal{H})$. Then, similar to the case of $k = 2$, by greedily choosing an edge e and removing all edges that have nonempty intersection with e, we obtain

$$\nu(\mathcal{H}) \geq \frac{m}{k(\lambda_{k-1}-1)\Lambda_{k-2}+1} > \frac{m}{\Lambda_{k-1}}$$

completing the proof. ∎

Let $\mathcal{H}$ be an $S_\lambda(\ell)$-free k-graph. Define

$$\mathcal{H}_{k-1} = \left\{ A \in \binom{V(\mathcal{H})}{k-1} : d_{\mathcal{H}}(A) \geq k\lambda \right\}.$$

If $\mathcal{H}$ is $\left\{S_{\lambda'_1}(k-1), \dots, S_{\lambda'_{k-k'-1}}(k'+1), S_\lambda(\ell)\right\}$-free for some $\ell < k' \leq k-2$, then also define

$$\mathcal{H}_{k'} = \left\{ A \in \binom{V}{k'} : d_{\mathcal{H}}(A) \geq k\lambda \prod_{i=1}^{k-k'-1} (i+1)\lambda'_i \right\}.$$

Lemma 9.1.21. *The hypergraphs $\mathcal{H}_{k'}$ and $\mathcal{H}_{k-1}$ defined above are $S_\lambda(\ell)$-free.*

Proof. We may only prove that $\mathcal{H}_{k'}$ is $S_\lambda(\ell)$-free, since the proof for $\mathcal{H}_{k-1}$ is basically the same. Suppose to the contrary that there exists $\{A_1, \ldots, A_\lambda\} \subset \mathcal{H}_{k'}$ forming a copy of $S_\lambda(\ell)$. Since $\mathcal{H}$ is $\{S_{\lambda'_1}(k-1), \ldots, S_{\lambda'_{k-k'-1}}(k'+1)\}$-free, the link $L_{\mathcal{H}}(A_i)$ is $\{S_{\lambda'_1}(k-k'-1), \ldots, S_{\lambda'_{k-k'-1}}(1)\}$-free for $i \in [\lambda]$. Let $\Lambda' = \prod_{i=1}^{k-k'-1}(i+1)\lambda'_i$. It follows from the definition of $\mathcal{H}_{k'}$ that $|L_{\mathcal{H}}(A_i)| \geq k\lambda\Lambda'$ for $i \in [\lambda]$. So, by Lemma 9.1.20, there are at least $k\lambda\Lambda'/\Lambda' \geq k\lambda$ pairwise disjoint edges in $L_{\mathcal{H}}(A_i)$ for $i \in [\lambda]$. Therefore, there exist λ pairwise disjoint $(k-k')$-sets $B_1, \ldots, B_\lambda$ such that $B_i \subset V \setminus \left(\bigcup_{i=1}^{\lambda} A_i\right)$ and $E_i = A_i \cup B_i \in \mathcal{H}$ for $i \in [\lambda]$. It is clear that $\{E_1, \ldots, E_\lambda\}$ is a copy of $S_\lambda(\ell)$ in $\mathcal{H}$, a contradiction. ∎

Recall that in the decomposition algorithm we defined

$$\mathcal{G}_{k'-i_0} = \left\{A \in \binom{V(\mathcal{H})}{k'-i_0} : d_{\mathcal{G}}(A) \geq \lambda_{i_0}\right\}, \quad and \quad \mathcal{G}_{k'} = \left\{B \in \mathcal{G} \colon \binom{B}{k'-i_0} \cap \mathcal{G}_{k'-i_0} = \emptyset\right\},$$

where $i_0 \in \{1, \ldots, k'-k_0\}$ is the smallest integer such that $\mathcal{G}$ contains a copy of $S_{\lambda_{i_0}}(k'-i_0)$ and $\lambda_{i_0} = (k\lambda)^{2^{i_0-1}}$. It is clear from the definition that $\mathcal{G}_{k'}$ is $S_{\lambda_{i_0}}(k'-i_0)$-free. On the other hand, Lemma 9.1.21 implies that both $\mathcal{G}_{k'}$ and $\mathcal{G}_{k'-i_0}$ are $S_\lambda(\ell)$-free. Therefore, the new hypergraphs $\mathcal{G}_{k'-i_0}$ and $\mathcal{G}_{k'}$ we added into $\mathcal{F}$ either have a smaller edge size (the case $\mathcal{G}_{k'-i_0}$) or forbid one more hypergraph (the case $\mathcal{G}_{k'}$). So the algorithm must terminate after finite many steps, and it is easy to see that the outputted family $\mathcal{F}$ has size at most 2^{k-k_0}. Indeed, the latter statement can be proved by associating a binary tree $T_{\mathcal{H}}$ to the algorithm: the vertex set of $T_{\mathcal{H}}$ is the collection of all hypergraphs (including $\mathcal{H}$) generated in each operation of the algorithm, the root of $T_{\mathcal{H}}$ is $\mathcal{H}$, and the children of a vertex $\mathcal{G}$ are $\mathcal{G}_{k'-i_0}$ and $\mathcal{G}_{k'}$ (if they are defined). It is easy to see that the height of $T_{\mathcal{H}}$ is at most $k-k_0$ and the outputted family $\mathcal{F}$ is the collection of hypergraphs that are leaf vertices of $T_{\mathcal{H}}$. Therefore, $|\mathcal{F}| \le 2^{k-k_0}$.

The following lemma shows that in order to find a large independent set in $\mathcal{H}$ it suffices to find a large common independent set of all hypergraphs in $\mathcal{F}$.

Lemma 9.1.22. *Let $\mathcal{H}$ be an $S_\lambda(\ell)$-free k-graph and $\mathcal{F}$ be the outputted family after applying the decomposition algorithm to $\mathcal{H}$. Then*

$$\alpha(\mathcal{H}) \ge \alpha\left(\bigcup_{\mathcal{G}\in\mathcal{F}} \mathcal{G}\right).$$

Proof. Suppose that $\mathcal{F} = \{\mathcal{H}_1, \dots, \mathcal{H}_m\}$ and $I \subset V(\mathcal{H})$ is independent in $\mathcal{H}_i$ for $i \in [m]$. It is clear from the definition that for every $E \in \mathcal{H}$ there is a subset $E' \subset E$ such that $E' \in \mathcal{H}_i$ for some $i \in [m]$. Since I is independent $\mathcal{H}_i$, $E' \not\subset I$ and it follows that $E \not\subset I$. Therefore, I is independent in $\mathcal{H}$. ∎

We also need the following lemma which gives an upper bound for the size of an indecomposable hypergraph.

Theorem 9.1.23 (Deza–Erdős–Frankl [52])**.** *Let $r \ge 1$, $t \ge 2$ be integers and $L = \{\ell_1, \dots, \ell_r\}$ be a set of integers with $0 \le \ell_1 < \cdots < \ell_r < k$. If an n-vertex k-graph $\mathcal{H}$ is $S_t(\ell)$-free for every $\ell \in [k] \setminus L$, then $|\mathcal{H}| = O(n^{r-1})$ unless $(\ell_2 - \ell_1) \mid \cdots \mid (\ell_r - \ell_{r-1}) \mid (k - \ell_r)$.*

Lemma 9.1.24. *Let $k \ge k_0 > \ell \ge 1$, $\lambda \ge 2$ be integers, $k > 2\ell + 1$, $k_0 \ge \ell + 3$, and $\mathcal{H}$ be a $S_\lambda(\ell)$-free (k_0, λ)-indecomposable k-graph with n vertices. Then there exists a constant $C_{k,\ell,\lambda}$ such that $|\mathcal{H}| \le C_{k,\ell,\lambda} n^{\min\{k_0-2, k-\ell-1\}}$.*

Proof. Since $\mathcal{H}$ is $S_\lambda(\ell)$-free and $k > 2\ell+1$, by the results in [99], $|\mathcal{H}| = O\left(n^{k-\ell-1}\right)$. On the other hand, since $\mathcal{H}$ is $\{S_{\lambda_1}(k-1), \ldots, S_{\lambda_{k-k_0}}(k_0), S_\lambda(\ell)\}$-free, applying Theorem 9.1.23 to $\mathcal{H}$ with $t = \max\{\lambda_1, \ldots, \lambda_{k-k_0}, \lambda\}$ and $L = \{0, 1, \ldots, \ell-1, \ell+1, \ldots, k_0-1\}$ we obtain $|\mathcal{H}| = O\left(n^{k_0-2}\right)$. ∎

Now we are ready to prove the lower bound in Theorem 9.1.4.

Proof of the lower bound in Theorem 9.1.4. We may assume that $k > 3\ell$ since otherwise by Equation 9.5 we are done. Let $\mathcal{H}$ be an $S_\lambda(\ell)$-free k-graph on n vertices and $V = V(\mathcal{H})$. Apply the decomposition algorithm to $\mathcal{H}$ with the threshold $k_0 = 2\ell+1$, and let $\mathcal{F}$ denote the outputted family. Suppose that $\mathcal{F} = \{\mathcal{H}_1, \ldots, \mathcal{H}_m\}$ for some integer m. For $i \in [m]$ let k_i denote the size of each edge in $\mathcal{H}_i$ and note from the definition of the algorithm that $2\ell+1 \le k_i \le k$. Let $C = \max\{C_{k_i,\ell,\lambda} \colon 2\ell+1 \le k_i \le k\}$, where $C_{k_i,\ell,\lambda}$ is the constant given by Lemma 9.1.24. Choose a set $I \subset V$ such that every vertex is included in I independently with probability $p = \delta n^{-\frac{2\ell-2}{3\ell-1}}$, where $\delta > 0$ is a small constant that satisfies $Cm\delta^{3\ell-2} \le 1/4$. Then by Lemma 9.1.24,

$$\begin{aligned}
\mathbb{E}\left[|I| - \sum_{i=1}^{m} |\mathcal{H}_i[I]|\right] &= \mathbb{E}[|I|] - \sum_{i=1}^{m} \mathbb{E}[|\mathcal{H}_i[I]|] \\
&\ge pn - \sum_{i=1}^{m} C p^{k_i} n^{\min\{2\ell-1, k_i-\ell-1\}} \\
&= pn - C\left(\sum_{i\in[m]:k_i\ge 3\ell} p^{k_i} n^{2\ell-1} + \sum_{i\in[m]:k_i\le 3\ell-1} p^{k_i} n^{k_i-\ell-1}\right) \\
&\ge \delta n^{\frac{\ell+1}{3\ell-1}} - Cm\delta^{3\ell} n^{\frac{\ell+1}{3\ell-1}} - Cm\delta^{3\ell-1} n^{\frac{\ell+1}{3\ell-1}} \ge \delta n^{\frac{\ell+1}{3\ell-1}}/2.
\end{aligned}$$

Therefore, there exists a set I of size $\Omega\left(n^{\frac{\ell+1}{3\ell-1}}\right)$ such that $\mathcal{H}_i[I]=\emptyset$ for $i\in[m]$, and it follow from Lemma 9.1.22 that $\alpha(\mathcal{H})\geq|I|=\Omega\left(n^{\frac{\ell+1}{3\ell-1}}\right)$. ∎

Remark. The lower bound $n^{\frac{\ell+1}{3\ell-1}}$ can be improved by optimizing the choice of k_0. Indeed, suppose that ℓ is sufficiently large. Let

$$k_0=\left(\frac{\sqrt{5}+1}{2}+o_\ell(1)\right)\ell,\quad s=\left(\frac{\sqrt{5}+3}{2}+o_\ell(1)\right)\ell,\quad \textit{and}\quad p=\delta n^{-\left(\frac{\sqrt{5}-1}{2}+o_\ell(1)\right)},$$

where $\delta>0$ is a sufficiently small constant. Repeating the argument above we obtain

$$\begin{aligned}\mathbb{E}\left[|I|-\sum_{i=1}^{m}|\mathcal{H}_i[I]|\right]&=\mathbb{E}[|I|]-\sum_{i=1}^{m}\mathbb{E}[|\mathcal{H}_i[I]|]\\&=pn-\left(\sum_{k_i\geq s}\mathbb{E}[|\mathcal{H}_i[I]|]+\sum_{2\ell+1\leq k_i\leq s}\mathbb{E}[|\mathcal{H}_i[I]|]+\sum_{k_0<k_i\leq 2\ell}\mathbb{E}[|\mathcal{H}_i[I]|]\right).\end{aligned}$$

By Lemma 9.1.24, we have

$$\sum_{k_i\geq s}\mathbb{E}[|\mathcal{H}_i[I]|]\leq C\sum_{k_i\geq s}p^{k_i}n^{k_0-2}.$$

By Theorem 9.1.1, we have

$$\sum_{2\ell+1\leq k_i\leq s}\mathbb{E}[|\mathcal{H}_i[I]|]\leq C\sum_{2\ell+1\leq k_i<s}p^{k_i}n^{k_i-\ell-1}\quad \textit{and}\quad \sum_{k_0<k_i\leq 2\ell}\mathbb{E}[|\mathcal{H}_i[I]|]\leq C\sum_{k_0\leq k_i\leq 2\ell}p^{k_i}n^{\ell}.$$

Therefore,

$$\begin{aligned}\mathbb{E}\left[|I| - \sum_{i=1}^{m}|\mathcal{H}_i[I]|\right] &\geq pn - C\left(\sum_{k_i \geq s} p^{k_i} n^{k_0-2} + \sum_{2\ell+1 \leq k_i < s} p^{k_i} n^{k_i-\ell-1} + \sum_{k_0 \leq k_i \leq 2\ell} p^{k_i} n^{\ell}\right) \\ &\geq pn - Cm\left(p^s n^{k_0-2} + p^{s-1} n^{s-\ell-2} + p^{k_0} n^{\ell}\right) \\ &\geq \delta n^{\left(\frac{3-\sqrt{5}}{2} + o_\ell(1)\right)}/2,\end{aligned}$$

which implies that $\mathcal{H}$ contains an independent set I of size $\Omega\left(n^{\left(\frac{3-\sqrt{5}}{2}+o_\ell(1)\right)}\right)$.

Similarly, the lower bound for $g(n,6,2)$ can be improved from $\Omega\left(n^{3/5}\right)$ to $\Omega\left(n^{2/3}\right)$ by letting $k_0 = 4$. Indeed, it is easy to see that when applying the decomposition algorithm to an n-vertex $S_\lambda(2)$-free 6-graph $\mathcal{H}$ with the threshold $k_0 = 4$, the outputted family $\mathcal{F}$ consists of three hypergraphs: an $S_\lambda(2)$-free $(4,\lambda)$-indecomposable 6-graph $\mathcal{H}_1$, an $S_\lambda(2)$-free $(4,\lambda)$-indecomposable 5-graph $\mathcal{H}_2$, and an $S_\lambda(2)$-free 4-graph $\mathcal{H}_3$. By Theorem 9.1.1 (the stronger version in [99]), $|\mathcal{H}_2| = O\left(n^2\right)$ and $|\mathcal{H}_3| = O\left(n^2\right)$. By Theorem 9.1.23, $\mathcal{H}_1 = O\left(n^2\right)$. So, it follows from a similar probabilistic argument as above that $\alpha(\mathcal{H}) \geq \alpha(\mathcal{H}_1 \cup \mathcal{H}_2 \cup \mathcal{H}_3) = \Omega\left(n^{2/3}\right)$.

9.1.3.2 Pseudorandom bipartite graphs

Our construction for the upper bound in Theorem 9.1.4 is related to some pseudorandom bipartite graphs, so it will be convenient to introduce some definitions and results related to pseudorandom bipartite graphs.

For a graph G on n vertices (assuming that $V(G) = [n]$) the adjacency matrix A_G of G is an $n \times n$ matrix whose (i, j)-th entry is

$$A_G(i,j) = \begin{cases} 1, & \text{if } \{i,j\} \in E(G), \\ 0, & otherwise. \end{cases}$$

Denote by $G(V_1, V_2)$ a bipartite graph with two parts V_1 and V_2, and that say $G(V_1, V_2)$ is (d_1, d_2)-regular if $d_G(v) = d_i$ for all $v \in V_i$ and $i = 1, 2$.

For a bipartite $G = G(V_1, V_2)$ denote by $\lambda(G)$ the second largest eigenvalue of A_G. Suppose that G is (d_1, d_2)-regular. Then we say G is pseudorandom if $\lambda(G) = O\left(\max\{\sqrt{d_1}, \sqrt{d_2}\}\right)$.

The Zarankiewicz number $z(m, n, s, t)$ is the maximum number of edges in a bipartite graph $G(V_1, V_2)$ with $|V_1| = m$, $|V_2| = n$ such that G contains no complete bipartite graph with s vertices in V_1 and t vertices in V_2.

Our construction of (n, k, ℓ)-systems is related to the lower bound (construction) for $z(m, n, s, t)$. More specifically, it is related to a construction defined by Alon, Mellinger, Mubayi and Verstraëte in [9], which was used to show that $z(n^{\ell/2}, n, 2, \ell) = \Omega\left(n^{(\ell+1)/2}\right)$.

Let q be a prime power and $\mathbb{F} = GF(q)$ be the finite field of size q. Denote by $\mathbb{F}[X]$ the collection of all polynomials over $\mathbb{F}$. The graph $G(q^\ell, q^2, 2, \ell)$ is a bipartite graph with two parts V_1 and V_2, where

$$V_1 = \{P(x) : P(x) \in \mathbb{F}[X], \deg(P(x)) \leq \ell - 1\}, \quad and \quad V_2 = \mathbb{F} \times \mathbb{F},$$

and for every $P(x) \in V_1$ and every $(x,y) \in V_2$, the pair $\{P(x),(x,y)\}$ is an edge in $G(q^\ell, q^2, 2, \ell)$ iff $y = P(x)$.

It is clear that $G(q^\ell, q^2, 2, \ell)$ does not contain a complete bipartite graph with two vertices in V_1 and ℓ vertices in V_2 since two distinct polynomials of degree at most $\ell-1$ over $\mathbb{F}$ can have the same value in at most $\ell-1$ points. It is also easy to see that $G(q^\ell, q^2, 2, \ell)$ is $(q, q^{\ell-1})$-regular.

The proof of the following result concerning the eigenvalues of $G(q^\ell, q^2, 2, \ell)$ can be found in [78].

Lemma 9.1.25 ([78])**.** *The eigenvalues of the adjacency matrix of $G(q^\ell, q^2, 2, \ell)$ are*

$$q^{\ell/2}, \underbrace{q^{(\ell-1)/2}, \ldots, q^{(\ell-1)/2}}_{q^2-q \text{ times}}, 0, \ldots, 0, \underbrace{-q^{(\ell-1)/2}, \ldots, -q^{(\ell-1)/2}}_{q^2-q \text{ times}}, -q^{\ell/2}.$$

In particular, $G(q^\ell, q^2, 2, \ell)$ is pseudorandom.

9.1.3.3 Upper bound

We prove the existence of (n,k,ℓ)-systems with independence number $O\left(n^{\frac{\ell+1}{2\ell}} (\log n)^{\frac{1}{\ell}}\right)$ in this section. Our construction is obtained from a random subgraph of the bipartite graph $G(q^\ell, q^2, 2, \ell)$ defined in the last section, and the method we used here is similar to that used in [149; 78].

First let us summarize the constructions used in [149] and [78] into a more general form.

Since we cannot ensure the random subgraph chosen from $G(q^\ell, q^2, 2, \ell)$ is exactly (d_1, d_2)-regular for some $d_1, d_2 \in \mathbb{N}$, it will be useful to consider the following more general setting.

Let $C, d_1, d_2 \geq 1$ be real numbers. A hypergraph $\mathcal{H}$ is

(a) (C,d_1)-uniform if $d_1/C \leq |E| \leq Cd_1$ for all $E \in \mathcal{H}$, and

(b) (C,d_2)-regular if $d_2/C \leq d_{\mathcal{H}}(v) \leq Cd_2$ for all $v \in V(\mathcal{H})$.

The edge density of a k-graph $\mathcal{H}$ with n vertices is $\rho(\mathcal{H}) = |\mathcal{H}|/\binom{n}{k}$. The bipartite incidence graph $G_{\mathcal{H}}$ of $\mathcal{H}$ is a bipartite graph with two parts $V_1 = E(\mathcal{H})$ and $V_2 = V(\mathcal{H})$, and for every $E \in E(\mathcal{H})$ and $v \in V(\mathcal{H})$ the pair $\{E, v\}$ is an edge in $G_{\mathcal{H}}$ iff $v \in E$. Denote by $A_{\mathcal{H}}$ the adjacency matrix of $G_{\mathcal{H}}$.

Let $n = |V(\mathcal{H})|$, $m = |\mathcal{H}|$ and labelling the edges in $\mathcal{H}$ with $E_1, \ldots, E_m$ We say a family $\mathcal{F}$ of hypergraphs fits $\mathcal{H}$ if $\mathcal{F} = \{\mathcal{G}_i : 1 \leq i \leq m\}$ and $\mathcal{G}_i$ is a hypergraph with $|V(\mathcal{G}_i)| = |E_i|$ for $i \in [m]$.

Given a hypergraph $\mathcal{H}$ and a family $\mathcal{F}$ that fits $\mathcal{H}$ we let $\mathcal{H}(\mathcal{F})$ be the random hypergraph obtained from $\mathcal{H}$ by taking independently for every $i \in [m]$ a bijection $\psi_i : E_i \to V(\mathcal{G}_i)$ and letting a set $S \subset E_i$ be an edge in $\mathcal{H}(\mathcal{F})$ if $\psi_i(S) \in \mathcal{G}_i$.

Let $\tau \geq 1$ be an integer and denote by $B_\tau(\mathcal{G})$ the collection of τ-subsets of $V(\mathcal{G})$ that are not independent in $\mathcal{G}$. Let $b_\tau(\mathcal{G}) = |B_\tau(\mathcal{G})|$ and $p_\tau(\mathcal{G}) = b_\tau(\mathcal{G})/\binom{v(\mathcal{G})}{\tau}$. In other words, $p_\tau(\mathcal{G})$ is the probability that a random τ-subset of $V(\mathcal{G})$ is not independent in $\mathcal{G}$. For a family $\mathcal{F}$ of hypergraphs define

$$p_\tau(\mathcal{F}) = \min\{p_\tau(\mathcal{G})\colon \mathcal{G} \in \mathcal{F}\}.$$

We extend the definition of $C_{\mathcal{G}}(2,j)$ in Section 9.1.2 by letting $C_{\mathcal{G}}(2,j)$ denote the number of pairs of edges $\{E, E'\}$ in a k-graph $\mathcal{G}$ with $|E \cap E'| = j$ for all $0 \leq j \leq k-1$.

The following lemma gives an upper bound for the independence number of $\mathcal{H}(\mathcal{F})$.

Lemma 9.1.26. *Let $C, d_1, d_2 \geq 1$ be real numbers and $k \geq 2$ be an integer. Suppose that $\mathcal{H}$ is a hypergraph with n vertices, m edges, and is (C, d_1)-uniform, (C, d_2)-regular. Let $\mathcal{F} = \{\mathcal{G}_i : i \in [m]\}$ be a family of k-graphs that fits $\mathcal{H}$. Suppose there exists $\lambda \geq 0$ such that the bipartite graph $G_{\mathcal{H}}$ satisfies*

$$\left| e_{G_{\mathcal{H}}}(X, Y) - \frac{d_1}{n}|X||Y| \right| \leq \lambda\sqrt{|X||Y|} \tag{9.6}$$

for all $X \subset V(\mathcal{H})$ and $Y \subset E(\mathcal{H})$. Then, w.h.p. $\alpha\,(H(\mathcal{F})) \leq 2\tau n/d_1$, if τ satisfies

$$\frac{p_\tau(\mathcal{F})}{\tau} \geq \frac{8C^2 \log n}{d_2} \quad \textit{and} \quad \tau \geq \frac{8C^2\lambda^2}{d_2}. \tag{9.7}$$

Proof. Let τ be a real number that satisfies Equation 9.7, $V = V(\mathcal{H})$, and $I \subset V$ be a set of size $2\tau n/d_1$.

Let $m = |\mathcal{H}|$ and label the edges in $\mathcal{H}$ by $\{E_1, \ldots, E_m\}$. Let $m_i = |E_i|$ for $i \in [m]$. Since $\mathcal{H}$ is (C, d_1)-uniform and (C, d_2)-regular, we obtain $d_1|\mathcal{H}|/C \leq \sum_{v \in V} d_{\mathcal{H}}(v) \leq Cd_1|\mathcal{H}|$, and consequently,

$$md_1/C^2 \leq nd_2 \leq C^2 md_1. \tag{9.8}$$

Define

$$\mathcal{E}_1 = \{E \in \mathcal{H} \colon |E \cap I| < \tau\}\,, \quad \text{and} \quad \mathcal{E}_2 = \{E \in \mathcal{H} \colon |E \cap I| > 3\tau\}\,.$$

Claim 9.1.27. *$|\mathcal{E}_i| \le 2C^2\lambda^2 m/d_2\tau \le m/4$ for $i = 1, 2$.*

Proof. It follows from Equation 9.6 that

$$\sum_{E \in \mathcal{E}_1} |E \cap I| = e_{G_{\mathcal{H}}}(I, \mathcal{E}_1) \ge d_1|I||\mathcal{E}_1|/n - \lambda\left(|I||\mathcal{E}_1|\right)^{1/2},$$

and by definition, $\sum_{E\in\mathcal{E}_1} |E \cap I| < \tau|\mathcal{E}_1|$. Therefore,

$$\tau|\mathcal{E}_1| > d_1|I||\mathcal{E}_1|/n - \lambda\left(|I||\mathcal{E}_1|\right)^{1/2}.$$

Since $|I| = 2\tau n/d_1$, we obtain

$$|\mathcal{E}_1| < \left(\frac{\lambda|I|^{1/2}}{d_1|I|/n - \tau}\right)^2 = \left(\frac{\lambda|I|^{1/2}}{d_1|I|/2n}\right)^2 = \frac{2\lambda^2 n}{\tau d_1},$$

which together with Equation 9.8 implies $|\mathcal{E}_1| < 2C^2\lambda^2 m/d_2\tau$. Notice that Equation 9.7 implies that $C^2\lambda^2/d_2\tau \le 1/8$, so $|\mathcal{E}_1| < m/4$.

Now consider $\mathcal{E}_2$. Similarly, By Equation 9.6,

$$\sum_{E \in \mathcal{E}_2} |E \cap I| = e_{G_{\mathcal{H}}}(I, \mathcal{E}_2) \le d_1|I||\mathcal{E}_1|/n + \lambda\left(|I||\mathcal{E}_1|\right)^{1/2},$$

and by definition, $\sum_{E\in\mathcal{E}_2} |E \cap I| > 3\tau|\mathcal{E}_2|$. Therefore,

$$3\tau|\mathcal{E}_2| < d_1|I||\mathcal{E}_2|/n + \lambda\left(|I||\mathcal{E}_2|\right)^{1/2},$$

Since $|I| = 2\tau n/d_1$, we obtain

$$|\mathcal{E}_2| < \left(\frac{\lambda|I|^{1/2}}{3\tau - d_1|I|/n}\right)^2 = \left(\frac{\lambda|I|^{1/2}}{d_1|I|/2n}\right)^2 = \frac{2\lambda^2 n}{\tau d_1} \leq \frac{2C^2\lambda^2 m}{\tau d_2} \leq \frac{m}{4}.$$

∎

For $i \in [m]$ let $I_i = I \cap E_i$. By Claim 9.1.27 the number of set I_i that satisfies $\tau \leq |I_i| \leq 3\tau$ (in fact, $|I_i| \geq \tau$ is sufficient for the proof) is at least $m - 2m/4 = m/2$. By the definition of $p_\tau(\mathcal{F})$, for every I_i that satisfies $\tau \leq |I_i| \leq 3\tau$ we have

$$P\left(\psi_i(I_i) \text{ is independent in } \mathcal{G}_i\right) \leq 1 - p_\tau(\mathcal{F}).$$

Since, by definition, the bijections $\{\psi_i \colon i \in [m]\}$ are mutually independent, the events

$$\{\psi_i(I_i) \text{ is independent in } \mathcal{G}_i \colon i \in [m]\}$$

are mutually independent. Therefore,

$$P\left(I \text{ is independent in } \mathcal{H}(\mathcal{F})\right) \leq P\left(\bigwedge_{i\in[m]} \psi_i(I_i) \text{ is independent in } \mathcal{G}_i\right)$$
$$= \prod_{i\in[m]} P\left(\psi_i(I_i) \text{ is independent in } \mathcal{G}_i\right) \leq \left(1 - p_\tau(\mathcal{F})\right)^{m/2}.$$

So the expected number of independent $2n\tau/d_1$-sets in $\mathcal{H}(\mathcal{G})$ is at most

$$\begin{aligned}(1 - p_\tau(\mathcal{F}))^{m/2}\binom{n}{2n\tau/d_1} &< \exp\left(-p_\tau(\mathcal{F})\frac{m}{2} + \frac{2\tau n}{d_1}\log\left(\frac{en}{2n\tau/d_1}\right)\right)\\ &< \exp\left(-p_\tau(\mathcal{F})\frac{m}{2} + \frac{2C^2\tau m}{d_2}\log n\right)\\ &< \exp\left(-p_\tau(\mathcal{F})\frac{m}{4}\right) \to 0 \quad as \quad m \to \infty.\end{aligned}$$

Therefore, $\alpha(\mathcal{H}(\mathcal{F})) \leq 2\tau n/d_1$ holds with high probability. ∎

The following corollary may be a simpler form to use Lemma 9.1.26.

Corollary 9.1.28. *Let $C, d_1, d_2 \geq 1$ be real numbers and $k \geq 2$ be an integer. Suppose that $\mathcal{H}$ is a hypergraph with n vertices, m edges, and is (C, d_1)-uniform, (C, d_2)-regular. Let $\mathcal{F} = \{\mathcal{G}_i : i \in [m]\}$ be a family of k-graphs that fits $\mathcal{H}$. Suppose there exists $\lambda \geq 0$ such that the bipartite graph $G_{\mathcal{H}}$ satisfies Equation 9.6 for all $X \subset V(\mathcal{H})$ and $Y \subset E(\mathcal{H})$. Suppose further that*

- *there exists $\rho > 0$ such that $\rho(\mathcal{G}_i) \geq \rho$ for $i \in [m]$, and*
- *$\lambda < (d_2\tau/8C^2)^{1/2}$, where $\tau = 2\left(16k!C^2 \log n/\rho d_2\right)^{1/(k-1)} \gg 1$, and*
- *$C_{\mathcal{G}_i}(2, j) \leq |\mathcal{G}_i| \left(v(\mathcal{G}_i)/3\tau\right)^{k-j}$ for $0 \leq j \leq k-1$ and $i \in [m]$.*

Then, w.h.p. $\alpha(\mathcal{H}(\mathcal{F})) \le 2\tau n/d_1$.

Proof. It suffices to show that $\tau = 2\left(16k!C^2 \log n/\rho d_2\right)^{1/(k-1)}$ satisfies Equation 9.7. First let us calculate $p_\tau(\mathcal{F})$. Fix $i \in [m]$ and for every edge set $\mathcal{E} \subset \mathcal{G}_i$ let $T_{\mathcal{E}}$ denote the collection of τ-sets in $V(\mathcal{G}_i)$ containing the vertex set $\bigcup_{E\in\mathcal{E}} E$. By the definition of $B_\tau(\mathcal{G}_i)$, we have

$$B_\tau(\mathcal{G}_i) = \bigcup_{E\in\mathcal{G}_i} T_{\{E\}}.$$

It follows from the Bonferroni inequalities [27] that

$$\begin{aligned} b_\tau(\mathcal{G}_i) = \left|\bigcup_{E\in\mathcal{G}_i} T_{\{E\}}\right| &\ge \sum_{E\in\mathcal{G}_i} |T_{\{E\}}| - \sum_{\{E,E'\}\in\binom{\mathcal{G}_i}{2}} |T_{\{E,E'\}}| \\ &= |\mathcal{G}_i|\binom{v(\mathcal{G}_i)-k}{\tau-k} - \sum_{j=0}^{k-1} C_{\mathcal{G}_i}(2,j)\cdot\binom{v(\mathcal{G}_i)-2k+j}{\tau-2k+j} \end{aligned}$$

Since $C_{\mathcal{G}_i}(2,j) \le |\mathcal{G}_i|\left(v(\mathcal{G}_i)/\tau\right)^{k-j}$ for $0 \le j \le k-1$, we obtain

$$\begin{aligned} \sum_{j=0}^{k-1} C_{\mathcal{G}_i}(2,j)\cdot\binom{v(\mathcal{G}_i)-2k+j}{\tau-2k+j} &\le \sum_{j=0}^{k-1} |\mathcal{G}_i|\left(\frac{v(\mathcal{G}_i)}{3\tau}\right)^{k-j}\binom{v(\mathcal{G}_i)-2k+j}{\tau-2k+j} \\ &= \sum_{j=0}^{k-1} |\mathcal{G}_i|\left(\frac{v(\mathcal{G}_i)}{3\tau}\right)^{k-j}\frac{(\tau-k)_{k-j}}{(v(\mathcal{G}_i)-k)_{k-j}}\binom{v(\mathcal{G}_i)-k}{\tau-k} \\ &\le \sum_{j=0}^{k-1} |\mathcal{G}_i|\left(\frac{v(\mathcal{G}_i)}{3\tau}\right)^{k-j}\left(\frac{\tau}{v(\mathcal{G}_i)}\right)^{k-j}\binom{v(\mathcal{G}_i)-k}{\tau-k} \\ &\le \frac{1}{2}|\mathcal{G}_i|\binom{v(\mathcal{G}_i)-k}{\tau-k}. \end{aligned}$$

Therefore, $b_\tau(\mathcal{G}_i) \geq \frac{1}{2}|\mathcal{G}_i|\binom{v(\mathcal{G}_i)-k}{\tau-k}$. Consequently,

$$p_\tau(\mathcal{G}_i) = \frac{b_\tau(\mathcal{G}_i)}{\binom{v(\mathcal{G}_i)}{\tau}} \geq \frac{1}{2}\frac{|\mathcal{G}_i|\binom{v(\mathcal{G}_i)-k}{\tau-k}}{\binom{v(\mathcal{G}_i)}{\tau}} = \frac{1}{2}\frac{|\mathcal{G}_i|}{\binom{v(\mathcal{G}_i)}{k}}\frac{(\tau)_k}{k!} = \frac{\rho(\mathcal{G}_i)}{2}\frac{(\tau)_k}{k!} \geq \frac{\rho}{2k!}(\tau)_k.$$

So we obtain

$$\frac{p_\tau(\mathcal{F})}{\tau} \geq \frac{\rho}{2k!}(\tau-1)_{k-1} \geq \frac{\rho}{2k!}\left(\frac{\tau}{2}\right)^{k-1} \geq \frac{8C^2\log n}{d_2}.$$

On the other hand, our assumption on λ clearly implies $\tau \geq 8C^2\lambda/d_2$. Therefore, by Lemma 9.1.26, *w.h.p.* $\alpha\left(\mathcal{H}(\mathcal{F})\right) \leq 2\tau n/d_1$. ∎

We will also need the following result in our proof.

Lemma 9.1.29 ([149])**.** *Let $\mathcal{H}$ be a d_1-uniform d_2-regular hypergraph on n vertices. Then for every $V' \subset V(\mathcal{H})$ and $\mathcal{E} \subset E(\mathcal{H})$,*

$$\left|\sum_{E\in\mathcal{E}}|E\cap V'| - \frac{d_1}{n}|V'||\mathcal{E}|\right| \leq \lambda(G_{\mathcal{H}})\sqrt{|V'||\mathcal{E}|}.$$

We also need the following Chernoff's inequality (e.g. see Theorem 22.6 in [107]).

Theorem 9.1.30 (Chernoff's inequality)**.** *Suppose that* $S_n = X_1 + \cdots + X_n$ *where* $0 \leq X_i \leq 1$ *for* $i \in [n]$ *are independent random variables. Let* $\mu = \mathbb{E}[X_1] + \cdots + \mathbb{E}[X_n]$*. Then for every* $0 \leq t \leq \mu$*,*

$$P\left(|S_n - \mu| \geq t\right) \leq e^{-\frac{t^2}{3\mu}}.$$

Now we are ready to prove the upper bound in Theorem 9.1.4.

Proof of the upper bound in Theorem 9.1.4. Let q be a prime power and $G = G(q^\ell, q^2, 2, \ell)$ be the bipartite graph on $V_1 \cup V_2$ with $|V_1| = q^\ell$ and $|V_2| = q^2$. Let $\mathcal{G}$ denote the hypergraph on q^2 vertices whose bipartite incident graph is G. Note that $\mathcal{G}$ is a $q^{\ell-1}$-regular q-graph, and by Lemmas 9.1.25 and 9.1.29,

$$\left| \sum_{E \in \mathcal{E}} |E \cap V'| - \frac{1}{q}|V'||\mathcal{E}| \right| \leq q^{(\ell-1)/2}\sqrt{|V'||\mathcal{E}|} \tag{9.9}$$

holds for all $V' \subset V(\mathcal{G})$ and $\mathcal{E} \subset \mathcal{G}$.

Let $U \subset V(\mathcal{G})$ be a random set such that every vertex in $V(\mathcal{G})$ is included in U independently with probability $p = q^{-\frac{2}{\ell+1}}$. Then $\mathbb{E}[|U|] = pq^2 = q^{\frac{2\ell}{\ell+1}}$, and by the Chernoff inequality,

$$P\left(\left||U| - pq^2\right| > pq^2/2\right) < e^{-\frac{\left(pq^2/2\right)^2}{3pq^2}} = e^{-pq^2/12} \to 0 \quad as \quad q \to \infty.$$

For every $E \in \mathcal{G}$ we have $\mathbb{E}[|E \cap U|] = pq = q^{\frac{\ell-1}{\ell+1}}$, and by the Chernoff inequality,

$$P\left(||E \cap U| - pq| > pq/2\right) < e^{-\frac{(pq/2)^2}{3pd_1}} = e^{-pq/12}.$$

Let B denote the collection of edges $E \in \mathcal{G}$ such that $||E \cap U| - pq| > pq/2$. Then

$$\mathbb{E}[|B|] \leq q^{\ell} e^{-pq/12} = q^{\ell} e^{-q^{\frac{\ell-1}{\ell+1}}/12} \to 0 \quad as \quad q \to \infty.$$

Therefore, *w.h.p.* the set U satisfies that $q^{\frac{2\ell}{\ell+1}}/2 \leq |U| \leq 3q^{\frac{2\ell}{\ell+1}}/2$ and $q^{\frac{\ell-1}{\ell+1}}/2 \leq |E \cap U| \leq 3q^{\frac{\ell-1}{\ell+1}}/2$ for all $E \in \mathcal{G}$.

Fix such a set U that satisfies the conclusion above, and let $c \in [1/2, 3/2]$ be the real number such that $|U| = cq^{\frac{2\ell}{\ell+1}}$. Let $n = |U| = cq^{\frac{2\ell}{\ell+1}}$, $m = |\mathcal{G}| = q^{\ell}$, $d_1 = q^{\frac{\ell-1}{\ell+1}}$, and $d_2 = q^{\ell-1}$. Let $\mathcal{H}$ be the hypergraph on U with

$$\mathcal{H} = \{E \cap U \colon E \in \mathcal{G}\}.$$

Since $d_1/2 \leq |E \cap U| \leq 3d_1/2$ for all $E \in \mathcal{G}$, the hypergraph $\mathcal{H}$ is a $(2, d_1)$-uniform. Moreover, for every pair of edges $E, E' \in \mathcal{G}$, since $|E \cap E'| < \ell < d_1/2$, we have $E \cap U \neq E' \cap U$. So, $d_{\mathcal{H}}(u) = d_{\mathcal{G}}(u) = d_2$ for all $u \in U$. In addition, Equation 9.9 also holds for all $V' \subset U$ and $\mathcal{E} \subset \mathcal{G}$.

Label the edges in $\mathcal{H}$ with $\{E_1, \ldots, E_m\}$ and let $m_i = |E_i|$ for $i \in [m]$. Let $\mathcal{F} = \{\mathcal{S}_i \colon i \in [m]\}$, where $\mathcal{S}_i$ is the k-graph on $[m_i]$ whose edge set is the collection of all k-subsets of $[m_i]$ that contain $[\ell+1]$. Our construction of the (n,k,ℓ)-omitting system is simply $\mathcal{H}(k,\ell)=\mathcal{H}(\mathcal{F})$, and indeed, one can easily check that $|e'\cap e'|\neq \ell$ for all distinct edges $e,e'\in\mathcal{H}(k,\ell)$.

Let $\tau = 100\,(\log n)^{1/\ell}$.

Claim 9.1.31. $p_\tau(\mathcal{F}) \geq \left(\frac{\tau}{3d_1/2}\right)^{\ell+1}/2$.

Proof. Fix $i\in[m]$ and let I be a random τ-subset of $[m_i]$. It is easy to see that I is not independent in $\mathcal{S}_i$ iff $[\ell+1]\subset I$. Since

$$P\left([\ell+1]\subset I\right) = \frac{\binom{m_i-\ell-1}{\tau-\ell-1}}{\binom{m_i}{\tau}} = \frac{\tau\cdots(\tau-\ell)}{m_i\cdots(m_i-\ell)} \geq (1-o(1))\left(\frac{\tau}{m_i}\right)^{\ell+1} > \frac{1}{2}\left(\frac{\tau}{3d_i/2}\right)^{\ell+1},$$

we obtain

$$p_\tau(\mathcal{F}) > \frac{1}{2}\left(\frac{\tau}{3d_i/2}\right)^{\ell+1}.$$

∎

Observe that τ satisfies

$$\frac{p_\tau(\mathcal{F})}{\tau} > \frac{\left(\frac{\tau}{3d_1/2}\right)^{\ell+1}/2}{\tau} = \frac{100^\ell \log n}{2(3/2)^{\ell+1}d_1^{\ell+1}} = \frac{100^\ell}{2(3/2)^{\ell+1}}\frac{\log n}{d_2} > \frac{32\log n}{d_2}$$

(here we used the fact that $d_2 = d_1^{\ell+1}$) and

$$\tau = 100(\log n)^{1/\ell} > \frac{32\left(q^{(\ell-1)/2}\right)^2}{q^{\ell-1}}.$$

We may therefore apply Lemma 9.1.26 with $C = 2$ to obtain

$$\alpha\left(\mathcal{H}(k,\ell)\right) \leq 2\tau n/d_1 \leq 400 n^{\frac{\ell+1}{2\ell}} (\log n)^{1/\ell}.$$

∎

9.1.4 Independent sets in (n,k,ℓ,λ)-systems

In this section we prove Theorem 9.1.6. Our proof is a direct application of Theorem 9.1.13.

Proof of Theorem 9.1.6. Fix $\delta > 0$, and let $\epsilon > 0$ be sufficiently small such that $\frac{\ell-1}{k-2} - \delta < \frac{(\ell-1)(1-\epsilon)}{k-2+\epsilon}$ holds. Let $t = \lambda^{\frac{1}{k-1}} n^{\frac{\ell-1}{k-1}}$ and $\mathcal{H}$ be a (n,k,ℓ,λ)-system, where $0 < \lambda < n^{\frac{\ell-1}{k-2}-\delta}$.

Let $j \in [\ell-1]$ and $S \subset V(\mathcal{H})$ be a set of size j. Since $\mathcal{H}$ is an (n,k,ℓ,λ)-system, $L_{\mathcal{H}}(S)$ is an $(n,k-j,\ell-j,\lambda)$-system. Therefore,

$$\begin{aligned} \Delta(\mathcal{H}) &\leq \lambda \binom{n}{\ell-1} \Big/ \binom{k-1}{\ell-1} < t^{k-1}, \quad \textit{and} \\ |\Delta_j(\mathcal{H})| &\leq \lambda \binom{n}{\ell-j} \Big/ \binom{k-j}{\ell-j} = O(\lambda n^{\ell-j}) \quad \textit{for} \quad 2 \leq j \leq \ell-1. \end{aligned}$$

It follows that

$$C_{\mathcal{H}}(2,j) = O\left(\lambda n^{\ell-j}|\mathcal{H}|\right) = O\left(\lambda^2 n^{2\ell-j}\right) \leq n t^{2k-j-1-\epsilon} \quad \textit{for} \quad 2 \leq j \leq \ell-1.$$

On the other hand, for $\ell \leq j' \leq k-1$ and a set $S \subset V(\mathcal{H})$ of size j' the link $L_{\mathcal{H}}(S)$ has size at most λ. Therefore,

$$C_{\mathcal{H}}(2,j') = O(\lambda|\mathcal{H}|) = O\left(\lambda^2 n^{\ell}\right) \leq nt^{2k-j'-1-\epsilon} \quad for \quad \ell \leq j' \leq k-1.$$

Therefore, by Theorem 9.1.13, $\alpha(\mathcal{H}) = \Omega\left((\log t)^{1/(k-1)} n/t\right) = \Omega\left(\lambda^{-\frac{1}{k-1}} n^{\frac{k-\ell}{k-1}} (\log n)^{\frac{1}{k-1}}\right)$. ∎

9.1.5 The Ramsey number of the k-Fan

In this section we prove Theorem 9.1.10. The lower bound (construction) is given by the so called L-constructions. These were introduced in [44], where they were used to answer an old Ramsey-type question of Ajtai–Erdős–Komlós–Szemerédi [3].

Let $m, n \geq 2$ and let $\mathcal{L}_{m,n}$ be the k-graph with vertex set $[m] \times [n]$ and edge set

$$\{\{(x_1,y_1),(x_1,y_2),\ldots,(x_{k-1},y_2)\} \colon x_1 < \cdots < x_{k-1}, y_1 > y_2\}.$$

Proposition 9.1.32. *For every $m, n \geq 2$ the hypergraph $\mathcal{L}_{m,n}$ is F^k-free.*

Proof. Suppose that $\mathcal{L}_{m,n}$ contains a copy of $F^k = \{E_1, \ldots, E_k, E\}$. Let $v = \bigcap_{i=1}^k E_i$ and assume that $v = (x_0, y_0)$, $E = \{(x_1,y_1),(x_1,y_2),\ldots,(x_{k-1},y_2)\}$, where $x_1 < \cdots < x_{k-1}$ and $y_1 > y_2$.

By the definition of F^k, for every vertex $u \in E$, there exists an edge E_i that contains both u and v. It is easy to see that if $x_1' < x_2'$ and $y_1' < y_2'$, then there is no edge in $\mathcal{L}_{m,n}$ containing both (x_1', y_1') and (x_2', y_2'). Therefore, we must have (see Figure 29)

(1) $x_0 \le x_1$ and $y_0 \ge y_1$, or

(2) $x_0 \ge x_{k-1}$ and $y_0 \le y_2$, or

(3) $x_0 = x_1$ and $y_2 < y_0 < y_1$, or

(4) $y_0 = y_2$ and $x_1 < x_0 < x_{k-1}$.

If $x_0 \le x_1$ and $y_0 \ge y_1$, then by the definition of $\mathcal{L}_{m,n}$, there is a $(k-1)$-set $J \subset [k]$ such that $\bigcap_{j \in J} E_j = (x_0, y_2)$, a contradiction. If $x_0 \ge x_{k-1}$ and $y_0 \le y_2$, then by the definition of $\mathcal{L}_{m,n}$, there exist $\{i, j\} \subset [k]$ such that $E_i \cap E_j = (x_1, y_0)$, a contradiction. Similarly, if Case (3) or Case (4) happens, then there exist $\{i, j\} \subset [k]$ such that $E_i \cap E_j = (x_1, y_2)$, which is a contradiction. ∎

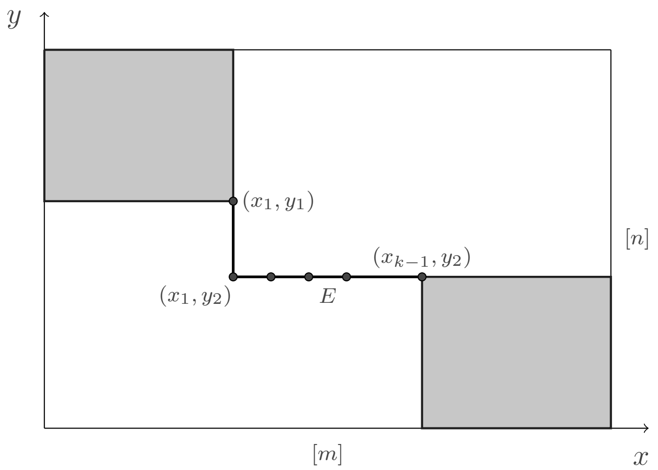


Figure 29. Only vertices that lie in these two shaded areas and the L-shaped path that connects these two areas can be adjacent to all vertices in E.

The following result gives an upper bound for the independence number of $\mathcal{L}_{m,n}$.

Proposition 9.1.33. *The hypergraph $\mathcal{L}_{m,n}$ satisfies $\alpha(\mathcal{L}_{m,n}) < m + (k-2)n$.*

Proof. Let I be an independent in $\mathcal{L}_{m,n}$. Remove the topmost vertex of each column and the $k-2$ rightmost vertices of each row in I. It is easy to see that we removed at most $m+(k-2)n$ vertices from I, and I has no vertex left since otherwise I would contain an edge in $\mathcal{L}_{m,n}$. Therefore, $\alpha(\mathcal{L}_{m,n}) < m + (k-2)n$. ∎

Now we finish the proof of Theorem 9.1.10.

Proof of Theorem 9.1.10. First we prove the lower bound. Let $m = \lfloor \frac{t}{2} \rfloor$ and $n = \lfloor \frac{t-1}{2(k-2)} \rfloor$. By Propositions 9.1.32 and 9.1.33, the k-graph $\mathcal{L}_{m,n}$ is F^k-free and $\alpha(\mathcal{L}_{m,n}) \leq m + (k-2)n < t$. So,

$$r_k(F^k, t) > mn = \left\lfloor \frac{t}{2} \right\rfloor \left\lfloor \frac{t-1}{2(k-2)} \right\rfloor .$$

To prove the upper bound, let us show that $r_k(F^k, t) \leq r_k(S_t^k, t)$ first. Indeed, let $\mathcal{H}$ be a k-graph on $r_k(S_t^k, t)$ vertices. We may assume that $\mathcal{H}$ does not contain an independent set of size t. Then, there exist t distinct edges $E_1, \ldots, E_t$ and a vertex v in $\mathcal{H}$ such that $E_i \cap E_j = \{v\}$ for $1 \leq i < j \leq t$. Let S be a set that contains exactly one vertex from each $E_i \setminus \{v\}$ for $i \in [t]$. Then S has size t and hence contains an edge in $\mathcal{H}$, and it implies that $\mathcal{H}$ contains a copy of F^k. So $r_k(F^k, t) \leq r_k(S_t^k, t)$, and it follows from Theorem 9.1.9 that $r_k(F^k, t) \leq t(t-1) + 1$. ∎

9.2 Explicit constructions of designs

9.2.1 Introduction

We start with the following definition.

Definition 9.2.1. *For fixed integers $r \geq s \geq 1$ we say there is an explicitly construction of an (n, r, s)-system with property $\mathcal{P}$ if there exists an algorithm $\mathcal{A}$ such that for every integer n as input, $\mathcal{A}$ runs in time poly(n) and outputs an (n, r, s)-system with property $\mathcal{P}$.*

Explicit constructions of (n, r, s)-systems with certain properties are very useful in theoretical computer science. For example, in the seminal work of Nisan and Wigderson [202], dense (n, r, s)-systems are used to construct pseudorandom generators (PRGs) (see also [240; 213] for more applications). More recently, explicit constructions of (n, r, s)-systems with small independence number were used to construct extractors for adversarial sources [36; 35].

In this note, we focus on the explicit constructions of (n, r, s)-systems with small independence number. Rödl and Šiňajová's proof of the existence of an (n, r, s)-system with small independence number uses the Lovász local lemma, and hence it does not provide an explicit way to construct them. Perhaps the first explicit construction of an $(n, 3, 2)$-system (also called a Steiner triple system) with independence number $O(n^{1-\epsilon})$ for some absolute constant $\epsilon > 0$ is due to Chattopadhyay, Goodman, Goyal, and Li [36]. Their proof uses results about cap sets (see [45; 55]).

Theorem 9.2.2 (Chattopadhyay–Goodman–Goyal–Li [36]). *There exists a constant $C \geq 1$ such that for every integer $n \geq 3$ there exists an explicit construction of an $(n, 3, 2)$-system with independence number at most $Cn^{0.9228}$.*

Later, using results about linear codes [125; 29] and Sidorenko's recent bounds on the size of sets in $\mathbb{Z}_2^n$ containing no r elements that sum to zero [229; 230], Chattopadhyay and Goodman [35] extended Theorem 9.2.2 to all integers $r > s \geq 2$ with $s \geq \lceil r/2 \rceil$.

Theorem 9.2.3 (Chattopadhyay–Goodman [35]). *There exists a constant $C \geq 1$ such that for every integer $s \geq 2$ and every even integer $r > s$ there exists an explicit construction of an (n, r, s)-system with independence number at most $Cr^4 n^{\frac{2(r-s)}{r}}$.*

Remark. For odd r they showed that there exists an explicit construction of an (n, r, s)-system with independence number at most $C(r+1)^4 n^{\frac{2(r+1-s)}{r+1}}$.

Our main results in this note extend Theorem 9.2.2 for certain values of r and s in the range $s < \lceil r/2 \rceil$ which was not addressed by Theorem9.2.3.

Our proof of the first theorem below is based on a recent result about the maximum size of a set in $\mathbb{Z}_6^n$ that avoids 6-term arithmetic progressions [207].

Theorem 9.2.4. *There exists a constant $C > 0$ such that for every integer $r \in \{4, 5, 6\}$ and every integer $n \geq r$ there exists an explicit construction of an $(n, r, 2)$-system $\mathcal{H}$ with $\alpha(\mathcal{H}) \leq Cn^{0.973}$.*

Using a lemma about the independence number of the product of two hypergraphs we are able to extend Theorem 9.2.4 to a wider range of r and s.

For every integer $s = 3^{\ell_1}4^{\ell_2}5^{\ell_3}6^{\ell_4} + 1$, where $\ell_1, \ell_2, \ell_3, \ell_4 \geq 0$ are integers, define

$$R(s) = \begin{cases} 6(s-1) & if \quad \ell_1 = \ell_2 = \ell_3 = 0 \\ 5(s-1) & if \quad \ell_1 = \ell_2 = 0 \quad and \quad \ell_3 \neq 0 \\ 4(s-1) & if \quad \ell_1 = 0 \quad and \quad \ell_2 \neq 0 \\ 3(s-1) & if \quad \ell_1 \neq 0 \end{cases}$$

Theorem 9.2.5. *For every integer s of the form $3^{\ell_1}4^{\ell_2}5^{\ell_3}6^{\ell_4}+1$, where $\ell_1, \ell_2, \ell_3, \ell_4 \geq 0$ are integers, and every integer r satisfying $2s \leq r \leq R(s)$ there exist constants $C = C(\ell_1, \ell_2, \ell_3, \ell_4), \epsilon = \epsilon(\ell_1, \ell_2, \ell_3, \ell_4) > 0$ such that for every integer $n \geq r$ there exists an explicit construction of an (n, r, s)-system with independence number at most $Cn^{1-\epsilon}$.*

The following result focusing on $(n, 5, 4)$-systems uses a different argument and it improves the bound $O(n^{2/3})$ given by Theorem 9.2.3.

Theorem 9.2.6. *There exists a constant $C > 0$ such that for every integer $n \geq 5$ there exists an explicit construction of an $(n, 5, 4)$-systems with independence number at most $Cn^{\log_3 2} \leq Cn^{0.631}$.*

9.2.2 Proofs of Theorems 9.2.4 and 9.2.5

9.2.2.1 Proof of Theorems 9.2.4

Let us first introduce a construction of r-graphs based on r-term arithmetic progressions (r-AP) over $\mathbb{Z}_r^k$. We do not allow trivial progressions so an r-AP has r distinct elements.

Construction $\mathcal{A}(r,k)$. Let $r \geq 3$ and $k \geq 1$ be integers. The hypergraph $\mathcal{A}(r,k)$ is the r-graph with vertex set $V = \mathbb{Z}_r^k$ and edge set

$$\left\{ \{v_1, \ldots, v_r\} \in \binom{V}{r} : v_1, \ldots, v_r \text{ form an } r\text{-AP} \right\}.$$

Remarks.

- It is clear that $\mathcal{A}(r,k)$ can be constructed in time *poly* $\left(r^k\right)$ for all integers $r, k \geq 1$.
- Even though we defined $\mathcal{A}(r,k)$ for all integers $r \geq 3$, in the proof of Theorem 9.2.4 we will consider only the case $r = 6$.

The following easy proposition shows that for every integer $r \geq 3$ the hypergraph $\mathcal{A}(r,k)$ is *linear*, i.e. every pair of edges has an intersection of size at most one.

Proposition 9.2.7. *Let $r \geq 3$, $k \geq 1$ be integers and $n = r^k$. Then $\mathcal{A}(r,k)$ is an $(n,r,2)$-system.*

Proof. Suppose to the contrary that there exist two distinct edges $E, E' \in \mathcal{H}$ such that $|E \cap E'| \geq 2$. Assume that $E = \{a, a+d, \ldots, a+(r-1)d\}$ for some $a, d \in \mathbb{Z}_r^k$ and d is not the zero vector. Without loss of generality we may assume that $a \in E \cap E'$ (otherwise we can choose an arbitrary element in $E \cap E'$ and rename it as a) and assume that $E' = \{a, a+id, \ldots, a+(r-1)id\}$ for some integer $i \in [r-1]$. Since $|E'| = r$, the set $\{0, id \pmod r, \ldots, (r-1)id \pmod r\}$ has size r. Therefore, sets $\{0, id \pmod r, \ldots, (r-1)id \pmod r\}$ and $\{0, 1, \ldots, r-1\}$ are identical, which implies that $E = E'$, a contradiction. Therefore, $\mathcal{A}(r,k)$ is an $(n,r,2)$-system. ∎

The next proposition shows that in order to prove Theorem 9.2.4 it suffices to find an explicit construction of an $(n, 6, 2)$-system with independence number $O(n^{1-\epsilon})$.

Proposition 9.2.8. *Suppose that there exists an (n, r, s)-system with independence number at most α. Then there exists an (n, r', s)-system with independence number at most α for every integer $r' \in [s + 1, r]$.*

Proof. Let $\mathcal{H}$ be an (n, r, s)-system with independence number at most α. Let $V = V(\mathcal{H})$. Fix an integer $r' \in [s + 1, r]$. Let the r'-graph $\mathcal{H}'$ be obtained from $\mathcal{H}$ in the following way: for every edge $E \in \mathcal{H}$ replace it by an arbitrary r'-set $E' \subset E$. It is clear that $\mathcal{H}'$ is an r'-graph on V. Now suppose that $S \subset V$ is a set of size strictly greater than α. Then, by assumption, there exists an edge $E \in \mathcal{H}$ such that $E \subset S$. It follows from the definition of $\mathcal{H}'$ that there exists $E' \in \mathcal{H}$ such that $E' \subset E \subset S$. So, S is not an independent set in $\mathcal{H}'$, which implies that $\alpha(\mathcal{H}') \leq \alpha$. ∎

Another ingredient we need for the proof of Theorem 9.2.4 is the following result due to Pach and Palincza [207].

Theorem 9.2.9 (Pach–Palincza [207])**.** *Suppose that k is a sufficiently large integer. Then every set of $\mathbb{Z}_6^k$ of size greater than $(5.709)^k$ contains a 6-AP.*

Now we are ready to prove Theorem 9.2.4.

Proof of Theorem 9.2.4. By Proposition 9.2.8, it suffices to prove that there exists an $(n, 6, 2)$-system $\mathcal{H}$ with $\alpha(\mathcal{H}) = O(n^{0.973})$.

First, for all integers n of the form 6^k we let the construction be $\mathcal{H} = \mathcal{A}(6, k)$. It follows from Proposition 9.2.7 that $\mathcal{H}$ is an $(n, 6, 2)$-system. On the other hand, it follows from the definition of $\mathcal{A}(6, k)$ that a set $S \subset V$ is independent in $\mathcal{A}(6, k)$ iff it does not contain a 6-AP. So, by Theorem 9.2.9, $|S| \leq (5.709)^k$. Therefore, $\alpha(\mathcal{H}) \leq (5.709)^k \leq n^{0.973}$.

Now suppose that n is not of the form 6^k. Then let k be the smallest integer such that $n \leq 6^k$. Let $\mathcal{H}$ be any n-vertex induced subgraph of $\mathcal{A}(6, k)$. Then $\alpha(\mathcal{H}) \leq \alpha(\mathcal{A}(6, k)) \leq (5.709)^k \leq 6n^{0.973}$. ∎

9.2.2.2 Proof of Theorem 9.2.5

Given two hypergraphs $\mathcal{H}_1$ and $\mathcal{H}_2$, the direct product of $\mathcal{H}_1$ and $\mathcal{H}_2$, denoted by $\mathcal{H}_1 \Box \mathcal{H}_2$, is the hypergraph on $V(\mathcal{H}_1) \times V(\mathcal{H}_2)$ with edge set

$$\{E_1 \times E_2 \colon E_1 \in \mathcal{H}_1 \text{ and } E_2 \in \mathcal{H}_2\},$$

where $\times$ denotes the usual cartesian product of sets.

Remark. It is clear that there exists an algorithm $\mathcal{A}'$ such that for every input $(\mathcal{H}_1, \mathcal{H}_2)$, $\mathcal{A}'$ runs in time *poly* $(|\mathcal{H}_1| \cdot |\mathcal{H}_2|)$ and outputs $\mathcal{H}_1 \Box \mathcal{H}_2$.

One nice property of the operation defined above is that the direct product of two designs is still a design.

Lemma 9.2.10. *Suppose that $\mathcal{H}_1$ is an (n_1, r_1, s_1)-system and $\mathcal{H}_2$ is an (n_2, r_2, s_2)-system. Then $\mathcal{H}_1 \Box \mathcal{H}_2$ is an $(n_1 n_2, r_1 r_2, \max\{r_1(s_2 - 1) + 1, r_2(s_1 - 1) + 1\})$-system.*

Proof. Let $n = n_1 n_2$, $r = r_1 r_2$, and $s = \max\{r_1(s_2 - 1) + 1, r_2(s_1 - 1) + 1\}$. It is clear that $\mathcal{H}_1 \square \mathcal{H}_2$ is an r-graph on n vertices. So it suffices to show that every s-set of $V(\mathcal{H}_1) \times V(\mathcal{H}_2)$ is contained in at most one edge in $\mathcal{H}_1 \square \mathcal{H}_2$.

Fix an s-set $S \subset V(\mathcal{H}_1) \times V(\mathcal{H}_2)$. Suppose to the contrary that there exist two distinct edges $E, E' \in \mathcal{H}_1 \square \mathcal{H}_2$ such that $S \subset E \cap E'$. Assume that $E = E_1 \times E_2$ and $E' = E_1' \times E_2'$, where $E_1, E_1' \in \mathcal{H}_1$, $E_2, E_2' \in \mathcal{H}_2$, and $(E_1, E_2) \neq (E_1', E_2')$. Since $E \cap E' = (E_1 \cap E_1') \times (E_2 \cap E_2')$, we have $|E \cap E'| = |E_1 \cap E_1'| \times |E_2 \cap E_2'|$. On the other hand, since $(E_1, E_2) \neq (E_1', E_2')$, we have either $E_1 \neq E_1'$ or $E_2 \neq E_2'$. In the former case we have $|E \cap E'| = |E_1 \cap E_1'| \times |E_2 \cap E_2'| \leq r_2(s_1 - 1) < s$, and in the latter case we have $|E \cap E'| = |E_1 \cap E_1'| \times |E_2 \cap E_2'| \leq r_1(s_2 - 1) < s$, both contradict the assumption that $S \subset E \cap E'$ and $|S| = s$. ∎

Using Lemma 9.2.10 we obtain the following corollary.

Corollary 9.2.11. *Let* $(\ell_1, \ell_2, \ell_3, \ell_4) \in \mathbb{N}^4$, $s = 3^{\ell_1} 4^{\ell_2} 5^{\ell_3} 6^{\ell_4} + 1$, $m \in \mathbb{N}$, $m_{i,j} \in \mathbb{N}$ *for* $i \in [\ell_j]$ *and* $j \in [4]$, *and* $M = m \prod_{j=1}^{4} \prod_{i=1}^{\ell_j} m_{i,j}$. *Suppose that* $\mathcal{H}_{i,j}$ *is an* $(m_{i,j}, 3, 2)$*-system for* $i \in [\ell_j]$ *and* $j \in [4]$, *and* $\mathcal{G} = \square_{j=1}^{4} \square_{i=1}^{\ell_j} \mathcal{H}_{i,j}$. *Then the following hold.*

(1) Suppose that $\ell_1 \neq 0$ *and* $\mathcal{H}$ *is an* $(m, 3, 2)$*-system, then* $\mathcal{H} \square \mathcal{G}$ *is an* $(M, 3(s-1), s)$*-system.*

(2) Suppose that $\ell_1 = 0$, $\ell_2 \neq 0$, *and* $\mathcal{H}$ *is an* $(m, 4, 2)$*-system, then* $\mathcal{H} \square \mathcal{G}$ *is an* $(M, 4(s-1), s)$*-system.*

(3) Suppose that $\ell_1 = \ell_2 = 0$, $\ell_3 \neq 0$, *and* $\mathcal{H}$ *is an* $(m, 5, 2)$*-system, then* $\mathcal{H} \square \mathcal{G}$ *is an* $(M, 4(s-1), s)$*-system.*

(4) Suppose that $\ell_1 = \ell_2 = \ell_3 = 0$, $\ell_4 \neq 0$, and $\mathcal{H}$ is an $(m, 6, 2)$-system, then $\mathcal{H} \Box \mathcal{G}$ is an $(M, 6(s-1), s)$-system.

The proof of Corollary 9.2.11 is just some simple but tedious calculations and we omit it here. Corollary 9.2.11 explains the reason we define $R(s)$ in the first section.

Next, we will show that the independence number of the direct product of two hypergraphs with small independence number is still relatively small. To prove this we will use the following bipartite version of the Dependent random choice lemma. Its proof is basically the same as proofs in [87; 152; 8; 236], and for the sake of completeness we include it here.

For a graph G and a set $T \subset V(G)$ we use $N(T)$ to denote the common neighbors of T in G.

Lemma 9.2.12 (Dependent random choice, see [87; 152; 8; 236])**.** *Let a, m, n_1, n_2, r be positive integers and $d_1 \geq 0$ be a real number. Let $G = G[V_1, V_2]$ be a bipartite graph with $|V_1| = n_1$, $|V_2| = n_2$, and $|G| \geq d_1 n_1$. If there exists a positive integer t such that*

$$\frac{n_1 d_1^t}{n_2^t} - \binom{n_1}{r} \left(\frac{m}{n_2}\right)^t \geq a.$$

Then there exists a subset $U \subset V(G)$ of size at least a such that every set of r vertices in U has at least m common neighbors.

Proof. Pick a set T of t vertices from V_2 uniformly at random with repetition. Set $A = N(T) \subset V_1$ and let X denote the cardinality of A. By the linearity of expectation,

$$\mathbb{E}[X] = \sum_{v \in V_1} \left(\frac{|N(v)|}{n_2} \right)^t = n_2^{-t} \sum_{v \in V_1} |N(v)|^t \geq n_2^{-t} n_1 \left(\frac{\sum_{v \in V_1} |N(v)|}{n_1} \right)^t \geq \frac{n_1 d_1^t}{n_2^t}.$$

Let Y be the random variable counting the number of subsets $S \subset A$ of size r with fewer than m common neighbors. For a given such subset S the probability that it is a subset of A equals $\left(\frac{|N(S)|}{n_2} \right)^t$. Since there are at most $\binom{n_1}{r}$ subsets $S \subset V_1$ of size r for which $|N(S)| < m$, it follows that

$$\mathbb{E}[Y] \leq \binom{n_1}{r} \left(\frac{m}{n_2} \right)^t.$$

By the linearity of expectation,

$$\mathbb{E}[X - Y] \geq \frac{n_1 d_1^t}{n_2^t} - \binom{n_1}{r} \left(\frac{m}{n_2} \right)^t \geq a.$$

Hence there exists a choice of T for which the corresponding set $A = N(T)$ satisfies $X - Y \geq a$. Deleting one vertex from each subset S of A of size r with fewer than m common neighbors. We let U be the remaining subset of A. The set U has at least $X - Y \geq a$ vertices and all subsets of size r have at least m common neighbors. ∎

The following lemma gives an upper bound for the independence number of the direct product of two hypergraphs.

Lemma 9.2.13. *Suppose that $\mathcal{H}_1$ is an r_1-graph on n_1 vertices with $\alpha(\mathcal{H}_1) < n_1/f(n_1)$ and $\mathcal{H}_2$ is an r_2-graph on n_2 vertices with $\alpha(\mathcal{H}_2) < n_2/g(n_2)$ for some real numbers $f(n_1), g(n_2) \geq 1$. Then $\mathcal{H}_1 \Box \mathcal{H}_2$ is an $r_1 r_2$-graph on $n_1 n_2$ vertices with $\alpha(\mathcal{H}_1 \Box \mathcal{H}_2) < n_1 n_2 / h(n_1, n_2)$, where $h(n_1, n_2) = (f(n_1)/2)^{1/t}$ and $t = \left\lceil \frac{\log\left(n_1^{r_1-1} f(n_1)/r_1!\right)}{\log g(n_2)} \right\rceil$.*

Proof. Let $f = f(n_1)$, $g = g(n_2)$, $t = \lceil \frac{\log\left(n_1^{r_1-1} f/r_1!\right)}{\log g} \rceil$, $h = h(n_1, n_2) = (f/2)^{1/t}$, $d_1 = n_2/h$, $m = n_2/g$, and $a = n_1/f$. Let $S \subset V(\mathcal{H}_1) \times V(\mathcal{H}_2)$ be a set of size $d_1 n_1 = n_1 n_2/h$. Define an auxiliary bipartite graph $G = G[V_1, V_2]$ with $V_1 = V(\mathcal{H}_1)$ and $V_2 = V(\mathcal{H}_2)$, and $u \in V_1$, $v \in V_2$ are adjacent iff $(u, v) \in S$. Since

$$\begin{aligned} \frac{n_1 d_1^t}{n_2^t} - \binom{n_1}{r_1}\left(\frac{m}{n_2}\right)^t - a &\geq \frac{n_1}{h^t} - \frac{n_1^{r_1}}{r_1!}\frac{1}{g^t} - \frac{n_1}{f} \\ &= n_1\left(\frac{2}{f} - \frac{n_1^{r_1-1}}{r_1!}\frac{1}{g^t} - \frac{1}{f}\right) \geq n_1\left(\frac{2}{f} - \frac{1}{f} - \frac{1}{f}\right) = 0, \end{aligned}$$

it follows from Lemma 9.2.12 that there exists a set $U \subset V_1$ of size n_1/f such that every r_1-subset of U has at least n_2/g common neighbors. Since $\alpha(\mathcal{H}_1) < n_1/f$, there exists an r_1-subset $E_1 \subset U$ such that $E_1 \in \mathcal{H}_1$. Let $W = N(E_1)$. Since $|W| \geq n_2/g > \alpha(\mathcal{H}_2)$, there exists an r_2-subset $E_2 \subset W$ such that $E_2 \in \mathcal{H}_2$. Since every pair $\{u, v\}$ with $u \in E_1$ and $v \in E_2$ is an edge in G, the set $E_1 \times E_2$ is contained in S. This implies that S is not an independent set in $\mathcal{H}_1 \Box \mathcal{H}_2$ as it contains the edge $E_1 \times E_2 \in \mathcal{H}_1 \Box \mathcal{H}_2$. Therefore, $\alpha(\mathcal{H}_1 \Box \mathcal{H}_2) < n_1 n_2/h$. ∎

Now we are ready to prove Theorem 9.2.5. As indicated by Corollary 9.2.11 our construction will be the direct product of some $(m_i, 3, 2)$-systems, $(m_j, 4, 2)$-systems, $(m_k, 5, 2)$-systems, and $(m_\ell, 6, 2)$-systems depending on the value of $\ell_1, \ell_2, \ell_3, \ell_4$, where the choice of integers m_i, m_j, m_k, m_ℓ can be optimized so that the independence number of the resulting construction is as small as possible. In the following proof we will use an inductive argument to show that such construction has a small independence number. In order to keep the argument simple, we will not try to optimize the choice of integers m_i, m_j, m_k, m_ℓ.

Proof of Theorem 9.2.5. We prove this theorem by induction on $\sum_{i\in[4]} \ell_i$. Theorem 9.2.4 shows that the base case $\sum_{i\in[4]} \ell_i = 0$ holds, so we may assume that $\sum_{i\in[4]} \ell_i \geq 1$. Let $s = 3^{\ell_1}4^{\ell_2}5^{\ell_3}6^{\ell_4} + 1$, and let us assume, for the sake of simplicity, that $\ell_1 \geq 1$ (the other cases can be proved using a similar argument). By Proposition 9.2.8 it suffices to show there is an explicit construction of an $(n, R(s), s)$-system with independence number $O(n^{1-\epsilon})$.

Fix n and let $m = \lceil \sqrt{n} \rceil$, $s_1 = 3^{\ell_1 - 1}4^{\ell_2}5^{\ell_3}6^{\ell_4} + 1$, $r_1 = 3(s_1 - 1)$. By the induction hypothesis, there exists an explicit construction $\mathcal{H}_1$ of an (m, r_1, s_1)-system with $\alpha(\mathcal{H}_1) \leq C_1 m^{1-\epsilon_1}$, where $C_1 > 0$ and $\epsilon_1 > 0$ are constants only related to r_1 and s_1. On the other hand, by Theorem 9.2.4, there exists an explicit construction $\mathcal{H}_2$ of an $(m, 3, 2)$-system with $\alpha(\mathcal{H}_2) \leq C_2 m^{1-\epsilon_2}$, where $C_2 > 0$ and $\epsilon_2 > 0$ are absolute constants. Let $C = C(C_1, C_2, \epsilon_1, \epsilon_2) > 0$ be a sufficiently large constant, $\epsilon = \epsilon(C_1, C_2, \epsilon_1, \epsilon_2) > 0$ be a sufficiently small constant (C and ϵ can be determined from the proof below), and let $\mathcal{H}_3 = \mathcal{H}_1 \Box \mathcal{H}_2$. Then by Lemma 9.2.10, $\mathcal{H}_3$ is an $(m^2, 3(s-1), s)$-system. Applying Lemma 9.2.13 to $\mathcal{H}_3$ with $f(m) = m^{\epsilon_1}/C_1$, $g(m) = m^{\epsilon_2}/C_2$ we obtain $t = \lceil \frac{\log(r_1! C_1)}{\log C_2} \frac{r_1 - 1 + \epsilon_1}{\epsilon_2} \rceil$, $h(m, m) = (m^{\epsilon_1}/2C_1)^{1/t}$, and $\alpha(\mathcal{H}_3) \leq m^2/h(m, m) \leq Cn^{1-\epsilon}$ (we can choose $C > 0$ to be sufficiently large and $\epsilon > 0$ to be sufficiently small such that the last inequality holds for all integers n). Finally, to obtain an explicit construction of an

$(n, 3(s-1), s)$-system with independent number at most $Cn^{1-\epsilon}$ one just needs to take any n-vertex induced subgraph of $\mathcal{H}_3$. ∎

Remark. As we mentioned before, one could change the number of vertices in each design in the proof above to get a better bound. For example, for $(\ell_1, \ell_2, \ell_3, \ell_4) = (2, 0, 0, 0)$, Theorem 9.2.2 with our proof above gives an $(n, 27, 10)$-design with independence number $O(n^{1-\epsilon})$, where $\epsilon \approx 6.8732 \times 10^{-6}$. On the other hand, if we take the direct production of three copies of $(\lceil n^{1/3} \rceil, 3, 2)$-systems with independence number $O(n^{1/3*0.9228})$, we obtain an $(n, 27, 10)$-design with independence number $O(n^{1-\epsilon'})$, where $\epsilon' \approx 3.5396 \times 10^{-5}$.

9.2.3 $(n, 5, 4)$-systems

We prove Theorem 9.2.6 in this section. We will show how to construct an $(n, 5, 4)$-system with small independence number inductively. More specifically, assuming that we have an $(m, 5, 4)$-system $\mathcal{H}$ with small independence number. Then we will construct a $(3m, 5, 4)$-system $\mathcal{H}'$ with small independence number by first taking three disjoint copies of $\mathcal{H}$, then embedding the vertex set of each copy of $\mathcal{H}$ into some finite field, and finally adding some crossing edges that satisfy certain equation. The set of crossing edges we add will be sparse enough to make sure the resulting construction is a $(3m, 5, 4)$-system, and it will also be dense enough to make sure the resulting construction has small independence number.

Proof of Theorem 9.2.6. We will show that it suffices to choose $C = 21$. Similar to the proof of Theorem 9.2.4 it suffices to show an explicit construction of an $(n, 5, 4)$-system with indepen-

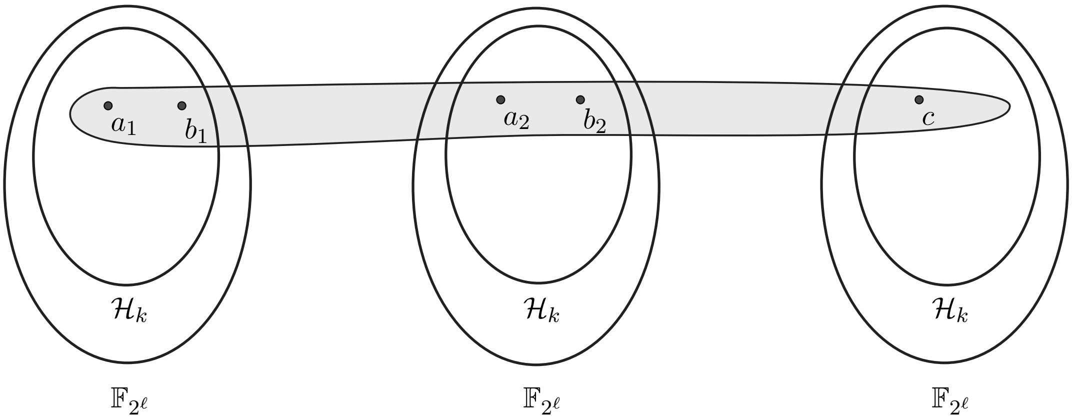


Figure 30. The induction step for constructing $\mathcal{H}_{k+1}$ using $\mathcal{H}_k$.

dence number at most $7n^{\log_3 2} - \frac{\sqrt{2}}{2-\sqrt{3}}n^{1/2}$ (this is slightly stronger that what we need) for all integers n of the form 3^k, and we will prove it by induction on k.

For $k \leq 3$ we have $7\left(3^k\right)^{\log_3 2} - \frac{\sqrt{2}}{2-\sqrt{3}}3^{k/2} \geq 3^k$, so we may assume that $k \geq 4$ and focus on the induction step. Fix an integer k and let $\mathcal{H}_k$ be a $(3^k, 5, 4)$-system with $\alpha(\mathcal{H}_k) \leq 7\left(3^k\right)^{\log_3 2} - \frac{\sqrt{2}}{2-\sqrt{3}}3^{k/2} = 7 \cdot 2^k - \frac{\sqrt{2}}{2-\sqrt{3}}3^{k/2}$. Let $\ell \in \mathbb{N}$ such that $2^\ell \geq 3^k > 2^{\ell-1}$. Let U_1, U_2, U_3 be three pairwise disjoint copies of $\mathbb{F}_{2^\ell} \setminus \{0\}$, where $\mathbb{F}_{2^\ell}$[1] is the finite field of order 2^ℓ with characteristic

[1] It is clear that $\mathbb{F}_{2^\ell}$ can be constructed in time $poly(2^\ell)$ for every integer $\ell \geq 1$.

2. For $i \in [3]$ let $\psi_i \colon V(\mathcal{H}_k) \to U_i$ be an injection and let $V_i = \psi_i(V(\mathcal{H}_k))$. Let $\mathcal{H}_{k+1}$ be the 5-graph on $V = V_1 \cup V_2 \cup V_3$ whose edge set is (see Figure 30)

$$\mathcal{H}_{k+1} = \left\{ \{a_1, b_1, a_2, b_2, c\} \in \binom{V}{5} \colon a_1, b_1 \in V_1, a_2, b_2 \in V_2, c \in V_3, a_1 + b_1 \cdot c = a_2 + b_2 \cdot c \right\}$$
$$\cup \left(\bigcup_{i \in [3]} \psi_i\left(\mathcal{H}_k\right) \right).$$

Claim 9.2.14. *$\mathcal{H}_{k+1}$ is a $(3^{k+1}, 5, 4)$-system.*

Proof. Let $S = \{a, b, c, d\} \subset V_1 \cup V_2 \cup V_3$ be a set of size 4. It is clear that if $|S \cap V_i| \geq 3$ for some $i \in [3]$ or $|S \cap V_3| \geq 2$, then S can be contained in at most one edge of $\mathcal{H}_{k+1}$. So we may assume that $|S \cap V_1|, |S \cap V_2| \leq 2$ and $|S \cap V_3| \leq 1$.

Suppose that $|S \cap V_1| = |S \cap V_2| = 2$, and without loss of generality we may assume that $S \cap V_1 = \{a, b\}$ and $S \cap V_2 = \{c, d\}$. By the definition of $\mathcal{H}_{k+1}$, every vertex $e \in V_3$ that satisfies $\{a, b, c, d, e\} \in \mathcal{H}_{k+1}$ must satisfy $a + c \cdot e = b + d \cdot e$ or $a + d \cdot e = d + c \cdot e$. Since both equations yield $e = \frac{a+b}{c+d}$ (here we used the fact that $x - y = x + y$ holds for all $x, y \in \mathbb{F}_{2^\ell}$), such vertex e is unique. Therefore, S is contained in at most one edge in $\mathcal{H}_{k+1}$.

Suppose that $|S \cap V_1| = 2$ and $|S \cap V_2| = |S \cap V_3| = 1$. Without loss of generality we may assume that $S \cap V_1 = \{a, b\}$, $S \cap V_2 = \{c\}$, and $S \cap V_3 = \{d\}$. It is easy to see that every vertex $e \in V$ that satisfies $\{a, b, c, d, e\} \in \mathcal{H}_{k+1}$ must satisfy

- $e \in V_2$, and
- $a + c \cdot d = b + e \cdot d$ or $a + e \cdot d = b + c \cdot d$.

Since both $a + c \cdot d = b + e \cdot d$ and $a + e \cdot d = b + c \cdot d$ imply $e = \frac{a+b}{d} + c$ (here we used the fact that $x - y = x + y$ holds for all $x, y \in \mathbb{F}_{2^\ell}$ again), such vertex e is unique. Therefore, S is contained in at most one edge in $\mathcal{H}_{k+1}$.

By symmetry, for the other cases one can show that S is contained in at most one edge in $\mathcal{H}_{k+1}$. Therefore, $\mathcal{H}_{k+1}$ is a $(3^{k+1}, 5, 4)$-system. ∎

Claim 9.2.15. $\alpha(\mathcal{H}_{k+1}) \leq 2\left(7 \cdot 2^k - \frac{\sqrt{2}}{2-\sqrt{3}} 3^{k/2}\right) + \sqrt{2} \cdot 3^{k/2}$.

Proof. Suppose to the contrary that there exists an independent set $S \subset V$ of size greater than $2\left(7 \cdot 2^k - \frac{\sqrt{2}}{2-\sqrt{3}} 3^{k/2}\right) + \sqrt{2} \cdot 3^{k/2}$. Let $S_i = S \cap V_i$ and $s_i = |S_i|$ for $i \in [3]$. Since S is independent in $\mathcal{H}_{k+1}$, S_i must be independent in $\psi_i(\mathcal{H}_k)$. Therefore, $s_i \leq \alpha(\mathcal{H}_k) \leq 7 \cdot 2^k - \frac{\sqrt{2}}{2-\sqrt{3}} 3^{k/2}$ for $i \in [3]$ and consequently, $s_i > \sqrt{2} \cdot 3^{k/2}$ for $i \in [3]$. Moreover, we have $s_1 + s_2 > 7 \cdot 2^k - \frac{\sqrt{2}}{2-\sqrt{3}} 3^{k/2} + \sqrt{2} \cdot 3^{k/2}$ and hence,

$$s_1 \cdot s_2 > \left(7 \cdot 2^k - \frac{\sqrt{2}}{2-\sqrt{3}} 3^{k/2}\right) \cdot \sqrt{2} \cdot 3^{k/2} \geq \sqrt{2}\left(7 - \frac{\sqrt{2}}{2-\sqrt{3}}\right) \cdot 2^k \cdot 3^{k/2} \geq 2 \cdot 3^k \geq 2^\ell.$$

Fix $z \in S_3$. Since $s_1 s_2 > 2^\ell$, by the Pigeonhole principle, there exists distinct elements $(a_1, b_1), (a_2, b_2) \in S_1 \times S_2$ such that $a_1 + b_1 \cdot z = a_2 + b_2 \cdot z$. It is easy to see that $a_1 \neq a_2$ and $b_1 \neq b_2$ since otherwise the equation $a_1 + b_1 \cdot z = a_2 + b_2 \cdot z$ would imply $(a_1, b_1) = (a_2, b_2)$, a contradiction. Therefore, $|\{a_1, a_2, b_1, b_2, z\}| = 5$ and hence, $\{a_1, a_2, b_1, b_2, z\} \in \mathcal{H}_{k+1}$. However, this implies that S contains an edge in $\mathcal{H}_{k+1}$, a contradiction. ∎

Remark. We may assume that $\alpha(\mathcal{H}_k) = \left\lceil 7 \cdot 2^k - \frac{\sqrt{2}}{2-\sqrt{3}} 3^{k/2} \right\rceil$ by removing some edges from $\mathcal{H}_k$ if necessary. If we let $n = 3^{k+1}$ and use $f(3^k)$ to denote the independence number of $\mathcal{H}_k$ for $k \in \mathbb{N}$. Then Claim 9.2.15 can be written as

$$f(n) \leq 2f(n/3) + \sqrt{2/3}\sqrt{n}.$$

By the master theorem, we have $f(n) = O(n^{\log_3 2})$. This explains the $\log_3 2$ in the exponent.

Claim 9.2.15 shows that

$$\alpha(\mathcal{H}_{k+1}) \leq 2\left(7 \cdot 2^k - \frac{\sqrt{2}}{2-\sqrt{3}} 3^{k/2}\right) + \sqrt{2} \cdot 3^{k/2} = 7 \cdot 2^{k+1} - \frac{\sqrt{2}}{2-\sqrt{3}} 3^{(k+1)/2}.$$

This completes the proof of the induction step.

Notice that given $\mathcal{H}_k$ the r-graph $\mathcal{H}_{k+1}$ can be constructed in time $poly(|\mathcal{H}_k|) + poly(2^\ell) = poly(3^k)$. So for every integer $k \geq 1$ the r-graph $\mathcal{H}_k$ can be constructed in time $poly(3^k)$.

∎

APPENDICES

.1 Proofs of lemmas in Section 6.1.2

Proof of Lemma 6.1.12. Fix $M \geq \max\{r, v\}$ such that $\mathrm{EX}(M, \mathcal{H}_v^e) \leq (\Pi(\mathcal{H}_v^e) + a/2)\binom{M}{r}$. Then there must be at least $(a/2)\binom{n}{M}$ M-sets $S \subset V(\mathcal{G})$ inducing an r-graph $\mathcal{G}[S]$ with $e(\mathcal{G}[S]) > (\Pi(\mathcal{H}_v^e) + a/2)\binom{M}{r}$. Otherwise, we would have

$$\sum_{S \in \binom{V(\mathcal{G})}{M}} e(\mathcal{G}[S]) \leq \binom{n}{M}\left(\Pi(\mathcal{H}_v^e) + \frac{a}{2}\right)\binom{M}{r} + \frac{a}{2}\binom{n}{M}\binom{M}{r} = (\Pi(\mathcal{H}_v^e) + a)\binom{n}{M}\binom{M}{r}.$$

However, we also have

$$\sum_{S \in \binom{V(\mathcal{G})}{M}} e(\mathcal{G}[S]) = \binom{n-r}{M-r} e(\mathcal{G}) > \binom{n-r}{M-r}(\Pi(\mathcal{H}_v^e) + a)\binom{n}{r} = (\Pi(\mathcal{H}_v^e) + a)\binom{n}{M}\binom{M}{r},$$

a contradiction. By the choice of M, every M-set S of $V(\mathcal{G})$ contains a copy of an element in $\mathcal{H}_v^e$. So the number of copies of elements in $\mathcal{H}_v^e$ is at least $\frac{a/2\binom{n}{M}}{\binom{n-v}{M-v}} = \frac{a/2}{\binom{M}{v}}\binom{n}{v}$. So b is at least $(a/2)/\binom{M}{v}$. ∎

Proof of Lemma 6.1.15. Let t be the number of $(k-2)$-sets $T \subset [n] - x$ satisfying

$$d_{\mathcal{F}(x)}(T) \geq n - k + 1 - \frac{\left(k^2/c + 2k\right) m}{\binom{n-1}{k-2}}.$$

Then

$$\begin{aligned}(k-1)|\mathcal{F}(x)| &= \sum_{T'\in\binom{[n]-x}{k-2}} d_{\mathcal{F}(x)}(T') \\ &\le t(n-k+1) + \left(\binom{n-1}{k-2} - t\right)\left(n-k+1-\frac{\left(k^2/c+2k\right)m}{\binom{n-1}{k-2}}\right),\end{aligned}$$

which implies that

$$\begin{aligned}\frac{\left(k^2/c+2k\right)m}{\binom{n-1}{k-2}}t &\ge (k-1)|\mathcal{F}(x)| - \binom{n-1}{k-2}\left(n-k+1-\frac{\left(k^2/c+2k\right)m}{\binom{n-1}{k-2}}\right) \\ &\ge (k-1)\left(\binom{n-k-1}{k-1} - m\right) - \binom{n-1}{k-2}\left(n-k+1-\frac{\left(k^2/c+2k\right)m}{\binom{n-1}{k-2}}\right) \\ &= (k-1)\left(\binom{n-k-1}{k-1} - \binom{n-1}{k-1}\right) + \left(k^2/c+2k-k+1\right)m \\ &\ge \left(k^2/c+k+1\right)m - (k-1)(k+1)\binom{n-k-1}{k-2} \ge km.\end{aligned}$$

Here we used the fact that $\binom{n-1}{k-1} - \binom{n-k-1}{k-1} \le (k+1)\binom{n-k-1}{k-2}$ holds for sufficiently large n, and $m \ge c\binom{n-1}{k-2}$. From the inequality above, we obtain

$$t \ge \frac{k}{k^2/c+2k}\binom{n-1}{k-2} = \frac{1}{k/c+2}\binom{n-1}{k-2}.$$

Now let us consider the family of all $(k-2)$-sets described above, and let $T_1, ..., T_\ell$ be a maximum matching in this family. Since any other set has non-empty intersection with $\bigcup_{i\in[\ell]} T_i$, we have $t \le \ell(k-2)\binom{n-1}{k-3}$. So we obtain $\ell \ge \frac{1}{(k-2)(k/c+2)}\binom{n-1}{k-2}/\binom{n-1}{k-3}$. When n is sufficiently large, we have $\ell \ge 3$, and this completes the proof. ∎

Proof of Lemma 6.1.16. By Lemma 6.1.15, there exist three disjoint $(k-2)$-sets $S_1, S_2, S_3 \subset [n] - x$ such that for each i we have

$$d_{\mathcal{F}(x)}(S_i) \geq n - k + 1 - \frac{\left(k^2/c + 2k\right) m}{\binom{n-1}{k-2}}.$$

Therefore, for each i we have

$$|\{y \in [n] \colon \{x, y\} \cup S_i \notin \mathcal{F}\}| < k + \frac{\left(k^2/c + 2k\right) m}{\binom{n-1}{k-2}}.$$

Let $B = \{y \in [n] \colon \{x, y\} \cup S_i \notin \mathcal{F}$ for some $i \in [3]\}$. Then, we have $|B| \leq 3k + \frac{\left(3k^2/c + 6k\right) m}{\binom{n-1}{k-2}}$. By adding vertices into B, we may assume that

$$|B| = 3k + \frac{\left(3k^2/c + 6k\right) m}{\binom{n-1}{k-2}}.$$

Since $m \geq c\binom{n-1}{k-2}$ holds for some constant $c > 0$, we have

$$\begin{aligned} |B| = 3k + \frac{\left(3k^2/c + 6k\right) m}{\binom{n-1}{k-2}} &\leq \frac{\left(6k^2/c + 6k\right) m}{\binom{n-1}{k-2}} \\ &\leq \frac{\left(6k^2/c + 6k\right) \delta \binom{n-1}{k-1}}{\binom{n-1}{k-2}} \leq \left(6k^2/c + 6k\right) \delta n \leq \frac{n-1}{2}. \end{aligned}$$

For each $i \in \{0, 1, \ldots, k\}$ define

$$\mathcal{T}_i = \{T \in \mathcal{F}(\bar{x}) \colon |T \cap B| = i\}.$$

Note that $\bigcup_{i=0}^{k} \mathcal{T}_i$ is a partition of $\mathcal{F}(\bar{x})$. First we show that $\mathcal{T}_0 = \mathcal{T}_1 = \mathcal{T}_2 = \emptyset$. Our first observation is that by definition $S_i \subset B$ for all $i \in [3]$. If $S \in \mathcal{T}_0 \cup \mathcal{T}_1 \cup \mathcal{T}_2$, then there is an i for which $S_i \cap S = \emptyset$. Choose $d-2 \leq k-2$ elements $y_1, \ldots, y_{d-2} \in S \backslash B$ and $y \in [n]-x-B-S$. Now the $d-2$ sets $\{x, y_j\} \cup S_i$ for $j \in [d-2]$, together with S and $\{x, y\} \cup S_i$ form a d-cluster in $\mathcal{F}$, a contradiction. Therefore, $\mathcal{T}_0 = \mathcal{T}_1 = \mathcal{T}_2 = \emptyset$. So, we have $\mathcal{F}(\bar{x}) = \bigcup_{i=3}^{k} \mathcal{T}_i$. We may assume that $|\mathcal{T}_p| \geq m/(k-2)$ holds for some $3 \leq p \leq k$. Applying Lemma 2.4 with $U_1 = B, U_2 = [n]-x-B$ and $u_1 = |U_1|, u_2 = |U_2|$, we obtain

$$\frac{m}{k-2} \leq |\mathcal{T}_p| \leq k u_1^{p-1} u_2^{k-p} \leq k \left(\frac{\left(6k^2/c + 6k\right) m}{\binom{n-1}{k-2}} \right)^{p-1} n^{k-p}.$$

Simplifying the inequality above, we obtain

$$m^{p-2} \geq \frac{\binom{n-1}{k-2}^{p-1}}{(k-2)k\left(6k^2/c + 6k\right)^{p-1} n^{k-p}}.$$

Since $m \leq \delta\binom{n-1}{k-1} \leq \delta n \binom{n-1}{k-2}$, we know that

$$\begin{aligned} \delta \geq \delta^{p-2} &\geq \frac{m^{p-2}}{n^{p-2}\binom{n-1}{k-2}^{p-2}} \geq \frac{\binom{n-1}{k-2}}{(k-2)k\left(6k^2/c + 6k\right)^{p-1} n^{k-2}} \\ &\geq \frac{\left(\frac{n-k}{n}\right)^{k-2}}{(k-2)!(k-2)k\left(6k^2/c + 6k\right)^{p-1}} \geq \frac{1}{2(k-2)!(k-2)k\left(6k^2/c + 6k\right)^{p-1}} \end{aligned}$$

holds for sufficiently large n.

Now choose $\delta > 0$ to be sufficiently small such that $\delta < \frac{1}{2(k-2)!(k-2)k(6k^2/c+6k)^{p-1}}$. Then we get a contradiction, and this completes the proof. ∎

.2 Validity of constructions in Section 6.1

It is easy to see that $\nu(\mathcal{L}_i) = \nu + 1$ holds for every $i \in [5]$. So, it suffices to show that the families $\mathcal{L}_i$ are d-cluster-free.

Claim .2.1. *The family $\mathcal{L}_1$ is 3-cluster-free and $\nu(\mathcal{L}_1) = \nu + 1$.*

Proof. Suppose there exist three sets $L_1, L_2, L_3 \in \mathcal{L}_1$ that form a 3-cluster. Since $L_1 \cap L_2 \cap L_3 = \emptyset$, one of these three sets must be C_i for some i, and we may assume that $L_1 = C_1$. On the other hand, since $|L_1 \cup L_2 \cup L_3| \leq 2k$, the two sets L_2 and L_3 must both contain y, and $L_2 \cap J, L_3 \cap J$ must be both contained in P_1. However, in this case, we would have $v_1 \in L_1 \cap L_2 \cap L_3$, a contradiction. Therefore, the family $\mathcal{L}_1$ is 3-cluster-free. ∎

Claim .2.2. *The family $\mathcal{L}_2$ is 3-cluster-free and $\nu(\mathcal{L}_2) = \nu + 1$.*

Proof. Suppose there exist three sets $L_1, L_2, L_3 \in \mathcal{L}_1$ that form a 3-cluster. Similar to the proof of Claim B.1, we may assume that $L_1 = C_1$. Since $|L_1 \cap L_2 \cap L_3| \leq 2k$, we know that $L_2 \cap J$ and $L_3 \cap J$ must be both nonempty. For every $i \in \{2, 3\}$ let $\mathcal{L}_2(i) = \{L \in \mathcal{L}_2 \colon |L \cap J| = i\}$. From the proof of Claim .2.1, we know that L_2 and L_3 cannot be both in $\mathcal{L}_2(2)$.

If $L_2 \in \mathcal{L}_2(2)$ and $L_3 \in \mathcal{L}_2(3)$, then we would have $|L_2 \cap L_3| \leq k - 3$ and this implies $|L_1 \cup L_2 \cup L_3| = 3k - (|L_1 \cap L_2| + |L_1 \cap L_3| + |L_2 \cap L_3|) \geq 2k + 1$, a contradiction.

So we may assume that L_2, L_3 are both contained in $\mathcal{L}_2(3)$. Let $I_2 = L_2 \cap J$ and $I_3 = L_3 \cap J$. By the definition of $\mathcal{L}_2$, we have $|I_2 \cap I_3| \leq 1$. Note that at least one of $L_1 \cap I_2, L_1 \cap I_3, I_2 \cap I_3$ must be the empty set, since otherwise we would have $L_1 \cap L_2 \cap L_3 \neq \emptyset$, a contradiction.

Therefore, we have $|L_1 \cap L_2| + |L_1 \cap L_3| + |L_2 \cap L_3| \leq k - 3 + 2 = k - 1$, and this implies that $|L_1 \cup L_2 \cup L_3| \geq 2k + 1$, a contradiction. Therefore, the family $\mathcal{L}_2$ is 3-cluster-free. ∎

Claim .2.3. *The family $\mathcal{L}_3$ is 4-cluster-free and $\nu(\mathcal{L}_3) = 2$.*

Proof. Suppose there exist four sets $L_1, L_2, L_3, L_4 \in \mathcal{L}$ that form a 4-cluster. Similar to the proof of Claim .2.1, we may assume that $L_1 = C_1$. Since $|L_1 \cup \cdots \cup L_4| \leq 2k$, there are at least two sets in $\{L_2, L_3, L_4\}$ containing v. We may assume that $v \in L_2$ and $v \in L_3$. Let $E_2 = L_2 \cap W$ and $E_3 = L_3 \cap W$, and note that $E_2, E_3 \in \mathcal{G}$. Since $\mathcal{G}$ is P_2^{k-2}-free, we have $|E_2 \cup E_3| \geq k$, but this contradicts our assumption that $|L_1 \cup \cdots \cup L_4| \leq 2k$. Therefore, the family $\mathcal{L}_3$ is 4-cluster-free. ∎

Claim .2.4. *The family $\mathcal{L}_4$ is 4-cluster-free and $\nu(\mathcal{L}_4) = \nu + 1$.*

Proof. Suppose there exist four sets $L_1, L_2, L_3, L_4 \in \mathcal{L}_4$ that form a 4-cluster. Similar to the proof of Claim .2.1, we may assume that $L_1 = C_1$. Since $|L_1 \cup L_2 \cup L_3 \cup L_4| \leq 2k$, the three sets L_2, L_3, L_4 must all contain y and all have nonempty intersection with J. For every $i \in \{2, 3, 4\}$ let $E_i = L_i \cap J$ and let $S_i = L_i \cap W$. The inequality $|L_1 \cup L_2 \cup L_3 \cup L_4| \leq 2k$ implies that $|S_2 \cup S_3 \cup S_4| \leq k - 2$.

Suppose that $|S_2 \cup S_3 \cup S_4| = k - 2$. Then, the inequality $|L_1 \cup L_2 \cup L_3 \cup L_4| \leq 2k$ implies that $E_2 = E_3 = E_4$, and $E_i \cap C_1 \neq \emptyset$ holds for every $i \in \{2, 3, 4\}$. However, in this case, we would have $L_1 \cap L_2 \cap L_3 \cap L_4 \neq \emptyset$, a contradiction. Therefore, we may assume that $S_2 = S_3 = S_4$. Since $|L_1 \cup L_2 \cup L_3 \cup L_4| \leq 2k$, at least two sets in $\{E_2, E_3, E_4\}$ have nonempty intersection with C_1, and we may assume that $E_2 \cap C_1 \neq \emptyset$ and $E_3 \cap C_1 \neq \emptyset$. Now we already have $|L_1 \cup L_2 \cup L_3| = 2k$, therefore, the set E_4 must be contained in $E_2 \cup E_3 \cup C_1$. However, by the definition of G, this is impossible. Therefore, the family $\mathcal{L}_4$ is 4-cluster-free. ∎

Claim .2.5. *The family $\mathcal{L}_5$ is d-cluster-free and $\nu(\mathcal{L}_5) = \nu + 1$.*

Proof. Suppose there exist d sets $L_1, ..., L_d \in \mathcal{L}_5$ that form a d-cluster. Similar to the proof of Claim .2.1, we may assume that $L_1 = C_1$. For every $i \in \{2, \ldots, d\}$, let $S_i = L_i \cap W$ and $T_i = L_i \cap J$, and note that some of the T_i's may be empty. The inequality $|L_1 \cup \cdots \cup L_d| \leq 2k$ implies that the $d-1$ sets $L_2, \ldots, L_d$ all contains y, and $|S_2 \cup \ldots \cup S_d| \leq k-1$. Let $S = S_2 \cup \cdots \cup S_d$.

If S is of size $k-1$, then $T_i \subset C_1$ holds for every $i \in \{2, \ldots, d\}$. Since at most one set in $\{T_2, \ldots, T_d\}$ is empty, the set S contains at least $d-2$ edges of $\mathcal{G}_1$ and, hence, we have $\mathcal{G}_1[S] \in \mathcal{H}_{k-1}^{d-2}$, a contradiction. Therefore, we may assume that $S_2 = \cdots = S_d$. Note that S is of size $k-2$ and every T_i is nonempty. Since $|L_1 \cup \cdots \cup L_d| \leq 2k$, at most one set in $\{T_2, \ldots, T_d\}$ is not contained in C_1. This implies that at least $d-2$ sets in $\{T_2, \ldots, T_d\}$ are contained in C_1. However, this implies that S is an edge in $\mathcal{G}_1$ with multiplicity at least $d-2$, a contradiction. Therefore, the family $\mathcal{L}_5$ is d-cluster-free. ∎

.3 Proof of Theorem 7.1.16

In this section we prove Theorem 7.1.16. We need the following lemmas.

Lemma .3.1 ([65]). *Let $0 < \alpha \leq 1$ and let G be a triangle-free graph on αn vertices with at least $(2\alpha - 1)n^2/4$ edges. Then G contains a matching with at least $(2\alpha - 1)n/2$ edges.*

Lemma .3.2. *Let $1/2 \leq \alpha \leq 1$, $n \in \mathbb{N}$ and $\alpha n \in \mathbb{N}$. Let G be bipartite graph on n vertices. If every vertex set of size αn in G spans at least $(2\alpha - 1)n^2/4$ edges, then $G \cong T_2(n)$.*

Proof. Let $V_1 \cup V_2 = V(G)$ be a partition such that G is a bipartite graph with parts V_1 and V_2. Let $x = |V_1|/n$ and we may assume that $x \geq 1/2$. By assumption, $x < \alpha$, since otherwise there would be a subset of A of size αn that spans zero edges. Now choose an arbitrary set $S \subset B$ with $|S| = (\alpha - x)n$. Then $|A \cup B| = \alpha n$ and $e(A \cup S) \leq x(\alpha - x)n^2 \leq (2\alpha - 1)n^2/4$. By assumption the inequality above must be tight, which means $x = 1/2$ and $G[A, S]$ is a complete bipartite graph. Since S was chosen randomly, G must be a complete bipartite graph with $|V_1| = n/2$. Therefore, $G \cong T_2(n)$. ∎

Now we prove Theorem 7.1.16.

Proof of Theorem 7.1.16. First one could see from Kriveleich's proof (i.e. the proof of Theorem 4) in [153] that if G does not contain an independent set of size $(1 - \alpha)n$, then there exists a vertex set of size αn in G that spans strictly less than $\frac{2\alpha-1}{4}n^2$ edges. So by assumption there exists an independent set in G whose size is $(1 - \alpha)n$. Next, we use the argument of Erdős et al. [65] to show that $G \cong T_2(n)$.

Let $A \subset V(G)$ be an independent set of size $(1-\alpha)n$. By Lemma .3.1, there exists a matching M in $G[V(G) \setminus A]$ with $(2\alpha-1)n/2$ edges. Let $C = V(M)$ and let $B = V(G) \setminus (A \cup C)$. Note that $|C| = (2\alpha - 1)n$.

Since G is triangle-free and M is a matching, every vertex in A is adjacent to at most half of the vertices in C. Therefore, $e(A, C) \leq (1-\alpha)(2\alpha-1)n^2/2$ and hence

$$e(A \cup C) = e(A, C) + e(C) \leq \frac{(1-\alpha)(2\alpha-1)n^2}{2} + \frac{(2\alpha-1)^2 n^2}{4} = \frac{2\alpha-1}{4}n^2.$$

Since $|A \cup C| = \alpha n$, by assumption, $e(A \cup C) \geq (2\alpha-1)n^2/4$. So all inequalities above must be tight, which means $G[C]$ is a balanced complete bipartite graph. and every vertex in A is adjacent to exactly half of the vertices in C.

Let $C_1 \cup C_2 = C$ be a partition such that $G[C] = G[C_1, C_2]$ and note that $|C_1| = |C_2| = |C|/2 = (2\alpha-1)n/2$. For $i \in \{1, 2\}$ let

$$A_i = \{u \in A \colon \exists v \in C_i, uv \in E(G)\} \quad \text{and} \quad B_i = \{u \in B \colon \exists v \in C_i, uv \in E(G)\},$$

and let $B_3 = B \setminus (B_1 \cup B_2)$. Since $G[C_1, C_2]$ is a complete bipartite graph and every $v \in A$ is adjacent to at least half vertices in $C_1 \cup C_2$, we have $uw \in E(G)$ for all $u \in A_i$ and $w \in C_i$ for $i \in \{1, 2\}$. Notice that $A_1 \cup A_2$ is a partition of A, and for $i \in \{1, 2\}$ we have $uv \notin E(G)$ for all $u \in B_i, v \in A_i$, since otherwise there exists $w \in C_1$ such that u, v, w induces a copy of K_3 in G, a contradiction. Therefore, if $B_3 = \emptyset$, then G is bipartite with two parts $V_1 = C_1 \cup A_2 \cup B_2$ and $V_2 = C_2 \cup A_1 \cup B_1$, and by Lemma .3.2, $G \cong T_2(n)$. So we may assume that $B_3 \neq \emptyset$.

Let $\widehat{C}_1 = C_1 \cup B_2$ and $\widehat{C}_2 = C_2 \cup B_1$. Let $x_i = |A_i|/n$, $y_i = |\widehat{C}_i|/n$ for $i \in \{1, 2\}$, and $z = |B_3|/n$. Since $|\widehat{C}_1 \cup \widehat{C}_2 \cup A_1 \cup B_3| = n - |A_2| \geq \alpha n$, there exists $U_1 \subset \widehat{C}_1$ with $|U_1| = \alpha n - |B_3 \cup A_1 \cup \widehat{C}_2| = (\alpha - z - x_1 - y_2)n$. Since $|B_3 \cup A_1 \cup \widehat{C}_2 \cup U_1| = \alpha n$, by assumption

$$\frac{2\alpha - 1}{4} n^2 \leq e(B_3 \cup A_1 \cup \widehat{C}_2 \cup U_1) \leq z x_1 n^2 + (x_1 + y_2)(\alpha - z - x_1 - y_2)n^2.$$

Similarly, there exists $U_2 \subset \widehat{C}_2$ with $|U_2| = (\alpha - z - x_2 - y_2)n$, and

$$\frac{2\alpha - 1}{4} n^2 \leq e(B_3 \cup A_2 \cup \widehat{C}_1 \cup U_2) \leq z x_2 n^2 + (x_2 + y_1)(\alpha - z - x_2 - y_1)n^2.$$

Adding up these two inequalities we obtain

$$\begin{aligned} \frac{2\alpha - 1}{2} &\leq z x_1 + (x_1 + y_2)(\alpha - z - x_1 - y_2) + z x_2 + (x_2 + y_1)(\alpha - z - x_2 - y_1) \\ &= \alpha(x_1 + x_2 + y_1 + y_2) - z(y_1 + y_2) - \left((x_1 + y_2)^2 + (x_2 + y_1)^2\right) \\ &\leq \alpha(1 - z) - z(\alpha - z) - \frac{(x_1 + x_2 + y_1 + y_2)^2}{2} \\ &= \alpha(1 - z) - z(\alpha - z) - \frac{(1 - z)^2}{2} = \frac{z^2}{2} - (2a - 1)z + \frac{2\alpha - 1}{2}, \end{aligned}$$

which implies that $z^2/2 - (2a - 1)z \geq 0$. However, since $0 < z \leq 1 - \alpha < 4\alpha - 2$ (here we used $\alpha > 3/5$ and $B_3 \neq \emptyset$),

$$\frac{z^2}{2} - (2a - 1)z = \frac{z}{2}(z - (4\alpha - 2)) < 0,$$

a contradiction. ∎

.4 Proofs of Claims 7.1.21 and 7.1.22

In this section we prove Claims 7.1.21 and 7.1.22.

Proof of Claim 7.1.21. Since $x_2^2 + x_3^2 \geq (x_2 + x_3)^2/2$, it suffices to show

$$\frac{(x_2 + x_3)^2/2 + x_4^2}{9x_1(1 - 2x_1)} - \frac{x_4^2}{6x_1} + \frac{1}{6} - c > 0.$$

Plugging $x_4 = 1 - g(c)$ and $x_2 + x_3 = g(c) - x_1$ into the inequality above, it becomes

$$\ell(c, x_1) := \frac{(g(c) - x_1)^2 + 2(1 - g(c))^2}{18x_1(1 - 2x_1)} - \frac{(1 - g(c))^2}{6x_1} + \frac{1}{6} - c > 0. \tag{.10}$$

Then with the aid of computer [183] one can see that

$$\min \{\ell(c, x) \colon x \in (0, 1/2), c \in [1/4, 1/3]\} > 0.003.$$

Therefore, Equation .10 is true. ∎

Proof of Claim 7.1.22. First since $1/(1/2 - x)$ is convex, by Jensen's inequality

$$\frac{x_4}{1/2 - x_1} + \frac{x_4}{1/2 - x_2} \geq \frac{4x_4}{1 - (x_1 + x_2)}.$$

Since $x_1 + x_2 \geq 1/2$ and $x_2 \leq x_1 < 1/2$,

$$\frac{x_2}{1/2 - x_1} + \frac{x_1}{1/2 - x_2} - \frac{2(x_1 + x_2)}{1 - (x_1 + x_2)} = \frac{2(x_1 - x_2)^2(2x_1 + 2x_2 - 1)}{(1 - 2x_1)(1 - 2x_2)(1 - x_1 - x_2)} \geq 0.$$

It suffice to show that

$$c + \frac{(1 - x_1 - x_2)x_4}{6(x_1 + x_2)} < \frac{1}{18}\left(\left(1 + \frac{1}{1 - 2x_3} - \frac{1}{x_1 + x_2}\right) \cdot \frac{x_1 + x_2}{1 - (x_1 + x_2)} \right.$$
$$\left. + \frac{1}{1 - 2x_3} + \frac{1}{2(x_1 + x_2)} \cdot \frac{4x_4}{1 - (x_1 + x_2)} + 1\right),$$

Let $x = x_1 + x_2$. Then $x_3 = g(c) - x$ and the inequality above can be simplified as

$$m(x, c) := \left(1 + \frac{1}{1 - 2(g(c) - x)} - \frac{1}{x}\right) \cdot \frac{x}{1 - x}$$
$$+ \frac{1}{1 - 2(g(c) - x)} + \frac{2(1 - g(c))}{x(1 - x)} + 1 - \frac{3(1 - x)(1 - g(c))}{x} - 18c > 0. \qquad (.11)$$

Then with the aid of computer [183] one can see that

$$\min\{m(x, c) \colon x \in [1/2, 1], c \in [1/4, 1/3]\} > 0.099.$$

Therefore, Equation .11 is true. ∎

CITED LITERATURE

1. R. Ahlswede, H. Aydinian, and L. H. Khachatrian. On shadows of intersecting families. *Combinatorica*, 24(4):555–566, 2004.

2. R. Ahlswede and L. H. Khachatrian. The complete nontrivial-intersection theorem for systems of finite sets. *J. Combin. Theory Ser. A*, 76(1):121–138, 1996.

3. M. Ajtai, P. Erdős, J. Komlós, and E. Szemerédi. On Turán's theorem for sparse graphs. *Combinatorica*, 1(4):313–317, 1981.

4. M. Ajtai, J. Komlós, J. Pintz, J. Spencer, and E. Szemerédi. Extremal uncrowded hypergraphs. *J. Combin. Theory Ser. A*, 32(3):321–335, 1982.

5. M. Ajtai, J. Komlós, and E. Szemerédi. A note on Ramsey numbers. *J. Combin. Theory Ser. A*, 29(3):354–360, 1980.

6. N. Alon and F. R. K. Chung. Explicit construction of linear sized tolerant networks. In *Proceedings of the First Japan Conference on Graph Theory and Applications (Hakone, 1986)*, volume 72, pages 15–19, 1988.

7. N. Alon, D. Hefetz, M. Krivelevich, and M. Tyomkyn. Edge-statistics on large graphs. *Combin. Probab. Comput.*, 29(2):163–189, 2020.

8. N. Alon, M. Krivelevich, and B. Sudakov. Turán numbers of bipartite graphs and related Ramsey-type questions. volume 12, pages 477–494. 2003. Special issue on Ramsey theory.

9. N. Alon, K. E. Mellinger, D. Mubayi, and J. Verstraëte. The de Bruijn-Erdős theorem for hypergraphs. *Des. Codes Cryptogr.*, 65(3):233–245, 2012.

10. N. Alon and C. Shikhelman. Many T copies in H-free graphs. *J. Combin. Theory Ser. B*, 121:146–172, 2016.

11. B. Andrásfai, P. Erdős, and V. T. Sós. On the connection between chromatic number, maximal clique and minimal degree of a graph. *Discrete Math.*, 8:205–218, 1974.

12. J. Balogh and F. C. Clemen. On stability of the Erdős-Rademacher problem. *arXiv preprint arXiv:2003.12917*, 2020.

13. J. Balogh, F. C. Clemen, M. Lavrov, B. Lidickỳ, and F. Pfender. Making K_{r+1}-free graphs r-partite. *Combin. Probab. Comput.*, pages 1–10, 2019.

14. Z. Baranyai. On the factorization of the complete uniform hypergraph. In *Infinite and finite sets (Colloq., Keszthely, 1973; dedicated to P. Erdős on his 60th birthday), Vol. I*, pages 91–108. Colloq. Math. Soc. János Bolyai, Vol. 10. 1975.

15. W. Bedenknecht, G. O. Mota, C. Reiher, and M. Schacht. On the local density problem for graphs of given odd-girth. *J. Graph Theory*, 90(2):137–149, 2019.

16. L. Bellmann and C. Reiher. Turán's theorem for the Fano plane. *Combinatorica*, 39(5):961–982, 2019.

17. A. Bene Watts, S. Norin, and L. Yepremyan. A Turán theorem for extensions via an Erdős-Ko-Rado theorem for Lagrangians. *Combinatorica*, 39(5):1149–1171, 2019.

18. P. Bennett and T. Bohman. A note on the random greedy independent set algorithm. *Random Structures Algorithms*, 49(3):479–502, 2016.

19. J.-C. Bermond and P. Frankl. On a conjecture of Chvátal on m-intersecting hypergraphs. *Bull. London Math. Soc.*, 9(3):310–312, 1977.

20. A. Bhattacharya, S. Friedland, and U. N. Peled. On the first eigenvalue of bipartite graphs. *Electron. J. Combin.*, 15(1):Research Paper 144, 23, 2008.

21. N. Biggs. *Algebraic graph theory.* Cambridge Mathematical Library. Cambridge University Press, Cambridge, second edition, 1993.

22. T. Bohman, X. Liu, and D. Mubayi. Independent sets in hypergraphs omitting an intersection. *Random Structures Algorithms*, 2021.

23. B. Bollobás. On generalized graphs. *Acta Math. Acad. Sci. Hung.*, 16:447–452, 1965.

24. B. Bollobás. Three-graphs without two triples whose symmetric difference is contained in a third. *Discrete Math*, 8(1):21–24, 1974.

25. B. Bollobás. Relations between sets of complete subgraphs. In *Proc. Fifth British Comb. Conference (Univ. Aberdeen, Aberdeen, 1975)*, pages 79–84, 1975.

26. B. Bollobás, D. Daykin, and P. Erdős. Sets of independent edges of a hypergraph. *Quart. J. of Math. Oxford Ser*, 27(1):25–32, 1976.

27. C. Bonferroni. Teoria statistica delle classi e calcolo delle probabilita. *Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commericiali di Firenze*, 8:3–62, 1936.

28. P. Borg and C. Feghali. On the hilton-spencer intersection theorems for unions of cycles. *arXiv preprint arXiv:1908.08825*, 2019.

29. R. C. Bose and D. K. Ray-Chaudhuri. On a class of error correcting binary group codes. *Information and Control*, 3:68–79, 1960.

30. A. Brandt, D. Irwin, and T. Jiang. Stability and Turán numbers of a class of hypergraphs via Lagrangians. *Combin. Probab. Comput.*, 26(3):367–405, 2017.

31. J. I. Brown and A. Sidorenko. The inducibility of complete bipartite graphs. *J. Graph Theory*, 18(6):629–645, 1994.

32. W. G. Brown. On an open problem of Paul Turán concerning 3-graphs. In *Studies in pure mathematics*, pages 91–93. Springer, 1983.

33. S. Bubeck and N. Linial. On the local profiles of trees. *J. Graph Theory*, 81(2):109–119, 2016.

34. M. Budden, J. Hiller, and A. Rapp. Generalized Ramsey theorems for r-uniform hypergraphs. *Australas. J. Combin.*, 63:142–152, 2015.

35. E. Chattopadhyay and J. Goodman. Explicit extremal designs and applications to extractors. *arXiv preprint arXiv:2007.07772*, 2020.

36. E. Chattopadhyay, J. Goodman, V. Goyal, and X. Li. Extractors for adversarial sources via extremal hypergraphs. In *Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing*, pages 1184–1197, 2020.

37. W. Chen, J. Liu, and L. Wang. Families of sets with intersecting clusters. *SIAM J. Discrete Math.*, 23(3):1249–1260, 2009.

38. F. Chung and L. Lu. An upper bound for the Turán number $t_3(n, 4)$. *J. Comb. Theory, Ser. A*, 87(2):381–389, 1999.

39. F. R. K. Chung and P. Frankl. The maximum number of edges in a 3-graph not containing a given star. *Graphs Combin.*, 3(2):111–126, 1987.

40. F. R. K. Chung and R. L. Graham. On graphs not containing prescribed induced subgraphs. In *A tribute to Paul Erdős*, pages 111–120. Cambridge Univ. Press, Cambridge, 1990.

41. V. Chvátal. An extremal set-intersection theorem. *J. London Math. Soc. (2)*, 9:355–359, 1974/75.

42. G. F. Clements and H.-D. O. F. Gronau. On maximal antichains containing no set and its complement. *Discrete Math.*, 33(3):239–247, 1981.

43. F. N. Cole. The triad systems of thirteen letters. *Trans. Amer. Math. Soc.*, 14(1):1–5, 1913.

44. J. Cooper and D. Mubayi. Sparse hypergraphs with low independence number. *Combinatorica*, 37(1):31–40, 2017.

45. E. Croot, V. F. Lev, and P. P. Pach. Progression-free sets in $\mathbb{Z}_4^n$ are exponentially small. *Ann. of Math. (2)*, 185(1):331–337, 2017.

46. R. Csákány and J. Kahn. A homological approach to two problems on finite sets. *J. Algebraic Combin.*, 9(2):141–149, 1999.

47. G. Currier. New results on simplex-clusters in set systems. *Combinatorica*, 41(4):495–506, 2021.

48. G. Currier. On the d-cluster generalization of Erdős-Ko-Rado. *J. Combin. Theory Ser. A*, 182:Paper No. 105464, 7, 2021.

49. D. de Caen. On upper bounds for 3-graphs without tetrahedra. volume 62, pages 193–202. 1988. Seventeenth Manitoba Conference on Numerical Mathematics and Computing (Winnipeg, MB, 1987).

50. D. De Caen and Z. Füredi. The maximum size of 3-uniform hypergraphs not containing a Fano plane. *J. Combin. Theory Ser. B*, 78(2):274–276, 2000.

51. L. de Oliveira Contiero, C. Hoppen, H. Lefmann, and K. Odermann. Stability results for two classes of hypergraphs. *SIAM J. Discrete Math.*, 33(4):2023–2040, 2019.

52. M. Deza, P. Erdős, and P. Frankl. Intersection properties of systems of finite sets. *Proc. London Math. Soc. (3)*, 36(2):369–384, 1978.

53. R. A. Duke, H. Lefmann, and V. Rödl. On uncrowded hypergraphs. In *Proceedings of the Sixth International Seminar on Random Graphs and Probabilistic Methods in Combinatorics and Computer Science, "Random Graphs '93" (Poznań, 1993)*, volume 6, pages 209–212, 1995.

54. G. P. Egorychev. The solution of van der Waerden's problem for permanents. *Adv. in Math.*, 42(3):299–305, 1981.

55. J. S. Ellenberg and D. Gijswijt. On large subsets of $\mathbb{F}_q^n$ with no three-term arithmetic progression. *Ann. of Math. (2)*, 185(1):339–343, 2017.

56. P. Erdős. On a theorem of Rademacher-Turán. *Illinois J. Math.*, 6:122–127, 1962.

57. P. Erdős. On the number of complete subgraphs contained in certain graphs. *Magyar Tud. Akad. Mat. Kutató Int. Közl.*, 7:459–464, 1962.

58. P. Erdős. On extremal problems of graphs and generalized graphs. *Israel J. Math.*, 2:183–190, 1964.

59. P. Erdős. A problem on independent r-tuples. *Ann. Univ. Sci. Budapest. Eötvös Sect. Math.*, 8:93–95, 1965.

60. P. Erdős. Some recent results on extremal problems in graph theory. Results. In *Theory of Graphs (Internat. Sympos., Rome, 1966)*, pages 117–123 (English); pp. 124–130 (French). Gordon and Breach, New York; Dunod, Paris, 1967.

61. P. Erdős. Problems and results in graph theory and combinatorial analysis. In *Proceedings of the Fifth British Combinatorial Conference (Univ. Aberdeen, Aberdeen, 1975)*, pages 169–192. Congressus Numerantium, No. XV, 1976.

62. P. Erdős. On the combinatorial problems which I would most like to see solved. *Combinatorica*, 1(1):25–42, 1981.

63. P. Erdős. Some old and new problems in various branches of combinatorics. volume 165/166, pages 227–231. 1997. Graphs and combinatorics (Marseille, 1995).

64. P. Erdős, R. Faudree, J. Pach, and J. Spencer. How to make a graph bipartite. *J. Combin. Theory Ser. B*, 45(1):86–98, 1988.

65. P. Erdős, R. J. Faudree, C. C. Rousseau, and R. H. Schelp. A local density condition for triangles. volume 127, pages 153–161. 1994. Graph theory and applications (Hakone, 1990).

66. P. Erdős, P. Frankl, and V. Rödl. The asymptotic number of graphs not containing a fixed subgraph and a problem for hypergraphs having no exponent. *Graphs Combin.*, 2(2):113–121, 1986.

67. P. Erdős, E. Győri, and M. Simonovits. How many edges should be deleted to make a triangle-free graph bipartite? In *Sets, graphs and numbers (Budapest, 1991)*, volume 60 of *Colloq. Math. Soc. János Bolyai*, pages 239–263. North-Holland, Amsterdam, 1992.

68. P. Erdős, A. Hajnal, V. T. Sós, and E. Szemerédi. More results on Ramsey-Turán type problems. *Combinatorica*, 3(1):69–81, 1983.

69. P. Erdős, C. Ko, and R. Rado. Intersection theorems for systems of finite sets. *Quart. J. Math. Oxford Ser. (2)*, 12:313–320, 1961.

70. P. Erdős, L. Lovász, and J. Spencer. Strong independence of graphcopy functions. Graph theory and related topics, Proc. Conf. Honour W. T. Tutte, Waterloo/Ont. 1977, 165-172 (1979)., 1979.

71. P. Erdős and M. Simonovits. A limit theorem in graph theory. *Studia Sci. Math. Hungar.*, 1:51–57, 1966.

72. P. Erdős and M. Simonovits. On a valence problem in extremal graph theory. *Discrete Math.*, 5:323–334, 1973.

73. P. Erdős and V. T. Sós. Some remarks on Ramsey's and Turán's theorem. In *Combinatorial theory and its applications, II (Proc. Colloq., Balatonfüred, 1969)*, pages 395–404, 1970.

74. P. Erdős and A. H. Stone. On the structure of linear graphs. *Bull. Amer. Math. Soc.*, 52:1087–1091, 1946.

75. P. Erdös. Some theorems on graphs. *Riveon Lematematika*, 9:13–17, 1955.

76. P. Erdős. A problem on independent r-tuples. *Ann. Univ. Sci. Budapest. Eötvös Sect. Math.*, 8:93–95, 1965.

77. P. Erdős. Topics in combinatorial analysis. In *Proceedings of the Second Louisiana Conference on Combinatorics, Graph Theory and Computing*, pages 2–20, 1971.

78. A. Eustis and J. Verstraëte. On the independence number of Steiner systems. *Combin. Probab. Comput.*, 22(2):241–252, 2013.

79. C. Even-Zohar and N. Linial. A note on the inducibility of 4-vertex graphs. *Graphs Combin.*, 31(5):1367–1380, 2015.

80. V. Falgas-Ravry and E. R. Vaughan. Applications of the semi-definite method to the Turán density problem for 3-graphs. *Combin. Probab. Comput.*, 22(1):21–54, 2013.

81. D. I. Falikman. Proof of the van der Waerden conjecture on the permanent of a doubly stochastic matrix. *Mat. Zametki*, 29(6):931–938, 957, 1981.

82. D. Fisher. Lower bounds on the number of triangles in a graph. *J. Graph Theory*, 13(4):505–512, 1989.

83. D. Fisher and J. Ryan. Bounds on the number of complete subgraphs. *Discrete Math*, 103(3):313–320, 1992.

84. D. G. Fon-Der-Flaass. On a method of construction of $(3, 4)$-graphs. *Mat. Zametki*, 44(4):546–550, 1988.

85. J. Fox. A new proof of the graph removal lemma. *Ann. of Math. (2)*, 174(1):561–579, 2011.

86. J. Fox and L. Sauermann. A completion of the proof of the edge-statistics conjecture. *Adv. Comb.*, pages Paper No. 4, 52, 2020.

87. J. Fox and B. Sudakov. Dependent random choice. *Random Structures Algorithms*, 38(1-2):68–99, 2011.

88. P. Frankl. On Sperner families satisfying an additional condition. *J. Combin. Theory Ser. A*, 20(1):1–11, 1976.

89. P. Frankl. On intersecting families of finite sets. *Bull. Austral. Math. Soc.*, 21(3):363–372, 1980.

90. P. Frankl. On a problem of Chvátal and Erdős on hypergraphs containing no generalized simplex. *J. Combin. Theory Ser. A*, 30(2):169–182, 1981.

91. P. Frankl. The shifting technique in extremal set theory. In *Surveys in combinatorics 1987 (New Cross, 1987)*, volume 123 of *London Math. Soc. Lecture Note Ser.*, pages 81–110. Cambridge Univ. Press, Cambridge, 1987.

92. P. Frankl. Shadows and shifting. *Graphs Combin.*, 7(1):23–29, 1991.

93. P. Frankl. Improved bounds for Erdős' matching conjecture. *J. Comb. Theory, Ser. A*, 120(5):1068–1072, 2013.

94. P. Frankl. A stability result for the Katona theorem. *J. Combin. Theory Ser. B*, 122:869–876, 2017.

95. P. Frankl. A structural result for 3-graphs. *Discrete Math.*, 340(5):1039–1041, 2017.

96. P. Frankl. Resilient hypergraphs with fixed matching number. *Combinatorica*, 38(5):1079–1094, 2018.

97. P. Frankl and Z. Füredi. A new generalization of the Erdős-Ko-Rado theorem. *Combinatorica*, 3(3-4):341–349, 1983.

98. P. Frankl and Z. Füredi. An exact result for 3-graphs. *Discrete Math.*, 50:323–328, 1984.

99. P. Frankl and Z. Füredi. Forbidding just one intersection. *J. Combin. Theory Ser. A*, 39(2):160–176, 1985.

100. P. Frankl and Z. Füredi. Extremal problems concerning Kneser graphs. *J. Combin. Theory Ser. B*, 40(3):270–284, 1986.

101. P. Frankl and Z. Füredi. Exact solution of some Turán-type problems. *J. Combin. Theory Ser. A*, 45(2):226–262, 1987.

102. P. Frankl and Z. Füredi. Extremal problems whose solutions are the blowups of the small Witt-designs. *J. Comb. Theory, Ser. A*, 52(1):129–147, 1989.

103. P. Frankl and Z. Füredi. A new short proof of the EKR theorem. *J. Combin. Theory Ser. A*, 119(6):1388–1390, 2012.

104. P. Frankl, T. Łuczak, and K. Mieczkowska. On matchings in hypergraphs. *Electron. J. of Combin.*, 19(2):42, 2012.

105. P. Frankl and V. Rödl. Hypergraphs do not jump. *Combinatorica*, 4:149–159, 1984.

106. P. Frankl and N. Tokushige. Some best possible inequalities concerning cross-intersecting families. *J. Combin. Theory Ser. A*, 61(1):87–97, 1992.

107. A. Frieze and M. Karoński. *Introduction to random graphs*. Cambridge University Press, Cambridge, 2016.

108. A. Frieze and D. Mubayi. Coloring simple hypergraphs. *J. Combin. Theory Ser. B*, 103(6):767–794, 2013.

109. Z. Füredi. On finite set-systems whose every intersection is a kernel of a star. *Discrete Math.*, 47(1):129–132, 1983.

110. Z. Füredi. A proof of the stability of extremal graphs, Simonovits' stability from Szemerédi's regularity. *J. Combin. Theory Ser. B*, 115:66–71, 2015.

111. Z. Füredi and L. Özkahya. Unavoidable subhypergraphs: a-clusters. *J. of Combin. theory. Ser. A*, 118(8):2246–2256, 2011.

112. Z. Füredi, O. Pikhurko, and M. Simonovits. On triple systems with independent neighbourhoods. *Combin. Probab. Comput.*, 14(5-6):795–813, 2005.

113. Z. Füredi and M. Simonovits. Triple systems not containing a Fano configuration. *Combin. Probab. Comput.*, 14(4):467–484, 2005.

114. R. Glebov, A. Grzesik, P. Hu, T. Hubai, D. Kral, and J. Volec. Densities of 3-vertex graphs. *arXiv preprint arXiv:1610.02446*, 2016.

115. A. Goodman. On sets of acquaintances and strangers at any party. *Amer. Math. Monthly*, 66(9):778–783, 1959.

116. H.-D. O. F. Gronau. On Sperner families in which no k sets have an empty intersection. *J. Combin. Theory Ser. A*, 28(1):54–63, 1980.

117. W. H. Haemers. Interlacing eigenvalues and graphs. *Linear Algebra Appl.*, 226/228:593–616, 1995.

118. M. Hall, Jr. and J. D. Swift. Determination of Steiner triple systems of order 15. *Math. Tables Aids Comput.*, 9:146–152, 1955.

119. J. Han and Y. Kohayakawa. The maximum size of a non-trivial intersecting uniform family that is not a subfamily of the Hilton-Milner family. *Proc. Amer. Math. Soc.*, 145(1):73–87, 2017.

120. H. Hatami and S. Norin. On the boundary of the region defined by homomorphism densities. *J. Comb.*, 10(2):203–219, 2019.

121. H. Hatami and S. Norine. Undecidability of linear inequalities in graph homomorphism densities. *J. Amer. Math. Soc.*, 24(2):547–565, 2011.

122. D. Hefetz and P. Keevash. A hypergraph Turán theorem via Lagrangians of intersecting families. *J. Combin. Theory Ser. A*, 120(8):2020–2038, 2013.

123. A. Hilton and E. Milner. Some intersection theorems for systems of finite sets. *Quart. J. Math. Oxford*, 18(1):369–384, 1967.

124. J. Hirst. The inducibility of graphs on four vertices. *J. Graph Theory*, 75(3):231–243, 2014.

125. A. Hocquenghem. Codes correcteurs d'erreurs. *Chiffers*, 2:147–156, 1959.

126. P. Hu, B. Lidickýv, T. Martins, S. Norin, and J. Volec. Large multipartite subgraphs in H-free graphs. *manuscript*.

127. H. Huang, N. Linial, H. Naves, Y. Peled, and B. Sudakov. On the 3-local profiles of graphs. *J. Graph Theory*, 76(3):236–248, 2014.

128. H. Huang, P. Loh, and B. Sudakov. The size of a hypergraph and its matching number. *Combin., Probab. and Comput.*, 21(3):442–450, 2012.

129. J. L. W. V. Jensen. Sur les fonctions convexes et les inégalités entre les valeurs moyennes. *Acta Math.*, 30(1):175–193, 1906.

130. T. Jiang, Y. Peng, and B. Wu. Lagrangian densities of some sparse hypergraphs and Turán numbers of their extensions. *European J. Combin.*, 73:20–36, 2018.

131. G. Katona. Intersection theorems for systems of finite sets. *Acta Math. Acad. Sci. Hungar.*, 15:329–337, 1964.

132. G. Katona. A theorem of finite sets. In *Theory of graphs (Proc. Colloq., Tihany, 1966)*, pages 187–207, 1968.

133. G. Katona, T. Nemetz, and M. Simonovits. On a problem of Turán in the theory of graphs. *Mat. Lapok*, 15:228–238, 1964.

134. P. Keevash. Shadows and intersections: stability and new proofs. *Adv. Math.*, 218(5):1685–1703, 2008.

135. P. Keevash. Hypergraph Turán problems. In *Surveys in combinatorics 2011*, volume 392 of *London Math. Soc. Lecture Note Ser.*, pages 83–139. Cambridge Univ. Press, Cambridge, 2011.

136. P. Keevash. The existence of designs. *arXiv preprint arXiv:1401.3665*, 2014.

137. P. Keevash. Counting Steiner triple systems. In *European Congress of Mathematics*, pages 459–481. Eur. Math. Soc., Zürich, 2018.

138. P. Keevash. The existence of designs II. *arXiv preprint arXiv:1802.05900*, 2018.

139. P. Keevash and D. Mubayi. Stability theorems for cancellative hypergraphs. *J. of Combin. Theory. Ser. B*, 92(1):163–175, 2004.

140. P. Keevash and D. Mubayi. Set systems without a simplex or a cluster. *Combinatorica*, 30(2):175–200, 2010.

141. P. Keevash and B. Sudakov. On a hypergraph Turán problem of Frankl. *Combinatorica*, 25(6):673–706, 2005.

142. P. Keevash and B. Sudakov. The Turán number of the Fano plane. *Combinatorica*, 25(5):561–574, 2005.

143. P. Keevash and B. Sudakov. Sparse halves in triangle-free graphs. *J. Combin. Theory Ser. B*, 96(4):614–620, 2006.

144. P. Keevash, B. Sudakov, and J. Verstraëte. On a conjecture of Erdős and Simonovits: even cycles. *Combinatorica*, 33(6):699–732, 2013.

145. N. Keller and N. Lifshitz. The junta method for hypergraphs and the Erdős-Chvátal simplex conjecture. *arXiv preprint arXiv:1707.02643*, 2017.

146. J. H. Kim. The Ramsey number $R(3,t)$ has order of magnitude $t^2/\log t$. *Random Structures Algorithms*, 7(3):173–207, 1995.

147. D. Korándi, A. Roberts, and A. Scott. Exact stability for Turán's theorem. *Adv. Comb.*, pages Paper No. 9, 17, 2021.

148. A. Kostochka and D. Mubayi. The structure of large intersecting families. *Proc. Amer. Math. Soc.*, 145(6):2311–2321, 2017.

149. A. Kostochka, D. Mubayi, and J. Verstraëte. Hypergraph Ramsey numbers: triangles versus cliques. *J. Combin. Theory Ser. A*, 120(7):1491–1507, 2013.

150. A. Kostochka, D. Mubayi, and J. Verstraëte. On independent sets in hypergraphs. *Random Structures Algorithms*, 44(2):224–239, 2014.

151. A. V. Kostochka. A class of constructions for Turán's (3,4)-problem. *Combinatorica*, 2:187–192, 1982.

152. A. V. Kostochka and V. Rödl. On graphs with small Ramsey numbers. *J. Graph Theory*, 37(4):198–204, 2001.

153. M. Krivelevich. On the edge distribution in triangle-free graphs. *J. Combin. Theory Ser. B*, 63(2):245–260, 1995.

154. J. B. Kruskal. The number of simplices in a complex. In *Mathematical optimization techniques*, pages 251–278. Univ. of California Press, Berkeley, Calif., 1963.

155. H. W. Kuhn and A. W. Tucker. Nonlinear programming. In *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, 1950*, pages 481–492. University of California Press, Berkeley and Los Angeles, 1951.

156. M. Kwan, B. Sudakov, and T. Tran. Anticoncentration for subgraph statistics. *J. Lond. Math. Soc. (2)*, 99(3):757–777, 2019.

157. N. Lifshitz. On set systems without a simplex-cluster and the junta method. *J. Combin. Theory Ser. A*, 170:105139, 20, 2020.

158. H. Liu, O. Pikhurko, and K. Staden. The exact minimum number of triangles in graphs with given order and size. *Forum Math. Pi*, 8:e8, 144, 2020.

159. X. Liu. Cancellative hypergraphs and steiner triple systems. *arXiv preprint arXiv:1912.11917*, 2019.

160. X. Liu. d-cluster-free sets with a given matching number. *European J. Combin.*, 82:103000, 19, 2019.

161. X. Liu. A note on hypergraphs without non-trivial intersecting subgraphs. *arXiv preprint arXiv:2007.11055*, 2020.

162. X. Liu. Structural results for conditionally intersecting families and some applications. *Electron. J. Combin.*, 27(2):Paper No. 2.33, 13, 2020.

163. X. Liu. New short proofs to some stability theorems. *European J. Combin.*, 96:Paper No. 103350, 8, 2021.

164. X. Liu and J. Ma. Sparse halves in K_4-free graphs. *J. Graph Theory*, 99(1):5–25, 2022.

165. X. Liu and D. Mubayi. A note on explicit constructions of designs. *Electron. J. Combin.* Accepted.

166. X. Liu and D. Mubayi. On a generalized Erdős-Rademacher problem. *J. Graph Theory*, 2020.

167. X. Liu and D. Mubayi. The feasible region of hypergraphs. *J. Comb. Theory, Ser. B*, 148:23–59, 2021.

168. X. Liu and D. Mubayi. Tight bounds for Katona's shadow intersection theorem. *European J. Combin.*, 97:Paper No. 103391, 17, 2021.

169. X. Liu and D. Mubayi. A hypergraph Turán problem with no stability. *Combinatorica*, pages 1–30, 2022.

170. X. Liu, D. Mubayi, and C. Reiher. Hypergraphs with many extremal configurations II. In preparation.

171. X. Liu, D. Mubayi, and C. Reiher. A unified approach to hypergraph stability. submitted.

172. X. Liu, D. Mubayi, and C. Reiher. The feasible region of induced graphs. *arXiv preprint arXiv:2106.16203*, 2021.

173. X. Liu, D. Mubayi, and C. Reiher. Hypergraphs with many extremal configurations. *arXiv preprint arXiv:2102.02103*, 2021.

174. P.-S. Loh. A note on embedding hypertrees. *Electron. J. Combin.*, 16(1):Note 18, 4, 2009.

175. L. Lovász. *Combinatorial Problems and Exercises.* North-Holland Publishing Co., Amsterdam, second edition, 1993.

176. L. Lovász. *Combinatorial Problems and Exercises*, volume 361. American Mathematical Soc., 2007.

177. L. Lovász. *Large networks and graph limits*, volume 60. American Mathematical Soc., 2012.

178. L. Lovász and M. Simonovits. On the number of complete subgraphs of a graph. II. In *Studies in pure mathematics*, pages 459–495. Birkhäuser, Basel, 1983.

179. L. Lu and L. Székely. Using Lovász local lemma in the space of random injections. *Electron. J. Combin.*, 14(1):Research Paper 63, 13, 2007.

180. C. M. Lüders and C. Reiher. The Ramsey-Turán problem for cliques. *Israel J. Math.*, 230(2):613–652, 2019.

181. C. M. Lüders and C. Reiher. Weighted variants of the Andrásfai-Erdős-Sós theorem. *J. Comb.*, 11(2):305–328, 2020.

182. J. Lyle. A note on independent sets in graphs with large minimum degree and small cliques. *Electron. J. Combin.*, 21(2):Paper 2.38, 17, 2014.

183. J. Ma and X. Liu. Calculations for "sprase halves in K_4-free graphs". `http://staff.ustc.edu.cn/~jiema/localdensitycal.pdf`. Accessed May 27, 2020.

184. J. Ma and Y. Qiu. Some sharp results on the generalized Turán numbers. *European J. Combin.*, 84:103026, 16, 2020.

185. A. Mammoliti and T. Britz. On Mubayi's conjecture and conditionally intersecting sets. *SIAM J. Discrete Math.*, 32(3):2361–2380, 2018.

186. W. Mantel. Problem 28. *Wiskundige Opgaven*, 10(60-61):320, 1907.

187. A. Martinsson, F. Mousset, A. Noever, and M. Trujić. The edge-statistics conjecture for $\ell \ll k^{6/5}$. *Israel J. Math.*, 234(2):677–690, 2019.

188. J. Moon and L. Moser. On a problem of Turán. *Magyar Tud. Akad. Mat. Kutató Int. Közl*, 7:283–286, 1962.

189. T. S. Motzkin and E. Straus. Maxima for graphs and a new proof of a theorem of Turán. *Canad. J. Math.*, 17:533–540, 1965.

190. D. Mubayi. Erdős-Ko-Rado for three sets. *J. of Combin. theory. Ser. A*, 113(3):547–550, 2006.

191. D. Mubayi. A hypergraph extension of Turán's theorem. *J. Combin. Theory, Ser. B*, 96(1):122–134, 2006.

192. D. Mubayi. An intersection theorem for four sets. *Adv. Math.*, 215(2):601–615, 2007.

193. D. Mubayi. Structure and stability of triangle-free set systems. *Trans. Amer. Math. Soc.*, 359(1):275–291, 2007.

194. D. Mubayi. Counting substructures I: color critical graphs. *Adv. Math.*, 225(5):2731–2740, 2010.

195. D. Mubayi and O. Pikhurko. A new generalization of Mantel's theorem to k-graphs. *J. Comb. Theory, Ser. B*, 97(4):669–678, 2007.

196. D. Mubayi, O. Pikhurko, and B. Sudakov. Hypergraph Turán problem: Some open questions, 2011.

197. D. Mubayi and R. Ramadurai. Set systems with union and intersection constraints. *J. of Combin. theory. Ser. B*, 99(3):639–642, 2009.

198. D. Mubayi and J. Verstraëte. Proof of a conjecture of Erdős on triangles in set-systems. *Combinatorica*, 25(5):599–614, 2005.

199. D. Mubayi and J. Verstraëte. A survey of Turán problems for expansions. In *Recent Trends in Combinatorics*, pages 117–143. Springer, 2016.

200. J. Nešetřil and V. Rödl, editors. *Mathematics of Ramsey theory*, volume 5 of *Algorithms and Combinatorics*. Springer-Verlag, Berlin, 1990.

201. V. Nikiforov. The number of cliques in graphs of given order and size. *Trans. Amer. Math. Soc.*, 363(3):1599–1618, 2011.

202. N. Nisan and A. Wigderson. Hardness vs. randomness. *J. Comput. System Sci.*, 49(2):149–167, 1994.

203. S. Norin and L. Yepremyan. Sparse halves in dense triangle-free graphs. *J. Combin. Theory Ser. B*, 115:1–25, 2015.

204. S. Norin and L. Yepremyan. Turán number of generalized triangles. *J. Combin. Theory Ser. A*, 146:312–343, 2017.

205. S. Norin and L. Yepremyan. Turán numbers of extensions. *J. Combin. Theory Ser. A*, 155:476–492, 2018.

206. D. Olpp. A conjecture of Goodman and the multiplicities of graphs. *Australas. J. Combin.*, 14:267–282, 1996.

207. P. P. Pach and R. Palincza. Sets avoiding six-term arithmetic progressions in $\mathbb{Z}_6^n$ are exponentially small. *arXiv preprint arXiv:2009.11897*, 2020.

208. O. Pikhurko. An exact Turán result for the generalized triangle. *Combinatorica*, 28(2):187–208, 2008.

209. O. Pikhurko. Exact computation of the hypergraph Turán function for expanded complete 2-graphs. *J. Combin. Theory, Ser. B*, 103(2):220–225, 2013.

210. O. Pikhurko. On possible Turán densities. *Israel J. Math.*, 201(1):415–454, 2014.

211. N. Pippenger and M. C. Golumbic. The inducibility of graphs. *J. Combinatorial Theory Ser. B*, 19(3):189–203, 1975.

212. D. K. Ray-Chaudhuri and R. M. Wilson. On t-designs. *Osaka Math. J.*, 12(3):737–744, 1975.

213. R. Raz, O. Reingold, and S. Vadhan. Extracting all the randomness and reducing the error in Trevisan's extractors. In *Annual ACM Symposium on Theory of Computing (Atlanta, GA, 1999)*, pages 149–158. ACM, New York, 1999.

214. A. Razborov. On 3-hypergraphs with forbidden 4-vertex configurations. *SIAM J. Discrete Math.*, 24(3):946–963, 2010.

215. A. A. Razborov. Flag algebras. *J. Symbolic Logic*, 72(4):1239–1282, 2007.

216. A. A. Razborov. On the minimal density of triangles in graphs. *Combin. Probab. Comput.*, 17(4):603–618, 2008.

217. C. Reiher. The clique density theorem. *Ann. of Math. (2)*, 184(3):683–707, 2016.

218. C. Reiher. K_4-free graphs have sparse halves. *arXiv preprint arXiv:2108.07297*, 2021.

219. C. Reiher and S. Wagner. Maximum star densities. *Studia Sci. Math. Hungar.*, 55(2):238–259, 2018.

220. R. A. Rödl, V. and A. Taraz. Hypergraph packing and graph embedding. *Combin. Probab. Comput.*, 8(4):363–376. Random graphs and combinatorial structures (Oberwolfach, 1997).

221. V. Rödl, A. Ruciński, and E. Szemerédi. Perfect matchings in large uniform hypergraphs with large minimum collective degree. *J. Combin. Theory Ser. A*, 116(3):613–636, 2009.

222. V. Rödl and E. Šiňajová. Note on independent sets in Steiner systems. In *Proceedings of the Fifth International Seminar on Random Graphs and Probabilistic Methods in Combinatorics and Computer Science (Poznań, 1991)*, volume 5, pages 183–190, 1994.

223. W. Rudin. Principles of mathematical analysis. 3rd ed. International Series in Pure and Applied Mathematics. Düsseldorf etc.: McGraw-Hill Book Company. X, 342 p. DM 47.80 (1976)., 1976.

224. A. Sali. Some intersection theorems. *Combinatorica*, 12(3):351–361, 1992.

225. R. H. Schelp and A. Thomason. A remark on the number of complete and empty subgraphs. *Combin. Probab. Comput.*, 7(2):217–219, 1998.

226. J. B. Shearer. On the independence number of sparse graphs. *Random Structures Algorithms*, 7(3):269–271, 1995.

227. J. B. Shearer. A new construction for cancellative families of sets. *Electron. J. Combin.*, 3(1):3, 1996.

228. A. Sidorenko. What we know and what we do not know about Turán numbers. *Graphs Comb.*, 11(2):179–199, 1995.

229. A. Sidorenko. Extremal problems on the hypercube and the codegree Turán density of complete r-graphs. *SIAM J. Discrete Math.*, 32(4):2667–2674, 2018.

230. A. Sidorenko. On generalized Erdős-Ginzburg-Ziv constants for $\mathbb{Z}_2^d$. *J. Combin. Theory Ser. A*, 174:105254, 20, 2020.

231. A. F. Sidorenko. The maximal number of edges in a homogeneous hypergraph containing no prohibited subgraphs. *Math Notes*, 41(3):247–259, 1987.

232. A. F. Sidorenko. On the maximal number of edges in a homogeneous hypergraph that does not contain prohibited subgraphs. *Mat. Zametki*, 41(3):433–455, 459, 1987.

233. M. Simonovits. A method for solving extremal problems in graph theory, stability problems. In *Theory of Graphs (Proc. Colloq., Tihany, 1966)*, pages 279–319. Academic Press, New York, 1968.

234. J. Spencer. Turán's theorem for k-graphs. *Discrete Math.*, 2:183–186, 1972.

235. E. M. Stein and R. Shakarchi. *Real analysis. Measure theory, integration, and Hilbert spaces.* Princeton, NJ: Princeton University Press, 2005.

236. B. Sudakov. A few remarks on Ramsey-Turán-type problems. *J. Combin. Theory Ser. B*, 88(1):99–106, 2003.

237. B. Sudakov. Making a K_4-free graph bipartite. *Combinatorica*, 27(4):509–518, 2007.

238. A. Thomason. A disproof of a conjecture of Erdős in Ramsey theory. *J. London Math. Soc. (2)*, 39(2):246–255, 1989.

239. F. Tian and Z.-L. Liu. Bounding the independence number in some (n, k, ℓ, λ)-hypergraphs. *Graphs Combin.*, 34(5):845–861, 2018.

240. L. Trevisan. Extractors and pseudorandom generators. *J. ACM*, 48(4):860–879, 2001.

241. P. Turán. Eine Extremalaufgabe aus der Graphentheorie. *Mat. Fiz. Lapok*, 48:436–452, 1941.

242. P. Turán. Eine Extremalaufgabe aus der Graphentheorie. *Mat. Fiz. Lapok*, 48:436–452, 1941.

243. R. Wilson. The early history of block designs. *Rend. Semin. Mat. Messina, Ser. II*, 9:267–276, 2004.

244. R. M. Wilson. An existence theory for pairwise balanced designs. I. Composition theorems and morphisms. *J. Comb. Theory, Ser. A*, 13:220–245, 1972.

245. R. M. Wilson. An existence theory for pairwise balanced designs. II. The structure of PBD-closed sets and the existence conjectures. *J. Comb. Theory, Ser. A*, 13:246–273, 1972.

246. R. M. Wilson. Nonisomorphic Steiner triple systems. *Math. Z.*, 135:303–313, 1973/74.

247. R. M. Wilson. An existence theory for pairwise balanced designs. III. Proof of the existence conjectures. *J. Comb. Theory, Ser. A*, 18:71–79, 1975.

248. E. Witt. Über Steinersche Systeme. *Abh. Math. Sem. Univ. Hamburg*, 12(1):265–275, 1937.

249. C. Xiao and G. O. H. Katona. The number of triangles is more when they have no common vertex. *Discrete Math.*, 344(5):Paper No. 112330, 6, 2021.

250. X. Y. Zha. On a conjecture on the Sperner property. *European J. Combin.*, 10(6):603–607, 1989.

251. V. A. Zorich. *Mathematical analysis I. Transl. from the 4th Russian editon by Roger Cooke.* Berlin: Springer, 2003.

252. A. A. Zykov. On some properties of linear complexes. *Mat. Sbornik N.S.*, 24(66):163–188.

VITA

Education

2018–2022: Ph.D. in Mathematics, University of Illinois at Chicago.

2018–2020: M.S. in Mathematics, University of Illinois at Chicago.

2014–2018: B.S. in Mathematics and Applied Mathematics, University of Science and Technology of China.

Publications

Hypergraphs with many extremal configurations II (with D. Mubayi and C. Reiher), in preparation.

The feasible region of induced graphs (with D. Mubayi and C. Reiher), arXiv:2106.16203, submitted.

A unified approach to hypergraph stability (with D. Mubayi and C. Reiher), arXiv:2104.00167, submitted.

Hypergraphs with many extremal configurations (with D. Mubayi and C. Reiher), arXiv:2102.02103, submitted.

A note on hypergraphs without non-trivial intersecting subgraphs, arXiv:2007.11055, submitted.

Stability theorems for some Kruskal–Katona type results (with S. Mukherjee), arXiv:2006.04848, submitted.

Cancellative hypergraphs and Steiner triple systems, arXiv:1912.11917, submitted.

A note on explicit constructions of designs (with D. Mubayi), accepted, Electronic Journal of Combinatorics.

Independent sets in hypergraphs omitting an intersection (with T. Bohman and D. Mubayi), accepted, Random Structures & Algorithm.

On a generalized Erdős–Rademacher problem (with D. Mubayi), accepted, Journal of Graph Theory.

Sparse halves in K_4-free graphs (with J. Ma), accepted, Journal of Graph Theory.

A hypergraph Turán problem with no stability (with D. Mubayi), accepted, Combinatorica.

Tight bounds for Katona's shadow intersection theorem (with D. Mubayi), *European J. Combin., 97:103391, 2021.*

The feasible region of hypergraphs (with D. Mubayi), *J. Combin. Theory, Ser. B, 148:23 - 59, 2021.*

New short proofs to some stability theorems, *European J. Combin., 96:103350, 2021*

Structural results for conditionally intersecting families and some applications, *Electron. J. Combin., 33(2), 2020.*

d-cluster-free sets with a given matching number, *European J. Combin., 82:103000, 2019.*

Publisher's Copyright Statements

Part of the materials in Chapter 5 was accepted in March 2021 for publication in Combinatorica, Springer. Their website states:

> Authors have the right to reuse their articles Version of Record, in whole or in part, in their own thesis. Additionally, they may reproduce and make available their thesis, including Springer Nature content, as required by their awarding academic institution. Authors must properly cite the published article in their thesis according to current citation standards.

Materials in Chapter 7 were published on the Journal of Graph Theory, John Wiley & Sons, in July 2021, and on the Journal of Graph Theory, John Wiley & Sons, in November 2021. Part of the materials in Chapter 9 was published in Random Structures & Algorithms, John Wiley & Sons, in December 2021. Their website states:

> Authors may post their submitted manuscript at any time on their personal website, in their company or institutional repository, in not-for-profit subject-based preprint servers or repositories, and on SCNs which have signed up to the STM sharing principles. Authors may wish to add a note about acceptance by the journal and upon publication it is recommended that authors add a Digital Object Identifier (DOI) link back to the final article so that users with access to Wiley Online Library may access the final version of record and make use of the fully networked research environment. The deposited version may not be updated to look like the final version of record. While this is the copyright policy, individual journals may operate

> different editorial policies with respect to preprints and authors should consult the relevant author guidelines.

Part of the materials in Chapter 6 was published in the Electronic Journal of Combinatorics in May 2020. Part of the materials in Chapter 9 was accepted for publication in December 2021 on the Electronic Journal of Combinatorics. Their website states:

> The journal is completely free for both authors and readers. Authors retain the copyright of their papers and are strongly encouraged to release their articles under a Creative Commons license. All papers have DOIs, are reviewed on MathSciNet and ZBMath, and are indexed by Web of Science.