2019

Extremal Problems for Forests in Graphs and Hypergraphs

Omid Khormali

Extremal Problems for Forests in Graphs and Hypergraphs

By

Omid Khormali

B.S., University of Tabriz, Tabriz, Iran
M.S., Tarbiat Modares University, Tehran, Iran
M.S., University of Montana, Missoula, USA

Dissertation

presented in partial fulfillment of the requirements
for the degree of

Doctor of Philosophy
in Mathematics

University of Montana
Missoula, MT

December 2019

Approved by:

Scott Whittenburg, Dean of the Graduate School
Graduate School

Dr. Cory Palmer, Chair
Department of Mathematical Sciences, University of Montana

Dr. Mark Kayll
Department of Mathematical Sciences, University of Montana

Dr. Brian Steele
Department of Mathematical Sciences, University of Montana

Dr. Daniel Johnston
Mathematics and Statistics Department, Skidmore College

Dr. John Chandler
Department of Management and Marketing, University of Montana

بِسْمِ ٱللَّهِ ٱلرَّحْمَٰنِ ٱلرَّحِيمِ

"In the name of God, the Most Gracious, the Most Merciful"

To My Beloved Wife

Khormali, Omid, Doctorate of Philosophy, December 2019 Mathematics

Extremal problems for forests in graphs and hypergraphs

Committee Chair: Cory Palmer, Ph.D.

The *Turán number*, $\mathrm{ex}_r(n, F)$, of an r-uniform hypergraph F is the maximum number of hyperedges in an n-vertex r-uniform hypergraph which does not contain F as a subhypergraph. Note that when $r = 2$, $\mathrm{ex}_r(n, F) = \mathrm{ex}(n, F)$ which is the Turán number of graph F. We study Turán numbers in the degenerate case for graphs and hypergraphs; we focus on the case when F is a forest in graphs and hypergraph. In the first chapter we discuss the history of Turán numbers and give several classical results. In the second chapter, we examine the Turán number for forests with path components, forests of path and star components, and forests with all components of order 5. In the third chapter we determine the Turán number of an r-uniform "star forest" in various hypergraph settings.

Acknowledgments

I would like to thank my advisor, Cory Palmer, for the unconditional support of my Ph.D study and related research, for his patience, motivation, and immense knowledge. Without his guidance and constant feedback this dissertation would not have been achievable.

Besides my advisor, I would like to thank the rest of my dissertation committee: Mark Kayll, Brian Steele, Daniel Johnston and John Chandler, for their insightful comments and encouragement.

I would like to thank the University of Montana (UM) for supporting me through a teaching assistantship during my studies, a UM Graduate Student Summer Research Award in four summers of my studies and a Bertha Morton Scholarship, and helping me to receive a funded summer internship in US Forestry Service. Also, I am very thankful for having the opportunity to getting a master degree in data science at UM beside of my Ph.D degree.

Last but not the least, I would like to express my sincere gratitude to my wife, Fereshteh, for her non-stopping encouragements, strong support, and extremely patience.

Contents

Chapter 1

Introduction

1.1 Graphs and hypergraphs

Informally, a graph is constructed from a set of points (called vertices) and lines joining pairs of vertices (called edges). Graphs are often used to model network-like structures. For example, a graph can be used to model the US highway system in the following way: major cities are represented by vertices and we join two cities by an edge if there is a highway between them. Another example is the social network Facebook where we can represent each user by a vertex and two users are joined by an edge if they are friends.

Mathematically, a *graph* G consists of a vertex set $V(G)$ and an edge set of $E(G)$ which is a subset of the collection of all pairs of elements from $V(G)$. The *order* of a graph is the number of its vertices and the *size* of a graph is the number of its edges. We denote the number of edges of a graph G by $e(G) = |E(G)|$. We often use the term *n-vertex* for a graph with n vertices. All graphs in this dissertation are *simple*, i.e., we do not allow multiple edges between the same pair of vertices and we do not allow an edge between a vertex and itself. The *complete* graph on n vertices, denoted by K_n, is an n-vertex graph with all $\binom{n}{2}$ possible edges between vertices. A *subgraph* H of a graph G is a graph whose vertex set and edge set are subsets of those of G; i.e., $V(H) \subseteq V(G)$ and $E(H) \subseteq E(G)$. A *clique* of a graph is a complete subgraph of the graph. The *complete* r*-partite* graph $K_{n_1,n_2,\ldots,n_r}$ is a graph with r parts with no edges in a part and all possible edges between parts. A *path* is a graph whose vertices can be linearly ordered so that two vertices are adjacent if and only if they are consecutive in the ordering. The *endpoints* of a path are the first and last vertices in such an order. We denote a path on n vertices by P_n and its *length* is the number of edges, i.e., $n-1$. A *cycle* is a graph with an equal number of vertices and edges whose vertices can be placed around a circle so that two vertices are adjacent if and only if they appear consecutively along the circle. The *degree* of a vertex in a graph is the number of edges incident to the vertex. A star, S_k, is a connected graph with $k+1$ vertices such that one vertex is of degree k and the other k vertices are of degree 1. A *tree* is a connected graph with no cycle. Alternatively, a tree is a graph in which any two vertices are connected by exactly one path. And a *forest* is a graph whose components are trees. The notation $k \cdot G$ denotes k vertex-disjoint copies of graph G. For two graphs G and H, $G \cup H$ is a graph with the vertex set $V(G \cup H) = V(G) \cup V(H)$ and the edge set $E(G \cup H) = E(G) \cup E(H)$. The graph $G + H$ has vertex set $V(G+H) = V(G) \cup V(H)$ and edge set $E(G+H) = E(G) \cup E(H) \cup \{xy \mid x \in V(G) \text{ and } y \in V(H)\}$. Also, $e(G, H)$ is the number of edges with one end in G and the other in H (for graphs G, H or even vertex sets G, H).

The *chromatic number* of a graph G is the smallest number of colors needed to color the vertices of G so that no two adjacent vertices share the same color; it is usually denoted by $\chi(G)$. For example, $\chi(P_n) = 2$ and $\chi(C_n) = 2$ when n is even and $\chi(C_n) = 3$ when n is odd. A *bipartite* graph is a graph whose vertices can be divided into two parts such that the edges are only between these two parts. Equivalently, a bipartite graph is a graph that does not contain an odd-length cycle. Note that trees and forests are bipartite graphs.

A *hypergraph* $\mathcal{H}$ is a generalization of a graph where edges (called *hyperedges*) are subsets of the vertex set of any size. More formally, a hypergraph $\mathcal{H}$ consists of a vertex set $V(\mathcal{H})$

and an edge set $E(\mathcal{H})$ made up of subsets of $V(\mathcal{H})$. The notion of *subhypergraph* is defined analgously to graphs. A hypergraph is *r-uniform* if all of its edges are of order r. Note that graphs are 2-uniform hypergraphs. A hypergraph $\mathcal{H}$ is *linear* if each pair of hyperedges intersects in at most one vertex. To distinguish graphs and hypergraphs, we represent graphs by uppercase letters and hypergraphs by script letters. For notation not otherwise defined see the monograph of Bollobás [2].

1.2 Extremal graph and hypergraph theory

Extremal graph theory is a branch of graph theory that has developed rapidly in recent decades. It deals with the problem of determining or estimating the maximum or minimum possible number of edges in a graph which satisfies certain requirements. Often these problems are related to other areas including theoretical computer science, information theory and number theory. The *Turán number*, $\mathrm{ex}(n, F)$, of a graph F is the maximum number of edges in an n-vertex graph which does not contain F as a subgraph. An *extremal graph* for F is an n-vertex graph containing no subgraph F with $\mathrm{ex}(n, F)$ many edges. We use the term *F-free* to describe a graph that does not contain F as a subgraph. The problem of determining Turán numbers is a long-standing field of study in graph theory and it has played an important role in the development of extremal combinatorics. The prototypical result in extremal graph theory is a theorem of Mantel that the maximum number of edges in an n-vertex triangle-free graph is $\lfloor \frac{n^2}{4} \rfloor$. Turán [26] proved a generalization of this result by characterizing the extremal graph for any complete graph K_k. The *Turán graph* $T_{k-1}(n)$ is the graph on n vertices divided into $k-1$ classes of orders as equal as possible where each class is an independent set and all edges between every pair of classes are present.

Theorem 1.2.1 (Turán, [26]). *The maximum number of edges in an n-vertex graph with no complete graph K_k subgraph is exactly $e(T_{k-1}(n))$. Furthermore, $T_k(n)$ is the unique graph attaining this maximum.*

A weaker version of Turán's Theorem states: the number of edges in a graph on n vertices with no K_k subgraph is at most

$$\mathrm{ex}(n, K_k) \le \left(1 - \frac{1}{k-1}\right)\frac{n^2}{2}.$$

The Erdős-Stone-Simonovits theorem states that asymptotically Turán's construction is best-possible for any k-chromatic graph F as long as $k > 2$.

Theorem 1.2.2 (Erdős and Stone, [13]; Erdős and Simonovits, [12])**.** *If F is a graph with chromatic number k, then*

$$\mathrm{ex}(n, F) = \left(1 - \frac{1}{k-1}\right)\frac{n^2}{2} + o(n^2).$$

Thus when F is bipartite, then $k = 2$ and the Erdős-Stone-Simonovits theorem states only that $\mathrm{ex}(n, H) = o(n^2)$. In other words, for each fixed positive real ϵ, there exists n_0 such that for $n > n_0$, if G is an n-vertex graph with at least ϵn^2 edges, then G contains F as a subgraph. Bipartite graphs are often called "degenerate" in extremal problems for graphs [15]. A major area of research is to determine the behavior of $\mathrm{ex}(n, F)$ for bipartite graphs F.

The even cycle theorem of Bondy and Simonovits [3] gives an upper bound for even cycles.

Theorem 1.2.3 (Bondy and Simonovits, [3])**.** *For each $\ell \ge 2$, there exists $c_\ell > 0$ such that*

$$\mathrm{ex}(n, C_{2\ell}) \le c_\ell n^{1+1/\ell}.$$

Erdős and Gallai [10] determined the Turán number for paths P_ℓ.

Theorem 1.2.4 (Erdős and Gallai, [10])**.**

$$\mathrm{ex}(n, P_\ell) \le \frac{\ell - 2}{2} n.$$

A matching lower-bound is given by disjoint copies of the complete graph $K_{\ell-1}$.

Erdős and Sós in [8] conjectured that the bound on the Turán number for paths should hold for trees on the same number of vertices.

Conjecture 1.2.1 (Erdős and Sós, [8])**.** *Let T be a tree on ℓ vertices. Then*

$$\mathrm{ex}(n, T) \leq \frac{\ell - 2}{2} n.$$

A proof of the conjecture for large trees was announced by Ajtai, Komlós, Simonovits and Szemerédi in an unpublished work. The conjecture is known to hold for various classes of small trees.

The *Turán number*, $\mathrm{ex}_r(n, F)$, of an r-uniform hypergraph F is the maximum number of hyperedges in an n-vertex r-uniform hypergraph that does not contain F as a subhypergraph. In contrast with the graph case, there is relatively little understanding of the hypergraph Turán problem. Turán posed the question to determine $\mathrm{ex}_r(n, F)$ when $F = K_k^r$ is a complete r-uniform hypergraph on k vertices. The case with $k > r > 2$ of this question is still open, and no asymptotically sharp bounds are known.

Fix a graph F and let $r \geq 3$ be an integer. The *r-uniform expansion* of F is the r-uniform hypergraph F^+ obtained from F by enlarging each edge of F with $r - 2$ new vertices disjoint from $V(F)$ such that each edge of F is enlarged by a distinct set of vertices. We write

$$\mathrm{ex}_r(n, F^+)$$

for the maximum number of edges in an n-vertex r-uniform hypergraph that does not contain a subhypergraph F^+. Mubayi and Verstraëte [25] showed the following general bounds.

Theorem 1.2.5 (Mubayi and Verstraëte, [25]). *If $\chi(F) = k > r$, then*

$$\mathrm{ex}_r(n, F^+) = \binom{k-1}{r}\left(\frac{n}{k-1}\right)^r + o(n^r),$$

and if $\chi(F) = k \leq r$ then

$$\mathrm{ex}_r(n, F^+) = o(n^r).$$

Kostochka, Mubayi and Verstraëte determined the Turán number of P_ℓ^+ and C_ℓ^+.

Theorem 1.2.6 (Kostochka, Mubayi and Verstraëte, [22]). *For fixed $r \geq 3$, $\ell \geq 4$ with $(\ell, r) \neq (4, 3)$, and for large enough n*

$$\mathrm{ex}_r(n, P_\ell^+) = \mathrm{ex}_r(n, C_\ell^+) = \binom{n}{r} - \binom{n - \lfloor \frac{\ell-1}{2} \rfloor}{r} + \begin{cases} 0 & \text{if } \ell \text{ is odd} \\ \binom{n - \lfloor \frac{\ell-1}{2} \rfloor - 2}{r-2} & \text{if } \ell \text{ is even.} \end{cases}$$

Another example is due to Chung and Frankl [5] who determined the Turán number of the expansion of a star S_ℓ^+ in 3-uniform hypergraphs for n large enough.

Theorem 1.2.7 (Chung and Frankl, [5]).

$$\mathrm{ex}_3(n, S_\ell^+) = \begin{cases} n\ell(\ell-1) + 2\binom{\ell}{3} & \text{if } n > \ell(\ell-1)(5\ell+2)/2 \text{ and } \ell \geq 3 \text{ is odd} \\ \frac{n\ell(2\ell-3)}{2} - \frac{2\ell^3 - 9\ell + 6}{2} & \text{if } n \geq 2\ell^3 - 9\ell + 7 \text{ and } \ell \geq 4 \text{ is even.} \end{cases}$$

Let F be a graph and $\mathcal{H}$ be a hypergraph. The hypergraph $\mathcal{H}$ is a *Berge-F* if there is a bijection $f : E(F) \to E(\mathcal{H})$ such that $e \subseteq f(e)$ for every $e \in E(F)$. Equivalently, $\mathcal{H}$ is *Berge-F* if we can embed a distinct graph edge into each hyperedge of $\mathcal{H}$ to obtain a copy of F. Note that a Berge-F refers to a class of hypergraphs. This definition was introduced by Gerbner and Palmer [17] to generalize the established concepts of "Berge path" and "Berge cycle" to general graphs.

For a fixed graph F, if a hypergraph $\mathcal{H}$ has no subhypergraph isomorphic to any Berge-F we say that $\mathcal{H}$ is *Berge-F-free*. We denote the maximum number of hyperedges in an n-vertex r-uniform Berge-F-free hypergraph by

$$\mathrm{ex}_r(n, \text{Berge-}F).$$

As the expansion F^+ is a specific Berge-F we have the following trivial inequality: $\mathrm{ex}_r(n, \text{Berge-}F) \leq \mathrm{ex}_r(n, F^+)$. Moreover, the construction used in Theorem 1.2.5 is Berge-F-free. Thus, for $\chi(F) = k > r$, we have

$$\mathrm{ex}_r(n, \text{Berge-}F) = \binom{k-1}{r}\left(\frac{n}{k-1}\right)^r + o(n^r).$$

Similar to the situation in graphs we call the case $\chi(F) = k \leq r$ the "degenerate case" for the Turán number of a Berge-F.

A theorem of Győri, Katona and Lemons [19] together with a theorem of Davoodi, Győri, Methuku and Tompkins [6] gives the Turán number for Berge-paths:

$$\mathrm{ex}_r(n, \text{Berge-}P_{\ell+1}) \leq \begin{cases} \frac{n}{\ell}\binom{\ell}{r} & \text{if } \ell \geq r+1 > 3 \\ \frac{n(\ell-1)}{r+1} & \text{if } r \geq \ell > 2. \end{cases}$$

Both bounds above are sharp when n has appropriate divisibility conditions.

Győri and Lemons gave upper-bounds for Berge-even and Berge-odd cycles [20].

Theorem 1.2.8 (Győri and Lemons, [20])**.** *For all $\ell \geq 2$ and $r \geq 3$,*

$$\mathrm{ex}_r(n, \textit{Berge-}C_{2\ell}) \leq O(n^{1+1/\ell})$$

and

$$\mathrm{ex}_r(n, \textit{Berge-}C_{2\ell+1}) \leq O(n^{1+1/\ell}).$$

Gerbner, Methuku and Palmer proved a sharp result for Berge-trees under the assumption that the Erdős-Sós conjecture 1.2.1 holds.

Theorem 1.2.9 (Gerbner, Methuku and Palmer, [16]). *Let us suppose that the Erdős-Sós conjecture holds for all trees. Let T be a tree on $\ell + 1$ vertices. If $\ell > r + 1 > 3$, then*

$$\mathrm{ex}_r(n, \textit{Berge-}T) \leq \frac{n}{\ell}\binom{\ell}{r}.$$

Moreover, if ℓ divides n, this bound is sharp. If $\ell \leq r + 1$, then

$$\mathrm{ex}_r(n, \textit{Berge-}T) \leq \frac{\ell}{2}n.$$

Without the assumption of Conjecture 1.2.1 they proved the following.

Theorem 1.2.10 (Gerbner, Methuku and Palmer, [16]). *Let T be a tree on ℓ vertices with maximum degree $\Delta(T)$. If $\ell - 1 \leq r$, then we have $\mathrm{ex}_r(n, \text{Berge-}T) \leq (\Delta(T) - 1)n$.*

Motivated by these results, we study Turán numbers in the degenerate case for graphs and hypergraphs. We primarily focus on the case when F is a forest. In the second chapter, we examine the Turán number for forests with path components, forests of path and star components, and forests with all components of order 5. In the third chapter we determine the Turán number of an r-uniform “star forest” in various settings.

Chapter 2

Graphs

2.1 Forest of linear paths

We begin by recalling the Erdős-Gallai theorem for the extremal number of paths.

Theorem 2.1.1 (Erdős-Gallai, [10])**.** *For any* ℓ *and* $n > 1$, $\mathrm{ex}(n, P_\ell) \leq \frac{\ell-2}{2}n$. *Moreover, equality holds when* $\ell - 1$ *divides* n *as seen by an* n*-vertex graph composed of vertex-disjoint copies of* $K_{\ell-1}$.

Faudree and Schelp [14] gave an extension of the Erdős-Gallai theorem. Here and throughout this section E_n denotes the graph on n vertices with no edges, i.e., the complement of the complete graph K_n.

Theorem 2.1.2 (Faudree and Schelp, [14])**.** *Let* G *be a graph with* $|V(G)| = n = q\ell + r$ *such that* $0 \leq q$ *and* $0 \leq r < \ell$. *If* G *contains no* $P_{\ell+1}$, *then*

$$e(G) \leq q\binom{\ell}{2} + \binom{r}{2}.$$

Furthermore, equality holds if and only if

1. $G \cong \cup_{i=1}^{q} K_\ell \cup K_r$, *or*

2. ℓ *is odd,* $q > 0$, $r = (\ell+1)/2$ *or* $(\ell-1)/2$, *and*

$$G \cong \cup_{i=1}^{s} K_\ell \cup \left(K_{(\ell-1)/2} + E_{(\ell+1)/2+(q-s-1)\ell+r}\right) \qquad \textit{(for all integers } s \in [0, q-1]\textit{)}.$$

A generalization of the problem is the determination of the Turán number of a forest whose components are paths. When all paths are of length 1, then we have a forest that is a matching. In this case, Erdős and Sós [8] proved that

$$\mathrm{ex}(n, k \cdot K_2) = \max\left\{\binom{k-1}{2} + (k-1)(n-k+1), \binom{2k-1}{2}\right\}.$$

An extremal graph when n is large enough compared to k is $K_{k-1} + E_{n-k+1}$, i.e., an independent set of order $n-k+1$ together with $k-1$ vertices each of degree $n-1$. We will sometimes call a vertex *universal* if it is adjacent to all other vertices, i.e. it has degree $n-1$. This construction forms a model for the constructions presented throughout this dissertation.

Gorgol [18] presented constructions giving the following lower bound for the Turán number of a forest of paths of length two:

$$\mathrm{ex}(n, k \cdot P_3) \geq \begin{cases} \binom{3k-1}{2} + \lfloor \frac{n-3k+1}{2} \rfloor & \text{if } 3k \leq n \leq 5k-1 \\ \binom{k-1}{2} + (k-1)(n-k+1) + \lfloor \frac{n-k+1}{2} \rfloor & \text{if } n \geq 5k. \end{cases}$$

Gorgol conjectured that this is the correct value of $\mathrm{ex}(n, k \cdot P_3)$ and proved it for $k = 2, 3$:

$$\mathrm{ex}(n, 2 \cdot P_3) = \left\lfloor \frac{n-1}{2} \right\rfloor + n - 1, \text{ for } n \geq 9$$

$$\mathrm{ex}(n, 3 \cdot P_3) = \left\lfloor \frac{n}{2} \right\rfloor + 2n - 4, \text{ for } n \geq 9.$$

Bushaw and Kettle [4] proved this conjecture for $n \geq 7k$ and showed that

$$\mathrm{ex}(n, k \cdot P_3) = \binom{k-1}{2} + (k-1)(n-k+1) + \left\lfloor \frac{n-k+1}{2} \right\rfloor \text{ for } n \geq 7k.$$

They also showed that the extremal graph is $K_{k-1} + M_{n-k+1}$ where M_{n-k+1} is a maximum matching graph on $n-k+1$ vertices.

Yuan and Zhang [27] completed the characterization for $\mathrm{ex}(n, k \cdot P_3)$ as follows.

Theorem 2.1.3 (Yuan and Zhang, [27])**.**

$$\mathrm{ex}(n, k \cdot P_3) = \begin{cases} \binom{n}{2} & \text{if } n < 3k \\ \binom{3k-1}{2} + \lfloor \frac{n-3k+1}{2} \rfloor & \text{if } 3k \leq n < 5k-1 \\ \binom{3k-1}{2} + k & \text{if } n = 5k-1 \\ \binom{k-1}{2} + (k-1)(n-k+1) + \lfloor \frac{n-k+1}{2} \rfloor & \text{if } n \geq 5k \end{cases}$$

Furthermore,

1. *If $n < 3k$, then the extremal graph is K_n.*
2. *If $3k \leq n < 5k-1$, then the extremal graph is $K_{3k-1} \cup M_{n-3k+1}$.*
3. *If $n = 5k-1$, then the extremal graphs are $K_{3k-1} \cup M_{2k}$ and $K_{k-1} + M_{4k}$.*
4. *If $n \geq 5k$, then the extremal graph is $K_{k-1} + M_{n-k+1}$.*

It should be pointed that the extremal graphs are not unique for $n = 5k-1$ and $k \geq 2$, while the extremal graphs are unique otherwise.

Bushaw and Kettle [4] also determined the extremal number for $k \cdot P_\ell$.

Theorem 2.1.4 (Bushaw and Kettle, [4]). *For $k \geq 2$, $\ell \geq 4$ and $n \geq 2\ell + 2k\ell(\lceil \frac{\ell}{2} \rceil + 1)\binom{\ell}{\lfloor \frac{\ell}{2} \rfloor}$,*

$$\mathrm{ex}(n, k \cdot P_\ell) = \left(k \left\lfloor \frac{\ell}{2} \right\rfloor - 1\right)\left(n - k\left\lfloor \frac{\ell}{2} \right\rfloor + 1\right) + \binom{k\lfloor \frac{\ell}{2} \rfloor - 1}{2} + c_\ell,$$

where $c_\ell = 1$ if ℓ is odd, and $c_\ell = 0$ if ℓ is even.

Lidický, Liu and Palmer [24] extended Bushaw and Kettle's result to paths of different lengths.

Theorem 2.1.5 (Lidický, Liu and Palmer, [24]). *Let F be a forest of paths $P_{v_1}, P_{v_2}, \ldots, P_{v_k}$. If at least one v_i is not 3, then for n sufficiently large,*

$$\mathrm{ex}(n, F) = \left(\sum_{i=1}^{k} \left\lfloor \frac{v_i}{2} \right\rfloor - 1\right)\left(n - \sum_{i=1}^{k} \left\lfloor \frac{v_i}{2} \right\rfloor + 1\right) + \binom{\sum_{i=1}^{k} \lfloor \frac{v_i}{2} \rfloor - 1}{2} + c,$$

where $c = 1$ if each v_i is odd, and $c = 0$ otherwise. Moreover, the extremal graph is unique.

We give a new proof of Theorem 2.1.5 in the case when F has $k \geq 2$ components each of even order at least 4. (See Theorem 2.1.7 below.) Our proof gives a significantly smaller lower bound requirement on n. We begin by defining a class of graphs and establishing a general extremal lemma.

Let H_q^* be the graph consisting of a set of $\lfloor \frac{q-1}{2} \rfloor$ universal vertices and a set B of $n - \lfloor \frac{q-1}{2} \rfloor$ vertices such that B is independent if q is odd and B contains exactly one edge if q is even.

Lemma 2.1.6. *Let H be a graph with a maximum length path P_q such that $H - P_q$ is P_t-free for $t \leq \frac{4}{3} \left\lfloor \frac{q-1}{2} \right\rfloor$. If $n \geq 2\binom{q}{2}$ and $e(H) \geq e(H_q^*)$, then H consists of a set U of $\lfloor \frac{q-1}{2} \rfloor$ universal vertices and set I of $n - \lfloor \frac{q-1}{2} \rfloor$ vertices such that I is independent if q is odd and I contains exactly one edge if q is even.*

Proof. Let H be as in the statement of the theorem. Let $v_1 v_2 \cdots v_q$ be a longest path in H. Denote its vertices by P. Let B be the vertices in H not on P. First observe that any vertex

x in B is adjacent to at most $\lfloor \frac{q-1}{2} \rfloor$ vertices of P as x cannot be adjacent to an endvertex of P or to two consecutive vertices in P. Thus, $e(x, P) \leq \lfloor \frac{q-1}{2} \rfloor$.

Now suppose that x and y are contained in a component in B. Observe that $e(x, P) \leq \lfloor \frac{q-1}{2} \rfloor - 2e(y, P)$ as otherwise x and y have neighbors on P of distance at most 2 and we can form a path longer than P. Likewise, we obtain $e(y, P) \leq \lfloor \frac{q-1}{2} \rfloor - 2e(x, P)$. Therefore, when x, y are in a component in B,

$$e(x, P) + e(y, P) \leq \frac{2}{3} \left\lfloor \frac{q-1}{2} \right\rfloor.$$

Let $S = S_1 \cup S_2 \cup \cdots \cup S_r$ be the non-singleton components in B; i.e., S contains all vertices in B that are adjacent to a vertex in B. Then $B - S$ forms an independent set. So assume $|S| \geq 1$. Now let us estimate the number of edges between S and P as follows. Observe

$$(|S_i| - 1)e(S_i, P) = \sum_{x,y \in S_i} (e(x, P) + e(y, P)) \leq \frac{2}{3} \left\lfloor \frac{q-1}{2} \right\rfloor \binom{|S_i|}{2}.$$

Thus,

$$e(S, P) = \sum_i e(S_i, P) \leq \frac{1}{3} \left\lfloor \frac{q-1}{2} \right\rfloor |S|.$$

Observe that there is no P_t in S, so $e(S) \leq \frac{t-2}{2}|S|$.

Now we show that there is a vertex $x \in B$ such that $e(x, P) = \lfloor \frac{q-1}{2} \rfloor$. Suppose not; then the number of edges in H is

$$
\begin{aligned}
e(H) &= e(P) + e(B) + e(P,B) = e(P) + e(S) + e(S,P) + e(B-S,P) \\
&\le \binom{q}{2} + \frac{t-2}{2}|S| + \frac{1}{3}\left\lfloor \frac{q-1}{2} \right\rfloor |S| + \left(\left\lfloor \frac{q-1}{2} \right\rfloor - 1 \right) (|B| - |S|) \\
&= \binom{q}{2} + \left(\frac{t-2}{2} - \frac{2}{3}\left\lfloor \frac{q-1}{2} \right\rfloor + 1 \right) |S| + \left(\left\lfloor \frac{q-1}{2} \right\rfloor - 1 \right) |B| \\
&= \binom{q}{2} + \left(\frac{t}{2} - \frac{2}{3}\left\lfloor \frac{q-1}{2} \right\rfloor \right) |S| + \left(\left\lfloor \frac{q-1}{2} \right\rfloor - 1 \right) (n-q) \\
&\le \binom{q}{2} + \left(\left\lfloor \frac{q-1}{2} \right\rfloor - 1 \right) (n-q)
\end{aligned}
$$

where the last inequality follows from $\frac{t}{2} \le \frac{2}{3}\left\lfloor \frac{q-1}{2} \right\rfloor$. On the other hand, the number of edges in H_q^* is

$$
e(H_q^*) \ge \binom{\left\lfloor \frac{q-1}{2} \right\rfloor}{2} + \left\lfloor \frac{q-1}{2} \right\rfloor \left(n - \left\lfloor \frac{q-1}{2} \right\rfloor \right).
$$

Given that $n \ge 2\binom{q}{2}$, a simple calculation gives $e(H) < e(H_q^*)$, a contradiction. So there exists a vertex x that is adjacent to $\left\lfloor \frac{q-1}{2} \right\rfloor$ vertices on P.

Let P' be the vertices of P not adjacent to x. If q is odd, then $P' = \{v_1, v_3, v_5, \ldots, v_q\}$ and if q is even, then (without loss of generality) $P' = \{v_1, v_3, v_5, \ldots, v_{q-1}, v_q\}$. Observe that if any two vertices in P' (aside from the pair v_{q-1}, v_q when q is even) are adjacent, then we have a longer path using the vertex x. Furthermore, no vertex in B is adjacent to a vertex of P' as again we can construct a longer path using x. Let I be the vertices of $B \cup P'$ and let U be the vertices in $P - P'$.

Any vertex $y \in S$ is adjacent to a vertex z in S. If y is adjacent to any vertex in P then we can form a longer path using z and x, a contradiction. Therefore, there is no edge leaving S.

Now let us estimate the number of edges in H as follows:

$$\begin{aligned}
e(H) &= e(U) + e(I) + e(S) + e(I,U) + e(S,U) + e(S,I) \\
&\le \binom{\lfloor \frac{q-1}{2} \rfloor}{2} + e(I) + \frac{t-2}{2}|S| + \left\lfloor \frac{q-1}{2} \right\rfloor |I| + 0 + 0 \\
&= \binom{\lfloor \frac{q-1}{2} \rfloor}{2} + e(I) + \frac{t-2}{2}|S| + \left\lfloor \frac{q-1}{2} \right\rfloor \left(n - \left\lfloor \frac{q-1}{2} \right\rfloor - |S|\right) \\
&= \binom{\lfloor \frac{q-1}{2} \rfloor}{2} + e(I) + \left\lfloor \frac{q-1}{2} \right\rfloor \left(n - \left\lfloor \frac{q-1}{2} \right\rfloor\right) - \left(\left\lfloor \frac{q-1}{2} \right\rfloor - \frac{t-2}{2}\right)|S|
\end{aligned}$$

The number of edges in I is 0 if q is odd and 1 if q is even. Now, as $\frac{t-2}{2} < \left\lfloor \frac{q-1}{2} \right\rfloor$, we have that if $|S| > 0$, then $e(H) < e(H_q^*)$, a contradiction. Thus, S is empty. Furthermore, if any vertex in I has fewer than $\lfloor \frac{q-1}{2} \rfloor$ neighbors in U or if U is not a complete graph, then $e(H) < e(H_q^*)$, a contradiction. Therefore, H has the exact same structure as H_q^*. $\square$

Theorem 2.1.7. *Let F be a linear forest with components of order $v_1, v_2, \ldots, v_k$ such that $k \ge 2$ and each v_i is even and at least 4. If $n \ge 2\binom{\sum v_i}{2}$ and G is an n-vertex F-free graph with the maximum number of edges, then G is isomorphic to the graph composed of a set of $\frac{1}{2}\sum v_i - 1$ universal vertices and a set of $n - \frac{1}{2}\sum v_i + 1$ vertices forming an independent set. Thus,*

$$\mathrm{ex}(n, F) = \left(\frac{1}{2}\sum v_i - 1\right)\left(n - \frac{1}{2}\sum v_i + 1\right) + \binom{\frac{1}{2}\sum v_i - 1}{2}.$$

Proof of Theorem 2.1.7. Let G^* be the graph composed of a set of $\frac{1}{2}\sum v_i - 1$ universal vertices and a set of $n - \frac{1}{2}\sum v_i + 1$ vertices forming an independent set. It is easy to see (and it was already observed in [24]) that G^* is F-free.

Now let $n \ge 2\binom{\sum v_i}{2}$ and G be an n-vertex F-free graph with the maximum number of edges. Thus, $e(G) \ge e(G^*)$. Put $q = \sum v_i - 1$ and observe that G is P_{q+1}-free as otherwise G contains a copy of F.

As $n \ge 2\binom{\sum v_i}{2} \ge 2\binom{\frac{1}{2}\sum v_i}{2}$, an easy computation gives $e(G) \ge e(G^*) > (\frac{\sum v_i - 3}{2})n = (\frac{q-3}{2})n$.

Thus G contains a path $P = P_q$. Let t be the order of the smallest component of F. Observe that $G - P$ is P_t-free as otherwise G contains F. Observe that $\frac{t}{2} \leq \frac{2}{3}\left\lfloor \frac{q-1}{2} \right\rfloor$ as long as $t \geq 4$ and $k \geq 2$. Therefore, by Lemma 2.1.6 we have that G is isomorphic to G^*. □

2.2 Forest of path and star components

Lidický, Liu and Palmer also determined the Turán number of a forest of stars.

Theorem 2.2.1 (Lidický, Liu and Palmer, [24]). *Let F be forest of stars $S_{d_1}, S_{d_2}, \ldots, S_{d_k}$ where $d_1 \geq d_2 \geq \cdots \geq d_k$. For n sufficiently large, the Turán number of F is*

$$\mathrm{ex}(n, F) = \max_{1 \leq i \leq k} \left\{ (i-1)(n-i+1) + \binom{i-1}{2} + \left\lfloor \frac{d_i - 1}{2}(n-i+1) \right\rfloor \right\}.$$

Moreover, the extremal graph is unique.

Furthermore, the authors found the Turán number of forests whose components are all of order 4. There are only two trees of order 4: the path P_4 and the star S_3. Put $F = a \cdot P_4 \cup b \cdot S_3$ and $n - b = 3d + r$ with $r \leq 2$. Let $G_F^1(n)$ be the n-vertex graph consisting of b universal vertices together with $K_r \cup d \cdot K_3$. Let $G_F^2(n)$ be the n-vertex graph containing $2a + b - 1$ universal vertices together with an independent set of order $n - 2a - b + 1$. Thus $G_F^2(n) = K_{2a+b-1} + E_{n-2a-b+1}$. It is easy to check that $G_F^1(n)$ and $G_F^2(n)$ are F-free. See Figure 2.1 for illustrations of these graphs.

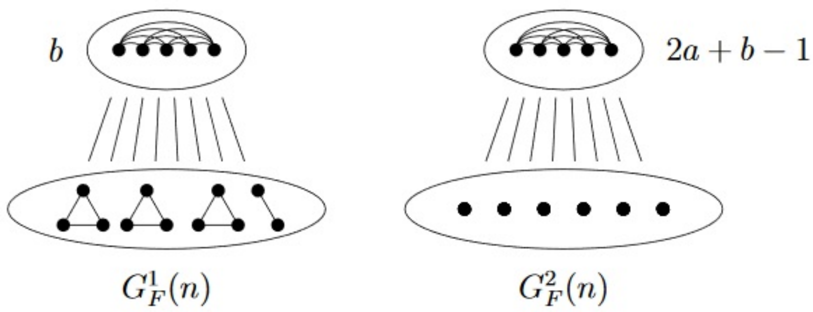

Figure 2.1: The extremal graphs for a forest with components of order 4.[1]

Theorem 2.2.2 (Lidický, Liu and Palmer, [24]). *If $F = a \cdot P_4 \cup b \cdot S_3$, and $n = 3d + r$ is sufficiently large, with $0 \leq r \leq 2$, then*

1. *if $a = 1$ and $r = 0$, then $G^1_F(n)$ is the unique extremal graph;*
2. *if $a = 1$ and $r \neq 0$, then $G^1_F(n)$ and $G^2_F(n)$ are the only extremal graphs;*
3. *if $a > 1$, the $G^2_F(n)$ is the unique extremal graph.*

We give a generalization of Theorem 2.2.2 and determine the Turán number of $F = a \cdot P_\ell \cup b \cdot S_t$. We keep $G^1_F(n)$ as before, slightly update the definition of $G^2_F(n)$, and define an additional graph $G^3_F(n)$.

Let $G^1_F(n)$ be defined as above. Let $G^2_F(n)$ be the n-vertex graph containing $a\lfloor \frac{\ell}{2} \rfloor + b - 1$ universal vertices and the remaining part graph is an independent set if ℓ is even and contains exactly one edge if ℓ is odd. The graph $G^3_F(n)$ is obtained by adding a set U of $b-1$ universal vertices to an extremal graph H (for the star S_t) on $n - b + 1$ vertices. In this way H is a $(t-1)$-regular graph if one of $(t-1)$ or $n - b + 1$ is even and n is large enough. If both of $(t-1)$ and $n-b+1$ are odd, then H has exactly one vertex of degree $(t-2)$ and the remaining vertices have degree $(t-1)$. See Figure 2.2 for illustrations of these three graphs.

[1]The figures are from [24].

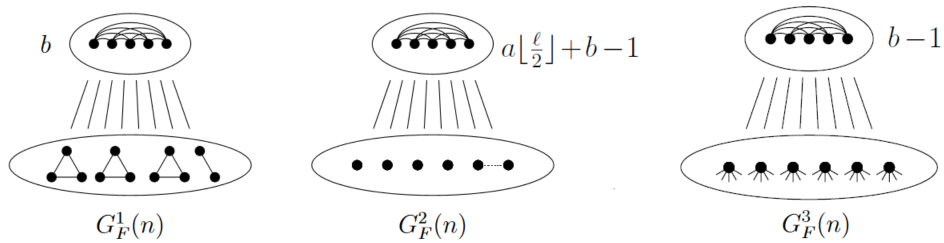

Figure 2.2: The extremal graphs for a forest with components of paths P_ℓ and stars S_t.[2]

The following observation gives the number of edges in the three constructions.

Observation 2.2.3. *Let $F = a \cdot P_\ell \cup b \cdot S_t$. Then*

$$e(G_F^1(n)) = \left(b + \frac{\ell-2}{2}\right) n + b\left(\frac{\ell+b-1}{2}\right) - r\left(\frac{\ell-r-1}{2}\right),$$

$$e(G_F^2(n)) = \left(a\left\lfloor\frac{\ell}{2}\right\rfloor + b - 1\right) n - \binom{a\left\lfloor\frac{\ell}{2}\right\rfloor + b}{2} + c$$

where $c = 0$ if ℓ is even and $c = 1$ if ℓ is odd, and

$$e(G_F^3(n)) = \left(b + \frac{t-1}{2} - 1\right) n - \frac{b-1}{2}(b+t-1) - c'$$

where $c' = 0$ if one of $(t-1)$ and $n-b+1$ is even, and $c' = \frac{1}{2}$ if both of $(t-1)$ and $n-b+1$ are odd.

For the case $a = 0$ and $b \geq 1$, by Theorem 2.2.1, $G_F^3(n)$ is the unique extremal graph. When $b = 0$ and $a > 1$, by Theorem 2.1.4, $G_F^2(n)$ is the unique extremal graph. When $b = 0$ and $a = 1$, by Theorem 2.1.2, the extremal graph is the graph of vertex-disjoint copies of $d \cdot K_{\ell-1} \cup K_r$ or (when ℓ is even and $r = \frac{\ell}{2}, \frac{\ell-2}{2}$) $s \cdot K_{\ell-1} \cup (K_{\lfloor\frac{\ell}{2}\rfloor-1} + E_{\lfloor\frac{\ell}{2}\rfloor+1+(d-s-1)(\ell-1)+r})$ where $0 \leq s < d$. So we may suppose that a and b are both nonzero.

Theorem 2.2.4. *Put $F = a \cdot P_\ell \cup b \cdot S_t$. Let $a \geq 2$, $b \geq 1$ and n be sufficiently large.*

[2]The figures are from [24].

1. *If* $a\lfloor \frac{\ell}{2} \rfloor > \frac{t-1}{2}$*, then*

$$\mathrm{ex}(n,F) = \left(a\left\lfloor \frac{\ell}{2}\right\rfloor + b - 1\right)n - \binom{a\lfloor \frac{\ell}{2}\rfloor + b}{2} + c$$

where $c = 1$ *if* ℓ *is odd and* $c = 0$ *if* ℓ *is even.*

2. *If* $a\lfloor \frac{\ell}{2} \rfloor \leq \frac{t-1}{2}$*, then*

$$\mathrm{ex}(n,F) = \left(b + \frac{t-1}{2} - 1\right)n - \frac{b-1}{2}(b+t-1) - c'$$

where $c' = 0$ *if at least one of* $(t-1)$ *and* $n-b+1$ *is even and* $c' = \frac{1}{2}$ *otherwise.*

3. *If* $a = 1$ *and for* $t \geq \ell + 1$*, then*

$$\mathrm{ex}(n,F) = \left(b + \frac{t-1}{2} - 1\right)n - \frac{b-1}{2}(b+t-1) - c'$$

where $c' = 0$ *if at least one of* $(t-1)$ *and* $n-b+1$ *is even and* $c' = \frac{1}{2}$ *otherwise.*

4. *If* $a = 1$ *and* $t < \ell + 1$ *and* $n - b = (\ell-1)d + r$ *with* $r \leq \ell - 2$*, then*

$$\mathrm{ex}(n,F) = \left(b + \frac{\ell-2}{2}\right)n + b\left(\frac{\ell+b-1}{2}\right) - r\left(\frac{\ell-r-1}{2}\right).$$

Proof of Theorem 2.2.4. The proof of all five cases is by induction on b. The inductive step of these proofs is the same. We give the full argument for (1) and prove only the base cases for (2)–(5). Let F be as in the statement of the theorem and suppose G is an n-vertex extremal graph for F. As $G_F^i(n)$ is F-free we may assume that $e(G) \geq G_F^i(n)$ for $i = 1, 2, 3$.

(1) For $a\lfloor \frac{\ell}{2} \rfloor > \frac{t-1}{2}$ we have that $e(G_F^2(n)) > e(G_F^3(n))$ when n is large enough. We proceed by induction on b. Let $b = 1$. Then $F = a \cdot P_\ell \cup S_t$. If G is S_t-free, then the maximum degree of G is $t-1$ and thus $e(G) \leq e(G_F^3(n))$, a contradiction. Thus G contains a copy of S_t. Now the graph $G - S_t$ is $a \cdot P_\ell$-free. Observe that the extremal construction in Theorem 2.1.4 is exactly $G^2_{a\cdot P_\ell}(n)$. Therefore, $e(G-S_t) \le e(G^2_{a\cdot P_\ell}(n-t-1))$. Let e_0 be the number of edges between S_t and $G-S_t$. Then

$$e_0 = e(G) - e(G-S_t) - e(S_t) \ge e(G^2_F(n)) - e(G^2_{a\cdot P_\ell}(n-t+1)) - e(S_t) = n + O(1) = \Omega(n).$$

Therefore, G contains a vertex x of degree $\Omega(n)$. The graph $G-x$ is $a\cdot P_\ell$-free as x is the center of an S_t. Thus $e(G-x) \le e(G^2_{a\cdot p_\ell}(n-1))$ by Theorem 2.1.4 and $d(x) \le n-1$ which implies that $e(G) \le e(G^2_F(n))$.

Now let $b > 1$ and suppose the statement (1) holds for smaller values. Put $F = a\cdot P_\ell \cup b\cdot S_t$ and let G be an n-vertex extremal graph for F. As before $e(G) \ge e(G^2_F(n))$ and G contains an S_t. Therefore, $G-S_t$ is $a\cdot P_\ell \cup (b-1)\cdot S_t$-free. Put $F' = a\cdot P_\ell \cup (b-1)\cdot S_t$. By induction we have $e(G-S_t) \le e(G^2_{F'}(n-t+1))$. Let e_0 be the number of edges between S_t and $G-S_t$. Then

$$e_0 = e(G) - e(G-S_t) - e(S_t) \ge e(G^2_F(n)) - e(G^2_{F'}(n-t+1)) - e(S_t) = n + O(1) = \Omega(n).$$

Therefore, G contains a vertex x of degree $\Omega(n)$. The graph $G-x$ is F'-free since x is the center of an S_t. Thus $e(G-x) \le e(G^2_{F'}(n-1))$ by induction and $d(x) \le n-1$ which implies that $e(G) \le e(G^2_F(n))$.

(2) For $a\lfloor \frac{\ell}{2} \rfloor \le \frac{t-1}{2}$ we have that $e(G^3_F(n)) > e(G^2_F(n))$ when n is large enough. Let $b=1$. Then $F = a\cdot P_\ell \cup S_t$. Suppose G is an extremal graph for F.

If G contains a copy of S_t, then the argument in (1) gives that $e(G) \le e(G^2_F(n)) < e(G^3_F(n))$, a contradiction. Therefore, G is S_t-free; i.e., G has maximum degree at most $t-1$ which completes the proof of the base case wherein the asserted value of $\mathrm{ex}(n,F)$ evaluates to $\frac{(t-1)n}{2}$.

(3) If G is S_t-free, then $e(G) \le e(G^3_F(n))$ and we are done. So assume G contains a copy of S_t. Now the graph $G-S_t$ is P_ℓ-free. Let e_0 be the number of edges between S_t and $G-S_t$.

Note that $G - S_t$ is P_ℓ-free. Suppose A is the vertex set of S_t. Then since $t > \ell + 1$,

$$e_0 = e(G) - e(G - S_t) - e(A) \geq e(G_F^3(n)) - e(G - S_t) - e(A)$$

$$= \frac{t-1}{2}n - (n - t - 1)\frac{\ell - 2}{2} - \binom{t+1}{2} = 2n + o(1) \sim \Omega(n).$$

So G contains a vertex x of of degree $\Omega(n)$. Then $G - x$ is P_ℓ-free as x is the center of an S_t. Therefore, $e(G) \leq \frac{\ell-2}{2}(n-1) + n - 1 < e(G_F^3(n))$, a contradiction.

(4) As $t < \ell + 1$ we have that $e(G_F^2(n)) \geq e(G_F^3(n))$ when n is large enough. Let $F = P_\ell \cup S_t$ and suppose G is an extremal graph for F. Since $G_F^1(n)$ and $G_F^2(n)$ are F-free we have $e(G) \geq e(G_F^1(n)) \geq e(G_F^2(n))$. If G is S_t-free, then $e(G) \leq \frac{t-1}{2}n < e(G_F^1(n))$, a contradiction. Thus G contains a copy of S_t. Now the graph $G - S_t$ is P_ℓ-free. Let e_0 be the number of edges between S_t and $G - S_t$, and A is the vertex set of S_t. Then

$$e_0 = e(G) - e(G - S_t) - e(A) \geq e(G_F^1(n)) - (n - t - 1)\frac{\ell - 2}{2} - \binom{t+1}{2} = 2n + o(1) = \Omega(n).$$

So G contains a vertex x of degree $\Omega(n)$. Since $G - x$ is P_ℓ-free, $e(G) \leq \frac{(n-1)(\ell-2)}{2} + n - 1$. Then $e(G) \leq e(G_F^1(n))$. $\square$

We conjecture that when $t < \ell - 1$ and $F = P_\ell \cup b \cdot S_t$, then $\mathrm{ex}(n, F) = e(G_F^2(n))$.

2.3 Forest with components of order 5

Following in the spirit of Theorem 2.2.2, we consider forests whose components are all of order 5. Suppose T_5 is the path P_4 in which an edge is connected to one of the vertices of degree 2. Observe that P_5, S_4 and T_5 are the only 5-vertex trees. Therefore, we would like to determine the Turán number of the graph $F = a \cdot P_5 \cup b \cdot T_5 \cup c \cdot S_4$. We begin with a lemma on the Turán number of T_5.

Lemma 2.3.1. *The unique extremal graph for P_5 is the unique extremal graph for T_5. This implies that* $\mathrm{ex}(n, T_5) = \mathrm{ex}(n, P_5) \leq \frac{3}{2}n$.

Proof of Lemma 2.3.1. By Theorem 2.1.2, the unique extremal graph for P_5 is the disjoint union of cliques, i.e., $G^* = \bigcup_{i=1}^{q} K_4 \cup K_r$ where $n = 4q + r$ and $r < 4$. Clearly this graph is also T_5-free. We now show that G^* is also the unique extremal graph for T_5. We proceed by induction on n. The result is immediate when $n \leq 4$. So let $n > 5$ and suppose the results holds for smaller values. Let G be an n-vertex extremal graph for T_5. Let x be a vertex of maximum degree in G. Denote the neighbors of x by $N(x)$ and let Y be the vertices of $G - N(x) - x$. If $d(x) < 3$, then $e(G) < e(G^*)$, a contradiction. So we may assume that $d(x) \geq 3$. If $d(x) \geq 4$, then the neighbors of x form an independent set and there is no edge between $N(x)$ and Y as otherwise we can find a copy of T_5. In this case, $x \cup N(x)$ form a component of G containing $d(x)$ edges. Applying induction to Y gives that $e(G) < e(G^*)$, a contradiction. Therefore, $d(x) = 3$ and no vertex in $N(x)$ is adjacent to a vertex in Y as otherwise we have a copy of T_5. Now by induction, Y is the graph $\bigcup_{i=1}^{q-1} K_4 \cup K_r$. The component on vertex set $x \cup N(x)$ has 4 vertices and is maximized when it is a K_4. Thus, $G = G^*$. $\square$

A consequence of Lemma 2.3.1 is that for any 5-vertex tree T we have $\mathrm{ex}(n, T) \leq \frac{3}{2}n$. We continue with a simple proposition for the forest $P_4 \cup T_5$. The argument serves as a simple sketch of the proofs later in this section.

Proposition 2.3.2. *Let $F = P_4 \cup T_5$ and let G be an n-vertex extremal graph for F. If n is large enough, then G is isomorphic to the extremal graph for $P_4 \cup P_4$; i.e., $G \simeq K_3 + E_{n-3}$.*

Proof. We will show that G contains no $P_4 \cup P_4$ which implies the result. Suppose G contains a copy of $P_4 \cup P_4$. Let A be the vertices of a copy of $P_4 \cup P_4$ and let B be the remaining vertices. It is easy to see that the end-vertices of the two P_4 paths send at most one edge to B and that the vertices of degree 2 of the two P_4 paths send no edges to B. Indeed, in either case we can find a copy of a $P_4 \cup T_5$ in G. Clearly, B must be T_5-free. Therefore, applying Lemma 2.3.1 to B gives that the number of edges in G is

$$e(G) = e(A) + e(B) + e(A, B) \leq \binom{8}{2} + \frac{3}{2}(n-8) + 4.$$

Now observe that as $P_4 \cup P_4$ is a subgraph of $P_4 \cup T_5$ we have

$$e(G) = \mathrm{ex}(n, P_4 \cup T_5) \geq \mathrm{ex}(n, P_4 \cup P_4) = 3n - 6,$$

where the last equality is by Theorem 2.1.5. Therefore, for n large enough, these two estimates for $e(G)$ give a contradiction. $\square$

Lemma 2.3.3. *Fix integers a, b, c such that $2a + 2b + c > \frac{7}{2}$. Let $F = a \cdot P_5 \cup b \cdot T_5 \cup c \cdot S_4$ and let F' be a graph resulting from the removal of a component from F. If G is an n-vertex extremal graph for F, then every copy of F' in G contains two vertices x, y with a common neighborhood of $\Omega(n)$ vertices.*

Proof. Let A be the vertices of a copy of F' in G and let B be the remaining vertices. Observe that $|A| = 5(a + b + c - 1)$. Let T be the component of F removed to create F'. Then B contains no copy of T. Thus $e(B) \leq \frac{3}{2}n$. The graph $K_{2a+2b+c-1} + E_{n-(2a+2b+c-1)}$ contains no F, so $e(G) \geq (2a + 2b + c - 1)n - O(1)$. Let us estimate the number of edges between A and B in two ways. First,

$$\begin{aligned} e(A, B) = e(G) - e(A) - e(B) &\geq (2a + 2b + c - 1)n - O(1) - \binom{5(a+b+c-1)}{2} - \frac{3}{2}n \\ &\geq \left(2a + 2b + c - 1 - \frac{3}{2}\right) n + O(1). \end{aligned}$$

Now let n_0 be the number of vertices in B that are adjacent to at least 2 vertices in A. Then

$$e(A, B) \leq 5(a + b + c - 1)n_0 + (n - n_0).$$

Combining these two estimates for $e(A,B)$ and solving for n_0 gives

$$n_0 \geq \left(\frac{2a+2b+c-2-\frac{3}{2}}{5a+5b+5c-6} \right) n + O(1)$$

which implies $n_0 = \Omega(n)$ as long as $2a+2b+c > \frac{7}{2}$. Now there are $\binom{5(a+b+c-1)}{2}$ pairs of vertices in A, so some pair of vertices x, y in A has at least $\Omega(n)$ common neighbors in B. □

Lemma 2.3.4. *Let T be a T_5 or P_5. Then, for large enough n, the extremal graph for $T \cup T_5$ is $K_3 + E_{n-3}$.*

Proof. Let G be an extremal graph for $T \cup T_5$. By Lemma 2.3.3, each copy of T_5 in G contains two vertices with a common neighborhood of order $\Omega(n)$.

Let us construct an auxiliary graph H on the same vertex set as G. Two vertices are adjacent in H if and only if they are in a copy of T_5 in G and have common neighborhood of size at least 7. Observe that any pair of vertices with three common neighbors span a copy of T_5 and P_5. This implies that if H contains two vertex-disjoint edges, then G contains $T \cup T_5$, a contradiction. Therefore, all edges in H are incident. This implies that H has a vertex x in every edge or the edges of H form a triangle x, y, z. We distinguish two cases.

Case 1: The graph H has a vertex x incident to every edge.

This case implies that $H - x$ contains no edge and therefore $G - x$ is T_5-free. Thus, we can estimate the number of edges in G as follows:

$$e(G) = e(x, G-x) + e(G-x) \leq (n-1) + \frac{3}{2}(n-1) < 3n - 6 \leq e(G)$$

which is a contradiction for $n > 7$.

Case 2: The edges in H form a triangle on vertices x, y, z.

Put $A = \{x, y, z\}$ and $B = V(G) - \{x, y, z\}$. Each vertex in A has at least 6 neighbors in B as every pair has at least 6 common neighbors in B. Let N be the vertices in B adjacent to a vertex of A. Suppose there is an edge uv in B with at least one vertex in N. Without loss of generality we may suppose that u is adjacent to x. We can form a copy of T_5 with x, u, v and another neighbor of x in N. We can form a copy of T that is vertex-disjoint from the T_5 by using y, z and their common neighbors, a contradiction. Therefore, there is no edge in B incident to N. Moreover, $B - N$ is T_5-free. Put $c = |B - N|$.

Therefore, applying Lemma 2.3.1 to $B - N$ gives that the number of edges in G is

$$\begin{aligned} e(G) = e(A) + e(B) + e(A, B) &= e(A) + e(B - N) + e(A, N) \\ &\leq 3 + \frac{3}{2}c + 3(n - 3 - c) = 3n - 6 - \frac{3}{2}c. \end{aligned}$$

This implies that $c = 0$ (as otherwise we have $e(G) < 3n - 6$), that G must contain all edges between A and B, and that A forms a triangle in G. Thus $G \simeq K_3 + E_{n-3}$ for n large enough. □

Theorem 2.3.5. *Fix integers a, b, c such that $a + b \geq 2$ and $b \geq 1$. Let $F = a \cdot P_5 \cup b \cdot T_5 \cup c \cdot S_4$. Then, for n large enough, the extremal graph for F is $K_{2a+2b+c-1} + E_{n-(2a+2b+c-1)}$.*

Proof of Theorem 2.3.5. We prove the result by induction on $a + b + c$. The base case is when $a + b = 2$ and $c = 0$ which is handled by Lemma 2.3.4. So assume $a + b + c \geq 3$ and that the theorem holds for forests with fewer components.

Let G be an extremal graph for F and put $G^* = K_{2a+2b+c-1} + E_{n-(2a+2b+c-1)}$ Thus,

$$e(G) \geq e(G^*) \geq (2a + 2b + c - 1)n - O(1).$$

As $a+b+c \geq 3$, there is a component T of F whose removal leaves a forest $F' = a' \cdot P_5 \cup b' \cdot T_5 \cup c' \cdot S_4$ such that $a'+b' \geq 2$ and $b' \geq 1$. First we claim that G contains F'. Indeed, if it does not, then by induction

$$e(G) \leq \mathrm{ex}(n, F') \leq (2a+2b+c-2)n - O(1),$$

a contradiction.

Let A be the vertices of F' and let B be the remaining vertices of G. By Lemma 2.3.3 there are two vertices x, y in A with $\Omega(n)$ common neighbors in B.

We distinguish two (similar) cases.

Case 1: T is an S_4.

Let G' be the graph resulting from the removal of x from G. Note that x has $\Omega(n)$ neighbors in B. Observe that

$$\begin{aligned} e(G') &\geq e(G) - (n-1) \geq e(K_{2a+2b+c-1} + E_{n-(2a+2b+c-1)}) - (n-1) \\ &= e(K_{2a+2b+c-2} + E_{n-2-(2a+2b+c-1)}) = \mathrm{ex}(n-1, F') \end{aligned}$$

where the last equality is by induction and the second to last equality comes from comparing the graphs $K_{2a+2b+c-1} + E_{n-(2a+2b+c-1)}$ and $K_{2a+2b+c-2} + E_{n-2-(2a+2b+c-1)}$.

Thus $e(G') \geq \mathrm{ex}(n-1, F')$. Now if G' contains a copy of F', then we can find a copy of S_4 that is vertex-disjoint from the F' by using x and its large neighborhood in B. Thus, G' is F'-free. Therefore, by induction G' is isomorphic to $K_{2a+2b+c-2} + E_{n-2-(2a+2b+c-1)}$ for n large enough. In order for G to have as many edges as $K_{2a+2b+c-1} + E_{n-(2a+2b+c-1)}$ then x must be a universal vertex in G. Therefore, G is isomorphic to $K_{2a+2b+c-1} + E_{n-(2a+2b+c-1)}$.

Case 2: T is a P_5 or T_5.

Let G' be the graph resulting from the removal of x and y from G. Observe that

$$\begin{aligned} e(G') &\geq e(G) - (n-1) - (n-2) \geq e(K_{2a+2b+c-1} + E_{n-(2a+2b+c-1)}) - (n-1) - (n-2) \\ &= e(K_{2a+2b+c-3} + E_{n-3-(2a+2b+c-1)}) = \mathrm{ex}(n-2, F') \end{aligned}$$

where the last equality is by induction and the second to last equality comes from comparing the graphs $K_{2a+2b+c-1} + E_{n-(2a+2b+c-1)}$ and $K_{2a+2b+c-3} + E_{n-3-(2a+2b+c-1)}$.

Thus $e(G') \geq \mathrm{ex}(n-2, F')$. Let T be the component removed from F to get F'. If G' contains a copy of F', then we can find a copy of T that is vertex-disjoint from the F' by using x, y and their large common neighborhood. Thus, G' is F'-free. Therefore, by induction G' is isomorphic to $K_{2a+2b+c-3} + E_{n-3-(2a+2b+c-1)}$ for n large enough. In order for G to have as many edges as $K_{2a+2b+c-1} + E_{n-(2a+2b+c-1)}$ both x and y must be universal vertices in G. Therefore, G is isomorphic to $K_{2a+2b+c-1} + E_{n-(2a+2b+c-1)}$. $\square$

It remains to handle the cases for $F = a \cdot P_5 \cup b \cdot T_5 \cup c \cdot S_4$ when $b = 0$ and when $a = 0$, $b = 1$. The case $a \geq 2$, $b = 0$ is handled by Case 1 of Theorem 2.2.4. Therefore, the remaining case is

Proposition 2.3.6. *Let $F = T \cup c \cdot S_4$ where T is P_5 or T_5 and $c \geq 1$. Then, for n large enough,*

$$\mathrm{ex}(n, F) \leq \left(c + \frac{3}{2}\right)(n - c) + \binom{c}{2}$$

Proof. Let G^* be constructed as follows. Put $n - c = 4q + r$ where $0 \leq r < 4$ and join c universal vertices to a collection of q vertex-disjoint copies of K_4 and (if $r \geq 1$) one K_r. Observe that G^* is F-free as each of the $c + 1$ components of F in G^* must use at least one of the c universal vertices.

Now let G be an n-vertex extremal graph for F. Thus,

$$e(G) \geq e(G^*) \geq \left(c + \frac{3}{2}\right) n + \binom{c}{2} - c(c + \frac{3}{2}) - \frac{r(4-r)}{2}.$$

We proceed by induction on c. For $c = 1$ we have $F = T \cup S_4$ and $e(G) \geq \frac{5}{2}n - \frac{5}{2} - \frac{r(4-r)}{2}$. Therefore, G contains a copy of T. Let A be the vertices of T and let B be the remaining vertices. Clearly, B is S_4-free; i.e., $e(B) \leq \frac{3}{2}(n-5)$. Therefore, the number of edges between A and B is

$$e(A, B) = e(G) - e(B) - e(A) \geq \frac{5}{2}n - \frac{5}{2} - \frac{r(4-r)}{2} - \frac{3}{2}n - \frac{15}{2} - \binom{5}{2} = n - O(1).$$

Therefore, A contains a vertex x of degree $\Omega(n)$. Now observe that the graph $G - x$ is T-free as otherwise we can find a copy of F using x as the center of S_4 vertex-disjoint from T. Therefore, Lemma 2.3.1 gives $e(G - x) \leq \frac{3}{2}(n-1)$. Therefore, in order for G to have as many edges as G^* we must have that x is a universal vertex. Then

$$e(G) \leq \frac{3}{2}(n-1) + (n-1) = \frac{5}{2}(n-1).$$

This completes the proof of the base case. The inductive step is similar, and we next consider it.

Let $c \geq 2$ and suppose the statement holds for forests on fewer components. First we claim that G contains a copy of $T \cup (c-1) \cdot S_4$. If it does not, then by induction

$$e(G) \leq (c + \frac{1}{2})(n - c + 1) + \binom{c-1}{2} < (c + \frac{3}{2})(n - c) + \binom{c}{2} - \frac{3}{2}r = e(G^*)$$

for n large enough, a contradiction. Let A be the vertices of $T \cup (c-1) \cdot S_4$ in G and let B be the set of remaining vertices. Clearly, B is S_4-free; i.e., $e(B) \leq \frac{3}{2}(n - 5c)$. Therefore, the number of edges between A and B is

$$e(A,B) = e(G) - e(B) - e(A) \geq \left(c + \frac{3}{2}\right) n + \binom{c}{2} - c(c + \frac{3}{2}) - \frac{r(4-r)}{2} - \frac{3}{2}n - \frac{15}{2}c - \binom{5c}{2}$$
$$= cn + O(1).$$

Therefore, A contains a vertex x of degree $\Omega(n)$. Now observe that the graph $G - x$ is $T \cup (c-1) \cdot S_4$-free as otherwise we can find a copy of F using x as the center of S_4 vertex-disjoint from $T \cup (c-1) \cdot S_4$. Therefore, by induction

$$e(G - x) \leq \mathrm{ex}(n-1, T \cup (c-1)S_4) \leq \left(c - 1 + \frac{3}{2}\right)(n-c) + \binom{c-1}{2}$$

for n large enough. Therefore, the number of edges in G is

$$e(G) \leq (n-1) + \left(c - 1 + \frac{3}{2}\right)(n-c) + \binom{c-1}{2} = \left(c + \frac{3}{2}\right)(n-c) + \binom{c}{2}$$

which completes the proof. $\square$

Chapter 3

Hypergraphs

3.1 Forest of expansions of stars

One of most beautiful results in extremal set theory is the so-called Sunflower Lemma discovered by Erdős and Rado in 1960. A *sunflower* with ℓ *petals* and a *core* Y is a collection of sets $A_1, \ldots, A_\ell$ such that $A_i \cap A_j = Y$ for all $i \neq j$; the sets $A_j - Y$ are *petals*, and we require that none of them is empty. The Sunflower Lemma states:

Theorem 3.1.1 (Erdős and Rado, [11])**.** *Let $\mathcal{F}$ be a family of sets each of cardinality r. If $|\mathcal{F}| > r!(\ell-1)^r$ then $\mathcal{F}$ contains a sunflower with ℓ petals.*

Note that the r-uniform expansion S_ℓ^+ of a star S_ℓ is a sunflower with ℓ petals and a core of size 1. Therefore, a result of Duke and Erdős [7] on sunflowers gives a bound on the Turán number of S_ℓ^+.

Theorem 3.1.2 (Duke and Erdős, [7])**.** *Let $r \geq 3$ and $\ell \geq 2$ be positive integers. There exist* *constants $c(r)$ and $n(\ell, r)$ such that for $n > n(\ell, r)$ we have*

$$\mathrm{ex}_r(n, S_\ell^+) \le c(r)\ell(\ell-1)n^{r-2}.$$

Let M_k be a set of k pairwise-disjoint r-edges, i.e., a matching of size k. A classic theorem of Erdős [9] gives the Turán number of M_k.

Theorem 3.1.3 (Erdős, [9])**.** *For $r \ge 2$ and n large enough we have*

$$\mathrm{ex}_r(n, M_k) = \binom{n}{r} - \binom{n-k+1}{r}.$$

As we may view M_k as a forest of stars each of size 1, i.e., $k \cdot S_1^+$, Theorem 3.1.3 serves as an initial case for the Turán number of a forest of pairwise disjoint stars S_ℓ^+. We give the following bounds in an attempt to generalize Theorem 3.1.2 and Theorem 3.1.3. Let $k \cdot S_\ell^+$ denote k pairwise vertex-disjoint copies of the r-uniform expansion of the star S_ℓ.

Theorem 3.1.4. *Fix integers $\ell, k \ge 1$ and $r \ge 2$. Then for n large enough,*

$$\mathrm{ex}_r(n, k \cdot S_\ell^+) = \binom{n}{r} - \binom{n-k+1}{r} + \mathrm{ex}_r(n-k+1, S_\ell^+).$$

Proof. For the lower bound, consider an r-uniform n-vertex hypergraph constructed as follows. Let A and B be sets of $k-1$ and $n-k+1$ vertices, respectively. First we embed an S_ℓ^+-free hypergraph with $\mathrm{ex}_r(n-k+1, S_\ell^+)$ hyperedges into B. Next we add every r-set that is incident to A to our hypergraph. It is easy to see that this hypergraph has exactly as many hyperedges as in the statement of the theorem. Moreover, as B contains no copy of S_ℓ^+ each such subgraph must contain at least one vertex of A. Therefore, there is no $k \cdot S_\ell^+$ subgraph.

We now continue with the upper bound. Let $\mathcal{H}$ be an r-uniform n-vertex hypergraph with

$$|E(\mathcal{H})| > \binom{n}{r} - \binom{n-k+1}{r} + \mathrm{ex}_r(n+k+1, S_\ell^+).$$

We will show that $\mathcal{H}$ contains a copy of $k \cdot S_\ell^+$. We proceed by induction on k. For $k = 1$ the base case is immediate as $|E(\mathcal{H})| > \mathrm{ex}_r(n, S_\ell^+)$. So let $k > 1$ and assume the statement holds for $k - 1$. We distinguish two cases based on the maximum degree $\Delta(\mathcal{H})$ of $\mathcal{H}$.

Case 1: The maximum degree satisfies

$$\Delta(\mathcal{H}) < \frac{1}{(k-1)((r-1)\ell+1)}\left(\binom{n}{r} - \binom{n-k+1}{r}\right).$$

Consider a copy of $t \cdot S_\ell^+$ in $\mathcal{H}$ such that t is maximal. We claim that $t \geq k$. Indeed, if $t < k$, then at most $(k-1)((r-1)\ell+1)$ vertices are spanned by the $t \cdot S_\ell^+$. Removing these vertices (and the incident hyperedges) leaves at least

$$\begin{aligned}|E(\mathcal{H})| - (k-1)((r-1)\ell+1)\cdot\Delta(\mathcal{H}) &> |E(\mathcal{H})| - \binom{n}{r} - \binom{n-k+1}{r}\\ &> \mathrm{ex}_r(n-k+1, S_\ell^+)\end{aligned}$$

hyperedges. Therefore, there is a copy of S_ℓ^+ that is vertex-disjoint from the $t \cdot S_\ell^+$. This violates the maximality of t, a contradiction.

Case 2: The maximum degree satisfies

$$\Delta(\mathcal{H}) \geq \frac{1}{(k-1)((r-1)\ell+1)}\left(\binom{n}{r} - \binom{n-k+1}{r}\right).$$

Let x be a vertex of maximum degree. Observe that $d(x) \leq \binom{n-1}{r-1}$, so

$$\begin{aligned}|E(\mathcal{H})| - d(x) &> \binom{n}{r} - \binom{n-1}{r-1} - \binom{n-k+1}{r} + \mathrm{ex}_r(n-k+1, S_\ell^+)\\ &= \binom{n-1}{r} - \binom{n-k+1}{r} - \mathrm{ex}_r(n-k+1, S_\ell^+).\end{aligned}$$

Therefore, if we remove x from $\mathcal{H}$ and apply induction to the resulting hypergraph we have a copy of $(k-1)\cdot S_\ell^+$. Now it remains to show that there is a copy of S_ℓ^+ with center x that is vertex-disjoint from the $(k-1)\cdot S_\ell^+$.

First observe that x and any vertex $y \in V((k-1)\cdot S_\ell^+)$ are contained in at most $\binom{n-2}{r-2}$ common hyperedges. Therefore, the number of hyperedges containing x and a vertex of the $(k-1)\cdot S_\ell^+$ is at most

$$(k-1)((r-1)\ell+1)\binom{n-2}{r-2} = O(n^{r-2}).$$

On the other hand, $d(x) = \Omega(n^{r-1})$. Therefore, if we we remove the hyperedges of the $(k-1)\cdot S_\ell^+$, we are still left with $\Omega(n^{r-1})$ hyperedges incident to x. Applying Theorem 3.1.2 to these hyperedges gives a copy of S_ℓ^+ that is vertex-disjoint from the $(k-1)\cdot S_\ell^+$; i.e, $\mathcal{H}$ contains a copy of $k\cdot S_\ell^+$. □

3.2 Forest of expansions of stars in linear hypergraphs

Recall that a hypergraph is *linear* if any pair of hyperedges intersect in at most 1 vertex. In this subsection we examine Turán numbers of linear hypergraphs. Denote by

$$\mathrm{ex}_r^{\mathrm{lin}}(n, F)$$

the maximum number of hyperedges in an n-vertex r-uniform linear hypergraph.

We begin with a general construction that builds linear hypergraphs from smaller linear hypergraphs.

Let $[r]^d$ denote the integer lattice formed by d-tuples from $\{1, 2, \ldots, r\}$. We can think of $[r]^d$ as a hypergraph in the following way: the collection of d-tuples that are fixed in all but one coordinate form a hyperedge. Thus $[r]^d$ is r-uniform and has r^d vertices and $d \cdot r^{d-1}$ hyperedges. Observe that $[r]^d$ is linear as two hyperedges are either disjoint or intersect in exactly one vertex. Furthermore, every vertex is included in exactly d hyperedges, so $[r]^d$ is d-regular. Finally, note that the hyperedges of $[r]^d$ can be partitioned into d classes each of which forms a matching. This gives a natural proper edge-coloring of $\mathcal{H}$. We call such a proper edge-coloring a *canonical coloring*. See Figure 3.1 for an illustration of the hypergraph $[4]^3$ with a canonical coloring.

Consider the Cartesian product of two integer lattices, i.e., $[r]^d \times [s]^t$. There are two hypergraphs naturally associated with $[r]^d \times [s]^t$. The first $\mathcal{Q}_1$ is an s-uniform hypergraph formed by r^d disjoint copies of the hypergraph $[s]^t$. The second $\mathcal{Q}_2$ is an r-uniform hypergraph formed by s^t disjoint copies of the hypergraph $[r]^d$. These two hypergraphs are defined on the same vertex set, but note that a hyperedge $q_1 \in \mathcal{Q}_1$ and a hyperedge $q_2 \in \mathcal{Q}_2$ share at most one vertex.

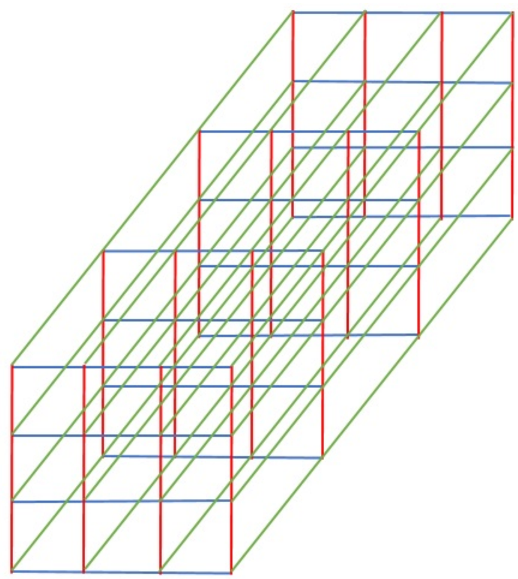

Figure 3.1: The hypergraph $[4]^3$. Each hyperedge is represented by a colored line.

We need a simple generalization of a lemma proved in [21].

Lemma 3.2.1 (Average Degree Lemma)**.** *Fix positive integers d and Δ and a constant $0 \leq \epsilon < 1$. If $\mathcal{G}$ is a hypergraph with average degree at least $d - \epsilon$ and maximum degree at most Δ, then the number of vertices in $\mathcal{G}$ of degree less than d is at most*

$$\frac{\Delta - d + \epsilon}{\Delta - d + 1} n.$$

In particular, the number of vertices in $\mathcal{G}$ of degree at least d is $\Omega(n)$ (i.e. at least Cn where $C = C(d, \Delta, \epsilon)$).

Proof. The sum of the degrees in $\mathcal{G}$ is at least $(d - \epsilon)n$. On the other hand, if x is the number of vertices of degree less than d in $\mathcal{G}$, then the sum of the degrees in $\mathcal{G}$ is at most $(d-1)x + \Delta(n-x)$. Combining these two estimates and solving for x gives the result. $\square$

Theorem 3.2.2. *For large enough n, we have*

$$\mathrm{ex}_r^{\mathrm{lin}}(n, k \cdot S_\ell^+) \leq \left(\frac{\ell-1}{r} + \frac{k-1}{r-1}\right)(n-k+1) + \frac{\binom{k-1}{2}}{\binom{r}{2}}.$$

Furthermore, this bound is sharp asymptotically.

Proof. Let $\mathcal{H}$ be a linear r-uniform n-vertex hypergraph with no $k \cdot S_\ell^+$ subhypergraph. Observe that $k \cdot S_\ell^+$ is simply k pairwise vertex-disjoint copies of S_ℓ^+. Let A be the vertices in $\mathcal{H}$ of degree at least some fixed (large enough) constant $d(r,k,\ell)$. If $|A| \geq k$, then we can greedily embed k pairwise vertex-disjoint copies of S_ℓ^+ into $\mathcal{H}$. Thus, $|A| \leq k-1$.

Let $\mathcal{H}'$ be the r-uniform hypergraph in $\mathcal{H} - A$. The maximum degree in $\mathcal{H}'$ is less than $d(r,k,\ell)$. If the average degree in $\mathcal{H}'$ is $\ell - \epsilon$ for any $\epsilon < 1$, then by Lemma 3.2.1 we have $\Omega(n)$ vertices of degree at least ℓ in $\mathcal{H}'$. In this case we can greedily embed k pairwise vertex-disjoint copies of S_ℓ^+ into $\mathcal{H}'$. Therefore, the average degree in $\mathcal{H}'$ is at most $\ell - 1$. Thus,

$$e(\mathcal{H}') \leq \frac{\ell-1}{r}(n - |A|).$$

Let B be the vertices $V(\mathcal{H}') = V(\mathcal{H}) - A$. Now let us count the r-edges of $\mathcal{H}$ that contain at least one vertex of A and one vertex of B. Denote this collection of edges by $E(A,B)$. To this end let us count the number of pairs $(h, \{x,y\})$ where h is a hyperedge of $\mathcal{H}$ and x is a vertex in $A \cap h$ and y is a vertex in $B \cap h$. Fixing an r-edge $h \in E(A,B)$ we have $|A \cap h|$ choices for x and $|B \cap h|$ choices for y. Thus the number of pairs $(h, \{x,y\})$ is

$$\sum_{h \in E(A,B)} |A \cap h||B \cap h| \geq \sum_{h \in E(A,B)} (r-1) = |E(A,B)|(r-1).$$

On the other hand, for a fixed x and y there is at most one hyperedge containing them as $\mathcal{H}$ is linear. Thus, the number of pairs $(h, \{x,y\})$ is at most $|A|(n-|A|)$. Combining these two estimates and solving for $|E(A,B)|$ gives

$$|E(A,B)| \leq \frac{|A|}{r-1}(n-|A|).$$

Finally, the maximum number of hyperedges contained completely in A is at most $\binom{k-1}{2}/\binom{r}{2}$. Therefore, the number of hyperedges in $\mathcal{H}$ is

$$e(\mathcal{H}) \leq \frac{\ell-1}{r}(n-|A|) + \frac{|A|}{r-1}(n-|A|) + \frac{\binom{k-1}{2}}{\binom{r}{2}}.$$

As $|A| \leq k-1$, we have that for n large enough,

$$e(\mathcal{H}) \leq \left(\frac{\ell-1}{r} + \frac{k-1}{r-1}\right)(n-k+1) + \frac{\binom{k-1}{2}}{\binom{r}{2}}.$$

Now suppose that $n-k+1$ is divisible by $(r-1)^{k-1}$ and $r^{\ell-1}$. Let us construct a hypergraph $\mathcal{H}^*$ as follows. The vertex set of $\mathcal{H}^*$ is partitioned into a set $A^* = \{a_1, a_2, \ldots, a_{k-1}\}$ of $k-1$ vertices and a set B^* of $n-k+1$ vertices partitioned into distinct copies of $[r-1]^{k-1} \times [r]^{\ell-1}$. The hyperedges of $\mathcal{H}^*$ consist of two types. The first type is every r-edge in the copies of $[r]^{\ell-1}$ in B^*. The second type consists of every $h \cup \{a_i\}$ where $a_i \in A^*$ and h is a hyperedge of color i in a copy of $[r-1]^{k-1}$ with a canonical edge-coloring. See Figure 3.2 for an illustration of this construction. Finally, depending on the size of $k-1$ compared to r we may embed a negligible number of hyperedges into A without violating the linear property of $\mathcal{H}^*$. However, we ignore these potential edges in our construction.

Let us confirm that $\mathcal{H}^*$ is linear. Two hyperedges of the first type are hyperedges from copies of the linear hypergraph $[r]^{\ell-1}$ and therefore intersect in at most one vertex. Two hyperedges of the second type are either the same color and therefore intersect in a vertex in A^* but not in B^* or are of different colors and may intersect in B^* in one vertex, but do not intersect in A^*. Finally, a hyperedge of the first type and a hyperedge of the second type intersect in at most one vertex by the construction of $[r-1]^{k-1} \times [r]^{\ell-1}$.

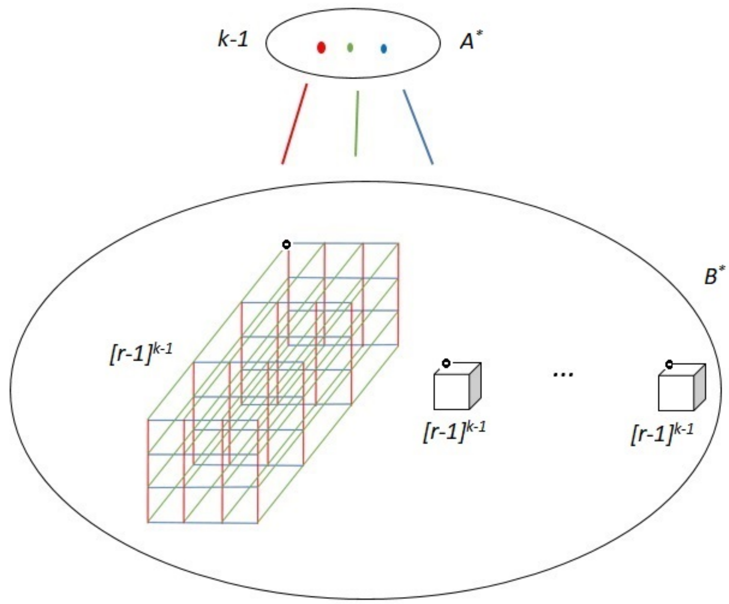

Figure 3.2: Construction of $\mathcal{H}^*$ for $k-1=3$ and $r=5$. A $[r]^{\ell-1}$ is constructed by taking $r^{\ell-1}$ vertices from $r^{\ell-1}$ copies of $[r-1]^{k-1}$ such that they are in the same $(r-1)$-tuple in $[r-1]^{k-1}$.

The number of hyperedges in $\mathcal{H}^*$ is

$$\begin{aligned}|\mathcal{H}^*| = e(A^*, B^*) + e(B^*) &\geq (k-1)\cdot(r-1)^{k-2}\frac{n-k+1}{(r-1)^{k-1}} + (\ell-1)\cdot r^{\ell-2}\frac{n-k+1}{r^{\ell-1}} \\ &= \left(\frac{\ell-1}{r} + \frac{k-1}{r-1}\right)(n-k+1).\end{aligned}$$

It now remains to show that $\mathcal{H}^*$ contains no $k\cdot S_\ell^+$. Every vertex in B^* is incident to exactly $\ell-1$ hyperedges that are contained completely in B^*. Therefore, any copy of S_ℓ^+ in $\mathcal{H}^*$ must use at least one vertex from A^*. Therefore, there are at most $k-1$ pairwise vertex-disjoint copies of S_ℓ^+. □

Theorem 3.2.2 gives the following corollary for a matching in the linear setting.

Corollary 3.2.3. *For $r, k \geq 2$, we have*

$$\mathrm{ex}_r^{\mathrm{lin}}(n, k \cdot K_2^+) = \frac{k-1}{r-1}n + O(1).$$

3.3 Berge star forests

A *System of Distinct Representative* (SDR) in a hypergraph $\mathcal{G}$ is a collection of c distinct hyperedges $h_1, h_2, \ldots, h_c$ and c distinct vertices $x_1, x_2, \ldots, x_c$ such that $x_i \in h_i$ for all i. We call x_i the *representative* of h_i. The *size* of an SDR is the number of hyperedges. We will also use the term *c-SDR* to refer to an SDR of size c.

Let $\mathcal{G}$ be a Berge-G and let f be an arbitrary bijection from $\mathcal{G}$ to G such that $f(h) \subset h$; i.e, f is a bijection that establishes that $\mathcal{G}$ is a Berge-G. A *skeleton* of $\mathcal{G}$ is the vertex set of the image of f (i.e., a copy of G embedded into $\mathcal{G}$). See Figure 3.3.

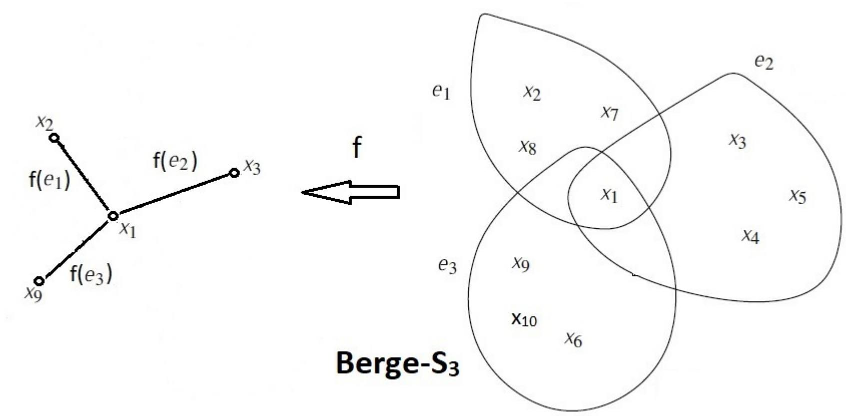

Figure 3.3: A Berge-S_3 hypergraph with skeleton $\{x_1, x_2, x_3, x_9\}$

Given a hypergraph $\mathcal{G}$ and a vertex x, the *link hypergraph* $\mathcal{G}_x$ is defined as

$$\mathcal{G}_x = \{h - x \mid h \in \mathcal{G}, x \in h\}.$$

Lemma 3.3.1. *Fix integers ℓ and $r \geq 2$, and let $\mathcal{G}$ be an r-uniform hypergraph.*

(1) *If $\ell > r$ and x is a vertex of degree $d(x) > \binom{\ell-1}{r-1}$, then there exists a Berge-S_ℓ with center x.*

(2) *If $\ell > r$ and x is a vertex of degree $d(x) = \binom{\ell-1}{r-1}$ such that the neighborhood of x contains at least ℓ vertices, then there exists a Berge-S_ℓ with center x.*

(3) *If $\ell \leq r$ and x is a vertex of degree $d(x) \geq \ell$, then there exists a Berge-S_ℓ with center x.*

Proof. Consider the $(r-1)$-uniform link hypergraph $\mathcal{G}_x$. By the degree condition on $d(x)$ in **(1)** or **(2)** we have $|E(\mathcal{G}_x)| \geq \binom{\ell-1}{r-1}$. Suppose the maximum SDR is of size $c < \ell$. Let $f_1, f_2, \dots, f_c$ be the vertices of a c-SDR and let S be the representatives. Let f be an arbitrary hyperedge distinct from $f_1, \dots, f_c$. Observe that f is contained in S as otherwise we can easily form a $(c+1)$-SDR with f which contradicts the maximality of c. Therefore, every hyperedge not part of the c-SDR must be contained in S. By condition **(1)** or **(2)** we have at least one f_i, say f_1, that is not contained in S. Let $y \in f_1 \setminus S$ and $z \in S$ be the representative of f_1. If f contains z, then we can form a $(c+1)$-SDR by changing the representative of f_1 to y and allowing z to be the representative of f, a contradiction. Thus, f does not contain z.

We distinguish two cases.

Case 1: f_1 is the only hyperedge of the c-SDR not contained in S.

If no hyperedge different from f_1 is incident to z, then the number of hyperedges is at most

$$1 + \binom{|S \setminus \{z\}|}{r-1} \leq 1 + \binom{\ell-2}{r-1} < \binom{\ell-1}{r-1}$$

which contradicts the bound on $|E(\mathcal{G}_x)|$. If there is a hyperedge (other than f_1) incident to z, then it must be a hyperedge of the c-SDR. So suppose f_2 is incident to z. Now if f is incident to the representative z' of f_2, then we can form a larger SDR by allowing z' to be the representative of f, z to be the representative of f_2 and y to be the representative of f_1, a contradiction. Therefore, f is disjoint from z and z'. Then the number of hyperedges is at most

$$c + \binom{|S \setminus \{z, z'\}|}{r-1} \leq (\ell - 1) + \binom{\ell - 3}{r-1} < \binom{\ell - 1}{r-1}$$

which contradicts the bound on $|E(\mathcal{G}_x)|$.

Case 2: There are at least two hyperedges of the c-SDR, say f_1 and f_2, not contained in S.

If z and z' are the representatives of f_1 and f_2, respectively, then f cannot contain z or z' as otherwise we can form a larger SDR. Therefore, the number of hyperedges is at most

$$c + \binom{|S \setminus \{z, z'\}|}{r-1} \leq (\ell - 1) + \binom{\ell - 3}{r-1} < \binom{\ell - 1}{r-1}$$

which contradicts the bound on $|E(\mathcal{G}_x)|$.

For case **(3)**, we will show that $\mathcal{G}_x$ contains a ℓ-SDR which implies that $\mathcal{G}$ contains a Berge-S_ℓ with center x. Suppose there is no ℓ-SDR. Then by Hall's theorem, there exists a collection of hyperedges $h_1, \ldots, h_t$ with $1 \leq t \leq \ell$ such that

$$\left| \bigcup_{i=1}^{t} h_i \right| < t.$$

Each h_i contains $r - 1$ vertices, so $|\cup h_i| \geq r - 1 \geq 2$. This implies that $t \geq 2$. As any pair of hyperedges span at least r vertices we have $|\cup h_i| \geq r \geq \ell$. This implies that $t > \ell$, a contradiction. □

Lemma 3.3.2. *Fix $\ell > r$ and suppose $\mathcal{G}$ is a Berge-S_ℓ-free r-uniform hypergraph. If x is a vertex of degree $d(x) = \binom{\ell-1}{r-1}$, then the link hypergraph $\mathcal{G}_x$ is a $K_{\ell-1}^{r-1}$.*

Proof. Suppose that the link hypergraph $\mathcal{G}_x$ is not a $K_{\ell-1}^{r-1}$; then the neighborhood of x must contain more than $\ell-1$ vertices as $d(x)=\binom{\ell-1}{r-1}$. Now we may apply Lemma 3.3.1 to find a Berge-S_ℓ in $\mathcal{G}$, a contradiction. □

We begin with a simple application of Lemma 3.3.1.

Theorem 3.3.3. *Fix integers ℓ and $r \geq 2$.*

(1) *If $\ell > r$, then*

$$\mathrm{ex}_r(n, \text{Berge-}S_\ell) \leq \binom{\ell}{r}\frac{n}{\ell}.$$

Furthermore, this bound is sharp whenever ℓ divides n.

(2) *If $\ell \leq r$, then*

$$\mathrm{ex}_r(n, \text{Berge-}S_\ell) \leq \frac{\ell-1}{r}n.$$

Furthermore, this bound is sharp whenever r divides n.

Proof. For the lower bound in Case **(1)** (i.e. for the sharpness assertion), consider the r-uniform hypergraph consisting of $\frac{n}{\ell}$ pairwise disjoint complete hypergraphs K_ℓ^r. A Berge-S_ℓ necessarily contains at least $r+1$ vertices (for $\ell \geq 2$), so no such complete hypergraph contains a Berge-S_ℓ. The number of hyperedges in this construction is exactly $\frac{n}{\ell}\binom{\ell}{r}$.

For the upper bound in Case **(1)** suppose that $\mathcal{H}$ is an r-uniform n-vertex hypergraph with $e(\mathcal{H}) > \frac{n}{\ell}\binom{\ell}{r}$ hyperedges. The average degree of $\mathcal{H}$ is

$$d(\mathcal{H}) > \frac{r}{n}\binom{\ell}{r}\frac{n}{\ell} = \binom{\ell-1}{r-1}.$$

This implies that $\mathcal{H}$ contains a vertex x of degree greater than $\binom{\ell-1}{r-1}$. Applying Lemma 3.3.1 to $\mathcal{H}_x$ gives a Berge-S_ℓ in $\mathcal{H}$, a contradiction.

The upper bound in Case **(2)** follows directly from the third part of Lemma 3.3.1. The lower bound (i.e. the sharpness statement) follows from the the fact that r-uniform $(\ell-1)$-regular n vertex hypergraphs exist when r divides n. We give a concrete example here. Arrange n vertices around a circle. By considering intervals of r consecutive vertices we can partition the n vertices into n/r classes each of size r. If we begin this partition from different starting vertices we may create $\ell-1<r$ different partitions such that no pair of partitions has a class in common. It is easy to see that this collection of $(\ell-1)\frac{n}{r}$ total partition classes forms an r-uniform $(\ell-1)$-regular hypergraph on n vertices. $\square$

We are now ready to prove the main theorem of this subsection. We begin with a construction of an r-uniform n-vertex Berge-$k \cdot S_\ell$-free hypergraph. Fix integers $r \geq 2$, $\ell \geq 1$ and $k \geq 1$ such that $\ell+k-1>r+1$. Put $n=s\cdot\ell+q$ for $0 \leq q < \ell$.

Let C^* be a set of $k-1$ vertices and B^* be a set of $n-k+1$ vertices. Partition the vertices of B^* into classes of size ℓ and (if $q>0$) a single class of size $q<\ell$. For each partition class S^* of B^* we form a complete r-uniform hypergraph $K^r_{\ell+k-1}$ (or K^r_{q+k-1}) on the vertices of $C^* \cup S^*$. Let $\mathcal{H}_{n,r,\ell,k}$ be the resulting hypergraph as Figure 3.4.

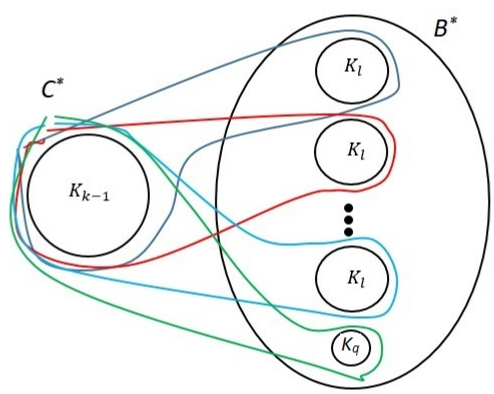

Figure 3.4: The r-uniform hypergraph $\mathcal{H}_{n,r,\ell,k}$

The number of hyperedges in $\mathcal{H}_{n,r,\ell,k}$ is exactly

$$\left(\binom{\ell+k-1}{r}-\binom{k-1}{r}\right)\frac{n-q-k+1}{\ell}+\binom{q+k-1}{r}.$$

The skeleton of any Berge-S_ℓ in $\mathcal{H}_{n,r,\ell,k}$ must use at least one vertex of C^*. Therefore, there are at most $k-1$ copies of a Berge-S_ℓ that have vertex-disjoint skeletons; i.e., $\mathcal{H}_{n,r,\ell,k}$ is Berge-$k\cdot S_\ell$-free.

Theorem 3.3.4. *For $\ell+k-2>r$ and n large enough and such that ℓ divides $n-k+1$,*

$$\mathrm{ex}_r(n, \text{Berge-}k\cdot S_\ell) \leq \left(\binom{\ell+k-1}{r}-\binom{k-1}{r}\right)\left\lceil\frac{n-k+1}{\ell}\right\rceil+\binom{k-1}{r}.$$

Furthermore, this bound is sharp and $\mathcal{H}_{n,r,\ell,k}$ is the unique hypergraph achieving this bound.

Proof. In order to prove the theorem it is enough to restrict ourselves to the case when ℓ divides $n-k+1$.

We proceed by induction on k and r. The case when $r=2$ and $k\geq 1$ follows from Theorem 2.2.1 as Berge-S_ℓ is simply a copy of the graph S_ℓ; i.e.,

$$\begin{aligned}\mathrm{ex}(n,k\cdot S_\ell) &\leq (k-1)(n-k+1)+\binom{k-1}{2}+\frac{\ell-1}{2}(n-k+1)\\ &=\frac{n-k+1}{\ell}\left(\binom{\ell+k-1}{2}-\binom{k-1}{2}\right)+\binom{k-1}{2}.\end{aligned}$$

The case when $r\geq 3$ and $k=1$ follows from Theorem 3.3.3 as $\ell>r$.

Now fix integers ℓ, r, k as in the statement of the theorem and suppose n is large enough.

Let $\mathcal{H}$ be an n-vertex r-uniform hypergraph with no Berge-$k\cdot S_\ell$ and at least $e(\mathcal{H}^*)$ hyperedges. Let C be the set of vertices in $\mathcal{H}$ of degree greater than $\binom{k(\ell+1)-1}{r-1}$. Put $c=|C|$ and let $B=V(\mathcal{H})-C$ be the set of remaining vertices.

Claim. $c \le k-1$.

Proof. Suppose $c \ge k$ and let C' be a set of k vertices of degree greater than $\binom{k(\ell+1)-1}{r-1}$. It is easy to see that there is a Berge-S_ℓ with center in C' but whose skeleton is otherwise disjoint from C'. Now suppose that we have found a Berge-$(k-1)\cdot S_\ell$ whose skeleton intersects C' in $(k-1)$ vertices. The skeleton of this Berge-$(k-1)\cdot S_\ell$ spans $(k-1)(\ell+1)$ vertices. Let x be the vertex in C' not in the skeleton of this Berge-$(k-1)\cdot S_\ell$. By Lemma 3.3.1 there is a Berge-$S_{k(\ell+1)}$ with center x. At most $(k-1)(\ell+1)$ of the vertices of the skeleton of this Berge-$S_{k(\ell+1)}$ intersect the skeleton of the Berge-$(k-1)\cdot S_\ell$. Therefore, there is a Berge-S_ℓ whose skeleton is disjoint from the skeleton of the Berge-$(k-1)\cdot S_\ell$. In particular, we have a Berge-$k\cdot S_\ell$, a contradiction. □

We distinguish two cases based on the value of c.

Case 1: $c < k-1$.

Claim. There exists a constant $\epsilon > 0$ such that the average degree of the vertices in B is at least $\binom{\ell+c-1}{r-1}+\epsilon$.

Proof. Suppose there is no such ϵ. Then the average degree of vertices in B is at most $\binom{\ell+c-1}{r-1}$. For each $x \in C$ consider the link hypergraph $\mathcal{H}_x = \{e - x \,|\, x \in e \in \mathcal{H}\}$. If $\mathcal{H}_x$ contains an $(r-1)$-uniform Berge-$(k-1)\cdot S_\ell$, then $\mathcal{H}$ contains a Berge-$(k-1)\cdot S_\ell$ whose skeleton does not use the vertex x. The degree of x is large enough so that we can find greedily a copy of a Berge-S_ℓ with center x that is edge-disjoint from the Berge-$(k-1)\cdot S_\ell$. Together we have a Berge-$k\cdot S_\ell$, a contradiction. Therefore, $\mathcal{H}_x$ is an $(r-1)$-uniform Berge-$(k-1)\cdot S_\ell$-free hypergraph on $n-1 \le n$ vertices. By induction we have

$$d(x) = |\mathcal{H}_x| \le \left(\binom{\ell+k-2}{r-1} - \binom{k-2}{r-1}\right)\frac{n-k+1}{\ell} + \binom{k-2}{r-1}.$$

Thus,

$$\sum_{x\in V(\mathcal{H})} d(x) = \sum_{x\in C} d(x) + \sum_{x\in B} d(x)$$
$$\leq c\left(\left(\binom{\ell+k-2}{r-1} - \binom{k-2}{r-1}\right)\frac{n-k+1}{\ell} + \binom{k-2}{r-1}\right) + (n-c)\binom{\ell+c-1}{r-1}.$$

On the other hand, for $\mathcal{H}^* = \mathcal{H}_{n,r,\ell,k}$ we have

$$\sum_{x\in V(\mathcal{H}^*)} d(x) = \sum_{x\in C^*} d(x) + \sum_{x\in B^*} d(x)$$
$$\geq (k-1)\left(\left(\binom{\ell+k-2}{r-1} - \binom{k-2}{r-1}\right)\frac{n-k+1}{\ell} + \binom{k-2}{r-1}\right) + (n-k+1)\binom{\ell+k-2}{r-1}.$$

It is clear that the coefficient of n in the second inequality is larger than that in the first inequality. Therefore, for n large enough we have $\sum_{x\in V(\mathcal{H})} d(x) < \sum_{x\in V(\mathcal{H}^*)} d(x)$ which implies $e(\mathcal{H}) < e(\mathcal{H}^*)$, a contradiction. □

Claim. There are $\Omega(n)$ many vertices of B with degree greater than $\binom{\ell+c-1}{r-1}$.

Proof. By the preceding Claim, the sum of the degrees in B is at least $(\binom{\ell+c-1}{r-1}+\epsilon)(n-c)$ for some $\epsilon > 0$. Now let b be the number of vertices of degree less than $\binom{\ell+c-1}{r-1}+1$ in B. Recall that the vertices in B have degree at most $D = \binom{k(\ell+1)-1}{r-1}$. Thus, the sum of degrees in B is at most $\binom{\ell+c-1}{r-1}b + D(n-c-b)$. Combining these estimates and solving for b gives

$$b < \frac{D - \binom{\ell+c-1}{r-1} - \epsilon}{D - \binom{\ell+c-1}{r-1}}(n-c).$$

Therefore, the number of vertices of degree greater than $\binom{\ell+c-1}{r-1}$ is $\Omega(n)$ □

Let us call two vertices $x, y \in B$ *far* if they do not share a common neighbor in B (they may still have a common neighbor in C). As the vertices in B have constant maximum degree we can find a subset B' of size $\Omega(n)$ such that all vertices have degree greater than $\binom{\ell+c-1}{r-1}$ and all pairs of vertices are far.

For each vertex x in B' there is a Berge-$S_{\ell+c}$ with center x. Therefore, there is a Berge-S_ℓ with center x whose skeleton is disjoint from C. As any two vertices in B' do not share a common neighbor in B we have a collection of hyperedge-disjoint Berge-S_ℓ's. As n is large enough we can find a Berge-$k \cdot S_\ell$, a contradiction.

Case 2: $c = k - 1$.

In this case we have that each vertex in B has degree at most $\binom{\ell+k-2}{r-1}$. Indeed, if there is a vertex x of degree greater than $\binom{\ell+k-2}{r-1}$, then there is a Berge-$S_{\ell+k-1}$ with center x. The skeleton of this Berge-$S_{\ell+k-1}$ uses at most $k-1$ vertices from C, so there remains a Berge-S_ℓ whose skeleton is contained in B. This Berge-S_ℓ together with $k-1$ more Berge-S_ℓ's constructed with the vertices in C as their centers form a copy of a Berge-$k \cdot S_\ell$, a contradiction.

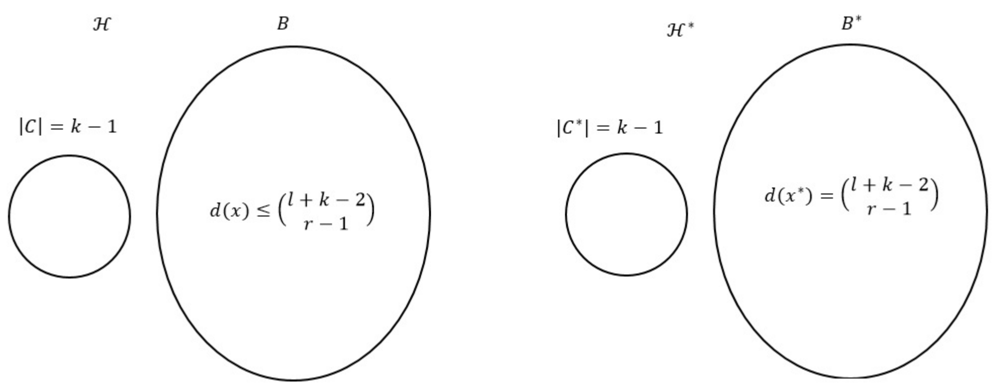

Figure 3.5: The r-uniform hypergraph $\mathcal{H}$ with $|C| = k - 1$, and r-uniform hypergraph $\mathcal{H}^*$

We distinguish two subcases.

Case 2.1: There exists a vertex $x \in B$ with $d(x) < \binom{\ell+k-2}{r-1}$.

Let us compare $\mathcal{H}$ to the construction $\mathcal{H}^*$. Every vertex in B has degree at most $\binom{\ell+k-2}{r-1}$

while every vertex in B^* has degree $\binom{\ell+k-2}{r-1}$. See Figure 3.5.

As $e(\mathcal{H}) \geq e(\mathcal{H}^*)$ this implies that there exists a vertex $y \in C$ and a vertex $y^* \in C^*$ such that $d(y) > d(y^*)$. Define two j-uniform multi-hypergraphs as follows:

$$E_j^y = \{e \setminus C \,:\, y \in e \text{ and } |e \setminus C| = j\} \text{ and } E_j^{y^*} = \{e \setminus C^* \,:\, y \in e \text{ and } |e \setminus C^*| = j\}.$$

Note that the hyperedges in E_j^y have multiplicity at most $\binom{k-2}{r-1-j}$ and those in $E_j^{y^*}$ have multiplicity exactly $\binom{k-2}{r-1-j}$.

Observe that when $k-1 \geq r-1$ each vertex of B^* is in a hyperedge with each subset of C^* of size $r-1$. This implies that $|E_1^{y^*}| \geq |E_1^y|$. When $k-1 < r-1$, then $E_1^{y^*} = E_1^y = \emptyset$. Therefore, as $d(y) > d(y^*)$, we have that $|E_j^y| > |E_j^{y^*}|$ for some $j \geq 2$. Now let E be the j-uniform hypergraph resulting from deleting all repeated hyperedges in E_j^y.

Thus, the number of hyperedges in the j-uniform family E is

$$\begin{aligned} |E| \geq \binom{k-2}{r-1-j}^{-1} |E_j^y| > \binom{k-2}{r-1-j}^{-1} |E_j^{y^*}| &= \binom{k-2}{r-1-j}^{-1} \binom{k-2}{r-1-j} \frac{n-k+1}{\ell} \binom{\ell}{j} \\ &= \frac{n-k+1}{\ell} \binom{\ell}{j} \geq \mathrm{ex}_j(n-k+1, \text{Berge-}S_\ell). \end{aligned}$$

Therefore, there is a j-uniform Berge-S_ℓ in E on the vertices of B. As each edge of E is contained in an edge of $\mathcal{H}$, this corresponds to a Berge-S_ℓ in $\mathcal{H}$ with skeleton contained in B. This Berge-S_ℓ together with $k-1$ more Berge-S_ℓ's constructed with the vertices in C as their centers form a copy of a Berge-$k \cdot S_\ell$, a contradiction.

Case 2.2: For every vertex $x \in B$ we have $d(x) = \binom{\ell+k-2}{r-1}$.

If the neighborhood $N(x)$ of x contains more than $\ell+k-2$ vertices, then by Lemma 3.3.1 there is a Berge-$S_{\ell+k-1}$ with center x. As before, we can use this Berge-$S_{\ell+k-1}$ to show the existence of a Berge-$k \cdot S_\ell$, a contradiction.

As $d(x) = \binom{\ell+k-2}{r-1} > \binom{\ell+k-3}{r-1}$, Lemma 3.3.1 implies that there is a Berge-$S_{\ell+k-2}$ with center x. The skeleton of this Berge-$S_{\ell+k-2}$ must use each vertex of C, for otherwise we have a Berge-S_ℓ whose skeleton is contained in B which we can combine with the Berge-$(k-1)\cdot S_\ell$ with centers in C to form a Berge-$k \cdot S_\ell$, a contradiction. Therefore, $N(x) \cap C = C$ and $|N(x) \cap B| = \ell - 1$. Now let y be a vertex in $N(x) \cap B$. For each z different from y in $N(x) \cup x$, there is a hyperedge containing x and y. Therefore, the neighborhood of y contains $N(x) \cup x$. If $N(y)$ is any larger, then by Lemma 3.3.1 we can find a Berge-$S_{\ell+k-1}$ with center y, a contradiction.

This implies that x is contained in a complete graph $K^r_{\ell+k-1}$ that intersects C in exactly $k-1$ vertices. This holds for every vertex $x \in B$, so $\mathcal{H}$ has the exact structure as the construction $\mathcal{H}^*$. □

As a consequence we have the Turán number of a k-matching in an r-uniform hypergraph when when $k > r + 1$.

Corollary 3.3.5. *For $k \geq r + 1$, we have*

$$\mathrm{ex}_r(n, \text{Berge-}k \cdot K_2) \leq \binom{k-1}{r-1}(n - k + 1) + \binom{k-1}{r}.$$

Chapter 4

Future Directions

4.1 Forest of Path and Star components

We will attempt to generalize Theorem 2.2.4 to a forest of path and star components of different orders. Let $F = F_1 \cup F_2$ where F_1 is a forest of paths $P_{v_1}, \ldots, P_{v_s}$ and F_2 is a forest of stars $S_{d_1}, \ldots, S_{d_k}$ and determine $\mathrm{ex}(n, F)$. This would also generalize Theorem 2.1.5 and Theorem 2.2.1.

In Theorem 2.1.7 we reproved Theorem 2.1.5 when ℓ is even. Our proof gives a significantly smaller lower bound requirement on n. We would like to use our approach to reprove Theorem 2.1.5 for ℓ odd.

The value of $\mathrm{ex}(n, k \cdot P_\ell)$ has also been considered for fixed values of k and ℓ and small values of n. For example,

Theorem 4.1.1 (Bielak and Kieliszek, [1])**.**

$$\mathrm{ex}(n, 2\cdot P_5) = \begin{cases} 3n-5 & \textit{if } n \geq 18 \\ 48 & \textit{if } n = 17 \\ 45 & \textit{if } n = 16 \\ 43 & \textit{if } n = 15 \\ 42 & \textit{if } n = 14 \\ 42 & \textit{if } n = 13 \\ 39 & \textit{if } n = 12 \\ 37 & \textit{if } n = 11 \\ 36 & \textit{if } n = 10 \\ \binom{n}{2} & \textit{if } n \leq 9 \end{cases}$$

Another example is as follows.

Theorem 4.1.2 (Lan, Qin and Shi, [23])**.**

$$\mathrm{ex}(n, 2\cdot P_7) = \begin{cases} 5n-14 & \textit{if } n \geq 22 \\ 94 & \textit{if } n = 21 \\ 93 & \textit{if } n = 20 \\ 93 & \textit{if } n = 19 \\ 88 & \textit{if } n = 18 \\ 84 & \textit{if } n = 17 \\ 81 & \textit{if } n = 16 \\ 79 & \textit{if } n = 15 \\ 78 & \textit{if } n = 14 \\ \binom{n}{2} & \textit{if } n \leq 13 \end{cases}$$

We would like to generalize both of these theorems. In particular, we conjecture:

Conjecture 4.1.3. ***(1)*** *Let $\ell \geq 4$ be even and n be a positive integer. Then*

$$\mathrm{ex}(n, 2\cdot P_\ell) = \begin{cases} (\ell-1)n - \frac{(\ell-1)(\ell-2)}{2} & \textit{if } n \geq 3\ell-2 \\ \binom{2\ell-1}{2} + \binom{n-2\ell+1}{2} & \textit{if } 2\ell \leq n < 3\ell-2 \\ \binom{n}{2} & \textit{if } n \leq 2\ell-1 \end{cases}$$

(2) Let $\ell \geq 7$ be odd and n be a positive integer. Then

$$\mathrm{ex}(n, 2 \cdot P_\ell) = \begin{cases} (\ell-2)n - \frac{(\ell-1)(\ell-2)}{2} + 1 & \text{if } n > 3\ell \\ \binom{2\ell-1}{2} + \binom{\ell-1}{2} + 1 & \text{if } 3\ell - 1 < n \leq 3\ell \\ \binom{2\ell-1}{2} + \binom{n-2\ell+1}{2} & \text{if } 2\ell \leq n \leq 3\ell - 1 \\ \binom{n}{2} & \text{if } n \leq 2\ell - 1 \end{cases}.$$

4.2 Hypergraph linear forests

We have extended Theorem 3.3.4 to the case when $r \geq \ell+k-1$. In particular, in a forthcoming manuscript we have shown:

Theorem 4.2.1. *Fix integers $\ell, k \geq 1$, $r \geq \ell + k - 1$ and n large enough,*

$$\mathrm{ex}_r(n, \text{Berge-}k \cdot S_\ell) \leq \frac{\ell-1}{r-k+1}(n-k+1).$$

Furthermore, this bound is sharp whenever $r - k + 1$ divides $n - k + 1$.

Following Theorem 3.1.4, Theorem 3.2.2 and Theorem 3.3.4, we will work on finding the Turán number of a linear forest in different host hypergraphs; i.e., suppose F is the r-uniform expansion of $k \cdot P_\ell$ or a Berge-$k \cdot P_\ell$ and we want to find $\mathrm{ex}_r(n, F)$. Note that when $\ell = 1$, $\mathrm{ex}_r(n, F)$ is calculated in Theorems 3.1.4, 3.2.2 and 3.3.4.

Győri, Katona and Lemons [19] determined $\mathrm{ex}_r(n, \text{Berge-}P_{\ell+1}) = \frac{n}{\ell}\binom{\ell}{r}$ where $P_{\ell+1}$ is the Berge-path on ℓ hyperedges and $\ell > r$. Then by Theorem 3.3.3, the extremal bounds of a Berge-S_ℓ and a Berge-$P_{\ell+1}$ are equal when $\ell > r$ as the graph version.

Our conjecture for Turán number of Berge-$k \cdot P_\ell$ is:

Conjecture 4.2.2. *Suppose F is $k \geq 2$ disjoint Berge paths $P_{\ell+1}$ and $\ell > 2r$. For large* *enough n,*

$$\mathrm{ex}_r(n,F)=\binom{k\lfloor\frac{\ell+1}{2}\rfloor-1}{r-1}(n-k\lfloor\tfrac{\ell+1}{2}\rfloor+1)+\binom{k\lfloor\frac{\ell+1}{2}\rfloor-1}{r}+c$$

where $c=1$ if ℓ is even and $c=0$ if ℓ is odd.

Bibliography

[1] H. Bielak and S. Kieliszek, *The Turán number of the graph* $2P_5$, Discussiones Mathematicae: Graph Theory, 36 (2016), pp. 683–694.

[2] B. Bollobás, *Modern Graph Theory 3rd edition*, Springer Science+Business Meda, Inc., USA, 1998.

[3] J. A. Bondy and M. Simonovits, *Cycles of even length in graphs*, J. Combin. Theory Ser. B, 16 (1974), pp. 97–105.

[4] N. Bushaw and N. Kettle, *Turán numbers of multiple paths and equibipartite forests*, Combinatorics, Probability and Computing, 20 (2011), pp. 837–853.

[5] 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) (1987), pp. 111–126.

[6] A. Davoodi, E. Győri, A. Methuku, and C. Tompkins, *An Erdős-Gallai type theorem for uniform hypergraphs*, European Journal of Combinatorics, 69 (2018), pp. 159–162.

[7] R. A. Duke and P. Erdős, *Systems of finite sets, having a common intersection, in: Proc. 8th Southeastern Conf. Combinatorics*, Congressus Numerantium, 19 (1977), pp. 247–252.

[8] P. Erdős, *Extremal problems in graph theory, Theory of Graphs and its Applications*, Proc. Symp. Smolenice, Plublishing House of the Czechoslovak Academy of Science, Prague, 1964.

[9] ——, *A problem on independent r-tuples*, Ann. Univ. Sci. Budapest, 8 (1965), pp. 93–95.

[10] P. Erdős and T. Gallai, *On maximal paths and circuits of graphs*, Acta Math. Acad. Sci. Hungar., 1 (1975), pp. 215–235.

[11] P. Erdős and R. Rado, *Intersection theorems for systems of sets*, Journal of the London Mathematical Society, Second Series, 35(1) (1960), pp. 85–90.

[12] P. Erdős and M. Simonovits, *A limit theorem in graph theory*, Studia Sci. Math. Hungar., 1 (1966), pp. 51–57.

[13] P. Erdős and A. H. Stone, *On the structure of linear graphs*, Bull. Amer. Math. Soc., 52 (1946), pp. 1087–1091.

[14] R. J. Faudree and R. H. Schelp, *Path Ramsey numbers in multicoloring*, J. Comb. Theory, Ser. B, 75 (1975), pp. 141–144.

[15] Z. Füredi and M. Simonovits, *A survey of Turán problems for expansions, Recent Trends in Combinatorics*, vol. 25, Erdős Centennial. Bolyai Society Mathematical Studies, 2013.

[16] D. Gerbner, A. Methuku, and C. Palmer, *General lemmas for Berge-Turán hypergraph problems*, arXiv:1808.10842v1, (2018).

[17] D. Gerbner and C. Palmer, *Extremal results for Berge-hypergraphs*, SIAM Journal on Discrete Mathematics, 31 (2017), pp. 2314–2327.

[18] I. Gorgol, *Turán numbers for disjoint copies of graphs*, Graphs and Combinatorics, 27 (2011), pp. 661–667.

[19] E. Győri, G. Katona, and N. Lemons, *Hypergraph extensions of the Erdos and Gallai theorem*, Electronic Notes in Discrete Mathematics, 36 (2010), pp. 655–662.

[20] E. Győri and N. Lemons, *Hypergraphs with no cycle of a given length*, Combin. Probab. Comput., 21(1) (2012), pp. 193–201.

[21] D. Johnston, C. Palmer, and A. Sarkar, *Rainbow Turán problems for paths and forests of stars*, Electronic Journal of Combinatorics, 24 (2017), p. 15.

[22] A. Kostochka, D. Mubayi, and J. Verstraëte, *Turán problems and shadows I: paths and cycles*, J. Combin. Theory Ser. A, 29 (2015), pp. 57–79.

[23] Y. Lan, Z. Qin, and Y. Shi, *The Turán number of* $2P_7$, arXiv:1711.07734v1, (2017).

[24] B. Lidický, H. Liu, and C. Palmer, *On the Turán number of forests*, Electronic Journal of Combinatorics, 20 (2013), p. # P62.

[25] D. Mubayi and J. Verstraëte, *A survey of Turán problems for expansions, Recent Trends in Combinatorics*, Springer, 2011.

[26] P. Turán, *On the theory of graphs*, Colloquium Math., 3 (1954), pp. 19–30.

[27] L. Yuan and X. Zhang, *The Turán number of disjoint copies of paths*, Discrete Mathematics, 340(2) (2017), pp. 132–139.