Spring 5-2022

Extremal Problems in Graph Saturation and Covering

Adam Volk
University of Nebraska-Lincoln, avolk@huskers.unl.edu

EXTREMAL PROBLEMS IN GRAPH SATURATION AND COVERING

by

Adam Volk

A DISSERTATION

Presented to the Faculty of

The Graduate College at the University of Nebraska

In Partial Fulfilment of Requirements

For the Degree of Doctor of Philosophy

Major: Mathematics

Under the Supervision of Professor Jamie Radcliffe

Lincoln, Nebraska

May, 2022

EXTREMAL PROBLEMS IN GRAPH SATURATION AND COVERING

Adam Volk, Ph.D.

University of Nebraska, 2022

Adviser: Jamie Radcliffe

This dissertation considers several problems in extremal graph theory with the aim of finding the maximum or minimum number of certain subgraph counts given local conditions. The local conditions of interest to us are saturation and covering. Given graphs F and H, a graph G is said to be F-*saturated* if it does not contain any copy of F, but the addition of any missing edge in G creates at least one copy of F. We say that G is H-*covered* if every vertex of G is contained in at least one copy of H. In the former setting, we prove results regarding the minimum number of copies of certain subgraphs, primarily cliques and stars. Special attention will be given to the somewhat surprising challenge of minimizing the number of cherries, i.e. stars with two vertices of degree 1, in triangle-saturated graphs and its connection to Moore graphs. In the latter setting, we are interested in maximizing the number of independent sets of a fixed size in H-covered graphs, primarily when H is a star, path, or disjoint union of edges. Along the way, we will introduce and prove several results regarding a new style of question regarding graph saturation, namely determining for which graphs F there exist trees that are F-saturated. We will call such graphs *tree-saturating*.

ACKNOWLEDGMENTS

There are many people in my life who have played an important role in where I am today and without whose support, you would not be reading this. I first want to thank my advisor Jamie Radcliffe who has worked with me over the past four years and whose discrete math sequence during my first year of graduate school cemented my interest in the field. On the academic side, I also want to thank the members of my supervisory committee: Christine Kelley, Xavier Pérez Giménez, Vinod Variyam, and Mark Walker. I especially want to thank Christine and Xavier for reading my dissertation. I also appreciate the faculty members at UNL who have helped me to grow as an instructor and the faculty members at the University of Dayton who helped cultivate my passion for mathematics and pushed me to pursue that passion.

Next, I want to thank my family back home in Ohio for their prayers and continued support over the past six years. I also don't know how I would have reached this moment without the support of my friends back home, especially my college roommates, and the friends I've made in Nebraska, especially my graduate school cohort and my academic siblings.

Lastly, I want to thank God for the strength He has given me through this process and for all of the people He has placed in my life. I am also incredibly grateful to my wife Dana for all of her encouragement over these past few years. While I am incredibly proud to be writing these pages, meeting you is why I am most grateful for moving to Nebraska.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

The problems considered in this dissertation belong to the field of extremal graph theory, an area of mathematics concerned with the way in which certain properties of graphs influence the counts of various substructures. In one of his books [6], Béla Bollobás says,

"Extremal graph theory, in its strictest sense, is a branch of graph theory developed and loved by Hungarians."

We will mention just a few of the Hungarians - and other mathematicians - whose work has shaped the history of extremal graph theory, leading up to the problems considered in these pages. This begins in 1941 when Turán determined the maximum number of edges in a graph with a fixed number of vertices n and no complete subgraphs K_t [33]. The unique optimal graph, known as the Turán graph, is given by the complete multipartite graph with parts as equal of size as possible. See Figure 1.1 for the optimal K_5-free graph on 12 vertices. Notice that any copy of K_5 must include at least two vertices of the same color, but no such pairs are adjacent. When $t = 3$, the optimal graph is a complete bipartite graph with sides as equal as possible. This fact, along with the optimal edge count of $\lfloor \frac{n^2}{4} \rfloor$ was known previously by Mantel [27].

The problem of determining the maximum number of edges allowed in an F-free

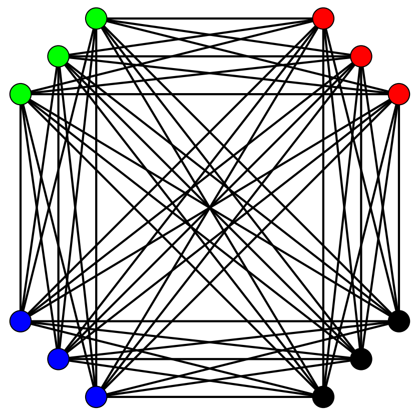

Figure 1.1: The K_5-free graph on 12 vertices with $\mathrm{ex}(12, K_5)$ edges

graph on n vertices received a significant amount of attention over the ensuing decades [31]. This maximum count is referred to as the extremal number of F and is denoted $\mathrm{ex}(n, F)$. The graph property that we are most interested in is directly related to this.

Definition 1. *We say that a graph G is F-**saturated** if it does not contain any copy of F, but the addition of any missing edge in G creates at least one copy of F.*

Note that every time we refer to a *copy* of F, we are saying that there is a subgraph of G isomorphic to F. This subgraph is not necessarily induced. That is, we are considering a subset of the vertices and a subset of the edges on those vertices. There is a version of graph saturation that looks at induced subgraphs [28], but that will not be considered in this dissertation.

Since F-saturated graphs are F-free, $\mathrm{ex}(n, F)$ is always attained by an F-saturated graph and has the most edges among such graphs. Restricting our scope from F-free graphs to F-saturated graphs also allows one to consider the minimization counterpart to the problem of finding an extremal number. We write $\mathrm{sat}(n, F)$ to denote the minimum number of edges among all F-saturated graphs on n vertices. Turán's Theorem provides a value for $\mathrm{ex}(n, K_t)$ for all $t \geq 3$ as well as the unique extremal

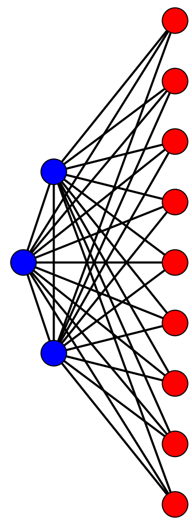

Figure 1.2: $K_3 + \overline{K}_9$; The K_5-saturated graph on 12 vertices with $\mathrm{sat}(12, K_5)$ edges

graph. The following theorem of Erdős, Hajnal, and Moon determines the value of $\mathrm{sat}(n, F)$ along with a unique optimal construction as well.

Theorem 1 (Erdős, Hajnal, and Moon, 1964)**.** *For every $n \geq t \geq 2$,*

$$\mathrm{sat}(n, K_t) = (n-t+2)(t-2) + \binom{t-2}{2}.$$

The graph $K_{t-2} + \overline{K}_{n-t+2}$ is the unique extremal example.

The graph in Figure 1.2 is an example of a complete split graph. $K_{t-2} + \overline{K}_{n-t+2}$ is the disjoint union of a clique on $t-2$ vertices and an independent set, a set of vertices with no edges, on $n-t+2$ vertices, along with all $(t-2)(n-t+2)$ edges between the two sets. As illustrated in the case where $t = 5$, a copy of K_t must include at least two vertices on the right, and the addition of any edge within that independent set will create a copy of K_t. We will refer to these graphs simply as *split graphs* for the remainder of this dissertation, and we will see that their property of minimizing edges among clique-saturated graphs is not the only thing that is special about them when it comes to graph saturation.

Saturation numbers have been well-studied and a collection of some of these results can be found in [17]. Now, in extremal graph theory, we are not only concerned with counting edges. In fact, counting edges is simply a special case of counting the number of copies of some given subgraph in a host graph. An important result of this type is a generalization of Turán's Theorem proved independently by Zykov in [35] and Erdős in [11], which states that the same graph which maximizes the number of edges among K_t-saturated graphs on n vertices also maximizes the number of copies of smaller cliques K_r among such graphs. On the other end, Bollobás proved that the same graph which minimizes the number of edges among K_t-saturated graphs also minimizes the number of copies of smaller cliques K_r among such graphs [4]. Interestingly, this minimization result extends to counting subgraphs in hypergraphs; however, the maximization problem for hypergraphs is notoriously open in even the smallest cases.

Other results regarding the number of copies of a given subgraph in graphs with certain properties can be found in [1, 15, 25]. More recently, Alon and Shikhelman began a systematic study of the minimum number of copies of a target graph H among F-saturated graphs on n vertices for various choices of H and F [3]. We will write $\mathrm{ex}_H(n, F)$ to denote this generalized extremal number. Note that by definition $\mathrm{ex}_{K_2}(n, F) = \mathrm{ex}(n, F)$.

Inspired by this, Kritschgau, Methuku, Tait, and Timmons [23] introduced the generalized saturation number, which we will denote $\mathrm{sat}_H(n, F)$, to give the minimum number of copies of H among F-saturated graphs on n vertices. In addition to proving general results, they focused on the cases where at least one of H and F was a clique or cycle. Extending one of their results and proving a conjecture of Kritschgau et al., Chakraborti and Loh [8] showed the following.

Theorem 2 (Chakraborti and Loh, 2019)**.** *For every $t > r \geq 2$, there exists a constant $n_{r,t}$ such that, for all $n \geq n_{r,t}$, we have*

$$\mathrm{sat}_{K_r}(n, K_t) = (n - t + 2)\binom{t-2}{r-1} + \binom{t-2}{r}.$$

Furthermore, for n sufficiently large, the (complete) split graph $K_{t-2} + \overline{K}_{n-t+2}$ is the unique extremal graph.

In addition, Chakraborti and Loh showed that for n sufficiently large the split graph minimizes copies of cycles among K_t-saturated graphs. Inspired by their results, Chakraborti and Loh asked if the split graph minimizes the number of copies of any F among K_t-saturated graphs. In Chapter 2 we demonstrate that this is not the case when F is a star S_r with $r \geq 3$. In fact, the split graph is far from optimal in this scenario. However, when we consider copies of S_2, which we will call cherries throughout, we prove the following.

Theorem 3. *For all $n \geq t \geq 3$,*

$$\mathrm{sat}_{S_2}(n, K_t) = \frac{t-2}{2}n^2 + O(n^{3/2}).$$

Chapter 2 deals with this question of minimizing stars in clique-saturated graphs as well as other problems concerned with the generalized saturation number $\mathrm{sat}_H(n, F)$. In particular, we look at the other variations involving stars and cliques, namely $\mathrm{sat}_{K_r}(n, S_t)$ and $\mathrm{sat}_{S_r}(n, S_t)$. Taking advantage of strong structural constraints on star-saturated graphs, we prove the following concerning which K_t-saturated graph minimizes the number of copies of S_r. The graph $\mathrm{KR}_{t,n}(m)$, which is a K_t-saturated graph on n vertices with an additional parameter m, will be described in detail in Chapter 2.

Theorem 4. *For all $n \geq 2t-1$ with $t \geq 2$ and $r < t$,*

$$\operatorname{sat}_{S_r}(n, S_t) = \min_{0 \leq m \leq t-1} s_r(\mathrm{KR}_{t,n}(m)).$$

Note also that

$$s_r(\mathrm{KR}_{t,n}(m)) = \begin{cases} m\binom{m-1}{r} + (n-m)\binom{t-1}{r} & \text{if } (t-1)(n-m) \text{ is even} \\ m\binom{m-1}{r} + (n-m)\binom{t-1}{r} + \binom{m-1}{r-1} & \text{if } (t-1)(n-m) \text{ is odd.} \end{cases}$$

When counting cliques in star-saturated graphs we will focus on determining values of n for which $\operatorname{sat}_{K_r+1}(n, S_t) = 0$. As a sufficient criteria for this to occur, we prove the following.

Theorem 5. *Let $r \geq 3$ and $t \geq 3$ be fixed. There exists an n-vertex, r-partite, S_t-saturated graph if*

$$n \geq \max\left(t+1, \min_{0 \leq c \leq r-2}\left\{(r-c)\left\lceil\frac{t-1}{r-c-1}\right\rceil + r - c\right\}\right). \tag{1.1}$$

We prove similar, but not quite matching, necessary conditions for the existence of such graphs and show that loosening the restriction from r-partite graphs to K_{r+1}-free graphs does not change the bounds on n asymptotically.

We also consider $\operatorname{sat}_{S_r}(n, K_t)$ in a more restricted setting, proving that the split graph is still not optimal for minimizing copies of S_r with $r \geq 3$ when we consider graphs with a linear maximum degree. We will also discuss the problem of counting cherries in triangle-free graphs and its connection to Moore graphs of diameter 2, and we will look at generalized saturation numbers involving paths.

Inspired by observations made while investigating these generalized saturation problems, we introduced a new style of problem related to graph saturation. That is, given a graph H, does there exist a tree T such that H is T-saturated? If so, we call H *tree-saturating*. Similar questions can be asked for other classes than trees such as triangle-free graphs or r-partite graphs. This topic is the focus of Chapter 3. Here, our primary result characterizes the spiders, trees with a single vertex of degree greater than 2, which are tree-saturating.

In Chapter 4 we direct our attention to a regime separate from but with similarities to graph saturation. Another way of stating our definition for graph saturation is that a graph G is F-saturated if every non-edge of G is contained in at least one copy of F when that edge is added. A similar property that we can consider is the following.

Definition 2. *We say that a graph G is H-**covered** if every vertex of G is contained in at least one copy of H.*

This concept was first formally discussed by Chakraborti and Loh in [7]. One of their primary results proves that for a given graph H the minimum number of copies of K_t among H-covered graphs on n vertices is given by the solution to an integer program. When $t = 2$, this provides the minimum number of edges in such graphs. Equivalently, it provides the maximum number of independent sets of size 2.

The authors comment that the structure of graphs minimizing independent sets of larger size seems to be quite different. This is where we pick up in Chapter 4. In addition to looking at small graphs, choices of H that we address include stars, paths, and disjoint sets of edges. We point out how some known results in extremal graph theory allow us to quickly address star-covered graphs, and we prove the following main result concerning paths and which path-covered graph maximizes independent sets of size t.

Theorem 6. *Let G be a P_k-covered graph on n vertices and let S be the spider with*

$n-k-1$ legs of length 1 and one leg of length $k-2$. Then for all $t \geq 3$,

$$i_t(G) \leq i_t(S).$$

In addition, we prove the following structure for an mK_2-covered graph on n vertices that maximizes independent sets of size t.

Theorem 7. *Let G be an mK_2-covered graph on n vertices with $m \geq 2$. Then for all $t \geq 3$ there exists a graph $H_\ell = K_{1,n-2\ell-1} \cup \ell K_2$ for some $\ell \geq m-1$ such that $i_t(G) \leq i_t(H_\ell)$.*

This chapter is partially dedicated to determining which value of ℓ is optimal for given values of n, m, and t.

We conclude in Chapter 5 with a discuss of problems for future work. Some of these are tangential to the work in this dissertation, and others are extensions and generalizations of our results.

1.1 Notation

We end our introduction by providing some notation that will be used throughout the dissertation as well as some important terminology, beginning with some standard graph theory.

Given a graph G and a vertex $v \in V(G)$, we write $d(G)$ to denote the *degree* of v in G. The maximum and minimum degrees of G are denoted by $\Delta(G)$ and $\delta(G)$, respectively. Given two vertices v and u, we write $d(v,u)$ to denote the *distance* between v and u; that is, the smallest number of edges on a path from v to u in G. The *diameter* of G is the maximum value of $d(v,u)$ for all pairs v and u. We will only consider diameters in connected graphs. The *complement* of G, denoted $\overline{G}$ is the graph on the same vertex set as G whose edge set is given by $\binom{V(G)}{2} \setminus E(G)$. We say that G is r-partite if there exists a partition of the vertices into r parts such that each part is an independent set. We write $N(v)$ to denote the neighborhood of v; that is, the set of vertices u such that u is adjacent to v. $N[v] = N(v) \cup \{v\}$ is called the closed neighborhood of v. We write $N_2(v)$ to denote the set of vertices u such that $d(u, v) = 2$. To avoid ambiguity at times, we will write $N_G(v)$ to denote the neighborhood of v in G.

Given a pair of graphs G and H, we write $G \cup H$ to denote the disjoint union of these graphs. That is, we take the union of their vertex sets and edge sets. The disjoint union of k copies of G is written kG. The *join* of G and H, written $G + H$, is the graph obtained from $G \cup H$ by adding edges between u and v for all $u \in V(G)$ and $v \in V(H)$.

When discussing asymptotic results, we will make use of standard notation to relate various functions. In what follows, all constants are taken to be positive real numbers. In particular, given two functions $f(n)$ and $g(n)$, we write $f(n) = O(g(n))$ if there exist constants $C > 0$ and n_0 such that $f(n) \leq C \cdot g(n)$ for all $n \geq n_0$. Slightly stronger than this big-O notation is little-o notation. We write $f(n) = o(g(n))$ if for every constant $C > 0$ there exists n_0 such that $f(n) < C \cdot g(n)$ for all $n \geq n_0$. These will frequently be used when considering upper bounds on various quantities. Note that when there are multiple variables in play, we will write O_k and o_k to indicate that we are considering behavior as k tends to infinity.

Similarly, we define big-Omega and little-omega as follows. We write $f(n) = \Omega(g(n))$ if there exist constants $c > 0$ and n_0 such that $f(n) \geq c \cdot g(n)$ for all $n \geq n_0$. We write $f(n) = \omega(g(n))$ if for every constant $c > 0$ there exists n_0 such that $f(n) > c \cdot g(n)$ for all $n \geq n_0$. Just as big-O and little-o allow us to consider upper bounds on long term behaviors, big-Omega and little-omega allow us to consider lower bounds.

We end this discussion of asymptotic terminology by defining big-Theta notation. That is, we write $f(n) = \Theta(g(n))$ if $f(n) = O(g(n))$ and $f(n) = \Omega(g(n))$.

In the following chapters, we will be counting many substructures within graphs. Given a graph G, we write $s_r(G)$ to denote the number of copies of S_r in G. Similarly, we write $k_r(G)$ and $i_r(G)$ to denote the number of cliques and independent sets of size r respectively.

Additional notation and definitions will be provided as need for them arises.

Chapter 2

Generalized Saturation Problems

In this chapter we prove several minimization results dealing primarily with clique-saturation and star-saturation. Most of the attention in this chapter will be dedicated to $\mathrm{sat}_{K_r}(n, S_t)$, $\mathrm{sat}_{S_r}(n, S_t)$, and $\mathrm{sat}_{S_r}(n, K_t)$. In the final of these, we discuss the remaining challenge regarding $\mathrm{sat}_{S_2}(n, K_3)$. We also briefly consider the problem of counting paths in clique-saturated graphs and a more general problem for trees and cliques. Recall that $\mathrm{sat}_H(n, F)$ denotes the minimum number of copies of H among all F-saturated graphs on n vertices.

2.1 Counting Cliques in Star-Saturated Graphs

This section is concerned with understanding $\mathrm{sat}_{K_r}(n, S_t)$; that is, the minimum number of copies of K_r among S_t-saturated graphs on n vertices. In what follows, we will focus more specifically on finding conditions for which $\mathrm{sat}_{K_r}(n, S_t) = 0$. We begin this endeavor by stating a structural lemma regarding star-saturated graphs, which was observed by Kaszonyi and Tuza [22]. We include a short proof since it provides further insight into the structure of these graphs.

Lemma 1 (Kaszonyi and Tuza, 1986)**.** *Let G be an S_t-saturated graph on n vertices. Then the maximum degree of G is $t-1$, and all vertices of degree less than $t-1$ form* *a clique.*

Proof. Since G is S_t-saturated, it must be S_t-free. Thus the maximum degree of G is at most $t-1$. Adding any missing edge can increase the maximum degree by at most 1. Since G is S_t-saturated, the new edge must force the maximum degree to become t. Thus G has maximum degree exactly $t-1$.

To prove the second part of the statement, suppose that u and v are vertices in G with degrees $d(u) < t-1$ and $d(v) < t-1$. If u is not adjacent to v, then the addition of the edge uv will only increase their degrees, and the resulting graph will still have maximum degree less than t, a contradiction to G being S_t-saturated. Therefore all vertices of degree less than $t-1$ must form a clique in G. □

Since we are interested in determining when $\mathrm{sat}_{K_r}(n, S_t) = 0$, it is natural to begin with the case where $r = 3$ since the existence of a triangle-free, S_t-saturated graph on n vertices means that $\mathrm{sat}_{K_r}(n, S_t) = 0$ for all $r \geq 3$. The following result provides us with a necessary and sufficient condition for such graphs to exist. Note that for $t = 1$ we need $n \geq 2$, and for $t = 2$ we need $n \geq 3$ for S_t-saturated graphs to exist. Moreover, there is a unique S_2-saturated graph on n vertices for all $n \geq 3$. When n is even, the graph is a collection of disjoint edges. When n is odd, it is a collection of disjoint edges and an isolated vertex. S_1-saturated graphs are simply independent sets on at least 2 vertices. Since these cases are trivial, we state our result for $t \geq 3$.

Proposition 1. *Let $t \geq 3$. There exists an S_t-saturated graph on n vertices that is K_3-free if and only $n \geq 2t-2$.*

Proof. If n is even and $n \geq 2t-2$, then there exists a $(t-1)$-regular bipartite graph on n vertices that can be constructed in the following way. Begin by considering the complete bipartite graph $K_{\frac{n}{2},\frac{n}{2}}$ and delete $\frac{n}{2} - (t-1)$ perfect matchings. Such matchings exist by Hall's Theorem. The resulting graph is certainly S_t-saturated because it is $(t-1)$-regular. When n is odd, consider $G \cup K_1$ where G is a $(t-1)$-regular bipartite graph on $n-1$ vertices. The resulting graph is still S_t-saturated since the disjoint union does not create a copy of S_t, but the addition of any missing edge must involve at least one vertex of degree $t-1$. $G \cup K_1$ is also bipartite, hence K_3-free.

For the other direction, let G be S_t-saturated and K_3-free. If G has no vertices of degree $t-1$, then by Lemma 1 G is complete and contains triangles. This means there is a vertex $v \in G$ with degree $t-1$. For any two vertices $x, y \in N(v)$, at least one of them must have degree $t-1$. Otherwise x is adjacent to y and $\{v, x, y\}$ form a copy of K_3. Without loss of generality, the degree of x is $t-1$. Since G is K_3-free, $N(v) \cap N(x) = \varnothing$ and $N(x)$ contains $t-2$ vertices outside of $N[v]$. Therefore $n \geq 2t-2$. □

Since triangle-free graphs are K_r-free for all $r \geq 3$, the following is immediate.

Corollary 1. *For all $r \geq 3$ and all $t \geq 3$,*

$$\mathrm{sat}_{K_r}(n, S_t) = 0$$

for $n \geq 2t-2$.

Proposition 1 gives us a cutoff for the values of n which require S_t-saturated graphs to contain at least one copy of K_3. It tells us, for instance, that any S_5-saturated graph that is K_3-free must have at least 8 vertices. Although there does not exist an S_5-saturated graph on 6 vertices that is K_3-free, there does exist an S_5-saturated graph on 6 vertices that is K_4-free. The graph in Figure 2.1 is an example.

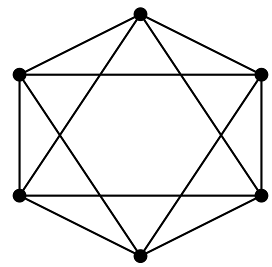

Figure 2.1: An example of a K_4-free, S_5-saturated graph on fewer than 8 vertices

As stated in the beginning of this section, we are interested in determining the values of n for fixed r and t such that $\mathrm{sat}_{K_r}(n, S_t) = 0$. We will provide a partial solution to this problem. Shifting to K_{r+1}-free graphs, we will focus especially on the existence of S_t-saturated graphs that are r-partite. Again due to the triviality of the cases where t is equal to 1 or 2, we will only consider $t \geq 3$. We remind the reader that r-partite graphs are always K_{r+1}-free; however K_{r+1}-free graphs are not necessarily r-partite.

Before stating our bounds on n, we state a theorem of Hoffman and Rodger, and we provide a construction that will aid us in our proof [21].

Theorem 8 (Hoffman and Rodger, 1992)**.** *Given a complete multipartite graph K, $\chi'(K) = \Delta(K)$ if and only if it is not overfull. Here χ' denotes the chromatic index of K, and we say that a graph G is overfull if*

$$|E(G)| > \Delta(G) \left\lfloor \frac{|V(G)|}{2} \right\rfloor .$$

In particular, the complete r-partite graph on ar vertices, $K_{a,\ldots,a}$, is overfull if and only if ar is odd.

Proposition 2. *Let a, r be positive integers. If ar is even, then there exists a k-regular spanning subgraph of the r-partite graph $K_{a,\ldots,a}$ for all $k \leq a(r-1)$.*

Proof. Consider the complete r-partite graph $K_{a,\ldots,a}$. If ar is even, then $K_{a,\ldots,a}$ is not overfull as

$$|E(K_{a,\ldots,a})| = a(r-1)\frac{ar}{2} = \Delta(K_{a,\ldots,a})\left\lfloor \frac{V(K_{a\ldots,a})}{2} \right\rfloor.$$

By Theorem 8, the chromatic index of this graph is equal to the maximum degree. That is, we can find a proper edge coloring using $a(r-1)$ colors. In particular, the color classes form a 1-factorization of $K_{a,\ldots,a}$. We delete perfect matchings until we are left with a k-regular subgraph for any $k \leq a(r-1)$. $\square$

We are now ready to state our first result regarding the existence of S_t-saturated, r-partite graphs.

Theorem 5. *Let $r \geq 3$ and $t \geq 3$ be fixed. There exists an n-vertex, r-partite, S_t-saturated graph if*

$$n \geq \max\left(t+1, \min_{0\leq c\leq r-2}\left\{(r-c)\left\lceil \frac{t-1}{r-c-1} \right\rceil + r - c\right\}\right). \tag{2.1}$$

Proof. Suppose Inequality 2.1 holds for some $0 \leq c \leq r-2$. Let a, b be non-negative integers such that $n = a(r-c) + b$ with $b < r - c$. That is, $a = \lfloor \frac{n}{r-c} \rfloor$. In addition, let k, d be non-negative integers such that $b = kt + d$ with $d < t$. We will consider two cases and exhibit an r-partite, S_t-saturated graph on n vertices in each case.

Begin by supposing $a(r-c)$ is even. By rewriting Inequality 2.1, focusing on the second expression on the right hand side, we see that $t - 1 \leq a(r - c - 1)$. By Proposition 2, there exists a $(t-1)$-regular, $(r-c)$-partite graph G on $a(r-c)$ vertices. Taking the disjoint union $G \cup kK_t \cup K_d$ yields an r-partite, S_t-saturated graph on n vertices.

Similarly, when $a(r-c)$ is odd, there exists a $(t-1)$-regular, $(r-c)$-partite graph G on $(a-1)(r-c)$ vertices. Thus $G \cup \ell K_t \cup K_m$ where $b + r - c = \ell t + m$ with $m < t$

provides an r-partite, S_t-saturated graph on n vertices. $\square$

With regards to the lower bounds on n, we note that if $n < t+1$, then there is no S_t-saturated graph on n vertices, and hence none that is r-partite. We now provide a necessary condition for the existence of S_t-saturated, r-partite graphs that is related to the second bound in Theorem 5.

Proposition 3. *For all $r \geq 3$ and $t \geq 3$, if there exists a graph G on n vertices that is S_t-saturated and r-partite, then $n \geq \frac{r(t-1)}{r-1}$.*

Proof. Suppose there exists an S_t-saturated, r-partite graph G on n vertices with $n < \frac{r(t-1)}{r-1}$. As remarked before stating this proposition, it must be the case that $n \geq t+1$. Let a, b be non-negative integers such that $n = ar + b$ with $b < r$. Since G is r-partite, there exists a partition $P_1, \ldots, P_r$ of the vertices of G with $|P_i| \leq |P_{i+1}|$ for all $1 \leq i \leq r-1$ such that no two vertices in a given P_i are adjacent. We have two cases to consider.

Begin by supposing that $|P_1| < |P_r|$. If $a = 0$, then G has b vertices with $b < r$. Since G is S_t-saturated, we have $b \geq t+1$, and, by assumption, $b < \frac{r(t-1)}{r-1}$. It follows that $t+1 < \frac{r(t-1)}{r-1}$. Rearranging, we obtain the following equivalent inequalities.

$$t+1 < \frac{r(t-1)}{r-1}$$

$$(t+1)(r-1) < r(t-1)$$

$$rt + r - t - 1 < rt - r$$

$$2r < t+1$$

This is a contradiction since $b < r$ and $b \geq t+1$. Thus it must be the case that $a \geq 1$.

Now, since $|P_1| < |P_r|$, it follows that

$$|P_r| \geq \left\lceil \frac{n}{r} \right\rceil = a + 1.$$

For any vertex $u \in P_r$, we have $d(u) \leq n - a - 1$. Since each P_i is an independent set, it can contain at most one vertex of degree less than $t - 1$ by Lemma 1. In particular, since $a \geq 1$, P_r must contain at least one vertex of degree exactly $t - 1$. Thus $n - a - 1 \geq t - 1$. Since $n < \frac{r(t-1)}{r-1}$ by assumption, we have the following equivalent inequalities:

$$\begin{aligned}
t + a &< \frac{r(t-1)}{r-1} \\
(t + a)(r - 1) &< r(t - 1) \\
rt + ar - t - a &< rt - r \\
a(r - 1) + r &< t \\
a(r - 1) + r - 1 &< a(r - 1) + b - 1.
\end{aligned}$$

Similarly, noting that $n = ar + b$, we having the following equivalences:

$$\begin{aligned}
n - a - 1 &\geq t - 1 \\
ar + b - a - 1 &\geq t - 1 \\
a(r - 1) + b &\geq t.
\end{aligned}$$

Taken all together, we see that $a(r - 1) + b \geq t > a(r - 1) + r$, contradicting the assumption that $b < r$. This completes the case where $|P_1| < |P_r|$.

Finally, we consider the case where each of the r parts in our partition have equal size. It follows that $b = 0$ and $n = ar$. Every vertex in G has degree at most $a(r-1)$.

Since $n < \frac{r(t-1)}{r-1}$, we have that $a(r-1) < t-1$. This means that if we add an edge to G, the maximum degree in the resulting graph will be at most $t-1$, a contradiction to the assumption that G is S_t-saturated. This completes the proof. $\square$

In addition to bridging the gap between these bounds, we would like to know which value of c for given r and t minimizes the lower bound on n in Theorem 5. If our bound did not include ceilings, this would be a straightforward computation as demonstrated in the lemma below. However, finding a general solution using the bound in the theorem is more complicated. Although we do not have a general solution, we determine which value of c provides the smallest bound on n in Theorem 5 for the existence of r-partite, S_t-saturated graphs in two special cases. These cases rely on the more straightforward situation which we now address.

Lemma 2. *Let $r \geq 3$ and $t \geq 3$ be fixed. Then*

$$n_2(c) = (r-c)\left(\frac{t-1}{r-c-1}\right) + r - c$$

is minimized on the interval $(-\infty, r-2]$ when $c = r-1-\sqrt{t-1}$.

Proof. Taking the derivative of n_2 with respect to c, we obtain

$$n_2'(c) = \frac{t-1}{(r-c-1)^2} - 1.$$

Setting this equal to 0, we have $c = r-1 \pm \sqrt{t-1}$. Since $c \leq r-2$ and

$$n_2''(r-1-\sqrt{t-1}) > 0,$$

our function is minimized at $c = r-1-\sqrt{t-1}$. $\square$

Proposition 4. *For all $r \geq 3$, if $t \leq (r-1)^2+1$ and $t = k^2+1$ for some integer k, then $r-1-\sqrt{t-1}$ is the minimizing c-value in Theorem 5. The corresponding lower bound on n is $n \geq (k+1)^2 = t + 2\sqrt{t-1}$.*

Proof. Let

$$n_1(c) = (r-c)\left\lceil \frac{t-1}{r-c-1} \right\rceil + r - c$$

and

$$n_2(c) = (r-c)\left(\frac{t-1}{r-c-1}\right) + r - c.$$

That is, $n_1(c)$ provides the lower bounds on n for the $(r-c)$-partite construction in Theorem 5. By Lemma 2, $n_2(c)$ is minimized when $c = r-1-\sqrt{t-1}$ on the interval $[0, r-2]$. Since $n_1(c) \geq n_2(c)$ for all c, it follows that $n_1(c)$ is minimized at $c = r-1-\sqrt{t-1}$ if $n_1(r-1-\sqrt{t-1}) = n_2(r-1-\sqrt{t-1})$. This is true precisely when $t = k^2+1$ for some integer k. We finally note that in this setting $n_1(r-1-\sqrt{t-1}) = t + 2\sqrt{t-1}$. $\square$

Proposition 5. *For all $r \geq 3$, if $t \geq (r-1)^2+1$ and $r-1$ divides $t-1$, then 0 is the minimizing c-value in Theorem 5. The corresponding lower bound on n is $n \geq r\left(\frac{t-1}{r-1}\right) + r$.*

Proof. Define $n_1(c)$ and $n_2(c)$ as in the previous proof. Since $t \geq (r-1)^2+1$, we have that $r-1-\sqrt{t-1} \leq 0$. Thus $n_2(c)$ is increasing on the interval $[0, r-2]$. It follows that $n_1(c)$ is minimized at $c = 0$, if $n_1(0) = n_2(0)$. This is true precisely when $\frac{t-1}{r-1}$ is an integer. That is, when $r-1$ divides $t-1$. We finally note that in this setting $n_1(0) = r\left(\frac{t-1}{r-1}\right) + r$. $\square$

Note that when $t \leq (r-1)^2+1$, Proposition 4 shows that the minimizing c-value for Theorem 5 is not always 0. That is, perhaps somewhat surprisingly, starting with fewer parts than what we're permitted leads to a smaller lower bound on n for infinitely many choices of r and t.

In what remains of this section, we provide necessary conditions for the existence of S_t-saturated graphs on n vertices that are K_{r+1}-free. We are continuing to consider cliques on $r+1$ vertices rather than r for ease of comparison with our bounds on n that guaranteed the existence of S_t-saturated, r-partite graphs.

Lemma 3. *Let G be an S_t-saturated graph on n vertices with m vertices of degree less than $t-1$ and no copy of K_{r+1}. Then*

$$n \geq \frac{r}{r-1}\left(\frac{t-1}{2} + \sqrt{\left(\frac{t-1}{2}\right)^2 - \frac{m(r-1)(t-m)}{r}}\right).$$

Proof. Since G is K_{r+1}-free, we have by Turán's Theorem that

$$e(G) \leq \frac{r-1}{r} \cdot \frac{n^2}{2}.$$

Using the structure of S_t-saturated graphs described in Lemma 1 at the beginning of this section, we also have that

$$e(G) \geq \frac{1}{2}(n-m)(t-1) + \binom{m}{2}$$

for some value of m, where m is the number of vertices of degree less than $t-1$. These two bounds together provide the desired inequality. □

Considering the case where $t \geq 2r$, we provide a general bound independent of the existence of S_t-saturated, K_r-free graphs with specific m. We also demonstrate when these bounds are at least $t+1$ since that is the trivial necessary condition for the existence of an S_t-saturated graph, and we want to be sure that the bound we provide has value.

Proposition 6. *Let G be an S_t-saturated graph on n vertices with no copy of K_{r+1} where $t \geq 2r$ and $r \geq 2$. Then*

$$n \geq \frac{r}{r-1}\left(\frac{t-1}{2} + \sqrt{\left(\frac{t-1}{2}\right)^2 - (r-1)(t-r)}\right).$$

Furthermore, this bound is at least $t+1$ whenever $t \geq r(r+1)-1$.

Proof. We begin by noting that the bound in Lemma 3 is minimized at $m = r$. This is because $m \leq r$ by Lemma 1 and because the bound in Lemma 3 is decreasing in m for $m < t/2$. Plugging this value in for m yields the desired bound. Letting $n(t)$ denote this bound, we observe that when $t = r(r+1)-1$, we have $n(t) = r(r+1) = t+1$. In addition,

$$n'(t) = \frac{r}{r-1}\left(\frac{1}{2} + \frac{1}{2}\frac{t-2r+1}{\sqrt{\frac{1}{4}(t-1)^2 - (r-1)(t-r)}}\right) \geq \frac{r}{r-1} > 1$$

for all $t \geq 2r$. By the racetrack principle, $n(t) \geq t+1$ whenever $t \geq r(r+1)-1$ as desired. □

Finally, we observe that the bound in Lemma 3 is asymptotically equivalent to the bound in Proposition 3. That is, our lower bound on n needed for the existence of S_t-saturated, r-partite graphs cannot be lowered too much when we loosen our restriction to K_{r+1}-free graphs.

Corollary 2. *Let r be fixed. If there exists an S_t-saturated graph on n vertices with* *no copy of K_{r+1} where $t \geq 2r$ and $r \geq 3$, then*

$$n \geq \frac{r(t-1)}{r-1} - O_t(1).$$

Proof. Note that the number of vertices, that is m, of degree less than $t-1$, in a K_{r+1}-free graph is at most r since these vertices form a clique in S_t-saturated graphs by Lemma 1. In particular, $m < \frac{t}{2}$. By Proposition 6, the following holds as $t \geq 2r$.

$$\begin{aligned}
n &\geq \frac{r}{r-1}\left(\frac{t-1}{2} + \sqrt{\left(\frac{t-1}{2}\right)^2 - \frac{m(r-1)(t-m)}{r}}\right) \\
&\geq \frac{r}{r-1}\left(\frac{t-1}{2} + \frac{t-1}{2} - \frac{m(r-1)t}{r(t-1)}\right) \\
&= \frac{r}{r-1}\left(t-1-O_t(1)\right) \\
&= \frac{r(t-1)}{r-1} - O_t(1).
\end{aligned}$$

We end by noting that the second inequality above can be seen from the fact that $\frac{3r}{2r+1}(2t-1) \leq t^2$ whenever $t \geq 2r$. $\square$

2.2 Counting Stars in Star-Saturated Graphs

2.2.1 General results for stars in star-saturated graphs

We now turn to counting copies of stars S_r in S_t-saturated graphs. We write $s_r(G)$ to denote the number of copies of S_r in a given graph G where S_r is the complete bipartite graph $K_{1,r}$. Note that if $t \leq r$, then $\mathrm{sat}_{S_r}(n, S_t) = 0$ trivially as an S_t-saturated graph must be S_t-free. That is, any S_t-saturated graph has no vertex of degree at least t, and hence none of degree at least r. Our focus is therefore on the situation where $t > r$, and we will consider the cases where t is odd and even separately. As in the previous section, we ignore the case where $t = 1$ since the only S_1-saturated graph on n vertices is an independent set.

As was the case in the previous section, we will give a construction that is especially useful for us. In particular, this construction will provide us with a candidate for an S_t-saturated graph with $\mathrm{sat}_{S_r}(n, S_t)$ copies of S_r. We will build this graph in two steps.

First, given $a < b$, we let $R_{a,b}$ denote a graph on b vertices that is as close to a-regular as possible. More specifically, when ab is even, $R_{a,b}$ is an a-regular graph on b vertices. When ab is odd, $R_{a,b}$ has one vertex of degree $a-1$ and $b-1$ vertices of degree a.

Lemma 4. *An $R_{a,b}$ exists if and only if $b \geq a+1$.*

Proof. It is well known that a-regular graphs exist on b vertices if and only if $b \geq a+1$ and ab is even. That is, a regular $R_{a,b}$ exists when ab is even. We now proceed to the irregular case. When $b \geq a+1$ and ab is odd, we can obtain a graph on b vertices that is a-regular with the exception of one vertex of degree $a-1$ in the following manner. Label the vertices $0, 1, \ldots, b-1$. Add edges between vertices with labels i and j if and only if

$$|i-j| \leq \frac{a-1}{2} \mod b.$$

This gives us an $(a-1)$-regular graph. Finally for $1 \leq i \leq \frac{b-1}{2}$, add an edge between the vertices labeled i and $i + \frac{b-1}{2} \mod b$. We aren't reusing any edges at this point since $b \geq a+1$, and the degree of every vertex increases by 1 when we add these edges, except for the vertex labeled 0. Its degree remains $a-1$. Thus every vertex has degree a with the exception of one vertex, namely the one labeled 0, of degree $a-1$, and hence $R_{a,b}$ exists. Necessity of the inequality $b \geq a+1$ is clear because a graph can not have any vertex of degree a if there are fewer than $a+1$ vertices. $\square$

Utilizing the regularity of $R_{a,b}$, or near regularity in the case where ab is odd, we define our candidate extremal graphs for $\mathrm{sat}_{S_r}(n, S_t)$. They are essentially the disjoint union of a clique and a regular graph. To be precise, for $m \leq t-1$ we let

$$\mathrm{KR}_{t,n}(m) = \begin{cases} K_m \cup R_{t-1,n-m} & \text{if } (t-1)(n-m) \text{ is even} \\ K_m \cup R_{t-1,n-m} + e & \text{if } (t-1)(n-m) \text{ is odd,} \end{cases}$$

where in the second case e is an edge between the vertex of degree $t-2$ in $R_{t-1,n-m}$ and an arbitrary vertex of the clique K_m. Note that $\mathrm{KR}_{t,n}(m)$ exists provided $R_{t-1,n-m}$ exists, which by Lemma 4 is the case precisely when $n - m \geq t$. Having constructed this graph $\mathrm{KR}_{t,n}(m)$, we now immediately state and prove our theorem showing this graph does in fact minimize the number of copies of small stars S_r among S_t-saturated graphs on n vertices for some choice of m.

Theorem 4. *For all $n \geq 2t-1$ with $t \geq 2$ and $r < t$,*

$$\mathrm{sat}_{S_r}(n, S_t) = \min_{0 \leq m \leq t-1} s_r(\mathrm{KR}_{t,n}(m)).$$

Note also that

$$s_r(\mathrm{KR}_{t,n}(m)) = \begin{cases} m\binom{m-1}{r} + (n-m)\binom{t-1}{r} & \text{if } (t-1)(n-m) \text{ is even} \\ m\binom{m-1}{r} + (n-m)\binom{t-1}{r} + \binom{m-1}{r-1} & \text{if } (t-1)(n-m) \text{ is odd.} \end{cases}$$

Proof. We begin by considering the case where t is odd. By Lemma 1, if G is S_t-saturated, then G contains a clique A containing all of the vertices with degree smaller than $t-1$. Let A have size m, and let $B = V(G) \setminus A$. We have two cases to consider. If G has no edges between A and B, then G contains exactly $m\binom{m-1}{r} + (n-m)\binom{t-1}{r}$ copies of S_r. The first term counts stars centered in A, and the second term counts stars centered in B. Since $t-1$ is even, there exists a $(t-1)$-regular graph $R_{t-1,n-m}$ on $n-m$ vertices for all $n-m \geq t$. This inequality holds since we assume $n \geq 2t-1$ and $m \leq t-1$. Thus an S_t-saturated graph with precisely the above count is given by $\mathrm{KR}_{t,n}(m)$.

Now, if there exist vertices $u \in A$ and $v \in B$ such that u is adjacent to v, then our graph contains all of the previously counted stars, along with at least $\binom{m-1}{r-1}$ stars centered at u containing the edge uv. This means than an S_t-saturated graph G with $|A| = m$ and any such edge must have at least as many copies of S_r as $\mathrm{KR}_{t,n}(m)$. Thus an S_t-saturated graph with minimum number of copies of S_r is given by some $\mathrm{KR}_{t,n}(m)$ for some $m \leq t-1$.

We now consider the case where t is even. For a given $m \leq t-1$, if $n-m$ is even, we can find a $(t-1)$-regular graph on $n-m$ vertices, and the argument is the same as before. That is, among S_t-saturated graphs with m vertices of degree less than $t-1$, $\mathrm{KR}_{t,n}(m)$ is a minimal example with respect to copies of S_r. When $n-m$ is odd, we can construct a graph $R_{t-1,n-m}$ on $n-m$ vertices that is $(t-1)$-regular with the exception of one vertex v of degree $t-2$. Thus we can construct an S_t-saturated graph $\mathrm{KR}_{t,n}(m)$ by taking the disjoint union of K_m with $R_{t-1,n-m}$ and adding an edge from v to an arbitrary vertex in the clique K_m. Since every S_t-saturated graph with m vertices of degree less than $t-1$ has at least $m\binom{m-1}{r} + (n-m)\binom{t-1}{r}$ many copies of S_r and there is no $(t-1)$-regular graph on $n-m$ vertices, this adds the fewest possible copies of S_r. That is, we must have at least one edge between A and B, introducing $\binom{m-1}{r-1}$ copies of S_r. Therefore the generalized saturation number $\mathrm{sat}_{S_r}(n, S_t)$ is obtained by minimizing $s_r(\mathrm{KR}_{t,n}(m))$ over all values of m between the two scenarios. $\square$

We note that the above theorem does not hold when $n < 2t-1$. This is because it is possible for an S_t-saturated graph to have fewer than t vertices of degree $t-1$. This means we can't consider the disjoint union of a small clique and a $(t-1)$-regular graph. In this setting, it turns out than an optimal graph does not need to have that structure. For example, a quick check shows that the graph in Figure 2.2 minimizes the number of copies of S_3 among S_5-saturated graphs on 6 vertices and has only 4 vertices of degree 4.

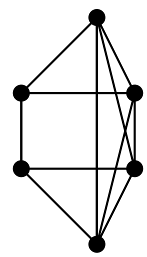

Figure 2.2: An S_5-saturated graph that does not satisfy the criteria in Theorem 4

In light of the fact that Theorem 4 is concerned with a minimum count over choices of m, we now introduce additional notation to more concisely discuss that optimal choice. Given n, r, and t with $n \geq \max\{2t-1, t+1\}$, we define

$$m_0(n,r,t) := \underset{m}{\operatorname{argmin}}\ s_r(\mathrm{KR}_{t,n}(m)).$$

That is, $m_0(n,r,t)$ is the value of m for which $\mathrm{KR}_{t,n}(m)$ attains the generalized saturation number $\mathrm{sat}_{S_r}(n, S_t)$. If the generalized saturation number is achieved for multiple values of m, we take the smallest one for definiteness. We shall see later in this section that there are instances where $\mathrm{KR}_{t,n}(m)$ is optimal for multiple choices of m. Our goal now is to identify the value of $m_0(n,r,t)$ for given values of n, r, and t. Kaszonyi and Tuza [22] showed that the number of edges in an S_t-saturated graph is minimized when $m = \lfloor \frac{t}{2} \rfloor$ or $\lfloor \frac{t+1}{2} \rfloor$, answering our question for $r = 1$. This solution does not hold for all $r \geq 1$ though. Rather $m_0(n,r,t)$ depends on both r and t. The value of n does not matter for $n \geq 2t-1$ when t is odd since we are simply getting more vertices of degree $t-1$ as we increase n. The particular value of n does matter when $t+1 \leq n < 2t-1$ though. This is because there may not exist a $(t-1)$-regular graph on $n-m$ as demonstrated by the example in Figure 2.2.

Although we are unable to provide a closed form for $m_0(n, r, t)$ for arbitrary pairs of r and t, there is more we can say about the optimal choice, or in some cases, choices. We will focus on the case where t is odd since we are guaranteed that $\mathrm{KR}_{t,n}(m)$ is the disjoint union of K_m and a $(t-1)$-regular graph $R_{t-1,n-m}$ We begin with the observation that

$$\mathcal{D}(m) := s_r(\mathrm{KR}_{t,n}(m+1)) - s_r(\mathrm{KR}_{t,n}(m)) = (r+1)\binom{m}{r} - \binom{t-1}{r}. \tag{2.2}$$

That is, $\mathcal{D}(m)$ denotes the change in the number of copies of S_r as we increase m, the number of vertices of degree less than $t-1$, by 1 from m to $m+1$. This is done by replacing K_m with K_{m+1} and $R_{t-1,n-m}$ with $R_{t-1,n-m-1}$. Note that in this difference, $(r+1)\binom{m}{r} = m\binom{m-1}{r-1} + \binom{m}{r}$ is the number of new stars centered in the clique K_{m+1}, and $\binom{t-1}{r}$ is the number of stars that were lost from the $(t-1)$-regular portion. The following is an immediate consequence of this observation.

Lemma 5. *If t is odd, then for all $n \geq 2t-1$ with $t > r$, we have $m_0(n, r, t) \geq r$. In particular, $m_0(n, t-1, t) = t-1$.*

Proof. If $m < r$, then $\binom{m}{r} = 0$ and so $\mathcal{D}(m) < 0$. That is,

$$s_r(\mathrm{KR}_{t,n}(m+1)) < s_r(\mathrm{KR}_{t,n}(m))$$

whenever $m < r$. Therefore the minimum must be attained when $m \geq r$. To prove the second statement, we observe that when $r = t-1$, we have $\mathcal{D}(t-1) = t-1 > 0$.

Since $\mathcal{D}$ is an increasing function, it follows that $m_0(n, t-1, t) = t-1$. $\square$

For fixed t and r, we extend the definition of $\mathcal{D}$ in Equation 2.2 to all real numbers as follows:

$$\mathcal{D}(x) := (r+1)\binom{x}{r} - \binom{t-1}{r}.$$

Now, $\binom{x}{r}$ is convex and increasing in x for all $x \geq r-1$. It follows that those properties hold for $\mathcal{D}(x)$ as well. Thus $\mathcal{D}(x)$ has a unique root on the interval $(r-1, \infty)$ as $\mathcal{D}(r-1) = -\binom{t-1}{r}$. Our next theorem takes advantage of this structure.

Theorem 9. *For fixed $n, r,$ and t with $n \geq 2t-1$ and $t > r$, let $\overline{x}$ denote the unique root of $\mathcal{D}(x)$ in the interval $(r-1, \infty)$. Then*

$$m_0(n, r, t) = \lceil \overline{x} \rceil.$$

Furthermore, when $\overline{x} \notin \mathbb{Z}$, this is the unique minimizing value of m in Theorem 4. When $\overline{x} \in \mathbb{Z}$, both $\overline{x}$ and $\overline{x}+1$ simultaneously minimize the number of copies of S_r among S_t-saturated graphs on n vertices.

Proof. As stated previously, $\mathcal{D}(x)$ has a unique root $\overline{x}$ and is increasing on the interval $(r-1, \infty)$. Let $m \geq r$ be an integer. If $m < \overline{x}$, then $\mathcal{D}(m) < 0$ and

$$s_r(\mathrm{KR}_{t,n}(m+1)) < s_r(\mathrm{KR}_{t,n}(m)).$$

If $m > \overline{x}$, then $\mathcal{D}(m) > 0$ and

$$s_r(\mathrm{KR}_{t,n}(m+1)) > s_r(\mathrm{KR}_{t,n}(m)).$$

If $\overline{x} \notin \mathbb{Z}$, then it follows that $s_r(\mathrm{KR}_{t,n}(m))$ is minimized when m is the first integer larger than $\overline{x}$, namely $\lceil \overline{x} \rceil$, and this choice of m is unique.

On the other hand, if $\overline{x} \in \mathbb{Z}$, then $\mathcal{D}(\overline{x}) = 0$ and

$$s_r(\mathrm{KR}_{t,n}(\overline{x}+1)) = s_r(\mathrm{KR}_{t,n}(\overline{x})).$$

Therefore $s_r(\mathrm{KR}_{t,n}(m))$ is minimized by $\overline{x}$ and $\overline{x}+1$ simultaneously. These are the only optimal choices for m as $\mathcal{D}(\overline{x}-1) < 0$ and $\mathcal{D}(\overline{x}+1) > 0$. □

Before proceeding to asymptotic results, we conclude the discussion on general results for $\mathrm{sat}_{S_r}(n, S_t)$ by providing two lower bounds on the value of $m_0(n,r,t)$ and giving a more precise answer for the case where $r = 2$. The second lower bound will be of additional interest in the following section on asymptotic results. Before stating these results, we prove a simple lemma regarding binomial coefficients.

Lemma 6. *If $a \geq c \geq 2$ and $b > 1$ where $a, b \in \mathbb{R}$ and $c \in \mathbb{Z}$, then*

$$b^c \binom{\lfloor a/b \rfloor}{c} < \binom{a}{c}.$$

Proof. Note that if $\lfloor a/b \rfloor < c$, then the inequality holds trivially as the left hand side of our inequality is equal to 0, and the right hand side is positive. Suppose then that $\lfloor a/b \rfloor \geq c$. Then we have the following

$$\begin{aligned} b^c(\lfloor a/b \rfloor)_c &= b^c(\lfloor a/b \rfloor)(\lfloor a/b \rfloor - 1)\cdots(\lfloor a/b \rfloor - (c-1)) \\ &\leq a(a-b)(a-2b)\cdots(a-b(c-1)) \end{aligned}$$

and

$$(a)_c = a(a-1)(a-2)\cdots(a-(c-1)).$$

Since $0 < a - bk < a - k$ for all $1 \leq k \leq c - 1$ and $b > 1$, the desired inequality holds. $\square$

With this lemma in hand, we are ready to state and prove our next lower bounds on the optimal choice for m in terms of minimizing copies of S_r among S_t-saturated graphs.

Corollary 3. *If $t \geq 3$ is odd and $r \geq 2$ with $t > r$, then for all $n \geq 2t - 1$, we have*

$$m_0(n, r, t) \geq \frac{t+1}{2}.$$

Proof. Since $r \geq 2$, we know that $2^r > r + 1$. Let $m \leq \frac{t-1}{2}$ be an integer. Applying Lemma 6 with $a = t - 1, b = 2$, and $c = r$, we obtain the following

$$(r+1)\binom{m}{r} < 2^r\binom{m}{r} \leq 2^r\binom{\frac{t-1}{2}}{r} < \binom{t-1}{r}.$$

Thus

$$\mathcal{D}(m) = (r+1)\binom{m}{r} - \binom{t-1}{r} < 0.$$

This means that $\mathrm{sat}_{S_r}(n, S_t)$ is not attained by $\mathrm{KR}_{t,n}(m)$, and $m_0(n, r, t) > \frac{t-1}{2}$. $\square$

Corollary 4. *If $t \geq 3$ is odd and $r \geq 2$ with $t > r$, then for all $n \geq 2t - 1$, we have*

$$m_0(n, r, t) > \frac{t-1}{(r+1)^{1/r}}.$$

Proof. Note that

$$\mathcal{D}\left(\left\lfloor \frac{t-1}{(r+1)^{1/r}} \right\rfloor\right) = (r+1)\binom{\left\lfloor \frac{t-1}{(r+1)^{1/r}} \right\rfloor}{r} - \binom{t-1}{r}.$$

Applying Lemma 6 with $a = t-1$, $b = (r+1)^{1/r}$, and $c = r$, we see that this quantity is strictly less than 0. Since $\frac{t-1}{(r+1)^{1/r}}$ is not an integer for all $r \geq 2$, it must be the case that $m_0(n, r, t) > \frac{t-1}{(r+1)^{1/r}}$. □

This is all we will say with regards to the value of m_0 for arbitrary r and t. However, when $r = 2$, finding $m_0(n, r, t)$ simply amounts to solving a quadratic equation and applying Theorem 9. The following is thus immediate.

Proposition 7. *For all $t \geq 3$ and $n \geq 2t - 1$, the value of $\overline{x}$ as in Theorem 9 for $r = 2$ is given by*

$$\overline{x} = \frac{1}{2} + \frac{1}{6}\sqrt{12t^2 - 36t + 33}.$$

We can say a little more when $r = 2$.

Proposition 8. *There are two optimal choices for m that simultaneously minimize the number of copies of S_2 among S_t-saturated graphs if and only if t is given by the following where $i \geq 0$ is some non-negative integer*

$$t(i) = \frac{1}{4}\left((1+\sqrt{3})(2+\sqrt{3})^i - (\sqrt{3}-1)(2-\sqrt{3})^i - 2\right) + 2.$$

Proof. By Theorem 9 and Proposition 7, there are two optimal choices for m precisely when $m = \frac{1}{2} + \frac{1}{6}\sqrt{12t^2 - 36t + 33}$ is an integer. This is the case when we can write $\sqrt{12t^2 - 36t + 33}$ in the form $6k + 3$ where k is an integer. Equivalently, we need $(t-1)(t-2) = 3(k+1)k$. Now, the second member of the Diophantine pair (x, y) that satisfies $3(x^2 + x) = y^2 + y$ is given by $y = y(i)$ where $y(i)$ satisfies the recurrence (see OEIS sequence A001571 [32])

$$y(i) = 4y(i-1) - y(i-2) + 1 \text{ with } y(0) = 0 \text{ and } y(1) = 2.$$

Solving the linear recurrence, we find that

$$y(i) = \frac{1}{4}\left((1+\sqrt{3})(2+\sqrt{3})^i - (\sqrt{3}-1)(2-\sqrt{3})^i - 2\right).$$

Therefore, since we are interested in $y = t - 2$, we see that the values of t for which there are two optimal choices of m are given by $t(i) = y(i) + 2$. □

2.2.2 Asymptotic results for stars in star-saturated graphs

In addition to general questions regarding the number of small stars in star-saturated graphs, we can address asymptotic questions. Here we focus on the case where t is odd and $n \geq 2t - 1$ for convenience. As noted earlier, this setting guarantees that $\mathrm{KR}_{t,n}(m)$ exists as the disjoint union of a clique K_m and a $(t-1)$-regular graph on the remaining $n - m$ vertices. Utilizing our general results from the previous section, we immediately proceed to our asymptotic results.

Theorem 10. *Let $r = o(\sqrt{t})$ with t odd and $r \geq 2$. Then for all $n \geq 2t - 1$, we have*

$$m_0(n,r,t) = (1 + o_t(1))\frac{t-1}{(r+1)^{1/r}}.$$

Proof. As in Theorem 9, we need to solve for the value of x such that

$$(r+1)\binom{x}{r} = \binom{t-1}{r}. \tag{2.3}$$

Let $\overline{x}$ be the unique solution to Equation 2.3 on the interval $(r-1, \infty)$. Observe that in order for this equality to hold, $\overline{x}$ must be less than $t - 1$. By Corollary 3, we know that $\overline{x} \geq \lfloor t/2 \rfloor$ and so $\overline{x} = \Theta(t)$. It is well known that when $k = o(\sqrt{n})$,

$$\binom{n}{k} = (1 + o_n(1))\frac{n^k}{k!}.$$

Since $r = o(\sqrt{t})$, $\overline{x} = \Theta(t)$, and $n \geq 2t - 1$, we can apply this to our equality and obtain the following.

$$(r+1)(1 + o_{\overline{x}}(1))\overline{x}^r = (1 + o_t(1))(t-1)^r.$$

Thus, since $\overline{x} = \Theta(t)$,

$$\overline{x} = (1 + o_t(1))\frac{t-1}{(r+1)^{1/r}}.$$

Applying Theorem 9 gives us the desired result. □

With slightly less precision than our previous theorem, we consider the more general case where $r = o(t)$. Before stating our theorem, we note the following useful result. A short proof can be found in [10]. See also [26] for a discussion of the result.

Lemma 7. *If $k = o(n)$, then $\log \binom{n}{k} = (1 + o_n(1))k \log \frac{n}{k}$.*

We are now ready to state and prove our final result concerning $\mathrm{sat}_{S_r}(n, S_t)$.

Theorem 11. *Let $r = o(t)$ with t odd and $r \geq 2$. Then for all $n \geq 2t - 1$, we have that*

$$m_0(n, r, t) = \left(\frac{t-1}{(r+1)^{1/r}}\right)^{1+o_t(1)}.$$

Proof. Let $\overline{x}$ be the unique solution to Equation 2.3 on the interval $(r-1, \infty)$. As mentioned in the proof of the previous theorem, $\overline{x} = \Theta(t)$. We now take the logarithm of both sides in our equality, and we observe the following.

$$\log\left((r+1)\binom{\overline{x}}{r}\right) = \log\binom{t-1}{r}, \text{ so}$$
$$\log(r+1) + \log\binom{\overline{x}}{r} = \log\binom{t-1}{r}$$

By Lemma 7,

$$\log(r+1) + (1+o_{\overline{x}}(1))r\log\left(\frac{\overline{x}}{r}\right) = (1+o_t(1))r\log\left(\frac{t-1}{r}\right)$$

Solving for $\overline{x}$ in this equation and using the fact that $\overline{x} = \Theta(t)$, we get

$$\overline{x} = \left(\frac{t-1}{(r+1)^{1/r}}\right)^{1+o_t(1)}.$$

By Theorem 9, $m_0(n,r,t) = \lceil \overline{x} \rceil$, and the result follows. □

This concludes our section concerning $\text{sat}_{S_r}(n, S_t)$, although determining the exact value of m_0 for arbitrary r and t is still very much of interest, even in the case where $n \geq 2t-1$ and t is odd.

2.3 Counting Stars in Clique-Saturated Graphs

Having considered the minimum number of cliques of a given size in star-saturated graphs and counting stars in star-saturated graphs, we move to counting stars in clique-saturated graphs. Clique-saturated graphs have been of interest since the proof of Turán's Theorem, which established the maximum number of edges among clique-saturated graphs. From the other end, Theorem 1 gives the minimum number of edges among clique-saturated graphs. This gives us the exact value of $\text{sat}_{S_1}(n, K_t)$ since S_1 is simply an edge. As we noted in Chapter 1, the split graph $K_{t-2} + \overline{K}_{n-t+2}$

attains the minimum edge count for Theorem 1. For n sufficiently large, Chakraborti and Loh showed that this graph is the unique K_t-saturated graph minimizing the number of copies of K_r for all $r < t$, as well as the number of copies of C_r [8].

We will show in this section that the split graph is far from optimal when we want to minimize stars S_r for $r \geq 3$; however, it is a good place to start. In particular, it gives us the following upper bound

$$\mathrm{sat}_{S_r}(n, K_t) \leq s_r(K_{t-2} + \overline{K}_{n-t+2}) = (n-t+2)\binom{t-2}{r} + (t-2)\binom{n-1}{r}.$$

Here the first term counts stars centered in the independent set $\overline{K}_{n-t+2}$ and the second counts stars centered in the clique K_{t-2}. Our first result in this section shows that this count is asymptotically best for $r = 2$ when $t \geq 3$. Using terminology introduced by Erdős and Rényi, we will refer to copies of S_2 as cherries [12]. We now prove Theorem 3

Theorem 3. *For all $n \geq t \geq 3$,*

$$\mathrm{sat}_{S_2}(n, K_t) = \frac{t-2}{2}n^2 + O(n^{3/2}).$$

Proof. Note that, since

$$\mathrm{sat}_{S_2}(n, K_t) \leq s_2(K_{t-2} + \overline{K}_{n-t+2}) = (n-t+2)\binom{t-2}{r} + (t-2)\binom{n-1}{r},$$

we have the desired upper bound on our generalized saturation number. it remains to be shown that we can not achieve a lower count asymptotically. To this end, let G be a K_t-saturated graph on n vertices with m edges, and suppose that G has $\mathrm{sat}_{S_2}(n, K_t)$ copies of S_2. We begin by noting two lower bounds on the number of cherries in a

K_t-saturated graph on n vertices. First, since G is K_t-saturated, the addition of any missing edge must create a copy of K_t using that new edge. Thus every pair of non-adjacent vertices must have at least $t-2$ common neighbors, giving us $t-2$ cherries with the non-adjacent pair as the endpoints. This gives us the bound

$$s_2(G) \geq (t-2)\left[\binom{n}{2} - m\right].$$

For our second bound, we note that we can also count cherries by pulling pairs of vertices from a given vertex's neighborhood. That is,

$$\begin{aligned}
s_2(G) &= \sum_{v\in V(G)} \binom{d(v)}{2} \\
&= \sum_{v\in V(G)} \frac{d(v)^2}{2} - m \\
&= \frac{n}{2}\left(\frac{1}{n}\sum_{v\in V(G)} d_i^2\right) - m \\
&\geq \frac{n}{2}\left(\frac{1}{n}\sum_{v\in V(G)} d_i\right)^2 - m \\
&= 2m^2/n - m.
\end{aligned}$$

The last inequality in this string uses Cauchy's inequality. We now have a lower bound on the number of cherries that is increasing in m and another bound that is decreasing in m. Since both bounds must be satisfied, it suffices to determine the value of m for which these bounds agree. This occurs when

$$m = \frac{1}{8}\left[2 - 2(t-2)n + \sqrt{4(t-2)^2n^2 - 8(t-2)n + 4 + 16(t-2)(n^3-n^2)}\right].$$

Evaluating our lower bounds at this choice of m, we see that $s_2(G) = \frac{t-2}{2}n^2 + O(n^{3/2})$ as desired. $\square$

It turns out that for $t \geq 4$, the split graph is not only asymptotically tight, but it is in fact the unique extremal graph minimizing the number of cherries among K_t-saturated graphs. This was recently proved by Ergemlidze, Methuku, Tait, and Timmons [14]. Now, in the situation where $t = 3$, the optimal number of edges with regards to the two bounds in the previous proof is $m = \frac{n}{2}\sqrt{n-1}$. Such a graph is $(\sqrt{n-1})$-regular and has the property that every pair of non-adjacent vertices has a unique common neighbor. The graph with this property is the Moore graph of diameter 2 and girth 5, where the girth of a graph is the length of its smallest cycle.

Definition 3. *A* ***Moore graph*** *is a d-regular graph of diameter k with*

$$1 + d\sum_{i=0}^{k-1}(d-1)^i$$

vertices. Equivalently, it is a graph of diameter k and girth $2k+1$.

That is, a Moore graph has the largest possible number of vertices among d-regular graphs with diameter k. The fact that this is the maximum number of vertices in such a graph can be seen by counting the number of vertices a given distance from a vertex v. In particular, there are d vertices distance 1 from v and at most $d(d-1)$ vertices distance 2 from v. This is obtained when the vertices with distance 2 from v have unique neighbors in $N(v)$. Similarly, there are at most $d(d-1)^{i-1}$ vertices at distance i from v.

In the case where $k = 2$, we see that the number of vertices is $n = d^2 + 1$. That is, $d = \sqrt{n-1}$, as in the situation previously described. Furthermore, these graphs are strongly regular since every pair of non-adjacent vertices have the same number

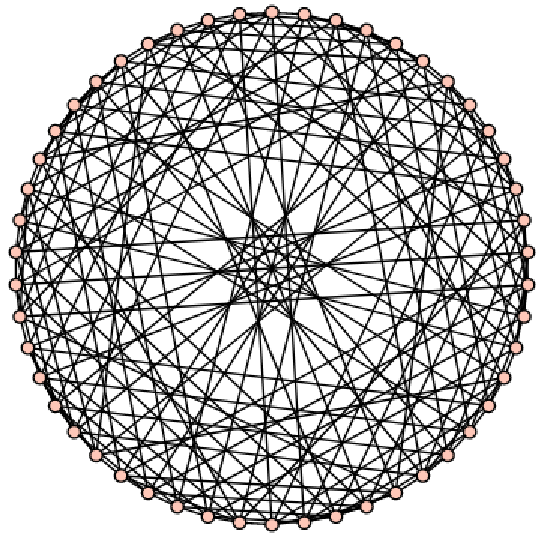

Figure 2.3: Hoffman-Singleton graph

of common neighbors, namely 1, and every pair of adjacent vertices have the same number of common neighbors, namely 0. The first of these gives the bijection between non-edges and cherries, and the second is a consequence of being triangle-free.

It is known that such Moore graphs exist on 5, 10, and 50 vertices, as well as possibly 3250 vertices [20]. The three known graphs are C_5, the Petersen graph P, and the Hoffman-Singleton graph. The last of these is a 7-regular graph on 50 vertices pictured in Figure 2.3. We will now discuss in more detail the connection between Moore graphs and the problem of determining the value of $\mathrm{sat}_{S_2}(n, K_3)$.

2.3.1 Cherries in Triangle-Saturated Graphs

As we just mentioned, Moore graphs of diameter 2 and girth 5, when they exist, minimize the number of cherries among K_3-saturated graphs. In particular, they contain fewer copies of S_2 than the split graph, which, in the case of K_3-saturated graphs, is itself a large star. Since there are few enough graphs on a small number of vertices, a quick check shows that for $n \in \{3, 4, 6, 7, 8, 9\}$, the K_3-saturated graph on n vertices with the minimum number of cherries is that large star $K_{1,n-1}$, or S_{n-1}. With the exception of $n = 6$, this is the unique extremal graph. This leads to the following question.

Question 1. *For what values of n, is*

$$s_2(K_{1,n-1}) = \mathrm{sat}_{S_2}(n, K_3)?$$

To better understand what is necessary for a K_3-saturated graph to contain fewer cherries than $K_{1,n-1}$, we can consider bounds on the number of edges m that such a graph can contain. In Figure 2.4, we present constraints on such a graph using the lower bounds from Theorem 3. The upper bound on the number of edges to allow for a graph to have fewer than $\binom{n-1}{2}$ cherries is provided by Proposition 9. Whether there exists a graph on n vertices in this window for given values of n is another matter. Connected to our earlier remarks, there exists a graph at the intersection of the red and blue curves precisely when there exists a Moore graph of diameter 2 on n vertices. More generally, a graph that falls on the blue curve must be regular, and a graph that falls on the red line must have a bijection between non-edges and cherries.

We now state an upper and lower bound on the number of edges for us to consider, the former being necessary to beat the split graph.

Proposition 9. *Let G be a K_3-saturated graph on n vertices and m edges such that $s_2(G) < s_2(K_{1,n-1})$. Then*

$$m < \frac{n}{4} + \frac{1}{4}\sqrt{4n^3 - 11n^2 + 8n}.$$

Proof. The second bound in our proof of Theorem 3 states that $s_2(G) \geq 2m^2/n - m$. Setting $s_2(G) < \binom{n-1}{2}$, the number of cherries in $K_{1,n-1}$, and solving for m yields the

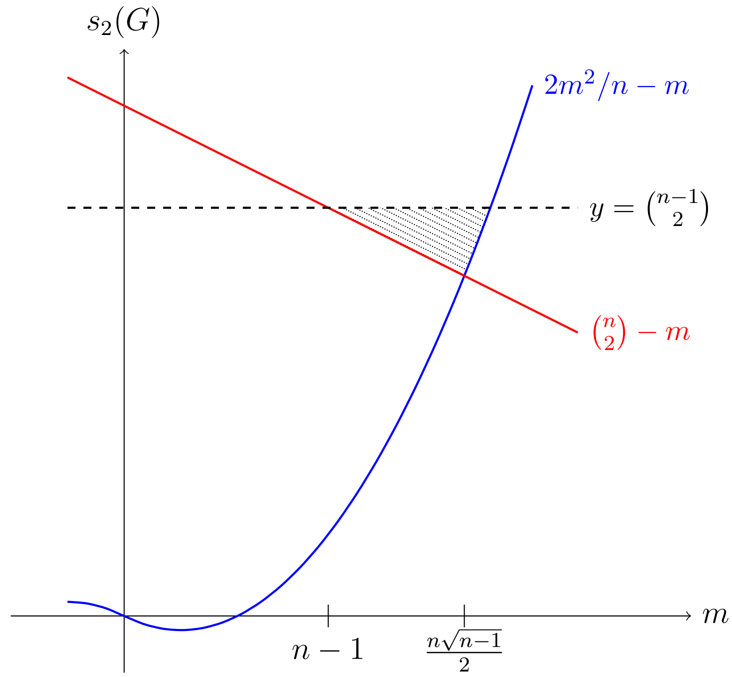

Figure 2.4: Window of interest where a graph on m edges contains fewer cherries than $K_{1,n-1}$

desired inequality. □

Proposition 10. *Let G be a K_3-saturated graph on n vertices and m edges with maximum degree Δ. Then*

$$m \geq \frac{(n-1)^2}{2\Delta} + \frac{1}{2}\Delta.$$

Proof. Let Δ and δ denote the maximum and minimum degrees of G respectively. Setting $k = 3$ in Theorem 4 of [2], we obtain the inequality

$$\delta\Delta \geq n - 1 \tag{2.4}$$

For our graph to have the smallest degree sum possible, we want $n - 1$ vertices of degree δ and a single vertex of degree Δ. Thus

$$2m = \sum_{v \in G} d(v) \geq (n-1)\delta + \Delta \geq \frac{(n-1)^2}{\Delta} + \Delta.$$

$\square$

We note that Inequality 2.4 in our previous proof presents a very powerful observation of Alon, Erdős, Holzmann, and Krivelevich concerning clique-saturated graphs in general. This observation that a small minimum degree forces a clique-saturated graph to have a large maximum degree will show up again in the conclusion of Section 2.3.

We now return our attention to Moore graphs and what makes them so special. As stately previously, these graphs satisfy the two lower bounds on the cherry count in the proof of Proposition 3 since they are regular and have the property that every non-edge corresponds to a unique cherry. To help describe how close a graph is to having this property, we introduce the following notation and definition.

Definition 4. *Let $x, y \in V(G)$ with x not adjacent to y. Let $c(xy)$ denote the number of cherries with x and y as endpoints. We say that the non-edge xy is a **flaw of order** i if $c(xy) = i$ with $i \geq 2$.*

With this terminology, we are ready to state a lower bound on the number of cherries in a K_3-saturated graph in terms of its edges and flaws.

Proposition 11. *Let G be a K_3-saturated graph on n vertices and m edges. If G contains a flaw of order i, then*

$$s_2(G) \geq \binom{n}{2} - m + (i-1) + \binom{i}{2}.$$

Proof. We begin by noting that there is an injection from the set of non-edges to cherries since G is K_3-saturated. That is,

$$s_2(G) \geq \binom{n}{2} - m.$$

We will now count cherries not included in this injection. To this end, let xy be a non-edge in G such that $c(xy) = i \geq 2$. That is, xy is a flaw of order i, and only one of the cherries involving xy has been accounted for in our initial count. Now, since G is K_3-free by assumption, $N(x) \cap N(y)$ must be an independent set. Furthermore, $c(uv) \geq 2$ for every pair $u, v \in N(x) \cap N(y)$ since $\{x, u, v\}$ and $\{y, u, v\}$ each induce a copy of S_2. Thus G has at least $\binom{i}{2}$ flaws of order 2, separate from the non-edge xy. For each of these flaws, only one cherry could have been included in the original count. Therefore we can guarantee that

$$s_2(G) \geq \binom{n}{2} - m + (i-1) + \binom{i}{2}$$

as desired. □

This tells us that having a single large flaw guarantees many flaws in our graph, pushing us further from the ideal situation presented by the Moore graph. In hopes of staying closer to to a Moore graph, it is natural to consider a construction for graphs that begins with a Moore graph. In particular, we shall consider blow-ups of a Moore graph.

Definition 5. *A* ***blow-up*** *B of a graph G is obtained by replacing one or more vertices of G with independent sets. We refer to vertices in an independent set that replaced a vertex v as* ***clones*** *of v. A pair of vertices x, y are adjacent in B if they were adjacent in G, if x is a clone of a vertex adjacent to y in G, or if x and y are*

n	Cherries	Extremal Graph(s)	Notes
3	1	$K_{1,2}$	$K_{1,2}$ is the only K_3-saturated graph on 3 vertices
4	3	$K_{1,3}$	Only options are C_4 and $K_{1,3}$
5	5	C_5	Only options are $K_{1,4}$, $K_{2,3}$, and C_5
6	10	$C_5^\star$ $K_{1,5}$	Only options are $K_{1,5}$, $K_{2,4}$, $K_{3,3}$, and $C_5^\star$
7	15	$K_{1,6}$	There are 6 candidate graphs
8	21	$K_{1,7}$	Verified with Sage
9	28	$K_{1,8}$	Verified with Sage
10	30	P	The Petersen graph P is a Moore graph

Table 2.1: K_3-saturated graphs on n vertices minimizing the cherry count

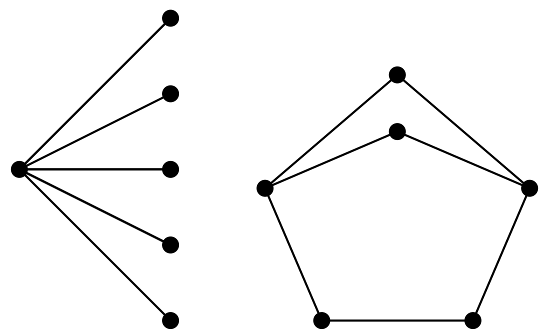

Figure 2.5: K_3-saturated graphs on 3 vertices with the minimum number of cherries

clones of vertices which were adjacent in G.

The size of the independent sets that replace vertices in G are not required to be the same size. Note that any blow-up of a Moore graph is K_3-saturated. In particular, adding an edge between vertices that are clones of the same vertex will create a triangle involving those two vertices and any common neighbor. Any other added edges will create the same triangles that they would have in the original Moore graph. At this point, we remind the reader of our remark that $K_{1,5}$ was not the unique K_3-saturated graph on 6 vertices with the minimum number of cherries. The other graph with the same cherry count, 10, is given by the single vertex blow-up of C_5, the smallest Moore graph of diameter 2. Call this graph $C_5^\star$. See Table 2.1 for a summary of the optimal graphs on up to 10 vertices.

In the case of $C_5^\star$, we see that cloning a single vertex in C_5 results in an optimal configuration; however, a 7-vertex blow-up of C_5, whether that means cloning two vertices once or a single vertex twice, results in a K_3-saturated graph with more cherries that $K_{1,6}$. As we move to larger Moore graphs, there is a little more wiggle room. We make this clear with the following pair of results. We write $\alpha(G)$ to denote the size of the largest independent set in G.

Proposition 12. *Let M be a Moore graph with diameter 2 on n vertices. Let M_k be a graph on $n+k$ vertices obtained by cloning k vertices (possibly with repetition) of M, and let $\alpha = \alpha(M)$. Then for all $k \leq \alpha$, $s_2(M_k)$ is minimized when the cloned vertices form an independent set $\mathcal{I}$ of size k. Furthermore, $s_2(M_k)$ is independent of the choice of $\mathcal{I}$.*

Proof. Let M_k be obtained from M by cloning k vertices, possibly with repetitions, where $k \leq \alpha$. Since M is strongly regular, $s_2(M_1)$ is independent of the cloned vertex. Let $k \geq 2$ and let M_k be obtained from an optimal choice of M_{k-1} by cloning a vertex v. Call the new vertex v'. Let a denote the number of times vertices in $N_M(v)$ have been cloned already. Similarly, let b denote the number of times vertices in $V(M) - N_M[v]$ have been cloned, and let c denote the number of times that v has already been cloned. By these definitions, $a + b + c = k - 1$. We now count the number of cherries that are created by the addition of v'. There are

$$\binom{\sqrt{n-1}+a}{2}$$

copies centered at v' since the degree of v was $\sqrt{n-1}+a$ in M_{k-1}. The number of cherries that use v' as a leaf is

$$\sum_{u\sim v} d_{M_{k-1}}(u) \geq (\sqrt{n-1}+a)(\sqrt{n-1}+c)+b. \tag{2.5}$$

Equality holds when there is a unique vertex in $N_{M_{k-1}}(v)$ that is adjacent to each of the cloned vertices from $V(M) - N_M[v]$. That is, when there is a unique cherry involving v and each of the clones of vertices in $V(M) - N_M[v]$. Observing that Inequality 2.5 holds with equality and is minimized when $a = c = 0$, we obtain the desired result. That is, the number of cherries is minimized when v' is not adjacent to any of the previously added vertices. By induction M_{k-1} is obtained by cloning an independent set, and hence so is M_k. Since the cherry counts obtained by adding v' only depend on it being non-adjacent to the other clones, it follows that $s_2(M_k)$ is independent of the independent set cloned. □

This proposition tells us that the optimal way of cloning vertices in a Moore graph to minimize the number of cherries is to clone the vertices of an independent set once. We are now interested in determining how large of an independent set we can blow up in a Moore graph to be obtain a graph on n vertices that contains fewer cherries than $K_{1,n-1}$. Our next proposition resolves this question.

Proposition 13. *Let M be a d-regular Moore graph with diameter 2. Let M_k be a graph on d^2+1+k vertices obtained by cloning an independent set of size $k \leq \alpha$ in M. Then $s_2(M_k) \leq s_2(K_{1,d^2+k})$ if and only if $k \leq d-1$. Equality holds only when $k = d-1$.*

Proof. We first observe that $\alpha \geq d$ since the neighborhood of any vertex in a K_3-saturated graph must be an independent set. We now note that M contains $(d^2+1)\binom{d}{2}$ cherries. Construct M_k by cloning an independent set $\mathcal{I}$ of size k. Call the set of newly added vertices $\mathcal{I}'$. Since all of the cloned vertices belong to an independent set, we gain $k\binom{d}{2}$ cherries centered in $\mathcal{I}'$. Since each pair of non-adjacent vertices in M has a unique common neighbor, there are $\binom{k}{2}$ cherries whose ends are both in $\mathcal{I}'$. The number of cherries with exactly one end in $\mathcal{I}'$ is d^2k since we have k choices for the vertex in $\mathcal{I}'$, d choices for its neighbor, and d choices for the second leaf using neighbors from the original vertices of M. Altogether, we have

$$s_2(M_k) = (d^2 + 1 + k)\binom{d}{2} + d^2k + \binom{k}{2}.$$

Setting $s_2(M_k) \leq s_2(K_{1,d^2+k})$, we obtain $k \leq d - 1$ as desired. We now observe that

$$s_2(M_{d-1}) = s_2(K_{1,d^2+d-1}).$$

Lastly, we note that

$$s_2(M_k) - s_2(M_{k-1}) = d^2 + k > d^2 + k - 1 = s_2(K_{1,d^2+k}) - s_2(K_{1,d^2+k-1}),$$

and therefore $s_2(M_k) < s_2(K_{1,d^2+k})$ whenever $k < d - 1$. $\square$

We note that when M is the Hoffman-Singleton graph, this proposition tells us that we can clone the vertices of an independent set of up to 5 vertices and still contain fewer copies of S_2 than a large star $K_{1,n-1}$. There is a tie if we clone an independent set of size 6. Although we do not know if there exists a Moore graph with diameter 2 on 3250 vertices, this result tells us that if such a graph exists, then we can clone the vertices of an independent set of up to 55 vertices and still contain fewer cherries than the star on as many vertices.

Perhaps the main takeaway from these results is that there exist K_3-saturated

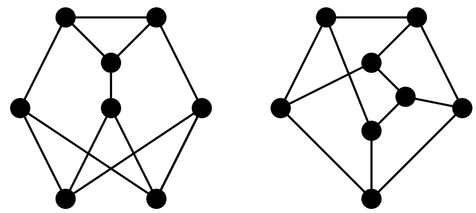

Figure 2.6: K_3-saturated graphs with defect 2

graphs G on n vertices which are not Moore graphs such that $s_2(G) < s_2(K_{1,n-1})$. We mentioned earlier that Moore graphs have the largest possible number of vertices among d-regular graphs of diameter k. We are only concerned with the case where $k = 2$, and since we are interested in values of n for which Moore graphs don't exist, it is of interest to consider graphs with almost as many vertices as possible given the constraints just mentioned. We will say that a graph with maximum degree d and diameter 2 has *defect* δ if it has $d^2 + 1 - \delta$ vertices. That is, it has δ fewer vertices than the maximum.

It turns out that you can't be too close to a Moore graph in this regard. For example, Erdős, Fajtlowicz, and Hoffman proved that the cycle C_4 is the only K_3-saturated graph with maximum degree d on d^2 vertices [13]. Thus we can't find a K_3-saturated graph with defect 1 on more than 4 vertices. Their proof, perhaps unsurprisingly considering standard proofs regarding Moore graphs, relies on an analysis of the eigenvalues of the adjacent matrix of a potential graph. They do however provide graphs with defect 2. See Figure 2.6. For additional results concerning graphs of diameter 2 and defect 2, see [9, 29, 30]

We end our discussions related to Moore graphs, defects, and cherries by proving the following results. We pay special attention to the presence of cycles on 4 vertices and the role that they play in counting cherries.

Proposition 14. *If G is a d-regular, K_3-saturated graph on n vertices such that every* *vertex is contained in a unique copy of C_4, then $G \cong C_4$.*

Proof. Since G is d-regular, it has $\frac{nd}{2}$ edges and $\binom{n}{2} - \frac{nd}{2}$ non-edges. Since every vertex is contained in exactly one copy of C_4, there are $\frac{n}{2}$ flaws of order 2 and none of higher order. That is, the remaining $\binom{n}{2} - \frac{nd}{2} - \frac{n}{2}$ non-edges give rise to unique cherries. Therefore

$$n\binom{d}{2} = s_2(G) = \frac{n}{2} + \binom{n}{2} - \frac{nd}{2}.$$

This holds only when $n = d^2$. Since K_3-saturated graphs have diameter 2, it follows from the aforementioned result of Erdős, Fajtlowicz, and Hoffman in [13] that $G \cong C_4$. □

Proposition 15. *Suppose G has diameter 2, maximum degree d, and defect δ. Then the following hold.*

1. *G is regular or $\delta \geq d$.*

2. *If G is K_3-free and $\delta < d$, then every vertex of G is contained in at least one copy of C_4.*

Proof. Let G be a graph with diameter 2, maximum degree d, and defect δ. If G is not regular, then there exists a vertex of degrees less than $d - 1$. Considering the vertices by their distance from said vertex, we have that

$$n \leq 1 + (d-1) + (d-1)^2 = d^2 + 1 - d.$$

This is less than $d^2 + 1 - \delta$ if $d > \delta$. Since G has defect δ, it follows that $\delta \geq d$.

Now, suppose G is K_3-free and has defect $\delta < d$. Since $\delta < d$, G is d-regular. If a vertex v is in no copy of C_4, then by considering its neighbors and those at distance

2, we have $n \geq 1 + d + d(d-1) = d^2 + 1$, a contradiction. Therefore every every vertex must be contained in a copy of C_4. $\square$

Proposition 16. *If G has diameter 2 and has $d^2 + 1 - \delta$ vertices with maximum degree d and minimum degree ℓ, then*

$$\ell \geq d - \frac{\delta}{d}.$$

Equivalently,

$$\delta \geq d^2 - d\ell = d(d - \ell).$$

Proof. If G has minimum degree ℓ and maximum degree d, then

$$n \leq 1 + \ell + \ell(d-1) = \ell d + 1.$$

Since $n = d^2 + 1 - \delta$, we have $d^2 + 1 - \delta \leq \ell d + 1$. Rearranging gives the desired inequalities. $\square$

Proposition 17. *Suppose G is d-regular with diameter 2 and $n = d^2 + 1 - \delta$ vertices such that some vertex is contained in exactly ρ copies of C_4. Then $\delta \geq \rho$ with equality when G is K_3-free.*

Proof. Let $v \in G$ be a vertex contained in ρ copies of C_4. Since v is in ρ copies of C_4, there exist exactly ρ vertices in $N_2(v)$ with two edges to vertices in $N(v)$. Every other vertex in $N_2(v)$ must have one such edge. Since G is d-regular and there are ρ pairs of vertices in $N(v)$ with a common neighbor in $N_2(v)$ vertices are adjacent to two vertices in $N(v)$, $|N(v)| = d$ and $|N_2(v)| \leq d(d-1) - \rho$. Adding the vertices based on their distance from v, we obtain $n \leq d^2 + 1 - \rho$. If G is K_3-free, then $N(v)$ is an independent set, and these inequalities are equalities, and the proof is complete. $\square$

Note that although copies of C_4 can overlap in different ways, it does not matter. Finally, we state one last result that connects to our mission of finding graphs with fewer cherries than a large star.

Proposition 18. *Let G be a d-regular graph on $d^2+1-\delta$ vertices with diameter* 2. *Then $s_2(G) < s_2(K_{1,d^2-\delta})$ if and only if $\delta \le d-2$.*

Proof. Note the following.

$$s_2(G) = (d^2+1-\delta)\binom{d}{2} \text{ and } s_2(K_{1,d^2-\delta}) = \binom{d^2-\delta}{2}.$$

Thus

$$s_2(K_{1,d^2-\delta}) - s_2(G) = \frac{1}{2}\left(\delta^3 + \delta(1-d-d^2) + d^3 - 2d^2 + d\right).$$

Setting this equal to 0, we have that

$$\delta = \frac{1}{2}\left(d^2 + d - 1 \pm \sqrt{d^4 - 2d^3 + 7d^2 - 6d + 1}\right).$$

Call these roots δ_+ and δ_-. We now note that $d^2 < \delta_+$ and $d-2 < \delta_- < d-1$. We end by observing that when $\delta = d-2$, we have

$$s_2(G) = (d^2-d+3)\binom{d}{2} = \frac{1}{2}(d^2-d+3)(d^2-d) = \frac{1}{2}(d^4-2d^3+4d^2-3d)$$
$$s_2(K_{1,d^2-\delta}) = \binom{d^2-d+2}{2} = \frac{1}{2}(d^2-d+2)(d^2-d+1) = \frac{1}{2}(d^4-2d^3+4d^2-3d+2).$$

That is, the d-regular graph has fewer copies of S_2 than the star for $\delta = d-2$ and therefore also for all $\delta \le d-2$. Similarly, the star contains fewer copies of S_2 for $\delta = d$ and hence all $\delta \ge d-1$. $\square$

2.3.2 Counting Larger Stars in Clique-Saturated Graphs

We end our section on $\text{sat}_{S_r}(n, K_t)$ by considering the number of copies of S_r in K_t-saturated graphs for $r \geq 3$. We begin this task with a pair of results proved by Alon, Erdős, Holzman, and Krivelevich in [2], the first of which appeared in less general form in the proof of Proposition 10. We provide a short proof due to its simplicity and the insight it provides for K_t-saturated graphs.

Lemma 8 (Alon et al., 1996)**.** *Let G be a K_t-saturated graph on n vertices with $\delta(G) = \delta$ and $\Delta(G) = \Delta$. Then*

$$\delta \geq \frac{(t-2)(n-1)}{\Delta + t - 3}.$$

Proof. Let v be a vertex of degree δ in G and let $A = N(v)$ and $B = V(G) \setminus N[v]$. Since G is K_t-saturated, $|N(v) \cap N(u)| \geq t-2$ for all $u \in B$. Also, every vertex in A has at most $\Delta - 1$ neighbors in B. Thus

$$(t-2)(n-\delta-1) \leq e(A, B) \leq \delta(\Delta - 1)$$

where $e(A, B)$ is the number of edges uv where $u \in A$ and $v \in B$. Rearranging yields the desired inequality. □

An immediate consequence of this result is that $\Delta(G) = \Omega(n^{1/2})$ when G is K_t-saturated. We now briefly describe a construction provided by the same authors that produces a K_t-saturated graph with $\Delta(G) = \Theta(n^{1/2})$. We build the vertex set as follows.

1. Begin with a projective plane P of order q where q is a power of a prime and $q \geq t-1$ ($q \geq 3$ for $t = 3$).

2. Label the points and lines of P. In particular, we label the points $p_0, \ldots, p_{q^2+q}$ and the lines $\ell_0, \ldots, \ell_{q^2+q}$ in such a way that $p_{q^2+q} \in \ell_i$ for all $0 \leq i \leq q$ and $p_{iq+j} \in \ell_i$ for all $0 \leq i \leq q$ and $0 \leq j \leq q-1$.

3. Remove the point p_{q^2+q} and the $q+1$ lines it is contained in to obtain a truncated projective plane P'.

4. Make $t-2$ copies of P'.

5. Blow up each point into $t-1$ vertices of the form (i, j, τ, s) where i is the level, j is the position, τ is the type, and s is the copy of P'. The type refers to which of the $t-1$ vertices we are considering from the blow-up of a given point. We call this vertex set V_0.

6. Introduce q^2 independent sets $V_1, \ldots, V_{q^2}$ of size about

$$\left\lfloor \frac{1}{q^2}[n - (q^2+q)(t-1)(t-2)] \right\rfloor .$$

This gives us our vertex set. We define the graph's adjacency rules as follows.

1. If $u \in V_0$ and $v \in V_k$ for some $k \neq 0$, then u is adjacent to v if and only if the point from which u originated belongs to the line ℓ_k.

2. For $u, v \in V_0$, say $u = (i, j, \tau, s)$ and $v = (i', j', \tau', s')$:

    a) If $i = i'$, then u is adjacent to v if and only if the following conditions hold:

        i. $\tau \neq \tau'$

        ii. $j \neq j'$ or $s \neq s'$

    b) If $i < i'$, then u is adjacent to v if and only if the following conditions hold:

        i. $s' = s + 1$ in $\mathbb{Z}_{t-2}$ and $j' = j + \alpha$ in $\mathbb{Z}_q$ where $\alpha \in \{1, \ldots, q-1\}$.

We now state their theorem, the proof of which amounts to showing that the graph described by the above construction is in fact K_t-saturated and has the appropriate maximum degree.

Theorem 12 (Alon et al., 1996)**.** *For every $t \geq 3$, there exists a K_t-saturated graph G on n vertices with*

$$\Delta(G) \leq \left(\frac{(t-2)(2t-3)+1}{\sqrt{(t-1)(t-2)+1}} + o(1) \right) \sqrt{n}.$$

The following is immediate.

Corollary 5. *For every $t \geq 3$ and $r \geq 3$,*

$$\mathrm{sat}_{S_r}(n, K_t) = O(n^{r/2+1}).$$

Proof. Let G be the graph given by Theorem 12. By the same theorem, G has maximum degree $\Delta = O(n^{1/2})$. Since

$$s_r(G) = \sum_{v \in G} \binom{d(v)}{r} \leq n \binom{\Delta}{r},$$

we have the desired upper bound on $\mathrm{sat}_{S_r}(n, K_t)$. □

Our upper bound on $\mathrm{sat}_{S_r}(n, K_t)$ for $r \geq 3$ and $t \geq 3$ comes from the very clever construction of a graph. To find a lower bound, one needs to say something about the general structure of K_t-saturated graphs and their degrees. Ergemlidze, Methuku, Tait, and Timmons, did this, proving that the upper bound of the previous theorem is of the correct order of n [14]. We now state their result.

Theorem 13 (Ergemlidze et al., 2021)**.** *For integers $n \geq t \geq 3$ and $r \geq 3$,*

$$\mathrm{sat}_{S_r}(n, K_t) = \Theta(n^{r/2+1})$$

We end Section 2.3 by demonstrating that the split graph $K_{t-2} + \overline{K}_{n-t+2}$ not only fails to minimize stars among all K_t-saturated graphs for $r \geq 3$ and $t \geq 3$; it is still not optimal for minimizing stars when considering families of K_t-saturated graphs with linear maximum degree. The graph in Figure 2.7 will be the starting point for our construction.

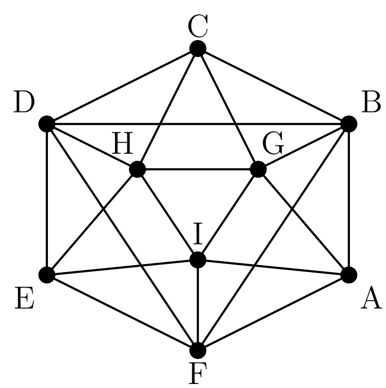

Figure 2.7: K_4-saturated graph $G_{4,9}$ on 9 vertices

Proposition 19. *Let $t \geq 4$ and $r \geq 3$. There exists a sequence $(G_{t,n})$ of K_t-saturated graphs on n vertices with $\Delta(G_{t,n}) = \Theta(n)$ and a constant $n_{r,t}$ such that*

$$s_r(G_{t,n}) < s_r(K_{t-2} + \overline{K}_{n-t+2})$$

for all $n \geq n_{r,t}$.

Proof. Consider the graph $G_{4,9}$ in Figure 2.7. We obtain $G_{4,n}$ for $n > 9$ by blowing up vertices A, C, E into independent sets of size as equal as possible. For $t > 4$, we define $G_{t,n}$ to be $G_{4,n-t+4} + K_{t-4}$. That is, we take the disjoint union of $G_{4,n-t+4}$

and K_{t-4} and add all possible edges between the two sets of vertices. Since $G_{4,n-t+4}$ is K_4-saturated, joining $t-4$ universal vertices, i.e. vertices adjacent to all other vertices, results in a K_t-saturated graph. Given r and t, we can find a constant $n_{r,t}$ such that $s_r(G_{t,n}) < s_r(K_{t-2} + \overline{K}_{n-t+2})$ for all $n \geq n_{r,t}$. Since $G_{t,n}$ is obtained by blowing up three vertices into independent sets of size as equal as possible, there are three cases to consider. We address one of those cases in detail now and omit the details for the other two cases.

We start by proving that $s_r(G(4,n)) < s_r(K_2 + \overline{K}_{n-2})$ for all $n \geq 9$ and $3 \leq r \leq n-1$ in the case where $n = 3k+6$. To this end, we observe that

$$s_3(K_2 + \overline{K}_{n-2}) = 2\binom{3k+5}{3} + (3k+4)\binom{2}{3} = 9k^3 + 36k^2 + 47k + 20$$

and

$$s_3(G_{4,n}) = 3k\binom{4}{3} + 6\binom{2k+3}{3} = 8k^3 + 24k^2 + 34k + 6.$$

Thus $s_3(G_{4,n}) < s_3(K_2 + \overline{K}_{n-2})$ for all $k \geq 1$. Similarly, an explicit count shows that $s_4(G_{4,n}) < s_4(K_2 + \overline{K}_{n-2})$ and $s_5(G_{4,n}) < s_5(K_2 + \overline{K}_{n-2})$. For $r \geq 5$, our counts are simplified as the vertices of small degree contribute no copies of S_r. Now, when $2k+3 < r < 3k+6$, the desired inequality is immediate as our star's degree exceeds the maximum degree of $G_{4,n}$. That is, $s_r(G_{4,n}) = 0$ for $r \geq 2k+4$. Assume then that $r \leq 2k+3$. We now observe the following for $r \geq 6$.

$$\begin{aligned}
s_{r-1}(K_2 + \overline{K}_{n-2}) - s_{r-1}(G_{4,n}) &= 2\binom{3k+5}{r-1} - 6\binom{2k+3}{r-1} \\
&= 2\binom{3k+5}{r} \cdot \frac{r}{3k+6-r} - 6\binom{2k+3}{r} \cdot \frac{r}{2k+4-r} \\
&= \frac{r}{3k+6-r}\left[s_r(K_2 + \overline{K}_{n-2}) - s_r(G_{4,n}) \cdot \frac{3k+6-r}{2k+4-r}\right] \\
&< \frac{r}{3k+6-r}\left[s_r(K_2 + \overline{K}_{n-2}) - s_r(G_{4,n})\right]
\end{aligned}$$

The inequality at the end holds since $r \leq 2k+3$. By induction, we have that $s_{r-1}(K_2 + \overline{K}_{n-2}) - s_{r-1}(G_{4,n}) > 0$. This, along with the string of inequalities above, gives us the desired result.

We now proceed to the more general case concerning $G_{t,n}$ for all $t \geq 4$ where $n = 3k+t+2$. We observe that

$$s_r(G_{t,n}) = (t-4)\binom{3k+t+1}{r} + 6\binom{2k+t-1}{r} + 3k\binom{t}{r}$$

and

$$s_r(K_{t-2} + \overline{K}_{n-t+2}) = (t-2)\binom{3k+t+1}{r} + (3k+4)\binom{t-2}{r}.$$

We proceed by induction on r and t. We have already shown that the desired inequality holds for all $r \geq 3$ when $t = 4$. We next show that the inequality holds in our case for all $t \geq 4$ when $r = 3$ and $n \geq \frac{11+3\sqrt{33}}{2}t + 2$. Substituting $r = 3$ in the above equations, we see that

$$s_3(K_{t-2} + \overline{K}_{n-t+2}) - s_r(G_{t,n}) = k^3 - 3k^2(t-8) + k(-6t^2 + 36t - 35) + 6t - 10.$$

Rearranging and setting $k = ct$ where $c > 0$ is a constant, we obtain

$$t^3(c^3 - 3c^2 - 6c) + t^2(24c^2 + 36c) + t(6 - 35c) - 10.$$

We want this difference to be positive. Note that

$$t^2(24c^2 + 36c) + t(6 - 35c) - 10 > 0$$

for any choice of c and t. Thus we are guaranteed a positive difference whenever

$c^3 - 3c^2 - 6c \geq 0$. This happens when $c \geq \frac{3+\sqrt{33}}{2}$. That is, when

$$n = 3k + t + 2 \geq \frac{11 + 3\sqrt{33}}{2} t + 2.$$

We now let $r \geq 4$. Since $G_{t,n}$ and $K_{t-2} + \overline{K}_{n-t+2}$ both have universal vertices, it is useful to look at how an individual vertex impacts our counts for stars. In particular, notice the following.

$$s_r(G_{t,n}) = s_r(G_{t-1,n-1}) + s_{r-1}(G_{t-1,n-1}) + \binom{n-1}{r}$$
$$s_r(K_{t-2} + \overline{K}_{n-t+2}) = s_r(K_{t-3} + \overline{K}_{n-t+2}) + s_{r-1}(K_{t-3} + \overline{K}_{n-t+2}) + \binom{n-1}{r}.$$

By induction, $s_r(G_{t-1,n-1}) < s_r(K_{t-3} + \overline{K}_{n-t+2})$ and $s_{r-1}(G_{t-1,n-1}) < s_{r-1}(K_{t-3} + \overline{K}_{n-t+2})$. This along with the previous relationships completes our proof. $\square$

2.4 Counting Paths in Clique-Saturated Graphs

Although most of our attention with regards to generalized saturation numbers $\mathrm{sat}_H(n, F)$ is directed to stars and cliques, we are interested in other choices of H and F as well. In particular, we are interested in the broader topic of saturation involving trees and cliques. Stars are a very simple tree structure to work with, which is part of why they were a great starting point for us. Paths are another important tree, especially since they are on the opposite end from stars in terms of trees with the largest (or smallest) maximum degrees and diameters. We will let P_{r+1} denote a path on $r+1$ vertices. As was the case with stars, clarifying this notation is important since different authors use different conventions. Before stating our first result concerning the number of paths of fixed size in clique-saturated graphs, we state a closely connected result of Kritschgau, Methuku, Tait, and Timmons [23] concerning cycles.

Proposition 20 (Kritschgau et al., 2020)**.** *For $t \geq 5$ and $r \leq 2t-4$,*

$$\mathrm{sat}_{C_r}(n, K_t) = \Theta(n^{\lfloor r/2 \rfloor}).$$

Since a path can be obtained from a cycle by deleting a single edge, it is not surprising that this result is useful when it comes to counting paths in clique-saturated graphs. Moreover, a slight refinement of this theorem's proof, along with a similar argument used by Chakraborti and Loh [8], leads us to the following result regarding the order of n when counting paths in clique-saturated graphs.

Theorem 14. *For $t \geq 4$ and $r \leq 2t-3$,*

$$\mathrm{sat}_{P_{r+1}}(n, K_t) = \Theta(n^{\lceil \frac{r+1}{2} \rceil}).$$

If $r \geq 2t-2$, the split graph is P_{r+1}-free.

Proof. We begin by showing that the generalized saturation number is $O(n^{\lceil \frac{r+1}{2} \rceil})$. We do so by considering the split graph $K_{t-2} + \overline{K}_{n-t+2}$. We can construct a path on $r+1$ vertices in this graph by using at most $\lceil \frac{r+1}{2} \rceil$ vertices from the independent set. The remaining vertices must come from the clique of order $t-2$. Thus the number of copies of P_{r+1} in $K_{t-2} + \overline{K}_{n-t+2}$ is

$$\frac{1}{2}(n-t+2)_{\lceil \frac{r+1}{2} \rceil}(t-2)_{\lfloor \frac{r+1}{2} \rfloor} + o(n^{\lceil \frac{r+1}{2} \rceil}).$$

Our first term is obtained by choosing and ordering $\lceil \frac{r+1}{2} \rceil$ vertices from the independent set and $\lfloor \frac{r+1}{2} \rfloor$ vertices from the clique. We multiply by $\frac{1}{2}$ so that we don't count paths of the form $v_1 v_2 \dots v_{\lceil \frac{r+1}{2} \rceil}$ and $v_{\lceil \frac{r+1}{2} \rceil} \dots v_2 v_1$ separately. The $o(n^{\lceil \frac{r+1}{2} \rceil})$

term accounts for any paths using fewer than k elements in the independent set. This gives us the appropriate upper bound.

For the lower bound, let G be a K_t-saturated graph on n vertices, and let I be an independent set of order $\lceil \frac{r+1}{2} \rceil$ in G. There are $\lceil \frac{r+1}{2} \rceil!$ ways to order the elements of I. Enumerate the elements $v_1, v_2, \ldots, v_{\lceil \frac{r+1}{2} \rceil}$. For each ordering we will give a lower bound on the number of copies of P_{r+1} containing it. For all $1 \leq i \leq \lceil \frac{r+1}{2} \rceil - 1$, let V_i be a set of vertices such that $V_i \subseteq N(v_i) \cap N(v_{i+1})$ and the subgraph induced by V_i is a copy of K_{t-2}. Such copies exist since $v_i, v_{i+1} \in I$ which is an independent set, and G is K_t-saturated. That is, they must have a copy of K_{t-2} in their common neighborhood. Since each V_i has $t-2$ elements, we can pick distinct $u_i \in V_i$ such that $v_1 u_1 v_2 \cdots v_{\lceil \frac{r+1}{2} \rceil - 1} u_{\lceil \frac{r+1}{2} \rceil - 1} v_{\lceil \frac{r+1}{2} \rceil}$ is a path in G. This gives us at least $\frac{1}{2} \lceil \frac{r+1}{2} \rceil! (t-2)_{\lceil \frac{r+1}{2} \rceil}$ copies of P_{r+1} involving every element of I. The factor of $\frac{1}{2}$ accounts for the double-counting of paths being read from left-to-right and right-to-left. Chakraborti and Loh showed in [8] that any K_t-saturated graph contains $\Theta(n^k)$ independent sets of order k for any given k. This gives us the corresponding lower bound, and we have that $\text{sat}_{P_{r+1}}(n, K_t) = \Theta(n^{\lceil \frac{r+1}{2} \rceil})$ as desired.

For the second statement in the theorem, we note that when $r \geq 2t-2$, as is true in general, a copy of P_{r+1} in $K_{t-2} + \overline{K}_{n-t+2}$ must use at least $\lfloor \frac{r+1}{2} \rfloor$ vertices from the clique K_{t-2}. Thus

$$2t - 2 = 2(t-2) + 2 \geq 2 \left\lfloor \frac{r+1}{2} \right\rfloor + 2 > r,$$

a contradiction. Therefore $K_{t-2} + \overline{K}_{n-t+2}$ is P_{r+1}-free. $\square$

Although we know the correct degree of n in $\text{sat}_{P_{r+1}}(n, K_t)$, we would still like to know its asymptotic value. Furthermore, Chakraborti and Loh [8] showed that for n sufficiently large in terms of r and t, the split graph minimizes the number of copies

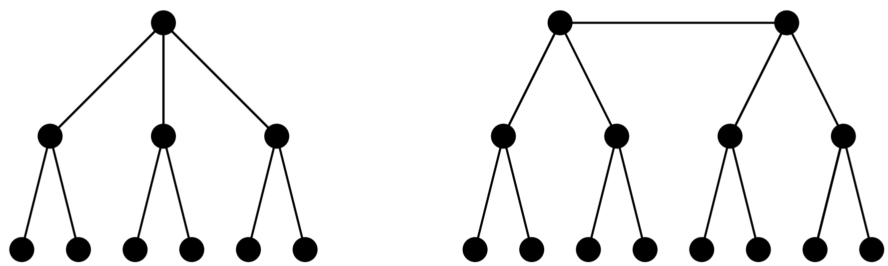

Figure 2.8: P_6-saturated graph on 10 vertices and a P_7-saturated graph on 14 vertices

of C_r in K_t-saturated graphs. This suggests the following question which is still open.

Question 2. *For n sufficiently large, is the number of copies of P_{r+1} in K_t-saturated graphs is minimized by $K_{t-2} + \overline{K}_{n-t+2}$?*

Having briefly considered the counting of paths in clique-saturated graphs, we end this section by commenting on a solution to the problem of counting cliques in path-saturated graphs. It turns out that for n sufficiently large, we can find path-saturated graphs which are trees. Since this immediately resolves our problem, we will describe this useful construction of Kaszonyi and Tuza [22].

When t is odd, consider a rooted tree with $\lfloor \frac{t+1}{2} \rfloor$ levels in which every vertex, except those at highest and lowest levels, have exactly two neighbors in the level below. We require the vertex in the highest level, the root, to have degree at least 3. The longest path in such a graph has t vertices, and a quick check reveals that the graph is in fact P_{t+1}-saturated. When t is even, we take two copies of such a graph and add an edge between the vertices at the highest level. We still require the vertices at the highest level to have degree at least 3, but we only require two neighbors in the level below. The resulting graph is P_{t+1}-saturated. A couple of examples are given in Figure 2.8 and will be important in Chapter 3 when discussing rooted and double-rooted trees.

We now state our proposition which follows immediately from the existence of

P_{t+1}-saturated graphs provided by the discussed construction from [22].

Proposition 21. *If $t \geq 3$ is odd, then $\mathrm{sat}_{K_r}(n, P_{t+1}) = 0$ for all $n \geq 3 \cdot 2^{\frac{t+1}{2}-1} - 2$ and all $r \geq 3$. If $t \geq 4$ is even, then $\mathrm{sat}_{K_r}(n, P_{t+1}) = 0$ for all $n \geq 2^{\frac{t}{2}+1} - 2$ and all $r \geq 3$.*

2.5 General Cases

We now reach the end of our chapter on generalized saturated problems; that is, determining the minimum number of copies of H in F-saturated graphs. To conclude this discussion, we consider the more general questions of counting trees in clique-saturated graphs and of counting cliques in tree-saturated graphs for arbitrary trees. Some results for the traditional saturation problem, i.e. counting edges, on trees were proved by Kaszonyi and Tuza [22]. Many additional results on this subject are due to Faudree, Faudree, Gold, and Jacobson [16]. Here we prove some initial results regarding the generalized saturation problem for trees and cliques. We let $n_T(G)$ denote the number of copies of a tree T in G.

Proposition 22. *Let T be a tree on $r \leq 2t - 4$ vertices. For n sufficiently large, there exists a K_t-saturated graph G on n vertices with minimum degree δ such that $n_t(G) > 0$ provided $\delta \in \{t-2, t-1\}$ or $\delta \geq 2t - 5$.*

Proof. Our aim is to show that for any given tree T and K_t-saturated graph G, these minimum degree restrictions guarantee at least one copy of T in G. Kritschgau et al. [23] demonstrated the exact structure of K_t-saturated graphs with minimum degree $t-2$ and $t-1$. In particular, when $\delta = t-2$, G is isomorphic to the split graph $K_{t-2} + \overline{K}_{n-t+2}$. When $\delta = t-1$, G is isomorphic to $(K_{t-1} - e) + \overline{K}_{n-t+1}$ or $W_t(m_1, 1, m_3, m_4, 1)$ for some $m_1 + m_3 + m_4 = n - t + 1$. Here e is any edge in

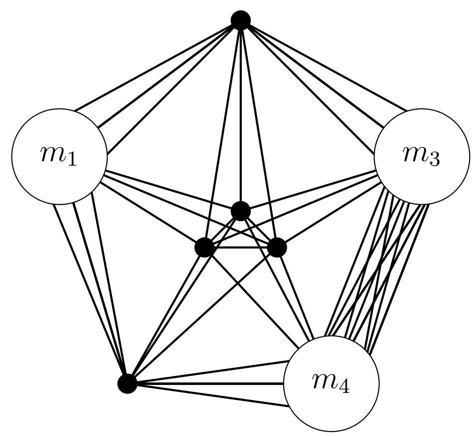

Figure 2.9: K_6-saturated graph $W_6(m_1, 1, m_3, m_4, 1)$

K_{t-1}, and $W_t(m_1, m_2, m_3, m_4, m_5)$ is the graph obtained by taking a wheel with five vertices on the outer cycle and replacing the central vertex with a clique of size $t-3$ and each vertex v_i of the outer cycle with an independent set of size m_i. Two vertices are adjacent in $W_t(m_1, m_2, m_3, m_4, m_5)$ if and only if they replaced adjacent vertices in the original wheel. See Figure 2.9 for an example with independent sets of size m_1, m_2, and m_3 labeled with these sizes.

We now note that each of the above graphs contains a complete bipartite graph $K_{t-2,m}$ for some m where the m vertices are taken from a large independent set of G. Since T is a tree, it is bipartite, and we can partition the vertices into two sets, one of which has size at most $t-2$. For $m \geq r-t+2$, we have that T is a subgraph of $K_{t-2,m}$, which in turn is a subgraph of our K_t-saturated graph. For n sufficiently large, our necessary condition on n is satisfied, and $n_T(G) > 0$.

Lastly, for the case where $\delta \geq 2t-5$, we note the well known result that if a graph has minimum degree at least δ, then it must contain any tree on $\delta+1$ vertices. $\square$

Mimicking a technique of Kaszonyi and Tuza in [22], we provide a lower bound on r for which $\mathrm{sat}_{K_r}(n, F) = 0$ for a t-vertex graph F in terms of its independence number. To this end, let $u(F) = t - \alpha(F) - 1$ where α is the independence number of F. Let d be the minimum number of edges in a subgraph of G induced by an independent set S of size α and one other vertex v. Note that the graph induced by v and S is the star S_d and some number of isolated vertices. It is trivially true that an F-saturated graph can not contain any copy of K_t since F has t vertices. The following proposition shows that we can find F-saturated graphs whose largest cliques are significantly smaller. Before proving our result, we state a key lemma of Kaszonyi and Tuza [22]. Here we say that a graph G is $\mathcal{F}$-saturated for a family $\mathcal{F}$ of forbidden subgraphs $F_1, \ldots, F_k$ if G contains no F_i but the addition of any missing edge creates at least one copy of some F_i.

Lemma 9 (Kaszonyi and Tuza, 1986)**.** *Let $\mathcal{F}' = \{F_i \setminus \{x\} : x \in V(F_i), F_i \in \mathcal{F}\}$ and suppose that some vertex $x \in V(G)$ has degree $d(x) = n-1$. Then G is $\mathcal{F}$-saturated if and only if $G \setminus \{x\}$ is $\mathcal{F}'$-saturated.*

We note that this notion of saturation in which a new edge creates a subgraph from some list of forbidden subgraphs is of further interest in its own right, and there are generalized saturation questions one could ask, but we will not go down that path here. Instead we will state our last result for this section, utilizing this notion and the previously stated lemma.

Proposition 23. *Let F be a graph on t vertices. For n sufficiently large,* $\mathrm{sat}_{K_r}(n, F) = 0$ *for all $r \geq t - \alpha + d$.*

Proof. As defined above, we let $u = t - \alpha(F) - 1$. Suppose G is F-saturated with u vertices of degree $n-1$. Pick u such vertices and remove them one by one. Setting $\mathcal{F} = \{F\}$, we have that G is $\mathcal{F}$-saturated since G is F-saturated. After repeated application of our most recent lemma, we obtain a graph G' that is $\mathcal{F}'$-saturated where $S_d \in \mathcal{F}'$. Thus the maximum degree of G' is $d-1$ and $\omega(G') \leq d$ where ω is the clique number of G. Hence the largest clique in G has size at most $t - \alpha + d - 1$. Therefore G is $K_{t-\alpha+d}$-free and consequently $\mathrm{sat}_{K_r}(n, F) = 0$ for all $r \geq t - \alpha + d$. $\square$

Chapter 3

Tree-saturating Graphs

In this chapter we will continue in the realm of graph saturation. Although the primary question considered here is inspired by the ideas of the preceding chapter, we will not be counting subgraphs. On the contrary, we will be focused solely on the existence of graphs satisfying particular criteria, and we will begin by discussing the reasoning that led to this line of investigation.

Based on the results in Chapter 2 for stars and paths, along with the fact that among all trees T, the star S_t has the largest saturation number [22], one may suspect that for any tree T there exist triangle-free graphs on n vertices for n sufficiently large that are T-saturated. In other words, is

$$\mathrm{sat}_{K_3}(n, T) = 0$$

for all trees T? A quick check shows that this is true for any tree T on $t \leq 6$ vertices. However, the graph T^* in Figure 3.1 shows that this does not hold for all trees. This statement is formalized in our first proposition of the chapter.

Proposition 24. *There does not exist a K_3-free graph that is T^*-saturated.*

Proof. Suppose G is a T^*-saturated graph that is K_3-free. Since T^* has 7 vertices,

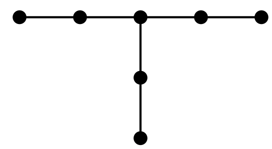

Figure 3.1: Graph T^* for which no K_3-free, T^*-saturated graph exists

G must have at least 7 vertices as well. Since G is K_3-free, the neighborhood of any vertex must be an independent set. We now consider the possible degrees of vertices in G. If G has a vertex v of degree 2, then adding the missing edge between its neighbors x and y must create a copy of T^*. In particular, this copy must use the added edge xy. There are two cases to consider here. Either the added edge is incident to the degree 3 vertex in the new copy of T^*, or it is incident to a degree 1 vertex in the created copy of T^*. If the former case, suppose x is that vertex of degree 3 in our copy of T^*. Then we can replace the edge xy and the second edge incident to y in T^* with the edges xv and vy, resulting in a copy of T^* that was already in G, a contradiction to the assumption that the graph is T^*-free. In the latter case, if y is the vertex of degree 1 in our copy of T^*, then we simply replace xy with xv in our copy of T^*, arriving at the same contradiction. Namely, the edges used in our copy of T^* were already in G, but G is T^*-free. Thus G has no vertex of degree 2.

Now, the maximum degree of G must be at least 3. Otherwise our graph is a matching along with some isolated vertices, but such a graph is not T^*-saturated. We will only focus on a component containing a vertex of degree at least 3 as any edge added within that component must create T^*. Let v be a vertex of degree at least 3 in G. If v has three neighbors a, b, c of degree at least 3, then their neighborhoods must be precisely v and two other common vertices x and y. Otherwise G already contains T^* using v as the vertex of degree 3 in our copy of T^*. We also know that there must be another vertex u in this component of G that does not belong to $\{v, a, b, c, x, y\}$. This is because the graph induced by this vertex set has missing edges and not enough

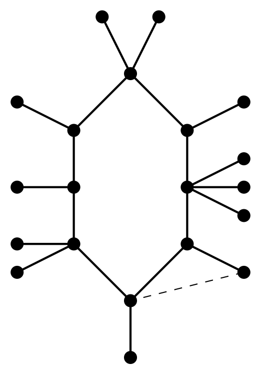

Figure 3.2: Example of a cycle with pendants. The dotted edge does not induce T^* when added.

vertices to create a copy of T^*. A quick check shows that no matter which vertex we join u to, we must already have a copy of T^* once again contradicting the condition that G is T^*-saturated.

We also note that if v has a unique neighbor of degree at least 3, then adding an edge between two of its neighbors of degree 1 will not create a copy of T^* because the use of this new edge requires v to play the role of the degree-3 vertex in T^*, but v originally has only 1 neighbor of degree greater than 1. Therefore every vertex of degree at least 3 is adjacent to at least one vertex of degree 1 and exactly two vertices of degree at least 3. Therefore G is isomorphic to a cycle whose vertices each have at least one pendant, a neighbor of degree 1. However, these graphs are not T^*-saturated. This can be seen by adding an edge between a vertex on the cycle and a pendant of a neighboring vertex as in Figure 3.2. This contradicts the only remaining case. Therefore no K_3-free graph G exists that is T^*-saturated. $\square$

The existence of graphs H for which there do not exist H-saturated graphs with certain properties, leads us to the following new definition.

Definition 6. *Given a graph H, the set $\boldsymbol{SAT}(n, H)$ is the set of all H-saturated graphs on n vertices. We say that H is* <u>*tree-saturating*</u> *if there exists a tree in*

SAT(n, H) *for* n *sufficiently large.*

Our motivating result above shows that there do not exist T^*-saturated graphs which are K_3-free. Consequently, T^* is not tree-saturating. Our hope is to completely characterize which connected graphs are tree-saturating. We do this fully when H is a spider or not a tree. We provide some necessary conditions for when H is a tree as well as some constructions for large classes of other trees including caterpillars.

Since we will dedicate our attention to trees, it is especially important to agree on some terminology. We will say that a tree is *rooted* if there is a designated vertex v that we call the root. This is beneficial as we can analyze the distance of a vertex from the root and whether a path beginning at a vertex moves towards or away from v. Given two vertices x and y, if there exists a path from x to v such that y is an internal vertex on that path, we say that y is an *ancestor* of x and that x is a *descendant* of y. Furthermore, if x is adjacent to y, then we say that y is the *parent* of x and that x is a *child* of y.

3.1 Tree-Saturating Spiders

A spider is a tree with a unique vertex of degree at least 3. That is, the graph consists of a star with a unique path beginning at each endpoint. Such graphs are called spiders because we can view these paths as legs stemming out from that single vertex of larger degree. In this section, we completely characterize the spiders S that are tree-saturating. We already know that there exist saturated trees with respect to paths [22] but none for stars with at least three edges. The latter half of this statement is a consequence of Lemma 1 and the requirement that vertices of small degree form a clique. In particular, you can have at most one vertex of degree 1 in S_t-saturated graphs with $t \geq 3$, and trees need leaves. Thus our only concern is

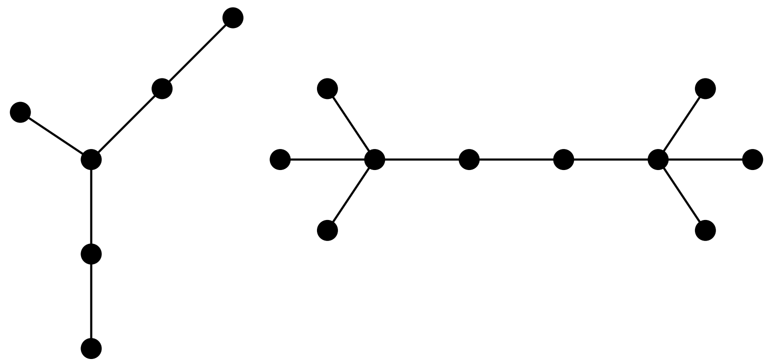

Figure 3.3: $(1, 2, 2)$ spider S with a tree that is S-saturated

spiders with $\ell \geq 3$ legs and longest leg $c \geq 2$.

Note that we can describe a spider by a weakly increasing integer sequence. For example, $(3, 4, 4, 6)$ corresponds to a spider with 4 legs whose lengths are $3, 4, 4$, and 6 respectively. Here the length of a leg is the number of edges on the path from the center vertex to the leaf. The special case where the sequence is all ones corresponds to stars. For another example, $(1, 2, 2)$ is a spider with legs of length $1, 2$, and 2. This spider is tree-saturating as demonstrated by the picture below. Note that we can blow up the leaves in the graph on the right to obtain stars of arbitrarily large size and preserve the property of being S-saturated for this spider S. This specific example will be of interest later on.

We now define several classes of spiders, after which we state our main result of the chapter.

Definition 7. *Let S be a spider with ℓ legs and longest leg which has length c.*

1. *We say that S is **Type 1** if S has a unique longest leg.*

2. *We say that S is **Type 2** if S has exactly two legs of length c, and the next longest is smaller than $c - 1$.*

3. *We say that S is* ***Type 2'*** *if $\ell \geq 4$, S has exactly two legs of length c, and the next longest is exactly $c-1$.*

4. *We say that S is* ***Type 3*** *if $\ell \geq 3$ and S has at least three legs of length c.*

Theorem 15. *Let S be a spider with ℓ legs and longest leg length c. S is tree-saturating if and only if S is Type 1, Type 2, Type 2', or the $(1,2,2)$-spider.*

Having already commented on the $(1,2,2)$-spider, we proceed with a set of lemmas that address the different types of spiders and whether or not they are tree-saturating.

Lemma 10. *Let S be a Type 1 spider. That is, S has ℓ legs and a unique longest leg which has c edges. Then for n sufficiently large, there exists an n-vertex tree that is S-saturated.*

Proof. Let b be the second longest length of a leg in S. We will consider two cases. Suppose $b+c$ is odd. We will build a rooted tree T, beginning with a vertex v. Let v have at least $\ell+1$ children, and let $u \in V(T)$ be another vertex. If the shortest path between u and v has fewer than $\frac{b+c-1}{2}$ edges, then $d(u) \geq \ell$. If the shortest path has $\frac{b+c-1}{2}$ edges, then u is a leaf. This graph T is a tree and is S-saturated. In particular, T is S-free as the longest path in T has $b+c$ vertices, while S has a path of length $b+c+1$. However, a quick check shows that the addition of any missing edge, whether within one of the $\ell+1$ branches or between two, creates a leg of length at least c centered at a vertex with $\ell-1$ additional legs of length b.

Otherwise $b+c$ is even. In this case, we take two copies of the previously described graph T with longest paths from the root vertices having $\frac{b+c}{2}-1$ edges. Finally, we add an edge between the roots of the two trees to complete the construction of T'. As in the previous case, T' has a longest path on $b+c$ vertices and can not contain S. The addition of any missing edge induces a copy of S as desired.

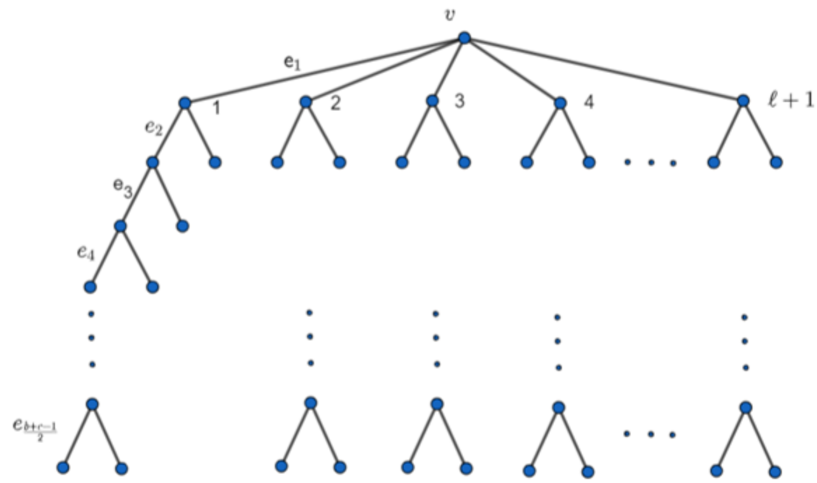

Figure 3.4: T as described in the previous proof where $b+c$ is odd; a quasi-uniform rooted tree of depth $\frac{b+c-1}{2}$ and degree 3

In either case, we have constructed a tree that is S-saturated. Noting that any leaf in our graph can be blown up into an independent set of arbitrarily large size, the proof is complete. □

Note that since the vertices at the second lowest level, i.e. vertices adjacent to leaves, have a lower bound on their degree, this construction provides S-saturated graphs on n vertices for all n sufficiently large.

We now continue with the spiders S that are Type 2 and Type 2'. The S-saturated graphs used in each situation will be very similar to the ones used in the previous proof with some minor structural differences.

Lemma 11. *Let S be a Type 2 spider. That is, S has ℓ legs, exactly two of longest length c, and the next longest is smaller than $c-1$. Then for n sufficiently large, there exists an n-vertex tree that is S-saturated.*

Proof. As in the proof of the previous lemma, we will build a rooted tree T starting with two adjacent root vertices u and v. Let u and v each have at least ℓ children.

Let $w \in T$ be another vertex. If the shortest path from w to u or v has fewer than $c-1$ edges, then $d(w) \geq \ell$. That is, w has at least $\ell - 1$ children. If the shortest path has exactly $c-1$ edges, then we make w a leaf. The graph T is a tree and is S-free since there is no path on $2c$ edges.

We now show that T is S-saturated by considering the missing edges. Adding an edge within the side hanging from u induces a copy of S rooted at u. Likewise for v. Adding an edge between the two sides creates a copy rooted at u and a copy at v. If we add an edge between v and u' where $u' \in N(u)$, we obtain a copy of S rooted at u' since we have legs of length c beginning with the edges $u'u$ and $u'v$, and the graph has enough depth for the remaining legs. Finally, if we add an edge between a descendent w' of u' and v, we obtain a copy rooted at v. One long leg, that is of length c, uses u and one of its legs that does not contain u', and the other long leg goes up from w' to u' and back down as far as needed. The remaining legs are in the v side of the graph.

We end by noting that any leaf can be blown up into a larger independent to generate S-saturated graphs on more vertices. □

Lemma 12. *Let S be a Type 2' spider. That is, S has $\ell \geq 4$ legs, exactly two of longest length c, and the next longest has length exactly $c-1$. Then for n sufficiently large, there exists a tree T such that T is S-saturated.*

Proof. Continuing the trend, we construct a tree T rooted at a vertex v of degree $\ell - 1$. For any vertex $u \neq v$, we have $d(u) \geq \ell + 1$ for any u whose shortest path to v has at most $c-1$ edges. If there are c edges, then u is a leaf. From this construction T is a tree that is S-saturated. In particular, adding an edge between v and any non-neighbor creates a copy rooted at v. A quick check shows that any other added edge creates a copy rooted at u for some neighbor u of v. □

Note that when $\ell = 3$, this construction forces v to have degree 2. We will show shortly that this does not result in an S-saturated graph when S has longest leg of length $c \geq 3$.

We now prove that these are the only spiders S for which an S-saturated tree T exists. To assist with this, we prove the following very useful and general lemma about graph saturation. While this result may be known, we include it for completeness.

Lemma 13. *Suppose G is an F-saturated graph and F' is a subgraph of F. Then G is F'-saturated or contains a copy of F'.*

Proof. Let G be an F-saturated graph, and suppose that G contains no copy of F' where F' is a subgraph of F. Since G is F-saturated, the addition of any missing edge will create a copy of F. Since F' is a subgraph of F, we must have created a copy of F' by adding this arbitrary missing edge. Since we must create F' by adding any missing edge, and G contained no copy of F' to begin with, we have by definition that G is F'-saturated. $\square$

We now prove a result concerning triangle-free graphs that are S-saturated for certain spiders S. The same result holds true for trees since they are triangle-free.

Lemma 14. *If G is a triangle-free graph that is S-saturated for a spider S with $\ell \geq 3$ legs, none of which has length 1, then G has no vertices of degree 2.*

Proof. Let G be a triangle-free S-saturated graph for some spider S with no legs of length 1, and suppose that a vertex $v \in V(G)$ have degree 2. Let x, y be the neighbors of v in G. Since G is triangle-free, we have that x is not adjacent to y. Since G is S-saturated, adding the edge xy must create a copy of S. Since S has at least 3 legs, v can not be the root of this copy of S. Now, we can view the copy of S as having directed edges from the root towards the leaves at the end of each leg. Without loss of generality, edge xy is directed from x to y. Then we have a directed path starting at y that ends somewhere in G that corresponds to a leaf in S. Since S has no legs of length 1 and a unique vertex of degree greater than 1, we are not using the edge xv in this copy of S, even if S is rooted at x. Replacing the edge xy and the final directed edge on the leg that uses xy with directed edges xv and vy, we obtain a copy of S that does not use the added edge. That is, our copy of S was already in G to begin with. This contradicts the fact that G was S-saturated. Therefore no vertex in G can have degree exactly 2. $\square$

We now address the remaining possible spiders: spiders with at least three legs of longest length c for any ℓ legs, and 3-legged spiders of the form $(c-1, c, c)$.

Lemma 15. *Let S be a Type 3 spider. That is, S has ℓ legs and at least 3 legs are of longest length c. Then there is no tree T that is S-saturated.*

Proof. We will perform induction on c. Note that when $c = 1$, the statement is true because there are no star-saturated trees for $\ell \geq 3$. Now assume $c \geq 2$ and that such a tree T exists. Let S' be the Type 3 spider obtained from S by cutting all legs of length c down by 1. By induction, T is not S'-saturated. It follows by Lemma 13 that T must contain S' as a subgraph.

Let v be the root of some copy of S'. Now, adding an edge to our graph can create at most one new leg of length at least c for a spider in T rooted at any given vertex since that leg must use the new edge. Since S has at least 3 legs of length c, the root of any created copy of S must have already had at least two legs of length at least c. At most one of these legs can include v. Since T is connected and acyclic, this means that there must be a leg of length at least c independent of v. Thus v has at least one leg of length at least c rooted to it.

Let v' be a vertex adjacent to v on a longest leg rooted at v, and let P denote the set of vertices v'' such that $d(v', v'') < d(v, v'')$; that is, the set of vertices which are closer to v' than v. Note that $v' \in P$ as $d(v', v') = 0$. We have two cases to consider. First, suppose there exist $x, y \in P$ with greatest depth relative to v such that $x, y \in N(u)$ for some $u \in P$.

Add the edge xy. This creates a copy of S rooted at some vertex r in P where r belongs to the unique path between x and v. See Figure 3.5. This is because x and y have distance at least c from v, and we must use the newly added edge in the created copy of S. The leg using xy has length at most c. But there can be only one leg using edges between r and v. Thus there must be another leg of length c going down from the root r. But the greatest depth relative to the root r is $c - 1$, a contradiction. Therefore xy does not induce a copy of S, and T is not S-saturated.

Otherwise any vertex x in P with greatest depth relative to v is adjacent to a vertex u of degree 2. Let w be the unique other neighbor. Add the edge wx. Unless, S has a longest leg of length c and at least two legs of length 1, then this edge can not help create a copy of S rooted at w since the edge can not contribute to a leg that was not already present. That is, the edges wu and ux could be used in place of wx and xu. Thus our copy of S is rooted at some vertex r in P that is above w. However, the added edge still can not contribute to a leg of S for the same reason. To complete the proof, we note that the remaining case of a spider centered at w with longest leg of length 2 and at least two legs of length 1 requires w to have another neighbor of degree 2. Adding an edge between leafs adjacent to the degree 2 neighbors of w does not induce a copy of S. $\square$

We end with the final special spider to consider.

Lemma 16. *Let S be the $(c-1, c, c)$ spider for some $c \geq 3$. There is no tree T such*

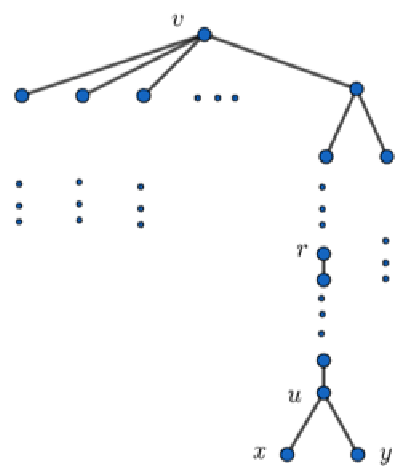

Figure 3.5: Sample T from previous proof with vertices on P

that T is S-saturated.

Proof. Suppose such a tree T exists. By the previous lemma, along with Lemma 13, T is not S'-saturated where $S' = (c-1, c-1, c-1)$ and therefore must contain S' as a subgraph. Let v be the root vertex in a copy of S'. As in the previous proof, there must be at least one leg of length at least c rooted at v. Since a second such leg would give us a copy of S, this is the only such leg. If the greatest depth relative to v is exactly c, then we add the edge xy where $x, y \in N(v)$, x is in the leg of length c, and y is in a shorter leg. This edge does not create a copy of S, meaning that T is not S-saturated.

So assume that the long leg has length at least $c+1$, again letting x be the neighbor of v in this leg. As with the previous proof, we consider two cases. Since T can not have any vertices of degree 2 by Lemma 14 as S has no leg of length 1, x has at least 2 children.

One is on a longest leg rooted at v, but the other belongs to a path of length at most $c-1$ rooted at v. Otherwise, we already have a copy of S rooted at x. Call the children u and w respectively. We now add the edge uw. This can't induce a copy of

S rooted at v because there is already a leg of length c using u that is rooted at v. But it also can't induce one at or below u because only one leg can go up from such a vertex, and using v along with any other leg of length $c-1$ gives us that long leg. That is, the new edge can't help. Therefore either T already had a copy of S or the new edge doesn't create a new copy. In either case T is not S-saturated. □

With all of these lemmas in hand, our main theorem falls into place.

Theorem 15. *Let S be a spider with ℓ legs and longest leg length c. S is tree-saturating if and only if S is Type 1, Type 2, Type 2', or the $(1, 2, 2)$-spider.*

Proof. The result follows immediately from the preceding lemmas. □

3.2 Arbitrary Tree-Saturating Trees

We move briefly into full generality and will return later to specific classes of trees.

Proposition 25. *Let T be an arbitrary tree. If there exists a tree that is T-saturated, then one of the following holds:*

(a) T has a leaf that is adjacent to a vertex of degree 2, or

(b) (i) T has vertices u, v, x, y such that $d(u) = d(v) = d(x) = d(y) = 2$ and $\{u, v, x, y\}$ forms a path of length 4, and

(ii) T has a vertex w with at least two leaf neighbors and no pair of disjoint paths on at least 4 vertices with w as an endpoint.

Proof. Let T' be a tree that is T-saturated, and suppose that T has no leaf that is adjacent to a vertex of degree 2. Since T' is T-saturated, joining any two leaves in T' must induce a copy of T that uses the new edge. If x, y are leaves with a common neighbor, then the edge xy can only be used to connect a leaf of T with a vertex of degree 2. Since T has no such vertices, every leaf in T' is adjacent to a vertex of degree 2.

Let v and x be leaves in T' with degree-2 neighbors u and y respectively. Adding the edge vx must induce a copy of T. By assumption, T has no leaf adjacent to a vertex of degree 2. Hence T must also contain edges uv and xy, along with the unique additional edges incident to u and y. Thus u, v, x, y have degree 2 in T and satisfy the desired adjacencies. That is, they form a path of length 4 in T and condition (b_i) holds.

In addition, if we let v' be an arbitrary vertex of T' and u' be a vertex of greatest distance from v' in T', we can force condition (b_{ii}) by considering the neighbors of u'. In particular, u' must be a leaf that is adjacent to a degree 2 vertex x' with another neighbor w. Add edge wu'. Since this must create a copy of T and T has no leaf adjacent to a degree 2 vertex, the copy of T must use edges wu' and wx'. If it only needed one of them, then the graph would have already contained a copy of T.

This gives us a vertex w with at least two leaf neighbors. Since u' is as far away from v' as possible, we can not find a path on 4 vertices with w as endpoint that is disjoint from the unique path whose first edge is in the direction of v'. That is, in our copy of T, the vertex w has two leaf neighbors and at most one path of at least 4 vertices with w as a endpoint. $\square$

We now briefly define a pair of graphs about which we have an immediate corollary. The double star $D_{r,k}$ is obtained by adding an edge between the central vertices of stars S_r and S_k. The (r, k)-banana tree is obtained from r copies of S_k by adding an edge from a leaf of each copy of S_k to an additional vertex. See Figure 3.6. The following is immediate.

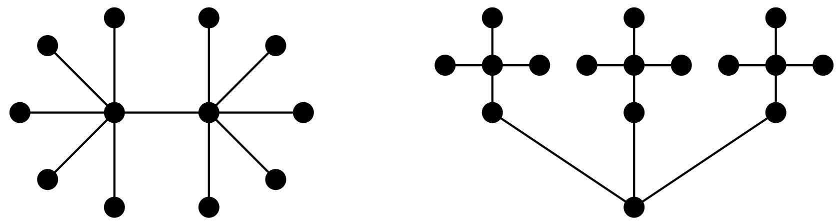

Figure 3.6: Double star $D_{5,5}$ (left) and $(3,4)$-banana tree (right)

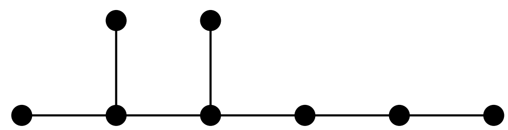

Figure 3.7: Non-spider that is tree-saturating

Corollary 6. *Double stars $D_{r,k}$ with $r, k \geq 2$ and (r,k)-banana trees with $r \geq 2$ and $k \geq 3$ are not tree-saturating.*

Our proposition gives us some restrictions on the possible trees that can be tree-saturating. If we look at small graphs (up to 7 vertices), the only tree-saturating graphs are all spiders. However, it turns out that there are tree-saturating trees which are not spiders once we look at larger numbers of vertices. Using Sage, we found two such trees on 8 vertices, one of which is pictured in Figure 3.7. The code used can be found at the following link `https://cocalc.com/share/public_paths/1ceae877274bbe6854d388524947034f85ac426c`.

This actually gives us a way to find tree-saturating non-spiders for all $n \geq 8$ by cloning the leaves which are adjacent to vertices of degree 3 as many times as we want. We can exhibit a saturated tree by taking a root vertex of degree at least 4 whose children all have degree at least 3. Non-neighbors of the root are leaves. We can make the non-leaves have arbitrarily high degree to account for the generalization to larger values of n. Moving forward, we are interested in developing a stronger picture of which trees are tree-saturating, and we will do so by beginning with caterpillars.

3.3 Tree-Saturating Caterpillars

Spiders are built from a star by building a path off of each leaf. Caterpillars, on the other hand, are obtained from a path by replacing interior vertices with stars S_t with $t \geq 2$. We focus on caterpillars using a different approach than what we used when analyzing spiders. In particular, we will start with some tree T as our host graph and provide some conditions for which caterpillars C it is the case that T is C-saturated. To aid with our approach and to more easily describe our host trees, we introduce some terminology, formalizing a description of the host graphs from our discussion of spiders.

Definition 8. *A* ***quasi-uniform rooted tree*** T *of depth* k *and degree* d *is a tree with a designated root vertex* v *of degree at least* 3 *such that* $d(w, v) = k - 1$ *for all leaves* w *and for which any vertex that is not a leaf has degree at least* d. *Similarly, a* ***quasi-uniform double-rooted tree*** T' *of depth* k' *and degree* d' *is a tree with a pair of adjacent root vertices* u *and* v *of degree at least* 3 *such that* $\min\{d(w, u), d(w, v)\} = k' - 1$ *for all leaves* w *and for which any vertex that is not a leaf has degree at least* d.

Note that we will refer to a vertex's distance from the nearest root as its depth, or level. Given a quasi-uniform rooted (or double-rooted) tree of depth k, we will characterize the caterpillars with maximum path on $2k$ and $2k + 1$ vertices that saturate those trees. For an example of a quasi-uniform rooted tree, see Figure 3.4 or the the figures in the following proof. We will begin with the even case.

Proposition 26. *Let* C *be a caterpillar with maximum path having* $2k$ *vertices with* $k \geq 3$ *and maximum degree* Δ. *For* n *sufficiently large, there exists a tree* T *on* n *vertices such that* T *is* C*-saturated if the following both hold:*

(a) C *has a degree* 2 *vertex that is* 1 *edge away from an end of a maximum path*

(b) C has a degree 2 vertex that is $k-d-1+i$ edges from some end of a maximum path for some $0 \leq i \leq k-d-3$ for all $0 \leq d \leq k-3$.

Proof. Let T' be a quasi-uniform rooted tree with depth k and degree Δ. We will show that the two conditions above are both necessary and sufficient for T' to be C-saturated. Consequently, we will have shown them to be sufficient conditions for a caterpillar C to be tree-saturating. We begin by noting that the addition of any missing edge between two internal vertices of T' will induce any caterpillar with maximum path having $2k$ vertices. This is because the edge creates such a path without using any leaves of T' internally and our degrees are sufficiently large.

Joining sibling leaves forces condition (a). That is, if T' is C-saturated, then C must have a degree 2 vertex that is adjacent to the end of a maximum path in C. Now, if we join a leaf w to any other vertex w' that is not an ancestor of w, then w is a degree 2 vertex in our copy of C that can be treated as being adjacent to an end of C with w' being the corresponding end. No additional restrictions are forced on us from what we have already stated. This is because remaining vertices in the created path have sufficiently large degrees.

The only other type of edge we can add is between w and an ancestor u of w. Suppose that u is at level d for some $0 \leq d \leq k-3$. Here v is the unique vertex at level 0. See Figure 3.8. Adding such an edge induces a copy of C with at least $k-d-1$ vertices below level d after using the edge uw. This is because the maximum path has $2k$ vertices, k of which are from a leaf in a different branch up to v, d to reach u, and the additional 1 for vertex w itself. Similarly, C uses at most $2k-2d-4$ vertices below level d after edge uw. This is done by going up from w to level $d+1$ and back down to a leaf in the same main branch. Since we induce a copy of C by the addition of any missing edge, one of these situations must yield a copy of C with w

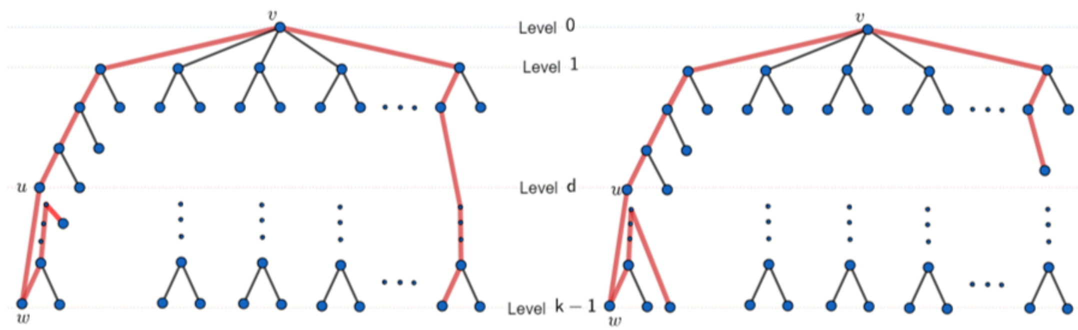

Figure 3.8: The two extreme cases in red for an edge between u and w in T'

serving as a vertex of degree 2 that is $k-d-1+i$ edges from some end of a maximum path in C for some $0 \leq i \leq k-d-3$ for all $0 \leq d \leq k-3$. Figure 3.8 illustrates what the maximum path looks like in each case. We have now shown that condition (b) is necessary. Since these conditions on C result in T' being C-saturated and since we can clone leaves in T' to obtain larger C-saturated trees, the proof is complete. □

We immediately proceed to the corresponding result concerning a caterpillar whose maximum path has an odd number of vertices. The conditions are identical, but the key construction in our proof is different.

Proposition 27. *Let C be a caterpillar with maximum path having $2k+1$ vertices with $k \geq 3$ and maximum degree Δ. For n sufficiently large, there exists a tree T on n vertices such that T is C-saturated if the following hold:*

1. *C has a degree 2 vertex that is 1 edge away from an end of a maximum path,*
2. *C has a degree 2 vertex that is $k-d-1+i$ edges from some end of a maximum path for some $0 \leq i \leq k-d-3$ for all $0 \leq d \leq k-3$,*

Proof. Let T'' be a quasi-uniform double-rooted tree with maximum depth $k-1$ from its roots v and u and with degree Δ. The only new edge of interest to consider is between u and a leaf w that is in a branch below v as all others have been accounted for in the previous proof. There must be at least 1 vertex below u that is used in an induced copy of C. This situation can arise by starting at a leaf in a branch of v that does not contain w and tracing all the way through to uw. On the other hand, there can be at most $k-1$ vertices used below u since that is the maximum depth. Thus C must have a degree 2 vertex that is j edges from some end of a maximum path for some $2 \leq j \leq k$. Here the role of the degree 2 vertex is played by w. This requirement along with those from the previous proposition provide necessary and sufficient conditions for T'' to be C-saturated where C has the appropriate diameter. This uses the fact that aside from vertex w, every vertex that we travel through in our path on $2k+1$ vertices has as large of degree as we want. □

We now consider caterpillars that don't saturate a quasi-uniform rooted tree or a quasi-uniform double-rooted tree for any choice of depth or degrees. In an attempt to show that caterpillars which don't saturate such graphs are in fact not tree-saturating, we prove the following lemma.

Lemma 17. *Every tree contains a pair of leaves u, v with at most one vertex of degree at least 3 on the unique path from u to v.*

Proof. Let T be a tree with vertices $u_1, \ldots, u_r$ of degree at least 3. Construct an auxiliary graph T^* with vertex set $\{u_1, \ldots, u_r\}$. A vertex u_i is adjacent to u_j in T^* if and only if there is no vertex of degree at least 3 on the unique path from u_i to u_j in T. Since T was a tree, so is T^*. In particular, it is connected and acyclic. Let u_s be a leaf in T^*. Then u_s is a vertex of degree at least 3 in T with a unique vertex of minimum distance in $\{u_1, \ldots, u_r\} \setminus \{u_s\}$, namely its neighbor in T^*. Thus there are at least two maximal paths beginning at u_s with no internal vertices of degree at least 3. Let u and v be the ends of these two paths that are distinct from u_s. Then u_s is the unique vertex of degree at least 3 on the path from u to v in T. □

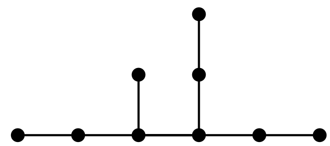

Figure 3.9: Tree-saturating graph T on 9 vertices that is neither a spider nor a caterpillar

With this lemma, we now prove that the first of our conditions for a caterpillar to be tree-saturating via a quasi-uniform rooted (or double-rooted) tree is in fact necessary for a caterpillar to be tree-saturating in general.

Proposition 28. *Suppose C is a caterpillar that is tree-saturating. Then at least one end of a maximum path in C is adjacent to a vertex of degree* 2.

Proof. Let C be a caterpillar that is tree-saturating, and let T be a tree that is C-saturated. By Lemma 17, there exist leaves u and v in T with at most one vertex w of degree at least 3 on the unique path connecting them. Add the edge uv to T. Since T is C-saturated, this must create a copy of C using the edge uv.

Let P be a maximum path in C. This necessarily uses the edge uv since the edge must be used in C and the degree of u and v is 2 in $T+uv$. One end of P must lie on the path from v to w in T or the path from u to w in T. Without loss of generality, it is the latter. Note that this vertex could be u itself. Call this vertex u'. Observe that u' is a leaf in C since it is an endpoint of P. Furthermore, the neighbor of u' in P has degree 2 in $T+uv$. Hence it must have degree 2 in C. $\square$

We note at this point that spiders and caterpillars are not the only trees which are potentially tree-saturating. For example, the graph T in Figure 3.9 is tree-saturating. In particular, a quasi-uniform rooted tree of depth 3. which is P_6-saturated, is also T-saturated.

3.4 Tree-Saturating Non-Trees

We end this chapter by briefly addressing graphs that are not trees in our context as well as some different directions that can be taken.

Lemma 18. *Suppose F is not a tree. If F contains a cycle with at least 4 vertices, then F is not tree-saturating.*

Proof. Let F be a graph with a cycle on at least 4 vertices, and suppose there exists a tree T that is F-saturated. If T has leaves x, y that are adjacent to a common vertex, then adding the edge xy can not induce a cycle with more than 3 vertices. If no such pair exists, then T has a degree 2 vertex u that is adjacent to a leaf v and some other vertex w. Adding the edge vw can only create a cycle with 3 vertices, giving us our desired contradiction. □

The same argument shows that if F is not a tree but is tree-saturating, then F can have at most one cycle of length 3 and no other cycles. This restrictive detail leads us to the following classification of tree-saturating graphs which are not themselves trees.

Proposition 29. *The only connected graphs F that are not trees but are tree-saturating are stars with a single added edge.*

Proof. We begin by noting that a star with a single added edge is tree-saturating by taking a star with at least as many vertices. The previous lemma and remark show that if F is tree-saturating, it must have only one cycle, and it must be of length 3.

If F is not a star with an added edge, then a star is not F-saturated. Thus an F-saturated tree must have a pair of leaves x, y that are separated by at least 3 edges. Thus adding the edge xy can not induce a copy of K_3 and hence no copy of F. □

We end this section with our lone result in which we consider tree-saturating graphs which are not connected.

Proposition 30. *Suppose F is tree-saturating and has multiple components. Then F is a forest.*

Proof. Let T be an F-saturated tree on n vertices where F has $k \geq 2$ components. As previously argued, F can have at most one cycle and such a cycle must be K_3. However, this requires that the addition of any missing edge in T must create a copy of K_3. That is, the diameter of T must be 2, and T must be the star $K_{1,n-1}$. But the only graphs H for which $K_{1,n-1}$ is H-saturated are stars with an added edge. This contradicts the assumption that F has $k \geq 2$ components. Therefore F must have no cycles and is a forest. □

3.5 Other Questions

In the first section of this chapter, we defined tree-saturating graphs. In a similar vein, for a collection $\mathscr{H}$ of graphs, we can say that a graph F is *$\mathscr{H}$-saturating* if there exists a graph $H \in \mathscr{H}$ on n vertices for n sufficiently large such that H is F-saturated. The study of $\mathscr{H}$-saturating graph may be of interest when $\mathscr{H}$ is the collection of triangle-free graphs, as hinted at in the motivation for this perspective on saturation, or r-partite graphs. In particular, our example in Figure 3.1 is not only not tree-saturating, but is not triangle-free-saturating and therefore not bipartite-saturating. On the flip side of this, every graph that we have shown to be tree-saturating is also bipartite-saturating and therefore triangle-free-saturating.

We end our pursuits regarding this topic with the following initial result concerning bipartite-saturating graphs.

Proposition 31. *Suppose H is not a bipartite graph. H is bipartite-saturating if and only if there exists an edge $e \in E(H)$ such that e belongs to every odd cycle in H.*

Proof. We first note that for n sufficiently large, the graph $K_{\lfloor n/2 \rfloor, \lceil n/2 \rceil}$ is H-saturated provided there exists an edge e that is used in every odd cycle of H. This is because the deletion of e in H destroys all odd cycles, resulting in a bipartite graph. Provided n is large enough, this bipartite graph is a subgraph of $K_{\lfloor n/2 \rfloor, \lceil n/2 \rceil}$.

Now, if H does not have an edge that is involved in every odd cycle, then for every edge $e \in E(H)$, there exists an odd cycle C such that $e \notin C$. Suppose a bipartite graph G is H-saturated. Add an edge e' to G. Since G is bipartite and originally has no odd cycles, e' must belong to an odd cycle in the created copy of H. However, e' can not not belong to every odd cycle in H. Either one of the odd cycles is missing, or G already had one. Either way, we have a contradiction. $\square$

Chapter 4

H-Covered Graphs

In this chapter, we return to counting substructures in various graphs, but we direct our attention specifically to counting independent sets, collections of vertices with no edges between any pairs of vertices. We also leave the domain of graph saturation and turn our attention to H-covered graphs. Recall that we say a graph G is H-covered if every vertex of G is contained in at least one copy of H. In order to maximize independent sets, we want to build H-covered graphs efficiently in the sense that they are fairly sparse. However, as noted in the introduction, maximizing independent sets of size at least 3 may require a different structure than what maximizes independent sets of size 2 among H-covered graphs for a given choice of H. Throughout we will write $i_t(G)$ to denote the number of independent sets of size t in G.

4.1 Independent Sets in Star-Covered Graphs

To begin our journey in this direction, we state a powerful result due to Gan, Loh, and Sudakov that will in many cases answer our question [19].

Theorem 16 (Gan, Loh, and Sudakov, 2013)**.** *Let $\delta \leq n/2$. For every $t \geq 3$, every n-vertex graph G with minimum degree at least δ satisfies $i_t(G) \leq i_t(K_{\delta,n-\delta})$, and when $t \leq \delta$, $K_{\delta,n-\delta}$ is the unique extremal graph.*

Since the case where $\delta = 1$ will be of special importance to us, we note that this case was proved earlier by Galvin in [18]. In particular, the following is an immediate consequence of this special case.

Corollary 7. *Fix n, d, and t with $n \geq 4$, $1 \leq d \leq n-1$, and $3 \leq t \leq n-1$. If G is an n-vertex $K_{1,d}$-covered graph, then $i_t(G) \leq i_t(K_{1,n-1})$ with equality if and only if $G \cong K_{1,n-1}$.*

Proof. Let G be a $K_{1,d}$-covered graph on n vertices. This means that every vertex in G is contained in a copy of $K_{1,d}$. Since the minimum degree of $K_{1,d}$ is 1, G has minimum degree at least 1. Setting $\delta = 1$ in the previous theorem implies that $i_t(G) < i_t(K_{1,n-1})$ whenever $G \ncong K_{1,n-1}$. Since $K_{1,n-1}$ is $K_{1,d}$-covered for all $n > d$, this upper bound is best possible amongst $K_{1,d}$-covered graphs. $\square$

While we generally focus on independent sets of size $t \geq 3$, we state the following result to illustrated the the difference between minimizing independent sets of size 2 and of size at least 3. Recall that for graphs G_1 and G_2, we write $G_1 \cup G_2$ to denote the disjoint union of G_1 and G_2. Given an integer k, we write kG_1 to denote the union of k disjoint copies of G_1.

Proposition 32. *Let G be a $K_{1,d}$-covered graph on $n = q(d+1) + r$ vertices where $q \geq 0$ and $d + 1 \leq r \leq 2d + 1$. Then*

$$i_2(G) \leq i_2(qK_{1,d} \cup K_{1,r-1}) = \binom{n}{2} - n + q + 1.$$

The maximum value is attained by G precisely when G is a forest with $q + 1$ components, each of which is $K_{1,d}$-covered. The extremal graph is unique only when $r \in \{d + 1, d + 2\}$.

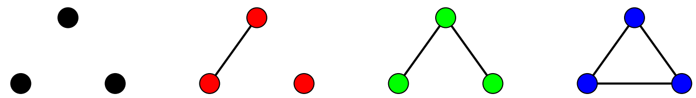

Figure 4.1: All graphs on 3 vertices up to isomorphism

Proof. Suppose G is $K_{1,d}$-covered and maximizes $i_2(G)$. Suppose further that G has k components $G_1, \ldots, G_k$. Each component must contain a copy of $K_{1,d}$ and hence at least $d+1$ vertices. The number of edges in G_i is at least $n_i - 1$ where G_i has n_i vertices for all $1 \leq i \leq k$. Since G is edge-minimal, each G_i has exactly $n_i - 1$ edges and must be a tree. Note that this is in fact possible by considering the case where $G_i \cong K_{1,n_i-1}$ since $n_i \geq d+1$. Thus G is a forest and contains $n-k$ edges. This is minimized when the number of components k is as large as possible. Since G is edge-minimal and each component has order at least $d+1$, we have that k must be $q+1$ and $i_2(G) = i_2(qK_{1,d} \cup K_{1,r-1})$. Uniqueness of G when $r \in \{d+1, d+2\}$ is immediate since every component must have $d+1$ vertices, with one exception of size $d+2$ when $r = d+2$. If $r > d+2$, we can distribute the surplus vertices to any of the G_i as we see fit. □

It is worth noting that this minimum number of edges agrees with the number given by the integer program of Chakraborti and Loh in [7].

4.2 H-Covered Graphs for Small H

We now turn out attention to some small choices of H, from which we will attempt to generalize some results. In particular, we shall consider graphs on 3 and 4 vertices. Since our graphs may be disconnected, we distinguish them by the color of their vertices in Figure 4.1.

We simply remark that an independent set on n vertices is $3K_1$-covered and trivially maximizes independent sets of all sizes. Very similarly, the disjoint union of an edge and an independent set on $n-2$ vertices maximizes independent sets of all sizes among $(K_2 \cup K_1)$-covered graphs. The optimal graph for $K_{1,2}$ is given by Corollary 7. Unfortunately, $K_{2,n-2}$ is not K_3-covered, so we are unable to use the theorem of Gan, Loh, and Sudakov to resolve this last graph on 3 vertices immediately. However, it turns out that the addition of a single edge to that graph does the trick. Without considering this isolate case first, we immediately state a recent result which gives the maximum independent set counts for all $t \geq 3$ among clique-covered graphs [34].

Theorem 17 (Wang, Hou, Liu, and Ma, 2020)**.** *For any positive integers, r, t, n with $t \geq 3$ and $n \geq r$, every K_r-covered graph G on n vertices satisfies $i_t(G) \leq \binom{n-r+1}{t}$, and when $n \geq r+t-1$, $K_{r-1} + \overline{K}_{n-r+1}$ is the unique extremal graph.*

Moving to graphs on 4 vertices, we have 11 choices of H to consider. Corollary 7 and the theorem of Wang, Hou, Liu, and Ma handle four of these graphs, namely the star $K_{1,3}$, the cycle C_4, the clique K_4, and $K_4 \setminus e$ for any edge e. We note that the extremal graph for $K_4 \setminus e$ is the same as that of K_3 since $K_4 \setminus e$ is itself K_3-covered and the optimal K_3-covered graph is also $(K_4 \setminus e)$-covered. The clique K_4 is accounted for with regards to independent sets of size at least 4.Four of the other graphs contain a vertex of degree 0, and the following Lemma addresses these and any other choices of H that include an isolated vertex.

Proposition 33. *Suppose G is an H-covered graph on n vertices where $H = L \cup K_1$ for some graph L on n_1 vertices. Then for all $t \geq 2$,*

$$i_t(G) \leq i_t(L \cup (n-n_1)K_1).$$

Proof. Let G be an H-covered graph on n vertices. Then it must contain H and hence L as a subgraph. No independent set in G can use a pair of vertices that are

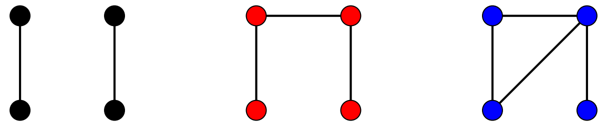

Figure 4.2: Graphs on 4 vertices for which Proposition 33 and Corollary 7 do not apply

adjacent in this copy of L. We now note that $L \cup (n - n_1)K_1$ is in fact H-saturated. Note that any set of t vertices in $L \cup (n - n_1)K_1$ that does not use a pair of adjacent vertices in the designated copy of L is in fact an independent set. Therefore $i_t(G)$ is maximized by this graph for all $t \geq 2$. $\square$

This leaves three graphs on 4 vertices for our consideration, and we display them in Figure 4.2, again coloring vertices to distinguish one graph from another. We begin with the rightmost graph in Figure 4.2, the one with blue vertices; that is, $K_{1,3} + e$ where e is an edge between any pair of vertices of degree 1 in $K_{1,3}$.

Proposition 34. *Suppose G is a $(K_{1,3} + e)$-covered graph on n vertices. Then for all $t \geq 3$,*

$$i_t(G) \leq i_t(K_{1,n-1} + e')$$

where e' is an edge between any pair of vertices of degree 1 in $K_{1,n-1}$.

Proof. Let G be a $(K_{1,3} + e)$-covered graph on n vertices. Then every vertex of G must be contained in a copy of $K_{1,3} + e$, but in particular G must contain a copy of $K_{1,3} + e$ as a subgraph. In this copy, label the vertices such that the vertex of degree 1 within this subgraph is v, the vertex of degree 3 is u, and the remaining vertices are x and y. Consider an independent set I of size $t \geq 3$ in G. I contains at most 2 vertices from the initial copy of $K_{1,3} + e$. In particular, the options for $I \cap \{u, v, x, y\}$ are the following:

$$\varnothing, \{v\}, \{u\}, \{x\}, \{y\}, \{v.x\}, \{v, y\}.$$

Now, consider another vertex $w \in V(G)$ that is not adjacent to u. We will count independent sets of size t in G that include w. Since $d(w) \geq 1$, w must have a neighbor in $\{v, u, x, y\}$ or $V(G) \setminus \{v, u, x, y, w\}$. In the latter case, there are at most $\binom{n-6}{t-1}$ independent sets of size t using w and no vertex from $\{v, u, x, y\}$, $4\binom{n-6}{t-2}$ using w and one of those vertices, and $2\binom{n-6}{t-3}$ using w and two of them.

If, on the other hand, w has a neighbor in $\{v, u, x, y\}$, then there are at most $\binom{n-5}{t-1}$ independent sets of size t using w and no vertex from $\{v, u, x, y\}$, $3\binom{n-5}{t-2}$ using w and one of those vertices, and $2\binom{n-5}{t-3}$ using w and two of them. Since this bound is greater and holds with equality when $N(w) = \{u\}$, it follows that $i_t(G)$ is maximized when $G \cong K_{1,n-1} + e$. $\square$

We now proceed to the remaining choices of H on 4 vertices, a path on 4 vertices and the disjoint union of two edges. Rather than proving results about these graphs in isolation, we will address their generalizations: paths and a disjoint union of edges.

4.3 Independent Sets in Path-Covered Graphs

Recall that we write P_k to denote a path on k vertices. Bearing in mind that the star maximizes independent set of all sizes $t \geq 3$, it is reasonable to begin our investigation of path-covered graphs by considering one with as large of a star as possible as a subgraph. In Chapter 3 we considered a class of graphs that lends itself very nicely for this purpose: the spider. In particular, consider the $(1, \ldots, 1, k-2)$ spider. This spider has $n-k+1$ legs of length 1 and a single leg of length $k-2$. By considering the path from any leaf on a leg of length 1 to the leaf on the leg of length $k-2$, we can see that this spider is in fact P_k-covered. We will now show that this graph does in fact maximize the number of independent sets of size t for all $t \geq 3$. Although we prove this for all $k \geq 4$, we prove the smallest case first as the proof strategy motivates and provides an introduction to what will be used in the general case.

Proposition 35. *Let G be a P_4-covered graph on n vertices and let S be the $(1,\ldots,1,2)$ spider with $n-3$ legs of length 1. Then for all $t \geq 3$,*

$$i_t(G) \leq i_t(S).$$

Proof. Since G is P_4-covered, there exists a copy of P_4 in G with ordered vertices v_1, v_2, v_3, v_4. Let w be another vertex in G. As in the previous section, we will consider the independent sets that contain w.

There are three cases to consider. Begin by supposing that $N(w) \cap \{v_1, v_2, v_3, v_4\} = \varnothing$. Then G contains at most $\binom{n-6}{t-1}$ independent sets of size t using w and no vertex from that set, at most $4\binom{n-6}{t-2}$ using w and one of those vertices, and at most $3\binom{n-6}{t-3}$ using w and any of the non-adjacent pairs in the set $\{v_1, v_2, v_3, v_4\}$. Using Pascal's identity, we find that the number of independent sets using G is at most

$$\binom{n-6}{t-1} + 4\binom{n-6}{t-2} + 3\binom{n-6}{t-3} = \binom{n-4}{t-1} + 2\binom{n-5}{t-2}.$$

For our second case, suppose w is adjacent to neither v_2 nor v_3 but at least one of v_1 or v_4. Then G contains at most $\binom{n-5}{t-1}$ independent sets of size t using w and none of the v_i, at most $3\binom{n-5}{t-2}$ using w and one of the v_i, and at most $\binom{n-5}{t-3}$ using a pair of vertices from that set. Again, we can rewrite our sum using Pascal's identity to obtain an upper bound of

$$\binom{n-5}{t-1} + 3\binom{n-5}{t-2} + \binom{n-5}{t-3} = \binom{n-3}{t-1} + \binom{n-5}{t-2}.$$

Lastly, we consider the case where w is adjacent to at least one of v_2 or v_3.

Without loss of generality, assume w is adjacent to v_3. Then G contains at most $\binom{n-3}{t-1}$ independent sets of size t using w but not v_2 and at most $\binom{n-4}{t-2}$ using w and v_2. Thus G has at most

$$\binom{n-3}{t-1} + \binom{n-4}{t-2}$$

independent sets of size t using w. Note that this bound is met with equality for every vertex $w \notin \{v_1, v_2, v_3, v_4\}$ in S. Furthermore, this is the largest bound provided by the three cases. Therefore $i_t(G)$ is maximized for all $t \geq 3$ when $G \cong S$. $\square$

In order to extend this to arbitrary paths, we state the following well-known and prove a lemma of our own.

Lemma 19. *[Hopkins and Staton, 1984] Let $k \geq 1$ and $0 \leq t \leq k+1$. Then*

$$i_t(P_k) = \binom{k+1-t}{t}.$$

Lemma 20. *For all $t \geq 1$ and for positive integers a and b,*

$$i_t(P_a \cup P_b) \leq i_t(P_{a+b-1} \cup K_1).$$

Proof. Let $t \geq 3$ be fixed. Label the vertices of $P_a \cup P_b$ as $v_1, \ldots, v_a, v_{a+1}, \ldots, v_{a+b}$ where $\{v_1, \ldots, v_a\}$ is the vertex set of P_a and $\{v_{a+1}, \ldots, v_{a+b}\}$ is the vertex set of P_b, each with vertices labeled in the order of the vertices on the respective paths. Label the vertices of $P_{a+b-1} \cup K_1$ as $u_1, \ldots, u_{a+b}$ where $\{u_1, \ldots, u_{a+b-1}\}$ is the vertex of set of P_{a+b-1} and u_{a+b} is the vertex for our isolated vertex.

We prove our inequality by exhibiting an injection f from the collection of independent sets of size t in $P_a \cup P_b$ to the collection of independent sets of size t in $P_{a+b-1} \cup K_1$. To this end, let $\mathcal{I}$ be an independent set in $P_a \cup P_b$. If $|\mathcal{I} \cap \{v_a, v_{a+1}\}| \leq 1$, let

$$f(I) = \{u_j : v_j \in \mathcal{I}\}.$$

Otherwise, $\{v_a, v_{a+1}\} \subseteq \mathcal{I}$. Then we define

$$f(I) = \{u_j : j \leq a-1 \text{ and } v_j \in \mathcal{I}\} \cup \{u_{a+b-j} : j \geq 0 \text{ and } v_{a+j} \in \mathcal{I}\}.$$

Each case is injective on its own and the definition of f is such that the two cases are disjoint. Therefore the number of independent sets of size t in $P_a \cup P_b$ is at most the number of independent sets of size t in $P_{a+b-1} \cup K_1$. □

We are now ready to prove Theorem 6.

Theorem 6. *Let G be a P_k-covered graph on n vertices and let S be the $(1, \ldots, 1, k-2)$ spider with $n-k-1$ legs of length 1. Then for all $t \geq 3$,*

$$i_t(G) \leq i_t(S).$$

Proof. Since G is P_k-covered, there exists a copy P of P_k in G with ordered vertices $v_1, v_2, \ldots, v_k$. Let w be another vertex in G. We will consider the independent sets that contain w, and we have two general cases to consider.

Begin by supposing that $N(w) \cap P = \varnothing$. Since $d(w) \geq 1$, there are at most $n-k-2$ non-neighbors of w that are outside of P. Let I be an independent set in G containing w. If $|I \cap P| = t-j-1$, then we can build I in at most $\binom{n-k-2}{j}\binom{k+1-(t-1-j)}{t-1-j}$ ways by choosing an independent set of size $t-1-j$ from P and the remaining j vertices from outside of P. Here the number of independent sets of size $t-1-j$ is given by Lemma 19. Summing over j, we see that the number of independent sets using w in this case is at most

$$\sum_{j=0}^{t-1} \binom{n-k-2}{j} \binom{k+1-(t-1-j)}{t-1-j}.$$

Otherwise w is adjacent to some v_c in P. Without loss of generality, we shall assume $c \geq k/2$, and we will show that the number of independent sets of size t using w is maximized when $c = k-1$. To this end, we first note that if $c = k$, then the vertices taken from P must be an independent set in a copy of P_{k-1}. On the other hand, if $c = k-1$, then the vertices from P are taking from a copy of $P_{k-2} \cup K_1$ where the disjoint copy of K_1 is the vertex v_k. Since $P_{k-2} \cup K_1$ is a proper subgraph of P_{k-1} on the same number of vertices, the former clearly contains at least as many independent sets of any fixed size as the latter.

We now compare the case where $c = k-1$ to the case where $c \leq k-2$. As previously stated, vertices in P that are used in an independent set with w are taken from $P_{k-2} \cup K_1$ when $c = k-1$. If $c \leq k-2$, then the vertices in P that are used in an independent set with w are taken from a copy of $P_{c-1} \cup P_{k-c}$ where P_{c-1} is given by the vertices $v_1, \ldots, v_{c-1}$ and P_{k-c} is given by the vertices $v_{c+1} \ldots, v_k$. Applying Lemma 20, with $a = c-1$ and $b = k-c$,

$$i_j(P_{c-1} \cup P_{k-c}) \leq i_j(P_{k-2} \cup K_1)$$

for all $0 \leq j \leq t-1$.

Since S meets the upper bound on the number of independent sets of size t using w where $N(w) \cap P = \{v_{k-1}\}$, the proof will be complete after establishing the following inequality where the summation on the right is the bound where $N(w) \cap P = \{v_{k-1}\}$.

$$\sum_{j=0}^{t-1} \binom{n-k-2}{j} \binom{k+1-(t-1-j)}{t-1-j} \le \sum_{j=0}^{t-1} \binom{n-k}{j} \binom{k-1-(t-1-j)}{t-1-j}.$$

Rather than showing this algebraically, we note that the bound on the left is the number of independent sets of size t in $P_k \cup \overline{K}_{n-k-2}$. The bound on the right is the number of independent sets of size t in $P_{k-2} \cup \overline{K}_{n-k}$. Since $P_{k-2} \cup \overline{K}_{n-k}$ is a proper, spanning subgraph of $P_k \cup \overline{K}_{n-k-2}$, every independent set in the larger graph is necessarily an independent set in its subgraph. Therefore our inequality is established, and the proof is complete. $\square$

4.4 Independent Sets in Matching-Covered Graphs

For our final choice of H, we now address the number of independent sets of size $t \ge 3$ in mK_2-covered graphs. To do this, we begin by proving Theorem 7, establishing a strong restriction on the possible structure of an optimal graph.

Theorem 7. *Let G be an mK_2-covered graph on n vertices with $m \ge 2$. Then for all $t \ge 3$ there exists a graph $H_\ell = K_{1,n-2\ell-1} \cup \ell K_2$ for some $\ell \ge m-1$ such that $i_t(G) \le i_t(H_\ell)$.*

Proof. We will first consider the case where $m = 2$. Let G be $2K_2$-covered and edge-minimal; that is, the removal of any edge in G results in a graph that is no longer $2K_2$-covered. Here G must contain a vertex-disjoint pair of edges v_1v_2 and v_3v_4. Pick these edges such that $d(v_1) + d(v_2) + d(v_3) + d(v_4)$ is maximum among such edges. Since G is edge-minimal, $e(\{v_1, v_2\}, \{v_3, v_4\}) = 0$. That is, there are no edges between the pairs. For $1 \le i \le 4$, let N_i denote the neighborhood of v_i in G, excluding the other v_j, and let $n_i = |N_i|$. Since each vertex in $N = \bigcup_{i=1}^{4} N_i$ is adjacent to some v_i, each vertex is $2K_2$-covered using one such edge and either v_1v_2 or v_3v_4. Since G is edge-minimal, we can assume then that there is exactly one such edge for each $v \in N$ and that N is an independent set.

We build a new graph G' on the same vertex set as follows. Take the graph G with v_1v_2 and v_3v_4 as chosen above. For every vertex $v \in N$, add the edge vv_1 and delete vv_i for all $i \geq 2$. Since every vertex $v \in N$ is still contained in a copy of $2K_2$ by considering the edges vv_1 and v_3v_4, and since no other edges are changed, we have that G' is also $2K_2$-covered.

We now show that $i_t(G) \leq i_t(G')$ for all $t \geq 3$. For all $J \subseteq N$ and for all $i \geq 0$, let $e_i(J)$ denote the number of independent sets I in G with $I \cap N = J$ and $|I| = |J| + i$. We define $e'_i(J)$ similarly for G'. We now prove the following claim.

Claim 1. *For all J and i, $e'_i(J) \geq e_i(J)$.*

Proof. Let $J \subseteq N$ and i be fixed. For a given J, an independent set I in G can include at most one of v_1, v_2 and at most one of v_3, v_4. If $J \neq \varnothing$, then there is some j such that any independent set I containing J in G cannot contain v_j. Thus there are at most 3 ways for I to include 1 element from $\{v_1, v_2, v_3, v_4\}$ and at most 2 ways to include two such elements. On the other hand, any independent set in G' can include v_2 and either v_3 or v_4. Thus we always have 3 ways to include 1 element from $\{v_2, v_3, v_4\}$ and 2 ways to include two such elements. The ways we can choose elements from the remaining vertices of the graph are unchanged between G and G'. Therefore $e_i(J) \leq e'_i(J)$. $\square$

By the above claim, we have that $i_t(G) \leq i_t(G')$ for all t. It follows that we can continue this process until there is no pair of disjoint edges with at least two vertices of degree greater than 1. At this point we are left with H_ℓ for some ℓ.

When $m > 2$, we take m disjoint copies of K_2. By the argument above, we can take any two disjoint copies of K_2 from the m that we are guaranteed and force the union of their neighborhoods to all be adjacent to the same vertex. In this manner, we obtain H_ℓ for some ℓ with $\ell \geq m-1$. □

This lemma tells us the general structure of the mK_2-covered graph that will maximize $i_t(G)$ among such graphs for a given t. In particular, the optimal graph is the disjoint union of a star and at least as many disjoint edges as needed to be mK_2-covered. Given this structure, the next question is with regards to which choice of $\ell \geq m-1$ results in H_ℓ having the most independent sets of size t. Is it better to have a large star with as few disjoint edges as possible, or is it better to have as many disjoint edges as possible? This requires some care, and we begin with the case where $t = 3$, with graphs on an even and odd number of vertices considered separately.

Lemma 21. *Let* $G \cong K_{1,2\alpha+1} \cup (\frac{n}{2} - \alpha - 1)K_2$ *and* $H \cong K_{1,2\alpha-1} \cup (\frac{n}{2} - \alpha)K_2$ *for some* $1 \leq \alpha < \frac{n}{2}$. *Then* $i_3(H) - i_3(G) = n - 4\alpha + 1$. *In particular,* $i_3(H) > i_3(G)$ *if and only if* $n \geq 4\alpha$.

Proof. Consider independent sets of size 3 that are in exactly one of G and H. Let uv be the extra copy of K_2 in H. Then G has an additional $n-3$ independent sets of size 3 that use both u and v. On the other hand, H has an additional $2(n-2\alpha-2)$ independent sets of size 3 using the center of the star and u or v. Thus

$$i_3(H) - i_3(G) = 2(n - 2\alpha - 2) - (n - 3) = n - 4\alpha + 1.$$

We end by noting that $i_3(H) > i_3(G)$ precisely when $n - 4\alpha + 1 > 0$; that is, when $n \geq 4\alpha$. □

Proposition 36. *Let G be an mK_2-covered graph on n vertices with $m \geq 2$ and n even. Then*

$$i_3(G) \leq i_3\left(\frac{n}{2}K_2\right).$$

Proof. Let G be mK_2-covered on n vertices with $i_3(G)$ maximum among mK_2-covered graphs. Suppose G is not a collection of disjoint edges. By Theorem 7, $G \cong K_{1,2\alpha+1} \cup (\frac{n}{2} - \alpha - 1)K_2$ for some $1 \leq \alpha \leq \frac{n}{2} - m$. By repeated application of Lemma 21,

$$i_3\left(\frac{n}{2}K_2\right) - i_3(G) = \sum_{j=1}^{\alpha}(n - 4j + 1) = \alpha(n - 2\alpha - 1).$$

This difference is greater than 0 whenever $\alpha < \frac{n}{2} - \frac{1}{2}$. Since $\alpha \leq \frac{n}{2} - m \leq \frac{n}{2} - 2$, the difference is always positive. Therefore $i_3(G) < i_3\left(\frac{n}{2}K_2\right)$, a contradiction to G having the maximum number of independent sets of size 3 among mK_2-covered graphs. Hence $G \cong \frac{n}{2}K_2$. □

With regards to the case where n is odd, we simply remark that our graph must contain $K_{1,2}$ as a subgraph.

Given this result and the fact that a disjoint union of edges will also maximize independent sets of size 2, it is reasonable to suspect that the pattern will continue. If $\frac{n}{2}K_2$ maximizes the number of independent sets of size 2 and 3 among mK_2-covered graphs for any choice of m, why not independent sets of larger size? Some computations in Sage using the 'IndependentSets' package reveal that this is not the case. Tables 4.1 and 4.2 present that the value of ℓ that optimizes $i_t(H_\ell)$ among mK_2-covered graphs for various choices of m, t, and n. We note that these choices were obtained by computing and comparing the number of independent sets of each size in these graphs using Sage.

Notice that in these examples, the optimal choice is always m or $n/2$. That is,

m	t=3	t=4	t=5	t=6	t=7	t=8	t=9	t=10
2	10	2	2	2	2	2	2	2
3	10	3	3	3	3	3	3	3
4	10	4	4	4	4	4	4	4
5	10	5	5	5	5	5	5	5
6	10	10	6	6	6	6	6	6
7	10	10	10	7	7	7	7	7
8	10	10	10	10	8	8	8	8
9	10	10	10	10	10	9/10	9	9
10	10	10	10	10	10	10	10	10

Table 4.1: Choices of ℓ that maximize $i_t(H_\ell)$ among mK_2-covered graphs on 20 vertices where $\ell \geq m-1$

m	t=3	t=4	t=5	t=6	t=7	t=8	t=9	t=10	t=11	t=12	t=13	t=14	t=15
2	15	2	2	2	2	2	2	2	2	2	2	2	2
3	15	3	3	3	3	3	3	3	3	3	3	3	3
4	15	4	4	4	4	4	4	4	4	4	4	4	4
5	15	5	5	5	5	5	5	5	5	5	5	5	5
6	15	6	6	6	6	6	6	6	6	6	6	6	6
7	15	15	7	7	7	7	7	7	7	7	7	7	7
8	15	15	8	8	8	8	8	8	8	8	8	8	8
9	15	15	9	9	9	9	9	9	9	9	9	9	9
10	15	15	15	10	10	10	10	10	10	10	10	10	10
11	15	15	15	11	11	11	11	11	11	11	11	11	11
12	15	15	15	15	12	12	12	12	12	12	12	12	12
13	15	15	15	15	15	15	13	13	13	13	13	13	13
14	15	15	15	15	15	15	15	15	15	14	14	14	14
15	15	15	15	15	15	15	15	15	15	15	15	15	15

Table 4.2: Choices of ℓ that maximize $i_t(H_\ell)$ among mK_2-covered graphs on 30 vertices where $\ell \geq m-1$

the optimal choice is either a large star with as few disjoint edges as required, or it is a disjoint union of edges.

We would like to determine if this is the case and, ideally, determine the threshold at which the optimal graph changes. We prove a generalization of Lemma 21 to point us in this direction and end the chapter with a conjecture. Recall that for a given n, we write H_ℓ to denote the graph $K_{1,n-2\ell-1} \cup \ell K_2$.

Lemma 22. *For all $t \geq 3$,*

$$i_t(H_{\ell+1}) - i_t(H_\ell) = \binom{\ell}{t-2}2^{t-1} - \sum_{j=0}^{t-2} \binom{\ell}{j}\binom{n-2\ell-3}{t-j-2}2^j.$$

Proof. Note that we can obtain $H_{\ell+1}$ from H_ℓ by deleting edges vx and vy from H_ℓ where v is the unique vertex of degree greater than 1 and x and y are two of its neighbors, and by adding the edge xy. We now consider which independent sets are gained under this transformation as well as those which are lost. In particular, the only independent sets that we gain are those which use v and one of x or y. By using v, we are forced to pick the remaining vertices from the other ℓ disjoint copies of K_2. Thus we select $t-2$ of these pairs as well as either of the two vertices from each of these pairs. Thus we gain

$$2\binom{\ell}{t-2}2^{t-2}$$

independent sets of size t.

On the other hand, we lose all independent sets of size t from H_ℓ that used both x and y. Suppose such an independent set uses j vertices from the ℓ disjoint edges. As before, we choose j of these ℓ pairs and then select a vertex from each. We then select the remaining $t-j-2$ vertices from the neighbors of v. That is, we lose

$$\binom{\ell}{j}\binom{n-2\ell-3}{t-j-2}2^j$$

independent sets of size t. We can do this for all j from 0 to $t-2$. Summing over j yields the desired difference. □

Conjecture 1. *Let G be an mK_2-covered graph on n vertices with n even. There exists a constant $t_{m,n}$ such that the following hold.*

1. $i_t(G) \leq i_t(K_{1,n-2m+1} \cup (m-1)K_2)$ *for all* $t \geq t_{m,n}$
2. $i_t(G) \leq i_t(\frac{n}{2}K_2)$ *for all* $2 \leq t < t_{m,n}$.

We are interested in determining, if this conjecture is true, what the value of this threshold $t_{m,n}$ is for all m and n, but proving this unimodality result on its own is still of interest.

Chapter 5

Future Work

We conclude with some remarks for future directions related to the problems discussed in the preceding chapters. There are many ways to proceed, but we will highlight a few specific problems that are especially intriguing.

In Chapter 2 we considered the generalized saturation problem of determining the value of $\mathrm{sat}_H(n, F)$ for various choices of H and F. This problem is still very young, and there are many other classes of graphs open for consideration. With regards to the classes considered in these pages, the following questions stand out.

Question 3. *For arbitrary r and t fixed, what value of c minimizes the lower bound on n in Theorem 5 for the existence of an n-vertex, r-partite, S_t-saturated graph?*

It would also be interesting to address the case where $n < 2t - 1$ and the optimal configuration of an S_t-saturated graph is less clear. The relationship between cherry-counting and Moore graphs is also a problem worth additional attention.

Question 4. *For $n \geq 2t - 1$ with $t \geq 2$ and $r < t$, what value(s) of m minimize $s_r(\mathrm{KR}_{t,n}(m))$? What structure minimizes $s_r(G)$ when G is S_t-saturated and has fewer than $2t - 1$ vertices?*

Question 5. *For what values of* n *is* $\mathrm{sat}_{S_2}(n, K_3) < \binom{n-1}{2}$*? That is, when does there exist a triangle-saturated graph with fewer cherries than a larger star?*

There is also a related notion not considered in our work which is that of *weak saturation*. Given graphs F and H, we say that a graph G that is a subgraph of H is weakly (F, H)-saturated if G is F-free and there exists an ordering $e_1, \ldots, e_k$ of the edges in $E(H) \setminus E(G)$ such that for all $i \in [k]$, there exists a copy of F in $G \cup \{e_1, \ldots, e_i\}$ using the edge e_i. This concept was introduced by Bollobás in [5]. We write $\mathrm{wsat}(F, H)$ to denote the minimum number of edges in a weakly (F, H)-saturated graph. Given this new setting and the problems addressed in Chapter 2, the following question is a natural one to ask.

Question 6. *For* $r < t \leq n$ *what is the value of* $\mathrm{wsat}_{K_r}(K_t, K_n)$*?*

This question is especially intriguing because $\mathrm{sat}_{K_r}(n, K_t)$ is attained by the split graph for all $r < t$ as discussed previously. In addition, $\mathrm{wsat}(K_t, K_n)$ is attained by the split graph [24]. Determining whether or not the split graph also maximizes copies of K_r for $r \geq 3$ among weakly (K_t, K_n)-saturated graphs is an exciting task. Doing this even in the case where $r = 3$ and $t = 4$ is open and of interest.

With regards to tree-saturating graphs, we would like to fully characterize which trees are tree-saturating. We completed this task for those trees which are spiders; that is, trees with a single vertex of degree greater than 2. We also provided some results concerning caterpillars that are tree-saturating, but we lack a complete characterization of them.

The natural next step would be to consider triangle-free-saturating graphs and then bipartite-saturating graphs beyond that. While these are a couple of important classes of graphs, there are many other options as well.

Finally, we turn to the topic of Chapter 4, H-covered graphs. We were interested in maximizing independent sets of fixed size $t \geq 3$ among star-covered graphs, path-covered graphs, and mK_2-covered graphs. We identified the extremal graphs for stars and paths, and we gave a list of candidates from which the optimal mK_2-covered graph resides. Unlike the other main classes of graphs we considered, we showed that the mK_2-covered graph on n vertices that maximizes independent sets of size t is dependent on n, m, and t. The most natural next step for us is to prove or disprove Conjecture 1 and, if it is true, to determine the value of the threshold $t_{m,n}$ at which the extremal mK_2-covered graph on n vertices changes. In addition to this direct continuation of our work, we can also count independent sets in H-covered graphs for other choices of H as well as other substructures in the way that we did in our generalized saturation problems considered in Chapter 2.

Bibliography

[1] N. Alon. On the number of certain subgraphs contained in graphs with a given number of edges. *Israel Journal of Mathematics*, 53:97–120, 1986.

[2] N. Alon, P. Erdős, R. Holzman, and M. Krivelevich. On k-saturated graphs with restrictions on the degrees. *J. Graph Theory*, 23:1–20, 1996.

[3] N. Alon and C. Shikhelman. Many T copies in H-free graphs. *Electronic Notes in Discrete Mathematics*, 49:683–689, 2015.

[4] B. Bollobás. On generalized graphs. *Acta Mathematica Academiae Scientiarum Hungarica*, 16:447–452, 1965.

[5] B. Bollobás. Weakly k-saturated graphs. *Beiträge zur Graphentheorie*, pages 25–31, 1968.

[6] B. Bollobás. *Extremal Graph Theory.* Academic Press, 1978.

[7] D. Chakraborti and P.-S. Loh. Extremal graphs with local covering conditions. *SIAM J. Discret. Math.*, 34:1354–1374, 2020.

[8] D. Chakraborti and P.-S. Loh. Minimizing the numbers of cliques and cycles of fixed size in an F-saturated graph. *Eur. J. Comb.*, 90:103185, 2020.

[9] J. Conde and J. Gimbert. On the existence of graphs of diameter two and defect two. *Discret. Math.*, 309:3166–3172, 2009.

[10] S. Das. A brief note on estimates of binomial coefficients. `http://page.mi.fu-berlin.de/shagnik/notes/binomials.pdf`, 2015.

[11] P. Erdős. On the number of complete subgraphs contained in certain graphs. *Magyar Tud. Akad. Mat. Kut. Intl. Kőzl*, 7:459–464, 1962.

[12] P. Erdős and A. Rényi. Asymmetric graphs. *Acta Mathematica Academiae Scientiarum Hungarica*, 14:295–315, 1963.

[13] P. Erdős, S. Fajtlowicz, and A. J. Hoffman. Maximum degree in graphs of diameter 2. *Networks*, 10:87–90, 1980.

[14] B. Ergemlidze, A. Methuku, M. Tait, and C. Timmons. Minimizing the number of complete bipartite graphs in a K_s-saturated graph. *Discussiones Mathematicae Graph Theory*, 0, 2021.

[15] R. Evans, J. Pulham, and J. Sheehan. On the number of complete subgraphs contained in certain graphs. *Journal of Combinatorial Theory, Series B*, 30(3):364–371, 1981.

[16] J. R. Faudree, R. J. Faudree, R. J. Gould, and M. S. Jacobson. Saturation numbers for trees. *Electron. J. Comb.*, 16, 2009.

[17] J. R. Faudree, R. J. Faudree, and J. R. Schmitt. A survey of minimum saturated graphs. *Electronic Journal of Combinatorics*, 1000, 2011.

[18] D. Galvin. Two problems on independent sets in graphs. *Discret. Math.*, 311:2105–2112, 2011.

[19] W. Gan, P.-S. Loh, and B. Sudakov. Maximizing the number of independent sets of a fixed size. *Combinatorics, Probability and Computing*, 24:521 – 527, 2014.

[20] A. J. Hoffman and R. R. Singleton. On Moore graphs with diameters 2 and 3. *IBM J. Res. Dev.*, 4:497–504, 1960.

[21] D. G. Hoffman and C. A. Rodger. The chromatic index of complete multipartite graphs. *J. Graph Theory*, 16:159–163, 1992.

[22] L. Kászonyi and Z. Tuza. Saturated graphs with minimal number of edges. *J. Graph Theory*, 10:203–210, 1986.

[23] J. Kritschgau, A. Methuku, M. Tait, and C. Timmons. Few H copies in F -saturated graphs. *J. Graph Theory*, 94:320–348, 2020.

[24] L. Lovász. Flats in matroids and geometric graphs. *Combinatorial Surveys*, pages 45–86, 1977.

[25] L. Lovász and M. Simonovits. On the number of complete subgraphs of a graph II. In *Studies in Pure Mathematics*, pages 459–495, 1983.

[26] D. MacKay. *Information Theory, Inference, and Learning Algorithms.* Cambridge University Press, 2003.

[27] W. Mantel. Problem 28. *Wiskundige Opgaven*, 10:60–61, 1907.

[28] R. R. Martin and J. J. Smith. Induced saturation number. *Discret. Math.*, 312:3096–3106, 2012.

[29] M. Miller, M. H. Nguyen, and G. Pineda-Villavicencio. On the nonexistence of graphs of diameter 2 and defect 2. *Journal of combinatorial mathematics and combinatorial computing*, 71:5–20, 2009.

[30] H. M. Nguyen and M. Miller. Structural properties of graphs of diameter 2 and defect 2. *AKCE International Journal of Graphs and Combinatorics*, 7:29–43, 2010.

[31] A. Sidorenko. What we know and what we do not know about Turán numbers. *Graphs and Combinatorics*, 11:179–199, 1995.

[32] N. J. A. Sloane and The OEIS Foundation Inc. The on-line encyclopedia of integer sequences, 2021.

[33] P. Turán. On an extremal problem in graph theory. *Matematikai és Fizikai Lapok (in Hungarian)*, 48:436–452, 1941.

[34] A. Wang, X. Hou, B. Liu, and Y. Ma. Maximizing the number of independent sets of fixed size in K_n-covered graphs. *arXiv*, 2020.

[35] A. A. Zykov. On some properties of linear complexes. *Mat. Sbornik N.S.*, 24(66):163–188, 1949.