Odd Covers and Graph Saturation

A Dissertation Presented

by

Calum Buchanan

to

The Faculty of the Graduate College

of

The University of Vermont

In Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy
Specializing in Mathematical Sciences

May, 2025

Defense Date: March 17, 2025
Dissertation Examination Committee:

Puck Rombach, DPhil, Advisor
Hamid R. Ossareh, Ph.D., Chairperson
Spencer Backman, Ph.D.
Gregory S. Warrington, Ph.D.
Holger Hoock, DPhil, Dean of the Graduate College

Abstract

The purpose of this dissertation is twofold: to introduce a unifying framework for what we call odd cover problems and to provide new insight into graph saturation.

An *odd cover* of a graph G is a collection of graphs such that every edge of G occurs in an odd number, and every nonedge in an even number, of graphs in the collection. We direct our interest towards finding the minimum cardinality of an odd cover of G with graphs from specific classes, in the vein of partitioning results like the Graham-Pollak theorem. Mainly, we focus on the classes of cliques and bicliques, but we also note results on odd covers with tricliques, paths (relating to Gallai's conjecture), and cycles. We find this value for various graphs G in each setting, including for all odd (and some even) cliques in the setting of bicliques, marking significant progress on a 1988 problem of Babai and Frankl. Deep relations to linear algebra are demonstrated: the minimum cardinality of an odd cover of a graph with cliques is either equal to or one more than its minimum rank over the binary field; and the minimum cardinality of an odd cover of a graph with bicliques is bounded above by the binary rank of its adjacency matrix and below by half this rank.

In Part II, we turn our attention to more extremal problems. The *saturation number* of a graph H is the minimum number of edges in an n-vertex graph which does not contain H as a subgraph, but to which the addition of any extra edge creates a copy of H. Saturation numbers of cliques were determined in 1964 by Erdős, Hajnal, and Moon, complementing one of the earliest extremal results: Turán's theorem. We prove a general lower bound on the saturation number and use it to determine the saturation numbers of unbalanced double stars asymptotically, resolving the last open cases of asymptotic saturation numbers of trees with diameter at most 3. We also provide upper bounds on the saturation numbers of certain trees of larger diameter, called caterpillars. Finally, we examine an edge-colored version of saturation, analogous to the rainbow Turán number, proving bounds on the (proper) *rainbow saturation numbers* of double stars.

Acknowledgements

I would foremost like to thank my advisor, Puck Rombach, not only for her support and advice throughout my graduate career, but for introducing me to the field of graph theory, seeing promise in me as an undergraduate, and inspiring my love for the subject. For bringing me into the fold of combinatorics, for being by my side as I journeyed into it, and for teaching me skills and perspectives I will never forget, I cannot express enough thanks.

For their contributions to the projects I discuss below, in addition to Puck Rombach, I would like to thank my collaborators Steffen Borgwardt, Neal Bushaw, Alexander Clifton, Eric Culver, Péter Frankl, Bryce Frederickson, Daniel P. Johnston, Jiaxi Nie, Jason O'Neill, Kenta Ozeki, Christopher Purcell, Mei Yin, and Youngho Yoo.

Over the course of my PhD, I have worked on a number of projects which are not included in this dissertation but have contributed to my growth as a researcher in mathematics and its applications. I would like to thank my collaborators on these projects as well: Mackenzie Carr, Rick Danner, Stephen Hartke, Paul Horn, Vesna Iršič, K. E. Perry, Brandon Du Preez, Nicholas Sieger, and Rebecca Whitman. Additionally, I have had the great privilege of applying my skills in graph theory to problems motivated by NASA applications via a Graduate Research Assistantship and a VT Space Grant Consortium Graduate Fellowship. For the opportunity to begin working on such projects, and for the incredible guidance I received along the way, I would like to thank Hamid Ossareh, James Bagrow, and Puck Rombach.

I have further had the opportunity to participate in wonderful research workshops as a graduate student, both the Graduate Research Workshop in Combinatorics and the Masamu Advanced Studies Institute of the Southern Africa Mathematical Sciences

Association. These experiences were invaluable to my growth as a researcher, and I would like to thank all of the organizers who made them possible, particularly those who guided me through them.

On a personal note, I would like to thank my family for the support they have lent throughout my life, both academic and otherwise. I would also like to thank my partner, Caitlin Seznec, who has been a constant source of joy, inspiration, comfort, and support over the course of my PhD. Finally, to all my friends, thank you for the light you bring to my life. And to Rosi (Joan Rosebush), thank you for supporting and encouraging me since the very first day of my undergraduate career. I will forever be grateful for the care and guidance you've provided.

Thank you to all of my fellow graduate students in mathematics and to my fellow union members who inspire me every day. I wish you all the best in obtaining a fair contract.

Finally, I extend my thanks to Hamid Ossareh for chairing my defense committee and to Spencer Backman and Gregory Warrington for serving on it. I am honored to be sent off by such supporting and caring faculty at UVM.

Statement of Collaboration

This dissertation would not have been possible without various collaborations, cited throughout the text. All results not cited are either unpublished results of the author or are the result of specified unpublished collaborations. The results discussed in Part I can (for the most part) be found in the papers *Subgraph complementation and minimum rank* with Christopher Purcell and Puck Rombach [19]; *Odd covers of graphs* with Alexander Clifton, Eric Culver, Jiaxi Nie, Jason O'Neill, Rombach, and Mei Yin [18]; *On odd covers of cliques and disjoint unions* with Clifton, Culver, Péter Frankl, Nie, Kenta Ozeki, Rombach, and Yin [17]; and *Path odd-covers of graphs* with Steffen Borgwardt, Culver, Bryce Frederickson, Rombach, and Youngho Yoo [15]. Many of the results in Part II can be found in *A lower bound on the saturation number and a strengthening for triangle-free graphs* with Rombach [20]. The results on rainbow saturation are due to an ongoing collaboration with Neal Bushaw, Daniel P. Johnston, and Rombach.

The dissertation author was one of the primary investigators in each of these collaborations and one of the primary authors on each of these papers.

We note that the results cited to [18] in Chapter 3 also appeared in the PhD dissertation of Eric Culver, a collaborator on this paper. They are, however, necessary to paint a complete picture of the odd cover problems we describe.

Table of Contents

List of Figures

List of Symbols

$\binom{V}{2}$	the set of all unordered pairs of distinct elements in the set V, 3
$V(G)$	vertex set of graph G, 3
$E(G)$	edge set of graph G, 3
$\|G\|$	order of graph G, 3
$\\|G\\|$	size of graph G, 3
$e(A,B)$	number of edges in a graph between disjoint subsets A and B of vertices, 4
$d(v)$	degree of vertex v, 4
$N(v)$	(open) neighborhood of vertex v, 4
$N[v]$	closed neighborhood of vertex v, 4
$G+H$	disjoint union of graphs G and H, 4
tG	disjoint union of t copies of G, 4
K_n	complete graph of order n, 4
$K_{a_1,\ldots,a_k}$	complete k-partite graph with partite sets of sizes $a_1,\ldots,a_k$, 4
P_n	path graph of order n, 5
C_n	cycle graph of order n, 5
$\mathcal{K}$	class of complete graphs, 5
$\mathcal{B}$	class of complete bipartite graphs, 5
$\mathcal{T}$	class of complete tripartite graphs, 5
$\mathcal{C}$	class of cycles, 5
$\mathcal{P}$	class of paths, 5
$O(g(n))$	function $f(n)$ with $\limsup_{n\to\infty} f(n)/g(n) < \infty$, 5
$\Omega(g(n))$	function $f(n)$ with $\liminf_{n\to\infty} f(n)/g(n) > 0$, 5
$o(g(n))$	function $f(n)$ with $\lim_{n\to\infty} f(n)/g(n) = 0$, 5
$\mathrm{cp}(G)$	clique partition number of G, 6
$\lfloor\cdot\rfloor$	floor function, 6
$\lceil\cdot\rceil$	ceiling function, 6
$\mathrm{cc}(G)$	clique covering number of graph G, 7
$A(G)$	adjacency matrix of graph G, 8
$M_{i,j}$	i,jth entry of matrix M, 8
$\tau(G)$	vertex cover number of graph G, 9
$a(G)$	arboricity of graph G, 10
$\mathrm{la}(G)$	linear arboricity of graph G, 11
$\Delta(G)$	maximum degree of graph G, 11

$\varrho_2(G,\mathcal{H})$	$\mathcal{H}$-odd cover number of graph G, 12
$\triangle$	symmetric difference operator, 12
$\mathbb{F}_2$	finite field of order 2, 13
rk_2	rank over $\mathbb{F}_2$, 13
$c_2(G)$	$\mathcal{K}$-odd cover number of graph G, 17
$G[U]$	induced subgraph of G with vertex set U, 18
$G-U$	induced subgraph of G with vertex set $V(G)-U$, 18
$\mathbb{R}$	field of real numbers, 23
$[n]$	set of positive integers at most n, 24
$\mathbb{R}^+$	set of positive real numbers, 24
$\mathbb{R}^-$	set of negative real numbers, 24
$d(G,\mathbb{F})$	minimum dimension of a faithful orthogonal representation of graph G over field $\mathbb{F}$, 24
$\mathrm{mr}(G,\mathbb{F})$	minimum rank of a symmetric matrix over field $\mathbb{F}$ which fits graph G, 26
M^{T}	transpose of matrix M, 27
$b_2(G)$	$\mathcal{B}$-odd cover number of graph G, 45
(X,Y)	biclique with partite sets X and Y, 46
(X,Y,Z)	triclique with partite sets X, Y, and Z, 46
$\oplus$	direct sum operator for matrices, 57
$A_k^{\leftrightarrow}$	adjacency matrix of a matching of size k, 57
$M_{\mathcal{O}}$	incidence matrix for a $\mathcal{B}$- or $\mathcal{T}$-odd cover, 58
T_k	universal graph for $\mathcal{T}$-odd covers, 66
B_k	universal graph for $\mathcal{B}$-odd covers, 66
$T^p(n)$	p-partite Turán graph of order n, 88
$\chi(G)$	chromatic number of graph G, 88
$\mathrm{ex}(n,H)$	extremal number of graph H, 89
$\mathrm{sat}(n,H)$	saturation number of graph H, 91
$\mathrm{ssat}(n,H)$	semisaturation number of graph H, 91
$\mathrm{wt}_0(uv)$	$\max\{d_H(x),d_H(y)\}-1$, 94
k_0	$\min\{\mathrm{wt}_0(uv) : uv\in E(H)\}$, 94
$G+xy$	graph obtained by joining vertices x and y in G, 95
$d(G)$	average degree of graph G, 96
$d(S)$	average degree over a subset S of $V(G)$, 96
$\mathrm{wt}_{\triangle}(uv)$	$\lvert N_H(u)\cap N_H(v)\rvert$, 96
$N(uv)$	$(N(u)-v)\cup(N(v)-u)$, 97
$\mathrm{wt}_1(uv)$	$\max\{d(w) : w\in N(uv), uv\in E(H)\}$, 98
k_1	$\min\{\mathrm{wt}_1(uv) : uv\in E(H)\}$, 98

k_0'	$\min\{\mathrm{wt}_0(uv) : uv \in E(H), \mathrm{wt}_1(uv) = k_1\}$, 98
k_1'	$\min\{\mathrm{wt}_1(uv) : uv \in E(H), \mathrm{wt}_0(uv) = k_0\}$, 98
$S_{s,t}$	double star with central vertices of degrees s and t, 122
P_ℓ^s	s-caterpillar of diameter $\ell - 1$, 128
T_m^k	almost k-ary tree of diameter $m - 1$, 129
$\mathrm{ex}^\star(n, H)$	rainbow Turán number of H, 135
$\mathrm{sat}^\star(n, H)$	rainbow saturation number of H, 135

Part I

Odd cover problems

Chapter 1

Introduction to odd covers

The first part of this dissertation introduces a unifying framework for a class of problems we call "odd cover problems:"

> Given a graph G and a class of graphs $\mathcal{H}$, what is the minimum number ϱ of graphs H_i in $\mathcal{H}$, $i \in \{1, \ldots, \varrho\}$, such that every edge in G is in an odd number of the H_i and every nonedge in an even number?

An odd cover of the square, C_4, with two triangles is depicted in Figure 1.1. This is a minimum odd cover of C_4 with cliques, for C_4 is not itself a clique. This is also an odd cover of C_4 with cycles, but it is not minimum in this case, for C_4 is itself a cycle.

Odd cover problems generalize well-studied graph partitioning problems, whereby

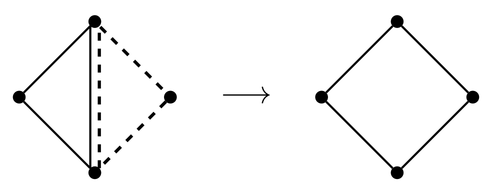


Figure 1.1: An odd cover of the square, C_4, with two triangles

the graphs $H_1, \ldots, H_\varrho$ are edge-disjoint. Tools from linear algebra are common in the study of such partitions, and we will see that they have a natural analogue in the setting of odd covers. Another well-studied class of problems, closely related to odd cover problems, are graph covering problems, whereby the graphs $H_1, \ldots, H_\varrho$ cover every edge in G at least once and do not contain any edges not present in G.

The term odd cover is inspired by the first such question in the literature, "the odd cover problem," due to Babai and Frankl in 1988 [6]. They asked for the minimum number of bicliques needed to cover every edge of the complete graph on n vertices an odd number of times. Until recently, the answer to this question was not known for any positive density subset of the integers. We will return to this problem in Chapter 3, but first, we step back to provide some basic definitions and notations.

1.1 General definitions and notations

This dissertation concerns itself with the combinatorial theory of graphs. Graphs underpin a wide variety of real-world problems, from the spread of diseases, to electrical circuits, to chemistry and molecular biology. By a *graph*, we refer to a pair of sets V and E. The former can be any finite set, whose elements we call *vertices*. The latter is a subset of $\binom{V}{2}$, the set of all unordered pairs of distinct vertices in V, whose elements we call *edges*. In the literature, these are sometimes called finite simple graphs, and the term "graph" can include infinite vertex sets, multiple edges joining the same pair of vertices, and edges whose endpoints are equal. When the vertex set or edge set of a graph G is not specified, we denote these sets respectively by $V(G)$ and $E(G)$. The *order* of G, denoted $|G|$, is the cardinality of its vertex set, and its *size* $\|G\|$ is the cardinality of its edge set. Given disjoint subsets A and B of V, we denote by $e_G(A, B)$, or simply by $e(A, B)$, the number of edges in G with one endpoint in A and the other in B.

The *degree* of a vertex v in a graph G is the number of edges incident to it (*i.e.*, the number of edges in which v occurs), denoted $d_G(v)$, or simply $d(v)$ when G is clear from context. The *open neighborhood*, or simply *neighborhood*, of v is the set of vertices adjacent to it, denoted $N_G(v)$ or $N(v)$. The *closed neighborhood* of v is $N(v) \cup v$, denoted $N_G[v]$ or $N[v]$. We give special names to those vertices whose neighborhoods are either empty or complete; an *isolated vertex* is one with no neighbors, and a *dominating vertex* v is one with $N[v] = V$. An *independent set* in G is a set of pairwise nonadjacent vertices.

A *union* of graphs G and H is the graph with vertex set $V(G) \cup V(H)$ and edge set $E(G) \cup E(H)$. If $V(G)$ and $V(H)$ are disjoint, we call it a *disjoint union*, denoted $G + H$. If we only write $G + H$, it is assumed that $V(G)$ and $V(H)$ are disjoint, and when we write $G + G$ or tG, we refer to the disjoint union of two or t copies of G, respectively.

A *graph class* is a collection of graphs sharing a specific property. We take this opportunity to define various graph classes appearing in this dissertation. Let V be a set of n vertices. The *complete graph on* V is the graph with edge set $\binom{V}{2}$. When it is not necessary to specify the vertex set, we denote an n-vertex complete graph by K_n. A graph G on V is called *bipartite* if V can be partitioned into two independent sets in G, *tripartite* if there is a partition into three independent sets, and *k-partite*, for a given positive integer k, if there is a partition into k independent sets. A *complete k-partite graph* $K_{a_1,\ldots,a_k}$ is a k-partite graph, with partite sets of cardinalities $a_1, \ldots, a_k$, which contains all edges between differing partite sets. We call a complete bipartite graph a *star* if either of its partite sets contains only one vertex. We also use the terms *cliques*, *bicliques*, and *tricliques* to refer to complete graphs,[1] complete bipartite graphs, and complete tripartite graphs, respectively. An *n-clique* is a clique of order n (not to be confused with a biclique or triclique).

A *path* P_n is a graph with edge set $\{v_1v_2, v_2v_3, \ldots, v_{n-1}v_n\}$ under some ordering of the vertices in V. The graph obtained by adding the edge v_1v_n to P_n is called a *cycle* C_n. We will make use of the notation $\mathcal{K}$ for the class of cliques, $\mathcal{B}$ for bicliques, and $\mathcal{T}$ for tricliques. The classes $\mathcal{P}$ of paths and $\mathcal{C}$ of cycles will also arise.

We will finally require notation to compare the asymptotic growth of two functions f and g of a variable n. We write $f = O(g)$ when $\limsup_{n\to\infty} \frac{f(n)}{g(n)} < \infty$ and conversely write $f = \Omega(g)$ when $\liminf_{n\to\infty} \frac{f(n)}{g(n)} > 0$. We also write $f = o(g)$ when $\lim_{n\to\infty} \frac{f(n)}{g(n)} = 0$.

1.2 A brief history of graph partitioning and covering problems

We begin with a summary of relevant results on graph partitions and coverings. By a *cover* of a graph G, we refer to a collection of graphs $H_1, \ldots, H_k$ such that $E(G) = \cup_1^k E(H_i)$. If the H_i are edge-disjoint, we call the cover a *partition* of G. A complete history of partitioning and covering problems would require more space than we have here, so we stick to those results most relevant to the odd cover problems we

[1] This is a slight abuse of terminology: a clique in a graph is technically a set of vertices which induces a complete subgraph, just as an independent set induces a graph with no edges.

study. We direct the reader to [86] for a more complete survey.

In the 1960's, a flurry of papers appeared on the subject of graph partitions and coverings. In 1966, Erdős, Goodman, and Pósa initiated the study of partitions and coverings with cliques [41].

Theorem 1.1 (Erdős-Goodman-Pósa theorem [41])**.** *Every graph G of order n can be partitioned into at most $\lfloor n^2/4 \rfloor$ cliques, and further, these can be taken to be edges and triangles.*

Sharpness of the Erdős-Goodman-Pósa theorem follows from a consideration of the biclique $K_{\lfloor n/2 \rfloor, \lceil n/2 \rceil}$ (also known as the bipartite Turán graph $T^2(n)$; see Section 4.1) which has $\lfloor n^2/4 \rfloor$ edges and no triangles. Note that this also solves the analogous problem for coverings. Győri and Tuza conjectured a stronger statement in 1987, confirmed in [69] only five years ago, that every graph of order n has a partition $H_1, \ldots, H_k$ into edges and triangles such that $\sum |H_i| \leq (1/2 + o(1))n^2$.

Let $\mathrm{cp}(G)$ denote the minimum cardinality of a *clique partition* of a graph G. While it easy to see that $\mathrm{cp}(K_n) = 1$, it is not obvious how many cliques of order strictly less than n are needed to partition K_n. The answer to this problem predates the study of clique partitions, and is known as the de Bruijn-Erdős theorem. We phrase their result in terms of clique partitions for consistency.

Theorem 1.2 (de Bruijn-Erdős theorem [34])**.** *Let n be an integer, $n \geq 3$. If $\{H_1, \ldots, H_k\}$ is a clique partition of K_n and $k \geq 2$, then $k \geq n$. Further, this bound is attained if and only if either $|H_1| = n - 1$ and $|H_i| = 2$ for $i \in \{2, \ldots, n\}$; or $n = q^2 - q + 1$, every $|H_i| = q$, and every $v \in V(K_n)$ is occurs in exactly q of the H_i.*

As a side note, when $q - 1$ is prime, the latter sharpness condition is equivalent to the existence of a projective plane of order $q - 1$. Indeed, the Fano plane provides a partition of K_7 into seven 3-cliques [25]. De Bruijn and Erdős proved their result in terms of set systems and derived as a corollary that, for any n pairwise connected points in the real projective plane, not all on a line, the number of lines is at least n.

The value of $\mathrm{cp}(G)$ is known for graphs G in various different classes. There is also a large amount of literature on the minimum cardinality $\mathrm{cc}(G)$ of a covering of G with cliques, for graphs G in various classes. As these results are less relevant to our current study, we direct the reader to [86].

The maximum difference $\mathrm{cp}(G) - \mathrm{cc}(G)$ and maximum ratio $\mathrm{cp}(G)/\mathrm{cc}(G)$ in terms of the order n of G have also been studied. For the interested reader, $\mathrm{cp}(G) - \mathrm{cc}(G) \geq \frac{n^2}{4} - \frac{n^{3/2}}{2} + \frac{n}{4}$ [23], and $\frac{n^2}{64} < \frac{\mathrm{cp}(G)}{\mathrm{cc}(G)} \leq \frac{n^2}{12}$ [40]. We demonstrate a much larger difference and ratio between $\mathrm{cp}(G)$ and the minimum cardinality of an odd cover with cliques in Corollary 2.24.

The problem of partitioning or covering the edges of a graph with a minimum number of bicliques has a rich history as well. The study of *biclique partitions* (collections of bicliques which partition the edges of a given graph) was initiated by Graham and Pollak [54] in the context of finding efficient "loop-switching" routing algorithms for the Bell System.

Theorem 1.3 (Graham-Pollak theorem [54])**.** *At least $n - 1$ bicliques are necessary to partition the edges of K_n.*

That $n - 1$ bicliques are sufficient to partition $E(K_n)$ is evident from a simple star partition, as depicted in Figure 1.2a. The Graham-Pollak theorem is a famous result in algebraic graph theory; the initial proof of the lower bound relies on the fact that there are $n - 1$ negative eigenvalues of the adjacency matrix of K_n. The

(a) A partition of K_5 into four bicliques *(b) An odd cover of K_5 with three bicliques*

Figure 1.2: A minimum partition, on the left, and a minimum odd cover, on the right, of K_5 with bicliques

adjacency matrix of a graph G, denoted $A(G)$ or simply A, is a square matrix whose rows and columns are indexed by the vertices of G. The entry $A_{i,j}$ is 1 if $ij \in E(G)$ and is 0 otherwise.[2] The adjacency matrix is an incredibly useful tool, not only for the Graham-Pollak theorem. To give a couple of examples, the entries of A^k count the number of walks of length k between pairs of vertices in G, and the eigenvalues of A provide a wealth of information about G, such as its size and whether or not it is bipartite. To this day, there is no known proof of the Graham-Pollak theorem which does not involve linear algebra in some form.

On the covering side of things, one can find much smaller collections of bicliques which cover K_n. Alon proved that the minimum number of bicliques needed to cover K_n is $\lceil \log_2 n \rceil$ [4]. Fan Chung wrote two lovely papers in 1980 and 1981 on coverings and partitions of graphs with cliques, bicliques, and forests [30, 31]. In the former paper (in which the notation $\varrho(G, \mathcal{H})$, which we will adapt for our odd cover purposes, is used for the minimum cardinality of a covering of G with graphs from $\mathcal{H}$), she proved that $\lim_{n\to\infty} \varrho(n)/n = 1$, where $\varrho(n)$ denotes the maximum value of $\varrho(G, \mathcal{B})$ over all graphs G of order n. This matches the natural upper bound afforded by the star partition of K_n, as in Figure 1.2a.

[2]"The" adjacency matrix is actually a class of matrices, but ordering the vertices does not concern us here.

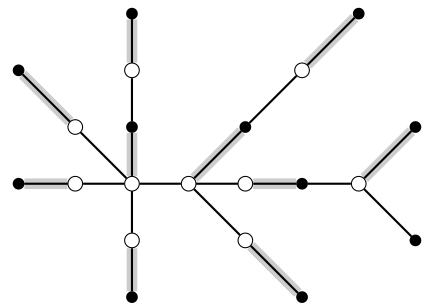


Figure 1.3: A minimum vertex cover of a tree is depicted with hollow vertices and a maximum matching with highlighted edges.

Coverings and partitions with stars are some of the oldest topics in the area, dating at least to two 1931 papers of Kőnig [62] and Egerváry [37]. A *vertex cover* of a graph G is a subset of $V(G)$ which contains at least one endpoint of every edge in G; equivalently, a vertex cover is the complement of an independent set in G. From a vertex cover U, we obtain a cover of G with stars $K_{1,d(u)}$, centered at each vertex u in U. By deleting some edges from such a star cover, we obtain a partition of $E(G)$ into stars.

In order to state the famed result known as the Kőnig-Egerváry theorem, which we will reference again in our study of biclique odd covers of trees, we define the *vertex cover number* $\tau(G)$ to be the minimum cardinality of a vertex cover of G. A *matching* in G is a set of edges which share no endpoints, and the *matching number* $m(G)$ is the maximum cardinality of a matching in G. Figure 1.3 depicts a maximum matching (in shaded edges) and minimum vertex cover (in hollow vertices) of a tree T. Note that every edge in the matching is incident to a distinct vertex in the vertex cover; indeed, $m(G)$ and $\tau(G)$ align in this case. Kőnig showed that these parameters align for all bipartite graphs.

Theorem 1.4 (Kőnig-Egerváry theorem [37,62])**.** *For any bipartite graph G, $\tau(G) =$*

$m(G)$.

Generalizing partitions with stars in a different direction than Graham and Pollak, in 1964, Nash-Williams determined the minimum number of acyclic graphs, or *forests*, needed to partition G. This is known as the *arboricity* of G, denoted $a(G)$. A certain notion of density in G provides a natural lower bound on $a(G)$: each forest in a partition contains at most $|H|-1$ edges from any subgraph H of G, and thus at least $\|H\|/(|H|-1)$ forests will be required. Nash-Williams proved that $a(G)$ is entirely determined by this density.

Theorem 1.5 ([76])**.** *For any graph G,*

$$a(G) = \max_{H \subseteq G} \left\lceil \frac{\|H\|}{|H|-1} \right\rceil .$$

Theorem 1.5 can be phrased in the setting of graphic matroids. Indeed, the natural analogue holds for the problem of partitioning the independent sets of an arbitrary matroid, as proven by Edmonds the following year [36].[3]

A forest which is *connected* (contains a path between any pair of vertices) is called a *tree*. Note that, for the complete graph K_n, Theorem 1.5 gives $a(K_n) = \left\lceil \binom{n}{2}/(n-1) \right\rceil = \lceil n/2 \rceil$. In particular, for even n, every forest in an optimal partition of K_n is an n-vertex tree, also known as a *spanning* tree. In 1978, Chung proved that every connected graph of order n has a partition into $\lceil n/2 \rceil$ trees [29].

In a different direction, rather than adding a connectivity restriction to the arboricity problem, one can restrict the tree components of a forest. A *component* of a graph is a maximal connected subgraph. The problem of partitioning a graph into a

[3]Recently, Frederickson and Michel studied circuit decompositions of Eulerian binary matroids in [49] and generalized the notion of odd covers of graphs with cycles to this setting.

minimum number of *linear forests*, forests in which every component is a path, was introduced in [1]. This minimum is called the *linear arboricity* of G, denoted $\mathrm{la}(G)$, conjectured to depend solely on the maximum degree $\Delta(G)$ of a vertex in G.

Conjecture 1.1 (Linear arboricity conjecture [1])**.** *For any graph G with maximum degree Δ, $\mathrm{la}(G) \leq \lceil(\Delta + 1)/2\rceil$.*

Alon showed that the linear arboricity conjecture holds asymptotically with Δ [3], but the conjecture remains open in general.

Adding both of the above restrictions to the arboricity problem, we obtain the problem of partitioning a graph with paths. According to a 1968 paper of Lovász [70], Erdős asked for the minimum number of paths needed to partition the edges of any connected n-vertex graph, and Gallai made the following conjecture:

Conjecture 1.2 (Gallai's conjecture)**.** *The edges of any graph G can be partitioned into at most $\lceil n/2 \rceil$ paths.*

This conjecture has been proven true for various classes of graphs, including planar graphs [11], graphs of maximum degree at most 5 [14], and graphs whose even-degree vertices induce a forest [82], yet it remains open in general. Lovász showed that $\lfloor n/2 \rfloor$ paths and cycles suffice [70] (which implies that n paths suffice). Fan proved that the analogue of Gallai's conjecture holds in the setting of path coverings [46].

In a similar vein to the result of Lovász above, a longstanding conjecture was made in [41], the same paper in which Theorem 1.1 was proved. Often called the Erdős-Gallai conjecture, it states that every graph of order n can be partitioned into $O(n)$ cycles and edges. It is well-known that every *Eulerian graph* (all vertices having even degree) can be partitioned into cycles; an equivalent conjecture states that $O(n)$ cycles suffice to partition an Eulerian graph.

We may have misled the reader in saying that the Kőnig-Egerváry theorem is one of the oldest results in the study of partitions and coverings. The study of partitioning an odd clique K_n into *Hamiltonian cycles* (spanning cycles), dates back to the *problème de ronde* posed by Lucas in 1883 [73] whose solution is attributed to Walecki in 1892. The problem is to arrange $2n + 1$ people around a single table on n successive nights so that nobody is seated next to the same person more than once. In graph theoretic terms, Walecki determined that every odd clique K_{2n+1} can be partitioned into n cycles.

1.3 Odd covers of graphs

Odd covers provide a natural generalization of graph partitions. We provide here some general definitions and observations before proceeding to analyze specific types of odd covers. Let G be a graph. We say that a collection of graphs $H_1, \ldots, H_k$ comprises an *odd cover* of G if every edge in G occurs in an odd number of the H_i and every nonedge in an even number. If the graphs H_i are all members of the same graph class $\mathcal{H}$, we call $\{H_1, \ldots, H_k\}$ an $\mathcal{H}$*-odd cover* of G. Note that an $\mathcal{H}$-odd cover of G always exists when $K_2 \in \mathcal{H}$, for we can partition G into its individual edges. The $\mathcal{H}$*-odd cover number of* G is the minimum cardinality of an $\mathcal{H}$-odd cover of G, denoted $\varrho_2(G, \mathcal{H})$.

There are a couple of different perspectives one can take on odd covers. The *symmetric difference* of two sets X and Y, denoted $X \triangle Y$, is the set $(X \cup Y) - (X \cap Y)$. Note that the symmetric difference operator is associative, and that $X_1 \triangle X_2 \triangle \cdots \triangle X_k$ consists of those elements occurring in an odd number of the X_i. For graphs H_1 and

H_2, we let $H_1 \triangle H_2$ denote the graph with vertex set $V(H_1) \cup V(H_2)$ and edge set $E(H_1) \triangle E(H_2)$. For a graph G without isolated vertices, an $\mathcal{H}$-odd cover of G is thus a set of graphs $H_1, \dots, H_k \in \mathcal{H}$ whose symmetric difference is G.

Alternatively, we might take an algebraic perspective. Let G be a graph on vertex set V. A collection of graphs $H_1, \dots, H_k$ on subsets of V is an odd cover of G if and only if $A_1 + \cdots + A_k \pmod 2 = A(G)$, where A_i denotes the adjacency matrix of the graph on V consisting of H_i and isolated vertices on $V - V(H_i)$. When the graphs in $\mathcal{H}$ have low rank over $\mathbb{F}_2$ (the finite field of order 2), the subadditivity of the rank function provides a helpful lower bound: the rank $\mathrm{rk}_2(A(G))$ of $A(G)$ over $\mathbb{F}_2$ is bounded below by the sum of the ranks of the $A(H_i)$. As we will show in Chapter 3, the resulting lower bound of $\mathrm{rk}_2(A(G))/2$ on the $\mathcal{B}$-odd cover number is sharp for many graphs G. In terms of tricliques, whose adjacency matrices also have binary rank 2, we have $\varrho_2(G, \mathcal{T}) = \mathrm{rk}_2(A(G))/2$ for *every* graph G (see Theorem 3.17). In Chapter 2, we will show that minimum odd covers with cliques also have deep algebraic roots.

1.4 A note on paths and cycles

Before diving into odd covers with cliques and bicliques, and their algebraic implications, we take a moment to address the problems of finding minimum odd covers with paths or cycles. These parameters, $\varrho_2(G, \mathcal{P})$ and $\varrho_2(G, \mathcal{C})$, are much less algebraic than the parameters $\varrho_2(G, \mathcal{K})$, $\varrho_2(G, \mathcal{B})$, and $\varrho_2(G, \mathcal{T})$, but we would be remiss not to mention some of the results obtained by the author, Steffen Borgwardt, Eric Culver, Bryce Frederickson, Puck Rombach, and Youngho Yoo in [15].

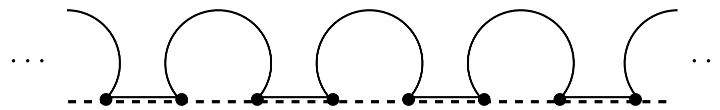


Figure 1.4: A disjoint union of cycles has an odd cover with two paths, though the minimum number of paths needed in a partition or covering can be arbitrarily large.

We recall Gallai's conjecture that the edges of any graph of order n can be partitioned into at most $\lceil n/2 \rceil$ paths. Note that this would be best possible by Theorem 1.5, since the arboricity $a(G) = \max_{H \subseteq G} \lceil \|H\|/(|H|-1) \rceil$ provides a lower bound on the number of paths in a partition, $\|P_n\| = n-1$, and $\|K_n\| = n(n-1)/2$. While $\varrho_2(G, \mathcal{H})$ can be significantly smaller than the minimum size of a path partition, as shown in Figure 1.4, each path in a $\mathcal{P}$-odd cover can still only contribute at most $|H|-1$ edges to any subgraph H of G, and thus $\varrho_2(G, \mathcal{P}) \geq a(G)$.

Two other relatively immediate lower bounds on $\varrho_2(G, \mathcal{P})$ derive from the number of odd-degree vertices in G, denoted $v_{\text{odd}}(G)$, and the maximum degree $\Delta(G)$. We note that, in a $\mathcal{P}$-odd cover of G, every path contributes at most 2 odd-degree vertices to G, even if those paths share edges. Further, every path contributes at most 2 to the degree of any vertex in G. Thus,

$$\varrho_2(G, \mathcal{P}) \geq \max\left\{ \frac{v_{\text{odd}}(G)}{2}, \left\lceil \frac{\Delta(G)}{2} \right\rceil \right\}. \tag{1.1}$$

While the minimum cardinality of a path partition can be arbitrarily far from either of these two lower bounds, we prove in [15] that $\varrho_2(G, \mathcal{P})$ is not more than a factor of two larger. In particular, one of the main results of our paper is the following theorem.

Theorem 1.6 ([15]). *For any graph G,*

$$\varrho_2(G,\mathcal{P}) \leq \max\left\{\frac{v_{odd}(G)}{2}, 2\left\lceil\frac{\Delta(G)}{2}\right\rceil\right\}.$$

Despite the constant factor of 2 on $\lceil \Delta(G)/2 \rceil$ in Theorem 1.6, we have yet to find a graph for which $\varrho_2(G,\mathcal{P})$ differs significantly from (1.1). Indeed, we do not know of any graph G with $\varrho_2(G,\mathcal{P}) > \max\{\frac{v_{\text{odd}}(G)}{2}, \lceil\frac{\Delta(G)+1}{2}\rceil\}$ [15, Problem 1].

A variant of $\varrho_2(G,\mathcal{P})$ provides an even closer relationship between path odd covers, $v_{\text{odd}}(G)$, and $\Delta(G)$. If we allow for edges in G to be *subdivided* (*i.e.*, replace uv by a path $uw_1\ldots w_k v$, $k \geq 1$), then we are able to obtain the lower bound (1.1) precisely for any graph G which is not the disjoint union of at least one a cycle with at most one path [15].

We note that the lower bound of $\lceil \Delta(G)/2 \rceil$ on $\varrho_2(G,\mathcal{P})$ is not far away from the conjectured upper bound for the linear arboricity of G (Conjecture 1.1). In fact, $\mathcal{P}$-odd covers are more closely related to linear forest partitions than one might expect at first glance. If $\mathcal{O}$ is a $\mathcal{P}$-odd cover of G with k paths, then by deleting from all paths in $\mathcal{O}$ any edges which are covered an even number of times, as well as from all but one path in $\mathcal{O}$ the edges covered an odd number of times, we obtain a collection of k linear forests which partition G. Thus, the linear arboricity of G provides yet another lower bound on $\varrho_2(G,\mathcal{P})$. We were able to prove that a second variant of $\varrho_2(G,\mathcal{P})$ aligns precisely with the linear arboricity of G in the case of Eulerian graphs. That is, the minimum value of $\varrho_2(G',\mathcal{P})$ over all graphs G' obtainable from an Eulerian graph G by adding isolated vertices is equal to the linear arboricity of G [15].

It is perhaps a surprising fact that adding isolated vertices to a graph can decrease the minimum cardinality of a $\mathcal{P}$-odd cover. However, we present examples of graphs for which this is the case in [15]. Indeed, we prove that for any odd integer k at least 3, there exists an Eulerian graph G with $\varrho_2(G,\mathcal{P}) = k+1$ for which, upon adding some number of isolates to obtain a graph G', the value $\varrho_2(G',\mathcal{P})$ drops to k. The proof of this fact used Walecki's cycle decomposition of an odd clique, discussed in Section 1.2.

We proved analogous results for these two variants of $\varrho_2(G,\mathcal{P})$ in the setting of cycle odd covers of Eulerian graphs: for some graph G' obtained by adding some number of isolates to G, $\varrho_2(G',\mathcal{C}) \leq \mathrm{la}(G')$; and, if G is not a disjoint union of two or more cycles, then $\varrho_2(G'',\mathcal{C}) = \Delta(G)/2$ for some subdivision G'' of G. The arguments used to prove Theorem 1.6 also lend themselves to $\mathcal{C}$-odd covers. We proved that $\varrho_2(G,\mathcal{C}) \leq \Delta(G)$ for every Eulerian graph G [15].

Chapter 2

Clique odd covers (subgraph complementation)

Here, we examine $\varrho_2(G, \mathcal{K})$. This parameter was introduced by the author, Christopher Purcell, and Puck Rombach in [19]. As shorthand, and in keeping with the notation (though not the terminology) we used in [19], we use $c_2(G)$ to denote the minimum number of cliques in a $\mathcal{K}$-odd cover of a graph G.

Note that taking the symmetric difference of a graph G with a clique is the same as complementing the edges and nonedges of an induced subgraph. This is known as a *subgraph complementation* of G, as defined by Kamiński, Lozin, and Milanič in [58]. Versions of this operation appeared earlier; for instance, Bouchet used successive complementations of neighborhoods of vertices to characterize the intersection graphs of chords on a circle [16]. A graph obtainable from G via successive local complementations and vertex deletions is known as a *vertex-minor* and has deep connections to rankwidth [79]. Recently, the problem of determining whether G is just one subgraph complementation away from being in a certain class, posed in [48], has also received considerable attention.

Since a graph G has $c_2(G) \leq k$ if and only if it can be obtained via at most k successive subgraph complementations from the graph on V with no edges, we originally introduced $\mathcal{K}$-odd covers under the name "subgraph complementation systems." However, we use the term odd cover here for consistency.

We came to this problem neither via a study of subgraph complementations, nor via Babai and Frankl's "odd cover problem" [6], but via a post on MathOverflow by Vincent Vatter [91]. Vatter asked whether anybody had studied the problem of "expressing the edges of a given graph as the sum of edge sets of graphs modulo 2." We came to find that this problem has deep algebraic roots, relating closely to the well-studied minimum rank problem for graphs. The results we describe have been applied in the context of invertibility of oriented graphs [10] and in the context of quantum networks [87].

2.1 Preliminary results and examples

We begin by examining some of the basic, purely combinatorial, properties of $\mathcal{K}$-odd covers. In a number of respects, the $\mathcal{K}$-odd cover number behaves nicely (in many others, it is likely to surprise you). For instance, it is monotone with respect to taking *induced subgraphs*; these are subgraphs obtained by deleting a subset of vertices from a graph, along with their incident edges. We denote by $G[U]$ the induced subgraph of G with vertex set $U \subset V(G)$. Alternatively, if $U' = V(G) - U$, we write $G - U'$ for $G[U]$. The following observation is made in passing in [19].

Proposition 2.1. *If F is an induced subgraph of a graph G, then $c_2(F) \leq c_2(G)$.*

Proof. If $\{H_1, \ldots, H_\varrho\}$ is a minimum $\mathcal{K}$-odd cover of G, and U is the subset of $V(G)$ such that $F = G[U]$, then $\{H_1[U], \ldots, H_\varrho[U]\}$ is a $\mathcal{K}$-odd cover of F. □

A graph class which is closed under taking induced subgraphs is called a *hereditary class*. Proposition 2.1 states that the class $\{G : c_2(G) \leq k\}$ is hereditary for any positive integer k. A hereditary class $\mathcal{X}$ can always be defined by a collection $\mathcal{F}$ of *forbidden induced subgraphs*; that is, $G \in \mathcal{X}$ if and only if G does not contain as an induced subgraph any graph F in $\mathcal{F}$. For instance, one can take $\mathcal{F}$ to be the set of all graphs not in $\mathcal{X}$. It is not always the case that there exists a finite set of forbidden induced subgraphs (to characterize the class of bipartite graphs, for example, one must forbid all odd cycles), but we prove this to be the case for $\{G : c_2(G) \leq k\}$ in Theorem 2.22 of Section 2.4. We explicitly state the minimum-cardinality sets $\mathcal{F}$ for $k \leq 3$, though a complete characterization seems unlikely as the cardinalities of the minimum sets of forbidden induced subgraphs grow quickly.

On the other hand, the class $\{G : c_2(G) \leq k\}$ is not closed under taking subgraphs which are not induced (consider, for instance, any n-vertex graph which is not complete as a subgraph of K_n). Further, unlike its partition and cover counterparts, $c_2(G)$ is not additive over disjoint unions.

Let W_5 denote the wheel graph on five vertices, consisting of a cycle C_4 and a dominating vertex. Figure 2.2a in Section 2.2 depicts an odd cover of W_5 with three cliques (which is the minimum, see Theorem 2.23), and every vertex of W_5 occurs in two of them. Adding two new vertices to each of the three cliques, we obtain an odd cover of $W_5 + K_2$, and thus $c_2(W_5) = c_2(W_5 + K_2) < c_2(W_5) + c_2(K_2)$. This is the smallest example of the lack of additivity of c_2 over disjoint unions.

We now state a few preliminary upper bounds on $c_2(G)$. The first, in terms of the order of G, follows from known algebraic results (see Theorem 2.7), but we include a combinatorial proof here as well, noted by Rombach in a comment on Vatter's original MathOverflow post.

Theorem 2.2 ([85])**.** *For any graph G of order n, $c_2(G) \leq n-1$.*

Proof. Let G be a graph with vertex set $V = \{v_1, \ldots, v_n\}$. We start with a clique H_1 on $N[v_1]$ and define $H_2, \ldots, H_{n-1}$ iteratively, letting G_i denote the graph for which $\{H_1, \ldots, H_i\}$ is a $\mathcal{K}$-odd cover. For each $i \in \{2, \ldots, n-1\}$, let H_i be a clique on $\{v_i\} \cup \{w \in V : v_i w \in E(G) \triangle E(G_{i-1})\}$. It is easy to check that, for every i and $j \leq i$, $N_{G_i}(v_j) = N_G(v_j)$ (indeed, every H_i contains only vertices from $\{v_i, \ldots, v_n\}$). Since every vertex in $\{v_1, \ldots, v_{n-1}\}$ has its correct neighborhood in G_{n-1}, so does v_n, and thus $G_{n-1} = G$, as desired. $\square$

Theorem 2.2, in comparison with the Erdős-Goodman-Pósa theorem from Section 1.2, marks a major distinction between $c_2(G)$ and $\mathrm{cp}(G)$: while the former is at most linear in n, the latter can be quadratic, as evidenced by $K_{\lfloor n/2 \rfloor, \lceil n/2 \rceil}$. The upper bound in Theorem 2.2 is also sharp, and holds if and only if $G = P_n$ [2] (see Theorem 2.7). Speaking of the difference between cp and c_2, we determine the maximum value of $\mathrm{cp}(G) - c_2(G)$ over all graphs G of order n to be $\lfloor n^2/4 \rfloor - 3$ in Corollary 2.24, following from the fact that $c_2(K_{3,3}) \geq 3$ (see Theorem 2.23) and Proposition 2.1.

Also in the vein of generalizing well-known results on clique partitions to the case of odd covers, one might pose the natural analogue of the de Bruijn-Erdős theorem of finding the minimum number of cliques of orders at most $n-1$ which comprise a $\mathcal{K}$-odd cover of K_n. On a trip to Budapest in 2024, the following observation arose in conversation between the author, Alexander Clifton, and Jiaxi Nie.

Proposition 2.3. *For any integer n at least 4, the cliques $[n]-1$, $[n]-2$, $\{1,2\}$, and $[n]-\{1,2\}$ comprise a $\mathcal{K}$-odd cover of the complete graph on vertex set $[n]$.*

Note that at least three cliques are necessary in a $\mathcal{K}$-odd cover of K_n with cliques of orders at most $n-1$; otherwise, the odd cover would be a partition, which would be impossible by the de Bruijn-Erdős theorem.

Proposition 2.4. *No fewer than four cliques of orders at most $n-1$ suffice to comprise a $\mathcal{K}$-odd cover of K_n when $n \geq 4$.*

Proof. By the observations above, it suffices to assume, for the sake of contradiction, that $k=3$. Since H_1, H_2, H_3 do not comprise a partition of K_n by the de Bruijn-Erdős theorem, there exists a pair of vertices u and v such that $\{u,v\} \subseteq V(H_i)$ for every i. Consider H_1 and H_2. Since $H_1 \neq H_2$, we assume, without loss of generality, that there is some vertex $i \in H_1 - H_2$. Let j be a vertex in $K_n - H_1$. If $H_1 \supset H_2$, then, in order that every edge jk is covered an odd number of times for $k \in V(K_n) - j$, we must have $H_3 = K_n$, a contradiction. Otherwise, we may assume $j \in H_2 - H_1$. In this case, we must have $ij \in E(H_3)$, and we know that $uv \in E(H_3)$, so $\{i,j,u,v\} \subseteq V(H_3)$. But now the edge iu is in both H_1 and H_3, but not H_2, a contradiction. □

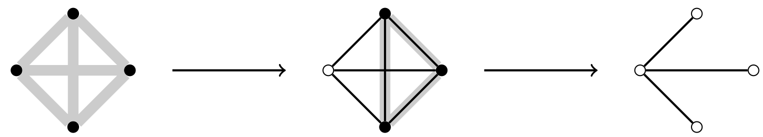
We now provide a final general upper bound on $c_2(G)$ in terms of the vertex cover number, $\tau(G)$.

Theorem 2.5 ([19])**.** *For any graph G, $c_2(G) \leq 2\tau(G)$.*

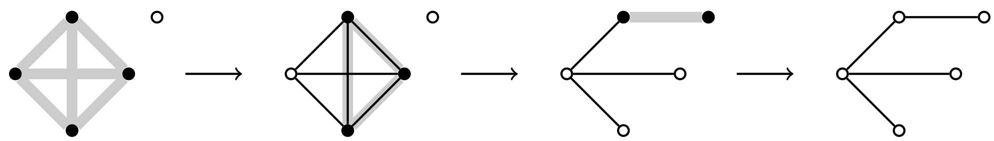
Proof. Let $U = \{u_1, \ldots, u_\tau\} \subset V$ be a minimum vertex cover of G. We iteratively construct a $\mathcal{K}$-odd cover $\mathcal{O}$ with at most 2τ cliques. We start with cliques on $N(u_1)$ and $N[u_1]$; these build the edges incident to u_1. Some of the edges incident to u_1 may also be incident to u_2. Thus, in order to obtain the remaining edges incident to u_2, we add cliques on $N(u_2) - u_1$ and $N[u_2] - u_1$ (if such a clique is empty or a singleton, we need not use it). For each $u_i \in U$, we add cliques on $N[u_i] - \{u_1, \ldots, u_{i-1}\}$ and $N(u_i) - \{u_1, \ldots, u_{i-1}\}$ to $\mathcal{O}$, thus obtaining the edges incident to u_i which have not already been constructed. Since every edge of G is incident to some vertex in U by definition of a vertex cover, and since at most two sets were needed to obtain the edges incident to each vertex in the cover, we have $c_2(G) \leq 2\tau(G)$. $\square$

Theorem 2.5 is sharp; for instance, a star has vertex cover number 1 and needs two cliques in an odd cover (see Figure 2.1a). On the other hand, if any of the sets $N(u_i) - \{u_1, \ldots, u_{i-1}\}$, $i \in \{1, \ldots, \tau\}$, in the proof of Theorem 2.5 are singletons (in which case the clique on $N[u_i] - \{u_1, \ldots, u_{i-1}\}$ is an edge), then $c_2(G) < 2\tau(G)$. By reordering, we see that the inequality is strict if there is a minimum vertex cover U of G containing a vertex with only one neighbor outside of U. For instance, if we subdivide one of the edges of $K_{1,3}$, we obtain a graph $K_{1,3}^+$ requiring two vertices in any vertex cover, the degree-3 vertex v and one other. Note that this other vertex, whichever one we choose, is incident to only one edge that isn't incident to v, and indeed, $c_2(K_{1,3}^+) = 3$ (see Figure 2.1b).

We will prove a third upper bound in terms of the size of G after noting an algebraic interpretation of $c_2(G)$ in the following section.



(a) An odd cover of $K_{1,3}$ with two cliques



(b) An odd cover of $K_{1,3}^{+}$ with three cliques

Figure 2.1: Graphs $K_{1,3}$ and $K_{1,3}^{+}$ exhibiting $c_2(G) = \tau(G)$ and $c_2(G) < \tau(G)$, respectively.

2.2 An algebraic interpretation and its implications

Let $\mathbb{F}$ be a field, and let G be a graph. A set of vectors over $\mathbb{F}$ labeled by the vertices of G is an *orthogonal representation of G* if nonadjacent vertices in G correspond to orthogonal vectors. Such representations were introduced by Lovász in 1979 [71]; when $\mathbb{F} = \mathbb{R}$ and the vectors are all of unit length, he called these orthonormal representations and used them to bound the Shannon capacity of G. An orthogonal representation of G is called *faithful* if adjacent vertices in G correspond to nonorthogonal vectors.

In 1989, Parsons and Pisanski introduced the more general notion of vector representations [80]. To quote from their paper, "vector representations. . . are of interest because they allow us to use linear algebra, the theory of bilinear forms, and geometry to study properties of the graphs being represented, and because they allow us to use these tools to construct interesting families of graphs."

Let G be a graph on vertex set $[n]$, shorthand for the set $\{1, \ldots, n\}$. Given a nondegenerate bilinear form $b : \mathbb{F}^d \times \mathbb{F}^d \to \mathbb{F}$ and subsets S, A, B, and C of $\mathbb{F}$, a *vector representation of G of dimension d* (with respect to all of these parameters) is a set of vectors $\boldsymbol{v}^{(1)}, \ldots, \boldsymbol{v}^{(n)} \in \mathbb{F}^d$ such that, for all i and j with $1 \leq i < j \leq n$,

1. the components of each $\boldsymbol{v}^{(i)}$ lie in S;
2. $b(\boldsymbol{v}^{(i)}, \boldsymbol{v}^{(i)}) \in A$ for all $i \in [n]$;
3. if $ij \in E(G)$, then $b(\boldsymbol{v}^{(i)}, \boldsymbol{v}^{(j)}) \in B$; and
4. if $ij \notin E(G)$, then $b(\boldsymbol{v}^{(i)}, \boldsymbol{v}^{(j)}) \in C$.

Note that orthogonal representations are those vector representations in which $b(\cdot, \cdot)$ is the standard dot product and $C = \{0\}$. Generalizing the notion of faithful orthogonal representations, we say that a vector representation is *faithful* is $C = \{0\}$ and $B \cap C = \varnothing$. Parsons and Pisanski were interested in the problem of finding the minimum dimension d of a vector representation over $\mathbb{R}$, where $b(\cdot, \cdot)$ is the standard dot product, A is some subset of $\mathbb{R}^+$, B some subset of $\mathbb{R}^-$, and $C = \{0\}$. The problem of minimizing the dimension of a faithful orthogonal representation over $\mathbb{R}$ was studied by Lovász, Saks, and Schrijver in [72].

Alekseev and Lozin examined the minimum dimension $d(G, \mathbb{F})$ of a vector representation of G over $\mathbb{F}$, under the standard dot product, with $S = A = \mathbb{F}$, $B = \{1\}$, and $C = \{0\}$ [2]. That is, $d(G, \mathbb{F})$ is the minimum dimension of a faithful orthogonal representation of G over $\mathbb{F}$ in which all pairs of nonorthogonal vectors have dot product 1. Note that, when $\mathbb{F} = \mathbb{F}_2$, this is simply the minimum dimension of a faithful orthogonal representation. This problem is, in fact, equivalent to that of finding $c_2(G)$, as we noted in [19].

Proposition 2.6. *For any graph G, the faithful orthogonal representations of G over $\mathbb{F}_2$ are in one-to-one correspondence with the $\mathcal{K}$-odd covers of G. In particular, we have $c_2(G) = d(G, \mathbb{F}_2)$.*

Proof. Let G be a graph. Given a collection of cliques $\{H_1, \ldots, H_d\}$ on subsets of $V(G)$, define an incidence vector $\boldsymbol{v} \in \mathbb{F}_2^d$ for each vertex v of G by $\boldsymbol{v}_i = 1$ if $v \in H_i$, and $\boldsymbol{v}_i = 0$ otherwise. Similarly, if $\{\boldsymbol{v} : v \in V(G)\} \subseteq \mathbb{F}_2^d$, we define a collection of cliques $H_1, \ldots, H_d$ by including v in H_i if and only if $\boldsymbol{v}_i = 1$. Note that $\boldsymbol{u} \cdot \boldsymbol{v} = 1$ if $\{u, v\} \subseteq V(H_i)$ for an odd number of H_i, and $\boldsymbol{u} \cdot \boldsymbol{v} = 0$ otherwise. It follows that $\{\boldsymbol{v} : v \in V(G)\}$ is a faithful orthogonal representation of G over $\mathbb{F}_2$ if and only if $\{H_1, \ldots, H_d\}$ is a $\mathcal{K}$-odd cover of G. □

We now have the luxury of borrowing results on $d(G, \mathbb{F})$ that apply to the case $\mathbb{F} = \mathbb{F}_2$. For instance, Alekseev and Lozin proved a stronger statement than Theorem 2.2.

Theorem 2.7 ([2])**.** *Let G be a graph of order n, $n > 2$, which is not a path. For any field $\mathbb{F}$ of characteristic 2, $d(G, \mathbb{F}) \leq n - 2$. Furthermore, $d(P_n, \mathbb{F}) = n - 1$.*

We can use this result to prove an upper bound on $c_2(G)$ in terms of the number of edges in G.

Theorem 2.8 ([19])**.** *For any graph G which is not a linear forest, $c_2(G) \leq \|G\| - 1$.*

Proof. Suppose that a graph G which is not a linear forest has a vertex v of degree $d(v) > 2$. The collection $\{N(v), N[v]\}$ is a $\mathcal{K}$-odd cover for the induced subgraph $G[N[v]]$. The remaining $m - d(v)$ edges of G may then be added one at a time to obtain a $\mathcal{K}$-odd cover for G of cardinality $m - d(v) + 2 \leq m - 1$.

Otherwise, G has maximum degree 2. Then G consists of disjoint cycles and paths. Since G is not a linear forest by assumption, it must contain a cycle. Theorem 2.7 completes the proof. □

We note that, if G is a linear forest, then $c_2(G) = \|G\|$. This will follow from Theorem 2.17 in Section 2.3 on forests.

The equivalence between faithful orthogonal representations over $\mathbb{F}_2$ and $\mathcal{K}$-odd covers leads to a close relationship between $c_2(G)$ and an algebraic parameter known as the minimum rank. A matrix is said to *fit* a graph G if its off-diagonal zeros match those of $A(G)$. Originally motivated by a discretization of the inverse Sturm-Liouville problem for vibrations of a string [8], the *inverse eigenvalue problem of a graph* G is to determine the sequences which can be the eigenvalues of a real symmetric matrix which fits G. Towards some understanding of this problem, many researchers have examined the problem of finding the maximum number of zeros which can be in such a sequence, that is, finding the maximum nullity of a real symmetric matrix which fits a given graph G. This, in turn, motivates a well-studied parameter known as the *zero forcing number of* G, which provides an upper bound on the maximum nullity of G. In the case of trees, the zero forcing number aligns with the maximum nullity, as well as the minimum number of components in a linear forest containing every vertex of the tree. As we will see in Theorem 2.17, this number is precisely $|G| - c_2(G)$ when G is a forest.

Equivalent to the maximum nullity, one can study the minimum rank of a matrix which fits G. We denote by $\mathrm{mr}(G, \mathbb{F})$ the minimum rank of a symmetric matrix over $\mathbb{F}$ which fits G. It has been noted, for example, in [81], that the minimum dimension of a faithful orthogonal representation of G over $\mathbb{F}$ is an upper bound on

$\text{mr}(G, \mathbb{F})$. Indeed, letting M be an $n \times d$ matrix whose rows are the vectors in a faithful orthogonal representation of G, we see that $MM^\intercal$ fits G, where $M^\intercal$ denotes the transpose of M. It is an easy exercise to show that the rank of $MM^\intercal$ is bounded above by the rank of M, from which the observation follows. By Proposition 2.6, when the field in question is $\mathbb{F}_2$, such a matrix M is an incidence matrix for a $\mathcal{K}$-odd cover of G with d cliques.[1] The following proposition was also noted in [19].

Proposition 2.9. *For any graph G, $c_2(G) \geq \text{mr}(G, \mathbb{F}_2)$.*

There are graphs, such as $K_{a,b}$ when $a, b \geq 3$, for which $c_2(G) > \text{mr}(G, \mathbb{F}_2)$. However, a result of Lempel from 1975 tells us these two parameters cannot be far apart.[2]

Theorem 2.10 (Lempel's lemma [68])**.** *Let A be a symmetric matrix over $\mathbb{F}_2$ of rank r. The minimum number of columns in a matrix M such that $MM^\intercal = A$ is $r+1$ if $A_{i,i} = 0$ for all i, and otherwise is r.*

Corollary 2.11 ([19, Corollary 12 and Theorem 13])**.** *For any graph G, $c_2(G) \in \{\text{mr}(G, \mathbb{F}_2), \text{mr}(G, \mathbb{F}_2) + 1\}$. Further, $c_2(G) = \text{mr}(G, \mathbb{F}_2) + 1$ if and only if $A(G)$ has minimum rank over all matrices fitting G over $\mathbb{F}_2$ and is the unique such matrix.*

Corollary 2.11 provides a characterization of the binary matrices of minimum rank which fit graphs with $c_2(G) > \text{mr}(G, \mathbb{F}_2)$. We are also able to provide a combinatorial perspective on the minimum $\mathcal{K}$-odd covers of such graphs. In particular, we will show

[1] As an aside, the rows of M also provide incidence vectors for a collection of sets whose intersection graph is G, an observation used by Erdős, Goodman, and Pósa in [41] along with their theorem on $\text{cp}(G)$ to prove that every graph is the intersection graph of a collection of sets with at most $\lfloor n^2/4 \rfloor$ elements each.

[2] We were unaware of Lempel's lemma when we wrote [19], in which we proved Corollary 2.11 using a result of Friedland [50].

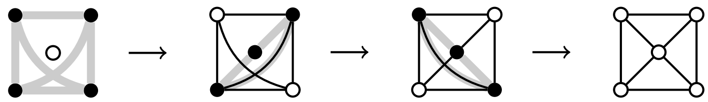
that they are the graphs possessing a minimum $\mathcal{K}$-odd cover in which every vertex occurs an even number of times. We require two lemmas to prove this.

Lemma 2.12 ([19]). *Let G be a graph with vertex set V, and let $\mathcal{O}$ be a $\mathcal{K}$-odd cover of G. If, for some vertex $v \in V$, every $u \in V - v$ occurs in an even number of cliques in $\mathcal{O}$, then the collection $\{H \triangle \{v\} : H \in \mathcal{O}\}$, is also a $\mathcal{K}$-odd cover of G.*

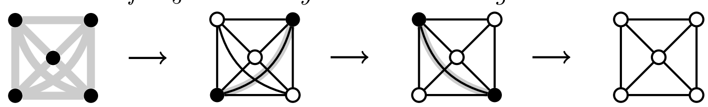
Proof. Suppose that $\mathcal{O}$ is a $\mathcal{K}$-odd cover of G such that every vertex in G, aside from possibly one vertex v, occurs in an even number of cliques in $\mathcal{O}$. Let $\mathcal{O}_v$ denote the collection of symmetric differences of $\{v\}$ with each clique in $\mathcal{O}$; *i.e.*, $\mathcal{O}_v = \{H \triangle \{v\} : H \in \mathcal{O}\}$. For any $u \in V - v$, if u and v occur together in an odd number of cliques in $\mathcal{O}$, then u is occurs without v in an odd number of cliques in $\mathcal{O}$, so u and v occur together an odd number of times in $\mathcal{O}_v$. Similarly, if u and v occur together an even number of times in $\mathcal{O}$, then u occurs an even number of times without v in $\mathcal{O}$, and thus an even number of times with v in $\mathcal{O}_v$. Also, any two vertices which are distinct from v occur together the same number of times in $\mathcal{O}_v$ as in $\mathcal{O}$. In other words, $\mathcal{O}_v$ is also a $\mathcal{K}$-odd cover for G, as desired. □

In a particular case of Lemma 2.12, if every vertex of G occurs in an even number of cliques in a $\mathcal{K}$-odd cover $\mathcal{O}$, then for any $v \in V$, the collection $\mathcal{O}_v$ is also a $\mathcal{K}$-odd cover for G. For example, Figure 2.2 depicts two $\mathcal{K}$-odd covers of the wheel graph W_5, related by Lemma 2.12.

Lemma 2.13 ([19]). *Let G be a graph with $c_2(G)$ even, and let $\mathcal{O}$ be a minimum $\mathcal{K}$-odd cover for G. Then there exists a vertex $v \in V$ such that v occurs in $\mathcal{O}$ an odd number of times.*



(a) A $\mathcal{K}$-odd cover $\mathcal{O}$ of W_5 with every vertex occurring in an even number of cliques



(b) The $\mathcal{K}$-odd cover $\{K \triangle \{v\} : K \in \mathcal{O}\}$, where v is the central vertex

Figure 2.2: An example of Lemma 2.12

Proof. Let G and $\mathcal{O}$ be as described. Suppose, for the sake of contradiction, that every vertex of G occurs an even number of times in $\mathcal{O}$. Let K be a clique in $\mathcal{O}$ on $\{u_1, \ldots, u_s\}$. Then $\mathcal{O}_{u_1}$ is a minimum $\mathcal{K}$-odd cover of G by Lemma 2.12. Furthermore, $\mathcal{O}_{u_1}$ maintains the property that every vertex occurs an even number of times. We can continue this process to find that $\mathcal{O}_{u_1,u_2} = (\mathcal{O}_{u_1})_{u_2}$ also maintains that property, and so on. Then $\mathcal{O}_{u_1,\ldots,u_s}$ is a minimum $\mathcal{K}$-odd cover of G, but it contains an empty clique (*i.e.*, a clique on $\varnothing$). This contradicts the minimality of $\mathcal{O}$. $\square$

If $\mathcal{O}$ is a $\mathcal{K}$-odd cover of odd cardinality in which every vertex occurs an even number of times, then the vertex v in Lemma 2.12 occurs an odd number of times in the $\mathcal{K}$-odd cover $\mathcal{O}_v$. Together with Lemma 2.13, we see that, for any graph G, there exists a minimum $\mathcal{K}$-odd cover in which some vertex occurs an odd number of times. We now show that this is the case for every minimum $\mathcal{K}$-odd cover if and only if $c_2(G) = \mathrm{mr}(G, \mathbb{F}_2)$.

Theorem 2.14 ([19])**.** *Let G be a graph with at least one edge. We have $c_2(G) > \mathrm{mr}(G, \mathbb{F}_2)$ if and only if G has a minimum $\mathcal{K}$-odd cover in which every vertex of G occurs an even number of times.*

Proof. Let G be a graph with at least one edge, and let $k = \mathrm{mr}(G, \mathbb{F}_2) > 0$. We begin by proving sufficiency. Suppose that there exists a minimum $\mathcal{K}$-odd cover $\mathcal{O}$ of G in which every vertex occurs an even number of times. Let M be the incidence matrix for $\mathcal{O}$. Each row of M contains an even number of 1's, so the columns of M are linearly dependent. Thus,

$$k \leq \mathrm{rk}(MM^{\mathsf{T}}) \leq \mathrm{rk}_2(M) < c_2(G).$$

Concerning the necessary condition, suppose that $c_2(G) \neq k$. Then $c_2(G) = k+1$ by Corollary 2.11. Further, the adjacency matrix A of G is the unique matrix which fits G and has minimum rank over $\mathbb{F}_2$. By Lempel's lemma, $A = MM^{\mathsf{T}}$ for some $n \times (k+1)$ matrix M of rank k over $\mathbb{F}_2$. Since $A_{i,i} = \sum_j M_{i,j}^2 = 0$ for every i, every row of M has an even number of 1's. Thus, M is an incidence matrix for a $\mathcal{K}$-odd cover for G in which every vertex occurs an even number of times, as desired. □

Note that we can slightly strengthen the converse of Theorem 2.14: if some minimum $\mathcal{K}$-odd cover $\mathcal{O}$ of G contains an odd subcollection in which every vertex in G occurs an even number of times, then the corresponding columns of the incidence matrix M for $\mathcal{O}$ are dependent. Thus, in this case, $\mathrm{mr}(G, \mathbb{F}_2) \leq \mathrm{rk}_2(M) < c_2(G)$.

In analyzing which graphs have $c_2(G) = \mathrm{mr}(G, \mathbb{F}_2) + 1$, we are able to restrict our attention to the class of connected graphs. We shall presently prove that, in order to determine whether a graph has $c_2(G) = \mathrm{mr}(G, \mathbb{F}_2)$, it suffices to determine whether any of its components have this property.

Theorem 2.15 ([19])**.** *A disconnected graph G has $c_2(G) > \mathrm{mr}(G, \mathbb{F}_2)$ if and only if, for every component G' of G, $c_2(G') > \mathrm{mr}(G', \mathbb{F}_2)$.*

Proof. Let $G = G_1 + \cdots + G_t$. If $\mathrm{mr}(G, \mathbb{F}_2) \neq c_2(G)$, by Corollary 2.11, the adjacency matrix $A = A(G)$ is the unique matrix of minimum rank which fits G over $\mathbb{F}_2$. Suppose, for the sake of contradiction, that there exists a component of G, say G_1, such that $\operatorname{mr}(G_1, \mathbb{F}_2) = c_2(G_1)$. Notice that every matrix which fits G is a block-diagonal matrix; let $A = \oplus_1^t A_i$ where $A_i = A(G_i)$. Furthermore, the rank of a block-diagonal matrix is minimized by minimizing the ranks of its blocks, so that $\operatorname{rk}_2(A_i) = \operatorname{mr}(G_i, \mathbb{F}_2)$ for each $i \in \{1, \ldots, t\}$. By Theorem 2.14, there exists a minimum $\mathcal{K}$-odd cover $\mathcal{O}$ for G_1 in which some vertex occurs in an odd number of cliques. Let $M = M(\mathcal{O})$ be the matrix associated to $\mathcal{O}$. Then MM^T fits G_1, is of rank $\operatorname{mr}(G_1, \mathbb{F}_2)$, and has some nonzero diagonal entry. We may thus replace A_k by MM^T to obtain a matrix fitting G of minimum rank over $\mathbb{F}_2$ with a nonzero diagonal entry, a contradiction.

On the other hand, if $\operatorname{mr}(G_i, \mathbb{F}_2) \neq c_2(G_i)$ for every $i \in \{1, \ldots, t\}$, then, for each i, the adjacency matrix $A_i = A(G_i)$ is the unique matrix of minimum rank over $\mathbb{F}_2$ which fits G_i. Thus, there is a unique matrix fitting G over $\mathbb{F}_2$ of minimum rank, and it is $\oplus_1^t A_i$. By Corollary 2.11, we have $c_2(G) \neq \operatorname{mr}(G, \mathbb{F}_2)$, as desired. □

Let us summarize our various characterizations of the graphs with $c_2(G) \neq \operatorname{mr}(G, \mathbb{F}_2)$.

Theorem 2.16 ([19]). *For any graph G with at least one edge, the following are equivalent:*

1. $c_2(G) \neq \operatorname{mr}(G, \mathbb{F}_2)$;
2. $c_2(G) = \operatorname{mr}(G, \mathbb{F}_2) + 1$;
3. *there is a unique matrix of minimum rank which fits G over $\mathbb{F}_2$, and it is $A(G)$;*
4. *every vertex in G occurs in an even number of cliques in some minimum $\mathcal{K}$-odd cover of G;*

5. for every component G' of G, $c_2(G') = \mathrm{mr}(G', \mathbb{F}_2) + 1$.

2.3 Forests

We now determine the value of $c_2(F)$ for an arbitrary forest F. Not only does $c_2(F)$ align with $\mathrm{mr}(F, \mathbb{F}_2)$, but it can be calculated using another combinatorial parameter which we will presently describe.

Let $p(F)$ denote the minimum cardinality of a set of vertex-disjoint paths in F which partition $V(F)$. Such a set of paths has been called a "path cover" in the literature, but we refrain from using this terminology to avoid confusion with the covering problems previously described. We prefer to think of $p(F)$ as the minimum number of components in a spanning linear forest in F.

Theorem 2.17 ([19]). *For any forest F of order n,*

$$c_2(F) = \mathrm{mr}(F, \mathbb{F}_2) = n - p(F).$$

To prove Theorem 2.17, we require two lemmas which reduce the problem of finding the minimum rank of a matrix (over an arbitrary field) which fits a forest F to that of finding the value of $p(F)$, which admits a relatively simple algorithm.

In general, there is no straightforward relationship between the minimum rank of a graph over differing fields. For example, the full house graph (depicted in Figure 2.4 of Section 2.4) has minimum rank 3 over $\mathbb{F}_2$, but minimum rank 2 over any other field. On the other hand, the triclique $K_{3,3,3}$ has minimum rank 2 over $\mathbb{F}_2$, but $\mathrm{mr}(K_{3,3,3}, \mathbb{R}) = 3$. Since $\mathrm{mr}(G, \mathbb{F})$ is additive over components, we can take disjoint unions of such examples to obtain graphs G in which $\mathrm{mr}(G, \mathbb{F}_2)$ and $\mathrm{mr}(G, \mathbb{R})$ are arbitrarily far apart. For trees, however, the parameters $\operatorname{mr}(G, \mathbb{F}_2)$ and $\operatorname{mr}(G, \mathbb{R})$ coincide.

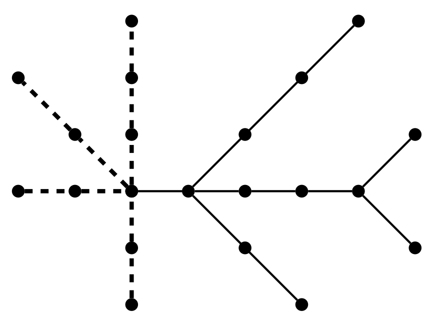
Lemma 2.18 ([28]). *The minimum rank of a forest is independent of the field.*

Lemma 2.19 ([57]). *For any tree T of order n, $\operatorname{mr}(T, \mathbb{R}) = n - p(T)$.*

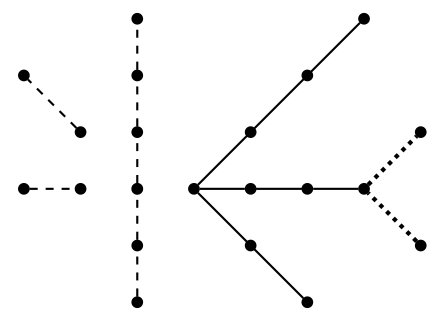
To prove Theorem 2.17, it now suffices to show that $c_2(T) \leq n - p(T)$ for an arbitrary tree T of order n, since $c_2(G) \geq \operatorname{mr}(G, \mathbb{F}_2)$ by Corollary 2.11, and the minimum rank function is additive over disjoint unions. We do so using an algorithm of Fallat and Hogben [45] for finding a spanning linear forest of T with $p(T)$ components. The algorithm can be briefly summarized as follows, and an example is depicted in Figure 2.3. If T is a *spider* (has at most one vertex of degree at least 3), take a maximal path in T (which passes through the high-degree vertex if it exists) and all remaining paths. Otherwise, T has a *pendent spider*: a spider subgraph S such that $T - S$ is disconnected. Take a maximal path P in S through this high-degree vertex (or take $P = S$ if S is itself a path). If $T - P$ is disconnected, then take each of the resulting path components as well (these are subgraphs of S). Repeat the process on $T - S$ until there are no more high-degree vertices left.

We use the resulting spanning linear forest of T, in addition to the technique used to prove Theorem 2.5, to prove Theorem 2.17.

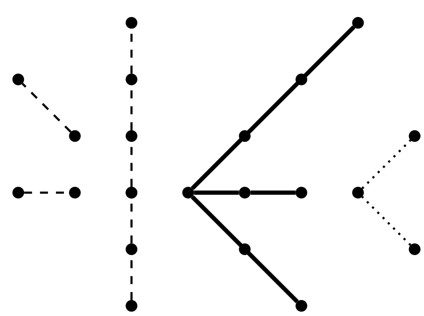
Proof of Theorem 2.17. By the additivity of the minimum rank function, we may assume F is a tree T. By Corollary 2.11, it suffices to show that $c_2(T) \leq n - p(T) = \operatorname{mr}(T, \mathbb{F}_2)$. Above, we described an algorithm from [45] for finding a spanning linear forest of T with the minimum number of components, $p(T)$. (See Figure 2.3 for an example.) In the linear forest L the algorithm outputs, every path contains at most



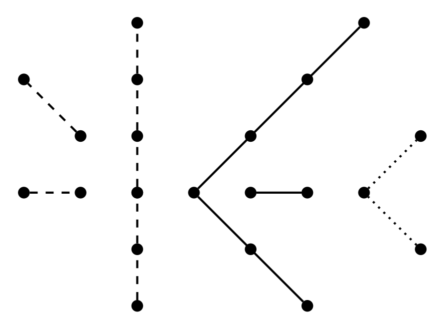
(a) A pendent spider S_1 in a tree T is depicted with dashed edges



(b) A maximal path in S_1, long with two edges, form three disjoint paths. A pendent spider in $T - S_1$ is depicted with dotted edges.



(c) The tree $T-(S_1 \cup S_2)$ is a spider, depicted with bolded edges.



(d) A spanning linear forest of T with six paths

Figure 2.3: Obtaining a minimum number of paths which partition the vertices of a forest via an algorithm of Fallat and Hogben [45]

one vertex of degree 3 or more in T, and these vertices are never endpoints of the paths in which they lie. If L also covers all of the edges of T, then T is a path and Theorem 2.7 completes the proof. Otherwise, any edges which are not in L are incident to high-degree vertices, which are internal in their respective paths. Denote these high-degree vertices by $v_1, v_2, \ldots, v_k$, and define $U = \{v \in V(T) : d_T(v) \leq 2\}$. Let $\mathcal{O}$ be the collection consisting of the $\|L\| - 2k$ sets of the form $\{u, v\}$ where $u, v \in U$ and $uv \in E(L)$, along with the sets $N_T(v_i)$ and $N_T[v_i]$ for $1 \leq i \leq k$. Then $\mathcal{O}$ is a $\mathcal{K}$-odd cover of T with cardinality $\|L\|$. Note that $\|L\| = n - p(T)$. Therefore, $c_2(T) \leq n - p(T) = \mathrm{mr}(T, \mathbb{F}_2)$, which completes the proof. □

2.4 Forbidden induced subgraphs

As discussed in Proposition 2.1 of Section 2.1, the class of graphs G with $c_2(G) \leq k$ is hereditary for every nonnegative integer k. Recall that every hereditary graph class can be defined by a collection of forbidden induced subgraphs. Here, we examine these collections. We note that, aside from some terminology and notational changes, much of this section is quoted directly from [19].

We saw in Proposition 2.9 that $\mathrm{mr}(G, \mathbb{F}_2) \leq c_2(G)$ for every graph G, and thus

$$\{G : c_2(G) \leq k\} \subseteq \{G : \mathrm{mr}(G, \mathbb{F}_2) \leq k\}. \tag{2.1}$$

It is known that the class of graphs $\{G : \mathrm{mr}(G, \mathbb{F}) \leq k\}$ is hereditary and finitely defined when $\mathbb{F}$ is a finite field [35]. For odd k, it follows from Corollary 2.11 that if $\mathrm{mr}(G, \mathbb{F}_2) = k$, then $c_2(G) = k$, and if $\mathrm{mr}(G, \mathbb{F}_2) < k$, then $c_2(G) \leq k$. Therefore, when k is odd, we also have $\{G : c_2(G) \leq k\} \supseteq \{G : \mathrm{mr}(G, \mathbb{F}_2) \leq k\}$.

Proposition 2.20 ([19]). *For any odd k,*

$$\{G : c_2(G) \leq k\} = \{G : \mathrm{mr}(G, \mathbb{F}_2) \leq k\}.$$

In particular, the classes $\{G : c_2(G) \leq k\}$ and $\{G : \mathrm{mr}(G, \mathbb{F}_2) \leq k\}$ for odd k are defined by the same finite set of minimal forbidden induced subgraphs. The two minimal forbidden induced subgraphs for $k = 1$ are evident, as a graph with $c_2(G) \leq 1$ consists of a single clique and/or isolated vertices. The class of graphs $\{G : c_2(G) \leq 1\}$ is then the class of $\{P_3, 2K_2\}$-free graphs. We obtain as a corollary to Proposition 2.20 that the set of minimal forbidden induced subgraphs for the property $c_2(G) \leq 3$ is the same set given in the following theorem and listed explicitly in [9].

Theorem 2.21 ([9]). *The class of graphs $\{G : \mathrm{mr}(G, \mathbb{F}_2) \leq 3\}$ is defined by forbidding a set of 62 minimal induced subgraphs, each of which has 8 or fewer vertices.*

On the other hand, when k is even, it does not follow from Proposition 2.20 that $\{G : c_2(G) \leq k\}$ is finitely defined.

Theorem 2.22 ([19]). *For any natural number k, the class of graphs $\{G : c_2(G) \leq k\}$ is defined by forbidding a finite set of induced subgraphs.*

Proof ([19]). Let F be a minimal forbidden induced subgraph for the property $c_2(G) \leq k$. First, we claim that $c_2(F) \leq k + 2$. Suppose, for the sake of contradiction, that $c_2(F) \geq k + 3$. Then, for any $v \in V(F)$ and $\mathcal{K}$-odd cover $\mathcal{O}'$ for $F - v$, we have that $\mathcal{O} = \mathcal{O}' \cup \{N(v), N[v]\}$ is a $\mathcal{K}$-odd cover for F, which implies that $c_2(F - v) \geq k + 1$. This contradicts the minimality of F.

Now, there exists a $\mathcal{K}$-odd cover $\mathcal{O}$ for F of cardinality $k + 2$. We can associate to F a vector of length $s = 2^{k+2}$, where each entry corresponds to an element of the powerset $2^{\mathcal{O}}$, such that each entry of the vector is a non-negative integer that counts the number of vertices of F that are in a given subcollection of $\mathcal{O}$. This vector defines the graph F up to isomorphism. It is easy to verify that, if two graphs F_a and F_b have vectors $(a_1, \ldots, a_s)$ and $(b_1, \ldots, b_s)$ such that $a_i \leq b_i$ for $1 \leq i \leq s$, then F_a is an induced subgraph of F_b. We now see that the poset of forbidden induced subgraphs for the property $c_2(G) \leq k$ ordered by the induced subgraph relation can be embedded in the poset $\mathbb{N}^s$, which is the direct product of the poset $\mathbb{N}$ ordered by $\leq$. It is known that a direct product of finitely many posets that are well-founded and that have no infinite anti-chains is itself well-founded and has no infinite anti-chains [63]. Furthermore, any restriction of such a poset has the same properties. This completes the proof to show that the poset of forbidden induced subgraphs for the property $c_2(G) \leq k$, ordered by the induced subgraph relation, is well-founded with a finite number of minimal elements. □

Theorem 2.22 only guarantees that the set of minimal forbidden induced subgraphs for the property $c_2(G) \leq k$ is finite; it does not provide an explicit upper bound. Based on the results concerning linear forests, we present the following conjecture.

Conjecture 2.1 ([19]). *A minimal forbidden induced subgraph for the property $c_2(G) \leq k$ has at most $2k+2$ vertices.*

By analyzing the structure of graphs with $c_2(G) \leq 2$, we can find the set of minimal forbidden induced subgraphs for this property. This is the set given in Theorem 2.23 and depicted by the set A in Figure 2.4.

Theorem 2.23 ([19]). *The class of graphs $\{G : c_2(G) \leq 2\}$ is the class of $\mathcal{F}$-free graphs, where $\mathcal{F}$ is the set A of graphs in Figure 2.4.*

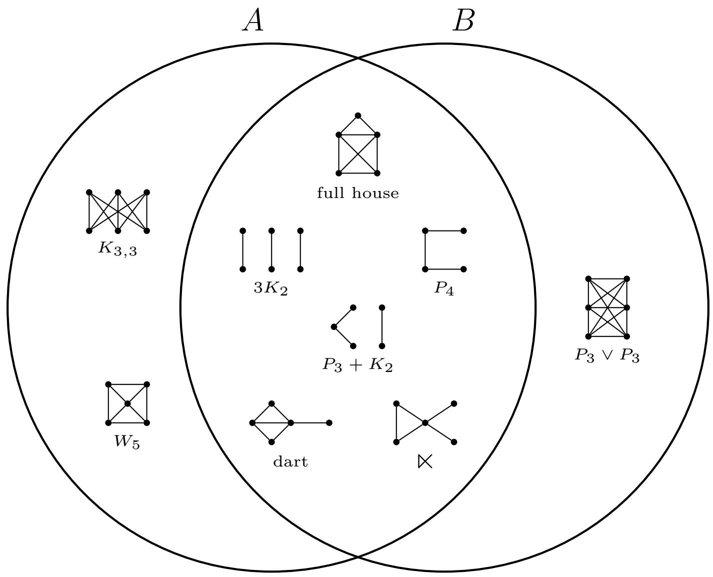


Figure 2.4: The sets of minimal forbidden induced subgraphs for the properties $c_2(G) \leq 2$ *(A) and* $\mathrm{mr}(G, \mathbb{F}_2) \leq 2$ *(B).*

Proof ([19]). Suppose, for the sake of contradiction, that there exists a graph $G = (V, E)$ such that $c_2(G) > 2$, and G does not contain any element of $\mathcal{F}$ as an induced subgraph. Furthermore, suppose that G is minimal with these qualities; that is, every proper induced subgraph H of G has $c_2(G) \leq 2$. Then G has no isolated vertices. Furthermore, $|G| \geq 5$ by Theorem 2.7.

The rest of the proof is outlined as follows. We show that there exists a vertex x for which $c_2(G - x) = 2$. Letting $\mathcal{O} = \{C_1, C_2\}$ be a minimum $\mathcal{K}$-odd cover for $G - x$, depicted in Figure 2.5, we then show that $C_1 \cap C_2$ is nonempty, and that $G - x$ has no isolated vertices. Finally, we split into two cases: either one of the sets in $\mathcal{O}$ contains the other, or not. Contradictions are derived by showing that either $c_2(G) \leq 2$ or G contains an induced subgraph in $\mathcal{F}$.

Firstly, we claim that there exists a vertex x for which $c_2(G - x) = 2$. We have $c_2(G - v) \geq c_2(G) - 2 \geq 1$ for all $v \in V$, since we can add $N(v)$ and $N[v]$ to any minimum $\mathcal{K}$-odd cover for $G - v$ to obtain one for G. Furthermore, if $c_2(G - v) = 1$ for all $v \in V$, then $\mathrm{mr}(G - v, \mathbb{F}_2) = 1$ for all $v \in V$, so G is a minimal forbidden induced subgraph for the property $\mathrm{mr}(G, \mathbb{F}_2) \leq 1$. These are the graphs P_3 and $2K_2$, which both have $\mathcal{K}$-odd covers of cardinality 2, which proves the claim.

Let $\mathcal{O} = \{C_1, C_2\}$ be a minimum $\mathcal{K}$-odd cover for $G - x$. Notice that both $|C_1| \geq 2$ and $|C_2| \geq 2$. We begin by showing that $C_1 \cap C_2$ is nonempty. Suppose $C_1 \cap C_2 = \emptyset$. The isolated vertices of $G - x$ are a subset of $N_G(x)$. If every neighbor of x is isolated in $G - x$, then G has an induced $3K_2$ or $P_3 + K_2$. Thus, x has a neighbor in at least one of C_1 and C_2. Without loss of generality, say x has a neighbor in C_1. Then x dominates C_1, otherwise G has an induced $P_3 + K_2$ (if x has no neighbor in C_2), or an induced P_4 (otherwise). If x has no neighbor in C_2, then either $c_2(G) \leq 2$, or G has an induced $P_3 + K_2$. In fact, x dominates C_2, otherwise G has an induced P_4. Then either $c_2(G) \leq 2$, or G has an induced $\ltimes$, a contradiction. Therefore, $C_1 \cap C_2$ is nonempty.

Suppose there exists an isolated vertex in $G - x$. Then, for each edge uv of $G - x$, either both or neither of u and v are neighbors of x, otherwise G has an induced P_4. If there are at least two isolated vertices, then for each edge uv of $G - x$, exactly one of u and v is a neighbor of x, otherwise G has an induced $P_3 + K_2$ or an induced $\ltimes$. We conclude there is exactly one isolated vertex in $G - x$. If x has no other neighbor, then G has an induced $P_3 + K_2$, since C_1 and C_2 are not disjoint. Without loss of generality, say x has a neighbor in C_1. In fact, we can conclude that x dominates C_1, otherwise G has an induced P_4. Then x has a neighbor in $C_1 \cap C_2$, so x dominates C_2 as well, and G has an induced dart. Therefore, $G - x$ has no isolated vertices.

Figure 2.5 represents a minimum $\mathcal{K}$-odd cover $\mathcal{O} = \{C_1, C_2\}$ of $G - x$. Without

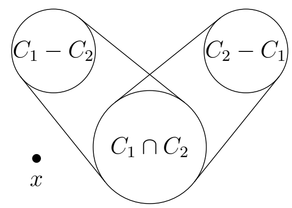


Figure 2.5: A $\mathcal{K}$-odd cover of $G - x$

loss of generality, we assume that $|C_1| \leq |C_2|$. One may imagine $G - x$ as disjoint cliques $C_1 - C_2$ and $C_2 - C_1$, and an independent dominating set $C_1 \cap C_2$. We now split into cases: either $C_1 - C_2$ and $C_2 - C_1$ are both nonempty, or $C_1 \subset C_2$. The former case is divided into subcases differentiating between the possible neighborhoods of x in G.

Case 1 ($C_1 - C_2$ and $C_2 - C_1$ both nonempty). Throughout Case 1, vertices in $C_1 - C_2$ are be denoted by $u = u_0, u_1, u_2, \ldots$, vertices in $C_1 \cap C_2$ by $w = w_0, w_1, w_2, \ldots$, and vertices in $C_2 - C_1$ by $z = z_0, z_1, z_2, \ldots$.

Suppose $N(x) \subseteq C_1 - C_2$. Then G has an induced P_4 on vertex set $\{x, u, w, z\}$, where $u \in N(x)$, $w \in C_1 \cap C_2$, and $z \in C_2 - C_1$. A similar contradiction is derived if $N(x) \subseteq C_2 - C_1$.

Suppose $N(x) \subseteq C_1 \cap C_2$, and let $w = w_0 \in N(x)$. If $|C_2 - C_1| \geq 2$, say $z_0, z_1 \in C_2 - C_1$, then G contains an induced $\ltimes$ on vertex set $\{x, z_0, z_1, w, u\}$, where $u \in C_1 - C_2$. Otherwise, since $|C_2| \geq |C_1|$ by assumption, $|C_1 - C_2| = |C_2 - C_1| = 1$. Let $C_1 - C_2 = \{u\}$, and let $C_2 - C_1 = \{z\}$. Since $|G| \geq 5$, we have $|C_1 \cap C_2| \geq 2$. If x has a non-neighbor in $C_1 \cap C_2$, say w_1, then G has an induced P_4 on $\{x, w_0, u, w_1\}$.

Otherwise, $N(x) = C_1 \cap C_2$. If $C_1 \cap C_2 = \{w_0, w_1\}$, then there is a $\mathcal{K}$-odd cover of G of cardinality 2: $\{\{x, w_0, u, z\}, \{x, w_1, u, z\}\}$. Thus, there exist vertices $w_0, w_1, w_2 \in N(x) \cap (C_1 \cap C_2)$, and G has an induced $K_{3,3}$ on $\{x, u, z, w_0, w_1, w_2\}$.

Suppose x has neighbors $u = u_0 \in C_1 - C_2$ and $w = w_0 \in C_1 \cap C_2$, but no neighbor in $C_2 - C_1$. Let $z \in C_2 - C_1$. If x has a non-neighbor $u_1 \in C_1 - C_2$, then G has an induced dart on $\{x, u, u_1, w, z\}$, and if x has a non-neighbor $w_1 \in C_1 \cap C_2$, then G has an induced P_4 on $\{x, u, w_1, z\}$. Thus, $C_1 - C_2 \subset N(x)$, and $C_1 \cap C_2 \subset N(x)$. Since $G - x$ has no isolated vertices, we have $N(x) = C_1$. But then G has a $\mathcal{K}$-odd cover of cardinality 2: $\{C_1 \cup \{x\}, C_2\}$. Thus, we arrive at a contradiction when x has neighbors in $C_1 - C_2$ and $C_1 \cap C_2$ but not $C_2 - C_1$. By similar arguments, we derive a contradiction if x has neighbors in $C_2 - C_1$ and $C_1 \cap C_2$ but none in $C_1 - C_2$.

Finally, suppose x has neighbors $u = u_0 \in C_1 - C_2$, $w = w_0 \in C_1 \cap C_2$, and $z = z_0 \in C_2 - C_1$. Since $|G| \geq 5$ and $|C_1| \leq |C_2|$, either $|C_1 \cap C_2| \geq 2$ or $|C_2 - C_1| \geq 2$. Suppose $|C_2 - C_1| \geq 2$. If x has a non-neighbor $z_1 \in C_2 - C_1$, then G has an induced P_4 on $\{u, x, z_0, z_1\}$. Otherwise, x dominates $C_2 - C_1$, and G has an induced full house on $\{u, x, w, z_0, z_1\}$. Thus, $|C_2 - C_1| = |C_1 - C_2| = 1$, and $|C_1 \cap C_2| \geq 2$. If x has 2 or more neighbors in $C_1 \cap C_2$, say $w_0, w_1 \in N(x) \cap C_1 \cap C_2$, then G has an induced W_5 on $\{x, u, w_0, w_1, z\}$. Thus, x has a non-neighbor w_1 in $C_1 \cap C_2$. Suppose $C_1 \cap C_2 = \{w_0, w_1\}$. Since $C_1 - C_2 = \{u\}$ and $C_2 - C_1 = \{z\}$, the two cliques on $\{x, u, w_0, z\}$ and $\{w_1, u, z\}$ comprise a $\mathcal{K}$-odd cover of G. Now suppose that $|C_1 \cap C_2| \geq 3$; say $w_0, w_1, w_2 \in C_1 \cap C_2$. We have seen that w_0 is the only neighbor of x in $C_1 \cap C_2$. Thus, G contains an induced $\ltimes$ on $\{x, u, w_0, w_1, w_2\}$. We conclude that x must not have neighbors in each of $C_1 - C_2$, $C_1 \cap C_2$, and $C_2 - C_1$. This is a contradiction, which concludes Case 1.

Case 2 ($C_1 - C_2 = \varnothing$; that is, $C_1 \subset C_2$). Let $u_0, u_1 \in C_1$ and $z = z_0 \in C_2 - C_1$. If $N(x) \subsetneq C_1$, say $u_0 \in N(x)$ and $u_1 \in C_1 - N(x)$, and if $z \in C_2 - C_1$, then G has an induced P_4 on $\{x, u_0, z, u_1\}$. If $N(x) = C_1$, then G has a $\mathcal{K}$-odd cover of cardinality 2: $\{C_2, C_1 \cup \{x\}\}$. Thus, x has a neighbor $z \in C_2 - C_1$. If $u_0 \in N(x)$ but $u_1, u_2 \in C_1$ are not neighbors of x, then G has an induced $\ltimes$ on $\{x, u_0, u_1, u_2, z\}$. If x has neighbors $u_0, u_1 \in C_1$, and a non-neighbor $u_2 \in C_1$, then G has an induced dart on $\{x, u_0, u_1, u_2, z\}$. Thus, x dominates C_1. If x also dominates C_2, then G has a $\mathcal{K}$-odd cover of cardinality 2: $\{C_1, C_2 \cup \{x\}\}$. Thus, x has a neighbor z_0 and a non-neighbor z_1 in $C_2 - C_1$, and G has an induced W_5 on $\{x, u_0, u_1, z_0, z_1\}$. This completes the proof of Case 2 and the proof of the theorem.

□

As a corollary, we see that the forbidden induced subgraphs for the property $c_2(G) \leq 2$ themselves have $c_2(G) = 3$.

Corollary 2.24. *The maximum difference* $\mathrm{cp}(G) - c_2(G)$ *over all graphs* G *of order* n *is* $\lfloor n^2/4 \rfloor - 3$.

Proof. The balanced biclique $G = K_{\lfloor n/2 \rfloor, \lceil n/2 \rceil}$ has $\mathrm{cp}(G) - c_2(G) = \lfloor n^2/4 \rfloor - 3$. The only graphs with $c_2(G) = 1$ are cliques plus some isolates, but these also have $\mathrm{cp}(G) = 1$. If $c_2(G) = 2$, then $\mathrm{cp}(G) < \lfloor n^2/4 \rfloor$, since no such graph contains $K_{3,3}$ as an induced subgraph, and the balanced biclique uniquely maximizes $\mathrm{cp}(G)$ in terms of n. The result follows. □

The graph $K_{\lfloor n/2 \rfloor, \lceil n/2 \rceil}$ also provides a lower bound on the maximum value of the ratio $\mathrm{cp}(G)/c_2(G)$.

Corollary 2.25. *The maximum value of* $\operatorname{cp}(G)/c_2(G)$ *over all graphs* G *of order* n *is at least* $\lfloor n^2/12 \rfloor$.

We leave open the problem of determining upper bounds on $\operatorname{cp}(G)/c_2(G)$.

Chapter 3

Biclique and triclique odd covers & a problem of Babai and Frankl

We now shift our attention from the class $\mathcal{K}$ of cliques to the class $\mathcal{B}$ of bicliques. We recall Graham and Pollak's celebrated result in algebraic graph theory, Theorem 1.3 of Section 1.2, that the complete graph K_n can be partitioned into no fewer than $n-1$ bicliques [54]. In their 1988 book, "Linear Algebra Methods in Combinatorics," Babai and Frankl posed the generalization of finding the minimum number of bicliques required to cover every edge of K_n an odd number of times [6]. In our notation, this is the value $\varrho_2(K_n, \mathcal{B})$. They posed as an exercise to show that $\varrho_2(K_n, \mathcal{B}) \geq \lfloor n/2 \rfloor$ (which will follow from Proposition 3.6), remarking that the precise value is unknown. It does not take long to find $\mathcal{B}$-odd covers of K_n, $n \geq 5$, with fewer than $n-1$ bicliques (see Figure 1.2 in Section 1.2), and in fact we will see that $\lfloor n/2 \rfloor + 1$ cliques always suffice in Theorem 3.31. Radhakrishnan, Sen, and Vishwanathan determined that $\varrho_2(K_n, \mathcal{B}) = n/2$ whenever $n = 2(q^2 + q + 1)$ for a prime power $q \equiv 3 \pmod 4$, or whenever there exists an $n \times n$ Hadamard matrix [83]. In Section 3.5, we make

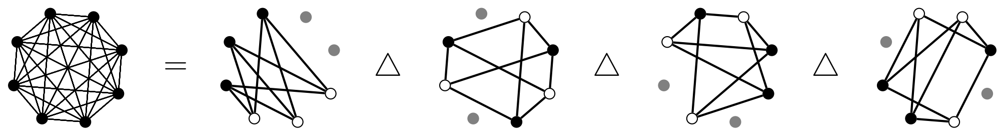


Figure 3.1: A $\mathcal{B}$-odd cover of K_8 with four bicliques. Partite sets are depicted by black and hollow vertices; gray vertices are not in either partite set.

significant progress on this problem, determining $\varrho_2(K_n, \mathcal{B})$ precisely when n is odd, a multiple of 8 (see Figure 3.1), or equivalent to 18 (mod 24). First, we analyze the more general problem of finding $\varrho_2(G, \mathcal{B})$, as many of the techniques we use for K_n apply more generally.

The results presented in this chapter are due to two collaborations with a large number of authors. In addition to the dissertation author, the former collaboration [18] comprised Alexander Clifton, Eric Culver, Jiaxi Nie, Jason O'Neill, Puck Rombach, and Mei Yin. The latter [17] comprised many of the same authors; the symmetric difference of the author sets is Jason O'Neill, Péter Frankl, and Kenta Ozeki. In keeping with the notation that we used in these papers, we write $b_2(G)$ for the parameter $\varrho_2(G, \mathcal{B})$.

We first began considering the parameter $b_2(G)$ as a variation of $c_2(G)$, before we knew about Babai and Frankl's problem. Our collaborator on the project discussed in the previous chapter, Christopher Purcell, pointed us towards another MathOverflow post, this time of Niel de Beaudrap [33], in which the problem of finding $b_2(G)$ was posed. The dissertation author brought this problem to the 2021 Graduate Research Workshop in Combinatorics, where we began our collaboration on [18].

There is a deep (algebraic) relationship between odd covers with bicliques and odd covers with tricliques, as we will see in Section 3.3. To make our lives easier, we allow our multipartite graphs to contain empty partite sets; in particular, bicliques also live in the class of tricliques. If a biclique has an empty partite set, then it has no edges at all, but this notion will still prove convenient (for instance, in the proof of Proposition 3.1). We denote a biclique with partite sets X and Y by (X, Y) and a triclique with partite sets X, Y, and Z by (X, Y, Z).

3.1 Preliminary results

We begin with a simple statement, analogous to Proposition 2.1.

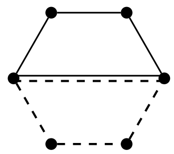
Proposition 3.1. *The class of graphs* $\{G : b_2(G) \leq k\}$ *is hereditary.*

Proof. We apply the same argument we used to prove the corresponding statement for c_2. Let $\{H_1, \ldots, H_\varrho\}$ be a minimum $\mathcal{B}$-odd cover of G, and let $U \subseteq V(G)$. Each induced subgraph $H_i[U]$ for $i \in \{1, \ldots, \varrho\}$ is a biclique, and it is easy to check that $\{H_1[U], \ldots, H_\varrho[U]\}$ is an odd cover of $G[U]$. □

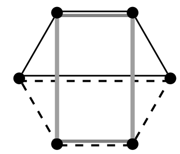
If G has *twin vertices* u and v, meaning that $N(u) = N(v)$, then deleting v from every biclique in which it occurs in a minimum $\mathcal{B}$-odd cover of G, as in the proof of Proposition 3.1, actually produces a minimum $\mathcal{B}$-odd cover of $G - v$.

Proposition 3.2. *If* u *and* v *are twin vertices in a graph* G*, then* $b_2(G - u) = b_2(G - v) = b_2(G)$.

Proof. That $b_2(G-u)$ and $b_2(G-v)$ are bounded above by $b_2(G)$ follows from Proposition 3.1. On the other hand, given a minimum $\mathcal{B}$-odd cover of $G-v$, let us add v to every partite set containing u. The edge uv is in none of the resulting bicliques, and the edge vw occurs in a biclique if and only if uw occurs. Thus, we have here a $\mathcal{B}$-odd



(a) An odd cover of C_6 with two bicliques



(b) An odd cover of $2K_3$ with three bicliques.

Figure 3.2: Minimum odd covers of C_6 and $2K_3$

cover of G of cardinality $b_2(G-v)$, so $b_2(G) \leq b_2(G-v)$. Similarly, $b_2(G) \leq b_2(G-u)$, which completes the proof. $\square$

For example, consider the square C_4. The two pairs of nonadjacent vertices in C_4, $\{x_1, x_2\}$ and $\{y_1, y_2\}$, are both twin pairs. Taking the $\mathcal{B}$-odd cover $\{(x_1, y_1)\}$ of K_2 and adding x_2 to x_1's partite set and y_2 to y_1's partite set, we obtain an odd cover of C_4 with one biclique; indeed, $C_4 = K_{2,2}$. Larger even cycles do not contain twin vertices, so we cannot apply this trick, but they do all possess minimum $\mathcal{B}$-odd covers with copies of C_4. We exhibit a construction of such an odd cover below (and in Figure 3.2a), and later prove that it is optimal in Corollary 3.7.

Proposition 3.3. *For any positive integer n, $b_2(C_{2n}) \leq n-1$.*

Proof. Let $V(C_{2n}) = \{x_1, \ldots, x_n, y_1, \ldots, y_n\}$ and $E(C_{2n}) = \{x_i x_{i+1}, y_i y_{i+1} : i \in [n-1]\} \cup \{x_1 y_1, x_n y_n\}$. For each $i \in [n-1]$, define $X_i = \{x_i, y_{i+1}\}$ and $Y_i = \{x_{i+1}, y_i\}$. Then $\{(X_i, Y_i) : i \in [n-1]\}$ is an odd cover of C_{2n} with squares, as in Figure 3.2a. $\square$

It follows from Proposition 3.1 that, for any component G' of a disconnected graph G, $b_2(G') \leq b_2(G)$. As in the case of $\mathcal{K}$-odd covers (and all the other types of odd covers we've encountered), however, the parameter b_2 is not additive over the components of G. For example, while it is clear that $b_2(K_3) = 2$, we have $b_2(2K_3) \leq 3$, as depicted in Figure 3.2b.

The odd cover of $2K_3$ depicted in Figure 3.2b also generalizes, but in a less obvious way than the one for C_6. We will generalize the following result yet again in Theorem 3.30 when we determine b_2 for an arbitrary disjoint union of cycles.

Proposition 3.4 ([18])**.** *For any positive integer t, $b_2(tK_3) \leq t+1$.*

Proof. Let the vertex set of tK_3 be $\{u_i, v_i, w_i : i \in [t]\}$, where each $\{u_i, v_i, w_i\}$ is a triangle. For each $i \in [t]$, define $X_i = \{v_i, w_i\}$ and $Y_i = \{v_j : j \neq i\} \cup u_i$. Note that $\{(X_1, Y_1), \ldots, (X_t, Y_t)\}$ is an odd cover of the graph G with edges $u_i v_i$ and $u_i w_i$ for $i \in [t]$ and $v_i w_j$ for $\{i, j\} \in \binom{[t]}{2}$. Now, define $X_{t+1} = \{v_i : i \in \{1, \ldots, t\}\}$ and $Y_{t+1} = \{w_i : i \in \{1, \ldots, t\}\}$. The edges in common between G and (X_{t+1}, Y_{t+1}) are the edges $v_i w_j$ for $\{i, j\} \in \binom{[t]}{2}$, and the edges present in (X_{t+1}, Y_{t+1}) which are not in G are those of the form $v_i w_i$ for $i \in [t]$. Thus, $G \triangle (X_{t+1}, Y_{t+1}) = tK_3$, which completes the proof. $\square$

We now construct odd covers of $G+G$, for any graph G of order n, using at most n bicliques.

Theorem 3.5 ([17])**.** *For any graph G of order n, $b_2(G+G) \leq n$.*

Proof. Let $\{u_1, \ldots, u_n\}$ and $\{v_1, \ldots, v_n\}$ be the vertex sets of two copies of G, denoted G_u and G_v, where $u_i u_j \in E(G_u)$ if and only if $v_i v_j \in E(G_v)$.

Claim. There exists a $\mathcal{B}$-odd cover $\{(X_i, Y_i) : i \in [n]\}$ of $G_u + G_v$, where $X_i = \{u_i, v_i\}$ for every i.

We prove the claim by induction. The base case, $n = 1$, is trivial. Suppose $n \geq 2$. Let $G'_u = G_u - u_n$ and $G'_v = G_v - v_n$. By induction, there is a $\mathcal{B}$-odd cover of $G'_u + G'_v$ with $n-1$ bicliques satisfying the claim. Let $\{(X'_i, Y'_i) : i \in [n-1]\}$ be such an odd cover. We extend this to an odd cover of $G_u + G_v$ as follows: let X_i be as defined in the claim for all $i \in [n]$. For each $i \in [n-1]$, we add u_n to Y_i' if and only if $u_i u_n$ in $E(G_u)$. That is, $Y_i = Y_i' \cup u_n$ if $u_i u_n \in E(G)$, and $Y_i = Y_i'$ otherwise. Clearly, the bicliques $(X_1, Y_1), \dots, (X_{n-1}, Y_{n-1})$ construct all of the edges in $G_u' + G_v'$, as well as the edges $u_i u_n \in E(G_u)$. We have also constructed the "wrong" edges $v_i u_n$ in place of the edges $v_i v_n$ we want. Let $Y_n = \{v_i : v_i v_n \in E(G_v)\}$. Note that (X_n, Y_n) consists of precisely these edges, as well as the missing edges $v_i v_n \in E(G_v)$. This proves the claim, and thus proves the theorem. □

We have now seen that the parameter $b_2(G)$ is not additive over disjoint unions. Surprisingly, we can even find graphs G and H for which $b_2(G+H) = c_2(G)$. We posed the question in [17] of determining whether, for any graph H, one can find a graph G such that $b_2(G) = b_2(G+H)$. We answer the question in the affirmative for $H = K_2$ or $H = K_3$. Figure 3.3 depicts a graph $G+K_2$ where $b_2(G+K_2) = b_2(G) = 4$. We have checked computationally that $b_2(G) = 4$, and an odd cover of $G + K_2$ with four bicliques $(X_1, Y_1), \dots, (X_4, Y_4)$ is encoded in the words which label the vertices. An ε in the ith entry of a word indicates that the vertex is not in the ith biclique, a 0 indicates that the vertex is in X_i, and a 1 indicates that the vertex is in Y_i. For example, the words 000ε and 1000 which label the endpoints of the isolated edge xy tell us that, say, x is in X_1, X_2, and X_3, but not in (X_4, Y_4), and y is in Y_1, X_2, X_3, and X_4. Similarly, Figure 3.4 depicts a graph $G+K_3$ having $b_2(G+K_3) = b_2(G) = 5$. We have checked computationally that no graph G with $b_2(G) < 4$ and $b_2(G+K_2) = b_2(G)$ exists, and neither does there exist a graph with $b_2(G) < 5$ and $b_2(G+K_3) = b_2(G)$.

Another interesting class of disjoint unions to consider would be those of the form $G + \bar{G}$, where $\bar{G}$ denotes the graph on $V(G)$ with edge set $\binom{V(G)}{2} - E(G)$. There is

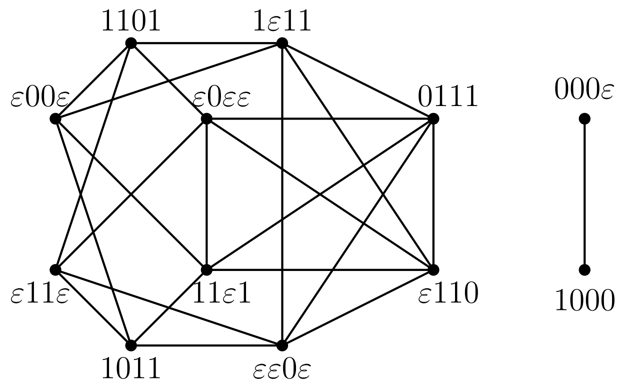


Figure 3.3: The 10*-vertex component,* G*, on the left satisfies* $b_2(G) = 4$*, and* $b_2(G{+}K_2) = 4$*.*

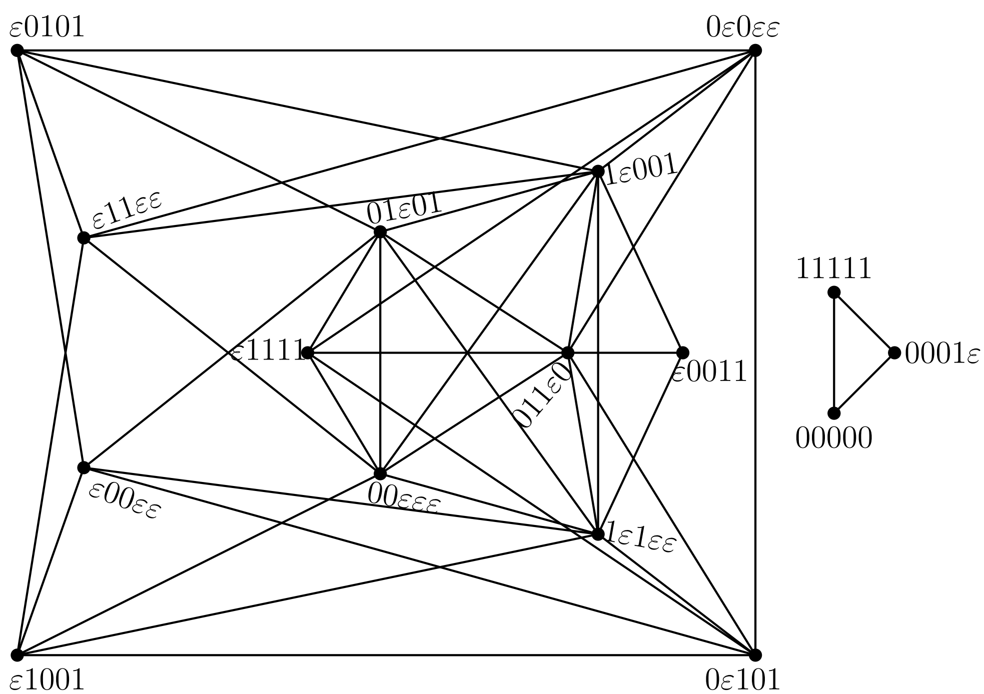


Figure 3.4: The 13*-vertex component,* G'*, on the left satisfies* $b_2(G') = 5$*, and* $b_2(G'{+}K_3) =$ 5*.*

a large amount of literature on what are called Nordhaus-Gaddum type inequalities. We pose one such question in [17]: is it always the case that, for a graph G of order n, $b_2(K_n) \leq b_2(G + \bar{G}) \leq n$? All the evidence we have at the moment suggests that this might be the case.

Note that each of the upper bounds on $b_2(G)$ we have seen thus far are no larger than $|G|/2 + 1$. In fact, we have yet to discover a graph G of order n for which $b_2(G) > n/2 + 1$. This is in sharp contrast to the minimum cardinality of a biclique partition, which can be as large as $n - 1$ by the Graham-Pollak theorem. We would be very interested to find a sharp upper bound on b_2 in terms of n (or, indeed, any upper bound in terms of n which does not hold for biclique partitions).

Problem 3.1 ([17]). *Does there exist an $\varepsilon > 0$ such that, for all graphs G of sufficiently large order n, $b_2(G) \leq (1 - \varepsilon)n$?*

To prove upper bounds on the minimum cardinality of an odd cover, we find constructions such as the ones given in Propositions 3.3 and 3.4. The lower bounds we obtained for c_2 were related to the algebraic parameter $\mathrm{mr}(G, \mathbb{F}_2)$. In the case of b_2, we also obtain an algebraic lower bound, thanks to a close relation to the rank of the adjacency matrix of G over $\mathbb{F}_2$. We denote $\mathrm{rk}_2(A(G))$ simply by $\mathrm{rk}_2(G)$, and sometimes refer to rank over $\mathbb{F}_2$ simply as rank, when the context is clear.

It is a well-known (perhaps folklore) result that the rank of any symmetric matrix over $\mathbb{F}_2$ with zero diagonal is even (see, for example, [59, p. 22]). That is, $\mathrm{rk}_2(G)$ is even for every graph G.

Proposition 3.6 ([18]). *For any graph G, $b_2(G) \geq \mathrm{rk}_2(G)/2$.*

Proof. Let $\{H_1, \ldots, H_\varrho\}$ be a minimum odd cover of G. To each H_i, we add isolated vertices on $V(G) - V(H_i)$ to obtain a graph H_i' on $V(G)$. As we saw in Section 1.3,

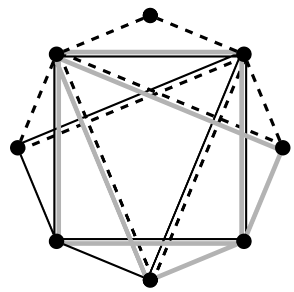


Figure 3.5: A minimum $\mathcal{B}$-odd cover of C_8 using three $K_{2,4}$s

letting $A_i = A(H_i)$ for each i, we have $\sum_1^{\varrho} A_i \pmod 2 = A(G)$. By the subadditivity of rank, and since the adjacency matrix of a nondegenerate biclique has rank 2, we have $\mathrm{rk}_2(G) \geq \sum_{i=1}^{\varrho} \mathrm{rk}_2(H_i) \geq 2\varrho = 2b_2(G)$. □

This simple bound provides the backbone for a large number of the values of b_2 we determine. For instance, $\mathrm{rk}_2(C_{2n}) = 2n - 2$, and thus the construction in Proposition 3.3 is optimal. Figure 3.5 depicts an odd cover of C_8 with three copies of $K_{2,4}$, a very different construction than the one in Proposition 3.3.

Corollary 3.7 ([18])**.** *For any positive integer n, $b_2(C_{2n}) = n - 1$.*

In Section 3.2, we prove that not only C_{2n}, but all bipartite graphs have $b_2(G) = \mathrm{rk}_2(G)/2$. We shall see that this is also the case for certain even cliques (see Theorem 3.31). For a number of other graph classes, such as (disjoint unions of) odd cliques (Theorem 3.32) or odd cycles (Theorem 3.30), the rank lower bound is off by just one. In Section 3.3, we demonstrate some nice properties possessed by minimum odd covers of graphs for which this bound is sharp. We call a $\mathcal{B}$-odd cover of G a *perfect odd cover* if its cardinality is exactly $\mathrm{rk}_2(G)/2$.

Before proceeding, we note a useful lemma concerning $\mathrm{rk}_2(G)$. We have seen that twin vertices in a graph G do not have any impact on $b_2(G)$. It is not hard to see that they do not impact $\mathrm{rk}_2(G)$ either, for they correspond to identical rows in $A(G)$. On the other hand, *adjacent twins* are pairs $u, v \in V(G)$ with $N[u] = N[v]$ (as opposed to $N(u) = N(v)$), and these may have a large impact on $b_2(G)$. The following lemma shows that disjoint pairs of adjacent twins contribute highly to $\mathrm{rk}_2(G)$.

Lemma 3.8 ([18])**.** *Let G be a graph on n vertices. If G contains a matching M on a set U of $2k$ vertices such that each edge $uv \in M$ is a pair of adjacent twins, then $\mathrm{rk}_2(G) = 2k + \mathrm{rk}_2(G - U)$.*

Proof [18]. It suffices to show that, by elementary operations, the adjacency matrix of G can be turned into a block diagonal matrix whose blocks are an identity matrix of size $2k$ and the adjacency matrix of $G - U$. Without loss of generality, we assume the first two rows r_1, r_2 correspond to two vertices v_1, v_2 that form an edge in M. By definition, $r_1 + r_2 = (1, 1, 0, \ldots, 0)$. The first two entries in each row other than r_1 and r_2 is either $(1, 1)$ or $(0, 0)$. So we can turn all entries in the first two columns, except for the two diagonal entries, into 0 by elementary row operations. And then we can turn all entries in the first two rows, except for the two diagonal entries, into 0 by elementary column operations. Similarly, assuming that the first $2k$ rows and columns correspond to vertices in U, we can turn all entries in the first $2k$ rows or the first $2k$ columns, except for the $2k$ diagonal entries, into 0 by elementary operations, while the entries in the last $(n-2k) \times (n-2k)$ diagonal block remain the same. This completes the proof. □

3.2 Bipartite graphs

In this section, we show that all bipartite graphs have perfect odd covers. First, we examine forests. Recall that a vertex cover of G of cardinality τ gives rise to a partition of $E(G)$ into τ stars. Since each star is a biclique, and a partition is an odd cover, the minimum cardinality $\tau(G)$ of a vertex cover of G provides an upper bound on $b_2(G)$. This was noted in passing in [18].

Proposition 3.9. *For any graph G, $b_2(G) \leq \tau(G)$.*

It is perhaps surprising that this trivial upper bound is sharp for all forests. However, an algebraic result of Mohammadian [75] makes the proof of this fact quick and easy. Recall that the maximum cardinality of a matching in a tree T, $m(T)$, is precisely $\tau(T)$ due to the Kőnig-Egerváry theorem.

Lemma 3.10 ([75])**.** *For any tree T, $\mathrm{rk}_2(T) = 2m(T)$.*

Proposition 3.11 ([18])**.** *For any forest F, $b_2(F) = \mathrm{rk}_2(F)/2 = \tau(F)$.*

Proof. The second equality follows from Lemma 3.10 and the observation that the adjacency matrix of a forest is the direct sum of the adjacency matrices of its tree components. The first equality follows from Proposition 3.6 (the rank lower bound on $b_2(F)$) and Proposition 3.9. □

Note that the maximum size of a matching in a path P_n is $n/2$ when n is odd, and $(n-1)/2$ when n is even.

Corollary 3.12 ([18])**.** *Let n be a positive integer. If n is even, then $b_2(P_n) = n/2$, and if n is odd, then $b_2(P_n) = (n-1)/2$.*

The main result of this section generalizes Proposition 3.11 to arbitrary bipartite graphs, not in terms of the vertex cover upper bound, but in terms of the rank lower bound. Given a bipartite graph G with partite sets X and Y, we we say that a biclique (X_i, Y_i) (or a collection of bicliques) *respects the bipartition of* G if $X_i \subseteq X$ and $Y_i \subseteq Y$.

Theorem 3.13 ([18])**.** *Every bipartite graph has a perfect odd cover which respects its bipartition.*

To simplify the proof of Theorem 3.13, and as it may be of independent interest, we begin with a lemma. Note that, for any graph G with $\text{rk}_2(G) < |G|$, there exists a vertex v such that $\text{rk}_2(G - v) = \text{rk}_2(G)$. That is, we can remove a row and column from $A(G)$ without reducing the rank. In this case, the row in $A(G)$ corresponding to v is the sum of some subset of other rows, corresponding to a subset S of $V(G) - v$. This is equivalent to having $N(v)$ be the symmetric difference of the neighborhoods of the vertices in S, or to having the symmetric difference of the neighborhoods of the vertices in $S \cup v$ be empty. This is true if and only if every vertex in G has an even number of neighbors in $S \cup v$. In other words, the sets of rows (or columns) in $A(G)$ which sum to the zero vector are in natural bijection with the subsets U of $V(G)$ such that $N(w) \cap U$ is even for every $w \in V(G)$. We call such a subset U an *even core* if it is nonempty.

Lemma 3.14 ([18])**.** *Let G be a graph with an even core U, and let $u \in U$. If $G - u$ has a minimum $\mathcal{B}$-odd cover $\mathcal{O}$ such that, for all $(X, Y) \in \mathcal{O}$, at least one of $X \cap (U - u)$ or $Y \cap (U - u)$ is even, then $\mathcal{O}$ can be extended to a $\mathcal{B}$-odd cover of G of the same cardinality. Hence, $b_2(G) = b_2(G - u)$.*

Proof. Suppose that $G - u$ has an odd cover $\mathcal{O}$ as described. For each biclique $(X, Y) \in \mathcal{O}$, we define a biclique (X', Y') by

$$X' = \begin{cases} X: & |X \cap (U-u)| \text{ even;} \\ X \cup u: & \text{otherwise,} \end{cases} \quad \text{and} \quad Y' = \begin{cases} Y: & |Y \cap (U-u)| \text{ even;} \\ Y \cup u: & \text{otherwise.} \end{cases}$$

The bicliques (X', Y') are well-defined, since no biclique (X, Y) has odd-odd intersection with $U - u$.

Let $\mathcal{O}' = \{(X', Y') : (X, Y) \in \mathcal{O}\}$. We claim that $\mathcal{O}'$ is an odd cover of G. Let $v \in V(G) - u$. Clearly, for any vertex $w \in V(G) - \{u, v\}$, v and w are joined by an edge in the same number of bicliques in $\mathcal{O}$ as in $\mathcal{O}'$. As far as the parity which with uv occurs in $\mathcal{O}'$, we claim that uv occurs an odd number of times if and only if $|N(v) \cap (U - u)|$ is odd. Indeed, $uv \in (X', Y')$ if and only if $v \in X$ and $|Y \cap (U - u)|$ is odd, or $v \in Y$ and $|X \cap (U - u)|$ is odd. Thus, the number of bicliques (X', Y') containing uv is odd if and only if v has an odd number of neighbors in $U - u$. Note that, since U is an even core, $uv \in E(G)$ if and only if v has an odd number of neighbors in $U - u$, completing the proof. □

We are now ready to prove Theorem 3.13.

Proof of Theorem 3.13 [18]. We proceed by induction on the order of G. The claim is easily verified for graphs of order at most 2. Now, let G be a bipartite graph on at least three vertices with bipartition (X, Y). Note that, if $\mathrm{rk}_2(G) = |G|$, then for every vertex $u \in V(G)$, $\mathrm{rk}_2(G - u) = \mathrm{rk}_2(G) - 2$. By the inductive hypothesis, $G - u$ has an odd cover with $\mathrm{rk}_2(G)/2 - 1$ bicliques respecting (X, Y). Adding a biclique with partite sets $\{u\}$ and $N(u)$, we obtain perfect odd cover of G which respects the bipartition.

On the other hand, if $\text{rk}_2(G) < |G|$, then G contains an even core U. Let $u \in U$, and without loss of generality, suppose $u \in X$. Note that $\text{rk}_2(G-u) = \text{rk}_2(G)$, since U is an even core, and that $U \subseteq X$, since $N(u) \subseteq Y$ and $N(u)$ is the symmetric difference of the neighborhoods of the vertices in $U - u$. By induction, there is a minimum odd cover $\mathcal{O}$ of $G - u$ that respects (X, Y). In particular, at least one partite set of each biclique in $\mathcal{O}$ contains no vertex in $U - u$. Thus, we may apply Lemma 3.14 to extend $\mathcal{O}$ to a minimum odd cover of G. This completes the proof. $\square$

3.3 Alternating vector representations

Recall the vector representations defined in Section 2.2. We consider here another vector representation over $\mathbb{F}_2$, this one defined for vectors $\boldsymbol{v}, \boldsymbol{w} \in \mathbb{F}_2^{2k}$ by $\boldsymbol{v}_1\boldsymbol{w}_2 + \boldsymbol{v}_2\boldsymbol{w}_1 + \cdots + \boldsymbol{v}_{2k-1}\boldsymbol{w}_{2k} + \boldsymbol{v}_{2k}\boldsymbol{w}_{2k-1}$. A bilinear form b is *symplectic* if it is alternating ($b(\boldsymbol{v}, \boldsymbol{v}) = 0$ for all vectors $\boldsymbol{v}$) and nondegenerate ($b(\boldsymbol{v}, \boldsymbol{w}) = 0$ for all $\boldsymbol{w}$ only if $\boldsymbol{v}$ is the zero vector). Up to isometry, there is a unique symplectic bilinear form over $\mathbb{F}_2^{2k}$ [88]. We thus refer to the bilinear form described above as "the" symplectic bilinear form.

Alternatively, we consider matrix factorizations of $A(G)$ of the form

$$A(G) = M\left(\oplus_1^k \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)\right) M^{\mathsf{T}},$$

where $\oplus$ denotes a direct sum of matrices. In other words, $\oplus_1^k \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ is the $2k \times 2k$ matrix with 1's on the upper and lower diagonal and 0's elsewhere. Note that $\oplus_1^k \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ is the adjacency matrix for a perfect matching on $2k$ vertices. For ease of notation, we denote this matrix by $A_k^{\leftrightarrow}$.

Let G be a graph of order n, and let $\mathcal{O}$ be a $\mathcal{B}$-odd cover of G, $\mathcal{O} = \{(X_i, Y_i) : i \in [k]\}$. For each $i \in [k]$, let $\boldsymbol{x}^{(i)} = (\boldsymbol{x}_v^{(i)})_{v \in V(G)}$ be the incidence vector for X_i. That is, $\boldsymbol{x}_v^{(i)} = 1$ if $v \in X_i$ and $\boldsymbol{x}_v^{(i)} = 0$ otherwise. Similarly, let $\boldsymbol{y}^{(i)}$ denote the incidence vector for Y_i. We form an $n \times 2k$ matrix $M_{\mathcal{O}}$ whose columns are the incidence vectors:

$$M_{\mathcal{O}} = \begin{pmatrix} | & | & & | & | \\ \boldsymbol{x}^{(1)} & \boldsymbol{y}^{(1)} & \cdots & \boldsymbol{x}^{(k)} & \boldsymbol{y}^{(k)} \\ | & | & & | & | \end{pmatrix}. \tag{3.1}$$

Proposition 3.15 ([18, 19])**.** *Let $\mathcal{O}$ be an odd cover of G, $\mathcal{O} = \{(X_i, Y_i) : i \in [k]\}$. The rows of $M_{\mathcal{O}}$ comprise a faithful vector representation of G over $\mathbb{F}_2$ defined by the symplectic bilinear form on $\mathbb{F}_2^{2k}$. That is,*

$$M_{\mathcal{O}} A_k^{\leftrightarrow} M_{\mathcal{O}}^{\mathsf{T}} = A(G).$$

Proof. Let $\boldsymbol{u} = (\boldsymbol{x}_u^{(1)}, \boldsymbol{y}_u^{(1)}, \ldots, \boldsymbol{x}_u^{(k)}, \boldsymbol{y}_u^{(k)})$ and $\boldsymbol{v} = (\boldsymbol{x}_v^{(1)}, \boldsymbol{y}_v^{(1)}, \ldots, \boldsymbol{x}_v^{(k)}, \boldsymbol{y}_v^{(k)})$ be (not necessarily distinct) rows of $M_{\mathcal{O}}$, corresponding to the vertices u and v of G. Then $b(\boldsymbol{u}, \boldsymbol{v}) = \sum_{i=1}^{k} (\boldsymbol{x}_u^{(i)} \boldsymbol{y}_v^{(i)} + \boldsymbol{y}_u^{(i)} \boldsymbol{x}_v^{(i)})$, where the sum is taken over $\mathbb{F}_2$. Note that the ith summand is 1 if u and v are in differing partite sets of (X_i, Y_i), and is 0 otherwise. Thus, $b(\boldsymbol{u}, \boldsymbol{v}) = 1$ if the edge uv occurs in an odd number of bicliques in $\mathcal{O}$, and $b(\boldsymbol{u}, \boldsymbol{v}) = 0$ otherwise, as desired. $\square$

We note that Proposition 3.15 provides an alternate proof of the rank lower bound, Proposition 3.6. As in Section 2.2, it is easy to show that $\mathrm{rk}(M A_k^{\leftrightarrow} M^{\mathsf{T}}) \leq \mathrm{rk}(M)$ for any $n \times 2k$ matrix M. Letting $\mathcal{O}$ be a minimum $\mathcal{B}$-odd cover of G, the matrix $M_{\mathcal{O}}$ has $2b_2(G)$ columns, so $2b_2(G) \geq \mathrm{rk}_2(M) \geq \mathrm{rk}_2(G)$.

Not every graph has a perfect odd cover, and thus we cannot always find an incidence matrix $M_{\mathcal{O}}$ with $2b_2(G)$ columns which factors $A(G)$ in this manner. However, the following theorem of Friedland tells us that some matrix M with $\mathrm{rk}_2(G)$ columns factors $A(G)$, even if it does not correspond to a $\mathcal{B}$-odd cover.

Theorem 3.16 ([50, p. 426-427])**.** *For any $n \times n$ symmetric matrix A with zero diagonal and rank $2r$ over a field of characteristic 2, there is an $n \times 2r$ matrix M such that*

$$A = MA^{\leftrightarrow}_{r}M^{\mathsf{T}}.$$

In particular, for any graph G of order n with $\mathrm{rk}_2(G) = 2r$, there is an $n \times 2r$ matrix M such that $A(G) = MA^{\leftrightarrow}_{r}M^{\mathsf{T}}$. Suppose that $b_2(G) > \mathrm{rk}_2(G)/2$. If this matrix M is not an incidence matrix for a $\mathcal{B}$-odd cover of G, then what is it?

Note that, if $M_{\mathcal{O}}$ is an incidence matrix for an odd cover $\mathcal{O}$ of G with k bicliques, then for every $i \in [k]$ and for every $v \in V(G)$, at least one of $\boldsymbol{x}_v^{(i)}$ and $\boldsymbol{y}_v^{(i)}$ is 0. That is, no vertex can be contained in both partite sets of a biclique. It follows that $b_2(G) > \mathrm{rk}_2(G)/2 = r$ if and only if, for every $n \times 2r$ matrix M (with entries denoted as in (3.1)) such that $MA^{\leftrightarrow}_{r}M^{\mathsf{T}} = A(G)$, at least one pair $(\boldsymbol{x}_v^{(i)}, \boldsymbol{y}_v^{(i)}) = (1,1)$. We interpret the case of $(1,1)$-pairs combinatorially using $\mathcal{T}$-odd covers.

Theorem 3.17 ([19])**.** *For every graph G, $\varrho_2(G, \mathcal{T}) = \mathrm{rk}_2(G)/2$.*

Proof. Let G be a graph, $|G| = n$ and $\mathrm{rk}_2(G) = 2r$. Let M be an $n \times 2r$ factor of the adjacency matrix A of G, as in Theorem 3.16. As before, let the rows of M be indexed by the vertices of G, and let the columns be denoted $\boldsymbol{x}^{(1)}, \boldsymbol{y}^{(1)}, \ldots, \boldsymbol{x}^{(r)}, \boldsymbol{y}^{(r)}$.

We obtain a collection of r tricliques (X_i, Y_i, Z_i) from M by setting

$$\begin{aligned} X_i &= \{v \in V(G) : (\boldsymbol{x}_v^{(i)}, \boldsymbol{y}_v^{(i)}) = (1,0)\}, \\ Y_i &= \{v \in V(G) : (\boldsymbol{x}_v^{(i)}, \boldsymbol{y}_v^{(i)}) = (0,1)\}, \text{ and} \\ Z_i &= \{v \in V(G) : (\boldsymbol{x}_v^{(i)}, \boldsymbol{y}_v^{(i)}) = (1,1)\}. \end{aligned}$$

Note that $A_{u,v} = \sum_{i=1}^{r}(\boldsymbol{x}_u^{(i)}\boldsymbol{y}_v^{(i)} + \boldsymbol{x}_v^{(i)}\boldsymbol{y}_u^{(i)})$, and the ith summand is 1 if and only if $(\boldsymbol{x}_u^{(i)}, \boldsymbol{y}_u^{(i)}) \neq (\boldsymbol{x}_v^{(i)}, \boldsymbol{y}_v^{(i)})$ and neither pair is $(0,0)$. Thus, we have that $uv \in E(G)$ if and only if uv occurs in an odd number of tricliques (X_i, Y_i, Z_i). In other words, $\{(X_i, Y_i, Z_i) : i \in [r]\}$ is a $\mathcal{T}$-odd cover of G. Further, it has minimum cardinality over all $\mathcal{T}$-odd covers of G since $\mathrm{rk}(M) \geq \mathrm{rk}(A)$ whenever $MA^{\leftrightarrow}M^{\mathsf{T}} = A$. This completes the proof. □

In other words, the matrices M such that $MA^{\leftrightarrow}M^{\mathsf{T}} = A(G)$ are in bijection with the $\mathcal{T}$-odd covers of G. Recall from Corollary 2.11 that $\varrho_2(G, \mathcal{K}) > \mathrm{mr}(G, \mathbb{F}_2)$ if and only if $A(G)$ uniquely minimizes the rank over all symmetric matrices which fit G over $\mathbb{F}_2$.

Corollary 3.18 ([19])**.** *For any graph G,* $\mathrm{mr}(G, \mathbb{F}_2) \in \{\varrho_2(G, \mathcal{K}), \varrho_2(G, \mathcal{T})\}$.

Theorem 3.17 also allows for an algebraic upper bound on $b_2(G)$.

Corollary 3.19 ([18])**.** *For any graph G, $b_2(G) \leq \mathrm{rk}_2(G)$.*

As opposed to the lower bound $b_2(G) \geq \mathrm{rk}_2(G)/2$, we do not know whether the bound in Corollary 3.19 is sharp. However, we believe that it is. In Section 3.3.2, we define a graph T_k for each positive integer k which maximizes b_2 over all graphs G with $\mathrm{rk}_2(G) = 2k$. For $k \in \{1, 2\}$, we have checked by computer that $b_2(T_k) = 2k$.

3.3.1 Bases and even cores

Here, we demonstrate what we previously called the "nice properties" possessed by perfect odd covers. Recall that a set of vectors is said to be *independent* if no subset sums to the zero vector. A *basis* is a maximally independent subset from a set of vectors; it is a standard exercise to show that all bases from the same set of vectors have the same cardinality. We have been using this fact all along, for the rank of a matrix is the size of a basis for its row space or column space.

The properties in question follow from a bijection between the independent sets of rows in M and the independent sets of rows in $A(G)$ when M is an $n \times \mathrm{rk}_2(G)$ matrix such that $MA^{\leftrightarrow}M^{\mathsf{T}} = A(G)$.

Theorem 3.20 ([17]). *Let G be a graph of order n and rank $2r$, and let $M \in \mathbb{F}_2^{n\times 2r}$ be such that $MA_r^{\leftrightarrow}M^{\mathsf{T}} = A(G)$, where $A_r^{\leftrightarrow} = \oplus_1^r \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$. For any subset S of $\{1, \ldots, n\}$, the set of rows in M indexed by S is independent if and only if the set of rows in $A(G)$ indexed by S is independent.*

Proof. Let G and M be as described, and let $A = A(G)$. For each vertex $v \in V(G)$, let $\boldsymbol{m}^{(v)}$ denote the row of M indexed by v and $\boldsymbol{a}^{(v)}$ the row of A indexed by v. First, suppose that a subset $\{\boldsymbol{m}^{(s)} : s \in S\}$ of rows of M sums to the zero vector $\mathbf{0}$. If $\boldsymbol{s}$ denotes the $1 \times n$ incidence vector for S, then $\boldsymbol{s}M = \mathbf{0}$, and thus $\boldsymbol{s}A = \mathbf{0}$. It follows that $\{\boldsymbol{m}^{(s)} : s \in S\}$ is independent whenever the set $\{\boldsymbol{a}^{(s)} : s \in S\}$ of rows in $A(G)$ is independent.

On the other hand, suppose that $\{\boldsymbol{m}^{(s)} : s \in S\}$ is independent. This set is contained in a basis for the row space of M which induces a $2r \times 2r$ full-rank submatrix

M_B. After a reordering of the vertices of G and the rows of M, we can write

$$M = \begin{pmatrix} M_B \\ RM_B \end{pmatrix} = \begin{pmatrix} I \\ R \end{pmatrix} M_B.$$

Letting $B = M_B A_r^{\leftrightarrow} M_B^{\mathsf{T}}$, we have

$$A = M A_r^{\leftrightarrow} M^{\mathsf{T}} = \begin{pmatrix} I \\ R \end{pmatrix} B \begin{pmatrix} I & R^T \end{pmatrix} = \begin{pmatrix} B & BR^{\mathsf{T}} \\ RB & RBR^{\mathsf{T}} \end{pmatrix}.$$

It follows that the first $2r$ rows in A, those corresponding to B, span the row space of A, and these contain every $\boldsymbol{a}^{(s)}$ for $s \in S$. Since $\mathrm{rk}_2(G) = 2r$, the proof is complete. □

Recall from Section 3.2 that a nonempty set of rows (or columns) in $A(G)$, indexed by a subset S of $V(G)$, sum to $\mathbf{0}$ if and only if S is an even core; that is, every vertex in G has an even number of neighbors in S. By virtue of Theorem 3.17, we obtain the following corollary.

Corollary 3.21 ([17])**.** *Let M be the incidence matrix for a minimum $\mathcal{T}$-odd cover of a graph G. A nonempty subset of $V(G)$ is an even core in G if and only if the corresponding rows in M sum to the zero vector.*

In the case that $M = M_{\mathcal{O}}$ for a perfect $\mathcal{B}$-odd cover $\mathcal{O}$ of G, there is a combinatorial interpretation of a set of rows in M which sum to 0: the corresponding subset S of $V(G)$ has even intersection with every partite set in $\mathcal{O}$. From this, we obtain the following corollary.

Corollary 3.22 ([17]). *If $\mathcal{O}$ is a perfect $\mathcal{B}$-odd cover of a graph G, then the even cores in G are precisely those nonempty subsets of $V(G)$ which have even intersection with both partite sets of every biclique in $\mathcal{O}$. In particular, no subgraph of G induced by an even core has an odd number of edges.*

Proof. If W is an even core such that $G[W]$ has an odd number of edges, then any $\mathcal{B}$-odd cover of G must contain some biclique (X, Y) where both $|X \cap W|, |Y \cap W|$ are odd. Thus, such an odd cover cannot be a perfect odd cover, so $b_2(G) \geq \frac{\mathrm{rk}_2(G)}{2} + 1$. □

As an aside, it follows from Corollary 3.22 that, for any graph G with a perfect odd cover $\mathcal{O}$, the even cores of cardinality 3 are independent sets (in the graph-theoretic sense). For, if W is an even core, and $w \in W$, then for any $(X, Y) \in \mathcal{O}$, at least one of $|X \cap (W - w)|$ or $|Y \cap (W - w)|$ is even since each of $|X \cap W|$ and $|Y \cap W|$ is even and $X \cap Y = \varnothing$. Thus, no biclique in $\mathcal{O}$ builds an edge between any pair of vertices in W.

We note one final relationship between perfect odd covers and linearly independent sets of rows in $A(G)$. Though we have not found a use for it yet, we believe it may come in handy some day. Let $\mathcal{A}$ be a collection of subsets of a set V. A *transversal* of $\mathcal{A}$ is a set of $|\mathcal{A}|$ elements from V, each one contained in a distinct set in $\mathcal{A}$. We now prove that the subsets of vertices which index bases for the row space of $A(G)$ are transversals for the partite sets in any perfect odd cover of G. Note that, when $\mathrm{rk}_2(G) = |G|$ and G has a perfect odd cover, this implies that each vertex of G can be assigned as a representative to a distinct partite set in which it occurs.

Theorem 3.23 ([17]). *Let G be a graph with $\mathrm{rk}_2(G) = 2r$ and a perfect odd cover $\{(X_i, Y_i) : i \in [r]\}$. If a set of rows in $A(G)$ is a basis for its row space, then the* *corresponding set of vertices in G is a transversal for the collection of partite sets* $\{X_1, Y_1, \ldots, X_r, Y_r\}$.

We present a proof using Hall's marriage theorem [56], a celebrated result in the study of set systems, which states that $\mathcal{A}$ has a transversal if and only if $|\mathcal{S}| \leq |\cup_{S \in \mathcal{S}}|$ for every subcollection $\mathcal{S}$ of $\mathcal{A}$. We note an alternate proof below it, using Theorem 3.20, an analysis of the even cores in $G \triangle (X_i, Y_i)$ for any $i \in [r]$, and using the augmentation property of independent sets.

Proof [17]. Suppose that a set of rows in $A(G)$, corresponding to a subset B of $V(G)$, is a basis for the row space of $A(G)$. For $1 \leq m \leq 2r$, any set of m vertices in B collectively appears in at least m partite sets of the perfect odd cover. This is due to the fact that if a vertex v is contained only within a subset of a fixed collection of $m-1$ partite sets, then the corresponding row of $A(G)$ is spanned by the indicator vectors of the $m-1$ opposite partite sets, so these m independent row vectors would lie in a dimension $m-1$ subspace of $\mathbb{F}_2^n$, giving a contradiction. Thus, by Hall's marriage theorem [56], there is an ordering $x_1, y_1, \ldots, x_r, y_r$ of the vertices in B so that $x_i \in X_i$ and $y_i \in Y_i$ for all $i \in \{1, \ldots, r\}$. □

Alternate proof of Theorem 3.23. Let $B \subseteq V(G)$ correspond to a basis of the row space of $A(G)$. By Theorem 3.20, the rows indexed by B in the incidence matrix M for $\mathcal{O}$ form a basis for its row space. Let $G' = G \triangle (X_i, Y_i)$, and let M' be the submatrix of M corresponding to the perfect odd cover $\mathcal{O} - (X_i, Y_i)$ of G'. Note that there exist $u, v \in B$ such that the set B' of rows indexed by $B - \{u, v\}$ in M' is a basis for its row space. Now $B' \cup u$ contains a unique even core I in G'; note that $u \in I$. Similarly, there is a unique even core J in $B' \cup v$, and $v \in J$. Note that $K = I \triangle J$ is also an even core in G', $u, v \in K$, and these three even cores are all distinct.

We now consider the sets I, J, and K in G. Each corresponds to an independent set of rows in $A(G)$, and the only possible sets S_I, S_J, and S_K of vertices w with an odd number of neighbors in I, J, and K in G, respectively, are X_i, Y_i, and $X_i \cup Y_i$. We claim that $\{S_I, S_J, S_K\} = \{X_i, Y_i, X_i \cup Y_i\}$. If, for a contradiction, we assume $S_I = S_J$, then that same set would have an even number of neighbors in $I \triangle J$ (since the parity of $N(v) \cap (I - J)$ is the same as that of $N(v) \cap (J - I)$ for any $v \in S$). Thus, K would be an even core in G, but $K \subseteq B$ and B corresponds to a basis of $A(G)$, a contradiction. The same argument holds if we assume $S_K \in \{S_I, S_J\}$.

We now claim that there exist vertices x, y in B such that $x \in X_i$, $y \in Y_i$, and $B - \{x, y\}$ is a basis for G'. If $u \in X_i$ and $v \in Y_i$, or vice-versa, then we are done. Otherwise, without loss of generality, we assume that I is the even core in G' such that Y_i is the set of vertices w with $|N_G(w) \cap I|$ odd. Note that I must contain a vertex $x \in X_i$, for otherwise $N_{G'}(w) \cap I = N_G(w) \cap I$ for every $w \in V(G)$. Note that $B' - x \cup u$ indexes a basis for the row space of $A(G')$ (this is a property of fundamental circuits; we delete x from I and augment it to a new basis). Now, both J and K contain v, and both contain a vertex from Y_i since $\{S_J, S_K\} = \{X_i, X_i \cup Y_i\}$. If, say, $y \in S_J \cap Y_i$, we obtain another basis for the row space of $A(G')$, $B' - y \cup v$. Now we use the basis exchange axiom: we can delete u from $B' - x \cup u$ and replace it with an element from $B' - y \cup v$ to obtain a third set indexing a basis of $A(G')$, and this set contains neither x nor y. Thus, we can delete x and y from G' and proceed by induction. □

(a) The graph T_2 (b) The graph B_2

Figure 3.6: The universal graphs T_2 and B_2

3.3.2 The universal graphs B_k and T_k

Recall Proposition 3.2, that if G is a graph with twin vertices u and v, then we can obtain a $\mathcal{B}$-odd cover of G from one of $G - v$ by including v in every partite set in which u occurs. Clearly, the same result holds for $\mathcal{T}$-odd covers. In this section, we define universal graphs for $\mathcal{B}$- and $\mathcal{T}$-odd covers, in the sense that they contain every twin-free graph with $\varrho_2(G, \mathcal{B}) \leq k$ (resp. $\varrho_2(G, \mathcal{T}) \leq k$) as an induced subgraph.

Definition 3.1 (B_k, T_k)**.** Let k be a positive integer, and let M_k denote the matrix which has for rows all distinct vectors over $\mathbb{F}_2^{2k}$. We define T_k to be the graph with adjacency matrix

$$A(T_k) = M_k A_k^{\leftrightarrow} M_k^{\mathsf{T}}.$$

Writing M_k as in (3.1), *i.e.*, with columns $\boldsymbol{x}^{(1)}, \boldsymbol{y}^{(1)}, \ldots, \boldsymbol{x}^{(k)}, \boldsymbol{y}^{(k)}$, we let N_k denote the submatrix obtained by deleting all rows having at least one pair $(\boldsymbol{x}^{(i)}, \boldsymbol{y}^{(i)}) = (1, 1)$. We define B_k to be the graph with adjacency matrix

$$A(B_k) = N_k A_k^{\leftrightarrow} N_k^{\mathsf{T}}.$$

In [53], Godsil and Royle also studied the graph T_k. Technically, they studied the graph obtained from T_k by deleting the isolated vertex (associated to the zero vector in M_k), but we find it simpler to leave it in. The graph T_k has many interesting properties.

Proposition 3.24 ([53]). *For any integer k, the graph T_k contains as an induced subgraph every twin-free graph G with $\mathrm{rk}_2(G) \leq 2k$.*

Combined with Theorem 3.17, we see that T_k contains as an induced subgraph every twin-free graph G with $\varrho_2(G, \mathcal{T}) \leq k$.

Corollary 3.25. *Let k be a positive integer. For any graph G with $\mathrm{rk}_2(G) \leq 2k$, $b_2(G) \leq b_2(T_k)$.*

This is in fact natural, as by definition T_k is obtained from a matrix M_k representing a $\mathcal{T}$-odd cover. In a more combinatorial light, if we take k tricliques (X_i, Y_i, Z_i) and ask for all possible combinations of tricliques that a vertex can be in (either in X_i, Y_i, or Z_i, or not in any of them, for each $i \in [k]$), there are 4^k such combinations, associated to the 4^k vertices of T_k. Similarly, B_k can be thought of as encoding all of the different ways that a vertex might take part in a collection of k bicliques. For this reason, we lay out T_2 like a grid in Figure 3.6a, thinking of the rows as corresponding to "in X_1," "in Y_1," "in Z_1," or "not in the triclique," and similarly for the columns.

The main result of [53] was that T_k maximizes the chromatic number over all graphs G with $\mathrm{rk}_2(G) \leq 2k$, and that $\chi(T_k) = 2^k + 1$. We are interested in T_k for a different reason: the value $b_2(T_k)$ is bounded asymptotically away from k, the rank lower bound. This fact will follow simply from the following observation.

Proposition 3.26 ([18]). *For any integer k, the graph B_k contains as an induced subgraph every twin-free graph G with $b_2(G) \leq k$.*

Proof. Suppose that G is a twin-free graph with $b_2(G) \leq k$. Let $\mathcal{O}$ be a $\mathcal{B}$-odd cover of G with k (possibly degenerate) bicliques, and let $M_{\mathcal{O}}$ be its incidence matrix, as in Proposition 3.15. Note that no two rows of $M_{\mathcal{O}}$ are identical, or else the corresponding vertices in G would be twins, and that $M_{\mathcal{O}}$ contains no $(1,1)$ pairs $(\boldsymbol{x}_v^{(i)}, \boldsymbol{y}_v^{(i)})$. Thus, $M_{\mathcal{O}}$ is a submatrix of N_k. It is easy to check that the rows of a submatrix of N_k correspond to the vertices in an induced subgraph of B_k, which completes the proof. □

Theorem 3.27 ([18]). *For any positive integer k, $b_2(T_k) \geq \log_3 4 \cdot k$.*

Proof. Recall that T_k is twin-free. If $b_2(T_k) = \ell$, then T_k is an induced subgraph of B_ℓ by Proposition 3.26. In particular, $|T_k| = 4^k \leq 3^\ell = |B_\ell|$. The desired result follows. □

We note that Theorem 3.27 can be used to find other graphs with $b_2(G) > \mathrm{rk}_2(G)/2 + r$ for any integer r by taking the symmetric difference of T_k with a graph on a subset of its vertices whose value of b_2 is sufficiently small.

We note one final property of T_k, which has proved useful in ongoing work with Clifton, Nie, and Rombach in improving the bound in Theorem 3.27.

Proposition 3.28. *The sum of any linearly independent set of rows in $A(T_k)$ contains exactly $4^k/2$ ones.*

Proof. Let M_k be the matrix whose rows are all distinct vectors from $\mathbb{F}_2^{2k}$, labeled by the vertices of T_k so that $A(T_k) = M_k A_{k}^{\leftrightarrow} M_k^T$. To fix some notation, write $A(T_k) =$

$(a_{v,w})$, and let the vth row of M be denoted $(\boldsymbol{x}_v^{(1)}, \boldsymbol{y}_v^{(1)}, \ldots, \boldsymbol{x}_v^{(k)}, \boldsymbol{y}_v^{(k)})$. Note that $A(T_k)_{v,w} = \sum_{i=1}^{k}(\boldsymbol{x}_v^{(i)}\boldsymbol{y}_w^{(i)} + \boldsymbol{y}_v^{(i)}\boldsymbol{x}_w^{(i)})$.

It is not hard to show that every vertex in T_k, aside from the isolate, has degree $4^k/2$. Now, suppose that $\{\boldsymbol{a}^{(u)} : u \in U\}$ is a linearly independent set of at least two rows from $A(T_k)$. Note that the corresponding set of rows in M, $\{\boldsymbol{m}^{(u)} : u \in U\}$, is independent by Theorem 3.20. Thus, there exists some $v \notin U$ such that $\boldsymbol{m}^{(v)} \neq 0$ and $\sum_{u\in U} \boldsymbol{m}^{(u)} = \boldsymbol{m}^{(v)}$. That is, $\boldsymbol{x}_v^{(i)} = \sum_{u\in U} \boldsymbol{x}_u^{(i)}$ and $\boldsymbol{y}_v^{(i)} = \sum_{u\in U} \boldsymbol{y}_u^{(i)}$ for all $i \in \{1, \ldots, k\}$. Then, for any $w \in \{1, \ldots, 4^k\}$,

$$\begin{aligned} A(T_k)_{v,w} = \sum_{i=1}^{k}(\boldsymbol{x}_v^{(i)}\boldsymbol{y}_w^{(i)} + \boldsymbol{y}_v^{(i)}\boldsymbol{x}_w^{(i)}) &= \sum_{i=1}^{k}\left(\boldsymbol{y}_w^{(i)} \sum_{u\in U} \boldsymbol{x}_u^{(i)} + \boldsymbol{x}_w^{(i)} \sum_{u\in U} \boldsymbol{y}_u^{(i)}\right) \\ &= \sum_{u\in U}\sum_{i=1}^{k}(\boldsymbol{x}_u^{(i)}\boldsymbol{y}_w^{(i)} + \boldsymbol{y}_u^{(i)}\boldsymbol{x}_w^{(i)}). \end{aligned}$$

Therefore, $\boldsymbol{a}^{(v)} = \sum_{u\in U} \boldsymbol{a}^{(u)}$. We know that $\boldsymbol{a}^{(v)}$, being nonzero, has $4^k/2$ ones, which completes the proof. □

Corollary 3.29. *Any set of rows in $A(T_k)$ sums either to the zero vector or to a vector with exactly $4^k/2$ ones over $\mathbb{F}_2$.*

3.4 Disjoint unions of cycles

In this short section, we use the relationship between even cores and perfect odd covers (Corollary 3.22) to derive a lower bound on the value of $b_2(G)$, which is one larger than $\mathrm{rk}_2(G)/2$, when G is a disjoint union of cycles. We also provide an upper bound to match.

Theorem 3.30 ([17])**.** *If G is a disjoint union of k odd cycles $C_{2n_1+1}, \ldots, C_{2n_k+1}$ and ℓ even cycles $C_{2m_1}, \ldots, C_{2m_\ell}$, then $b_2(G) = \sum n_i + \sum m_i - \ell + 1$.*

Proof. Note that $\mathrm{rk}_2(C_{2m}) = 2m-2$ and $\mathrm{rk}_2(C_{2n+1}) = 2n$. Thus, $\mathrm{rk}_2(G)/2 = \sum_1^k n_i + \sum_1^\ell m_i - \ell$. We obtain an odd cover of $C_{2m_1} + \cdots + C_{2m_\ell}$ using $\sum_1^\ell (m_i - 1)$ bicliques as in Proposition 3.3. If $k = 0$, then we have found a perfect odd cover, and we are done. Otherwise, there is at least one odd cycle in the union. This odd cycle is an even core with an odd number of edges, so Corollary 3.22 proves the lower bound. For the upper bound, we note that an odd cycle C_{2n+1} with vertex set $\{x_1, \ldots, x_n, y_1, \ldots, y_n, z\}$ has an odd cover with $n+1$ bicliques: $(\{x_i, y_{i+1}\}, \{x_{i+1}, y_i\})$ for $i \in [n-1]$ (as in Proposition 3.3), $(\{x_n, y_n\}, \{z\})$, and $(\{x_n\}, \{y_n\})$. Note that the last two bicliques form a triangle. In this manner, we can extend the odd cover of kK_3 with $k+1$ bicliques from Proposition 3.4 to an odd cover of $C_{2n_1+1} + \cdots + C_{2n_k+1}$ using an extra $\sum_1^k (n_i - 1)$ bicliques. Thus, $b_2(G) \leq k + 1 + \sum_1^k n_i - k + \sum_1^\ell m_\ell - \ell$, which simplifies to the desired upper bound. $\square$

3.5 Babai and Frankl's odd cover problem

Let us now turn our attention back to Babai and Frankl's problem of determining $b_2(K_n)$. We summarize our results in the following theorem.

Theorem 3.31 ([17, 18])**.** *For any $n \geq 3$, we have*

$$\left\lceil \frac{n}{2} \right\rceil \leq b_2(K_n) \leq \left\lfloor \frac{n}{2} \right\rfloor + 1.$$

Further, whenever $n \equiv 0 \pmod 8$ *or* $n \equiv 18 \pmod{24}$, K_n *has a perfect odd cover.*

We prove Theorem 3.31 in parts. That $b_2(K_{2n+1}) = \lceil n/2 \rceil$ will follow from Theorem 3.32.[1] That $b_2(K_{2n}) \geq n/2$ will follow from Proposition 3.6, and that $b_2(K_{2n}) \leq n/2 + 1$ will follow from Theorem 3.34. The values of $b_2(K_{8n})$ and $b_2(K_{24n+18})$ are proved in Theorems 3.35 and 3.36, respectively. Along the way, we reveal some interesting properties of perfect odd covers of even cliques and generalize our result for odd cliques to disjoint unions. Many of the proofs in this section are quoted directly from the collaboration in which they appear (either [18] or [17]).

Theorem 3.32 ([17]). *Let* $m_1, \ldots, m_j$ *be positive integers. We have*

$$b_2(K_{2m_1+1} + \cdots + K_{2m_j+1}) = \sum_{i=1}^{j} m_i + 1$$

Proof [17]. We first demonstrate the lower bound. Since $\mathrm{rk}_2(K_{2m_1+1} + \cdots + K_{2m_j+1}) = \sum 2m_i$, it suffices to show that no perfect odd cover exists. Suppose, for the sake of contradiction, that $B = (X, Y)$ is a biclique in a perfect odd cover. For each K_{2m_i+1}, there are a_i vertices in X, b_i vertices in Y, and c_i vertices in $Z = V(K_{2m_1+1} + \cdots + K_{2m_j+1}) - (X \cup Y)$. Every K_{2m_i+1} is an even core, and thus every a_i and b_i is even by Corollary 3.22. For each clique, we pair up vertices that are in the same one of X, Y, or Z to get $a_i/2 + b_i/2 + \lfloor c_i/2 \rfloor$ pairs of adjacent twins in $(K_{2m_1+1} + \cdots + K_{2m_j+1}) \triangle B$. Note that $m_i = a_i/2 + b_i/2 + \lfloor c_i/2 \rfloor$, so we have $\sum m_i$ pairs of adjacent twins in $(K_{2m_1+1} + \cdots + K_{2m_j+1}) \triangle B$. By Proposition 3.8, the rank of this graph is $2 \sum m_i$, which implies $b_2((K_{2m_1+1} + \cdots + K_{2m_j+1}) \triangle B) \geq \sum m_i$, a contradiction.

Now we will give a construction that provides a matching upper bound. Let

[1]Leader and Tan independently determined $b_2(K_{2n+1})$ in a parallel work [67]. Their construction arose from an analysis of the corresponding odd cover problem for complete hypergraphs.

$u_i, v_{i,1}, \ldots, v_{i,m_i}, w_{i,1}, \ldots, w_{i,m_i}$ be the vertices of the i^{th} complete graph K_{2m_i+1}, $1 \leq i \leq j$. By Theorem 3.5, there exist $m_1 + \cdots + m_j$ complete bipartite graphs $B'_{i,k}$ with parts $(X'_{i,k},\ Y'_{i,k})$, $1 \leq i \leq j$, $1 \leq k \leq m_i$, such that: (i) they form an odd cover of $2K_{m_1} + \cdots + 2K_{m_j}$ where one copy of K_{m_i} is induced on the $v_{i,k}$'s and the other is induced on the $w_{i,k}$'s; and (ii) $X'_{i,k} = \{v_{i,k}, w_{i,k}\}$ for all $1 \leq i \leq j$, $1 \leq k \leq m_i$.

Now we construct an odd cover of $K_{2m_1+1} + \cdots + K_{2m_j+1}$ as follows. For all $1 \leq i \leq j$, $1 \leq k \leq m_i$, let $V_i = \{v_{i,1},\ \ldots,\ v_{i,m_i}\}$, $W_i = \{w_{i,1},\ \ldots,\ w_{i,m_i}\}$,

$$X_{i,k} = X'_{i,k},$$

$$Y_{i,k} = Y'_{i,k} \cup \{u_i\} \cup \left(\bigcup_{1 \leq t \leq j,\ t \neq i} V_t \right).$$

Further, let

$$X_0 = \cup_{1 \leq i \leq j} V_i,$$

$$Y_0 = \cup_{1 \leq i \leq j} W_i.$$

Then it is not hard to check that the $m_1 + \cdots + m_j + 1$ bicliques $(X_0, Y_0), (X_{i,k}, Y_{i,k})$, $1 \leq i \leq j$, $1 \leq k \leq m_i$, form an odd cover of $K_{2m_1+1} + \cdots + K_{2m_j+1}$. This shows that $b_2(K_{2m_1+1} + \cdots + K_{2m_j+1}) \leq m_1 + \cdots + m_j + 1$, which matches the lower bound. □

Corollary 3.33 ([17]). *For any odd integer n, $n \geq 3$, $b_2(K_n) = \lceil n/2 \rceil$.*

Let us now consider $b_2(K_n)$ for even n. We will first show that, if an even clique does not have a perfect $\mathcal{B}$-odd cover (of cardinality $n/2$), then it requires at most one extra biclique. We will then determine that K_n has a perfect odd cover for n meeting certain divisibility conditions. What exactly the conditions on n are to guarantee a perfect odd cover, however, remains an open problem. We make progress towards this question in Theorem 3.38. We now proceed to prove an upper bound of $n+1$

on $b_2(K_{2n})$. Note that this is only one off from the rank lower bound of n from Proposition 3.6.

Theorem 3.34 ([18]). *Let G be a graph of order $2n$. If G contains a perfect matching M such that each edge $uv \in M$ is a pair of adjacent twins, then* $\mathrm{rk}_2(G) = 2n$ *and* $b_2(G) \leq n + 1$.

Proof. By Lemma 3.8, we have $\mathrm{rk}_2(G) = 2n$. We can show $b_2(G) \leq n + 1$ by construction. Let $M = \{a_i b_i\}_{1 \leq i \leq n}$. We prove the following two statements by induction:

(I) If n is odd, there exists a $\mathcal{B}$-odd cover $\{(X_1, Y_1), \ldots, (X_{n+1}, Y_{n+1})\}$ of G such that each pair a_i, b_i is joined by an edge in every biclique except (X_i, Y_i), and neither a_i nor b_i is in (X_i, Y_i).

(II) If n is even, there exists a $\mathcal{B}$-odd cover $\{(X_1, Y_1), \ldots, (X_{n+2}, Y_{n+2})\}$ of G such that each pair a_i, b_i is joined by an edge in every biclique except (X_i, Y_i), which contains neither a_i nor b_i, and every a_i is contained in X_{n+2}.

We note that the $\mathcal{B}$-odd covers described above may contain degenerate bicliques, with either one or both partite sets being empty, and thus are not necessarily optimal.

For the base case, let $(X_2, Y_2) = (\{a_1\}, \{b_1\})$. When $n = 1$, let $X_1 = Y_1 = \varnothing$; (I) is satisfied. When $n = 2$, let $(X_1, Y_1) = (\{a_2\}, \{b_2\})$,

$$(X_3, Y_3) = \begin{cases} (\{a_1, a_2\}, \{b_1, b_2\}) : & a_1 a_2 \notin E(G), \\ (\{a_1, b_2\}, \{b_1, a_2\}) : & a_1 a_2 \in E(G), \end{cases}$$

and $(X_4, Y_4) = (\{a_1, a_2\}, \{b_1, b_2\})$. One can check that (II) is satisfied.

Case 1 ($n = 2k+1$). First, we assume that n is an odd integer at least 3, and that (II) holds for $n-1$. Let $\mathcal{O} = \{(X_1, Y_1), \ldots, (X_{2k+2}, Y_{2k+2})\}$ be an odd cover of $G - \{a_n, b_n\}$ meeting (II). We will show that (I) holds for G by adding a_n and b_n to opposite partite sets of every biclique in $\mathcal{O}$ except (X_n, Y_n).

Consider the first $2k$ bicliques in $\mathcal{O}$. Let $\mathcal{O}_{2k} = \{(X_1, Y_1), \ldots, (X_{2k}, Y_{2k})\}$ and $\mathcal{O}_{2k}^Y = \{Y_1, \ldots, Y_{2k}\}$. For each $i \in [2k]$, we add a_{2k+1} to Y_i if either (i) a_i occurs an even number of times in $\mathcal{O}_{2k}^Y$ and $a_i a_{2k+1} \in E(G)$, or (ii) a_i occurs an odd number of times in $\mathcal{O}_{2k}^Y$ and $a_i a_{2k+1} \notin E(G)$. Otherwise, we add a_{2k+1} to X_i. Add b_{2k+1} to the opposite partite in each case. Let $\mathcal{O}'_{2k}$ denote the resulting collection $\{(X'_1, Y'_1), \ldots, (X'_{2k}, Y'_{2k})\}$.

Suppose that a_{2k+1} occurs an even number of times in $\{Y'_1, \ldots, Y'_{2k}\}$. We claim that all pairs $a_i a_{2k+1}$ are correct in $\mathcal{O}'_{2k}$; that is, $a_i a_{2k+1}$ occurs in an odd number of bicliques in $\mathcal{O}'_{2k}$ if and only if $a_i a_{2k+1} \in E(G)$. This can be checked as follows. Suppose $a_{2k+1} \in Y'_i$ and a_i occurs an even number of times in $\mathcal{O}_{2k}^Y$, as in case (i), so $a_i a_{2k+1} \in E(G)$. Since a_i is not in Y_i by assumption, it also occurs an even number of times in $\mathcal{O}_{2k}^Y - Y_i$. If $\{a_i, a_{2k+1}\} \in Y'_t$ for an even (resp. odd) number of $t \in [2k] - i$, then an even (resp. odd) number of these Y'_t contain a_i but not a_{2k+1}, and an odd (resp. even) number of these Y'_t contain a_{2k+1} but not a_i. Since both a_i and a_{2k+1} occur in every biclique in $\mathcal{O}'_{2k}$ except (X'_i, Y'_i), the edge $a_i a_{2k+1}$ occurs in an odd number of bicliques in $\mathcal{O}'_{2k}$. In other words, the pair $a_i a_{2k+1}$ is correct. By identical reasoning, if $a_{2k+1} \in Y'_i$ and a_i occurs an odd number of times in $\mathcal{O}_{2k}^Y$ (as in case (ii), so that $a_i a_{2k+1} \notin E(G)$), the edge $a_i a_{2k+1}$ occurs in an even number of bicliques in $\mathcal{O}'_{2k}$. This proves the claim.

In the case described above, where a_{2k+1} occurs an even number of times in $\{Y'_1, \ldots, Y'_{2k}\}$, we define $X'_{2k+2} = X_{2k+2} \cup a_{2k+1}$ and $Y'_{2k+2} = Y_{2k+2} \cup b_{2k+1}$. On the other hand, if a_{2k+1} occurs an odd number of times in $\{Y'_1, \ldots, Y'_{2k}\}$, a similar argument to the one above shows that the pairs $a_i a_{2k+1}$, $i \in [2k]$, are all incorrect in $\mathcal{O}'_{2k}$. In this case, we define $X'_{2k+2} = X_{2k+2} \cup b_{2k+1}$ and $Y'_{2k+2} = Y_{2k+2} \cup a_{2k+1}$. Since every $a_i \in X_{2k+2}$, $i \in [2k]$, by the inductive hypothesis, all pairs $a_i a_{2k+1}$ are now correct in the resulting collection $\mathcal{O}' = \mathcal{O}'_{2k} \cup \{(X_{2k+1}, Y_{2k+1}), (X'_{2k+2}, Y'_{2k+2})\}$. By symmetry, all pairs $b_i b_{2k+1}$ are also correct in $\mathcal{O}'$. Note that the edge $a_{2k+1} b_{2k+1}$ occurs an odd number of times in $\mathcal{O}'$, that $a_i b_{2k+1}$ occurs an odd number of times if and only if $a_i a_{2k+1}$ occurs an odd number of times for $i \in [2k]$, and similarly for the pairs $b_i a_{2k+1}$ and $b_i b_{2k+1}$. Since the pairs a_i, b_i are all adjacent twins in G, (I) is satisfied.

Case 2 ($n = 2k + 2$). Let $\mathcal{O}' = \{(X'_i, Y'_i) : i \in [n]\}$ be the odd cover of $G - \{a_n, b_n\}$ obtained from $G - \{a_{n-1}, b_{n-1}, a_n, b_n\}$ in the manner described in Case 1. We will add a_n and b_n to each biclique in $\mathcal{O}'$ except (X'_n, Y'_n), as well as two new bicliques, to obtain an odd cover of G which satisfies (II).

For each $i \in [n-1]$, we add a_n to Y'_i if either a_i occurs an even number of times in $\{Y'_1, \ldots, Y'_{n-1}\}$ and $a_i a_n \in E(G)$, or if a_i occurs an odd number of times in $\{Y'_1, \ldots, Y'_{n-1}\}$ and $a_i a_n \notin E(G)$. Otherwise, we add a_n to Y'_i. Add b_n to the opposite partite set in any case. Let $\mathcal{O}'' = \{(X''_i, Y''_i) : i \in [n]\}$ denote the resulting collection of bicliques.

Using similar reasoning as in Case 1, one can check that the pairs $a_i a_n$ for $i \in [n-1]$ are all correct in $\mathcal{O}''$ if a_n occurs an even number of times in $\{Y''_i : i \in [n]\}$, or are all incorrect otherwise. In the former case, we define $X_{n+1} = \{a_i : i \in [n]\}$ and $Y_{n+1} = \{b_i : i \in [n]\}$. In the latter, we define $X_{n+1} = \{a_i : i \in [n-1]\} \cup b_n$ and $Y_{n+1} = \{b_i : i \in [n-1]\} \cup a_n$. Now, defining $X_{n+2} = \{a_i : i \in [n]\}$ and $Y_{n+2} = \{b_i :$

$i \in [n]\}$, one can check that the collection $\mathcal{O}'' \cup \{(X_{n+1}, Y_{n+1}), (X_{n+2}, Y_{n+2})\}$ meets (II), completing the inductive step.

When n is odd, (I) provides us with an odd cover of G using $n+1$ bicliques. This implies $b_2(G) \leq n+1$. On the other hand, when n is even, (II) only provides us with an odd cover of G using $n+2$ bicliques. However, we notice that the symmetric difference $(X_{n+1}, Y_{n+1}) \triangle (X_{n+2}, Y_{n+2})$ is also a biclique; it is either degenerate, or $(\{a_n, b_n\}, \{a_i, b_i : i \in [n-1]\})$. Therefore, this also implies $b_2(G) \leq n+1$. □

We now show that, when n is a multiple of 8, K_n has a perfect odd cover. See Figure 3.1 for an example of our construction. Although the phrasing of the proof below differs from [18], the construction is the same.

Theorem 3.35 ([18]). *If n is a multiple of 8, then $b_2(K_n) = n/2$.*

Proof. By Proposition 3.6, it suffices to prove the upper bound. Let V denote the vertex set of K_n, $V = \{v_1, \ldots, v_{4k}, w_1, \ldots, w_{4k}\}$. We construct a $\mathcal{B}$-odd cover $\{(X_i, Y_i) : i \in [4k]\}$ as follows. For each $i \in \{1, \ldots, 4k\}$, (X_i, Y_i) contains every vertex except v_i and w_i. We define only the sets X_i below, as $Y_i = V - (X_i \cup \{v_i, w_i\})$.

For each i divisible by 4, we let

$$X_i = \{v_j : j \leq i-1 \text{ or } j = i+1\} \cup \{w_j : j \geq i+2\}.$$

For each $i \equiv 1 \pmod 4$, we let

$$X_i = \{v_j : j \leq i-1\} \cup \{w_j : j \geq i+1\}.$$

For each $i \equiv 2 \pmod 4$, we let

$$X_i = \{v_j : j \leq i-2\} \cup \{w_j : j = i-1 \text{ or } j \geq i+1\}.$$

Finally, for each $i \equiv 3 \pmod 4$, we let

$$X_i = \{v_j : j = i+1 \text{ or } j \leq i-2\} \cup \{w_j : j = i-1 \text{ or } j \geq i+2\}.$$

It is clear to see that $\{v_1, v_2, v_3, v_4\}$ is a clique. Further, for each $j \in [4k]$, the edge $v_j w_j$ is in every biclique except (X_j, Y_j), and thus occurs an odd number of times. Secondly, for each $j \in [4k-4]$, the edges uv_j and uv_{j+4} occur in the same number of bicliques for each $u \in V - \{v_j, v_{j+4}\}$, and uw_j and uw_{j+4} occur in the same number of bicliques for each $u \in V - \{w_j, w_{j+4}\}$. It remains to show that the edges $v_j v_{j+4}$ and $w_j w_{j+4}$ occur an odd number of times when $j \in [4k-4]$, as well as the edges $v_j w_k$ when $j \notin \{k-4, k, k+4\}$ or $k \notin \{k-4, k, k+4\}$. By the symmetry of our construction, for every vertex $u \in V - \{v_i, w_i\}$, the edge uv_j occurs in the same number of bicliques as uw_j. In particular, uv_j occurs in an odd number of (X_i, Y_i) if and only if uw_j also occurs in an odd number of (X_i, Y_i). This completes the proof. □

We shall now prove that cliques of order $18 \pmod{24}$ also have perfect odd covers. We use a pairs construction, as defined by Radhakrishnan, Sen, and Vishwanathan to determine perfect odd covers of K_n when $n = 2(q^2+q+1)$ and $q \equiv 3 \pmod 4$ or when there exists an $n/2 \times n/2$ Hadamard matrix [83].

In a *pairs construction* for an even clique K_n, the vertices are partitioned into ordered pairs (u, v) and, for each biclique (X, Y), $u \in X$ if and only if $v \in Y$ and $v \in X$ if and only if $u \in Y$. The construction is described by an $n/2 \times n/2$ matrix

M with entries from $\{0, \pm 1\}$, where $M_{i,j} = 0$ if the ith pair (u, v) of vertices does not occur in the jth biclique (X_j, Y_j); $M_{i,j} = 1$ if $u \in X_j$ and $v \in Y_j$; and $M_{i,j} = -1$ if $v \in X_j$ and $u \in Y_j$. Rephrasing Lemma 1 of [83] slightly, we see that a pairs construction yields a perfect odd cover of K_n if and only if every row of the matrix M contains an odd number of nonzero entries and, for any pair of distinct rows of M, the number of entries in which one row has a 1 and the other has a -1, as well as the number of entries in which both are 1 or both are -1, is odd.

Theorem 3.36 ([17])**.** *Let k be a nonnegative integer. If $n = 24k+18$, then $b_2(K_n) = n/2$.*

Proof [17]. Note that n is divisible by 6. We construct M as a block matrix

$$M = \begin{pmatrix} A & C & B \\ C & B & A \\ B & A & C \end{pmatrix}$$

where each block is an $(n/6) \times (n/6)$ matrix, A consists entirely of 1's, B consists entirely of 0's, and C is defined as follows:

$$c_{ij} = \begin{cases} 0: & i = j, \\ 1: & (j > i \text{ and } j - i \text{ odd}) \text{ or } (j < i \text{ and } j - i \text{ even}), \\ -1: & (j > i \text{ and } j - i \text{ even}) \text{ or } (j < i \text{ and } j - i \text{ odd}). \end{cases}$$

To verify that this construction yields a perfect odd cover, we note that each row of M has $n/6$ ± 1's from A, and an additional $n/6 - 1$ from C, for a total of $n/3 - 1$, which is odd. Now for each pair of distinct rows, we must verify that there are the correct number of entries where both are 1 or both are -1 and the correct number of entries where one row has a 1 and the other a -1.

First consider rows i and j which are both in the first $n/6$ rows, both in the next $n/6$ rows, or both in the last $n/6$ rows. Without loss of generality, we may assume that $i < j \leq n/6$. Then the total number of entries where both are 1 or both are -1 is $n/6$ (from the A block), plus the number of columns in C where rows i and j have the same ± 1 entry. The number of columns where they have different ± 1 entries is the number of columns in C where rows i and j have different ± 1 entries. As $n/6$ is odd, it suffices to verify that two distinct rows of C have both an even number of columns where they are the same ± 1 entry and an odd number of columns where they are different ± 1 entries. Note that for two distinct rows of C, there are only two columns where one row or the other has a zero. As there are an odd number, $n/6$, of total columns, having an even number of columns where the two rows have the same ± 1 entry guarantees that there are also an odd number of columns where the two rows have different ± 1 entries. Thus it suffices to check that there are an even number of columns such that rows i and j of C have the same ± 1 entry. If j and i have the same parity, this happens precisely for columns after column j and those before column i, for a total of $n/6 - (j - i + 1)$. Since $n \equiv 18 \pmod{24}$, this is even. If instead j and i have opposite parity, then this happens precisely for columns in between i and j for a total of $j - i - 1$, which is again even.

Otherwise, without loss of generality, we have that i is in the first $n/6$ rows and j is in the next $n/6$ rows. For any column aside from the first $n/6$ columns, at least one of these rows has a 0. For all the remaining columns, row i corresponds to an A block, so its only relevant entries are 1's. Thus it suffices to check that for row $j - n/6$

of C that there are both an odd number of 1's and an odd number of -1's. Indeed, any row of C has $\frac{n/6-1}{2}$ of each. Since $n \equiv 18 \pmod{24}$, this is odd, as desired. $\square$

We noted in [17] that the above construction, but replacing $n \equiv 18 \pmod{24}$ with $n \equiv 6 \pmod{24}$, gives a perfect odd cover of $3K_{n/3}$. That is, for any nonnegative integer k, $3K_{8k+2}$ has a perfect odd cover. It is also worth noting that $b_2(2K_{2n_1} + \cdots + 2K_{2n_k}) = 2\sum n_i$ regardless of the values of the n_i by Theorem 3.5 and Proposition 3.6. We posed the following question regarding an odd number of copies of an even cliques. Note that, if tK_{2n} has a perfect odd cover for some odd t, then so does $(t+i)K_{2n}$ for $i \geq 2$, for if i is odd, then we are back in the case of an even number of copies of K_{2n}, and if i is even, we take a perfect odd cover of iK_{2n} and a perfect odd cover of tK_{2n}.

Question 3.1 ([17]). *For every value of n, is there some odd t where tK_{2n} has a perfect odd cover?*

We conclude with a few final notes on $\mathcal{B}$-odd covers of even cliques. Firstly, we included in [17] a proof of István Tomon that $b_2(K_{3^k-1}) = \frac{3^k-1}{2}$ for any nonnegative integer k. We also generalized this construction to find perfect odd covers of even cliques which are distinctly different than those given in Theorems 3.35 and 3.36. However, these cases are all handled by the theorems we have already proven, and we do not include them here.

Let us now examine a few properties of perfect odd covers of even cliques which may be useful for future research in the area. We begin with an implication of Corollary 3.22.

Corollary 3.37 ([17]). *Suppose that $\mathcal{O}$ is a perfect odd cover of an even clique. If I is a nonempty subset of vertices in the clique, then at least one partite set of a biclique*

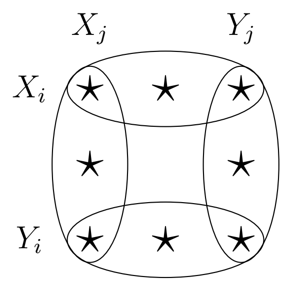


Figure 3.7: A pair of bicliques $(X_i, Y_i), (X_j, Y_j)$ in a perfect odd cover of an even clique (see items (i) and (ii) of Theorem 3.38). A $\star$ denotes an odd number of vertices.

in $\mathcal{O}$ has odd intersection with I.

We now prove more specific properties possessed by perfect odd covers of even cliques, supporting our belief that K_{2n} does not have a perfect odd cover when n is 2 or 3 modulo 4. This is conjectured in [17]. It is noted that our construction for K_{8n} can also be phrased as a pairs construction, but that no pairs construction can exist for K_{2n} when n is 2 or 3 modulo 4 [17, Theorem 18].

Figure 3.7 illustrates some of the information contained in the first two parts of the following theorem, our final result in this chapter.

Theorem 3.38 ([17])**.** *Suppose that K_{2k} has a perfect odd cover with bicliques $\{(X_i, Y_i) : i \in [k]\}$. The following conditions hold:*

(i) If k is odd, then $|X_i|, |Y_i| \equiv 1 \pmod 4$ for all $1 \le i \le k$. If k is even, then $|X_i|, |Y_i| \equiv 3 \pmod 4$ for all $1 \le i \le k$.

(ii) For all $i, j \in \{1, \ldots, k\}$ with $i \neq j$, $|X_i \cap X_j|$, $|X_i \cap Y_j|$ and $|Y_i \cap Y_j|$ are all odd.

(iii) Let $U_i = X_i \cup Y_i$, $1 \le i \le k$. Each vertex is contained in odd number of U_i's.

(iv) For any integer s equivalent to 2 or 3 modulo 4, and for any set $A \subseteq V(K_{2k})$ of size s, there exists an i such that $|A \cap X_i|$ and $|A \cap Y_i|$ are odd.

Proof [17]. We prove items (i)–(iv) in order.

(i) We begin by proving that $|X_i|, |Y_i|$ are odd. For a given i, let $Z_i := V(K_{2k}) - (X_i \cup Y_i)$. Note that in the graph $K_{2k} \triangle B_i$, any two vertices in the same one of X_i, Y_i, or Z_i are adjacent twins. Thus, if $|X_i|, |Y_i|$ are both even, we can form a perfect matching where each edge of the matching is between a pair of adjacent twins. Thus by Lemma 3.8, $\text{rk}_2(K_{2k} \triangle B_i) = 2k$. Therefore, it takes at least a further k bicliques to complete the odd cover of K_{2k} and thus it will not be a perfect odd cover. Suppose instead that one of $|X_i|, |Y_i|$ is odd and the other is even. Without loss of generality, $|X_i|$ is odd and thus, so is $|Z_i|$. We can pair up all but one vertex in X_i, all vertices in Y_i, and all but one vertex in Z_i to form a matching M of $k-1$ edges where each edge is between a pair of adjacent twins. By Lemma 3.8, $\text{rk}_2(K_{2k} \triangle B_i) = 2(k-1) + \text{rk}_2(K_{2k} \triangle B_i - V(M))$. Note that $K_{2k} \triangle B_i - V(M)$ has just two vertices, but as one is in X_i and the other is in Z_i, there is an edge between them. Thus $\text{rk}_2(K_{2k} \triangle B_i - V(M)) = 2$, and $\text{rk}_2(K_{2k} \triangle B_i) = 2k$, meaning again it will take at least a further k bicliques to complete the odd cover of K_{2k}. The only remaining option for K_{2k} to have a perfect odd cover is if $|X_i|, |Y_i|$ odd for each biclique in the odd cover.

Now we determine $|X_i|, |Y_i| \pmod 4$. Consider the graph $H := K_{2k} \triangle (X_i, Y_i)$. Both the induced graphs $H[X_i \cup Z_i]$ and $H[Y_i \cup Z_i]$ are complete graphs with an odd number of vertices. Thus, any vertex of H has an even number of neighbors in each of $X_i \cup Z_i, Y_i \cup Z_i$, so $X_i \cup Z_i, Y_i \cup Z_i$ are both even cores of H. If either $|X_i \cup Z_i|$ or $|Y_i \cup Z_i|$ is $3 \pmod 4$, then the complete graph on that vertex set has an odd number of edges, so by Corollary 3.22, we would then get that an odd cover of H requires at least $\frac{\text{rk}_2(H)}{2} + 1$ bicliques. Note that $\text{rk}_2(H) \geq \text{rk}_2(K_{2k}) - 2$, so an odd cover of H would require at least $\mathrm{rk}_2(K_{2k})/2$ bicliques, meaning an odd cover of K_{2k} would require more than that. Therefore, both $|X_i \cup Z_i|, |Y_i \cup Z_i|$, which are odd, must be $1 \pmod 4$. If k odd, this yields $|X_i|, |Y_i| \equiv 1 \pmod 4$, while if k even, this yields $|X_i|, |Y_i| \equiv 3 \pmod 4$.

(ii) For a given pair $i \neq j$, vertices that are in both the same one of X_i, Y_i, Z_i and the same one of X_j, Y_j, Z_j are adjacent twins. Suppose that not all of $|X_i \cap X_j|, |X_i \cap Y_j|, |Y_i \cap X_j|$, and $|Y_i \cap Y_j|$ are odd. Without loss of generality, there are five ways for this to happen:

(I) All are even.

(II) All but $|Y_i \cap Y_j|$ are even.

(III) $|X_i \cap X_j|, |X_i \cap Y_j|$ are even and the other two are odd.

(IV) $|X_i \cap X_j|, |Y_i \cap Y_j|$ are odd and the other two are even.

(V) All but $|Y_i \cap Y_j|$ are odd.

In each case, we will form a matching of adjacent twins and then apply Lemma 3.8 to determine $\mathrm{rk}_2(K_{2k} \triangle B_i \triangle B_j)$. In Case (I), we obtain a matching with $k-2$ edges. The remaining vertices are in $|Z_i \cap X_j|, |Z_i \cap Y_j|, |X_i \cap Z_j|, |Y_i \cap Z_j|$, so they constitute a C_4, which has rank 2. Thus, we get $\mathrm{rk}_2(K_{2k} \triangle B_i \triangle B_j) = 2(k-2)+2 = 2k-2$, meaning at least $k-1$ further bicliques are required to complete the odd cover of K_{2k}. In Case (II), we obtain a matching with $k-2$ edges. The remaining vertices, which are in $|Z_i \cap X_j|, |Y_i \cap Y_j|, |X_i \cap Z_j|, |Z_i \cap Z_j|$ constitute a C_3 with a pendant edge, which has rank 4. Thus, we get $\mathrm{rk}_2(K_{2k} \triangle B_i \triangle B_j) = 2(k-2)+4 = 2k$, meaning at least k further bicliques are required to complete the odd cover of K_{2k}. In Case (III), we obtain a matching with $k-2$

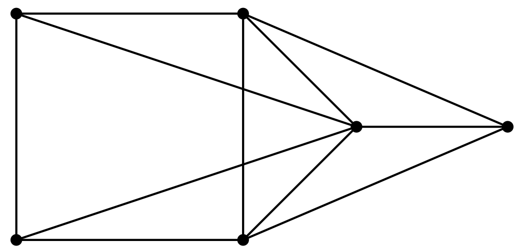


Figure 3.8: The graph obtained in Case (V) *of the proof of Theorem 3.38* (ii) *after removing a maximum matching of adjacent twins from* $K_{2k} \triangle B_i \triangle B_j$ [17].

edges. The remaining vertices, which are in $|Y_i \cap X_j|, |Y_i \cap Y_j|, |Y_i \cap Z_j|$, and $|X_i \cap Z_j|$ constitute a $K_{1,2}$ with an isolated vertex, which has rank 2. Thus, we get $\mathrm{rk}_2(K_{2k} \triangle B_i \triangle B_j) = 2(k-2)+2 = 2k-2$, meaning at least $k-1$ further bicliques are required to complete the odd cover of K_{2k}. In Case (IV), we obtain a matching with $k-1$ edges. The remaining vertices are in $|X_i \cap X_j|, |Y_i \cap Y_j|$, so they constitute a K_2, which has rank 2, Thus, $\mathrm{rk}_2(K_{2k} \triangle B_i \triangle B_j) = 2(k-1)+2 = 2k$, meaning at least k further bicliques are required to complete the odd cover of K_{2k}. In Case (V), we obtain a matching with $k-3$ edges. The remaining vertices are in $|X_i \cap X_j|, |X_i \cap Y_j|, |X_i \cap Z_j|, |Y_i \cap X_j|, |Z_i \cap X_j|, |Z_i \cap Z_j|$, which form a graph of rank 4, in particular, the complement of the graph consisting of a path of length 4 and an isolated vertex; see Figure 3.8. Thus, $\mathrm{rk}_2(K_{2k} \triangle B_i \triangle B_j) = 2(k-3)+4 = 2k-2$, meaning at least $k-1$ further bicliques are required to complete the odd cover of K_{2k}. Thus, it is only possible to form a perfect odd cover of K_{2k} if $|X_i \cap X_j|, |X_i \cap Y_j|, |Y_i \cap X_j|$, and $|Y_i \cap Y_j|$ are all odd.

(iii) Each vertex has odd degree in K_{2k}, so must be contained in an odd number of edges across all bicliques. However, each vertex has odd degree in each biclique in which it occurs, so it must appear in an odd number of bicliques.

(iv) Note that the number of edges in A is $\binom{s}{2}$, which is an odd number. To cover each edge in A odd number of times, we need odd number of edges. Hence there exists a biclique B_i with odd number of edges in A. This is only possible when $|X_i \cap A|$ and $|Y_i \cap A|$ are both odd.

$\square$

Part II

Graph saturation problems

Chapter 4

Introduction to graph saturation

Saturation problems concern the possible sizes of graphs which are maximal with respect to a given property, in the sense that no edge can be added between nonadjacent vertices while retaining said property. Classically, the property in question is to avoid some fixed forbidden subgraph. Such problems date back to the very beginnings of extremal graph theory, as we shall presently describe. We direct the reader to Section 1.1 for definitions and notations which we refrain from redefining in this part.

4.1 Extremal graph theory

Extremal combinatorics can be broadly described as the study of global parameters of combinatorial objects subject to (typically local) constraints. In the aptly named field of extremal graph theory, the combinatorial objects in question are graphs. Although Paul Erdős attributed the initiation of the field to a 1941 paper of Paul Turán [39,89], the study of Turán-type problems, as they are now known, dates at least to 1906 when

Willem Mantel asked for the maximum number of edges in a graph on n vertices which contains no triangle. The first published solution is due to Willem Wythoff in 1907, who determined the answer to be $\lfloor n^2/4 \rfloor$ [74].

The aforementioned 1941 paper of Turán concerns the more general problem of avoiding p-cliques for any $p \geq 3$. Note that a graph whose vertices can be partitioned into $p-1$ independent sets (a $(p-1)$-partite graph) contains no p-clique by the pigeonhole principle. Such a graph G is said to be $(p-1)$-*chromatic*, and the *chromatic number* $\chi(G)$ is the smallest value χ for which G is χ-chromatic. To state Turán's theorem, we define the eponymous *Turán graph* $T^p(n)$ to be the complete p-partite graph of order n whose partite sets are all of size $\lfloor n/p \rfloor$ or $\lceil n/p \rceil$. Note that $T^p(p) = K_p$; as a convention, we let $T^p(n) = K_n$ for $n \leq p$. The graph $T^4(9)$ is depicted in Figure 4.1a. Note that $\|T^p(n)\| = (1 - \frac{1}{p})\frac{n^2}{2} - \frac{s(p-s)}{2p}$.

Theorem 4.1 (Turán's theorem [89]). *Let n and p be positive integers, $p \geq 2$. Over all graphs of order n which do not contain a p-clique, $T^{p-1}(n)$ uniquely contains the maximum number of edges.*

In 1946, Erdős and Stone generalized Turán's theorem, showing that any graph of sufficiently large order n with asymptotically more edges than $T^{p-1}(n)$ contains a copy of $T^p(n')$, where $n' = \sqrt{\ln^{r-1}(n)}$ [44]. Then, in 1966, it was noted by Erdős and Simonovits that the methods in the former paper generalize again to finding copies of *any* graph in a sufficiently dense host graph.

Theorem 4.2 (Erdős-Stone-Simonovits theorem [43]). *Any graph with asymptotically more edges than $T^{\chi-1}(n)$ contains every graph with chromatic number χ. That is, for any $\varepsilon > 0$, there exists n_0 such that, if G is a graph of order $n \geq n_0$ with at least $\|T^{\chi-1}(n)\| + \varepsilon n^2$ edges, then G contains every graph with chromatic number χ.*

This resolves the problem of finding what is now known as the extremal number, or Turán number, $\text{ex}(n, H)$ asymptotically. Given graphs G and H, we say that G is *H-free* if it does not contain H as a subgraph. We always assume that H has at least one edge, or else there are no H-free graphs on at least $|H|$ vertices. The *extremal number* $\text{ex}(n, H)$ is the maximum number of edges in an H-free graph of order n. By the Erdős-Stone-Simonovits theorem, for any graph H with chromatic number χ,

$$\text{ex}(n, H) = \left(1 - \frac{1}{\chi - 1}\right)\frac{n^2}{2} + o(n^2).$$

As a note, when $\chi = 2$, their theorem simply says that $\text{ex}(n, H) = o(n^2)$, and this case remains (for the most part) wide open. The problem of determining the asymptotics of $\text{ex}(n, C_{2k})$, for example, is known only for $k \in \{2, 3, 5\}$ [52] and is one of the most famous open problems in the area.

Variations and generalizations of the extremal number abound in the literature, from finding the maximum number of copies of a fixed graph other than K_2 in an H-free graph of order n [5,38,94], to a spectral version of the extremal number [77], to an edge-colored version known as the *rainbow Turán number* [61] which will be relevant in Section 6.3. Here, we are only grazing the surface of Turán-type problems, yet the field of extremal graph theory has grown to include many other types of problems as well. Included in these are a wealth of problems stemming from the famous Ramsey's theorem [84].

4.2 Saturation problems

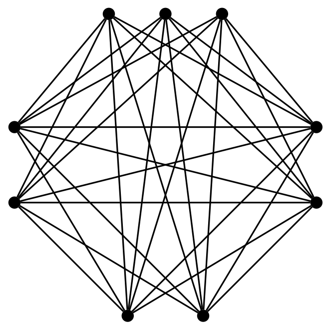
In 1949, Soviet mathematician A. A. Zykov [94] (perhaps unknowingly) reproved Turán's theorem, along with a notable generalization, using a method now known as Zykov symmetrization. In doing so, he considered not only graphs which have the maximum number of edges while avoiding a p-clique, but all of the graphs whose edge sets are maximal with respect to avoiding a p-clique. Zykov called these graphs *p-saturated.*

In 1964, Erdős, Hajnal, and Moon studied the problem of minimizing the number of edges in a graph of order n to which the addition of any edge increases the number of p-cliques. We call such a graph *p-semisaturated.*[1] Note that a p-semisaturated graph which does not contain a p-clique is p-saturated. They determined that the minimum size of a p-semisaturated graph of order n is attained by a unique graph, consisting of a single $(p-2)$-clique joined to an independent set of size $n-p+2$. This is the complete $(p-1)$-partite graph whose partite sets are as unbalanced as possible, as opposed to the balanced complete $(p-1)$-partite graph, $T^{p-1}(n)$. Figure 4.1 depicts both graphs for $p = 4$ and $n = 9$.

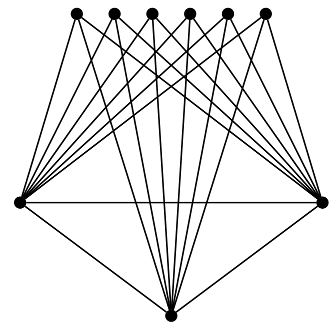
Theorem 4.3 ([42]). *The number of edges in a p-(semi)saturated graph of order n is minimized by a unique graph with $(p-2)(n-p+2)+\binom{p-2}{2}$ edges.*

Both p-saturation and p-semisaturation were quickly generalized to other forbidden graphs, in the vein of the function $\mathrm{ex}(n, H)$, as well as to host graphs other than K_n. As an instance of the latter, the problem of finding the minimum size a subgraph of $K_{a,b}$ which is edge-maximal with respect to not containing $K_{s,t}$ has been

[1] Such graphs have also been called strongly p-saturated, *e.g.*, in [13].



(a) The Turán graph $T^4(9)$



(b) The graph in Theorem 4.3

Figure 4.1: The graphs of maximum and minimum size over all K_5-saturated graphs on nine vertices.

studied [12, 93]). We focus on the former generalization. Let H be a graph with at least one edge. A graph G is said to be *H-semisaturated* if the addition of any edge to G increases the number of copies of H (*i.e.*, the number of subgraphs isomorphic to H). If G is both H-free and H-semisaturated, we say that G is *H-saturated*. Equivalently, G is H-saturated if it is maximally H-free. We denote by $\mathrm{ssat}(n, H)$ the minimum number of edges in an H-semisaturated graph of order n, called the *semisaturation number of H*. Similarly, $\mathrm{sat}(n, H)$ denotes the minimum number of edges in an H-saturated graph, the *saturation number of H*.

The graph saturation has received considerable attention over the years. We direct the reader to [32] for a survey. In 1972, Ollmann determined the saturation numbers of squares C_4; $\mathrm{sat}(n, C_4) = \lfloor (3n-5)/2 \rfloor$ [78]. Bollobás asked about saturation numbers of larger cycles in [13], but these are still not known in general. In fact, it was not until 2009 that $\mathrm{sat}(n, C_5)$ was determined to be $10n/7 + O(1)$ [26] (and then determined more precisely in 2011 [27]), and only in 2021 was $\mathrm{sat}(n, C_6)$ determined to be $4n/3 + O(1)$ [64]. For larger cycles C_k, $k \geq 7$, Füredi and Kim showed that $\left(1 + \frac{1}{k+2}\right)n - 1 < \mathrm{sat}(n, C_k) < \left(1 + \frac{1}{k-4}\right)n + \binom{k-4}{2}$ for $n \geq 2k-5$ [51]. The authors also studied semisaturation numbers of cycles, showing that $\left(1+\frac{1}{2k-2}\right)n-2 <$ $\mathrm{ssat}(n, C_k) < \left(1+\frac{1}{2k-10}\right)n+k-1$ for $n \geq k \geq 6$. Thus, for $k > 12$, $\mathrm{ssat}(n, C_k)$ is asymptotically smaller than $\mathrm{sat}(n, C_k)$. They also showed this to be the case for C_5, and noted that they believed it likely to be true for $k \in \{6, \ldots, 12\}$ as well. We will see that a similar result holds for paths in Chapter 6, due to a result of Burr [21].

In a 1986 paper [60], Kászonyi and Tuza laid much of the ground work for the study of saturation numbers. Indeed, they proved a sharp upper bound which remains (asymptotically) the best general upper bound today. In particular, they proved $\mathrm{sat}(n, H) = O(n)$ for every graph H. This marks a major difference between $\mathrm{sat}(n, H)$ and $\mathrm{ex}(n, H)$, for we recall that the extremal number is quadratic for all graphs which are not bipartite by the Erdős-Stone-Simonovitz theorem. Saturation numbers can also be constant. For instance, when H has an isolated edge (*i.e.*, a pair of adjacent degree-1 vertices), an n-vertex graph consisting of a clique of order $|H|-1$ and $n-|H|+1$ isolated vertices is H-saturated. Thus, $\mathrm{sat}(n, H) \leq \binom{|H|-1}{2}$ for every n. Kászonyi and Tuza also proved that such graphs H are the only graphs with constant saturation numbers [60]. The analogous result holds for semisaturation numbers, and thus to determine a saturation or semisaturation number asymptotically is to find an appropriate constant w for which $\mathrm{sat}(n, H) = wn + o(n)$.

In fact, it is not known that such a constant w exists for every graph H. Saturation (and semisaturation) numbers are volatile in comparison to their counterpart, extremal numbers. While $\mathrm{ex}(n, H)$ is easily seen to be monotone with respect to subgraphs ($\mathrm{ex}(n, H') \leq \mathrm{ex}(n, H)$ whenever $H' \subseteq H$) and with respect to n ($\mathrm{ex}(n, H) \leq \mathrm{ex}(n+1, H)$), neither $\mathrm{sat}(n, H)$ nor $\mathrm{ssat}(n, H)$ possess either of these properties. For instance, $\mathrm{sat}(2n-1, P_4) = n$ while $\mathrm{sat}(2n, P_4) = n-1$ [60] (see

(a) A P_4-saturated graph on nine vertices of minimum size

(b) A P_4-saturated graph on ten vertices of minimum size

Figure 4.2: Two P_4-saturated graphs, on nine and ten vertices, respectively, illustrate a lack of monotonicity of $\mathrm{sat}(n, H)$ *with respect to* n.

(a) $K_{1,5}^{+} + K_3$

(b) An $(K_{1,5}^{+} + K_3)$*-saturated graph*

Figure 4.3: We have $\mathrm{sat}(n, K_{1,5}^{+} + K_3) = 5n/6 + O(1)$ *[20] while* $\mathrm{sat}(n, K_{1,5}^{+}) = 6n/7 + O(1)$ *[47], illustrating a lack of monotonicity of* $\mathrm{sat}(n, H)$ *with respect to subgraph inclusion.*

Figure 4.2). Despite this, it remains an open conjecture of Tuza that the saturation number is at least close to being monotone with n; that is, that $\lim_{n\to\infty} \mathrm{sat}(n, H)/n$ exists for every graph H [90].

With regards to subgraph inclusion, the saturation number can indeed vary asymptotically in either direction. For instance, it is noted in [47] that the tree of order p with the largest saturation number is $K_{1,p-1}$ (Kászonyi and Tuza also determined $\mathrm{sat}(n, K_{1,p-1}) = (p-2)n/2 + O(1)$ in [60]), while the smallest saturation number of a tree with order p is $n - \lfloor (n + p - 2)/p \rfloor$, witnessed uniquely by the graph $K_{1,p-2}^{+}$ obtained by subdividing a single edge of a star (see Figure 2.1b). Another example of nonmonotonicity with respect to subgraph inclusion, this time by restricting to a single connected component of a disconnected graph, is depicted in Figure 4.3.

Chapter 5

A lower bound on the saturation number and a strengthening for triangle-free graphs

Here, we prove various lower bounds on the semisaturation number of an arbitrary graph H using the degrees of endpoints of edges in H as well as the degrees of their neighbors. We recall that, since every H-saturated graph is H-semisaturated, we have $\mathrm{ssat}(n, H) \leq \mathrm{sat}(n, H)$, and thus these are also lower bounds on $\mathrm{sat}(n, H)$. The theorems proved herein are the result of a collaboration with Puck Rombach [20].

Though it does not appear to have been directly stated in the literature before our paper [20], a relatively trivial bound on semisaturation (Proposition 5.1) follows as an easy corollary as the main result of [24]. We state it formally here, as the idea behind it serves as the foundation for each of the subsequently described lower bounds.

Definition 5.1 (wt_0, k_0)**.** Let H be a graph. For each edge uv in H, define $\mathrm{wt}_0(uv) = \max\{d(u), d(v)\} - 1$, and let $k_0 = \min_{uv \in E(H)} \{\mathrm{wt}_0(uv)\}$.

A graph is said to be *d-regular* if every vertex has degree d. We now show that an H-semisaturated graph cannot have many fewer edges than a k_0-regular graph.

Proposition 5.1. *For any graph H and integer $n \geq |H|$,*

$$\mathrm{ssat}(n, H) \geq k_0 n/2 - (k_0 + 1)^2/8.$$

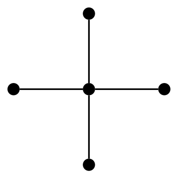
Proof. Let G be an H-semisaturated graph of order n. If G is complete, then $\|G\| = n(n-1)/2 > k_0 n/2$ since $k_0 \leq |H| - 2 < n - 1$. In this case, we are done, so we assume G has a pair of nonadjacent vertices x and y. By definition, xy is contained in a copy of H in $G + xy$, the graph obtained by adding the edge xy to G. Let uv be the edge in H mapped to xy in one such copy. Since $\max\{d_H(u), d_H(v)\} \geq k_0 + 1$, at least one of x or y has degree at least $k_0 + 1$ in $G + xy$, and thus has degree at least k_0 in G.

It follows that at least one out of any pair of nonadjacent vertices in G has degree at least k_0. In other words, the set of vertices with degree strictly less than k_0 in G form a clique A. Therefore, letting $a = |A|$,

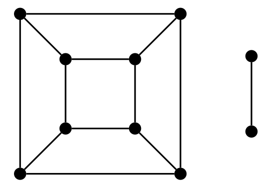
$$\begin{aligned}\sum_{v \in V(G)} d_G(v) \geq (n-a)k_0 + a(a-1) = k_0 n + a^2 - (k_0+1)a \\ \geq k_0 n - \frac{(k_0+1)^2}{4}.\end{aligned}$$

The well-known "handshake lemma" states that, since every edge contributes 2 to the degree sum of G, $\|G\| = \sum d(v)/2$. Therefore, $\|G\| \geq k_0 n/2 - (k_0+1)^2/8$, as desired. □

The *average degree* of a graph G of order n is $\sum_{v \in V(G)} d_G(v)/n$. We frequently use



(a) The star $K_{1,4}$ has $k_0 = 3$



(b) A minimum $K_{1,4}$-semisaturated graph

Figure 5.1: Evidence that the bound in Proposition 5.1 is sharp; the disjoint union of a k_0-regular graph and a clique of order $(k_0+1)/2$ meets the bound.

the handshake lemma in this chapter and in Section 6.3, making statements along the lines of "the average degree of an H-semisaturated graph cannot be much less than X" to mean that $\mathrm{ssat}(n,H) \geq \mathrm{X}n/2 - c$ for a constant c which does not depend on n. We denote the average degree of G by $d(G)$. As it will make our lives easier, we conflate this notation for subsets S of $V(G)$ as well; that is, $d(S) = \sum_{v \in S} d(v)/|S|$. By convention, we let $d(\varnothing) = 0$.

While Proposition 5.1 provides a relatively trivial general lower bound on $\mathrm{ssat}(n,H)$, it is still sharp for some fundamental classes of graphs, like stars [60] (see Figure 5.1). A nontrivial lower bound was proven by Cameron and Puleo in 2022 [24]. It combines the idea behind Proposition 5.1, that the endpoints of an edge added to an H-semisaturated graph G must have sufficiently large degree to play the role of an edge in H, with the idea that, if those endpoints are contained in a triangle in a resulting copy of H, then they must have neighbors in common.

More precisely, for each edge uv in H, let $\mathrm{wt}_\triangle(uv) = |N(u) \cap N(v)|$, that is, the number of triangles in H containing uv. Although the following theorem is phrased in terms of saturation numbers in [24], the argument generalizes to semisaturation numbers.

Theorem 5.2 ([24])**.** *Let H be a graph, and let $w = \min_{uv \in E(H)} \{\mathrm{wt}_0(uv) + \mathrm{wt}_\triangle(uv)\}$.*

There is a constant c depending only on H such that, for any integer $n \geq |H|$,

$$\mathrm{ssat}(n, H) \geq wn/2 - c.$$

Theorem 5.2 is asymptotically sharp for a number of graphs. Included in these are *threshold graphs*, obtained from a single vertex by iteratively adding isolated or dominating vertices, and disjoint unions of cliques [24]. Indeed, we note that every edge in K_p has $\mathrm{wt}_0 = \mathrm{wt}_\triangle = p - 2$, and thus we recover the asymptotic lower bound for Theorem 4.3. On the other hand, whenever H has an edge minimizing wt_0 which is not contained in any triangle, Cameron and Puleo's bound reduces to the one in Proposition 5.1.

In what follows, we address the question: how much larger than k_0 must the average degree of an H-semisaturated graph be when there exists an edge minimizing wt_0 which is not contained in any triangles? We first provide an answer in the form of a general lower bound (Theorem 5.4) using the degrees of neighbors of a given edge in H, in addition to the degrees of its endpoints. We then provide a stronger lower bound (Theorem 5.8) for two different classes of graphs H, both containing the class of triangle-free graphs. Section 5.1 is devoted to proving the general lower bound, and Section 5.2 the strengthening.

5.1 A general lower bound

Our answer to the question "how much larger than k_0 must the average degree of an H-semisaturated graph be?" depends on the degrees of neighbors of edges uv in H. By the *neighborhood* $N_H(uv)$ of an edge uv, we mean the set of neighbors of u or v

other than u and v themselves; that is, $N(uv) = (N(u) - v) \cup (N(v) - u)$.

Definition 5.2 (wt_1, k_1)**.** Let H be a graph. For each edge uv in H with a nonempty neighborhood, define $\mathrm{wt}_1(uv) = \max_{w \in N(uv)} \{d(w)\}$, and define $\mathrm{wt}_1(uv) = 0$ for any isolated edges. Further, let $k_1 = \min_{uv \in E(H)} \{\mathrm{wt}_1(uv)\}$.

Let G be an H-semisaturated graph. For an edge added to G to be contained in a copy of H, not only must one of its endpoints have degree at least k_0, but also one of its endpoints must have a neighbor of degree k_1 (the smallest possible degree of a neighbor of an edge in H). It is possible, however, that all of the edges uv in H having $\mathrm{wt}_1(uv) = k_1$ also have $\mathrm{wt}_0(uv) > k_0$. In this case, if both endpoints of the edge added to G have degree k_0, at least one of these endpoints must have a neighbor of degree strictly larger than k_1. To describe exactly what this larger degree should be, and to determine how large of a degree the endpoints should have to ensure a neighbor of degree strictly larger than k_1, we introduce two more parameters.

Definition 5.3 (k_0', k_1')**.** Let wt_0, k_0, wt_1, and k_1 be as in Definitions 5.1 and 5.2. We define

$$k_0' = \min_{\substack{uv \in E(H) \\ \mathrm{wt}_1(uv) = k_1}} \{\mathrm{wt}_0(uv)\} \qquad \text{and} \qquad k_1' = \min_{\substack{uv \in E(H) \\ \mathrm{wt}_0(uv) = k_0}} \{\mathrm{wt}_1(uv)\}.$$

Note that $k_0 = k_0'$ if and only if $k_1 = k_1'$. Otherwise, $k_0 < k_0'$ and $k_1 < k_1'$. We can now say more precisely which pairs of nonadjacent vertices in G need neighbors of which degrees. The following proposition summarizes a number of observations made in [20].

Proposition 5.3. *For any pair of nonadjacent vertices x, y in an H-semisaturated* *graph G, we have* $\max\{d(x), d(y)\} \geq k_0$*, and, for some* $z \in N(x) \cup N(y)$*,*

$$d(z) \geq \begin{cases} k_1' : & \max\{d(x), d(y)\} \leq k_0; \\ k_1 + 1 : & \max\{d(x), d(y)\} < k_0'; \\ k_1 : & \max\{d(x), d(y)\} \geq k_0'. \end{cases}$$

Hence, the subsets of vertices x in G

- $A = \{x : d(x) < k_0\}$*;*
- $B = \{x : d(x) \leq k_0$ *and x has no neighbor of degree at least* $k_1'\}$*;*
- $C = \{x : d(x) < k_0'$ *and x has no neighbor of degree strictly larger than* $k_1\}$*; and*
- $D = \{x : x$ *has no neighbor of degree at least* $k_1\}$

are (not necessarily disjoint) cliques, of orders at most k_0*,* $k_0 + 1$*,* $\min\{k_0', k_1 + 1\}$*, and* k_1*, respectively.*

We will refer to the cliques A, B, C, and D above throughout this chapter. Since each has size bounded by a parameter depending only on H, they contribute negligibly to the average degree of an H-semisaturated graph of large order n. In other words, an H-semisaturated graph G cannot have average degree much less than a graph G' with minimum degree k_0 in which every degree-k_0 vertex has a neighbor of degree at least k_1', every vertex of degree less than k_0' has a neighbor of degree larger than k_1, and every vertex of degree at least k_0' has a neighbor of degree at least k_1. The bulk of this section is devoted to finding the minimum average degree of such a graph G'. We also find specific graphs G' which attain this minimum for each set of possible relations between k_0, k_1, k'_0, and k'_1. These minimum sizes are reflected in the following theorem, summarizing our general lower bounds on semisaturation numbers.

Theorem 5.4 ([20]). *Let H be a graph with at least one edge and no isolated edges. There is a constant c depending on H such that, for any $n \geq |H|$,*

$$\operatorname{ssat}(n, H) \geq \left(k_0 + \frac{k'_1 - k_0}{k'_1 + 1}\right)\frac{n}{2} - c.$$

Further, if $k_1 > k_0$, then $\operatorname{ssat}(n, H) \geq \Big(k_0 + (k'_1 - k_0)/k'_1\Big)n/2 - c$, and if $k_0 = k_1 < k'_1 < k'_0$, then

$$\operatorname{ssat}(n, H) \geq \begin{cases} \left(k_0 + \frac{k'_0 - k_0}{k'_0 + 1}\right)\frac{n}{2} - c: & k'_0 \leq k'_1 + \frac{k'_0 - k_0}{k_0 + 1}; \\ \left(k_0 + \frac{k'_1 - k_0}{k'_1}\right)\frac{n}{2} - c: & \textit{otherwise}. \end{cases}$$

We prove Theorem 5.4 in two parts. In Lemma 5.6, we deal with the cases which do not involve k'_0, proving first two lower bounds. In Lemma 5.7, we handle the remaining cases (those where $k'_0 > k'_1$), noting a transition between the constructions which minimize the average degree at $k'_0 = k'_1 + (k'_0 - k_0)/(k_0 + 1)$.

5.1.1 Let's not worry about k'_0

We begin with a warm up. Rather than diving directly into semisaturation, we first determine the minimum average degree of a graph in which every vertex of minimum degree δ has a neighbor of degree at least k. Trivially, if $k \leq \delta$, this minimum is δ, so we suppose $k > \delta$. We also determine the minimum average degree of such a graph in which every vertex of degree k also has a neighbor of degree strictly larger than δ. (See the discussion of the graph G' preceding the statement of Theorem 5.4, replacing δ with k_0 and k with k_1'.)

Proposition 5.5 ([20])**.** *Let δ and k be positive integers with $\delta < k$. If G is a graph with minimum degree δ in which every vertex of degree δ has a neighbor of degree at least k, then $d(G) \geq \delta + (k-\delta)/(k+1)$. If, in addition, every vertex in G of degree at least k has a neighbor of degree strictly larger than δ, then $d(G) \geq \delta + (k-\delta)/k$.*

Proof. We partition the vertex set V of G as follows: let $S = \{v \in V : d(v) = \delta\}$, $M = \{v \in V : \delta < d(v) < k\}$, and $L = \{v \in V : d(v) \geq k\}$. By assumption, every vertex in S has a neighbor in L, so $e(L,S) \geq |S| = |L \cup S| - |L|$. Since $e(L,S) \leq \sum_{v \in L} d(v) = d(L)|L|$, we have $|L \cup S| \leq (d(L)+1)|L|$. Let $\ell = d(L)$. Combining the aforementioned inequalities yields

$$|L| \geq \frac{1}{\ell+1}|L \cup S| \quad \text{and} \quad |S| \leq \frac{\ell}{\ell+1}|L \cup S|.$$

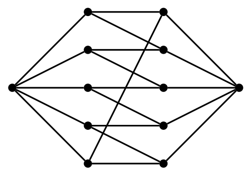
Thus,

$$\begin{aligned}\sum_{v \in V} d(v) &\geq \ell|L| + \delta|S| + (\delta+1)|M| \geq \frac{\ell(\delta+1)}{\ell+1}|L \cup S| + (\delta+1)|M| \\ &\geq \left(\delta + \frac{\ell-\delta}{\ell+1}\right)|G|.\end{aligned}$$

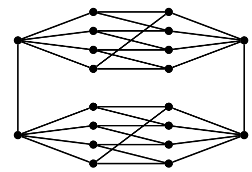
Since $\ell \geq k$ and

$$\frac{\ell-\delta}{\ell+1} = \frac{k-\delta}{k+1} + \frac{(\delta+1)(\ell-k)}{(\ell+1)(k+1)} \tag{5.1}$$

we have $d(G) = \sum d(v)/|G| \geq \delta + (k-\delta)/(k+1)$, as desired.



(a) A graph minimizing the average degree over all graphs in which every minimum-degree (3) vertex has a degree-5 neighbor



(b) A graph minimizing the average degree over all graphs with minimum degree 3 in which every vertex has a degree-5 neighbor

Figure 5.2: Graphs whose average degrees match the lower bounds given in Proposition 5.5

For the second statement, if every vertex in L has a neighbor in $V - S$, then $|S| \leq e(L, S) \leq \sum_{v \in L}(d(v) - 1) = (\ell - 1)|L|$. In this case, $|L| \geq |L \cup S|/\ell$ and $|S| \leq (\ell - 1)|L \cup S|/\ell$. By a similar argument as before, we have

$$\sum_{v \in V} d(v) \geq \left(\delta + \frac{\ell - \delta}{\ell}\right)|G|.$$

Since

$$\frac{\ell - \delta}{\ell} = \frac{k - \delta}{k} + \frac{\delta(\ell - k)}{k\ell} \tag{5.2}$$

we have $d(G) \geq \delta + (k - \delta)/k$, as desired. □

Constructions of graphs that minimize the average degree under the conditions in Proposition 5.5 fall right out of the proof. We simply need to find graphs with one degree-k vertex for every k degree-δ vertices, in the first case, or one degree-k vertex for every $k - 1$ degree-δ vertices in the second. Figure 5.2 depicts examples of such graphs. Essentially, we take an even number of copies of $K_{1,k}$ and add a $(\delta - 1)$-regular graph on the set of leaves (Figure 5.2a), or we take an even number of copies of $K_{1,k-1}$, add a matching on their centers, and again add a $(\delta - 1)$-regular graph on the leaves (Figure 5.2b).

Lemma 5.6 ([20]). *For any graph H with $k_1' > k_0$, and for any $n \geq |H|$,*

$$\mathrm{ssat}(n, H) \geq \left(k_0 + \frac{k_1' - k_0}{k_1' + 1}\right)\frac{n}{2} - c_1.$$

If, in addition, $k_1 > k_0$, then

$$\mathrm{ssat}(n, H) \geq \left(k_0 + \frac{k_1' - k_0}{k_1'}\right)\frac{n}{2} - c_2,$$

where $c_1 = \frac{(k_0+1)(k_1'-k_0)}{2k_1'+2} + \frac{(k_0+1)^2}{8}$ and $c_2 = \frac{(k_0+2)(k_1'-k_0)}{2k_1'} + \frac{(k_0+1)^2}{8}$.

Proof. Let G be an H-semisaturated graph of order n. Partition the vertex set V of G as follows: let $S = \{v \in V : d(v) \leq k_0\}$, $M = \{v \in V : k_0 < d(v) < k_1'\}$, and $L = \{v \in V : d(v) \geq k_1'\}$. We may assume S is nonempty, or else the statement is trivial. Let A and B be the cliques in S given by Proposition 5.3; that is, $A = \{v \in S : d(v) < k_0\}$ and $B = \{v \in S : N(v) \cap L = \varnothing\}$. Since every vertex in $S - B$ has a neighbor in L, $e(L, S) \geq |S - B| = |L \cup S| - |L| - |B|$, and clearly $e(L, S) \leq \sum_{v \in L} d(v) = |L|d(L)$. Letting $\ell = d(L)$, it follows that $|L \cup S| - |B| \leq |L|(\ell + 1)$, so

$$|L| \geq \frac{1}{\ell+1}|L \cup S| - \frac{|B|}{\ell+1} \quad \text{and} \quad |S| \leq \frac{\ell}{\ell+1}|L \cup S| + \frac{|B|}{\ell+1}.$$

Thus,

$$\ell|L| + k_0|S| \geq \frac{\ell + k_0\ell}{\ell+1}|L \cup S| - \frac{\ell - k_0}{\ell+1}|B| = \left(k_0 + \frac{\ell - k_0}{\ell+1}\right)|L \cup S| - \frac{\ell - k_0}{\ell+1}|B|.$$

Using equation (5.1) and noting that $|L \cup S| \geq |B|$, it follows that

$$\ell|L| + k_0|S| \geq \left(k_0 + \frac{k_1' - k_0}{k_1' + 1}\right)|L \cup S| - \frac{k_1' - k_0}{k_1' + 1}|B|.$$

Since $|B| \leq k_0 + 1$, we have

$$\begin{aligned}\sum_{v \in L \cup S} d(v) &= \ell|L| + k_0|S - A| + \sum_{v \in A} d(v) \geq \ell|L| + k_0|S| + |A|(|A| - 1 - k_0) \\ &\geq \left(k_0 + \frac{k_1' - k_0}{k_1' + 1}\right)|L \cup S| - \frac{(k_0 + 1)(k_1' - k_0)}{k_1' + 1} - \frac{(k_0 + 1)^2}{4}.\end{aligned}$$

Every vertex in M has degree at least $k_0 + 1$ by definition, and S, M, and L partition V, so the degree sum of G is at least $\big(k_0 + (k_1' - k_0)/(k_1' + 1)\big)n - 2c_1$.

For the second statement, suppose $k_1 > k_0$. Letting D be the clique of vertices without a neighbor of degree at least k_1, as in Proposition 5.3, we note that $|D \cap L| \leq 1$ since $k_1 \leq k_1'$. Thus, $e(L, S) \leq \sum_{v \in L}(d(v) - 1) + 1 = (\ell - 1)|L| + 1$. Since $e(L, S) \geq |L \cup S| - |B| - |L|$, we now have $|L \cup S| - |B| \leq \ell|L| + 1$. If $\ell = 0$ (*i.e.*, if $L = \varnothing$), then $S = B$. Otherwise,

$$|L| \geq \frac{1}{\ell}|L \cup S| - \frac{1}{\ell}(|B| + 1) \quad \text{and} \quad |S| \leq \frac{\ell - 1}{\ell}|L \cup S| + \frac{1}{\ell}(|B| + 1).$$

Also, in this case, $|L \cup S| \geq |B| + 1$, so that using equation (5.2) we have

$$\ell|L| + k_0|S| \geq \left(k_0 + \frac{k_1' - k_0}{k_1'}\right)|L \cup S| - \frac{(k_0 + 2)(k_1' - k_0)}{k_1'}.$$

Note that the above inequality still holds (and is strict) when $L = \varnothing$. Thus, by the same reasoning as before, the degree sum of G is at least $\big(k_0 + (k_1' - k_0)/k_1'\big)n - 2c_2$.

The handshake lemma completes the proof. □

5.1.2 Now we worry about k_0'

Let H be a graph with $k_0 = k_1 < k_1' < k_0'$. Recall from Proposition 5.3 that, in an H-semisaturated graph G, almost every vertex of degree at most k_0 has a neighbor of degree at least k_1', and almost every vertex of degree less than k_0' (including those of degree k_1') have a neighbor of degree larger than k_1 (and thus larger than k_0). The constructions discussed in the previous section (depicted in Figure 5.2) give us two ideas for such a graph of minimum size: either all vertices have degree either k_0 or k_1', with two degree-k_1' vertices for every $2(k_1' - 1)$ degree-k_0 vertices (as in Figure 5.2b); or all vertices have degree either k_0 or k_0', with one degree-k_0' vertex for every k_0' degree-k_0 vertices (as in Figure 5.2a).

Example 5.1 ([20])**.** Let us compare, as k_0' varies, the average degree of a graph of the form given in Figure 5.2a with vertices of degree k_0 and k_0' to the average degree of one as in Figure 5.2b with vertices of degree k_0 and k_1'. Note that the former graph has average degree $k_0+(k_0'-k_0)/(k_0'+1)$ and the latter $k_0+(k_1'-k_0)/k_1'$. Suppose that $k_1' = 4$ and $k_0 = k_1 < k_1'$. If $k_0' = 6$, then $(k_0'-k_0)/(k_0'+1) = 5/7 < (k_1'-k_0)/k_1' = 3/4$. However, if $k_0' = 8$, then $(k_0' - k_0)/(k_0' + 1) = 7/9 > 3/4$. If instead $k_0' = 7$, then the two quantities are equal. In general, we have

$$\frac{k_0' - k_0}{k_0' + 1} \leq \frac{k_1' - k_0}{k_1'} \quad \text{if and only if} \quad k_0' - k_1' \leq \frac{k_0' - k_0}{k_0 + 1}. \tag{5.3}$$

We conclude this section, and the proof of Theorem 5.4, by determining that these constructions are optimal. That is, when $k_0' > k_1'$, these graphs have minimum average degree over all graphs with minimum degree k_0 in which every degree-k_0 vertex has a neighbor of degree at least k_1', and every vertex of degree strictly less than k_0' has a neighbor of degree strictly greater than k_0.

Lemma 5.7 ([20])**.** *For any graph H with $k_0 = k_1 < k_1' < k_0'$, and for any $n \geq |H|$,*

$$\mathrm{sat}(n, H) \geq \begin{cases} \left(k_0 + \frac{k_0'-k_0}{k_0'+1}\right)\frac{n}{2} - c_1 : & k_0' \leq k_1' + \frac{k_0'-k_0}{k_0+1}; \\ \left(k_0 + \frac{k_1'-k_0}{k_1'}\right)\frac{n}{2} - c_2 : & k_0' \geq k_1' + \frac{k_0'-k_0}{k_0+1}, \end{cases}$$

where $c_1 = \frac{(k_0+1)(k_0'-k_0)}{2k_0'+2} + \frac{(k_0+1)^2}{8}$ *and* $c_2 = \frac{(k_0+2)(k_1'-k_0)}{2k_1'} + \frac{(k_0+1)^2}{8}$.

Proof. Let G be an H-semisaturated graph of order n. Partition the vertex set V of G as follows: let $S = \{v \in V : d(v) \leq k_0\}$, $M = \{v \in V : k_0 < d(v) < k_1'\}$, $L = \{v \in V : k_1' \leq d(v) < k_0'\}$, and $XL = \{v \in V : d(v) \geq k_0'\}$. Let A and B be as in Proposition 5.3. We partition $S - B$ into subsets S_L and S_{XL} of vertices with a neighbor in L or XL, respectively (if a vertex has neighbors in both L and XL, assign it to either set arbitrarily). We will show that $d(L \cup S_L)$ is not much less than $k_0 + (k_1' - k_0)/k_1'$ if L is nonempty, and that $d(XL \cup S_{XL})$ is not much less than $k_0 + (k_0' - k_0)/(k_0' + 1)$ if XL is nonempty.

First, suppose $L \neq \varnothing$ and consider $L \cup S_L$. At least one out of any pair of nonadjacent vertices in L has a neighbor in $V - S$, since $d(v) < k_0'$ for all $v \in L$. It follows that at most one vertex in L has all of its neighbors in S, so that $|S_L| \leq e(L, S_L) \leq \sum_{v \in L}(d(v) - 1) + 1$. That is, $|L \cup S_L| - |L| \leq |L|d(L) - |L| + 1$. Letting $\ell = d(L)$, we have

$$|L| \geq \frac{1}{\ell}|L \cup S_L| - \frac{1}{\ell} \quad \text{and} \quad |S_L| \leq \frac{\ell - 1}{\ell}|L \cup S_L| + \frac{1}{\ell}.$$

Thus,

$$\begin{aligned}\ell|L| + k_0|S_L| &\geq \frac{\ell + k_0(\ell-1)}{\ell}|L \cup S_L| - \frac{\ell - k_0}{\ell} \\ &\geq \left(k_0 + \frac{k_1' - k_0}{k_1'}\right)|L \cup S_L| - \frac{k_1' - k_0}{k_1'}.\end{aligned}$$

Note that if $L = \varnothing$, the final inequality above still holds, and is strict.

Now consider $XL \cup S_{XL}$. We have $|S_{XL}| \leq e(XL, S_{XL}) \leq \sum_{v \in XL} d(v)$. Letting $x = d(XL)$, we have

$$|XL| \geq \frac{1}{x+1}|XL \cup S_{XL}| \quad \text{and} \quad |S_{XL}| \leq \frac{x}{x+1}|XL \cup S_{XL}|.$$

Thus,

$$x|XL| + k_0|S_{XL}| \geq \frac{x(k_0+1)}{x+1}|XL \cup S_{XL}| \geq \left(k_0 + \frac{k_0' - k_0}{k_0'+1}\right)|XL \cup S_{XL}|.$$

We have

$$\sum_{v \in V-M} d(v) = \Big(x|XL| + k_0|S_{XL}|\Big) + \Big(\ell|L| + k_0|S_L|\Big) + k_0|B| - \sum_{s \in A}(k_0 - d(s)).$$

If $k_0' - k_1' \geq (k_0' - k_0)/(k_0+1)$, then by (5.3),

$$\sum_{v \in V-M} d(v) \geq \frac{k_1' + k_0(k_1'-1)}{k_1'}|V - M| - \frac{k_1' - k_0}{k_1'}(|B|+1) - |A|(k_0 + 1 - |A|).$$

It follows that the degree sum of G is at least $\big(k_0 + (k_1' - k_0)/k_1'\big)n - 2c_2$. Otherwise, if $k_0' - k_1' \leq (k_0' - k_0)/(k_0 + 1)$, then

$$\sum_{v \in V-M} d(v) \geq \frac{k_0'(k_0+1)}{k_0'+1}|V-M| - \frac{k_0'-k_0}{k_0'+1}|B| - |A|(k_0+1-|A|).$$

In this case, the degree sum of G is at least

$$\left(k_0 + \frac{k_0'-k_0}{k_0+1}\right)n - \frac{(k_0+1)(k_0'-k_0)}{k_0'+1} - \frac{(k_0+1)^2}{4}.$$

The handshake lemma completes the proof. □

5.2 Strengthenings

For a graph H with $k_1' > k_0$ in which some edge minimizing wt_0 is not contained in any triangles, Theorem 5.4 provides a stronger lower bound on $\mathrm{ssat}(n, H)$ than Cameron and Puleo's. Conversely, if every edge minimizing wt_0 is contained in at least one triangle, then $\mathrm{wt}_0(uv) + \mathrm{wt}_\triangle(uv) \geq k_0 + 1$ for every $uv \in E(H)$. All of the asymptotic lower bounds proven here on the average degree of an H-semisaturated graph are bounded (strictly) above by $k_0 + 1$, so Cameron and Puleo's bound outperforms ours in this case. However, the case which motivates our work is that of triangle-free graphs, and for these we can improve upon Theorem 5.4. In fact, our improvements hold for larger classes of graphs H than triangle-free graphs. As the classes can be a bit unwieldy to state, we phrase our results in terms of the more well-studied class of triangle-free graphs, noting the precise classes with the corresponding lemmas which make up the proof of our main result.

Theorem 5.8 ([20])**.** *Let H be a triangle-free graph, and let $n \geq |H|$. If $k_1' \geq$*

$k_0 + \sqrt{2k_0+1}$, or if $k_1' \geq k_0+2$ and at least one degree-(k_0+1) endpoint of every edge in H minimizing wt_0 *has a neighbor of degree at least k_1', then there is a constant c depending only on H such that*

$$\mathrm{ssat}(n, H) \geq \left(k_0 + \frac{k_1' + 1 - k_0}{k_1' + 2}\right)\frac{n}{2} - c. \tag{5.4}$$

If, in addition to either of the above conditions, $k_1 > k_0$, then

$$\mathrm{ssat}(n, H) \geq \left(k_0 + \frac{k_1' + 1 - k_0}{k_1' + 1}\right)\frac{n}{2} - c. \tag{5.5}$$

Essentially, Theorem 5.8 states that, under the given conditions, the average degree of an H-semisaturated graph cannot be much less than that of a graph as depicted in Figure 5.2a, in the first case, or Figure 5.2b in the second, but with high-degree vertices of degree $k_1' + 1$ and low-degree vertices of degree k_0. As we will see in Chapter 6, the former construction provides an upper bound on the saturation number to match Theorem 5.8 for certain trees called unbalanced double stars.

We prove Theorem 5.8 in Sections 5.2.1 and 5.2.2. In Section 5.2.3, we note that, as a corollary of our proof techniques, one can also obtain strengthenings of Theorem 5.4 for triangle-free graphs H which do not meet the conditions on k_1' in terms of k_0 in Theorem 5.8. This last result will also be used in Chapter 6 to provide an improved lower bound on the semisaturation numbers of certain trees called caterpillars.

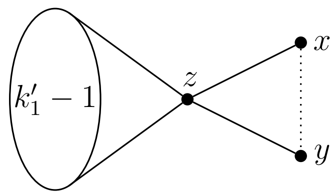


Figure 5.3: Nonadjacent low-degree vertices x and y which share their high-degree neighbor(s) in an H-semisaturated graph when H is triangle-free

5.2.1 Extra-high-degree neighbors, pt. 1

Let H be a graph in which none of the edges minimizing wt_0 are contained in any triangles. Let G be an H-semisaturated graph, and let x and y be nonadjacent vertices in G whose degrees are at most k_0. We shall call such vertices *low-degree*, and vertices of degree at least k'_1 *high-degree*. Recall that, in any copy of H containing xy in $G + xy$, xy plays the role of an edge $uv \in E(H)$ with $\mathrm{wt}_0(uv) = k_0$. Thus, there exists a neighbor $z \in N(x) \cup N(y)$ such that $d_G(z) \geq k'_1$. However, not any old high-degree neighbor z will do, for if z is adjacent to both x and y, then it must have at least $k'_1 - 1$ *other* neighbors in G in order to play the role of a neighbor w of uv in H with $d(w) = k'_1$ (see Figure 5.3). In other words, G has the following property:

for any pair of nonadjacent vertices x, y with degrees at most k_0, there exists $z \in N(x) \cup N(y)$ such that $|N(z) - \{x, y\}| \geq k'_1 - 1$. (P1)

We can make a number of similar (stronger) claims about G, x, and y when the endpoints of the minimum-wt_0 edges in H are also not contained in any triangles. In this case, since x and y are low-degree, whichever one plays the role of a degree-$(k_0 + 1)$ endpoint of an edge in a copy of $H \subseteq G + xy$ containing xy, say y, must use all of its incident edges in this copy. And, since these edges are not contained in any triangles, the vertex z playing the role of a high-degree neighbor of y in H must have $k_1' - 1$ edges outside of $N(y) \cup \{x, y\}$. In other words, for any pair of nonadjacent low-degree vertices x, y in an H-semisaturated graph G, there exists $z \in N(x) \cup N(y)$ such that $|N(z) - (N(x) \cup y)| \geq k_1'$ or $|N(z) - (N(y) \cup x)| \geq k_1'$. We will use a similar idea to this one in the following subsection to prove Lemma 5.10.

For now, we only use the simpler property (P1), along with similar techniques used for Lemma 5.6, to prove part of Theorem 5.8. The important implication of (P1) is that, for any vertex z in an H-semisaturated graph with $d(z) = k_1'$, the low-degree neighbors of z which only have z for a high-degree neighbor form a clique.

Lemma 5.9 ([19]). *Let H be a graph in which none of the edges minimizing* wt_0 *are contained in any triangles, and let $n \geq |H|$. If $k_1' \geq k_0 + \sqrt{2k_0 + 9/4} - 1/2$, then the inequality* (5.4) *holds. If $k_1' \geq k_0 + \sqrt{2k_0 + 1}$ and $k_1 > k_0$, then* (5.5) *holds.*

Proof. Let G be an H-semisaturated graph on vertex set V, $|V| = n$. We partition V into sets S, M, L, and XL, letting $S = \{v : d(v) \leq k_0\}$, $M = \{v : k_0 < d(v) < k_1'\}$, $L = \{v : d(v) = k_1'\}$, and $XL = \{v : d(v) > k_1'\}$. Note that L and XL differ than in the proof of Lemma 5.7. Let A and B be the cliques in Proposition 5.3; $A = \{v \in S : d(v) < k_0\}$ and $B = \{v \in S : N(v) \cap (L \cup XL) = \varnothing\}$.

We begin with the first statement. Suppose $k_1' \geq k_0 + \sqrt{2k_0 + 9/4} - 1/2$. We handle the degree sum over XL and the set S_{XL} of vertices in S with a neighbor in XL in a nearly identical manner as we proved the first statement of Lemma 5.6 or the second statement of Lemma 5.7. We have $|S_{XL}| \leq e(XL, S_{XL}) \leq d(XL)|XL|$ and $d(XL) \geq k_1' + 1$ so that

$$\sum_{v \in XL \cup S_{XL}} d(v) \geq \left(k_0 + \frac{k_1' + 1 - k_0}{k_1' + 2}\right) |XL \cup S_{XL}| - |A \cap S_{XL}|(k_0 + 1 - |A|). \quad (5.6)$$

We now restrict our attention to L and the set S_L of vertices in S whose only high-degree neighbors lie in L; $S_L = S - (B \cup S_{XL})$. By the property (P1), if x and y are vertices in S_L which share all of their neighbors in L, then $xy \in E(G)$. It follows that, for any $z \in L$, the set of vertices x in $N(z) \cap S_L$ whose only high-degree neighbor is z form a clique (of order at most k_0). Thus, at most $k_0|L|$ vertices in S_L have exactly one neighbor in L, so $2|S_L| - k_0|L| \leq e(L, S_L) \leq k_1'|L|$. This gives $|L| \geq \frac{2}{k_1'+k_0+2}|L \cup S_L|$ and $|S_L| \leq \frac{k_1'+k_0}{k_1'+k_0+2}|L \cup S_L|$. Therefore,

$$k_1'|L| + k_0|S_L| \geq \frac{2k_1' + k_0k_1' + k_0^2}{k_1' + k_0 + 2}|L \cup S_L|.$$

Note that

$$\frac{2k_1' + k_0k_1' + k_0^2}{k_1' + k_0 + 2} \geq \frac{(k_0+1)(k_1'+1)}{k_1'+2}$$

if and only if $k_1' \geq k_0 + \sqrt{2k_0 + 9/4} - 1/2$. This is precisely our supposition, and therefore

$$\sum_{v \in L \cup S_L} d(v) \geq \left(k_0 + \frac{k_1' + 1 - k_0}{k_1' + 2}\right)|L \cup S_L| - |A \cap S_L|(k_0 + 1 - |A|).$$

The degree sum over $S \cup L \cup XL$ is the degree sum over $L \cup S_L$, $XL \cup S_{XL}$ and B, so

$$\sum_{v \in S \cup L \cup XL} d(v) \geq \left(k_0 + \frac{k_1' + 1 - k_0}{k_1' + 2}\right)|L \cup XL \cup S| - \frac{k_1' + 1 - k_0}{k_1' + 2}|B| - |A|(k_0 + 1 - |A|).$$

Since $d(v) \geq k_0 + 1$ for all $v \in M$, and since S, M, L, and XL partition V, we have

$$\sum_{v \in V} d(v) \geq \left(k_0 + \frac{k_1' + 1 - k_0}{k_1' + 2}\right)n - \frac{(k_0+1)(k_1'+1-k_0)}{k_1'+2} - \frac{(k_0+1)^2}{4}.$$

Therefore, inequality (5.4) holds.

Now, we prove the second statement. Suppose $k_1 > k_0$ and $k_1' \geq k_0 + \sqrt{2k_0+1}$. Let D be as in Proposition 5.3. In particular, we note that $|D \cap (L \cup XL)| \leq 1$ since $k_1' \geq k_1$; that is, at most one vertex in $L \cup XL$ has all of its neighbors in S. Thus, $|S_{XL}| \leq e(XL, S_{XL}) \leq (d(XL)-1)|XL| + 1$, and

$$d(XL)|XL| + k_0|S_{XL}| \geq \left(k_0 + \frac{k_1'+1-k_0}{k_1'+1}\right)|XL \cup S_{XL}| - \frac{k_1'+1-k_0}{k_1'+1}. \qquad (5.7)$$

As before, $2|S_L| - k_0|L| \leq e(L, S_L)$, but this time $e(L, S_L) \leq (k_1'-1)|L| + 1$ since all but at most one vertex in L have neighbors outside of S when $k_1 > k_0$ (see the clique C in Proposition 5.3). This gives $2|L \cup S_L| \leq (k_1' + k_0 + 1)|L| + 1$, and thus

$$|L| \geq \frac{2|L \cup S_L| - 1}{k_1' + k_0 + 1} \quad \text{and} \quad |S| \leq \frac{(k_1' + k_0 - 1)|L \cup S_L| + 1}{k_1' + k_0 + 1}.$$

Therefore,

$$k_1'|L| + k_0|S_L| \geq \frac{(k_0+2)k_1' + k_0(k_0-1)}{k_1'+k_0+1}|L \cup S_L| - \frac{k_1'-k_0}{k_1'+k_0+1}.$$

Note that, since $k_1' \geq k_0 + \sqrt{2k_0+1}$,

$$\frac{(k_0+2)k_1' + k_0(k_0-1)}{k_1'+k_0+1} \geq k_0 + \frac{k_1'+1-k_0}{k_1'+1}.$$

Thus,

$$k_1'|L| + k_0|S_L| \geq \left(k_0 + \frac{k_1'+1-k_0}{k_1'+1}\right)|L \cup S_L| - \frac{k_1'-k_0}{k_1'+k_0+1}. \qquad (5.8)$$

Now, since we can't have both a vertex in L and a vertex in XL with all of its neighbors in S, we can eliminate the negative constant from at least one of (5.7) or (5.8); we eliminate the latter since $\frac{k_1'+1-k_0}{k_1'+1} > \frac{k_1'-k_0}{k_1'+k_0+1}$. Now, using this and the bounds (5.7) and (5.8), summing over L, XL, S_L, and S_{XL} (*i.e.*, over $V-(M \cup B)$), we obtain

$$\sum_{V-(M\cup B)} d(v) \geq \left(k_0 + \frac{k_1'+1-k_0}{k_1'+1}\right)|V-(M\cup B)| - \frac{k_1'+1-k_0}{k_1'+1} - |A|(k_0+1-|A|).$$

Finally, since $d(v) \geq k_0+1$ for all $v \in M$, $d(v) = k_0$ for all $v \in B$, and $|A|(k_0+1-|A|) \leq (k_0+1)^2/4$, we have

$$\sum_{v\in V} d(v) \geq \left(k_0 + \frac{k_1'+1-k_0}{k_1'+1}\right)n - \frac{(k_0+2)(k_1'+1-k_0)}{k_1'+1} - \frac{(k_0+1)^2}{4}.$$

Therefore, inequality (5.5) holds, completing the proof. □

Note that $\sqrt{2k_0+1} \geq \sqrt{2k_0+9/4} - 1/2$, and the difference between these two values is at most 1/2. In particular, Lemma 5.9 is only slightly stronger than Theorem 5.8 for graphs H in which $k_1 \leq k_0$ and some edge minimizing wt_0 does not have a degree-(k_0+1) endpoint with a neighbor of degree at least k_1'.

5.2.2 Extra-high-degree neighbors, pt. 2

We now complete the proof of Theorem 5.8. Our final lemma in this chapter applies to a smaller class of graphs H than Lemma 5.9. Here, we let H be a graph such that, for any edge minimizing wt_0, at least one higher-degree endpoint has a neighbor of degree k_1' and is not contained in any triangles. We will show that, if $k_1' > k_0+1$, then almost every low-degree vertex in a minimum H-semisaturated graph has a neighbor of degree at least $k_1'+1$; further, if $k_1 > k_0$ as well, then every vertex has a neighbor of degree at least $k_1'+1$. This is analogous to what we prove in Lemma 5.9.

Note that we could prove a similar statement by defining a new weight function wt_1^* on $E(H)$ which counts the maximum degree of a neighbor of a higher-degree endpoint of any edge in H. However, this would not result in a stronger bound; if the minimum value of wt_1^* over all edges minimizing wt_0 is strictly less than that of wt_1, then our improved statement would be no stronger than Theorem 5.4. This allows us to (thankfully) avoid using a third weight function to prove the following lemma.

Lemma 5.10 ([20])**.** *Let H be a graph in which, for every edge minimizing* wt_0*, at least one degree-*(k_0+1) *endpoint has a neighbor of degree k_1' and is not contained in any triangles. If $k_1' > k_0+1$, then the inequality* (5.4) *holds, and if we also have $k_1 > k_0$, then* (5.5) *holds.*

To prove Lemma 5.10, we require a stronger property than (P1) for graphs H in which every edge minimizing wt_0 has a degree-(k_0+1) endpoint which is not contained in any triangles and has a neighbor of degree k_1'.

Let G be an H-semisaturated graph, and let x and y be nonadjacent low-degree vertices in G. Consider a copy of H in $G+xy$ which uses the edge xy. Since xy plays the role of an edge minimizing wt_0 in H, at least one of x or y, say y, plays the role of a degree-(k_0+1) endpoint of this edge with a neighbor of degree k_1' and no common neighbors with this high-degree neighbor. It follows that all of the k_0+1 edges incident to y in $G+xy$ are used in the copy of H, and none of these make a triangle with the high-degree neighbor z of y in the copy of H. Thus, for some $z \in N(y)$, $|N(z) - N_{G+xy}(y)| \geq k_1'$. See Figure 5.4. In other words, G has the

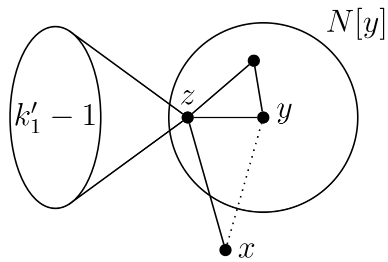


Figure 5.4: A high-degree neighbor z of y needs at least $k_1' - 1$ neighbors outside of $N(y) \cup \{x, y\}$ (see the property (P2)*).*

following property.

$$\text{for any pair of nonadjacent vertices } x, y \text{ with degrees at most } k_0, \text{ there exists either } z \in N(y) \text{ such that } |N(z) - (N(y) \cup x)| \geq k_1' \text{ or } z' \in N(y) \text{ such that } |N(z') - (N(y) \cup x)| \geq k_1'. \tag{P2}$$

With this in hand, we are ready to prove Lemma 5.10, finishing off the proof of Theorem 5.8.

Proof of Lemma 5.10. Let G be an H-saturated graph on vertex set V, $|V| = n$. Let S, M, L, and XL be the partition of V defined in the proof of Lemma 5.9, and let A and B be the cliques in Proposition 5.3 (as they have been throughout). Letting S_{XL} also be as in the proof of Lemma 5.9, we note that the bound (5.6) still holds.

Recall that the set S_L in the proof of Lemma 5.9 consisted of all vertices in S which are not in S_{XL} or B. Note that the vertices v in S_L which share a common neighbor with each of their neighbors in L form a clique by the property (P2) of G; indeed, if v' is another such vertex, then $|N(w) - (N(v) \cup v')| < k_1'$ and $|N(w') - (N(v') \cup v)| < k_1'$ for all $w \in N(v) \cap L$ and $w' \in N(w') \cap L$. Property (P2) also implies that, for any vertex $u \in B$, we must have $uv \in E(G)$. Let B' denote the clique in G consisting of B and the clique described above. Let $S_L' = S - (S_{XL} \cup B')$.

Let $v \in L$. If v has two neighbors in S'_L, each of them only having v for a high-degree neighbor, then these vertices are adjacent by property (P1), but then these vertices lie in B' by the discussion in the previous paragraph. Thus, at most $|L|$ vertices in S'_L have exactly one edge to L, so $2|S'_L| - |L| = 2|L \cup S'_L| - 3|L| \leq e(L, S'_L) \leq k'_1|L|$. It follows that $|L| \geq \frac{2}{k'_1+3}|L \cup S'_L|$ and $|S'_L| \leq \frac{k'_1+1}{k'_1+3}|L \cup S'_L|$. Thus,

$$k'_1|L| + k_0|S'_L| \geq \frac{2k'_1 + k_0(k'_1+1)}{k'_1+3}|L \cup S'_L|.$$

Note that $\frac{2k'_1+k_0(k'_1+1)}{k'_1+3} \geq k_0 + \frac{k'_1+1-k_0}{k'_1+2}$ if and only if $k'_1 \geq k_0 + 2$, which is true by supposition. Thus,

$$\sum_{v \in L \cup S'_L} d(v) \geq \left(k_0 + \frac{k'_1+1-k_0}{k'_1+2}\right)|L \cup S'_L| - |A \cap S'_L|(k_0+1-|A|).$$

Now, the degree sum over $S \cup L \cup XL$ is the degree sum over $L \cup S'_L$, $XL \cup S_{XL}$, and B', so

$$\sum_{v \in S \cup L \cup XL} d(v) \geq \left(k_0 + \frac{k'_1+1-k_0}{k'_1+2}\right)|L \cup XL \cup S| - \frac{k'_1+1-k_0}{k'_1+2}|B'| - |A|(k_0+1-|A|).$$

Noting that $d(v) \geq k_0+1$ for all $v \in M$ and that S, M, L, and XL partition V, we have

$$\sum_{v \in V} d(v) \geq \left(k_0 + \frac{k'_1+1-k_0}{k'_1+2}\right)n - \frac{(k_0+1)(k'_1+1-k_0)}{k'_1+2} - \frac{(k_0+1)^2}{4}.$$

That is, inequality (5.4) holds.

We now show that inequality (5.5) holds, supposing that $k_1 > k_0$. Note that the lower bound (5.7) on the degree sum over XL and S_{XL} in the proof of Lemma 5.9 holds for the same reasons. Also by similar reasoning, at most one vertex in L has all of its neighbors in S, so $e(L, S'_L) \leq (k'_1 - 1)|L| + 1$. We have seen that $2|L \cup S'_L| - 3|L| \leq e(L, S'_L)$. It follows that, $|L| \geq \frac{2}{k'_1+2}|L \cup S'_L| - \frac{1}{k'_1+2}$ and $|S'_L| \leq \frac{k'_1}{k'_1+2}|L \cup S'_L| + \frac{1}{k'_1+2}$. Therefore,

$$\begin{aligned} k'_1|L| + k_0|S'_L| &\geq \frac{2k'_1 + k_0k'_1}{k'_1+2}|L \cup S'_L| - \frac{k'_1 - k_0}{k'_1+2} \\ &\geq \left(k_0 + \frac{k'_1 + 1 - k_0}{k'_1 + 1}\right)|L \cup S'_L| - \frac{k'_1 - k_0}{k'_1 + 2}, \end{aligned}$$

where the second inequality holds by our supposition $k'_1 > k_0 + 1$.

As in the proof of the second statement of Lemma 5.9, since at most one vertex in $L \cup XL$ has all of its neighbors in S, we can eliminate one of the negative constants in our lower bounds on the degree sum over $L \cup S'_L$ and $XL \cup S_{XL}$. As such, we obtain

$$\sum_{V-(M\cup B')} d(v) \geq \left(k_0 + \frac{k'_1 + 1 - k_0}{k'_1 + 1}\right)|V - (M \cup B')| - \frac{k'_1 + 1 - k_0}{k'_1 + 1} - |A|(k_0 + 1 - |A|).$$

Finally, since $d(v) \geq k_0 + 1$ for all $v \in M$, $d(v) = k_0$ for all $v \in B'$, and $|A|(k_0 + 1 - |A|) \leq (k_0 + 1)^2/4$, we have

$$\sum_{v\in V} d(v) \geq \left(k_0 + \frac{k'_1 + 1 - k_0}{k'_1 + 1}\right)n - \frac{(k_0 + 2)(k'_1 + 1 - k_0)}{k'_1 + 1} - \frac{(k_0 + 1)^2}{4}.$$

Thus, inequality (5.5) holds, completing the proof. □

5.2.3 One more slight improvement

Concerning the triangle-free graphs H which do not meet the conditions on k_1' in terms of k_0 in Lemmas 5.9 and 5.10, we can still obtain improvements over our general lower bound Theorem 5.4. Indeed, with a bit of extra arithmetic, these improvements can be read off directly from the proofs. We do not state each of these improvements in detail (the interested reader can easily go back and obtain the bounds they desire from the corresponding proofs). We note one particular such bound below as it will be useful in the following chapter.

Corollary 5.11 ([20]). *Let H be a graph in which at least one higher-degree endpoint of any edge minimizing wt_0 has a neighbor of degree k_1' and is not contained in any triangles. If $k_1' = k_1 = k_0 + 1$, then for any $n \geq |H|$,*

$$\mathrm{ssat}(n, H) \geq \left(k_0 + \frac{2}{k_0+3}\right)\frac{n}{2} - \frac{2k_0+3}{2k_0+6} - \frac{(k_0+1)^2}{8}.$$

Proof. Before we used the assumption $k_1' > k_0+1$ in the proof of the second statement of Lemma 5.10, we had

$$\sum_{L \cup S_L'} d(v) \geq \frac{2k_1' + k_0 k_1'}{k_1'+2}|L \cup S_L'| - \frac{k_1' - k_0}{k_1'+2} - |A \cap S_L'|(k_0 + 1 - |A|).$$

In this case, $d(L \cup S_L') < d(XL \cup S_{XL})$ when XL is nonempty (see (5.7)). It follows that

$$\sum_{V-(M\cup B')} d(v) \geq \left(k_0 + \frac{2(k_1'-k_0)}{k_1'+2}\right)|V - (M \cup B')| - \frac{k_1'-k_0}{k_1'+2} - \frac{(k_0+1)^2}{4}.$$

Substituting $k_1' = k_0 + 1$, and noting $d(v) = k_0$ for all $v \in B'$, we have

$$\sum_{v \in V} d(v) \geq \left(k_0 + \frac{2}{k_0 + 3}\right)|G| - \frac{2|B'| + 1}{k_0 + 3} - \frac{(k_0 + 1)^2}{4},$$

and $2|B'| + 1 \leq 2k_0 + 3$, which completes the proof. □

Chapter 6

Double stars and caterpillars

In this section, we prove upper bounds on saturation numbers of certain trees. The first saturation numbers of trees were determined by Kászonyi and Tuza in [60]. In particular, they determined the saturation numbers of paths (see Theorem 6.5) and of stars (recall that these agree asymptotically with the lower bound in Proposition 5.1, $\mathrm{sat}(n, K_{1,t}) = (t-1)n/2 + O(1)$).

In 2009, Faudree, Faudree, Gould, and Jacobson began a more systematic study of saturation numbers of trees [47]. Among other results, they determined that $K^+_{1,p-2}$ has the smallest saturation number out of all trees of order p, found the exact saturation numbers of all trees of order at most 7, and the asymptotic saturation numbers of trees in a number of different classes. For other classes of trees, they were able to obtain bounds but were not able to fully resolve the asymptotics.

(a) The double star $S_{4,5}$

(b) An $S_{4,5}$-saturated graph

Figure 6.1: The double star $S_{4,5}$ on the left and an $S_{4,5}$-saturated graph on the right of order $n = 18$ and size $(12n - 6)/7 = 30$.

6.1 Double stars

One class examined in [47] is that of *double stars*. The *diameter* of a graph is the maximum distance between two of its vertices, and a tree of diameter 3 is called a *double star*. Let $S_{s,t}$ denote the double star whose central vertices have degrees s and t, respectively. We assume $s \leq t$. If $s = t$, we call the double star *balanced*, and if $s < t$, it is *unbalanced*. For example, the balanced double star $S_{2,2}$ is the path P_4, and the unbalanced double star $S_{2,t}$ is the subdivided star $K^+_{1,t-2}$. The unbalanced double star $S_{4,5}$ is depicted in Figure 6.1a.

In addition to determining $\operatorname{sat}(n, S_{2,t})$ exactly, the authors of [47] determined the saturation numbers of balanced double stars asymptotically and bounded the saturation numbers of unbalanced double stars.

Theorem 6.1 ([47])**.** *For $n \geq s^3$,*

$$\frac{s-1}{2}n \leq \operatorname{sat}(n, S_{s,s}) \leq \frac{s-1}{2}n + \frac{s^2-1}{2}, \quad \textit{and}$$
$$\frac{s-1}{2}n \leq \operatorname{sat}(n, S_{s,t}) \leq \frac{s}{2}n - \frac{(s-1)^2+8}{8}.$$

For an unbalanced double star $S_{s,t}$, we have $k_0 = s - 1$ and $k'_1 = t \geq k_0 + 2$.

Further, any edge uv minimizing wt_0 in $S_{s,t}$ has only one endpoint of degree s, and this endpoint has a neighbor of degree t. Theorem 5.8 thus provides an improved lower bound for unbalanced double stars $S_{s,t}$.

Corollary 6.2 ([20])**.** *For any positive integers s and t, $2 \le s < t$, and for any $n \ge s+t$,*

$$\mathrm{ssat}(n, S_{s,t}) \ge \frac{s(t+1)n - s(t-s+2)}{2t+4} - \frac{s^2}{8}.$$

We now determine that the saturation number of $S_{s,t}$ is $\frac{s(t+1)}{2t+4}n + O(1)$. Our upper bound is based upon the observation that a graph G_0 obtained from two copies of $K_{1,t+1}$ by joining their sets of leaves with an $(s-2)$-regular bipartite graph, as in the larger component of Figure 6.1b, is $S_{s,t}$-saturated and has average degree exactly $s(t+1)/(t+2)$. Further, any graph consisting of disjoint copies of G_0 is $S_{s,t}$-saturated. We are able to add a disjoint clique of cardinality s to such a graph to obtain another $S_{s,t}$-saturated graph G whose average degree is a little bit smaller than $s(t+1)/(t+2)$. More precisely, such a graph G has

$$\frac{s(t+1)n - s(t-s+2)}{2t+4} \tag{6.1}$$

edges. In fact, we will prove a slight improvement upon Corollary 6.2 in Theorem 6.4, determining that (6.1) is precisely the (semi)saturation number of $S_{s,t}$ when n is large and equivalent to s modulo $2t+4$. When $n \not\equiv s \pmod{2t+4}$, we add vertices to non-clique components in a manner described below.

Theorem 6.3 ([20])**.** *Let s and t be positive integers, $2 \le s < t$, and let $q =$*

max $\{1, \lfloor s/2 \rfloor - 1\}$. *For any* $n \geq q(2t+4)+s$,

$$\operatorname{sat}(n, S_{s,t}) \leq \frac{s(t+1)n + s(s-1)}{2t+4} + \left\lceil \frac{s}{2} \right\rceil.$$

Proof of Theorem 6.3. Let $S_{s,t}$ be a double star with $s < t$. For $n \geq q(2t+4)+s$ where $q = \max\{1, \lfloor (s-2)/2 \rfloor\}$, we construct an n-vertex graph G with the following properties.

(i) We have $V(G) = S \cup L$. For all $v \in S$, $d(v) = s-1$. For all $v \in L$, $d(v) \geq t+1$.

(ii) For all $v \in L$, $N(v) \subseteq S$, and every $w \in N(v)$ is contained in an independent set of cardinality $t+1$ in $N(v)$.

(iii) Aside from a clique B of order s, every vertex in S has a neighbor in L, and at most one vertex in S has two or more neighbors in L.

We claim that G is $S_{s,t}$-saturated. Since there are no vertices of degree at least t adjacent to any vertices of degree at least s, G is $S_{s,t}$-free. Let x and y be nonadjacent vertices in G. If $x, y \in L$, then both have degree at least $t+1$ by (i), and they have at most one common neighbor by (iii), so x and y are the internal vertices of a copy of $S_{s,t}$ in $G + xy$. If $x \in S - B$, let $z \in N(x) \cap L$. By (ii), there is an independent set I_z of cardinality $t+1$ in $N(z)$ which contains x. There are $t-1$ vertices in $I_z - \{x, y\}$ and $s-1$ neighbors of x which are not in I_z. Therefore, x and z are the internal vertices of a copy of $S_{s,t}$ in $G + xy$. If $x \in B$, we may assume $y \in L$, in which case $B - x$ serves as a set of $s-1$ leaves, and y has a set of $t-1$ neighbors disjoint from B, resulting in a copy of $S_{s,t}$.

We construct G as follows. Let L and S partition the vertex set of G with $|L| = 2\lfloor (n-s)/(2t+4) \rfloor$. Let r be the remainder of $(n-s)/(2t+4)$, and let R be a set of r vertices in S. Let B be a clique of order s in S. Let every vertex in L be adjacent to $t+1$ distinct vertices in $S-(B\cup R)$ so that $V(G)-(B\cup R)$ induces a set of at least $2q$ copies of $K_{1,t+1}$. This partitions $S-(B\cup R)$ into classes.

If r is even, make two of these stars into copies of $K_{1,t+1+r/2}$, and put an $(s-2)$-regular bipartite graph on the two sets of $t+1+r/2$ vertices in S. Since $|L|$ is even, we can pair up the remaining classes in $S-(B\cup R)$, and put an $(s-2)$-regular bipartite graph on each pair.

If r is odd, let $v\in R$, and repeat the steps in the previous paragraph for $R-v$. If s is even, give v a single neighbor in L, and if s is odd, give v two neighbors in L. If $s>3$, then take an adjacent pair in $S-B$, delete the edge between them, and give each an edge to v. Repeat this, choosing a different pair of classes at each step for the adjacent pair to ensure condition (ii), until v has degree $s-1$. By our assumption on n, this is always possible, as there are at least $\lfloor s/2\rfloor-1$ pairs of classes to choose from.

The resulting graph G meets conditions (i)–(iii). Further, for even r,

$$\begin{aligned}\|G\| &= \left(\frac{s(t+1)}{t+2}\right)\frac{n-r}{2}-\frac{s(t-s+2)}{2t+4}+\frac{sr}{2t+4}\\ &\le \left(\frac{s(t+1)}{t+2}\right)\frac{n}{2}+\frac{s(s+t)}{2t+4},\end{aligned}$$

and for odd r,

$$\begin{aligned}\|G\| &= \left(\frac{s(t+1)}{t+2}\right)\frac{n-1}{2}-\frac{s(t-s+2)}{2t+4}+\frac{s(r-1)}{2t+4}+\left\lceil\frac{s}{2}\right\rceil\\ &\le \left(\frac{s(t+1)}{t+2}\right)\frac{n}{2}+\frac{s(s-1)}{2t+4}+\left\lceil\frac{s}{2}\right\rceil.\end{aligned}$$

This completes the proof. $\square$

We now prove that this upper bound construction is best possible for certain values of n which are sufficiently large and meet divisibility conditions. Recall from equation (6.1) that there are $S_{s,t}$-saturated graphs with precisely $\frac{s(t+1)n-s(t-s+2)}{2t+4}$ edges when $n \equiv s \pmod{2t+4}$.

Theorem 6.4 ([20]). *Let s and t be positive integers, $2 \leq s < t$. There exists $n_0 = n_0(s,t)$ such that, for all $n \geq n_0$,*

$$\mathrm{ssat}(n, S_{s,t}) \geq \frac{s(t+1)n - s(t-s+2)}{2t+4},$$

and this is sharp when $n \equiv s \pmod{2t+4}$.

Proof. Suppose that G is an $S_{s,t}$-saturated graph of order n and that the clique A of vertices in G with degree at most $s-2$ is nonempty. Let $v \in A$. If w is a nonneighbor of v, then w must be the image of either the degree-s or degree-t vertex in the copy of $S_{s,t}$ in $G + vw$, and v must be the image of a *leaf*, a vertex of degree 1. Thus, w has a neighbor of degree at least s.

Let S, L, and XL be as in the proofs of Lemmas 5.9 and 5.10; that is, $S = \{v : d(v) < s\}$, $L = \{v : d(v) = t\}$, and $XL = \{v : d(v) > t\}$. Further, let S'_L, S_{XL}, and B' partition S in the same manner as Lemma 5.10. The vertex v in A has at most $s - 1 - |A|$ neighbors in $L \cup XL$. Let N denote this set of high-degree neighbors of v. We have $e(L, S_L) \leq (t-1)|L| + |N \cap L|$ and $e(XL, S_{XL}) \leq (x-1)|XL| + |N \cap XL|$ where $x = d(XL)$. By similar reasoning to the proof of Lemma 5.10, we have

$$\sum_{v \in V(G)} d(v) \geq \left(s - 1 + \frac{t-s+2}{t+1}\right)n - \frac{|B \cup N|(t-s+2)}{t+1} - \frac{s^2}{4},$$

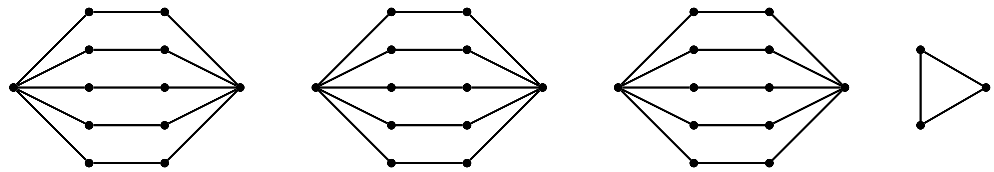


Figure 6.2: A graph of minimum size over all $S_{3,4}$-(semi)saturated graphs of order 39

and the right side of this inequality is strictly larger than

$$\frac{s(t+1)n - s(t-s+2)}{t+2},$$

when n is sufficiently large. Thus, in a minimum $S_{s,t}$-saturated graph G of large order, the set A is empty, and the first statement follows from the proof of Lemma 5.9. Tightness when $n \equiv s \pmod{2t+4}$ follows from the upper bound construction in Theorem 6.3. □

It is not hard to bound the function $n_0(s,t)$ using the upper bounds $|N| \leq s-2$ and $|B'| \leq s$ in the proof of Theorem 6.4, but it is ugly:

$$n_0 \geq \frac{s^2(t^2-t-10) + 4s(t^2+7t+10) - 8(t+2)^2}{4(t-s+2)}.$$

To give some idea of what this looks like, for $s = 3$ and $t = 4$, any $n \geq 32$ will do; and for $s = 4$ and $t = 5$, any $n \geq 74$ will do. In particular, the graph depicted in Figure 6.2 is an $S_{3,4}$-(semi)saturated graph of minimum size, and any graph consisting of at least five copies of the larger component in Figure 6.1b and one copy of K_4 is an $S_{4,5}$-(semi)saturated graph of minimum size. The author believes that the bound in Theorem 6.4 is likely sharp whenever $n = (2t+4)q + s$ for some positive integer q, but we have not attempted to prove this.

6.2 Caterpillars

A *caterpillar* is a tree obtained from a path (the *body*) by adding pendent edges (*feet*) to its internal vertices. If the same number s of edges are appended to each internal vertex of the body, we call it an s-*caterpillar*, and if the body is a path on ℓ vertices, we call this s-caterpillar P_ℓ^s. In [20], we nicknamed P_5^s "Shorty the caterpillar."

Note that the graphs P_4^s are balanced double stars. The saturation number of P_4^s aligns asymptotically with the lower bound in Proposition 5.1 [47], as we noted in Theorem 6.1 of the previous section. Faudree, Faudree, Gould, and Jacobson also provided upper and lower bounds on $\text{sat}(n, P_\ell^s)$ for any $\ell \geq 4$ [47]. Corollary 5.11 improves upon their lower bound; in fact, though obtained in a different manner, their lower bound aligns exactly with the one Proposition 5.1. We also demonstrated an improved upper bound for $\ell = 5$ in [20]. In particular, there is a constant c depending only on s such that, for any $n \geq |P_\ell^{s-1}|$,

$$\text{sat}(n, P_\ell^{s-1}) \geq \left(s + \frac{2}{s+3}\right)\frac{n}{2} - c \tag{6.2}$$

and a constant d depending only on s such that, for $n \geq q(2s+4) + s + 1$, where $q = \max\{2, \lfloor (s-1)/2 \rfloor\}$,

$$\text{sat}(n, P_5^{s-1}) \leq \left(s + \frac{2}{s+2}\right)\frac{n}{2} + d. \tag{6.3}$$

Rather than reproduce the proof of (6.3) from [20], we shall prove a generalization for any $\ell \geq 5$.

We note that $P_\ell = P_\ell^0$, and recall that Kászonyi and Tuza characterized the

(a) T_5^3 *(b)* T_6^3

Figure 6.3: Two almost ternary trees

minimum path-saturated graphs in [60]. In particular, they determined that a disjoint union of "almost binary" trees of depth $\lfloor \ell/2 \rfloor$ is a P_ℓ-saturated graph of minimum size.

Definition 6.1 (T_m^k)**.** For positive integers k, m with $m \geq 3$, let T_m^k denote the *almost k-ary tree* with diameter $m-1$, defined as follows: the vertex set of T_m^k is partitioned into $\lceil m/2 \rceil$ levels; all leaves of T_m^k are in the bottom level, and all other vertices have degree $k+1$; the top level contains either one vertex or a pair of adjacent vertices, depending on whether m is odd or even, respectively; the vertices in the top level have the rest of their $k+1$ neighbors in the level below, and so do the vertices in subsequent levels (except the bottom one, of course).

We acknowledge that notation may be getting confusing, with T_k, $T^p(n)$, and T_m^k all referring to different graphs. We encourage the reader to not distress, for we shall not reference T_k or $T^p(n)$ in this section. Indeed, no section contains any two of these three special classes of graphs.

Two almost ternary trees, T_5^3 and T_6^3, are depicted in Figure 6.3. We note that the order of T_{2m+1}^k is

$$1 + (k+1) + k(k+1) + \cdots + k^{m-1}(k+1) = \frac{(k+1)k^m - 2}{k-1}$$

and the order of T_{2m}^k is

$$2(1+k+\cdots+k^{m-1}) = \frac{2(k^m-1)}{k-1}.$$

We now state Kászonyi and Tuza's result more precisely.

Theorem 6.5 ([60]). *Let ℓ be an integer, $\ell \geq 3$, and let $a_\ell = |T_{\ell-1}^2|$. For any $n \geq a_\ell$, every P_ℓ-saturated graph of order n with minimum size is a forest with $\lfloor n/a_\ell \rfloor$ components. Hence $\mathrm{sat}(n, P_\ell) = n - \lfloor n/a_\ell \rfloor$. Further, every P_ℓ-saturated tree contains $T_{\ell-1}^2$ as a subgraph.*

In what follows, we generalize the upper bound, using the trees $T_{\ell-1}^{s+2}$ to construct P_ℓ^s-saturated graphs.

Before doing so, we note an important difference between semisaturation numbers and saturation numbers of paths. It was observed by Burr in [21] that a disjoint union of paths P_r, $r = r_\ell = \lfloor 3(\ell-1)/2 \rfloor$, is P_ℓ-semisaturated. Noting that

$$a_\ell = |T_{\ell-1}^2| = \begin{cases} 3 \cdot 2^{m-1} - 2: & \ell = 2m; \\ 4 \cdot 2^{m-1} - 2: & \ell = 2m+1, \end{cases}$$

a bit of basic arithmetic shows that $r_\ell < a_\ell$ for $\ell \geq 6$. Thus, in this case, P_ℓ-semisaturated graphs have asymptotically fewer edges than P_ℓ-saturated graphs.

Theorem 6.6 ([21]). *Let ℓ be an integer, $\ell \geq 2$, and let $r = \lfloor 3(\ell-1)/2 \rfloor$. For any $n \geq 2r$,*

$$n - \left\lfloor \frac{n-1}{r} \right\rfloor - 1 \leq \mathrm{ssat}(n, P_\ell) \leq n - \left\lfloor \frac{n}{r} \right\rfloor.$$

We now return to proving an upper bound on $\mathrm{sat}(n, P_\ell^s)$ for $\ell \geq 5$, $s \geq 0$. We begin with a certain connected P_ℓ^s-saturated graph, defined as follows.

Definition 6.2 (G_ℓ^s)**.** For positive integers ℓ, s with $\ell \geq 5$, let G_ℓ^s denote a graph whose vertex set can be partitioned into sets A and B such that each of the induced subgraphs $G_\ell^s[A]$ and $G_\ell^s[B]$ is isomorphic to $T_{\ell-1}^{s+2}$. The remaining edges in G_ℓ^s form your favorite s-regular bipartite graph between the sets of leaves in $G_\ell^s[A]$ and $G_\ell^s[B]$.

For example, the graph G_6^1 is depicted in Figure 6.4b. As we will presently show, the graph G_ℓ^s is P_ℓ^s-saturated, and so is any graph consisting of disjoint copies of G_ℓ^s. For a value of n not divisible by the order of G_ℓ^s, we can add extra leaves to the bottom level of one pair of $T_{\ell-1}^{s+2}$'s (plus one vertex with two high-degree neighbors if n and s are both odd) to obtain a P_ℓ^s-saturated graph of order n.

Lemma 6.7. *For any $\ell \geq 5$ and $s \geq 0$, the graph G_ℓ^s is P_ℓ^s-saturated.*

Proof. Let $G = G_\ell^s$. Every vertex in G is either of degree $s+1$ (*low-degree*) or $s+3$ (*high-degree*). There are no $\ell-2$ consecutive high-degree vertices in G, so it is P_ℓ^s-free. Let x and y be nonadjacent vertices in G. It remains to show that $P_\ell^s \subseteq G + xy$.

Recall the partition A, B of $V(G)$ from Definition 6.2, where $G[A] \cong G[B] \cong T_{\ell-1}^{s+2}$.

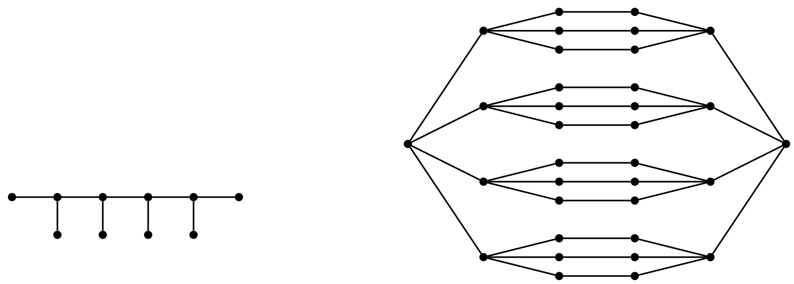
Case 1 ($x, y \in A$). We first suppose that x and y are on the same side of the partition A, B of $V(G)$, say A. Let $L_1, \ldots, L_{\lfloor \ell/2 \rfloor}$ denote the levels of $G[A]$, as described in Definition 6.1, where $|L_1| > \cdots > |L_{\lfloor \ell/2 \rfloor}|$.

Suppose $x \in L_i$ and $y \in L_j$ with $i \leq j$. Let P denote the unique x, y-path in $G[A]$; write $P = xx' \cdots y'y$. Note that x and y are in distinct components of the graph $G' = G[A] - yy'$. We will find an s-caterpillar in each component of G', one having x and the other having y as an endpoint, which we join with the edge xy to create a copy of P_ℓ^s in G.

First, suppose that x is not a descendant of y; that is, the path P is not monotone with respect to the levels $L_1, L_2, \dots$ in $G[A]$. In particular, the level containing y' is L_{j+1}, or is L_j if $j = \lfloor \ell/2 \rfloor$ and ℓ is odd. In this case, y is the endpoint of a P_j^s whose body is a path from y to a leaf vertex in G', and x is the endpoint of a $P_{\ell-i}^s$ in G' whose body is a path from x to L_1 (visiting both vertices if ℓ is odd and $y \notin L_1$) and back to a leaf. If at least one of x or y is high-degree, each of them has at least s neighbors in G which are not in the other's component of G' (if one of x or y is low-degree, then its neighbors are not in A). Thus, in this case, x and y are internal vertices on a copy of $P_{j+\ell-i}^s$ in $G + xy$, and so on a copy of P_ℓ^s. In the case that both x and y are low-degree vertices, we have $x = j = 1$, and y is a terminal vertex on the resulting copy of P_ℓ^s.

On the other hand, suppose that x is a descendant of y in $G[A]$. In G', x is the endpoint of a copy of P_{2j-i-2}^s whose body follows P from x to y' and then from y' to a leaf of G'. Since x is a nonadjacent descendant of y, we have $j - i \geq 2$ and $2j - i - 2 \geq j$. Also, y is the endpoint of a copy of $P_{\ell-j}$ whose body is a path from y to L_1 and back to a leaf of G'. We again join these with the edge xy, finding s neighbors for x from the bipartite graph between A and B if necessary, to make x and y internal vertices on a copy of $P_{\ell'}^s$, where $\ell' = \ell + j - i - 2 \geq \ell$. This completes the proof of Case 1.

Case 2 ($x \in A, y \in B$). We now suppose that x and y are on different sides of the partition, say $x \in A$ and $y \in B$. Again, suppose that x is in level i of $G[A]$ and y is in level j of $G[B]$, where $i \leq j$. Note that x is an endpoint of a copy of $P_{\ell-i}$ in $G[A]$ and y an endpoint of a $P_{\ell-j}$ in $G[B]$. If $j \geq 2$ (*i.e.*, if y is a high-degree vertex), then the sets $N[x]$ and $N[y]$ are certainly disjoint, and x and y are internal vertices on a



(a) P_6^1 (b) G_6^1

Figure 6.4: The caterpillar P_6^1 on the left and a P_6^1-saturated graph on the right

copy of $P^s_{2\ell-i-j}$ in $G+xy$. Since $i \le j \le \lfloor \ell/2 \rfloor$, we have $2\ell - i - j \ge \ell$. On the other hand, if $i = j = 1$, then since $xy \notin G$, we can make y the terminal vertex, and x the first internal vertex, on a copy of P^s_ℓ in $G + xy$.

This completes the proof. □

Theorem 6.8. *For any positive integers s and $t \ge 2$,*

$$\text{sat}(n, P^{s-1}_{2t+1}) \le \left(s + 2\frac{(s+1)^{t-1} - 1}{(s+1)^t - 1} \right) \frac{n}{2} + c$$

and

$$\text{sat}(n, P^{s-1}_{2t+2}) \le \left(s + 2\frac{(s+2)(s+1)^{t-1} - 2}{(s+2)(s+1)^t - 2} \right) \frac{n}{2} + c.$$

Proof. Note that the average degree of G^{s-1}_ℓ agrees with the upper bound ($c = 0$) when $\ell \in \{2t+1, 2t+2\}$.

Write $n = q|G^{s-1}_\ell| + r$ where $r < |G^{s-1}_\ell|$. Start by taking $q-1$ copies of G^{s-1}_ℓ. For the remaining $|G^{s-1}_\ell| + r$ vertices, we take two copies of $T^{s+1}_{\ell-1}$ as in the definition of G^{s-1}_ℓ. In each copy of $T^{s+1}_{\ell-1}$, choose a degree-$(s+2)$ vertex in the level just above the leaves. To one of these we append $\lfloor r/2 \rfloor$ pendent edges, and to the other we append $\lceil r/2 \rceil$ pendent edges. If $s-1$ and r are both odd, then we add an extra edge from the leaf endpoint x of one of these pendent edges to a high-degree vertex in the other copy of $T_{\ell-1}^{s+1}$. Then, we give every low-degree vertex degree s by filling in an $(s-1)$-regular bipartite graph between the leaves, if one of $s-1$ or r is even, or by giving x $s-2$ edges to any of the leaves, and then filling in the rest of the edges with a bipartite graph if both $s-1$ and r are odd. Call the resulting graph G.

Having no $\ell-2$ consecutive vertices of degree $s+1$, G is P_ℓ^{s-1}-free. Each component of G isomorphic to G_ℓ^{s-1} is P_ℓ^{s-1}-saturated by Lemma 6.7. The component with two vertices of degree greater than $s+2$ is P_ℓ^{s-1}-saturated by the same arguments, with the exception of the case in which n is odd and one low-degree vertex has these two high-degree neighbors. For any possible nonneighbor of this vertex, the cases follow similarly to the low-degree vertex cases above.

A similar argument to Case 2 in the proof of Lemma 6.7 shows that adding an edge between components of G creates a copy of P_ℓ^{s-1}. This completes the proof. □

6.3 Rainbow saturation numbers of double stars

Here we consider an edge-colored version of the saturation problem. In analogy with proper vertex colorings and the chromatic number, a *proper edge-coloring* of a graph G is an assignment of colors to its edges so that no two incident edges receive the same color. This can be thought of as a partition of G with matchings (indeed, the linear arboricity conjecture discussed in Chapter 1 can be seen as a generalization of a classical theorem of Vizing [92]). Clearly, a graph with maximum degree Δ will need at least Δ colors in a proper edge coloring, as every edge incident to a given vertex needs a different color. Vizing showed that this natural lower bound is off by at most 1 from the minimum number of colors needed in general.

Theorem 6.9 ([92])**.** *Every graph with maximum degree Δ can be properly edge colored with at most $\Delta + 1$ colors.*

An edge-colored graph is called *rainbow* if all of its edges receive different colors. In 2007, Keevash, Mubayi, Sudakov, and Verstraëte introduced the *rainbow Turán number* of a graph H, denoted $\mathrm{ex}^\star(n, H)$ [61]. The rainbow Turán number of H is the maximum number of edges in a graph of order n which can be properly edge-colored in a manner that avoids a rainbow copy of H. We call such an edge coloring *rainbow H-free*. As well as being a natural meeting point of two well-studied types of problems, Turán-type problems and edge-coloring problems, the study of rainbow Turán numbers was first motivated by an application to additive number theory [61]. In the same paper, these four authors proved that $\mathrm{ex}^\star(n, H) = \mathrm{ex}(n, H) + o(n^2)$, showing that equality holds for color-critical graphs H, and they made progress on the bipartite case. Notably, they proved that $\mathrm{ex}^\star(n, C_{2k}) \geq cn^{1+1/k}$ for an absolute constant c.

In 2022, Bushaw, Johnston, and Rombach introduced an analogous notion in the realm of graph saturation [22]. A graph is said to be *rainbow H-saturated* if it is edge-maximal with respect to the property of possessing a rainbow H-free edge coloring. As in the case of classical saturation, this definition allows one to ask not only for the maximum number of edges in a rainbow H-saturated graph of order n ($\mathrm{ex}^\star(n, H)$), but also for the *minimum* number of edges in an H-saturated graph of order n. This minimum is called the *rainbow saturation number* of H, denoted $\mathrm{sat}^\star(n, H)$. We note that $\mathrm{sat}^\star(n, H)$ is sometimes referred to as the proper rainbow saturation number to avoid confusion with a homonymous parameter.

Bushaw, Johnston, and Rombach proved that $\mathrm{sat}^\star(n, H) = O(n)$ for graphs H without an induced even cycle. It has since been observed by various sources that Kászonyi and Tuza's proof that $\mathrm{sat}(n, H)$ is linear [60] extends to prove this statement for arbitrary graphs H (see, for example, [65]). Thus, just like the classical saturation number and semisaturation number, the problem of determining $\mathrm{sat}^\star(n, H)$ comes down to determining a constant c such that $\mathrm{sat}^\star(n, H) = cn + o(n)$.

Proposition 6.10. *Every rainbow H-saturated graph is H-semisaturated.*

Proof. Suppose G possesses a rainbow H-free edge-coloring c, and let x and y be nonadjacent vertices in G. If the edge xy is not contained in any copy of H in $G+xy$, then extending c to an edge-coloring of $G + xy$ in any admissible manner, we obtain a rainbow H-free coloring in $G + xy$. □

While the lower bound $\mathrm{ssat}(n, H) \leq \mathrm{sat}^\star(n, H)$ is trivial, it is not known whether $\mathrm{sat}(n, H) \leq \mathrm{sat}^\star(n, H)$ in general. However, in all known nontrivial cases (*i.e.*, other than stars and triangles where every proper edge coloring is rainbow), these two parameters differ asymptotically. For instance, consider the graph P_4. A minimum P_4-saturated graph has about $n/2$ edges (see Figure 4.2), but it is clear that a matching is not rainbow P_4-saturated, for P_4 has a proper edge-coloring with 2 colors. For the same reason, any graph with two isolated edges is not P_4-saturated. On the other hand, a disjoint union of copies of $K_{1,4}$ is rainbow P_4-saturated and has average degree 4/5. This turns out to be asymptotically optimal.

Theorem 6.11 ([22])**.** *For each $n \geq 16$, $\lfloor 4n/5 \rfloor \leq \mathrm{sat}^\star(n, P_4) \leq 4n/5 + O(1)$.*

Recently, the rainbow saturation number has received considerable attention. The authors of [55] determined $\mathrm{sat}^\star(n, C_4) = 11n/6 + o(n)$ and provided bounds for C_5 and C_6. The value of $\mathrm{sat}^\star(n, P_\ell)$ was independently determined to be $n + O(1)$ for $\ell \geq 5$ in [7] and [66]. The former authors also determined $\mathrm{sat}^\star(n, K_4) = 7n/2 + O(1)$, and the latter authors studied many other classes of trees. Included in these were the double stars $S_{2,t}$. They determined that a disjoint union of copies of $K_{1,t+2}$ asymptotically minimizes $\mathrm{sat}^\star(n, S_{2,t})$; more precisely, $\mathrm{sat}^\star(n, S_{2,t}) = n - \lfloor (n+t+1)/(t+3) \rfloor$ [66]. In the same paper, the authors provided upper bounds on saturation and semisaturation for double stars which are not so far off from the ones described in Section 6.1. In rainbow case, they proved that $\mathrm{sat}^\star(n, S_{s,t}) \leq \frac{(\lceil t/(s-1) \rceil + 2)(s-1)}{(\lceil t/(s-1) \rceil + 2)(s-1) + 1} \cdot \frac{sn}{2} + O(1)$.

In this section, we prove that $\mathrm{sat}^\star(n, S_{s,t}) \leq \frac{s+t}{s+t+1} \cdot \frac{sn}{2} + O(1)$. First, we note that our lower bound on $\mathrm{ssat}(n, S_{s,t})$ also holds in the rainbow case by Proposition 6.10, improving upon previous bounds.

Corollary 6.12. *For any $s < t$ and $n \geq s+t$, we have $\mathrm{sat}^\star(n, S_{s,t}) \geq s\left(1 - \frac{1}{t+2}\right)\frac{n}{2} - c$, where $c = \frac{s(t-s+2)}{2t+4} + \frac{s^2}{8}$.*

We now prove an upper bound, reminiscent of our upper bound for $\mathrm{sat}(n, S_{s,t})$ in Theorem 6.3. This result is based on joint work with Bushaw, Johnston, and Rombach.

Theorem 6.13. *For any $s \leq t$ and $n \geq 2(s+t+1)$, we have*

$$\mathrm{sat}^\star(n, S_{s,t}) \leq s\left(1 - \frac{1}{s+t+1}\right)\frac{n}{2} + O(1).$$

Proof. For n divisible by $2(s+t+1)$, we construct a rainbow $S_{s,t}$-saturated graph G whose vertices have degree either $s-1$ or $s+t$ and with $s+t$ vertices of degree $s-1$

for each vertex of degree $s+t$. Such a graph has $\frac{1}{2} \cdot \frac{n}{s+t+1} \cdot s(s+t)$ edges, matching the claimed upper bound.

Our graph G is constructed by pairing up an even number of copies of $K_{1,s+t}$ and, for each pair, adding an $(s-2)$-regular bipartite graph between the partite sets of size $s+t$. The graph G is $S_{s,t}$-free, and thus rainbow $S_{s,t}$-free as well. To see that G is rainbow $S_{s,t}$-saturated, let x, y be nonadjacent vertices. If $d(x) = d(y) = s+t$, then this is easy to see. Otherwise, without loss of generality, $d(x) = s-1$. Let z be the neighbor of x in G with degree $s+t$. All of the edges incident to x in $G+xy$ receive different colors in a proper edge coloring by definition, which leaves at least $s+t-(s-1)-2 = t-1$ other colors incident to z which do not go to $\{x, y\} \cup N_G(x)$, and thus we find a rainbow copy of $S_{s,t}$ in every proper edge coloring of $G+xy$.

For even values of n not divisible by $2(s+t+1)$, we take a graph G as previously described, but with one connected component obtained from two copies of $K_{1,s+t+r/2}$ and an $(s-2)$-regular bipartite graph, where r is the remainder of $n/2(s+t+1)$. For odd n, add a new vertex to the $(n-1)$-vertex graph constructed as above. Join that new vertex to a single high-degree vertex if s is even, or to two high-degree vertices if s is odd. If $s \geq 4$, then delete an edge joining a pair of low-degree vertices in $\lfloor s/2 \rfloor - 1$ different components, and add edges from the new vertex to each previously adjacent pair of low-degree vertices so that the new vertex has degree $s-1$. A similar argument to the one above shows that the resulting graph is rainbow $S_{s,t}$-saturated. □

INDEX

Bibliography

[1] J. Akiyama, G. Exoo, and F. Harary. Covering and packing in graphs. III: Cyclic and acyclic invariants. *Mathematica Slovaca*, 30(4):405–417, 1980.

[2] V. E. Alekseev and V. V. Lozin. On orthogonal representations of graphs. *Discrete Mathematics*, 226(1-3):359–363, 2001.

[3] N. Alon. The linear arboricity of graphs. *Israel Journal of Mathematics*, 62(3):311–325, 1988.

[4] N. Alon. Neighborly families of boxes and bipartite coverings. *The Mathematics of Paul Erdős II*, 27–31, 1997.

[5] N. Alon and C. Shikhelman. Many T copies in H-free graphs. *Journal of Combinatorial Theory, Series B*, 121:146–172, 2016.

[6] L. Babai and P. Frankl. *Linear Algebra Methods in Combinatorics: Part I*. Department of Computer Science, The University of Chicago, 1988. (Preliminary version).

[7] D. Baker, E. Gomez-Leos, A. Halfpap, E. Heath, R. R. Martin, J. Miller, A. Parker, H. Pungello, C. Schwieder, and N. Veldt. On the proper rainbow saturation numbers of cliques, paths, and odd cycles. *arXiv preprint arXiv:2409.15258*, 2024.

[8] W. Barrett, S. Butler, S. M. Fallat, H. T. Hall, L. Hogben, J. C.-H. Lin, B. L. Shader, and M. Young. The inverse eigenvalue problem of a graph: Multiplicities and minors. *Journal of Combinatorial Theory, Series B*, 142:276–306, 2020.

[9] W. Barrett, J. Grout, and R. Loewy. The minimum rank problem over the finite field of order 2: minimum rank 3. *Linear Algebra and its Applications*, 430(4):890–923, 2009.

[10] N. Behague, T. Johnston, N. Morrison, and S. Ogden. A note on the invertibility of oriented graphs. *arXiv:2404.10663*, 2024.

[11] A. Blanché, M. Bonamy, and N. Bonichon. Gallai's path decomposition for planar graphs. In *Extended Abstracts EuroComb 2021*, pages 758–764. Springer, 2021.

[12] B. Bollobás. On a conjecture of Erdős, Hajnal and Moon. *American Mathematical Monthly*, pages 178–179, 1967.

[13] B. Bollobás. *Extremal graph theory.* Academic Press, London, 1978.

[14] M. Bonamy and T. J. Perrett. Gallai's path decomposition conjecture for graphs of small maximum degree. *Discrete Mathematics*, 342(5):1293–1299, 2019.

[15] S. Borgwardt, C. Buchanan, E. Culver, B. Frederickson, P. Rombach, and Y. Yoo. Path odd-covers of graphs. *arXiv:2306.06487*, 2023.

[16] A. Bouchet. Circle graph obstructions. *Journal of Combinatorial Theory, Series B*, 60(1):107–144, 1994.

[17] C. Buchanan, A. Clifton, E. Culver, P. Frankl, J. Nie, K. Ozeki, P. Rombach, and M. Yin. On odd covers of cliques and disjoint unions. *arXiv:2408.08598*, 2024.

[18] C. Buchanan, A. Clifton, E. Culver, J. Nie, J. O'Neill, P. Rombach, and M. Yin. Odd covers of graphs. *Journal of Graph Theory*, 104(2):420–439, 2023.

[19] C. Buchanan, C. Purcell, and P. Rombach. Subgraph complementation and minimum rank. *The Electronic Journal of Combinatorics*, 29(1):P1.38, 2022.

[20] C. Buchanan and P. Rombach. A lower bound on the saturation number and a strengthening for triangle-free graphs. *arXiv preprint arXiv:2402.11387*, 2024.

[21] E. Burr. A study of saturation number. Master's thesis, University of Alaska Fairbanks, 2017.

[22] N. Bushaw, D. Johnston, and P. Rombach. Rainbow saturation. *Graphs and Combinatorics*, 38(5):166, 2022.

[23] L. Caccetta, P. Erdos, E. T. Ordman, and N. J. Pullman. The difference between the clique numbers of a graph. *Ars Combinatoria A*, 19:97–106, 1985.

[24] A. Cameron and G. J. Puleo. A lower bound on the saturation number, and graphs for which it is sharp. *Discrete Mathematics*, 345(7):112867, 2022.

[25] M. S. Cavers. Clique partitions and coverings of graphs. Master's thesis, Waterloo, Ontario, Canada, December 2005.

[26] Y.-C. Chen. Minimum C_5-saturated graphs. *Journal of Graph Theory*, 61(2):111–126, 2009.

[27] Y.-C. Chen. All minimum C_5-saturated graphs. *Journal of Graph Theory*, 67(1):9–26, 2011.

[28] N. Chenette, S. Droms, L. Hogben, R. Mikkelson, and O. Pryporova. Minimum rank of a graph over an arbitrary field. *The Electronic Journal of Linear Algebra*, 16, 2007.

[29] F. Chung. On partitions of graphs into trees. *Discrete Mathematics*, 23(1):23–30, 1978.

[30] F. Chung. On the coverings of graphs. *Discrete Mathematics*, 30(2):89–93, 1980.

[31] F. Chung. On the decomposition of graphs. *SIAM Journal on Algebraic Discrete Methods*, 2(1):1–12, 1981.

[32] B. L. Currie, J. R. Faudree, R. J. Faudree, and J. R. Schmitt. A survey of minimum saturated graphs. *The Electronic Journal of Combinatorics*, page DS19, 2021.

[33] N. de Beaudrap. Decomposition of graphs as symmetric differences of copies of $K_{a,b}$. MathOverflow. URL:https://mathoverflow.net/q/76043 (version: 2011-10-27).

[34] N. G. de Bruijn and P. Erdős. On a combinatorial problem. In *Proceedings of the Section of Sciences of the Koninklijke Nederlandse Akademie van Wetenschappen te Amsterdam*, volume 51, pages 1277–1279, 1948.

[35] G. Ding and A. Kotlov. On minimal rank over finite fields. *The Electronic Journal of Linear Algebra*, 15, 2006.

[36] J. Edmonds. Minimum partition of a matroid into independent subsets. *J. Res. Nat. Bur. Standards Sect. B*, 69:67–72, 1965.

[37] E. Egerváry. On combinatorial properties of matrices. *Mat. Lapok*, 38:16–28, 1931. (Hungarian with German summary).

[38] P. Erdős. On the number of complete subgraphs contained in certain graphs. *Magyar Tud. Akad. Mat. Kutató Int. Közl*, 7(3):459–464, 1962.

[39] P. Erdős. Paul Turán, 1910–1976: his work in graph theory. *Journal of Graph Theory*, 1(2):97–101, 1977.

[40] P. Erdős, R. Faudree, and E. T. Ordman. Clique partitions and clique coverings. *Discrete Mathematics*, 72(1-3):93–101, 1988.

[41] P. Erdős, A. W. Goodman, and L. Pósa. The representation of a graph by set intersections. *Canadian Journal of Mathematics*, 18:106–112, 1966.

[42] P. Erdős, A. Hajnal, and J. W. Moon. A problem in graph theory. *The American Mathematical Monthly*, 71(10):1107–1110, 1964.

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

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

[45] S. M. Fallat and L. Hogben. The minimum rank of symmetric matrices described by a graph: a survey. *Linear Algebra and its Applications*, 426(2-3):558–582, 2007.

[46] G. Fan. Subgraph coverings and edge switchings. *Journal of Combinatorial Theory, Series B*, 84(1):54–83, 2002.

[47] J. Faudree, R. J. Faudree, R. J. Gould, and M. S. Jacobson. Saturation numbers for trees. *The Electronic Journal of Combinatorics*, 16(R91), 2009.

[48] F. V. Fomin, P. A. Golovach, T. J. Strømme, and D. M. Thilikos. Subgraph complementation. *Algorithmica*, 82(7):1859–1880, 2020.

[49] B. Frederickson and L. Michel. Circuit decompositions of binary matroids. *SIAM Journal on Discrete Mathematics*, 38(2):1193–1201, 2024.

[50] S. Friedland. Quadratic forms and the graph isomorphism problem. *Linear Algebra and its Applications*, 150:423–442, 1991.

[51] Z. Füredi and Y. Kim. Cycle-saturated graphs with minimum number of edges. *Journal of Graph Theory*, 73(2):203–215, 2013.

[52] Z. Füredi and M. Simonovits. The history of degenerate (bipartite) extremal graph problems. In *Erdős centennial*, pages 169–264. Springer, 2013.

[53] C. D. Godsil and G. F. Royle. Chromatic number and the 2-rank of a graph. *Journal of Combinatorial Theory, Series B*, 81(1):142–149, 2001.

[54] R. Graham and H. Pollak. On embedding graphs in squashed cubes. *Graph Theory and Appl., Springer Lecture Notes in Math.*, 303:99–110, 1972.

[55] A. Halfpap, B. Lidickỳ, and T. Masařík. Proper rainbow saturation numbers for cycles. *arXiv preprint arXiv:2403.15602*, 2024.

[56] P. Hall. On representatives of subsets. *Journal of the London Mathematical Society*, 10:26–30, 1935.

[57] C. R. Johnson and A. L. Duarte. The maximum multiplicity of an eigenvalue in a matrix whose graph is a tree. *Linear and Multilinear Algebra*, 46(1-2):139–144, 1999.

[58] M. Kamiński, V. V. Lozin, and M. Milanič. Recent developments on graphs of bounded clique-width. *Discrete Applied Mathematics*, 157(12):2747–2761, 2009.

[59] I. Kaplansky. *Linear algebra and geometry: a second course.* Chelsea Publishing Company, 2 edition, 1974.

[60] L. Kászonyi and Z. Tuza. Saturated graphs with minimal number of edges. *Journal of Graph Theory*, 10(2):203–210, 1986.

[61] P. Keevash, D. Mubayi, B. Sudakov, and J. Verstraëte. Rainbow Turán problems. *Combinatorics, Probability and Computing*, 16(1):109–126, 2007.

[62] D. König. Graphen und Matrizen. *Mat. Fiz. Lapok*, 38(10), 1931.

[63] J. B. Kruskal. The theory of well-quasi-ordering: A frequently discovered concept. *Journal of Combinatorial Theory, Series A*, 13(3):297–305, 1972.

[64] Y. Lan, Y. Shi, Y. Wang, and J. Zhang. The saturation number of C_6. *arXiv preprint arXiv:2108.03910*, 2021.

[65] A. Lane and N. Morrison. Improved bounds for proper rainbow saturation. *arXiv preprint arXiv:2409.15444*, 2024.

[66] A. Lane and N. Morrison. Proper rainbow saturation for trees. *arXiv preprint arXiv:2409.15275*, 2024.

[67] I. Leader and T. S. Tan. Odd covers of complete graphs and hypergraphs. *arXiv preprint arXiv:2408.05053*, 2024.

[68] A. Lempel. Matrix factorization over $GF(2)$ and trace-orthogonal bases of $GF(2^n)$. *SIAM Journal on Computing*, 4(2):175–186, 1975.

[69] B. Lidický, T. L. Martins, and Y. Pehova. Decomposing graphs into edges and triangles. *Combinatorics, Probability and Computing*, 28(3):465–472, 2019.

[70] L. Lovász. On covering of graphs. In *Theory of Graphs (Proc. Colloq., Tihany, 1966)*, pages 231–236. Academic Press New York, 1968.

[71] L. Lovász. On the Shannon capacity of a graph. *IEEE Transactions on Information Theory*, 25(1):1–7, 1979.

[72] L. Lovász, M. Saks, and A. Schrijver. Orthogonal representations and connectivity of graphs. *Linear Algebra and its Applications*, 114-115:439–454, 1989.

[73] É. Lucas. *Récréations mathématiques*, volume 2. Gauthier-Villars et fils, 1883.

[74] W. Mantel. Problem 28. In *Wiskundige Opgaven met de Oplossingen*. 1910. (solved by H. Gouwentak, W. Mantel, J. Teixeira de Mattos, Dr. F. Schuh, and Dr. W. A. Wythoff).

[75] A. Mohammadian. Trees and acyclic matrices over arbitrary fields. *Linear and Multilinear Algebra*, 64(3):466–476, 2016.

[76] C. S. J. Nash-Williams. Decomposition of finite graphs into forests. *Journal of the London Mathematical Society*, 1(1):12–12, 1964.

[77] V. Nikiforov. The spectral radius of graphs without paths and cycles of specified length. *Linear algebra and its applications*, 432(9):2243–2256, 2010.

[78] L. T. Ollmann. $K_{2,2}$-saturated graphs with a minimal number of edges. In *Proc. 3rd Southeastern Conference on Combinatorics, Graph Theory, and Computing*, pages 367–392, 1972.

[79] S.-i. Oum. Rank-width and vertex-minors. *Journal of Combinatorial Theory, Series B*, 95(1):79–100, 2005.

[80] T. Parsons and T. Pisanski. Vector representations of graphs. *Discrete Mathematics*, 78(1-2):143–154, 1989.

[81] R. Peeters. Orthogonal representations over finite fields and the chromatic number of graphs. *Combinatorica*, 16(3):417–431, 1996.

[82] L. Pyber. An Erdős-Gallai conjecture. *Combinatorica*, 5:67–79, 1985.

[83] J. Radhakrishnan, P. Sen, and S. Vishwanathan. Depth-3 arithmetic circuits for $S_n^2(x)$ and extensions of the Graham-Pollack theorem. In S. Kapoor and S. Prasad, editors, *FST TCS 2000: Foundations of Software Technology and Theoretical Computer Science. FSTTCS 2000. Lecture Notes in Computer Science*, volume 1974. Springer, Berlin, Heidelberg, 2000.

[84] F. Ramsey. On a problem of formal logic. *Proceedings of the London Mathematical Society*, 2(1):264–286, 1930.

[85] P. Rombach. Terminology for expressing a graph as a sum of cliques (mod 2). MathOverflow. URL:https://mathoverflow.net/q/318771 (version: 2018-12-16).

[86] S. Schwartz. An overview of graph covering and partitioning. Technical Report 20-24, ZIB, Takustr. 7, 14195 Berlin, 2020.

[87] A. Sen, K. Goodenough, and D. Towsley. Multipartite entanglement in quantum networks using subgraph complementations. In *2023 IEEE International Conference on Quantum Computing and Engineering (QCE)*, volume 2, pages 252–253. IEEE, 2023.

[88] D. E. Taylor. *The geometry of the classical groups*, volume 9 of *Sigma Series in Pure Mathematics*. Heldermann Berlin, 1992.

[89] P. Turán. Eine Extremalaufgabe aus der Graphentheorie. *Mat. Fiz. Lapok*, 48:436–452, 1941.

[90] Z. Tuza. A generalization of saturated graphs for finite languages. *Tanulmányok—MTA Számitástech. Automat. Kutató Int. Budapest,(185)*, pages 287–293, 1986.

[91] V. Vatter. Terminology for expressing a graph as a sum of cliques (mod 2). MathOverflow. URL:https://mathoverflow.net/q/317716 (version: 2018-12-15).

[92] V. G. Vizing. On an estimate of the chromatic class of a p-graph. *Diskret analiz*, 3:25–30, 1964.

[93] W. Wessel. Über eine Klasse paarer Graphen, i: Beweis einer Vermutung von Erdős, Hajnal and Moon. *Wiss. Z. Techn. Hochsch. Ilmenau*, 12:253–256, 1966.

[94] A. A. Zykov. On some properties of linear complexes. *Amer. Math. Soc. Translation*, (79):33, 1952. (Original work: *Mat. Sbornik N.S.* 24(66), 163–188 (1949); MR0035428).