8-2015

Graph centers, hypergraph degree sequences, and induced-saturation

Sarah Lynne Behrens

University of Nebraska-Lincoln, s-sbehren7@math.unl.edu

GRAPH CENTERS, HYPERGRAPH DEGREE SEQUENCES, AND

INDUCED-SATURATION

by

Sarah Lynne Behrens

A DISSERTATION

Presented to the Faculty of

The Graduate College at the University of Nebraska

In Partial Fulfilment of Requirements

For the Degree of Doctor of Philosophy

Major: Mathematics

Under the Supervision of Professor Stephen Hartke

Lincoln, Nebraska

August, 2015

GRAPH CENTERS, HYPERGRAPH DEGREE SEQUENCES, AND

INDUCED-SATURATION

Sarah Lynne Behrens, Ph.D.

University of Nebraska, 2015

Adviser: Stephen Hartke

The *center* of a graph is the set of vertices whose distance to other vertices is minimal.
The *centralizing number* of a graph G is the minimum number of additional vertices in
any graph H where $V(G)$ is the center of H. Buckley, Miller, and Slater and He and
Liu provided infinite families of graphs with each centralizing number. We show the
number of graphs with each nonzero centralizing number grows super-exponentially
with the number of vertices. We also provide a method of altering graphs without
changing the centralizing number and give results about the centralizing number of
dense graphs.

The *degree sequence* of a (hyper)graph is the list of the number of edges containing
each vertex. A *t-switch* replaces t edges with t new edges while maintaining the same
degree sequence. For graphs, it has been repeatedly shown that any realization of a
degree sequence can be turned into any other realization by a sequence of 2-switches.
However, Gabelman provided an example to show 2-switches are not sufficient for k-
graphs with $k \geq 3$. We classify all pairs of 3-graphs that do not admit a 2-switch but
differ by a 3-switch. We use this to provide support that 2-switches and a 3-switch
are sufficient for 3-graphs.

Given graphs G and H, G is *H-saturated* if G does not contain H as a subgraph,
but H is a subgraph of $G + e$ for any $e \notin E(G)$. While this is well defined for
subgraphs, the similar definition is not well defined for induced subgraphs. To avoid

this, Martin and Smith defined the *induced-saturation number* using trigraphs. We show that the induced-saturation number of stars is zero. This implies the existence of graphs that are star induced-saturated. We introduce the parameter indsat$^*(n, H)$ which is the minimum number of edges in an H-induced-saturated graph, when one exists. We provide bounds for indsat$^*(n, K_{1,3})$ and compute it exactly for infinitely many n.

COPYRIGHT

ACKNOWLEDGMENTS

This dissertation would not have been completed without the support of many friends and family. My parents were extremely supportive throughout my entire time as a graduate student. My friends and adoptive family at First Christian Church made Lincoln very quickly feel like home.

I would like to thank my advisor Stephen Hartke for his guidance over the past few years. I would also like to thank my committee members Jamie Radcliffe, Christine Kelley, Mohammad Rammaha, and Berthe Choueiry for their support.

I would not have gone to graduate school without the encouragement of my undergraduate advisor Jonathan White. He introduced me to mathematical research and inspired me to pursue an advanced degree.

As a graduate student I was fortunate enough to have opportunities to do research with graduate students from other universities. In particular I would like to thank Douglas West for being able to participate in the Combinatorics Research Experience for Graduate Students at the University of Illinois Urbana-Champaign. Through that program I met most of my collaborators. In particular I met Catherine Erbes, with whom I collaborated on two separate research papers. I enjoyed writing (and rewriting) those papers with Cathy.

When I first moved to Lincoln, it was excellent being able to live with Christina Edholm and Julia St. Goar. It was pleasant both academically and socially having other graduate students around. For day to day sanity, I am forever grateful to my officemates Caitlyn Parmelee and Kat Shultis. We regularly helped each other with a variety of things such as proofreading and teaching tips. I also enjoyed our calendar trivia and snowflakes, which we maintained even when we were in different states. If I ever hit a wall with research, I was always able to go bounce ideas off of James

Carraher or Lauren Keough. I am very appreciative of the tips and tricks they have shared with me through the years. Possibly the most tedious part of graduate school was grading exams. Those were made much more pleasant when I was able to grade with other graduate students. The most common one was with Ethan Twisdale. I am thankful for the company and moral support he provided over the years. Special thanks to Marilyn Johnson, who helped me get all of my paperwork turned in on time. Additionally, she was always willing to listen to how things were going, even if I just needed to destress.

Finally I would like to thank all of the faculty and graduate students that I have met while at UNL. I am sure there are more names that I have missed. I truly appreciate all the help that I have received over the years.

GRANT INFORMATION

Contents

List of Figures

List of Tables

List of Algorithms

Chapter 1

Introduction

This thesis has two major mathematical topics. The first topic is about understanding graphs that contain a subgraph in a specified way. Chapter 2 focuses on the subgraph being contained as the center, while Chapter 4 focuses on containment of induced copies of the specified subgraph. The second topic, which we study in Chapter 3, is examining hypergraph realizations of degree sequences. We are interested in specific methods to move among these realizations by making small changes and we focus on obstructions to this ability.

A recent trend in mathematics is the use of experimental mathematics: applying computational techniques to suggest theorem statements or provide counterexamples. Much of the work within this thesis is motivated by experimental results generated with a computer. These computations led to insights that in turn suggested conjectures and theorems. In addition, computations can be used within proofs of theorems, such as the Four Color Theorem [3, 77], Hales' proof of the Kepler conjecture [45], and the nonexistence of the projective plane of order ten [60]. In this thesis, we also use computation to prove some of our results.

In Chapter 2, we investigate centers of graphs. A *graph* G is an ordered pair of a vertex set $V(G)$ and edge set $E(G)$ such that each edge consists of two distinct vertices. The *center* of a graph is the set of vertices whose distance to other vertices is minimal.

Centers of graphs have been of interest for such applications as facility location and social networks. For example, when building an emergency facility such as a hospital, it should be placed so it is not too far from any one location. Similarly, consumers want to be located in places that are close to their desired services. For some recent papers, see [20, 21, 80]. See Buckley [16] for a survey, as well as [17, 18, 22, 23, 46, 48, 50, 61, 75].

Graphs whose center is the entire vertex set are called *self-centered*. Not all graphs

are self-centered. Buckley, Miller, and Slater [18] introduced the *centralizing number* $A(G)$ of a graph G, which is the minimum number of vertices that must be added to G to form a supergraph H where the subgraph induced by the center of H is isomorphic to G. Hedetniemi (see [18]) gave a construction showing the centralizing number is at most four. The centralizing number acts as a measure of how close a graph is to being self-centered.

Buckley et al. [18] determined the centralizing number of trees, classifying every tree with at least three vertices to have centralizing number two or four. However, they did not know if there existed any graphs with centralizing number three. Chen [24] and Bielak [11] provided the earliest examples of graphs with centralizing number three. He and Liu [67, 63] showed that, in fact, there are infinitely many graphs with centralizing numbers three and four. The natural next question is how rare these graphs are. In particular, what proportion of graphs on n vertices have centralizing number three or four? We develop an algorithm that we use to compute the exact number of graphs on at most twelve vertices with each centralizing number. This data suggests the number of graphs with each centralizing number is increasing for all n. We prove this statement, showing that the rate of increase is super-exponential.

Buckley et al. [18] gave a classification of the centralizing number for the sparsest graphs, trees. For very dense graphs, we prove the centralizing number is two unless the graph is one of two special cases and is self-centered. Additionally, we conjecture that the densest graphs with centralizing number three and four belong to specific families, and provide evidence toward this conjecture.

In Chapter 3 we investigate degree sequences of hypergraphs. A *hypergraph* is an ordered pair of a vertex set and an edge set, where edges are subsets of the vertices. Note that we only consider simple hypergraphs, where no edge can be repeated. A hypergraph is *k-uniform* if every edge contains exactly k vertices. Graphs are 2-

uniform hypergraphs. The *degree* of a vertex is the number of edges containing that vertex. The *degree sequence* of a hypergraph H is the list of all vertex degrees with their multiplicities, listed in nonincreasing order. We say H *realizes* that sequence. A sequence of nonnegative integers is *k-graphic* if there exists a k-uniform hypergraph realizing that sequence.

The fundamental question in the area of hypergraph degree sequences is which sequences of nonnegative integers are k-graphic. In the case of $k = 2$, this problem is well understood, with many characterizations of 2-graphic sequences. Two of the most well known characterizations are the Havel-Hakimi [47, 44] and the Erdős-Gallai [33] characterizations. Other characterizations are surveyed by Sierskma and Hoogeveen [79] and Mahadev and Peled [68]. However, there are significantly fewer results for $k \geq 3$.

Achuthan, Achuthan and Simanihuruk [1], Billington [13], and Choudum [25] gave some necessary conditions for a sequence to be k-graphic. Unfortunately, Achuthan et al. showed that none of these necessary conditions from [1, 13, 25] are sufficient. In fact, there are surprisingly few sufficient conditions for a sequence to be k-graphic for any $k \geq 3$.

Finding a minimal set of switches that connects the space of realizations has applications in network science; specifically, generating a realization chosen uniformly at random, which can be done using Markov chain Monte Carlo methods. The connectivity of the space for 2-graphs allows for uniform random sampling. For some results of using Markov chains for 2-graphs, see [35, 43, 52, 76]. Additionally, the Markov Chain Monte Carlo methods can be used to give an approximation of the number of realizations for a degree sequence [51]. If we have set of switches connecting the space of 3-graph realizations, then we can use Markov Chain Monte Carlo with the realizations of a 3-graphic sequence as well.

We make small changes to move from one realization to another by replacing a set of edges with another set of edges. A *t-switch* on a hypergraph H is replacing t edges with t other edges such that the resulting hypergraph has the same degree sequence as the original hypergraph. If a t-switch can be performed on a hypergraph, we say the hypergraph *admits* a t-switch. Given a k-graphic degree sequence π and a collection of switches S, we say the space of realizations of π is connected by S if for any two k-graph realizations of π, one can be obtained from the other by performing a sequence of switches from S.

Petersen [71] and subsequently others (for instance, Fulkerson, Hoffman, and McAndrew [39]) showed that, for any two realizations of a 2-graphic sequence, one realization can be obtained from the other by a sequence of 2-switches. However, Gabelman [40] demonstrated that 2-switches are not enough for hypergraphs. These examples were further generalized showing that a k-switch is needed for k-graphs, which leads to the questions of what other examples are there that are not connected by 2-switches and what is the minimal set of switches required to connect the space of realizations. Many results about 2-graphic sequences rely on the fact that the space of 2-graphic realizations is connected by 2-switches. If we know a minimum set of moves that connect the space of realizations for 3-graphic sequences, then it is likely that similar results can be proven for 3-graphic sequences.

The main result of Chapter 3 is a characterization of all pairs of 3-graphs that differ by a 3-switch but do not admit 2-switches. These 3-graphs are all related to the generalized Gabelman examples. This result provides evidence that the collection of 2-switches along with a specific 3-switch is enough to connect the space of realizations for 3-graphic sequences.

Since the previous result restricts to 3-graphs that do not admit a 2-switch, we also examine sequences similar to the degree sequences that we know have a realiza-

tion space not connected by 2-switches. In particular, we look at sequences that do not have repeated terms and few realizations. These sequences seem like the most likely candidates for degree sequences whose realization space is not connected by 2-switches, as fewer realizations have less room to move. Additionally, if there are repeated terms, then there is an increase in the number of realizations by the number of isomorphic copies. This provides more support that the collection of 2-switches along with a specific 3-switch (which we call a transversal 3-switch) is enough to connect the space of realizations for 3-graphic sequences. As an application of using switches, we also provide a result about packing hypergraphs.

In Chapter 4, we investigate a new version of the saturation number. The saturation number of a graph is a very classic parameter. The *saturation number* $\mathrm{sat}(n, H)$ for parameters n and graph H is the minimum number of edges in any graph G on n vertices such that H is not a subgraph of G, but H is a subgraph if any edge is added to G.

Erdős, Hajnal, and Moon [34] proved the earliest saturation result, computing the saturation number for complete graphs. Kászonyi and Tuza [54] expanded upon their work and showed that the saturation number for any graph on t vertices is less than the saturation number of the complete graph on t vertices. A survey of results related to the saturation number is provided by Faudree, Faudree, and Schmitt [36].

H is an induced subgraph of G if the vertices that H appears on have no other edges. For any graph H, adding enough edges to G eventually results in G containing H as a subgraph, but not necessarily as an induced subgraph. This lack of monotonicity leads to the need of a different definition for an induced subgraph version of the saturation number. Recently Martin and Smith [69] introduced a new definition relating induced graphs to trigraphs instead of graphs. A *trigraph* has a vertex set and each pair of vertices is either an edge, nonedge, or potential edge. A *realiza-*

tion of a trigraph is a graph with the same vertex set and whose edge set consists of all edges of the trigraph and any subset of the potential edges. A trigraph T is *H-induced-saturated* if no realization of T contains an induced H, but changing any edge or nonedge to a potential edge creates a realization containing H as an induced subgraph. The *induced-saturation number* of a graph H with respect to n is the minimum number of potential edges in a trigraph on n vertices that is H-induced saturated. Martin and Smith computed the induced-saturation number for the path on four vertices.

We determine the induced-saturation number for stars to be zero for sufficiently large n. This means that there are star-induced-saturated graphs that do not have potential edges in the trigraph. In such a case, we introduce the parameter $\mathrm{indsat}^*(n, H)$ to be the minimal number of edges in any H-induced-saturated graph on n vertices. We bound this parameter for the claw (a specific star) and provide the exact value for infinitely many values of n.

Chapter 2

Graph centers

2.1   Introduction

Given a graph G, the *centralizing number* of G is the minimum number of additional vertices in a graph H where G is isomorphic to the center of H. Kopylov and Timofeev [58] stated that for any graph G there exists such a graph H. Hedetniemi (see [18]) gave the construction of Figure 2.1 showing that the centralizing number of any graph is at most four. Buckley, Miller, and Slater [18] formally introduced the
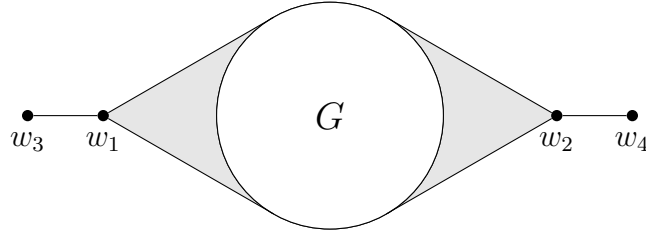
Figure 2.1: Hedetniemi's construction showing any graph G is the center of another graph. Given a graph G, add vertices w_1, w_2, w_3, and w_4 with w_1 and w_2 adjacent to every vertex of G, w_3 adjacent to w_1, and w_4 adjacent to w_2. In particular, this construction shows $A(G) \leq 4$.

concept of centralizing number and showed that every tree has centralizing number two or four. However, they did not find a graph with centralizing number three, and they asked if such a graph exists. Chen [24] and Bielak [11] answered affirmatively, Liu and He [67] provided an infinite family of graphs with centralizing number three, and Liu [63] provided an infinite family of graphs with centralizing number four.

In this chapter, we investigate the number of connected graphs on n vertices with a given centralizing number. We develop a practical algorithm for computing the centralizing number of a given graph, and use an implementation of the algorithm to determine the exact number of connected graphs on at most twelve vertices with each centralizing number. We examine centralizers generated by the algorithm to study properties of centralizers, such as their radius. In order to generate infinite families of graphs with centralizing number three or four, we provide results on the

effect twinning vertices has on the centralizing number of a graph. As an application of twinning vertices, we create families of graphs which we conjecture to be the densest graphs with centralizing number three or four. We conclude by providing a lower bound of the asymptotics for the number of graphs on n vertices with a given centralizer.

Centers of graphs have been of interest in topics such as facility location problems and social network problems. For example, when building an emergency facility such as a hospital, it should be placed so it is not too far from any one location. Similarly, consumers want to be located in places that are close to their desired services. For some recent papers, see [20, 21, 80]. Additionally, see Buckley [16] for a survey, as well as [17, 18, 22, 23, 46, 48, 50, 61, 75]. Since not all graphs are self-centered, we use the centralizing number as a measure of how close a graph is to being self-centered. We primarily focus on graphs with nonzero centralizing number.

Furthermore, it is not always ideal to have resources located in the center of a graph as described above. As suggested by Klavžar, Narayankar, and Walikar [55], a variant of this is placing two resources for the network that can efficiently serve the entire network, but need to be kept far apart due to interference. Klavžar, Narayankar, and Walikar referred to graphs containing all but two vertices in their center as almost self-centered graphs (ASC graphs). They provided a construction generating ASC graphs that is similar to twinning vertices. Balakrishnan, Brešar, Changat, Klavžar, Peterin, and Subhamathi [4] studied median and chordal ASC graphs. Note the center of any ASC graph has centralizing number at most two.

The centralizing number of several families of graphs have been determined. Buckley, Miller, and Slater [18] determined the centralizing number of trees, and Bielak [11] identified the centralizing number of all block graphs. Other infinite families of graphs with centralizing number two and three were found by He and Liu [67, 63]. We use

twinning of vertices to generate distinct infinite families of graphs with centralizing number three and four.

Kopylov and Timofeev [58] also focused on centralizers as their own entity instead of relating the centralizers to their central subgraph. They bounded the number of edges in a centralizer with respect to the number of vertices in its periphery. We pursue bounds on the number of edges in a graph with given centralizing number.

This chapter is organized as follows. Relevant definitions and notation are given in Section 2.1.1. In Section 2.2 we present an algorithm for computing the centralizing number of a graph. Section 2.3 contains the details of our implementation of the algorithm and the exact number of connected graphs on up to twelve vertices with each centralizing number, as well as questions that arose from examining examples generated by computations. The effect that twinning vertices has on the centralizing number of graphs is examined in Section 2.4. Section 2.5 contains an application of twinning vertices to create dense graphs with centralizing number greater than two. We further conjecture that these families are the densest such graphs and provide support for the conjecture. We conclude with the asymptotics of the number of graphs on n vertices with a given centralizer in Section 2.6.

2.1.1 Preliminaries

In general we follow the terminology and notation of West [84]. Let G be a graph with $u, v \in V(G)$. The *eccentricity* of v in G, denoted $\epsilon_G(v)$, is the maximum distance in G from v to any other vertex. The *horizon* of v in G, $\text{hor}_G(v)$, is the set of vertices in G that are at distance $\epsilon_G(v)$ from v in G. If v is adjacent to every vertex in $V(G) \setminus \{v\}$, then v is a *dominating vertex* of G, or v *dominates* G, and $\epsilon_G(v) = 1$. The *radius* of G, denoted $\text{rad}(G)$, is the minimum eccentricity of any vertex of G, while the

diameter of G, denoted diam(G), is the maximum eccentricity of any vertex of G. Vertices u and v are *diametral* if $d_G(u, v) = $ diam(G). Additionally, v is a *pendant* in G if $\deg_G(v) = 1$.

The *center* of a graph G is the set of vertices whose eccentricity is the radius of G. The *central subgraph $C(G)$* is the subgraph induced by the vertices in the center of G. The *periphery $P(G)$* of G is the set of vertices whose eccentricity equals the diameter of G.

Definition 2.1.1. The *centralizing number $A(G)$* of a graph G is the minimum number of additional vertices in a graph H where G is isomorphic to the center of H. In other words,

$$A(G) = \min_H \{|V(H)| - |V(G)| : C(H) \cong G\}.$$

A graph H is called a *centralizer* of a graph G if $C(H) = G$ and $|V(H) \backslash V(G)| = A(G)$. Note that in general, a centralizer of G is not unique.

If a graph has centralizing number zero, then we say the graph is *self-centered*. Alternatively, a graph G is self-centered if rad(G) = diam(G). Note that if G is a disconnected graph, then technically G is self-centered since every vertex has infinite eccentricity. Henceforth we will assume all graphs are connected.

We also introduce the following notation for convenience throughout the paper. Let G be a graph with centralizing number two or three, and let H be a centralizer of G. If $A(G) = 2$, then let w_1 and w_2 be the peripheral vertices of H; if $A(G) = 3$, then let w_1, w_2, and w_3 be the peripheral vertices of H. Since the farthest vertex from a vertex in $P(H)$ must be another vertex of $P(H)$, at least two pairs of the three peripheral vertices must be diametral if $A(G) = 3$. Throughout this paper we assume that if $P(H) = \{w_1, w_2, w_3\}$, then diam(H) $= d_H(w_1, w_2) = d_H(w_1, w_3)$.

Define $N_i = N_H(w_i)$ for $1 \leq i \leq A(G)$. The *peripheral neighborhood distance* is $d_H(N_1, N_2)$, which we denote by the letter p.

2.1.2 Background

As shown by Hedetniemi's construction (Figure 2.1), the centralizing number of a graph is at most four. Given a graph G with nonzero centralizing number and centralizer H, any farthest vertex from a vertex in $P(H)$ must be another vertex of $P(H)$. Hence there are no graphs with centralizing number one. Thus the centralizing number of any graph is either zero, two, three, or four.

Buckley, Miller, and Slater [18] determined the centralizing number of every tree, showing that only some of the numbers are attained.

Theorem 2.1.2 ([18])**.** *Let T be a tree with at least three vertices. If any two leaves of T are equidistant from the center of T, then $A(T) = 2$. Otherwise, $A(T) = 4$.*

This yields an infinite class of graphs with centralizing number two and another infinite class with centralizing four, but no examples with centralizing number three. Chen [24] proved the following lemma for graphs with centralizing number two, applying it to prove a specific graph had centralizing number three. Liu [63] adapted Chen's lemma to include graphs with centralizing number three.

Lemma 2.1.3 ([24, 63])**.** *Let G be a graph with centralizing number two or three, and let H be a centralizer of G. If $A(G) = 2$, then let w_1 and w_2 be the peripheral vertices of H; if $A(G) = 3$, then let w_1, w_2, and w_3 be the peripheral vertices of H with $\mathrm{diam}(H) = d_H(w_1, w_2) = d_H(w_1, w_3)$.*

1. $\mathrm{diam}(H) = p + 2$ *and* $\mathrm{rad}(H) = p + 1$. *Thus for any* $x \in V(H)$,

$$\epsilon_H(x) = \begin{cases} p + 2, & \text{for } x \in P(H); \\ p + 1, & \text{for } x \in V(G). \end{cases}$$

2. $1 \le p \le \mathrm{rad}(G)$.

3. *Let* $x_1 \in N_1$, *and for fixed* i *where* $2 \le i \le A(G)$, *let* $x_i \in N_i$. *Then*

   a) $d_H(x_1, x_i) = d_G(x_1, x_i) \in \{p, p+1\}$.

   b) $p = d_H(N_1, N_i) = d_G(N_1, N_i)$
   $= d_H(x_1, N_i) \ = d_G(x_1, N_i)$
   $= d_H(x_i, N_1) \ = d_G(x_i, N_1)$.

   Moreover, if $A(G) = 3$ *then* $p = d_H(x_1, N_2 \cup N_3) \ = d_G(x_1, N_2 \cup N_3)$.

4. *Suppose that* $v \in V(G)$ *has degree one in* G *and* u *is its neighbor in* G. *Then*

   a) v *is an element of some* N_i, *and*

   b) u *and* v *cannot belong to the same* N_i.

5. *Suppose* $\mathrm{diam}(G) > \mathrm{rad}(G) + 1$ *and* u, v *are vertices of* G *such that* $d_G(u, v) = \mathrm{diam}(G)$. *If* $A(G) = 2$ *then* u *and* v *cannot belong to different* N_i. *If* $A(G) = 3$ *then we cannot have* $u \in N_1$ *with* $v \in N_2 \cup N_3$, *nor* $v \in N_1$ *with* $u \in N_2 \cup N_3$.

Observe that if a graph G has centralizing number two or three, then the periphery of any centralizer H of G consists of exactly the vertices that are not in $C(H)$. Additionally, observe that $p \ge 1$ from statement (2) implies that w_1 and w_2 are not adjacent and have no common neighbors. If $A(G) = 3$, the same also holds for w_1 and w_3. However, w_2 and w_3 may be adjacent and may have common neighbors.

Furthermore, we can improve the lower bound on p when $A(G) = 3$.

Lemma 2.1.4. *If G has centralizing number 3 and H is a centralizer, then $p \geq 2$.*

Proof. Suppose H is a centralizer of G with $p = 1$. Construct the graph H' from H by identifying w_2 and w_3: $V(H') = V(G) \cup \{w_1, w_2\}$ and $E(H') = E(G) \cup \{w_1 v : v \in N_1\} \cup \{w_2 v : v \in M\}$. We show that H' contains G as its center, contradicting the fact that $A(G) = 3$. Note that $d_{H'}(w_1, w_2) = d_H(w_1, w_2) = 3$, and so $\epsilon_{H'}(w_1) = \epsilon_{H'}(w_2) = 3$.

It is left to show that $\epsilon_{H'}(v) = 2$ for $v \in V(G)$. Let $u \in V(H')$, and let P be a shortest u, v-path in H. If P does not contain w_3, then P is also a path in H'. If P does contain w_3, then we construct a u, v-path in H' by replacing w_3 with w_2 and shortening if P already contained w_2. Hence $\epsilon_{H'}(v) \leq \epsilon_H(v) = 2$. Since $p \geq 1$, v cannot be adjacent to both w_1 and w_2, and so $\epsilon_{H'}(v) = 2$. $\square$

It is worth noting that if G has centralizing number at most three and H is a centralizer of G, then each vertex of $P(H)$ is adjacent to some vertex of $C(H)$. However, this is not the case in Hedetniemi's construction of a centralizer for any graph G with $A(G) = 4$.

2.2 Algorithm for the centralizing number of a graph

We present a practical algorithm to compute the centralizing number of small graphs. The fundamental idea of the algorithm is to first check if a graph G is self-centered. If not, we add two vertices w_1 and w_2 to G in all possible ways in order to create a centralizer H. If still no centralizer is found, then three vertices are added to G in all possible ways to create a centralizer. If still no centralizer is found, then G has centralizing number four. A description in pseudocode of this method for checking if

a graph has centralizing number at most two is in Algorithm 1. The two main ideas that make the algorithm practical are an extension of the Floyd-Warshall algorithm and the selection of feasible neighborhoods of w_2. We describe each step in detail in the remainder of this section.

Algorithm 1 Computes if a graph has centralizing number zero, two, or greater than two.

Input: A graph G with n vertices.
Output: Returns whether the centralizing number is 0, 2, or greater than 2.

Use the Floyd-Warshall algorithm on G to compute pairwise distances.
if G is self-centered **then return** $A(G) = 0$
for $S \subseteq V(G)$ with $|S| \leq \frac{n}{2}$ **do**
    Create G' from G by adding a vertex w_1 with $N_{G'}(w_1) = S$.
    Use the Floyd-Warshall extension on G' to update pairwise distances.
    for feasible $T \subseteq \text{hor}_{G'}(w_1)$ with $|T| \geq |S|$ **do**
        Create H from G' by adding vertex w_2 with $N_H(w_2) = T$.
        Use the Floyd-Warshall extension on H to update pairwise distances.
        if $C(H) \cong G$ **then return** $A(G) = 2$
return $A(G) > 2$

2.2.1 Floyd-Warshall algorithm and extension

In order to compute the eccentricity of every vertex of a graph G, we compute the pairwise distances by applying the well known Floyd-Warshall algorithm [37, 83] to G. The output of the Floyd-Warshall algorithm is a pairwise distance matrix D_G, where $D_G[i, j]$ is the distance from vertex i to vertex j in G. The main idea of the Floyd-Warshall algorithm is to compute intermediate matrices D_G^k where $D_G^k[i, j]$ represents the distance from vertex i to vertex j by using only the first k vertices as intermediate vertices. Hence $D_G^n = D_G$ when G has n vertices. The Floyd-Warshall algorithm runs in $O(n^3)$ time; see [30] for more information. We use an optimized implementation of the Floyd-Warshall algorithm from Valerio [78].

When we compute the matrices $D_{G'}$ and D_H, we could apply the Floyd-Warshall algorithm again. However, we have already computed D_G, and so we use an extension of the Floyd-Warshall algorithm to compute $D_{G'}$ from D_G. We begin by computing $D_{G'}^n$, where G has n vertices and G' has $n+1$ vertices. After copying the pairwise distances $D_G[i,j]$ to $D_{G'}^n[i,j]$ for $1 \leq i,j \leq n$, we need only compute entries of $D_{G'}^n$ that are the distances to w_1. For each vertex i, $1 \leq i \leq n$, we have that $d(i,w_1) = 1 + \min\{d(i,j) : jw_1 \in E(G')\}$. Next, we compute $D_{G'}^{n+1}$ from $D_{G'}^n$ with one iteration of the Floyd-Warshall algorithm. The creation of $D_{G'}$ from D_G runs in $O(n^2)$ time.

We use this same extension to compute D_H from $D_{G'}$. This extension can be generalized to computing D_H for any graph H from D_G where G is an induced subgraph of H.

2.2.2 Selecting N_1

Given a centralizer H, we arbitrarily name the peripheral vertices w_1 and w_2. Hence we may assume that $\deg_H(w_1) \leq \deg_H(w_2)$. Since Lemma 2.1.3 (2) implies that w_1 and w_2 do not have any common neighbors, we conclude that $|N_1| \leq \frac{n}{2}$. Furthermore, we check possible neighborhoods of w_1 in order of increasing size. This heuristic improves the running time since most small graphs have at least one centralizer with a peripheral vertex of low degree.

2.2.3 Selecting N_2

As described in Algorithm 1, we add vertices w_1 and w_2 to G in separate steps instead of simultaneously. This allows us to eliminate some vertices of G as possible neighbors for w_2 based on the neighborhood of w_1.

Theorem 2.2.1. *Suppose G is a graph with centralizing number two. Let H be a centralizer of G with peripheral vertices w_1 and w_2. Let G' be the subgraph of H induced by $V(G) \cup \{w_1\}$, and let $S = \{v \in V(G) : d_G(v, u) \leq \epsilon_{G'}(w_1) \; \forall u \in N_{G'}(w_1)\}$. Then $N_2 \subseteq \hor_{G'}(w_1) \cap S$.*

Proof. There exists $y \in V(G)$ such that $d_{G'}(y, w_1) = \epsilon_{G'}(w_1)$. Suppose that $x \in N_1$ but $d_{G'}(x, w_1) \neq \epsilon_{G'}(w_1)$. Since H is a centralizer, $d_H(w_1, w_2) \geq d_H(w_1, x)$. Also, since every path in G' is also a path in H, we have that $d_H(w_1, x) \leq d_{G'}(w_1, x)$. If $d_H(w_1, x) < d_{G'}(w_1, x)$, then w_2 must be on a shortest w_1, x-path in H. This would imply that $d_H(w_1, w_2) < d_H(w_1, x)$, which is a contradiction. Hence we have $d_H(w_1, x) = d_{G'}(w_1, x)$. Similarly $d_H(w_1, y) = d_{G'}(w_1, y)$. Then $d_H(w_1, w_2) \leq d_H(w_1, y) + 1 \leq d_H(w_1, x)$, which is a contradiction. Hence $N_2 \subseteq \hor_{G'}(w_1)$.

Let $v \in N_2$. Since $N_2 \subseteq \hor_J(w_1)$, then $d_H(v, N_1) = \epsilon_J(w_1) - 1$. By Lemma 2.1.3 (3b), $p = \epsilon_J(w_1) - 1$. Lemma 2.1.3 (3a) implies $d_G(u, v) \in \{\epsilon_J(w_1) - 1, \epsilon_J(w_1)\}$ for any $u \in N_1$. Therefore $v \in S$. $\qquad\square$

Thus, the neighbors of w_2 must be in the set $\hor_{G'}(w_1) \cap S$. It is worth noting that we do not eliminate a vertex v as a possible neighbor of w_2 by checking if $d_G(v, u) \geq \epsilon_{G'}(w_1) - 1$ for all $u \in N_{G'}(w_1)$, since this is implied from $v \in \hor_{G'}(w_1)$. Furthermore, since we assumed that $\deg_H(w_1) \leq \deg_H(w_2)$, it is often the case that $|\hor_{G'}(w_1) \cap S|$ is smaller than the degree of w_1, and the algorithm simply moves on to the next neighborhood for w_1.

2.2.4   Beyond adding two vertices

After checking that a graph is not self-centered, we compute the diameter of the graph. If it is not self-centered but has a diameter of two, we know that the graph must have centralizing number two from Lemma 2.5.5.

When attempting to create a centralizer by adding three vertices instead of two, the same ideas apply. Recall that $\{w_1, w_2\}$ and $\{w_1, w_3\}$ are diametral pairs. Instead of concluding that $N_2 \subseteq S \cap \mathrm{hor}_{G'}(w_1)$, we have that $N_2 \cup N_3 \subseteq S \cap \mathrm{hor}_{G'}(w_1)$ where N_3 is the neighborhood of w_3. As a result, we cannot make any assumptions about the degree of w_1 in comparison to the degree of w_2, but we can assume that $\deg_H(w_2) \leq \deg_H(w_3)$. The other main difference when adding three vertices instead of two is that it is possible to have the edge $w_2 w_3$. Hence once N_1, N_2, and N_3 have been determined, we must check the eccentricities of H twice: once with the edge $w_2 w_3$ and once without.

2.2.5 Complexity

Algorithm 1 is efficient in practice, but has exponential running time in the worst case, which leads to the following question:

Question 2.2.2. *What is the computational complexity of computing the centralizing number of a graph?*

Computing whether or not a graph is self-centered is in P by using the Floyd-Warshall algorithm. Computing that a graph G has centralizing number at most t is an NP problem because we can verify a graph H is a centralizer for G in polynomial time by using the Floyd-Warshall algorithm.

2.3   Computational results

2.3.1   Centralizing number of small graphs

We implemented the algorithm of Section 2.2 in C++, and the code is available online.[1] We used `geng` from McKay's `nauty` package [70] to generate every connected graph (up to isomorphism) on at most twelve vertices, which was then piped to our program to compute the centralizing number. For graphs with eleven and twelve vertices, we performed the computations in parallel on the Open Science Grid [74]. The total computation time for each number of vertices is listed in Table 2.2.

The total number of connected graphs with each centralizing number for each n are listed in Table 2.1, and lists for all graphs with centralizing number three or four with at most twelve vertices are posted online.[2] Note that the smallest graph with centralizing number four has six vertices. The graph is a tree whose leaves are not equidistant from its center, so by Theorem 2.1.2 this graph has centralizing number four. The smallest graphs with centralizing number three are shown in Figure 2.2. The graph in Figure 2.2a was previously discovered by Bielak [11] and Liu and He [67]. Also note that the percentage of self-centered graphs increases with n, while the percentage of graphs with each nonzero centralizing number decreases for large enough n. We will discuss the growth rate of the number of graphs with each centralizing number in Section 2.6.

2.3.2   Radius and diameter of centralizers

Examining small graphs and their centralizers suggests comparing the radius and diameter of a centralizer to the radius and diameter of the original graph. Hedetniemi's

---

[1]http://www.math.unl.edu/~s-sbehren7/main/Data.html
[2]http://www.math.unl.edu/~s-sbehren7/main/Data.html

n	A(G)=0		A(G)=2		A(G)=3		A(G)=4	
1	1	100.0%						
2	1	100.0%						
3	1	50.0%	1	50.0%				
4	2	33.3%	4	66.7%				
5	5	23.8%	16	76.2%				
6	28	25.0%	83	74.1%			1	0.9%
7	223	26.1%	611	71.6%	5	0.6%	14	1.6%
8	3,151	28.3%	7,655	68.9%	127	1.1%	184	1.7%
9	79,673	30.5%	174,925	67.0%	3,988	1.5%	2,492	1.0%
10	3,853,870	32.9%	7,656,973	65.4%	163,049	1.4%	42,679	0.4%
11	357,767,742	35.5%	637,989,564	63.4%	9,924,086	1.0%	1,019,173	0.1%
12	63,104,972,067	38.5%	99,950,290,463	60.9%	967,896,377	0.6%	36,671,569	0.02%

Table 2.1: $A(G)$ distributions for fixed n. The first number is the total number of connected graphs on n vertices with $A(G) = t$ while the second number is a percentage out of all connected graphs on n vertices.
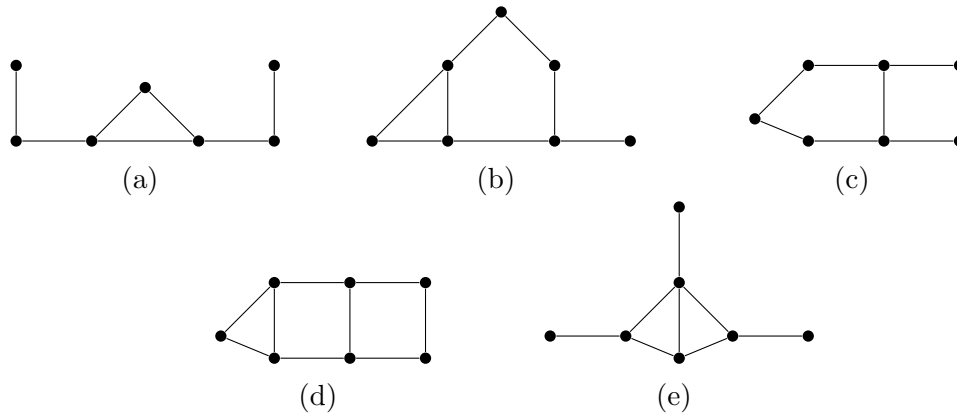
Figure 2.2: The smallest graphs with centralizing number three.

n	time
7	0s
8	3s
9	2m 10s
10	3h 17m 42s
11	13d 9h 36m 57s
12	10y 6d 23h 34m 05s

Table 2.2: The total computation time for fixed n.

construction shows that for any graph G with centralizing number four, there exists a centralizer H with $\mathrm{rad}(H) = 2$ and $\mathrm{diam}(H) = 4$. What bounds can be obtained on the radius of the centralizer of a graph with centralizing number two or three? If useful bounds could be obtained, we may be able to increase the computation speed of Algorithm 1.

We begin with some trivial bounds on the radius and diameter of centralizers.

Proposition 2.3.1. *Let G be a graph with centralizer H. If $A(G) = 2$, then $2 \leq \mathrm{rad}(H) \leq \mathrm{rad}(G) + 1$ and $3 \leq \mathrm{diam}(H) \leq \mathrm{rad}(G) + 2$. If $A(G) = 3$, then $3 \leq \mathrm{rad}(H) \leq \mathrm{rad}(G) + 1$ and $4 \leq \mathrm{diam}(H) \leq \mathrm{rad}(G) + 2$. The upper bounds in each case are sharp.*

Proof. Suppose G is a graph with centralizing number two and centralizer H. We use the bounds on p from Lemma 2.1.3 to obtain bounds on the radius and diameter of H. Since $1 \leq p \leq \mathrm{rad}(G)$ with $\mathrm{diam}(H) = p + 2$ and $\mathrm{rad}(H) = p + 1$, we conclude that $2 \leq \mathrm{rad}(H) \leq \mathrm{rad}(G) + 1$ and $3 \leq \mathrm{diam}(H) \leq \mathrm{rad}(G) + 2$. One sharpness example for the upper bounds is a graph G that is a star with central vertex x. Stars have a radius of one and diameter of two. A centralizer for G is the graph H where $V(H) = V(G) \cup \{w_1, w_2\}$ and $E(H) = E(G) \cup \{w_1, x\} \cup \{w_2, v : v \in V(G) \setminus \{x\}\}$. Then H has a radius of two and a diameter of three.

Similarly, suppose G is a graph with centralizing number three and centralizer H. We use the bounds on p from Lemmas 2.1.3 and 2.1.4 to obtain bounds on the radius and diameter of H. Since $2 \leq p \leq \mathrm{rad}(G)$ with $\mathrm{diam}(H) = p + 2$ and $\mathrm{rad}(H) = p + 1$, we conclude that $3 \leq \mathrm{rad}(H) \leq \mathrm{rad}(G) + 1$ and $4 \leq \mathrm{diam}(H) \leq \mathrm{rad}(G) + 2$. A sharpness example for the upper bounds is the centralizer H pictured in Figure 2.3 for the graph G induced by the solid vertices. The graph G has radius three while H has radius four and diameter five. $\square$
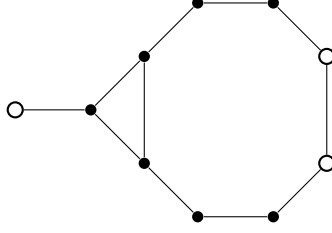
Figure 2.3: An example showing the upper bound of Proposition 2.3.1 for $A(G) = 3$ is sharp.

Our lack of a sharpness example for the lower bounds from Proposition 2.3.1 leads us to conjecture the lower bounds may be increased. In order to determine how much we could increase the lower bounds, we searched for any graph G with a centralizer H where $\mathrm{rad}(H) < \mathrm{rad}(G)$. Below are two examples of families of graphs G_t with a centralizer H_t where $\mathrm{rad}(H_t) = \mathrm{rad}(G_t) - 1$.

Proposition 2.3.2. *For $t \geq 2$, let C be a cycle on $2t + 4$ vertices with x and y a diametral pair, and let P be a path with t vertices whose endpoints are y and z. Let G_t be the graph formed by C and P, where y is the only common vertex of C and P. Let H_t be the graph containing G_t as a subgraph as well as vertices w_1 and w_2 with w_1 adjacent to x and z and w_2 adjacent to the four vertices of C that are a distance of one or two from y. Then H_t is a centralizer of G_t with $\mathrm{rad}(G_t) = t + 2$ and $\mathrm{rad}(H_t) = t + 1$.*

Proof. First observe that every vertex of C has eccentricity at least $t + 2$ in G_t in order to reach all vertices of C and every vertex of P has eccentricity at least $t + 2$ in G_t since any shortest path to x must use a x, y-path. In particular, $d_{G_t}(y, z) = t - 1$ and every vertex of G_t lies on a shortest x, y-path or y, z-path. Hence $e_{G_t}(y) = t + 2$ and $d_G(x, z) = 2t - 1$, so $\mathrm{rad}(G_t) = t + 2$ and G_t is not self-centered.

Hence in order to show $A(G_t) = 2$, it is left to show that $C(H_t) = G_t$. Note that $d_{H_t}(w_1, w_2) = t + 2$, so we need to show every vertex of G_t has eccentricity $t + 1$ in

H_t. Let C^1 be the cycle in H_t using a x, y-path in C, P, and w_1, and let C^2 be the cycle using the other x, y-path of C, P and w_1. Thus C^1 and C^2 are cycles on $2t + 3$ vertices. Let $v \in V(G_t)$. Then for any $u \in V(H_t)$, the pair u and v either lie in C, C^1, or C^2. If u and v are in C^1 or C^2, then $d_{H_t}(u, v) \le t + 1$. In particular, there exists a vertex in C^1 or C^2 whose distance in H_t from u is $t + 1$. If u and v are in C then either $d_C(u, v) \le t + 1$ or u and v are a diametral pair in C. If $d_C(u, v) = t + 2$, then there exists a shorter path in H_t by using w_2. Using the previous statements, we may conclude that at least one of the four neighbors of w_2 is a distance of at most t from u in H_t. Therefore, for any $u \in V(G_t)$, $\epsilon_{H_t}(u) = t + 1$. $\qquad\square$

A computer search has revealed that $G = G_1$ in the previous example, a graph with 9 vertices, is the smallest graph with a centralizer H and $\mathrm{rad}(H) < \mathrm{rad}(G)$.

Proposition 2.3.3. *For $t \ge 1$, let G_t be the graph formed by a disjoint union of a cycle on $4t + 6$ vertices with diametral pair x and y, and two paths with t vertices. Additionally, include the edges from x to an endpoint of one path and y to an endpoint of the other path. Let H_t be the graph containing G_t as a subgraph and the vertices w_1 whose neighbors are the two pendants of G_t and w_2 whose neighbors are the four vertices of the cycle that are a distance $t + 1$ from x or y. Then $\mathrm{rad}(G_t) = 2t + 3$ and $\mathrm{rad}(H_t) = 2t + 2$.*

We omit the proof of Proposition 2.3.3 because it is similar to the proof of Proposition 2.3.2.

The previous results lead us to conjecture the following.

Conjecture 2.3.4. *Suppose G is a graph with centralizing number two and $\mathrm{rad}(G) \ge 4$. If H is a centralizer of G, then $\mathrm{rad}(H) \ge \mathrm{rad}(G) - 1$.*

Thus if Conjecture 2.3.4 is true, the families of graphs from Propositions 2.3.2 and 2.3.3 show it is also sharp. Observe that Conjecture 2.3.4 is equivalent to $\mathrm{diam}(H) \geq \mathrm{rad}(G)$ since $\mathrm{diam}(H) = \mathrm{rad}(H) + 1$.

We modified Algorithm 1 to check the radius of every centralizer for each graph with centralizing number two and at most ten vertices. This computer search showed that Conjecture 2.3.4 holds for all graphs with at most ten vertices.

In order to provide support for Conjecture 2.3.4, we examine properties of a centralizer H for a graph G when $\mathrm{rad}(H) < \mathrm{rad}(G)$. In particular, we consider the role of the peripheral vertices in a graph's centralizer. Given a graph G with centralizer H, we define a vertex $w \in P(H)$ to be used *for transportation* if there exists a vertex $v \in V(G)$ with $u \in \mathrm{hor}_G(v)$ such that $\epsilon_H(v) < \epsilon_G(v)$ and a shortest u, v-path in H uses w. If G is a graph with centralizing number two and H a centralizer, then most centralizers have one peripheral vertex that is used for transportation while the other peripheral vertex keeps the first peripheral vertex out of the center. However, if $\mathrm{rad}(H) < \mathrm{rad}(G)$, this cannot be the case.

Proposition 2.3.5. *Let G be a graph with $A(G) = 2$ and centralizer H. If $\mathrm{rad}(H) < \mathrm{rad}(G)$, then both vertices in $P(H)$ are used for transportation.*

Proof. Let w_1 and w_2 be the peripheral vertices of H, and define
$$T_1 = \{(u,v) | u, v \in V(G), \epsilon_H(u) < \epsilon_G(u) = d_G(u,v), \text{ and } w_1 \text{ lies on a shortest } u, v\text{-path}\} \text{ and}$$
$$T_2 = \{(u,v) | u, v \in V(G), \epsilon_H(u) < \epsilon_G(u) = d_G(u,v), \text{ and } w_2 \text{ lies on a shortest } u, v\text{-path}\}.$$
Hence T_1 is the set of vertices in G using w_1 for transportation and T_2 is the set of vertices using w_2 for transportation.

If w_2 is never used for transportation, then $T_2 = \emptyset$. Since $\mathrm{rad}(H) < \mathrm{rad}(G)$, this implies that for every $v \in V(G)$, there exists some $u \in V(G)$ so that $(u,v) \in T_1$. Let $v \in V(G)$ be a vertex in the neighborhood of w_2, so $\epsilon_H(v) = \mathrm{rad}(H)$. Then $(u,v) \in T_1$

for some u, so there is a u, v-path P containing w_1 of length at most $\text{rad}(H)$. Hence $d_H(v, w_1) < \text{rad}(H)$ and thus $d_H(w_1, w_2) \leq \text{rad}(H)$, a contradiction. □

By Proposition 2.3.1, for any graph with radius two and centralizing number two, the radius of any centralizer for the graph is either two or three. However, if G is a graph with $A(G) = 2$, $\text{rad}(G) = 3$, and a centralizer H, then we have $2 \leq \text{rad}(H) \leq 4$. Hence the conclusion of Conjecture 2.3.4 holds for graphs with radius at most three. We conjecture that the lower bound can be further increased for graphs with radius three.

Conjecture 2.3.6. *Suppose G is a graph with centralizing number two and $\text{rad}(G) = 3$. If H is a centralizer of G, then $\text{rad}(H) \geq \text{rad}(G)$.*

In particular, if there is a counterexample to Conjecture 2.3.6, then it is a graph with diameter exactly four.

Proposition 2.3.7. *Let G be a graph with centralizing number two and $\text{rad}(G) = 3$. If there exists a centralizer H of G with $\text{rad}(H) = 2$, then $\text{diam}(G) = 4$.*

Proof. Since $A(G) = 2$, we have $\text{diam}(G) > \text{rad}(G)$. Suppose such an H exists with peripheral vertices w_1 and w_2. Suppose $u, v \in V(G)$ with $d_G(u, v) \geq 3$. Since $d_H(u, v) = 2$, either $u, v \in N_H(w_1)$ or $u, v \in N_H(w_2)$. Assume for the sake of contradiction that $\text{diam}(G) \geq 5$. Then there exist $x, y \in V(G)$ with $d_G(x, y) \geq 5$. Without loss of generality, $x, y \in N_H(w_1)$. For any $v \in V(G)$, either $d_G(x, v) \geq 3$ or $d_G(y, v) \geq 3$, and thus $v \in N_H(w_1)$. In particular, $N_H(w_1) = V(G)$, which contradicts Lemma 2.1.3. □

We also conjecture the lower bounds of Proposition 2.3.1 can be increased for graphs with centralizing number three.

Conjecture 2.3.8. *Suppose G is a graph with centralizing number three and* $\operatorname{rad}(G) \geq 4$. *If H is a centralizer of G, then $\operatorname{rad}(H) \geq \operatorname{rad}(G)$.*

From our computer search, the conjecture holds for all graphs with at most eleven vertices. By Proposition 2.3.1, if G is a graph with $A(G) = 3$ and radius two or three, then any centralizer H of G has $\operatorname{rad}(H) \geq 3$. Hence the conclusion of Conjecture 2.3.8 holds when $\operatorname{rad}(G) \leq 3$.

2.4   Effect of cloning and twinning vertices on $A(G)$

As stated in the introduction, there are examples of infinite families of graphs with each possible centralizing number. We introduce a method of constructing new graphs from some base graph, and then focus to the specific cases of cloning and twinning vertices. We use twinning and cloning vertices to create a new infinite family of graphs with centralizing number three and a new infinite family with centralizing number four. First, we introduce a sufficient condition for graphs to have centralizing number at most two.

Definition 2.4.1. A *clique* is a pairwise adjacent set of vertices. A clique K in a graph is a *dominating clique* if every vertex not in K is adjacent to at least one vertex in K. Note that a dominating clique is not necessarily a maximal clique.

Proposition 2.4.2. *Let G be a graph that is not complete and that contains a dominating clique. Then $A(G) \leq 2$. If G has diameter 3, then $A(G) = 2$.*

Proof. Let S be the vertices in the dominating clique that are adjacent to at least one vertex not in the clique. Construct H from G by adding vertex w_1 whose neighborhood

is S and adding vertex w_2 with neighborhood $V(G) \setminus S$. Then for $v \in S$, $\epsilon_H(v) = 2$ since v is adjacent to w_1, v can reach any vertex of $V(G) \setminus S$ via some vertex of S, and can reach w_2 via a neighbor in $V(G) \setminus S$. For any $v \in V(G) \setminus S$, $\epsilon_H(v) = 2$ because v is adjacent to w_2, v may reach any vertex of $V(G) \setminus S$ via w_2, and can use a neighbor in S to reach w_1 or any other vertex of S. Vertices w_1 and w_2 are in the periphery since every w_1, w_2-path has length at least three. Thus $C(H) = G$. Therefore $A(G) \leq 2$.

Since G has a dominating clique, the diameter of G is at most three. If the diameter of G is exactly three, then G has some vertex with eccentricity 3 in G. However any vertex of K has an eccentricity in G of at most two, so G is not self-centered. Thus $A(G) = 2$. $\qquad\square$

Corollary 2.4.3. *If G has a dominating vertex and is not complete, then $A(G) = 2$.*

Proof. If G has a dominating vertex d, then we apply Proposition 2.4.2 with $K = \{d\}$. Since G is not complete, there exists some vertex with eccentricity two, but $\epsilon_G(d) = 1$. Hence G is not self-centered, so $A(G) = 2$. $\qquad\square$

Now we introduce the graph operation $G \oplus_{u,S} J$, which generalizes a construction of Klavžar, Narayankar, and Walikar [55]. We use Corollary 2.4.3 to prove special cases about centralizing numbers of graphs $G \oplus_{u,S} J$.

Definition 2.4.4. Let G and J be graphs, with $u \in V(G)$ and $S \subseteq V(J)$ (S may be empty). Define $G \oplus_{u,S} J$ to be the graph obtained from the disjoint union of G and J by joining each vertex of J to all vertices in the open neighborhood of u in G, and joining each vertex of S to u.

The next lemma shows that the operation $G \oplus_{u,S} J$ essentially preserves distances from G.

Lemma 2.4.5. *Given a graph G with at least two vertices and a graph J with $u \in V(G)$ and $S \in V(J)$, let $G' = G \oplus_{u,S} J$. Then for any $v \in V(J)$, $x \in V(G) \setminus \{u\}$, and $y \in V(G)$, we have $d_{G'}(x, y) = d_G(x, y)$ and $d_{G'}(v, x) = d_{G'}(u, x)$. Hence $\epsilon_G(y) = \epsilon_{G'}(y)$. Furthermore, if u does not dominate G then $\epsilon_{G'}(u) = \epsilon_{G'}(v)$. When u is a dominating vertex of G, if $S = V(J)$ then $\epsilon_{G'}(u) = 1$, else $\epsilon_{G'}(u) = 2$. If $v \in S$ is also a dominating vertex of J, then $\epsilon_{G'}(v) = 1$, else $\epsilon_{G'}(v) = 2$ for any $v \in V(J)$.*

Proof. Let P be a shortest x, y-path in G'. If P does not have any vertices in J then $d_{G'}(x, y) = d_G(x, y)$. If P contains a vertex of J, then we may swap that vertex for u to create a path in G, and since P was a shortest path then $d_{G'}(x, y) = d_G(x, y)$. Hence $\epsilon_{G'}(u) = \epsilon_G(u)$ since (in G') u tis adjacent to every vertex of J.

Let Q be a shortest x, v-path in G'. Then v must be the only vertex of J in Q and the vertex next to v in Q must be a neighbor of u. Thus $d_{G'}(x, v) = d_{G'}(x, u)$. Hence $\epsilon_{G'}(x) = \epsilon_G(x)$.

This leaves the eccentricities in G' of vertices of J. The distance in G' between any two vertices of $V(J) \cup \{u\}$ is at most two. If u is not dominating in G, then $\epsilon_G(u) > 1$ and no vertex of J can dominate G', so $\epsilon_{G'}(u) = \epsilon_{G'}(v)$. Suppose u dominates G. If $S = V(J)$, then u is adjacent to every vertex of G', so $\epsilon_{G'}(u) = 1$. Otherwise there exists some vertex v of G that is nonadjacent to u, so $\epsilon_{G'}(u) = 2$. If some $v \in S$ is also dominating in J, then it is adjacent to every vertex of G', so $\epsilon_{G'}(v) = 1$. Otherwise for any $v \in V(J)$ there exists a vertex of $J \cup \{u\}$ that is not adjacent to v, so $\epsilon_{G'}(v) = 2$. $\square$

Theorem 2.4.6. *Given a connected graph G and a graph J, let $u \in V(G)$ be a non-dominating vertex of J, let $S \subseteq V(J)$, and let $G' = G \oplus_{u,S} J$. If $A(G) = 0$ or $A(G) = 2$, then $A(G') = A(G)$. If $A(G) = 3$ or $A(G) = 4$, then $2 \leq A(G') \leq A(G)$.*

If u dominates G and $S = V(G)$ then either both G and J are complete and $A(G') = 0$, or $A(G') = 2$.

If u dominates G and $S \neq V(J)$, then one of the following occurs:

1. *If $G = K_n$ for $n \geq 2$, then $A(G) = 0$ and $A(G') = 2$.*

2. *If u is the only dominating vertex of G and some $v \in S$ dominates J, then $A(G') = A(G) = 2$.*

3. *If u is the only dominating vertex of G and no $v \in S$ dominates J, then $A(G) = 2$ and $A(G') = 0$.*

4. *Otherwise $A(G) = A(G') = 2$.*

Proof. Let H be a centralizer of G and let $H' = H \oplus_{u,S} J$. First suppose u does not dominate G. In order to show $A(G') \leq A(G)$, it is enough to show that $C(H') = G'$ since H' has $A(G)$ more vertices than G'. By Lemma 2.1.3, H does not have a dominating vertex. Thus $\epsilon_{G'}(y) = \epsilon_G(y)$ for all $y \in V(G)$ and $\epsilon_{G'}(u) = \epsilon_{G'}(v)$ for all $v \in V(J)$ by Lemma 2.4.5. Hence all vertices in G' have equal eccentricities in H' and any vertices in $V(H') \setminus V(G')$ have higher eccentricities in H'. Therefore $C(H') = G'$.

Suppose for the sake of contradiction that $A(G) \geq 2$, but $A(G') = 0$. Then there exists at least one vertex of G whose eccentricities in G and G' are different, which contradicts Lemma 2.4.5. Thus $A(G) \geq 2$ implies that $A(G') \geq 2$.

Now suppose u is a dominating vertex of G with $S = V(J)$. Then u also dominates G'. If both G and J are complete, then G' is also complete, so $A(G) = A(G') = 0$. If either G or J is not complete then G' also has a non-dominating vertex, so $A(G') = 2$ by Corollary 2.4.3.

Finally suppose u is a dominating vertex of G with $S \neq V(J)$. Then u is not a dominating vertex of G'. If every vertex of G is dominating, then $G = K_n$, so G is self-

centered. Any vertex of $G \setminus \{u\}$ is dominating in G', so $A(G') = 2$ by Corollary 2.4.3. If u is the only vertex of G that dominates G, then by Corollary 2.4.3 we know that $A(G) = 2$. Furthermore, any vertex of $G \setminus \{u\}$ has eccentricity two in both G and G'. Additionally, $\epsilon_{G'}(u) = 2$ and any vertex of J has eccentricity at most two in G'. If there exists some $v \in S$ that dominates J, then v is also dominating in G', so $A(G') = 2$ be Corollary 2.4.3. Otherwise every vertex of G' has eccentricity two, so G' is self-centered. Otherwise G must have at least one other dominating vertex $x \neq u$, but G is not the complete graph. Then Lemma 2.4.5 tells us that x is also a dominating vertex of G', so applying Corollary 2.4.3 we conclude that $A(G) = A(G') = 2$. $\qquad\square$

The case of Theorem 2.4.6 where $A(G) = 2$ and $S = V(J)$ was proved by Klavžar, Narayankar, and Walikar [55]. Next we define cloning and twinning vertices, and show how the centralizing number of a graph changes after cloning or twinning a single vertex.

Definition 2.4.7. Given a graph G and a vertex v, the result of *cloning* v is a graph G' formed by adding a vertex v' whose neighborhood is $N_G(v)$. We say that v' is a *clone* of v in G'. The result of *twinning* v is a graph G'' formed by adding a vertex v'' whose neighborhood is $N_G(v) \cup \{v\}$. We say that v'' is a *twin* of v in G''.

Note that in the case where J is a single vertex v, if $S = \{v\}$ then $G \oplus_{u,S} J$ is the graph formed from G by twinning u, and otherwise $G \oplus_{u,S} J$ is the graph formed from G by cloning u. Hence we have the following immediate corollaries of Lemma 2.4.5 and Theorem 2.4.6.

Corollary 2.4.8. *Given a graph G with at least two vertices, let G' be any graph formed by cloning or twinning any vertex v of G. Let v' be the clone or twin of v.*

Then for any vertex $x \in V(G') \setminus \{v, v'\}$ and $y \in V(G') \setminus \{v'\}$, $d_{G'}(x, y) = d_G(x, y)$ and $d_{G'}(x, v') = d_{G'}(x, v)$. Hence $\epsilon_{G'}(x) = \epsilon_G(x)$ and $\epsilon_{G'}(v) = \epsilon_{G'}(v')$. If v' is a clone of a vertex that dominates G, then $\epsilon_{G'}(v) = 2 = \epsilon_G(v) + 1$. Otherwise, $\epsilon_{G'}(v) = \epsilon_G(v)$.

Corollary 2.4.9. *Given a graph G, let G' be any graph formed by twinning any vertex of G or cloning any non-dominating vertex of G. If $A(G) = 0$ or $A(G) = 2$, then $A(G') = A(G)$. If $A(G) = 3$ or $A(G) = 4$, then $2 \leq A(G') \leq A(G)$.*

If G' is a graph formed from G by cloning a dominating vertex v, then one of the following occurs:

1. *If $G = K_n$, then $A(G) = 0$ and $A(G') = 2$.*

2. *If v is the only dominating vertex of G, then $A(G) = 2$ and $A(G') = 0$.*

3. *Otherwise $A(G) = A(G') = 2$.*

Corollary 2.4.9 raises the question of whether twinning or cloning preserves the centralizing number. However, there exist graphs where cloning or twinning a vertex decreases the centralizing number. An example is shown in Figure 2.4. In Figure 2.4a, we provide the centralizer for the graph induced by the solid vertices, which has centralizing number three. When we clone vertex v, the resulting graph G' has centralizing number two, as shown in Figure 2.4b. Similarly, the graphs in Figures 2.5a and 2.5c have centralizing number four, while the cloning of vertex v results in a graph whose centralizing number is two (Figure 2.5b) and three (Figure 2.5d).

We checked graphs with at most eleven vertices and found examples where twinning a vertex causes a graph with centralizing number three to have centralizing number two, as shown in Figure 2.6. However, we did not find any graphs where twinning a vertex in a graph with centralizing number four resulted in a graph with lower centralizing number.
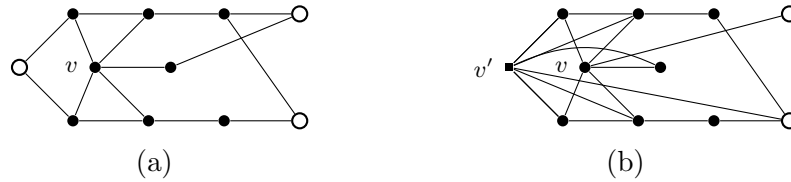
Figure 2.4: An example where cloning a vertex causes the centralizing number to decrease from three to two.

Question 2.4.10. *Can twinning a vertex in a graph with centralizing number four decrease the centralizing number?*
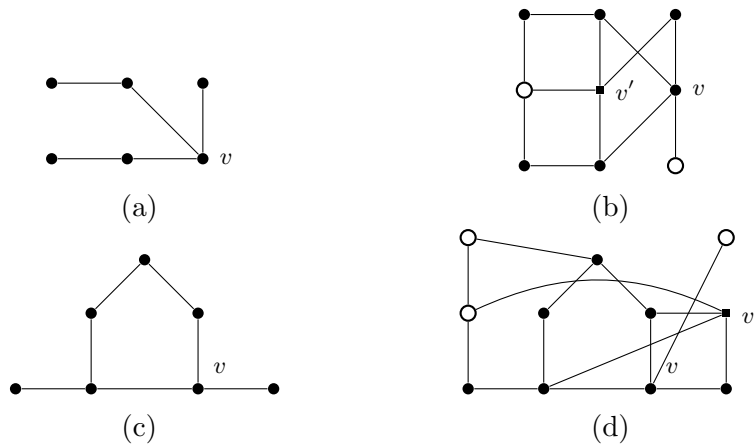
Figure 2.5: The graphs shown in (a) and (c) have centralizing number four. When cloning vertex v of (a), the resulting graph has centralizing number two. A centralizer is shown in (b). Similarly, cloning vertex v of (c) results with a graph with centralizing number three, and a centralizer is shown in (d).

Although cloning or twining a vertex may decrease the centralizing number, if a vertex has been cloned or twinned sufficiently many times then additional cloning or twinning will not cause the centralizing number to decrease.

Proposition 2.4.11. *Given a graph G, let G' be the graph formed by of cloning vertex v at least once. Let G'' be the graph formed by cloning v once more. If $A(G') \geq 3$, then $3 \leq A(G'') \leq A(G')$. The same result holds if twinning is used instead of cloning.*
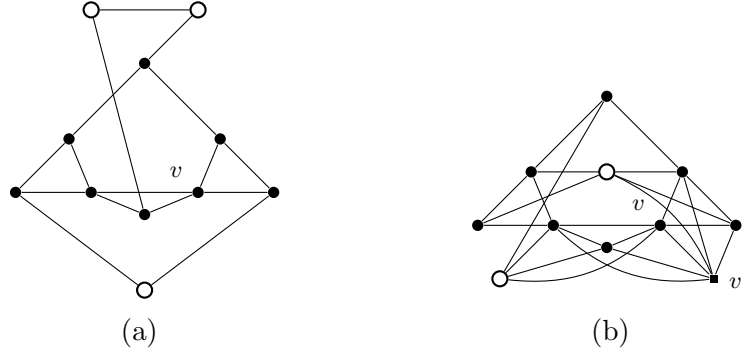
Figure 2.6: The graph in (a) is the centralizer for the subgraph induced by the solid vertices. The graph formed by twinning vertex v has lower centralizing number, as shown in (b).

Proof. By Corollary 2.4.9 $2 \leq A(G'') \leq A(G')$. Let $S = (V(G'') \setminus V(G)) \cup \{v\}$, so $|S| \geq 3$. Suppose to the contrary that $A(G'') = 2$ and let H be a centralizer of G'' with $P(H) = \{w_1, w_2\}$. Then by Lemma 2.1.3, w_1 and w_2 cannot have any common neighbors. Hence each vertex of S is adjacent to either w_1, w_2, or neither. Suppose some $s \in S$ is not adjacent to w_1 nor w_2. Let H' be the induced subgraph of H on $V(H) \setminus \{s\}$. Since $A(G) = 3$, Corollary 2.4.3 implies that G' does not have a dominating vertex. Hence G'' also does not have a dominating vertex. Thus Corollary 2.4.8 implies that $\epsilon_H(u) = \epsilon_{H'}(u)$ for all $u \in V(H')$, so H' is a centralizer for G with $P(H) = 2$, a contradiction.

Hence any $s \in S$ must be adjacent to w_1 or w_2. By the pigeonhole principle, we have at least two vertices in S adjacent to the same peripheral vertex. Let t be one of those vertices. Hence the induced subgraph H'' of H on $V(H) \setminus \{t\}$ has $\epsilon_H(u) = \epsilon_{H''}(u)$ for all $u \in V(H'')$ by Corollary 2.4.8. Therefore H'' is a centralizer for G with $|P(H'')| = 2$, a contradiction. □

Proposition 2.4.12. *Given a graph G, let G' be the graph formed by a sequence of cloning a vertex v at least three times. Let G'' be the graph formed by cloning v once more. If $A(G') = 4$, then $A(G'') = 4$. The same result holds if twinning is used*

instead of cloning.

Proof. By Proposition 2.4.11, $3 \leq A(G'') \leq 4$. Let $S = (V(G'') \setminus V(G)) \cup \{v\}$, so $|S| \geq 5$. Suppose to the contrary that $A(G'') = 3$ and let H be a centralizer of G'' with $P(H) = \{w_1, w_2, w_3\}$. Then by Lemma 2.1.3, w_1 and w_2 cannot have any common neighbors, and w_1 and w_3 cannot have any common neighbors. Hence each vertex of S must be adjacent to either w_1, w_2, w_3, both w_2 and w_3, or none of w_1, w_2, w_3. Suppose some $s \in S$ is not adjacent to no vertex of $P(H)$. Let H' be the induced subgraph of H on $V(H) \setminus \{s\}$. Since $A(G) = 4$, Corollary 2.4.3 implies that G' does not have a dominating vertex. Hence G'' also does not have a dominating vertex. Thus Corollary 2.4.8 implies that $\epsilon_H(u) = \epsilon_{H'}(u)$ for all $u \in V(H')$, so H' is a centralizer for G with $P(H) = 3$, a contradiction.

Hence any $s \in S$ must be adjacent to w_1, w_2, w_3, or both w_2 and w_3. By the pigeonhole principle, we have at least two vertices in S with the same set of neighbors in $P(H)$. Let t be one of those vertices. Hence the induced subgraph H'' of H on $V(H) \setminus \{t\}$ has $\epsilon_H(u) = \epsilon_{H''}(u)$ for all $u \in V(H'')$ by Corollary 2.4.8. Therefore H'' is a centralizer for G with $|P(H'')| = 3$, a contradiction. $\square$

As mentioned in the introduction, it has been previously shown that there are infinitely many graphs with centralizing number three and infinitely many with centralizing number four (see [11, 67, 87, 63]). Propositions 2.4.11 and 2.4.12 give us another way to generate infinitely many graphs with centralizing number three or four.

Proposition 2.4.13. *An infinite family of graphs with centralizing number four can be formed by cloning vertices.*

Proof. The graph in Figure 2.7a is a centralizer for the subgraph G_1 induced by the solid vertices. Since u is a clone of v in G_1 and has centralizing number three, itera-

tively applying Proposition 2.4.11 yields an infinite family of graphs with centralizing number three. □

Proposition 2.4.14. *An infinite family of graphs with centralizing number four can be formed by cloning vertices.*

Proof. The graph G_2 in Figure 2.7b is a graph with centralizing number 4. Observe that G_2 can be formed from the graph without vertices s,t, and u by twinning vertex v three times. Hence, continuing to twin vertex v we iteratively apply Proposition 2.4.12 to obtain an infinite family of graphs with centralizing number four. □
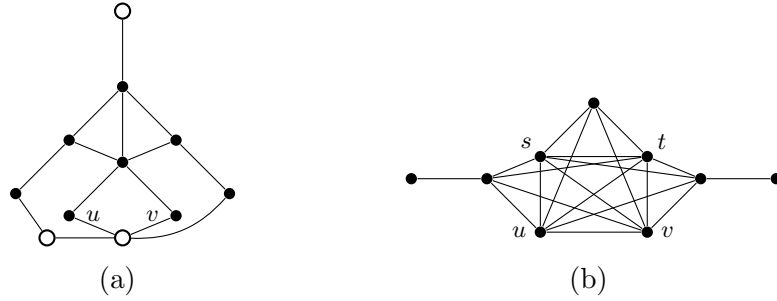
Figure 2.7: These graphs can be used to generate an infinite family of graphs with centralizing number three or four.

We conclude this section with another specific case of centralizing numbers with respect to twinned vertices.

Proposition 2.4.15. *Let G be a graph with a pendant vertex v, and let G' be the graph where v has been twinned. If $A(G) \leq 3$, then $A(G) = A(G')$. If $A(G) = 4$, then $A(G') \geq 3$.*

Proof. By Corollary 2.4.9, if $A(G) \geq 3$, then $A(G') \geq 2$. Hence it is left to show that if $A(G) \geq 3$, then $A(G') \geq 3$. So suppose $A(G') = 2$ with H a centralizer, and let u be the neighbor of v. By Lemma 2.1.3(4a), each of v and v' are adjacent to some vertex of $P(H) = \{w_1, w_2\}$. As argued in the proof of Proposition 2.4.11, without

loss of generality, v is adjacent to w_1 and v' is adjacent to w_2. Since v is adjacent to v', we conclude that $d_H(w_1, w_2) = 3$, so $\epsilon_H(w_1) = \epsilon_H(w_2) = 3$ and $\epsilon_H(y) = 2$ for $y \notin P(H)$. Since $A(G) > 2$, G cannot have a dominating vertex (Corollary 2.4.3). In particular, there exists some $x \in V(G)$ where $xu \notin E(G)$. Hence $xu \notin E(H)$. As $\epsilon_H(v) = \epsilon_H(v') = 2$, then $xw_1, xw_2 \in E(H)$, which contradicts $d_H(w_1, w_2) = 3$. □

2.5 Centralizing number of dense graphs

Theorem 2.1.2 [18] states that the sparsest connected graphs for a fixed number of vertices (trees) have either centralizing number two or four. We examine the opposite end of the spectrum and classify centralizing numbers of very dense graphs.

The following theorem is similar to a theorem of Buckley [15]. However, Buckley's focus was on self-centered graphs while we consider graphs that also have centralizing number two.

Theorem 2.5.1. *Let $\mathcal{G}_n$ be the collection of all graphs on n vertices with at least $\binom{n}{2} - \lceil \frac{n}{2} \rceil$ edges. Then there are exactly two graphs in $\mathcal{G}_n$ that are self-centered. The remaining graphs of $\mathcal{G}_n$ have centralizing number two.*

Proof. Suppose G is a graph with exactly $\binom{n}{2} - \lceil \frac{n}{2} \rceil$ edges, so the complement of G has $\lceil \frac{n}{2} \rceil$ edges. Recall that Corollary 2.4.3 states that any graph that is not complete and contains a dominating vertex has centralizing number two. Note that G has a dominating vertex if and only if the complement of G has an isolated vertex. The only way to avoid an isolated vertex in the complement is if the complement is a matching (for n even) or a matching plus an edge (for odd n). In this case, the eccentricity of every vertex in G is two, so G is self-centered.

The complete graph K_n is self-centered, so let G be a graph $G \neq K_n$ whose complement has at at least one edge and at most $\lfloor \frac{n-1}{2} \rfloor$ edges. This means there is at least one isolated vertex in the complement, so G has a dominating vertex. Then $A(G) = 2$ by Corollary 2.4.3. □

Thus, with the exception of the complete graph, the densest self-centered graph has $\binom{n}{2} - \lceil \frac{n}{2} \rceil$ edges, and any graph with more edges that is not the complete graph has centralizing number two. This leads us to ask what are the densest graphs with centralizing number three and four.

Definition 2.5.2. For $n \geq 6$, let $G_4(n)$ consist of a clique on $n - 5$ vertices, and three vertices that are adjacent to every vertex of the clique, two of which also have a pendant neighbor, as shown in Figure 2.8.

Proposition 2.5.3. *For $n \geq 6$, $A(G_4(n)) = 4$.*

Proof. When $n = 6$, we have a tree, which by Theorem 2.1.2 has a centralizing number of four. For $n \in \{7, 8, 9\}$, we confirm with a computer check that $G_4(n)$ also has a centralizing number of four. For $n > 9$, we apply Proposition 2.4.12 to conclude that $G_4(n)$ has a centralizing number of four because we can construct $G_4(n+1)$ by twinning a vertex in the maximum clique of $G_4(n)$. □

If we add edge $x'y'$ to $G_4(n)$, then adding w_1 with neighbors z, x', and y' and adding w_2 adjacent to the rest of the vertices in $G_4(n)$ yields a centralizer for $G_4(n)$ with the extra edge. If we add an edge from z to one of x, y, x', or y', then adding w_1 with neighbors x' and y' and adding w_2 whose neighbors are the vertices of the K_{n-5} yields a centralizer. Notice that by adding any remaining possible edge to $G_4(n)$, we may apply Proposition 2.4.2. In any case, adding an edge to $G_4(n)$ yields a graph with centralizing number two.
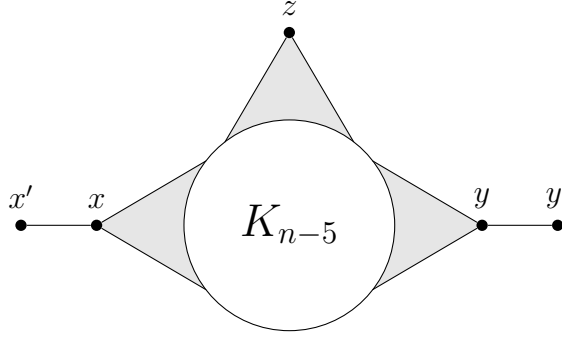
Figure 2.8: The graph $G_4(n)$. The circle indicates a clique of $n-5$ vertices, with x, y, and z adjacent to each vertex in the clique.

It is worth observing that $G_4(6)$ is the unique graph on 6 or fewer vertices with a centralizing number greater than two. A computer search has shown that $G_4(n)$ is the unique graph with the largest number of edges and centralizing number greater than two for $n \leq 11$.

Conjecture 2.5.4. *$G_4(n)$ is the unique graph with the largest number of edges on n vertices with a centralizing number larger than two.*

Since $G_4(n)$ has centralizing number four, Conjecture 2.5.4 implies that $G_4(n)$ is the densest graph with centralizing number four. The following statements provide support for Conjecture 2.5.4.

Lemma 2.5.5. *All graphs of diameter two have centralizing number at most two.*

Proof. If G is a graph with diameter two, then every vertex of G has eccentricity at most two. If every vertex has eccentricity two, then G is self-centered, so $A(G) = 0$. Otherwise there exists a vertex v with $\epsilon_G(v) = 1$, so v is a dominating vertex, and by Corollary 2.4.3, $A(G) = 2$. □

Lemma 2.5.6. *If G is a graph with $A(G) > 2$, then $\Delta(G) \leq n - 3$, where $\Delta(G)$ is the maximum degree of G.*

Proof. Suppose G is a graph with centralizing number greater than two. By Corollary 2.4.3, G cannot have a dominating vertex. Hence $\Delta(G) \leq n - 2$. If G has a vertex v of degree $n - 2$, then there is exactly one vertex u that is not adjacent to v. As G is a connected graph, there exists some vertex x that is adjacent to u. Assuming G has at least four vertices, then $\{v, x\}$ is a dominating clique satisfying Proposition 2.4.2. If G has less than four vertices, it is easy to check that $A(G) \leq 2$. Hence $\Delta(G) \leq n - 3$. $\qquad\square$

Theorem 2.5.7. *If G is a graph with $A(G) > 2$, then G has at most $\binom{n}{2} - \left(\frac{3}{2}n - 2\right)$ edges.*

Proof. By Lemma 2.5.5, if $A(G) > 2$, then G has diameter at least three. Hence there exists vertices x and y with $d_G(x, y) = 3$. Let $S = V(G) \setminus \{x, y\}$. Then for any $v \in S$, at least one of xv and yv is in $E(\overline{G})$. Additionally, $xy \in E(\overline{G})$. Since $A(G) > 2$, $\Delta(G) \leq n - 3$ (Lemma 2.5.6). Hence each vertex of S is not adjacent to at least two vertices of G. In order to satisfy this while maximizing the number of edges in G, we must have that one of those vertices is x or y and the other vertex is in S. Hence $|E(\overline{G})| \geq (n - 2) + 1 + \frac{1}{2}(n - 2) = \frac{3}{2}n - 2$. $\qquad\square$

In order to provide support for Conjecture 2.5.4, we also construct a necessary condition for a graph with n vertices to have at least $\binom{n}{2} - (2n - 2)$ edges and centralizing number at least two.

Recall from Proposition 2.4.2 that if a graph contains a dominating clique then the graph has centralizing number at most two.

Definition 2.5.8. If a vertex v is not adjacent to any vertex of some clique, then v prevents the clique from being dominating and so we say that v is a *blocking* vertex for the clique. Let $\beta(C)$ be the number of blocking vertices for a clique C.

Similarly, we say v is an *almost blocking* vertex for some clique if v is adjacent to exactly one vertex of the clique.

Hence if a graph G has centralizing number greater than two, every clique in G has at least one blocking vertex, so $\beta(C) \geq 1$ for any clique C of G.

Proposition 2.5.9. *Let G be a graph with $n \geq 20$ vertices, at least $\binom{n}{2} - (2n - 2)$ edges, and $A(G) > 2$. Let C be a maximal clique in G where $\beta(C)$ is minimized. Then C has only one blocking vertex, so $\beta(C) = 1$.*

Proof. Let $\beta = \beta(C)$, let $\gamma = |C|$, and let B be the set of blocking vertices for C. Let $P = V(G) \setminus (B \cup C)$ with $\rho = |P|$, so C, B, and P partition $V(G)$. In other words, P is the pool of vertices that are not in C nor B. Additionally, let A be the set of almost blocking vertices, and note $A \subseteq P$.

First we use Turán's theorem [82] to determine a bound on γ. By Turán's Theorem, if G is K_{r+1}-free then G has at most $(1 - \frac{1}{r})\frac{n^2}{2}$ edges. Setting $r = \frac{n}{5}$ yields $(1 - \frac{1}{r})\frac{n^2}{2} < \binom{n}{2} - (2n - 2)$. Hence G has a clique with at least $\lfloor \frac{n}{5} \rfloor + 1$ vertices, i.e. $\overline{G}$ has an independent set with at least $\lfloor \frac{n}{5} \rfloor + 1$ vertices. We conclude $\gamma \geq 5$ since $n \geq 20$.

Consider $\overline{G}$, the complement of G. In $\overline{G}$, C is an independent set, a blocking vertex of C is adjacent to every vertex of C, and an almost blocking vertex of C is adjacent to every vertex of C except one. Hence a dominating clique in G corresponds to an independent set in $\overline{G}$ where every vertex in $\overline{G}$ is nonadjacent to some vertex in the independent set. Additionally, since G is a graph with n vertices and at least $\binom{n}{2} - (2n - 2)$ edges, then $\overline{G}$ is a graph with n vertices and at most $2n - 2$ edges. Unless otherwise stated, assume all adjacency relations in the remainder of this proof are in $\overline{G}$.

Notice that every vertex of P is adjacent to at least one vertex of C, else we would contradict the maximality of C. Suppose a vertex p_0 in P is adjacent to exactly one vertex c_0 of C and is nonadjacent to at least one vertex of B. Then $(C \cup \{p_0\}) \setminus \{c_0\}$ is an independent set with γ vertices. Since β was minimized, then for every vertex of B that is not adjacent to p_0 there exists some vertex x in P that is adjacent to p_0 and every vertex of $C \setminus \{c_0\}$. Hence x is an almost blocking vertex.

Every almost blocking vertex has $\gamma - 1 \geq 4$ edges to C. Any vertex of $P \setminus A$ either has at least two edges to $B \cup C$ or has one edge to C and one edge to an almost blocking vertex. Every vertex of B is adjacent to every vertex of C. Hence we have accounted for at least $2\rho + \gamma\beta = 2n - 2\gamma - 2\beta + \gamma\beta$ edges in $\overline{G}$. Thus

$$2n - 2\gamma - 2\beta + \gamma\beta \leq 2n - 2,$$

which simplifies to

$$(\beta - 2)(\gamma - 2) \leq 2.$$

Since $\gamma \geq 5$ we have that $\beta \leq 2$.

Assume for the sake of contradiction that $\beta = 2$, and let $B = \{b_1, b_2\}$. Then every vertex of P is adjacent to at least one vertex of C, and every vertex of B is adjacent to every vertex of C. Hence we have accounted for $\rho + 2\gamma$ edges, and the number of remaining edges in $\overline{G}$ is at most

$$(2n - 2) - (\rho + 2\gamma) = n - \gamma = \rho + 2. \tag{2.5.1}$$

We next consider the following cases, one of which must occur.

Case 1: Some vertex p_0 of P is adjacent to exactly one vertex of C and is nonadjacent to at least one vertex of B. In this case, there are more edges that we can

account for. As described above, p_0 must be adjacent to an almost blocking vertex. Each almost blocking vertex in P is adjacent to all but one vertex of C, so we have an additional $\gamma - 2 \geq 3$ edges. Each vertex p_0 of P that is adjacent to exactly one vertex of C and is nonadjacent to at least one vertex of B must be adjacent to at least two vertices of $A \cup B$, and each remaining vertex of P is adjacent to at least one more vertex of $B \cup C$. Hence we have accounted for at least $\rho + 3$ more edges, which is larger than the value from Equation 2.5.1, and hence is too many.

Case 2: Every vertex of P is adjacent to at least two vertices in C or is adjacent to both b_1 and b_2 as well as a vertex of C. In this case, each vertex of P has at least one more edge leaving P than in the previous total. Since there were $\rho + 2$ edges left after Equation 2.5.1, at most two vertices of P can be adjacent to a third vertex in $B \cup C$. In particular, at most two vertices of P can be adjacent to both b_1 and b_2. Hence we will break this case into three subcases based on the number of vertices in P that are adjacent to both b_1 and b_2.

First observe that there cannot exist $p_0 \in P$ nonadjacent to $B \cup P \setminus \{p_0\}$. If such a p_0 exists, then p_0 has some non-neighbor $c_0 \in C$. Then $\{p_0, c_0\}$ form a dominating clique in G.

Subcase 1: Exactly two vertices of P are adjacent to both b_1 and b_2. Let p_1 and p_2 be the two vertices of P that are adjacent to both b_1 and b_2. If there exists a third vertex $p_3 \in P$, p_3 is nonadjacent to b_1 nor b_2 and there exists some vertex $c_1 \in C$ that is also nonadjacent to p_3. Hence $\{p_3, c_1\}$ is a dominating clique in G. Thus $\rho = 2$ and we have accounted for exactly the number of edges from Equation 2.5.1. Note b_1 and b_2 are adjacent to every vertex except for each other, which implies the edge $b_1 b_2$ in G is its own component. However, G cannot be disconnected if G has a nonzero centralizing number.

Subcase 2: Exactly one vertex $p_1 \in P$ is adjacent to both b_1 and b_2. Then by

Equation 2.5.1, $\overline{G}$ can have one more edge e that has not yet been described. Since every vertex of P must be adjacent to some vertex of $B \cup P$, if e is not contained in $B \cup P$ then $\rho = 1$. This implies that the edge $b_1 b_2$ is its own component in G, contradicting that $A(G) > 2$. If $\rho = 1$ and the remaining edge e is contained in $B \cup P$, the pair of vertices that we have not declared adjacent are b_1 and b_2. Thus $e = b_1 b_2$, which makes b_1 dominating, contradicting that G is connected. Hence $\rho \geq 2$, and e is used in $B \cup P$. If $\rho > 2$, then since every vertex of P must be adjacent to some vertex of $B \cup P$, we must have $P = \{p_1, p_2, p_3\}$ with $e = p_2 p_3$. Since $\gamma \geq 5$, there exists an $c_1 \in C$ that is not adjacent to p_2 nor p_3, thus $\{p_2, c_1\}$ is a dominating clique in G. Hence $\rho = 2$ with $P = \{p_1, p_2\}$ and p_2 has two neighbors c_1 and c_2 in C.

Since p_2 must have a neighbor in $B \cup P$, then e must be from p_2 to p_1 or from p_2 to a vertex in B, say b_1. In the former case $\gamma \geq 5$ implies that there exists some $c_3 \in C$ that is not adjacent to p_1 nor p_2. Thus p_2, c_3 is a dominating clique in G. Hence the latter case holds. Either $N_C(p_1) \subsetneq \{c_1, c_2\}$, say c_1, or p_1 has a distinct neighbor c_3 in C. This completely determines $\overline{G}$, so we examine the centralizing number of G. G is the graph in Figure 2.9a or Figure 2.9b, respectively. For either graph in Figure 2.9, add vertices $\{w_1, w_2\}$ with $N(w_1) = \{p_2\}$ and $N(w_2) = \{b_1, c_1\}$. This yields a graph with two more vertices than G and whose central subgraph is G. Thus $A(G) \leq 2$.

Subcase 3: Every vertex of P is adjacent to at most one of b_1 and b_2. Hence we assume every vertex of P has two neighbors in C. Thus we have accounted for ρ of the edges in Equation 2.5.1 and hence have at most two edges remaining before counting edges within $B \cup P$. Since every vertex in P has a neighbor in $B \cup P$, at least one edge must be used in $B \cup P$ that has at least one endpoint in P. Suppose only one edge is used. If it is contained entirely in P, say $p_1 p_2$, then since $\gamma \geq 5$, there exists a vertex $c_0 \in C$ that is not adjacent to p_1 nor p_2 (and hence p_1, c_0 is a dominating clique in G) unless $\gamma = 5$ and p_1 has that third edge to a vertex in C and p_1 and
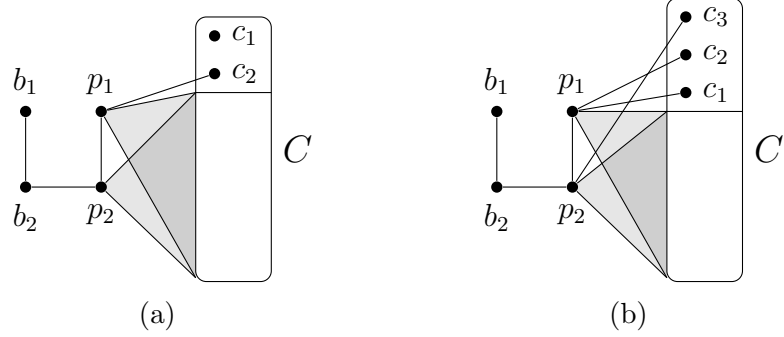
Figure 2.9: Graphs described in Subcase 2. In both graphs, C is a clique and p_2 is adjacent to every vertex of $C \setminus \{c_1, c_2\}$. In (a), p_1 is adjacent to every vertex of $C \setminus \{c_1\}$. In (b), p_1 is adjacent to every vertex of $C \setminus \{c_3\}$.

p_2 have disjoint neighborhoods in C. In this special case $\overline{G}$ has been determined, so we examine G, which is the self-centered graph in Figure 2.10a. Hence the single edge must be $p_1 b_1$ (or $p_1 b_2$). Once again, the edges of $\overline{G}$ have been determined, so we examine G. G is the graph in Figure 2.10b. Add vertices $\{w_1, w_2\}$ with $N(w_1) = \{p_1\}$ and $N(w_2) = \{b_1, c_1, c_2\}$. This is a centralizer for G, so $A(G) = 2$.

If one edge is in B, then the other edge must be $p_1 p_2$, else B would contain a dominating vertex. (Therefore we must use both edges.) Since $\gamma \geq 5$, there exists some $c_0 \in C$ that is not adjacent to p_1 nor p_2, and c_0, p_1 is a dominating clique in G.

Thus both edges have at least one endpoint in P. Hence we have two independent vertices p_1 and p_2 in P. Since $\gamma \geq 5$, there exists some $c \in C$ that is not adjacent to p_1 nor p_2. The only way p_1, p_2, c is not a dominating clique in G is if, without loss of generality, b_1 is adjacent to both p_1 and p_2. Then p_1, p_2, b_2 is a dominating clique in G unless p_1 and p_2 have a common neighbor in C. As this determines $\overline{G}$, we examine G. In $\overline{G}$, if p_1 and p_2 have the same two neighbors in C, then G is the graph in Figure 2.10c. Adding vertices $\{w_1, w_2\}$ with $N(w_1) = \{p_1, p_2\}$ and $N(w_2) = \{c_1, c_2, b_1\}$ yields a centralizer for G, so $A(G) = 2$. Otherwise p_1 and p_2 have exactly one common neighbor in C with respect to $\overline{G}$. In this case, G is the graph

(a) (b) (c) (d)
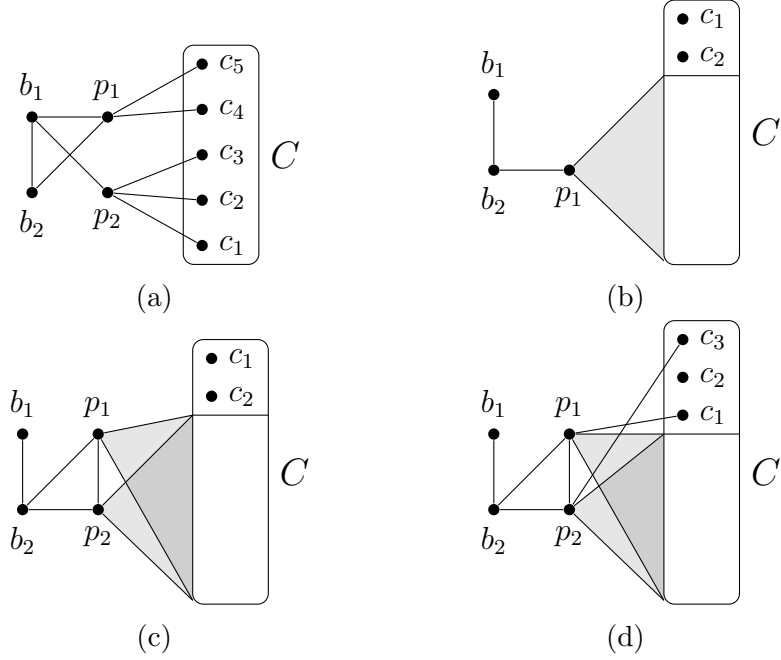
Figure 2.10: Graphs described in Subcase 3. In each graph, C is a clique.

in Figure 2.10d. Add vertices $\{w_1, w_2\}$ with $N(w_1) = \{p_1, p_2\}$ and $N(w_2) = \{b_1, c_2\}$. This yields a centralizer for G, so $A(G) = 2$.

We have concluded that each of the cases cannot happen. Therefore $\beta = 1$. □

Observe the graph $G_4(n)$ satisfies the hypothesis of Proposition 2.5.9 and has a clique with only one blocking vertex. One way to partition the vertices of G into $B \cup C \cup P$ as described within Proposition 2.5.9 is to let y' be the blocking vertex, let $P = \{x', y, z\}$, and let C be the remaining vertices. Proposition 2.5.9 shows that any dense graph with centralizing number greater than two must be structurally similar to $G_4(n)$.

We also conjecture that the densest graphs with centralizing number three are subgraphs of $G_4(n)$.

Definition 2.5.10. Let $G_3^1(n)$ be the subgraph of $G_4(n)$ where one edge to vertex z has been deleted. Let $G_3^2(n)$ be the subgraph of $G_4(n)$ where one edge within the

initial clique of $n - 5$ vertices has been deleted.

Proposition 2.5.11. *For $n \geq 8$, $A(G_3^1(n)) = A(G_3^2(n)) = 3$.*

Proof. For $n = 8$ and $n = 9$, we confirm with Algorithm 1 that $G_3^1(n)$ and $G_3^2(n)$ have centralizing number three. For $n \geq 9$, we can construct $G_3^1(n)$ from $G_3^1(n)$ and $G_3^2(n)$ from $G_3^2(n)$ by twinning a vertex. Hence we apply Proposition 2.4.11 to conclude that $G_3^1(n)$ and $G_3^2(n)$ have centralizing number three for $n \geq 10$. □

Conjecture 2.5.12. *For $n \geq 9$, $G_3^1(n)$ and $G_3^2(n)$ are the only two graphs of maximal size on n vertices with centralizing number three.*

We showed in Proposition 2.5.11 that $G_3^1(n)$ and $G_3^2(n)$ have centralizing number three. If Conjecture 2.5.4 is proven, then these are examples of maximal size with centralizing number three and only uniqueness would remain to be proven.

2.6 The asymptotic number of graphs on n vertices with given centralizing number

In Section 2.3.1 we determined the number of graphs with each centralizing number with up to twelve vertices. The number of graphs with each centralizing number increases with the number of vertices, but each of the percentages for nonzero centralizing numbers decrease. We proceed by giving more general results on the asymptotic number of graphs with each centralizing number for a fixed n.

Theorem 2.6.1. *Almost all graphs have centralizing number zero.*

Proof. Bollobás [14] proved that almost all graphs have diameter 2 but do not have a dominating vertex. By Lemma 2.5.5, such graphs are self-centered. □

This leaves us with the question of how many graphs with a fixed number of vertices actually have a nonzero centralizing number. We use dominating vertices to construct a lower found on the number of graphs on n vertices with centralizing number two.

Let $g(n)$ represent the number of unlabeled graphs on n vertices. Note that $g(n) \geq \frac{2^{\binom{n}{2}}}{n!}$ since each unlabeled graph corresponds to at most $n!$ labeled graphs.

Theorem 2.6.2. *The number of graphs on $n \geq 2$ vertices with centralizing number two is bounded below by $g(n-1) - 1$.*

Proof. Let G be a graph on $n-1$ vertices that is not complete, and let G' be G with a dominating vertex. By Corollary 2.4.3, $A(G') = 2$. In order to recover the original graph G from any graph that has been created by adding a dominating vertex, we remove a dominating vertex. $\square$

We use a construction of Liu to determine a lower bound for the number of graphs on n vertices with centralizing number three or four.

Theorem 2.6.3 ([65]). *Let G be a graph with $x, y \in V(G)$ where $d_G(x,y) = \mathrm{diam}(G) = 2m - 1$, and let $n > m$. Construct G' from G by taking the disjoint union of G and two P_ns and adding an edge from x to an endpoint of the first P_n and an edge from y to an endpoint of the second P_n.*

1. *If $\forall u \in V(G)$ we have $d_G(x,u) + d_G(u,y) \leq 2m - 1$, then $A(G') = 2$.*

2. *If $\forall u \in V(G)$ we have $d_G(x,u) + d_G(u,y) \leq 2m$ and there exists a vertex v with $d_G(x,v) = d_G(v,y)$ then $A(G') = 3$.*

3. *Otherwise, $A(G') = 4$.*

The proof of Theorem 2.6.3 heavily relies on Lemma 2.1.3. Using Theorem 2.6.3, we will construct families of graphs which then can be enumerated to obtain a lower bound on the number of graphs on n vertices with centralizing number three or four.

Theorem 2.6.4. *The number of graphs on $n \geq 11$ vertices with centralizing number three is bounded below by $g(n - 10)$, while the number of graphs on $n \geq 12$ vertices is bounded below by $g(n - 11)$.*

Proof. Given a graph G, fix $s \in V(G)$. We define graph $J_3(G)$ that satisfies part (2) of Theorem 2.6.3 and graph $J_4(G)$ that satisfies part (3) of Theorem 2.6.3. Graphs $J_3(G)$ and $J_4(G)$ are shown in Figure 2.11.

Define the graph $J_3(G)$ by $V(J_3(G)) = \{x, y, t, z\} \cup V(G)$ and $E(J_3(G)) = E(G) \cup \{xv, tv : v \in V(G)\} \cup \{yt, zs, zt\}$. Observe that $d_{J_3(G)}(x, y) = 3$ via the path x, v, t, y for any $v \in V(G)$. There is no shorter x, y-path since x is not adjacent to t and y is not adjacent to any vertex of G. Since z and each vertex of G is adjacent to t, we have $\epsilon_{J_3(G)}(x) = \epsilon_{J_3(G)}(y) = 3$, $\epsilon_{J_3(G)}(z) = 2$, and $\epsilon_{J_3(G)}(v) = 2$ for any $v \in V(G)$. Hence $\text{diam}(J_3(G)) = 3$, and for any $v \in V(J_3(G)) \setminus \{z\}$, we have $d_{J_1}(x, v) + d_{J_1}(v, y) = 3$. Furthermore $d_{J_3(G)}(x, z) = d_{J_3(G)}(z, y) = 2$, so $J_3(G)$ satisfies the hypothesis of Theorem 2.6.3 (2).

Next we define the graph $J_4(G)$ by $V(J_4(G)) = V(J_3(G)) \cup \{z'\}$; $E(J_4(G)) = E(J_3(G)) \cup \{zz'\}$. Let $u, v \in V(J_3(G))$ and note that $d_{J_4(G)}(u, v) \leq 3$ since $J_3(G)$ is a subgraph of $J_4(G)$. Also $d_{J_4(G)}(z', v)$ for any $v \in V(J_4(G) \setminus \{z'\}$. Thus $\text{diam}(J_4) = 3$ with $d_{J_4(G)}(x, z') + d_{J_4(G)}(y, z') = 6$. Hence $J_4(G)$ satisfies the hypothesis of Theorem 2.6.3 (3).

Finally, for $i \in \{3, 4\}$ we define $J_i'(G)$ by taking the disjoint union of $J_i(G)$ and two P_3s, adding an edge from an endpoint of one P_3 to x, and an edge from an endpoint of the other P_3 to y. By Theorem 2.6.3, $A(J_i'(G)) = i$ for $i \in \{3, 4\}$. $J_3'(G)$ has 10

vertices that are not in $V(G)$, and $J_4'(G)$ has 11 vertices that are not in $V(G)$. Since G was an arbitrary graph, then for $i \in \{3, 4\}$, the graph G can be any graph on the remaining $n - (7 + i)$ vertices. Now J_i' is a graph with n vertices and has centralizing number i. Hence for $i \in \{3, 4\}$, there are at least as many graphs on n vertices with centralizing number i as there are unlabeled graphs on $n - (7 + i)$ vertices.

In order to show that each $J_3'(G)$ and $J_4'(G)$ is unique, it is left to show that we can recover G. Assume G has at least two vertices. The graph $J_3'(G)$ has two pendant vertices. One pendant is part of a P_5 where the intermediate vertices have degree two in $J_3'(G)$, while the other pendant is part of a P_4 where the intermediate vertices have degree two in $J_3'(G)$. The non-pendant endpoint of the P_4 is x. Then G is the subgraph induced by the neighbors of x in $J_3'(G)$ except for the neighbor that is also on the P_4. The same argument holds for $J_4'(G)$, except that it has a third pendant vertex whose neighbor has degree three. □
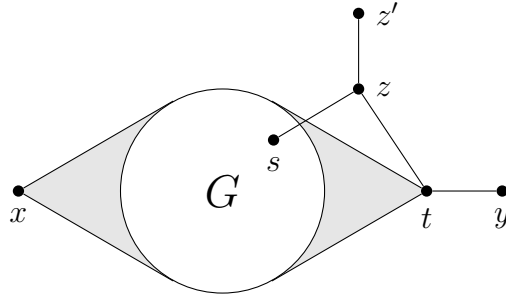
Figure 2.11: The graph above is the graph $J_4(G)$ as described in Theorem 2.6.4. Vertices x and t are adjacent to every vertex of G. Delete vertex z' to obtain the graph $J_3(G)$.

Note that the methods in Thereom 2.6.4 can be used to obtain a lower bound for graphs with centralizing number two, but Theorem 2.6.2 yields a stronger bound.

We conclude by making a conjecture about how the number of graphs on n vertices with each nonzero centralizing number compare to each other.

Conjecture 2.6.5. *Almost all graphs with nonzero centralizing number have central-izing number two, and almost all graphs with centralizing number greater than two have centralizing number three.*

Acknowledgements

Chapter 3

Hypergraphs[1]

3.1 Introduction

3.1.1 Realizability

A *hypergraph H* is an ordered pair of a vertex set and an edge set where edges are subsets of the vertices. A *simple* hypergraph contains no repeated edges, while a *multihypergraph* allows for repeated edges. In a *k-uniform* hypergraph, or *k*-graph, every edge contains exactly *k* edges. A 2-graph is also known as a graph. Unless otherwise stated, we assume all *k*-graphs are simple.

The *degree* of a vertex in a hypergraph *H* is the number of edges containing *v*, and the *degree sequence* of *H* is the list of degrees for each vertex of *H*. Degree sequences are typically written in nonincreasing order. A sequence of nonnegative integers π is *k-graphic* if there exists a hypergraph *H* that has π as its degree sequence. Furthermore, we say *H realizes* π.

Given a nonincreasing sequence π, it is a natural question to ask whether or not π is *k*-graphic.

Problem 3.1.1. *Find an efficient characterization for the k-graphic sequences.*

In the case of $k = 2$, Problem 3.1.1 is well understood. In fact, there are many characterizations of graphic sequences. Two of the most famous characterizations are the Havel-Hakimi [47, 44] and the Erdős-Gallai [33] characterizations. Sierskma and Hoogeveen [79] listed seven different characterizations and proved they are equivalent. More information on degree sequence and additional criterion were given by Mahadev and Peled [68]. Additionally, Kleitman and Wang [56] gave a characterization for directed graphs that is similar to the characterization of Havel and Hakimi. With this much information for graphic sequences, it is natural to extend the problem to *k*-graphic sequences. However, there are significantly fewer results for $k \geq 3$.

Dewdney [31] gave the only known characterization of k-graphic sequences. His characterization is similar to Havel and Hakimi's characterization, but it is inefficient as it requires testing the graphicality of many smaller sequences.

Bhave, Bam and Deshpande [10] also gave a characterization k-graphic sequences with respect to linear hypergraphs that is similar to the Erdős-Gallai characterization. Linear hypergraphs are hypergraphs where every pair of edges share at most one vertex. Others have worked towards Problem 3.1.1 by giving necessary conditions. Some of these necessary conditions are from Achuthan, Achuthan and Simanihuruk [1], Billington [13], and Choudum [25]. Unfortunately, Achuthan, Achuthan and Simanihuruk [1] showed that none of the necessary conditions from [1, 13, 25] are sufficient.

There are surprisingly few sufficient conditions for a sequence to be k-graphic for any $k \geq 3$. Billington [13] gave a heuristic for constructing a hypergraph from a 3-graphic degree sequence. Billington's algorithm is polynomial, but it does not always produce a hypergraph from a sequence that is 3-graphic. Behrens, Erbes, Ferrara, Hartke, Reininger, Spinoza, and Tomlinson [6] gave several sufficient conditions that are inspired by the results of Yin and Li [86], Aigner and Triesch [2] and Barrus, Hartke, Jao, and West [5] for 2-graphic sequences. Their sufficient conditions depend on the sum of the sequence and the maximum and minimum terms of the sequence.

Instead of providing necessary or sufficient conditions, Colbourn, Kocay, and Stinson [29] presented evidence that Problem 3.1.1 could be NP-complete for $k = 3$, showing that several related problems for degree sequences are in fact NP-complete.

3.1.2 Connectedness

As a related problem, there is interest in moving among realizations of a degree sequence by making small changes to the hypergraphs. A *t-switch* on a hypergraph H is replacing t edges with t other edges such that the resulting hypergraph has the same degree sequence as the original hypergraph. If a t-switch can be performed on a hypergraph, we say the hypergraph *admits* a t-switch. Given a k-graphic degree sequence π and a collection of switches S, we say the space of realizations is connected by S if for any two k-graph realizations of π, one can be obtained from the other by performing a sequence of switches from S. In other words, if the space of realizations is connected by S then any realization can be transformed into another realization by only using switches from S and any intermediate k-graph is also a realization of π (and only having simple k-graphs for intermediate graphs). Similarly, the space of multigraph realizations is connected by S if for any two multi-k-graph realizations of π, one can be obtained from the other by performing a sequence of switches from S.

Petersen [71] and subsequently others (for instance, Fulkerson, Hoffman, and McAndrew [39]) showed that for any two realizations of a 2-graphic sequence, one realization can be obtained from the other by a sequence of 2-switches. Kocay and Li [57] proved the space of multi-3-graph realizations is connected by the set of all possible 2-switches, and recently Behrens, Erbes, Ferrara, Hartke, Reininger, Spinoza, and Tomlinson [6] extended Kocay and Li's result to show the space of multi-k-graph realizations is connected by the set of all possible 2-switches for any k. Behrens et al. [6] also generalized an example of Gabelman [40] to show there exist sequences whose realization graphs for simple hypergraphs is not connected by using only t-switches for $t < k$. This result leads to the following problem.

Problem 3.1.2. *Determine a minimum family of switches such that for any π, the*

space of simple k-uniform realizations is connected by those switches.

In fact, it is not known if the set of all t-switches for $i \leq k$ is sufficient to connect the space.

Finding a minimal set of switches that connects the space of realizations has applications in network science; specifically, generating a realization chosen uniformly at random, which can be done using Markov chain Monte Carlo methods. The connectivity of the space for 2-graphs allows for uniform random sampling. For some results of using Markov chains for 2-graphs, see [35, 43, 52, 76]. Additionally, the Markov chain Monte Carlo methods can be used to give an approximation of the number of realizations for a degree sequence [51]. If we have set of switches connecting the space of 3-graph realizations, then we can use Markov chain Monte Carlo with the realizations of a 3-graphic sequence as well.

3.1.3   Results

We approach Problem 3.1.2 for the case of $k = 3$ from two different directions. We extend this classification in Section 3.3 to all 3-graphs that do not admit a 2-switch but differ by a 3-switch. Our investigation of slanted sequences is in Section 3.4, where we broaden the search to realizations that admit 2-switches. As an application of using switches, we provide a result about packing in Section 3.5. In further joint work [6], we provide sufficient conditions for a sequence to be k-graphic, show the space of multi-k-graph realizations is connected by 2-switches for any k, and generalize the example from Gabelman [40].

3.2   Classification of pseudo-threshold 3-graphs

One way of approaching Problem 3.1.2 for $k = 3$ is to first look for sequences whose space of realizations is not connected by using only 2-switches. Behrens et al. [6] generalize an example of Gabelman [40] by defining a matrix A

$$A = \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,k-1} & -y_1 \\ x_{2,1} & x_{2,2} & \cdots & x_{2,k-1} & -y_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ x_{k-1,1} & x_{k-1,2} & \cdots & x_{k-1,k-1} & -y_{k-1} \\ -z_1 & -z_2 & \cdots & -z_{k-1} & w \end{bmatrix}$$

where

$$y_j = \sum_{i=1}^{k-1} x_{j,i}, \qquad z_j = \sum_{i=1}^{k-1} x_{i,j}, \qquad \text{and} \qquad w = \sum_{i,j=0}^{k-1} x_{i,j}.$$

and each set of k entries sums to zero if and only if the set forms a row or column in A. One way to find such a matrix is to let the $x_{i,j}$'s be a power of some sufficiently small ϵ. They define a k-graph H on k^2 vertices by assigning each vertex the weight of a different entry of A. Any k-set is an edge of H if it has positive total weight or corresponds to a row in A. Since a t-switch takes t edges and replaces them with t non-edges, the switch takes t k-sets with nonnegative total weight to t k-sets with nonpositive total weight. As the degrees remain the same, the two sets of t edges have the same weight. Hence the total sum of vertex weights in each set of t edges is zero, and thus the t-switch takes k-sets with a total weight of zero and replaces them with k-sets that also have a total weight of zero. Since degrees of each vertex remain the same, this means the t-switch is the k-switch taking the k rows to the k columns.

We formally define this type of example as a pseudo-threshold hypergraph.

Definition 3.2.1. A $k \times k$ matrix M is a *pseudo-threshold matrix* if

1. The entries in any row or column sum to zero.

2. No other set of k entries of M sum to zero.

Definition 3.2.2. A k-uniform hypergraph H on k^2 vertices is *pseudo-threshold* if there exists a weight assignment to the vertices so that weights are the entries of a pseudo-threshold matrix M and the edges of H meet the following criteria:

1. Any k-set of vertices with positive sum forms an edge.

2. Any k-set of vertices with negative sum does not form an edge.

3. The rows of M correspond to edges of H and the columns are non-edges, or vice versa.

We say the matrix M is *associated* with H. Let $H_R(M)$ be the hypergraph with rows as edges and $H_C(M)$ be the hypergraph with columns as edges.

In particular, observe that every pseudo-threshold matrix is associated with two pseudo-threshold hypergraphs $H_R(M)$ and $H_C(M)$. However, every pseudo-threshold hypergraph has infinitely many pseudo-threshold matrices associated with it. For example, if M is associated with the pseudo-threshold hypergraph H, then $2M$ is also associated with H, where $2M$ is the matrix formed by multiplying every entry of M by two.

Lemma 3.2.3. *If M is a pseudo-threshold matrix, then neither $H_R(M)$ nor $H_C(M)$ admit a 2-switch. Furthermore, $H_R(M)$ and $H_C(M)$ differ by a 3-switch.*

We classify all pseudo-threshold k-graphs in the case $k = 3$. We provide a characterization of all pseudo-threshold 3-graphs by determining a complete list of matrices

M that generate distinct pseudo-threshold hypergraphs. Every pseudo-threshold 3-graph gives a degree sequence whose space of realizations is not connected by using only 2-switches. In Section 3.3 we discover this list of pseudo-threshold 3-graphs yields the complete list of hypergraphs that differ by a 3-switch but admit no 2-switch.

 We take the following steps to obtain the classification of pseudo-threshold 3-graphs.

1. Set a standard form for pseudo-threshold matrices.

2. Consider different cases of the standard form based on vertex weights.

3. Combine redundant cases.

4. For each remaining case, find the list of distinct weights of triples whose sign is not determined by the sign of the vertex weights.

5. For each list, check each possible sign assignment with a linear program to see if there is a pseudo-threshold matrix satisfying that assignment.

6. Eliminate matrices that generate isomorphic graphs.

 In order to set a standard form for pseudo-threshold matrices, we consider operations on matrices that preserve the property of being pseudo-threshold.

Proposition 3.2.4. *Let M be a pseudo-threshold matrix associated with a pseudo-threshold hypergraph H. Let M' be a matrix resulting from performing row-swaps and/or column-swaps on M and let M^T be the transpose of M. Then M' and M^T are also associated with H. Furthermore, $H_R(M) \simeq H_R(M')$ and $H_R(M) \simeq H_C(M^T)$.*

Proof. Since M, M', and M^T have the same set of entries, each matrix has exactly $2k$ k-sets that sum to zero. Thus in order to show M' and M^T are associated with H,

then it is enough to check that the rows and columns of M' and M^T sum to zero. Let $(a_1, a_2, \ldots, a_k)$ be a row of M'. Observe that row-swaps change the order of the rows, but the contents of the individual rows. Also, column-swaps change the order of the contents of the rows, but does not change the actual content. Hence $a_1, a_2, \ldots, a_k$ are the entries of some row of M, so their sum is zero. Similarly, if $(a_1, a_2, \ldots, a_k)$ is a column of M', then $a_1, a_2, \ldots, a_k$ are the entries of some column of M and have a sum of zero. Since the rows and columns of M^T are the columns and rows of M, we also have that the rows and columns of M^T sum to zero.

To show the isomorphisms, let e be an edge of $H_R(M)$. Then e corresponds to a k-set $(a_1, a_2, \ldots, a_k)$ in M that is either a row or has positive weight. If the k-set has positive weight, follow the transformation from M to M' and it corresponds to an edge in $H_R(M')$. If the k-set is a row, then as argued above it is also a row of M', so it also corresponds to an edge in $H_R(M)$. Similarly, non-edges of $H_R(M)$ are mapped non-edges of $H_R(M')$ under the transformation given by changing M into M'. Thus $H_R(M) \simeq H_R(M')$. A similar argument applies for $H_R(M) \simeq H_C(M^T)$, with the exception that rows get mapped to columns and vice versa. $\qquad \square$

Step 1: Set a standard form for pseudo-threshold matrices.

As we seek a classification of pseudo-threshold 3×3 matrices that produce different pseudo-threshold hypergraphs, the first step towards this goal is to obtain a standard form for the matrices. By Proposition 3.2.4, we can use row-swaps, column swaps, and transposition to assume that any pseudo-threshold 3x3 matrix M has the form

$$M = \begin{pmatrix} a & b & -(a+b) \\ c & d & -(c+d) \\ -(a+c) & -(b+d) & a+b+c+d \end{pmatrix}$$

with $|a| \leq |b| \leq |c|$ and $|a| \leq |d|$. Note this is an assumption on the 2×2 sub-matrix in the upper left corner of M. The entries in the third row and column are a consequence of the requirement that rows and columns sum to zero. Observe that no two entries can be equal since the only 3-sets with total weight zero are the rows and columns. For example, if $a = b$, then the triple $(b, c, -(a + c))$ has weight $b + c - a - c = a + c - a - c = 0$, but the triple is neither a row nor column of M.

Step 2: Consider different cases of the standard form based on vertex weights.

To continue the classification of pseudo-threshold matrices, we break the format of the matrix into cases based on the signs of the vertex weights. If we look at the sign of the vertex weights, we can sometimes determine if a triple is an edge or nonedge without knowing the exact value of the weights. Notice that any three nonnegative weighted vertices must form an edge, while any three nonpositive weighted vertices cannot form an edge. However, if we have a mixture of positive and negative weights, then we must know the exact weights to determine whether or not those vertices form an edge. Hence, in order to determine all pseudo-threshold 3-graphs, we can focus on determining which combinations of positive weight mixed sign triples can occur.

Recall we assume $|a| \leq |b| \leq |c|$ and $|a| \leq |d|$. If we also place assumptions on whether each of $a, b, c,$ and d are nonnegative $(+)$ or nonpositive $(-)$, we obtain the sixteen cases shown in Figure 3.1. Observe that the sign of the remaining five entries of the matrix are frequently determined by the combination of assumptions on the signs and relative magnitudes of a, b, c and d. For example, if $a > 0$ and $d < 0$ then $a + d > 0$ and hence the entry $-(a + d)$ must be positive. However there are times when the sign of the weight of a vertex is not determined. Entries whose signs are not determined are represented by a question mark.

$$\begin{pmatrix} + & + & - \\ + & + & - \\ - & - & + \end{pmatrix}$$
(a)

$$\begin{pmatrix} - & + & - \\ + & + & - \\ - & - & + \end{pmatrix}$$
(b)

$$\begin{pmatrix} + & - & + \\ + & + & - \\ - & ? & + \end{pmatrix}$$
(c)

$$\begin{pmatrix} + & + & - \\ - & + & ? \\ + & - & ? \end{pmatrix}$$
(d)

$$\begin{pmatrix} + & + & - \\ + & - & ? \\ - & ? & ? \end{pmatrix}$$
(e)

$$\begin{pmatrix} - & - & + \\ + & + & - \\ - & ? & + \end{pmatrix}$$
(f)

$$\begin{pmatrix} - & + & - \\ - & + & ? \\ + & - & ? \end{pmatrix}$$
(g)

$$\begin{pmatrix} - & + & - \\ + & - & ? \\ - & ? & ? \end{pmatrix}$$
(h)

$$\begin{pmatrix} + & - & + \\ - & + & ? \\ + & ? & ? \end{pmatrix}$$
(i)

$$\begin{pmatrix} + & - & + \\ + & - & ? \\ - & + & ? \end{pmatrix}$$
(j)

$$\begin{pmatrix} + & + & - \\ - & - & + \\ + & ? & - \end{pmatrix}$$
(k)

$$\begin{pmatrix} + & - & + \\ - & - & + \\ + & + & - \end{pmatrix}$$
(l)

$$\begin{pmatrix} - & + & - \\ - & - & + \\ + & ? & - \end{pmatrix}$$
(m)

$$\begin{pmatrix} - & - & + \\ + & - & ? \\ - & + & ? \end{pmatrix}$$
(n)

$$\begin{pmatrix} - & - & + \\ - & + & ? \\ + & ? & ? \end{pmatrix}$$
(o)

$$\begin{pmatrix} - & - & + \\ - & - & + \\ + & + & - \end{pmatrix}$$
(p)

Figure 3.1: Assume $|a| \le |b| \le |c|$ and $|a| \le |d|$. Given the sign (nonnegative or nonpositive) of the values $a, b, c,$ and d, the signs of most of the entries in the third row and third column are also determined. The question mark represents entries with nondetermined signs.

Step 3: Combine redundant cases.

Figure 3.1 gives a total of $2^4 = 16$ cases with additional subcases for each undetermined entry. However, these cases are not unique; in fact, they are extremely redundant. As an example, let $a = 1$, $b = 2$, $c = 3$, and $d = 4$, which yields the following pseudo-threshold matrix.

$$\begin{pmatrix} 1 & 2 & -3 \\ 3 & 4 & -7 \\ -4 & -6 & 10 \end{pmatrix}$$

This matrix is in the form of Figure 3.1a. Applying row-swaps, column-swaps, and transpositions, we also obtain the following matrices.

$$
\begin{pmatrix} 4 & -6 & 2 \\ -7 & 10 & -3 \\ 3 & -4 & 1 \end{pmatrix}
\qquad
\begin{pmatrix} 3 & 4 & -7 \\ -4 & -6 & 10 \\ 1 & 2 & -3 \end{pmatrix}
\qquad
\begin{pmatrix} 2 & -3 & 1 \\ 4 & -7 & 3 \\ -6 & 10 & 4 \end{pmatrix}
$$

These three matrices match the form of Figures 3.1i, 3.1k and 3.1j, respectively. Thus one pseudo-threshold matrix can be written in the form of at least four of the cases listed in Figure 3.1. In order to combine redundant cases, we apply Proposition 3.2.4 to show that every case is equivalent to either Figure 3.1h or 3.1i

Proposition 3.2.5. *Any pseudo-threshold 3-graph is associated with a pseudo-threshold matrix M with the form of Figure 3.1i or Figure 3.1i.*

Proof. Observe that every matrix in Figure 3.1 except for 3.1e and 3.1o contain $\left(\begin{smallmatrix} + & - \\ - & + \end{smallmatrix}\right)$ or $\left(\begin{smallmatrix} - & + \\ + & - \end{smallmatrix}\right)$ as a submatrix. In the matrix of Figure 3.1e, if the question mark in the second row or column was nonnegative, we would have a desired submatrix. If both are nonpositive, then the remaining question mark must be nonnegative, which also obtains a desired submatrix. The case for the matrix in Figure 3.1o follows symmetrically. Thus any pseudo-threshold matrix contains $\left(\begin{smallmatrix} + & - \\ - & + \end{smallmatrix}\right)$ or $\left(\begin{smallmatrix} - & + \\ + & - \end{smallmatrix}\right)$ as a submatrix.

Recall that we assume a pseudo-threshold matrix has the standard form

$$
\begin{pmatrix} a & b & -(a+b) \\ c & d & -(c+d) \\ -(a+c) & -(b+d) & a+b+c+d \end{pmatrix}
$$

with $|a| \leq |b| \leq |c|$ and $|a| \leq |d|$ and we achieved this by using row-swaps, column-

swaps, and transpositions to place the four entries in the desired order. Let M be a pseudo-threshold matrix, and let M' be matrix obtained from M by applying row-swaps and column swaps to put a $\left(\begin{smallmatrix} + & - \\ - & + \end{smallmatrix}\right)$ or $\left(\begin{smallmatrix} - & + \\ + & - \end{smallmatrix}\right)$ submatrix in the upper left corner. By Proposition 3.2.4, M' is a pseudo-threshold matrix and M and M' are associated with the same pair of pseudo-threshold hypergraphs. Apply row-swaps, column-swaps, and transpositions to M' to put the four entries of the upper left corner in the desired order and let M'' be the resulting matrix. Observe the upper left corner of M'' still has either the $\left(\begin{smallmatrix} + & - \\ - & + \end{smallmatrix}\right)$ or $\left(\begin{smallmatrix} - & + \\ + & - \end{smallmatrix}\right)$ pattern. Note the entries in the third row and column of the matrices in Figure 3.1 are a direct result of the assumption made on the four entries in the upper right corner. Hence M'' has the form of either Figure 3.1i or Figure 3.1i and is associated with the same pseudo-threshold hypergraphs as M. $\square$

Corollary 3.2.6. *Any pseudo-threshold hypergraph H is associated with a pseudo-threshold matrix M where M or $-M$ has the form of Figure 3.1i.*

Proof. If H is not associated with a matrix in the form of Figure 3.1i, then by Proposition 3.2.5 it must be associated with a matrix M in the form of Figure 3.1h. Observe Figures 3.1h and 3.1i share nondetermined entries a every determined entry in Figure 3.1h is the negative of the corresponding entry in Figure 3.1i. Thus $-M$ (the matrix formed by taking the negative entries of M) is in the form of Figure 3.1i. $\square$

Since we have limited the cases to only one of the matrices from Figure 3.1, we determine the possible subcases of Figure 3.1i. If we assume that the question mark in the second column of Figure 3.1i is nonnegative, then we have $|d| \leq |b| \leq |c|$, so the question mark in the second row is also nonnegative. In order for the third row or column to have a sum of zero the question mark in the third row and third column must be nonpositive. This yields the sign matrix of Figure 3.2a. In order for the third column to have a total weight of zero, at least one of the two question marks in the

third row must be nonpositive. Thus if the question mark in the second column is nonpositive, then the sign matrix must be one of Figure 3.2b, 3.2c, or 3.2d.

$$
\begin{pmatrix} + & - & + \\ - & + & + \\ + & + & - \end{pmatrix} \qquad \begin{pmatrix} + & - & + \\ - & + & + \\ + & - & - \end{pmatrix} \qquad \begin{pmatrix} + & - & + \\ - & + & - \\ + & - & + \end{pmatrix} \qquad \begin{pmatrix} + & - & + \\ - & + & - \\ + & - & - \end{pmatrix}
$$
$$
\text{(a)} \qquad\qquad\qquad \text{(b)} \qquad\qquad\qquad \text{(c)} \qquad\qquad\qquad \text{(d)}
$$

Figure 3.2: The four ways to fill out the question marks from Figure 3.1i.

Step 4: For each remaining case, find the list of distinct weights of triples whose sign is not determined by the sign of the vertex weights.

By Corollary 3.2.6, every pseudo-threshold 3-graph can be represented by a matrix M or $-M$ with the form of one of the four matrices of Figure 3.2. Observe the complement of $H_R(M)$ is $H_C(-M)$. In other words, the list of all pseudo-threshold matrices M that are in the form of Figure 3.1i and that generate distinct hypergraphs along with $-M$ will generate the complete list of pseudo-threshold 3-graphs. For each of the four matrices in Figure 3.2, we seek the distinct pseudo-threshold 3-graphs with a matrix of that form. In order to accomplish this, we now focus on the sign of the weight of each triple, which then allows us to focus on the possible edge combinations. For example, in every matrix with the form of Figure 3.2a, $a + d - (c + d) > 0$ as each of a, d, and $-(c + d)$ are nonnegative, so the associated hypergraph must have the edge $(a, d, -(c + d))$. Similarly, $(b, c, a + b + c + d)$ is a nonedge. However (a, b, c) may or may not be an edge, based on the actual values of a, b, and c. For all triples, we use SageMath [81] version 5.8 to determine whether or not the sign of the total weight is known based on the assumptions about the sign of a, b, c, and d as well of the relationships between the absolute values of a, b, c, and d. The SageMath code is given in Algorithm 2.

Any triple whose weight is not determined by those assumptions is recorded. Additionally we notice relationships among the list of undetermined triples. In particular, we found triples whose weight was the negative of the weight of another triple on the list. For example, in Figure 3.2a, exactly one of $(a, b, -(c + d))$ and $(-(a + b), c, d)$ is an edge since their weights are $a + b - c - d$ and $-(a + b - c - d)$. Hence only one of the two triples needs to be listed as the edge status of the second triple is determined by the first triple.

Algorithm 2 Finds triples whose weight is not determined for Case 3.2a

a, b, c, d=var('a, b, c, d')
assume($c < b, b < 0, 0 < a, a < d, a < -b$)
assume($a + c < 0, b + d < 0, c + d < 0, a + b + c + d < 0$)
pos=$[a, d, -(a + b), -(c + d), -(a + c), -(b + d)]$
neg=$[b, c, a + b + c + d]$
for $i, j \in$ pos, $k \in$ neg, $i \neq j$ **do**
    wt=$i + j + k$
    if not bool(wt> 0) and not bool(wt< 0) and not bool(wt$== 0$): **then**
        print (i, j, k)
for $i \in$ pos, $j, k \in$ neg, $j \neq k$ **do**
    wt=$i + j + k$
    if not bool(wt> 0) and not bool(wt< 0) and not bool(wt$== 0$): **then**
        print (i, j, k)

For each matrix in Figure 3.2, the weights for each undetermined triple is listed in Table 3.1.

Step 5: For each list, check each possible sign assignment with a linear program to see if there is a pseudo-threshold matrix satisfying that assignment.

At this point we have the list of weights of triples whose sign is not determined. In order to obtain a complete list of pseudo-threshold 3-graphs, we take each list of nondetermined weights and check each possible sign assignment to the weights of the

Case 3.2a	Case 3.2b	Case 3.2c	Case 3.2d
$a + b + d$	$a + c + d$	$-b - c - d$	$-2b - d$
$a + c + d$	$2a + b + c + 2d$	$-2b - d$	$-2a - c$
$2a + b + c + 2d$	$-2b - d$	$-2c - d$	$-2a - b$
$a + b - c - d$	$a + b - c - d$	$2a + 2b + c + d$	$-a + b - 2c - d$
$2a + b$	$2a + b$	$2a + b + 2c + d$	$a + 2b + c + 2d$
$2a + c$	$a - b - c - 2d$	$2a + b$	$a + b + 2c + 2d$
$c + 2d$	$c + 2d$	$2a + c$	$a - b - c - 2d$
$b + 2d$	$-a - 2b - c - 2d$	$a + 2b + 2c + d$	
$-a - 2b + c - d$	$-2a - c$	$-a + b - 2c - d$	
$-a + b - 2c - d$	$-a + b - 2c - d$	$a - b - c - 2d$	

Table 3.1: List of weights of triples whose sign is not determined.

triples to check if there exists a matrix satisfying the sign assumptions. For example, the Linear Program 3 with $B = 1111000000$ checks if there exists a way of assigning values to a, b, c, and d satisfying the assumptions on a, b, c, and d such that the first four sums in the first column of Table 3.1 are positive and the remaining sums are negative.

Linear Program 3 Finding possible vertex weights for Case 3.2a

Solve: $\quad a, b, c, d$

Subject to: $\quad a \geq 0$
$\qquad\qquad b \leq 0$
$\qquad\qquad c < b$
$\qquad\qquad d > a$
$\qquad\qquad b + d \leq 0$
$\qquad\qquad a, b, c, d \in \mathbb{R}$

Additional constraints:

    Input: $B \in \{0, 1\}^{10}$
      for $i \in \{1, 2, \ldots, 10\}$ **do**
         Let y be the ith entry of Case 3.2a in Table 3.1
         if $B[i] = 1$ **then**
            $y > 0$
         else
            $y < 0$

The matrices generated in this manner gives us a partially complete list of pseudo-threshold hypergraphs. As mentioned earlier, we ignored the matrix of Figure 3.1h. Hence for each matrix M on this list, we must also add the matrix $-M$ to the list. Now we have a list of matrices representing every pseudo-threshold hypergraph. Note this list produces isomorphic pseudo-threshold hypergraphs. One way this arises is that $\left(\begin{smallmatrix} + & - \\ - & + \end{smallmatrix}\right)$ or $\left(\begin{smallmatrix} - & + \\ + & - \end{smallmatrix}\right)$ appears multiple times in the matrices of Figure 3.2.

Step 6: Eliminate matrices that generate isomorphic graphs.

For each hypergraph from this list, no two vertices have the same degree. Hence we check for isomorphic hypergraphs by ordering the vertices by degree and searching for identical edge assignments. After removing isomorphic copies, a list of matrices representing every pseudo-threshold hypergraph appears in Table 3.2.

41 edges				42 edges				42 edges (ctd)				43 edges			
a	b	c	d	a	b	c	d	a	b	c	d	a	b	c	d
0	4	6	-1	0	1	7	-3	0	-1	-7	3	0	-4	-6	1
0	5	6	-2	0	2	9	-5	0	-2	-9	5	0	-5	-6	2
0	6	7	-4	0	3	6	-1	0	-3	-6	1	0	-6	-7	4
0	7	9	-4	0	3	6	-2	0	-3	-6	2	0	-7	-9	4
-2	3	11	-7	0	5	7	-4	0	-5	-7	4	2	-3	-11	7
-3	4	19	-9	0	6	9	-4	0	-6	-9	4	3	-4	-19	9
-4	5	14	-11	-2	3	12	-9	2	-3	-12	9	4	-5	-14	11
-4	6	11	-10	-2	3	16	-7	2	-3	-16	7	4	-6	-11	10

Table 3.2: The complete list of non-isomorphic pseudo-threshold 3-graphs, each represented by the matrix with the given weights.

Theorem 3.2.7. *The complete list of pseudo-threshold 3-graphs is given by Table 3.2.*

Observe all of the pseudo-threshold 3-graphs have between 41 and 43 edges, which is within one of half of the number of possible triples. Additionally, none of these 3-graphs are self-complementary.

3.3 Hypergraphs with no 2-switch

Recall our focus is on Problem 3.1.2. In this section, we focus on hypergraphs that do not admit a 2-switch and seek switches necessary to connect the space of realizations.

If two hypergraphs differ by a 3-switch, it does not imply that the two hypergraphs are in separate components of the space of realizations via 2-switches. For example, Figure 3.3 shows how to achieve a 3-switch by using only 2-switches. However, if the hypergraph already contains any of the intermediate edges of Figure 3.3, then this sequence cannot be used to obtain the 3-switch by 2-switches. To ensure the 3-switch is required, we focus on 3-graphs that do not admit a 2-switch. This guarantees that there is no way to obtain a t-switch for $t \geq 3$ by using 2-switches. For example, if (x, y, z) and (u, v, w) are edges in a 3-graph H that does not admit 2-switches, then either (x, v, w) or (u, y, z) must also be in $E(H)$. Otherwise H would admit the 2-switch from (x, y, z) and (u, v, w) to (x, v, w) and (u, y, z). More specifically, for any two disjoint edges of H and any way of splitting those six vertices into two triples, at least one of the triples must be in $E(H)$. Additionally, for any two edges of H that share one vertex z and any way of splitting the remaining four vertices into pairs (x, y) and (u, v), at least one of (x, y, z) and (u, v, z) must be in $E(H)$.

It may be possible that there exists a pair of hypergraphs on the same degree sequence that are in separate components of the space of realizations over 2-switches where at least one of the hypergraphs admits a 2-switch. However, the only currently known examples where the space of realizations is disconnected is with a pair of hypergraphs where neither admits a 2-switch.

Now if H_1 and H_2 are two 3-graphs that differ by a 3-switch but neither 3-graph admits a 2-switch, we have that H_1 and H_2 are isolates in the space of realizations over 2-switches. However, they are connected in the space of realizations over 2-switches
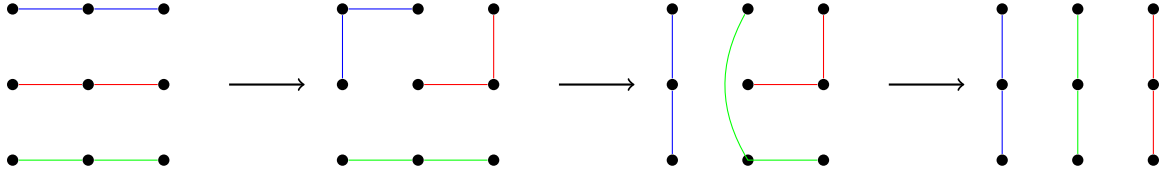
Figure 3.3: An example of how a 3-switch can be accomplished by using 2-switches.

and 3-switches. Hence we say H_1 and H_2 are an isolated pair.

Definition 3.3.1. An *isolated pair* is a pair of 3-graphs H_1 and H_2 such that neither hypergraph admits a 2-switch and the two hypergraphs have the same degree sequence.

For example, every matrix M represented in Table 3.2 yields the isolated pair $H_R(M)$ and $H_C(M)$. We prove that every isolated pair that differ by a 3-switch contains a 3-graph from Table 3.2 as an induced subhypergraph.

For any isolated pair, the two 3-graphs must differ by a t-switch for some $t \geq 3$. While some 3-switches involve nine distinct vertices, such as the 3-switch in Figure 3.3, some 3-switches only involve eight or fewer vertices. For example, the 3-switch exhibited in Figure 3.4 involves six distinct vertices. We show that any isolated pair is on at least nine vertices.

Figure 3.4: An example of a 3-switch on 6 vertices.

Lemma 3.3.2. *There are no isolated pairs on eight or fewer vertices.*

Proof. To determine that no isolated pairs on eight or fewer vertices exist, we solve Integer Program 4.

Integer Program 4 Isolated pairs on 8 vertices

Define:	$V = \{0, 1, \ldots, 7\}$
	$T = \{\{i, j, k\} : i, j, k \in V, |\{i, j, k\}| = 3\}$
Solve:	$\{x_t\}_{t \in T} \; \{y_t\}_{t \in T}$
Subject to:	1. $x_{\{0,1,2\}} = 1, y_{\{0,1,2\}} = 0$

$$\text{2. } x_s + x_t - (x_a + x_b) \leq 1 \qquad \forall s, t, a, b \in T : |s \cap t| \in \{0, 1\},$$
$$a \cap b = s \cap t, a \cup b = s \cup t$$

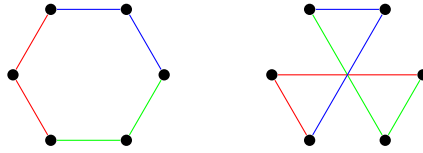$$\text{3. } y_s + y_t - (y_a + y_b) \leq 1 \qquad \forall s, t, a, b \in T : |s \cap t| \in \{0, 1\},$$
$$a \cap b = s \cap t, a \cup b = s \cup t$$

$$\text{4. } \sum_{t \ni i} (x_t - y_t) = 0 \qquad \forall i \in V$$

$$\text{5. } x_t, y_t \in \{0, 1\} \qquad \forall t \in T$$

The two sets of variables, $\{x_t\}_{t \in T}$ and $\{y_t\}_{t \in T}$ correspond to triples of vertices in the hypergraphs H_1 and H_2. In particular, $x_t = 1$ if and only if $t \in E(H_1)$, and $y_t = 1$ if and only if $t \in E(H_2)$. Setting $x_{\{0,1,2\}} = 1$ and $y_{\{0,1,2\}} = 0$ guarantees that H_1 and H_2 are different hypergraphs.

Constraints 2 and 3 of Integer Program 4 enforce that neither hypergraph admits a 2-switch. Suppose that $s, t \in E(H_1)$ share at most one vertex, and that a and b are triples with $a \cap b = s \cap t$ and $a \cup b = s \cup t$. Then at least one of a or b is an element of $E(H_1)$, else H_1 would admit a 2-switch. Observe $x_s + x_t - (x_a + x_b) > 1$ if and only if $s, t \in E(H_1)$ and $a, b \notin E(H_1)$. Thus the constraint $x_s + x_t - (x_a + x_b) \leq 1$ prevents H_1 from having a 2-switch from s, t to a, b. Note that no 2-switch is possible involving s and t if s and t have more than one vertex in common.

In order to ensure the two hypergraphs are a isolated pair, they must have the same degree sequence and neither hypergraph can admit a 2-switch. Note that $\sum_{t \ni i} x_t$ is the degree of vertex i in H_1, so constraint 4 enforces that H_1 and H_2 have the same degree sequence.

We solved Integer Program 4 using CPLEX [49] and GLPK [41] on a linux-based computer powered by a quad-core Intel CPU running at 2.4GHz, and it finished in

two hours. The integer program revealed there is no solution to the given parameters, so there does not exist an isolated pair on eight vertices. If there was an isolated pair with fewer vertices, then the integer program would have found a solution where both hypergraphs have an isolated or dominating vertex. Thus every isolated pair has at least nine vertices. $\qquad\square$

Proposition 3.3.3. *If H_1 and H_2 are an isolated pair on nine vertices, then there exists a matrix M with entries given in Table 3.2 so that $H_1 \simeq H_R(M)$ and $H_2 \simeq H_C(M)$.*

Proof. We run Integer Program 5 to determine the existence of isolated pairs that are not pseudo-threshold hypergraphs.

Integer Program 5 Isolated pairs on 9 vertices

Define:	$V = \{0, 1, \ldots, 8\}$
	$T = \big\{\{i, j, k\} : i, j, k \in V, |\{i, j, k\}| = 3\big\}$
Solve:	$\{x_t\}_{t \in T}\ \{y_t\}_{t \in T}$

Subject to:

1. $x_{\{0,1,2\}} = 1, y_{\{0,1,2\}} = 0$
2. $x_s + x_t - (x_a + x_b) \leq 1$ $\forall s, t, a, b \in T : |s \cap t| \in \{0, 1\}, \ a \cap b = s \cap t, a \cup b = s \cup t$
3. $y_s + y_t - (y_a + y_b) \leq 1$ $\forall s, t, a, b \in T : |s \cap t| \in \{0, 1\}, \ a \cap b = s \cap t, a \cup b = s \cup t$
4. $\displaystyle\sum_{t \ni i}(x_t - y_t) = 0$ $\forall i \in V$
5. $\displaystyle\sum_{t \ni i} x_t - \sum_{t \ni i+1} x_t \geq 0$ $\forall i \in V \setminus \{2, 8\}$
6. $\displaystyle\sum_{t \ni 0} x_t - \sum_{t \ni 3} x_t \geq 0$
7. $\displaystyle\sum_{t \in T} x_t = m$
8. $\displaystyle\sum_{t \in E(H)} x_t < m$ $\forall$ pseudo-threshold hypergraphs H on m edges
9. $x_t, y_t \in \{0, 1\}$ $\forall t \in T$

Integer Program 5 is made from Integer Program 4 by adding constraints 5-8 and increasing the number of vertices to nine. To further break symmetries, we place an order on the vertex degrees. Lemma 3.3.2 implies that all nine vertices are involved in the t-switch. Hence any maximum degree vertex must appear in some edge in

the first hypergraph that is not an edge in the second hypergraph. Thus we assume vertex 0 is a vertex with maximum degree. Additionally, we may assume that vertex 1 has higher degree than vertex 2, and vertex i has higher degree than vertex $i + 1$ for $i \in \{3, 4, 5, 6, 7\}$. This is reflected in constraints 5 and 6.

We split the search into cases by total number of edges m, as set in constraint 7. First, we set it to look for isolated pairs with at most 40 edges, and again with at least 44 edges (changing constraint 7 to an inequality). These searches yield no examples. Next, we examine the three cases of 41, 42, and 43 edges. For any of the pseudo-threshold matrices, there are exactly two ways of placing an associated pseudo-threshold 3-graph onto vertices following the assumptions on vertex degrees (one for the row and one for the column). Hence in our search for a fixed number of edges, we eliminate the previously found pseudo-threshold 3-graph examples by adding the constraint that the sum of the triples forming edges in those examples is less than the total number of edges. This is reflected in constraint 8.

The integer program returns that there is no integer solution to the given parameters. Hence there are no additional isolated pairs on nine vertices. The integer program for each case by number of edges was run on a single computer using GLPK [41] and CPLEX [49]. The GLPK version finished in four hours for each case, while CPLEX finished in about twenty minutes for each case.

$\square$

Note that we have been focusing on 3-graphs on nine vertices, but Definition 3.3.1 allows for more vertices. In fact, it is possible that an isolated pair H_1 and H_2 are hypergraphs on more than nine vertices. For example, for some pseudo-threshold 3×3 matrix M, let H_1 be the hypergraph defined with vertex set $V(H_R(M)) \cup \{u, v\}$ and edge set $E(H_R(M)) \cup \{(u, v, x) : x \in V(H_R(M))\}$ and H_2 on the same vertices with

edge set $E(H_C(M)) \cup \{(u,v,x) : x \in V(H_C(M))\}$. Then H_1 and H_2 are an isolated pair. Hence the statement of Theorem 3.3.4 only requires the pseudo-threshold 3-graphs to be induced.

Theorem 3.3.4. *If H_1 and H_2 are an isolated pair that differ by a 3-switch, then H_1 and H_2 each contain as an induced subhypergraph a pseudo-threshold 3-graph on the 9 vertices involved in the 3-switch.*

Proof. Suppose H_1 and H_2 are an isolated pair on at least ten vertices. By Lemma 3.3.2, H_1 and H_2 must differ on at least nine vertices. Thus the 3-switch must occur on exactly nine distinct vertices as any 3-switch involves at most nine vertices. Let G_1 and G_2 be the induced subhypergraphs of H_1 and H_2 on those nine vertices. Then neither G_1 nor G_2 admits a 2-switch, else H_1 or H_2 would admit a 2-switch. Hence G_1 and G_2 are a isolated pair on nine vertices. In particular, G_1 and G_2 are pseudo-threshold 3-graphs. □

Suppose a 3-switch replaces the edges $\{e_1, e_2, e_3\}$ with the edges $\{e_4, e_5, e_6\}$. The 3-switch is a *transversal 3-switch* if whenever vertices u and v are in the same edge from $\{e_1, e_2, e_3\}$ then they are in different edges of $\{e_4, e_5, e_6\}$ and vice versa. An example of a transversal 3-switch is given in Figure 3.5. Since $H_R(M)$ and $H_C(M)$ differ by a transversal 3-switch for any pseudo-threshold matrix M, we obtain the following corollary.

Corollary 3.3.5. *Suppose H_1 and H_2 are an isolated pair that differ by a 3-switch. Then the 3-switch connecting them must be a transversal switch.*

Since Proposition 3.3.3 did not depend on a 3-switch, the proof of Theorem 3.3.4 also yields the following result.
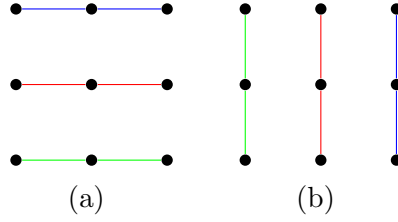
Figure 3.5: A transversal 3-switch.

Corollary 3.3.6. *If H_1 and H_2 are an isolated pair that differ on at most nine vertices, then H_1 and H_2 differ by a 3-switch.*

3.4   Slanted sequences

In the previous section, we found specific examples of spaces of realizations that are not connected by 2-switches. However, each of those examples became connected if we added the transversal 3-switch. In this section, we continue the search for additional examples of when the space of realizations is not connected by 2-switches. There are many sequences of length nine, so we need to narrow down this list.

There is a dominance order on sequences with a fixed sum. In this poset, Aigner and Triesch [2] showed that 2-graphic sequences form a downset, while Billington [12] shows that k-graphic sequences also form a downset. Berger [9] showed that if π' is below π, then π' has at least as many 2-graphic realizations as π. Barrus, Hartke, Jao, and West [5] proved that sequences with gaps, and in particular sequences without repeated values, are also less likely to be 2-graphic. Using these results about 2-graphs suggest we should examine sequences that do not have repeated terms.

Theorem 3.3.4 gives examples of sequences whose space of realizations is not connected by 2-switches and does not have repeated terms. Similar sequences seem like the most likely candidates for degree sequences whose realization space is not connected by 2-switches, as fewer realizations have less room to move. Additionally, if

two vertices have the same degree and there exist realizations where switching the vertices creates a distinct realization, then the number of realizations is increased with more isomorphic copies.

We say a k-graphic sequence is *slanted* if it does not have any repeated values. We used a Python program to generate the list of slanted sequences in colexicographic order. Each sequence was then tested with the necessary conditions of Achuthan, Achuthan and Simanihuruk [1], Billington [13], and Choudum [25] so it could be quickly eliminated if it was not 3-graphic. For each remaining slanted sequence π, we used the libexact library of Kaski and Pottonen [53] to generate the 3-graph realizations of π. Libexact uses the Knuth's dancing links technique for updating the data structure to solve combinatorial exact covering problems. If libexact finds over 50,000 realizations, we move on to the next sequence. For each sequence with at most 50,000 realizations, we use a breadth-first search to check that the realizations are connected by a 2-switch. We performed the computations in parallel on the Open Science Grid [74]. This search yields the following result.

Theorem 3.4.1. *Every 3-graphic slanted sequence of length at most 8 with at most 50,000 realizations is connected by the set of 2-switches. Every 3-graphic slanted sequence of length 9 with at most 50,000 realizations is either connected by the set of 2-switches and the transversal 3-switch.*

In particular, there were exactly 45 sequences that required a transversal 3-switch in addition to the 32 sequences given by pseudo-threshold 3-graphs. Additionally, there exists a pseudo-threshold 3-graphs with sequence π if and only if π is a slanted sequence with a realization space consisting of two isolates. This theorem provides more support that the collection of 2-switches along with the transversal 3-switch is enough to connect the space of realizations for 3-graphic sequences.

Conjecture 3.4.2. *The space of realizations for 3-graphic sequences is connected by the set of 2-switches and the transversal 3-switch.*

3.5 Applications

In this section we use 2-switches to gain further insight with k-graphic sequences. We introduce notation to refer to a specific 2-switch. Let e and e' be distinct edges in a k-graph H, and choose vertices $u \in e \setminus e'$ and $v \in e' \setminus e$. The operation $e \overset{u}{\underset{v}{\rightleftharpoons}} e'$ deletes the edges e and e' and adds the edges $e - u + v$ and $e' - v + u$ (where $e - u + v$ denotes the set $e - \{u\} \cup \{v\}$).

3.5.1 Obtaining a "good" realization

One consequence of the Havel-Hakimi characterization of 2-graphic sequences is that any graphic sequence has a realization in which a specified vertex v is adjacent to vertices whose degrees are the highest-degree vertices in the graph. This elementary fact has been proved in several places, for instance [39]. Motivated by this, we prove the following.

Theorem 3.5.1. *Let $\pi = (d_1, \ldots, d_n)$ be a nonincreasing k-graphic sequence, and let H be a k-realization of π on vertices $\{v_1, \ldots, v_n\}$ such that $d(v_i) = d_i$ for each $i, 1 \leq i \leq n$. Let $i < j$ and suppose there is an edge e in H such that v_j is in e but v_i is not in e. Then there is a realization H' of π such that $e - v_j + v_i$ is an edge in H'.*

Proof. If $e - v_j + v_i$ is already an edge in H, we are done. So we can assume this edge does not exist. Since $d_i \geq d_j$, there is an edge f such that v_i is in f but v_j is not. Additionally, some such f has the property that $f - v_i + v_j$ is not an edge in H.

Perform the exchange $e \overset{v_j}{\underset{v_i}{\rightleftharpoons}} f$. This does not create any multi-edges, so we have the desired realization. □

An immediate corollary of this result is that for any vertex v of positive degree, there is a k-realization of π such that v is in an edge with the $k-1$ remaining vertices of highest degree. Thus, there is always a realization of π in which the k vertices of highest degree are in a single edge. If we could prove the existence of a k-realization in which the link of a vertex contains only the highest degree vertices, then we would be able to obtain a Havel-Hakimi-type characterization of k-graphic sequences.

3.5.2   Packing k-graphic sequences

Two n-vertex graphs G_1 and G_2 *pack* if they can be expressed as edge-disjoint subgraphs of the complete graph K_n. Kostochka, Stocker, and Hamburger [59], and Pilśniak and Woźniak [72, 73] recently studied packing of hypergraphs. Busch et al. [19] extended the idea of graph packing to graphic sequences. We utilize edge exchanges to examine related questions for hypergraphic sequences.

Let π_1 and π_2 be k-graphic sequences with $\pi_1 = (d_1^{(1)}, \ldots, d_n^{(1)})$ and $\pi_2 = (d_1^{(2)}, \ldots, d_n^{(2)})$. We say that π_1 and π_2 *pack* if there exist edge-disjoint k-graphs G_1 and G_2 on vertex set $\{v_1, \ldots, v_n\}$ such that $d_{G_1}(v_i) = d_i^{(1)}$ and $d_{G_2}(v_i) = d_i^{(2)}$ for all i. When we discuss packing of graphic sequences, the sequences need not be nonincreasing; however, no reordering of the indices is allowed.

Dürr, Guiñez, and Matamala [32] showed that the problem of packing two graphic sequences is NP-complete, and we show that the same conclusion holds when considering k-graphic sequences for $k \geq 3$.

Theorem 3.5.2. *The degree-sequence packing problem for k-graphs is NP-complete for all $k \geq 2$.*

Proof. The degree-sequence packing problem for $k \geq 2$ is in NP since the certificate giving realizations that pack can easily be checked in polynomial time. NP-hardness for $k = 2$ is shown in [32]. For $k \geq 3$ we show that any instance of the degree-sequence packing problem for 2-graphs can be reduced to an instance of the degree-sequence packing problem for k-graphs. Given 2-graphic sequences π_1 and π_2, add $k - 2$ new entries to each sequence to create sequences π_1^k and π_2^k, with each new entry of π_i^k equal to $\frac{1}{2}\sigma(\pi_i)$. Then, any k-realization of π_i^k has the same number of edges as a 2-realization of π_i, and each of the $k - 2$ vertices associated with the new entries must appear in every edge. Hence there is a one-to-one correspondence between 2-realizations of the original sequences and k-realizations of the new sequences. $\square$

Given the computational complexity of the overarching problem, it is natural to seek sufficient conditions that ensure a pair of k-graphic sequences pack. Busch et al. showed that if π_1 and π_2 are graphic sequences and $\Delta \leq \sqrt{2\delta n} - (\delta - 1)$, where Δ and δ are the largest and smallest entries in $\pi_1 + \pi_2$, then π_1 and π_2 pack. We prove a similar result for k-graphic sequences when $k \geq 3$.

For a vertex v in a k-graph H, we define the neighborhood of v, $N_H(v)$, to be the set of vertices that are in at least one edge with v. Similarly, for a set $S = \{v_1, \ldots, v_m\}$ of vertices in H, the neighborhood of S is $N_H(S) = \cup_{i=1}^m N_H(v_i)$. When H is understood, we write $N(v)$. Also, let $H[S]$ denote the subgraph of H induced by the vertices in S.

Theorem 3.5.3. *Fix an integer $k \geq 2$. There exist constants c_1, c_2 (depending only on k) such that if π_1 and π_2 are k-graphic sequences each with length n that satisfy*

$$n > c_1 \frac{\Delta^{k/(k-1)}}{\delta} + c_2 \Delta, \tag{3.5.1}$$

where Δ and δ are the maximum and minimum entries of $\pi_1 + \pi_2$, then π_1 and π_2 pack.

Proof. Among all k-realizations of π_1 and π_2, let H_1 and H_2 be k-realizations such that the number of double edges in $H_1 \cup H_2$ is minimized. We may assume that $H_1 \cup H_2$ has at least one multiple edge, lest H_1 and H_2 give rise to a packing. Let $H = H_1 \cup H_2$, $e = \{v_1, \ldots, v_k\}$ be a double edge in H, and $I = V(H) \setminus \bigcup_{i=1}^{k} N_H(v_i)$. Taking $c_2 > k^2 - k$, inequality $(3.5.1)$ implies that $I \neq \emptyset$. Let $Q = N_H(I)$.

If there is some edge f that contains more than one vertex of I, say i_1 and i_2, then the 2-switch $e \underset{i_1}{\overset{v_1}{\rightleftharpoons}} f$ reduces the number of double edges, contradicting the choice of H_1 and H_2. Hence, each edge including a vertex of I consists of that vertex and $k-1$ vertices of Q.

Let $Q_i = N_{H_i}(I)$ for $i \in \{1, 2\}$. Suppose Q_1 is not a clique in H. That is, let $A = \{y_1, \ldots, y_k\}$ be a set of vertices in Q_1 that is not an edge in H. Since each y_j is in Q_1, for each j with $1 \leq j \leq k$ there is an edge $f_j \in H_1$ that contains both y_j and some vertex of I. Let $\mathcal{E} = \{f_1, \ldots, f_k\}$ be a set of such edges in H_1, where it is possible that some f_j's are equal. Now we can repeatedly perform 2-switches of the form $e \underset{y_j}{\overset{v_j}{\rightleftharpoons}} f_j$ until one copy of e is replaced by the new edge $\{y_1, \ldots, y_k\}$, in the following way. First, do the exchange $S_1 = e \underset{y_1}{\overset{v_1}{\rightleftharpoons}} f_1$ to obtain edges $e_1 = e - v_1 + y_1$ and $f_1' = f_1 - y_1 + v_1$. The edge e_1 may already exist in H, but it will be removed in the next step. The edge f_1' cannot exist in H, as it contains both a vertex of e and a vertex of I. Having performed edge exchanges S_1 through S_j, the next exchange is $S_{j+1} = e_{j+1} \underset{y_{j+1}}{\overset{v_{j+1}}{\rightleftharpoons}} f_{j+1}$, unless $f_{j+1} = f_p$ for some $p \leq j$. In that case, $f_{j+1} = f_p$ is no longer an edge, but has been transformed into the edge $f_p' = f_p - y_p + v_p$. Then $S_{j+1} = e_{j+1} \underset{y_{j+1}}{\overset{v_{j+1}}{\rightleftharpoons}} f_p'$, and the new edges created in this exchange are $e_{j+1} = e_j - v_j + y_j$ and $f_j' = f_p' - y_j + v_j$. After the kth iteration of this process, we have created the

edge consisting of the vertices in A, and removed one of the copies of e, while no new double edges have been created. Since this contradicts our choice of H_1 and H_2, the vertices of A must already form an edge, so Q_1 is a clique. The same argument shows that Q_2 is a clique.

Let $v_i \in e$ and $x \in Q$, and suppose that $e - v_i + x$ is not an edge in H. Let f be an edge containing x and a vertex of I. Then the switch $e \overset{v_i}{\underset{x}{\rightleftharpoons}} f$ reduces the number of double edges in H. Hence every vertex of Q is in an edge with each of the $(k-1)$-subsets of e.

Let $q = |Q|$ and $r = |E(H[Q])|$. Since Q_1 and Q_2 are cliques, $r \geq 2\binom{q/2}{k}$. Counting the degrees of vertices in Q, we have

$$\Delta q \geq kq + (k-1)\delta|I| + kr$$
$$\geq kq + (k-1)\delta|I| + 2k\binom{q/2}{k}.$$

Rearranging gives

$$|I| \leq \frac{(\Delta - k)q - 2k\binom{q/2}{k}}{(k-1)\delta}. \tag{3.5.2}$$

By the principle of inclusion-exclusion, we also know that

$$|I| = n - \left| \bigcup_{i=1}^{k} N_H(v_i) \right| \tag{3.5.3}$$
$$= n + \sum_{s=1}^{k} (-1)^s \sum_{\substack{B \subseteq e \\ |B|=s}} \left| \bigcap_{v \in B} N_H(v) \right|.$$

For any subset B of e, we have that all of Q and $e \setminus B$ are in the common neighborhood

of B in H; thus

$$q + k - |B| \leq \left| \bigcap_{v \in B} N_H(v) \right|.$$

On the other hand, the size of this common neighborhood is maximized when all vertices in B have the same neighborhood; hence

$$\left| \bigcap_{v \in B} N_H(v) \right| \leq (k-1)(\Delta - 2) + k - |B|.$$

Using these inequalities in (3.5.3), we have

$$|I| \geq n - \sum_{s \text{ odd}} \binom{k}{s} \left((k-1)(\Delta - 2) + k - s \right) + \sum_{s \text{ even}} \binom{k}{s} (q + k - s)$$

$$= n + \sum_{s=1}^{k} (-1)^s (k-s) \binom{k}{s} - (k-1)(\Delta - 2) \sum_{s \text{ odd}} \binom{k}{s} + q \sum_{s \text{ even}} \binom{k}{s}.$$

Applying the binomial theorem, this becomes

$$|I| \geq n - k - (\Delta - 2)(k-1)\left(2^{k-1} \right) + q \left(2^{k-1} - 1 \right) \tag{3.5.4}$$

$$= n - \Delta(k-1)\left(2^{k-1} \right) + q \left(2^{k-1} - 1 \right) + (k-1)\left(2^k - 1 \right) - 1.$$

Combining equations (3.5.2) and (3.5.4) yields

$$(k-1)\delta\left(n - \Delta(k-1)\left(2^{k-1} \right) + (k-1)\left(2^k - 1 \right) - 1 \right)$$

$$\leq (\Delta - k)q - (k-1)\delta \left(2^{k-1} - 1 \right) q - 2k \binom{q/2}{k} \tag{3.5.5}$$

$$\leq \Delta q - 2k \frac{q^k}{(2k)^k}.$$

Without loss of generality, suppose $|Q_1| \geq |Q_2|$, and let $q_1 = |Q_1|$. Since Q_1 is

a clique, $\binom{q_1-1}{k-1} \leq \Delta$, so $q_1 \leq c'\Delta^{1/(k-1)}$ for some constant c' depending only on k. Then, since $Q = Q_1 \cup Q_2$, we have $q \leq 2q_1 \leq 2c'\Delta^{1/(k-1)} = c\Delta^{1/(1-k)}$. Inequality (3.5.5) now becomes

$$n \leq \left(c - \frac{c^k}{(2k)^{k-1}}\right) \frac{\Delta^{k/(k-1)}}{\delta} + \left((k-1)2^{k-1}\right)\Delta.$$

This establishes the theorem, with $c_1 = \left(c - \frac{c^k}{(2k)^{k-1}}\right)$ and $c_2 = \left((k-1)2^{k-1}\right)$. □

When $\delta = o\left(\Delta^{1/(k-1)}\right)$, the bound in Theorem 3.5.3 reduces to

$$n > c\frac{\Delta^{k/(k-1)}}{\delta}$$

for $c = c_1 + c_2$. We show that for δ in this range, Theorem 3.5.3 is best possible up to the choice of c.

Fix k and δ and choose an integer $x \gg \delta$ such that $\frac{x-k}{\delta(k-1)}$ is an integer. Form a complete k-graph on x vertices; set aside k of these vertices to form the set B, and let T be the set of remaining vertices. Add an independent set I of order

$$\frac{(x-k)}{\rho(k-1)\delta}\binom{x-1}{k-1},$$

where $\rho > 1$ is chosen such that $\frac{1}{\rho}\binom{x-1}{k-1}$ is an integer. Partition T into sets $T_1, \ldots, T_r$, each of size $\delta(k-1)$, where $r = \frac{x-k}{\delta(k-1)}$, and partition I into sets $I_1, \ldots, I_r$ of size $\frac{1}{\rho}\binom{x-1}{k-1}$. For each vertex $v \in I_j$, create edges $e_1, \ldots, e_\delta$, where each edge consists of v and $k-1$ distinct vertices of T_j. Thus, $N(v) = T_j$ and each vertex in T_j is in exactly one edge with each vertex of I_j. Finally, add an independent set of size $x - k + |I|$.

We now have a k-graph H where each vertex in T has degree

$$\binom{x-1}{k-1} + \frac{1}{\rho}\binom{x-1}{k-1} = \left(1 + \frac{1}{\rho}\right)\binom{x-1}{k-1},$$

each vertex in B has degree $\binom{x-1}{k-1}$, and each vertex in I has degree δ.

Consider two orderings of the degree sequence of H:

$$\pi_1 = \left(\binom{x-1}{k-1}^k, \left(\left(1 + \frac{1}{\rho}\right)\binom{x-1}{k-1}\right)^{x-k}, 0^{x-k}, \delta^{|I|}, 0^{|I|}\right)$$

$$\pi_2 = \left(\binom{x-1}{k-1}^k, 0^{x-k}, \left(\left(1 + \frac{1}{\rho}\right)\binom{x-1}{k-1}\right)^{x-k}, 0^{|I|}, \delta^{|I|}\right).$$

Note that n, the length of sequences π_1 and π_2, is

$$n = 2x - k + 2|I|$$
$$= 2x - k + \frac{2(x-k)}{\rho(k-1)\delta}\binom{x-1}{k-1}$$
$$= \Theta(x^k/\delta).$$

In $\pi_1 + \pi_2$ the minimum degree is δ and the maximum degree is $\Delta = 2\binom{x-1}{k-1} = \Theta(x^{k-1})$. Hence $\Delta = \Theta((\delta n)^{(k-1)/k})$.

In any realization of π_1, Lemma 3.5.4 implies that the vertices of degree greater than δ must form a clique. Since the k vertices of B must be in this clique, those vertices must form an edge in any realization of π_1. The same argument applies to π_2, hence the sequences do not pack.

Lemma 3.5.4 ([6]). *If $\pi = (d_1, \ldots, d_n)$ is a k-graphic sequence, then*

$$\sum_{i=1}^{t} d_i \leq k \binom{t}{k} + (k-1) \sum_{j=t+1}^{n} d_j$$

for $k \leq t \leq n$. If equality holds, then the t vertices of highest degree in any k-realization of π form a clique, and any edge not contained in the clique contains exactly one vertex outside the clique.

Acknowledgements

Chapter 4

Induced-saturation[1]

4.1 Background and introduction

A well known graph parameter is the *saturation number*, defined for a graph H and a whole number n as the minimum number of edges in a graph G on n vertices such that H is not a subgraph of G, but H occurs if any edge of $\overline{G}$ is added to G. Formally,

$$\mathrm{sat}(n, H) = \min\{|E(G)| : \mathrm{G \ has} \ n \ \mathrm{vertices}, H \nsubseteq G, \ \mathrm{and} \ \forall e \notin E(G), H \subseteq G + e\}.$$

Determining the saturation number for a given graph H has proven, in general, quite difficult. For more information on the saturation number, see the dynamic survey of Faudree, Faudree, and Schmitt [36].

A natural attempt at defining an induced variant of graph saturation would be to state that an n-vertex graph G is H-induced-saturated if G is H-free and for all $e \notin E(G)$, $G+e$ contains H as an induced subgraph. Unfortunately, this is not always well defined. That is, there exist graphs H and values of $n \geq |V(H)|$ for which every n-vertex graph G either contains H as induced subgraph, or there exists $e \notin E(G)$ such that $G + e$ is H-free. A simple example is $n = 4$ and $H = K_{1,3}$.

In this chapter, we consider a variant of the saturation number introduced by Martin and Smith in 2012 that looks for induced copies of H, and considers deleting as well as adding edges. To create a well defined parameter, Martin and Smith [69] make use of *trigraphs*, objects also used by Chudnovsky and Seymour in their structure theorems on claw-free graphs [27].

Definition 4.1.1. A *trigraph* T is a quadruple $(V(T), E_B(T), E_W(T), E_G(T))$, where $V(T)$ is the vertex set and the other three elements partition $\binom{V(T)}{2}$ into a set $E_B(T)$ of black edges, a set $E_W(T)$ of white edges, and a set $E_G(T)$ of gray edges. These can be thought of as edges, nonedges, and potential edges, respectively. For any

$e \in E_B(T) \cup E_W(T)$, let T_e denote the trigraph where e is changed to a gray edge, i.e. $T_e = (V(T), E_B(T) - e, E_W(T) - e, E_G(T) + e)$.

A *realization* of T is a graph $G = (V(G), E(G))$ with $V(G) = V(T)$ and $E(G) = E_B(T) \cup S$ for some $S \subseteq E_G(T)$. Let $\mathcal{R}(T)$ be the family of graphs that are a realization of T.

A trigraph T is H-*induced-saturated* if no realization of T contains H as an induced subgraph, but H occurs as an induced subgraph of some realization whenever any black or white edge of T is changed to gray. Formally,

$$\text{indsat}(n, H) = \min\{|E_G(T)| : |V(T)| = n, \forall G \in \mathcal{R}(T), H \not\leq G,$$

$$\text{and } \forall e \in E_B(T) \cup E_W(T), H \leq G'$$

$$\text{where } G' \in \mathcal{R}(T_e)\}.$$

The *induced saturation number* of a graph H with respect to n, written $\text{indsat}(n, H)$, is the minimum number of gray edges in an H-induced-saturated trigraph with n vertices.

Notice that a trigraph with $E_G(T) = \emptyset$ has a unique realization, so if $\text{indsat}(n, H) = 0$, there is a graph G that has no induced copy of H yet adding or removing any edge creates an induced copy of H. We will call such a graph H-*induced-saturated.*

The *complement* of a trigraph T, denoted $\overline{T}$, is the trigraph with $V(\overline{T}) = V(T)$, $E_B(\overline{T}) = E_W(T)$, $E_W(\overline{T}) = E_B(T)$, and $E_G(\overline{T}) = E_G(T)$.

4.1.1 Notation

For graphs G and H, we let $G \cup H$ denote the disjoint union, $G \vee H$ denote the join, and $G \square H$ denote the Cartesian product of the two graphs. A trivial component of a graph is an isolated vertex. For a graph G, we use $n(G)$ for the number of vertices and $e(G)$ for the number of edges in G. We let P_n denote the path on n vertices and C_n the cycle on n vertices. K_n is the complete graph on n vertices, and for $k \geq 2$, $K_{a_1,\ldots,a_k}$ is the complete multipartite graph with parts of size $a_1, \ldots, a_k$. $K_{1,3}^+$ is the paw, which is obtained by adding a single edge to $K_{1,3}$. For a set $S \subseteq V(G)$, $G[S]$ is the subgraph of G induced by S, and if $S = \{v_1, \ldots, v_p\}$, we will sometimes write $G[v_1, \ldots, v_p]$. For a vertex $v \in V(G)$, $N_G(v)$ (or $N(v)$, if G is clear from context) is the set of neighbors of v in G, and $N[v] = N(v) \cup \{v\}$. We use $\deg_G(v)$ or $\deg(v)$ to denote the degree of v, that is, $|N(v)|$. In a trigraph, the black (resp. gray) degree of a vertex is the number of black (resp. gray) edges incident to that vertex. We say a set S of vertices *dominates* G, and we call S a *dominating set*, if every vertex of $G - S$ is adjacent to some vertex in S; if $S = \{v\}$, we say v is a *dominating vertex*. Similarly, a vertex u *dominates* a vertex set S if u is adjacent to every vertex in S. Finally, for an integer n, we let $[n] = \{1, \ldots, n\}$. Other notation will be defined as it is used, or see [84] for any undefined terms.

4.1.2 Observations and previous results

By definition, the only trigraphs on fewer than $v(H)$ vertices that are H-induced-saturated are those in which all edges are gray. Thus we will usually assume that $n \geq |V(H)|$ when we discuss indsat(n, H).

The following theorem summarizes the results of Martin and Smith [69]:

Theorem 4.1.2. *Let H be a graph.*

- *For all $n \geq |V(H)|$, $\mathrm{indsat}(n, H) \leq \mathrm{sat}(n, H)$. By [54], $\mathrm{sat}(n, H) \in O(n)$, so in particular $\mathrm{indsat}(n, H) \in O(n)$.*

- *For all $n \geq m \geq 3$, $\mathrm{indsat}(n, K_m) = \mathrm{sat}(n, K_m)$. (Note that $\mathrm{sat}(n, K_m)$ was determined by Erdős, Hajnal, and Moon in [34].)*

- *For all $n \geq m \geq 2$, and for $e \in E(K_m)$, $\mathrm{indsat}(n, K_m - e) = 0$. In particular, for all $n \geq 3$, $\mathrm{indsat}(n, P_3) = 0$.*

- *For all $n \geq 4$, $\mathrm{indsat}(n, P_4) = \left\lceil \frac{n+1}{3} \right\rceil$.*

We also make the following observation:

Observation. A trigraph T is H-induced-saturated if and only if $\overline{T}$ is $\overline{H}$-induced-saturated. In particular, $\mathrm{indsat}(n, H) = \mathrm{indsat}(n, \overline{H})$.

Proof. Suppose a trigraph T has a realization G such that H is an induced subgraph of G. Then $\overline{H}$ is an induced subgraph of $\overline{G}$. Using the definition of $\overline{T}$, $\overline{G}$ is a realization of $\overline{T}$. It follows that a trigraph T is H-induced-saturated if and only if $\overline{T}$ is $\overline{H}$-induced-saturated. □

4.1.3   Minimally H-induced-saturated graphs

In this chapter we show that $\mathrm{indsat}(n, H)$ is zero for stars, which as noted above, means that there exists a graph that is H-induced-saturated. This leads to the natural question: What is the minimum number of edges in such a graph?

Definition 4.1.3. For a graph H and whole number n with $\mathrm{indsat}(n, H) = 0$, we define

$$\mathrm{indsat}^*(n, H) := \min\{e(G) : |V(G)| = n \text{ and } G \text{ is } H\text{-induced-saturated}\}.$$

We say a graph G on n vertices with indsat$^*(n, H)$ edges is *minimally H-induced-saturated.*

By Observation 4.1.2, the *maximum* number of edges in an n-vertex H-induced-saturated graph is $\binom{n}{2} - $ indsat$^*(n, \overline{H})$.

In this chapter we show that stars $(K_{1,t})$ have induced-saturation number zero for n sufficiently large. We also determine indsat$^*(n, K_{1,t})$ within a factor of 2 and show that the upper bound is correct for $K_{1,3}$. Finally, we introduce the induced-saturation number of a family of graphs and show that while every graph in a family may have induced-saturation number zero, the family itself could have nonzero induced-saturation number. In further joint work [7], we show that the following graphs have induced-saturation number zero for n sufficiently large: $K_{1,3}^+$, stars , C_4, odd cycles, some modifications of even cycles, and matchings. Additionally, we provide bounds on indsat$^*(n, H)$ for the graphs listed above. In particular, we characterize the $K_{1,3}^+$-induced-saturated graphs, which in turn completely determines indsat$^*(n, K_{1,3}^+)$.

4.2 Stars

Recall that $K_{1,2} = P_3$, and indsat$(n, P_3) = 0$ for $n \geq 3$, as established in [69]. In this section we provide a construction extending this result, to show that for fixed $k \geq 2$ and n sufficiently large, indsat$(n, K_{1,k+1}) = 0$. Additionally, our construction, together with a simple argument, determines indsat$^*(n, K_{1,k+1})$ within a factor of two. The case when $k = 2$, which refers to the graph $K_{1,3}$, commonly known as the claw, will be addressed in further detail in Section 4.3.

Construction 4.2.1. Fix $k \geq 2$ and $n \geq 3^k$. Let z, R be positive integers such that $n = z3^k + R$ with $0 \leq R < 3^k$. Let H be the graph $K_3^1 \square K_3^2 \square \cdots \square K_3^k$, where K_3^i

denotes a single copy of K_3. In other words, $V(H) = \{(\alpha_1, \ldots, \alpha_k) : \alpha_i \in [3]\}$, and $(\alpha_1, \ldots, \alpha_k)(\beta_1, \ldots, \beta_k) \in E(H)$ if and only if $\sum |\{i : \alpha_i \neq \beta_i\}| = 1$. Define H' where $V(H')$ is the disjoint union of $V(H)$ and $V(K_R)$, and $E(H')$ consists of $E(H)$, $E(K_R)$ and the edges between H and K_R satisfy: for each $v \in V(K_R)$, $v\alpha \in E(H')$ if and only if $\alpha = (\alpha_1, 1, 1, \ldots, 1)$, $\alpha_1 \in [3]$. We now define G to be the disjoint union of $z - 1$ copies of H and a single copy of H'.

Proposition 4.2.2. *The graphs in Construction 4.2.1 are $K_{1,k+1}$-induced-saturated.*

Proof. Given fixed n and k, let G and R be as defined in Construction 4.2.1. Let F denote the subgraph of H' isomorphic to H. Suppose that G contains an induced $K_{1,k+1}$ with center x, and suppose first that x is in a copy of H. Since $V(G)$ can be represented by k-dimensional vectors as described in the construction, any set of $k + 1$ neighbors of x contains two vertices with vectors that differ in exactly one coordinate. Thus, x cannot have $k + 1$ neighbors which form an independent set, and H is $K_{1,k+1}$-free.

If H' contains an induced $K_{1,k+1}$, then x cannot be in the K_R as the neighborhood of x would be a clique. Thus, x is in F. If this induced $K_{1,k+1}$ contains no vertices from the copy of K_R, then the above argument produces a contradiction. Thus, this $K_{1,k+1}$ contains a vertex from the copy of K_R, and without loss of generality, we may assume that x is represented by $(1, 1, \ldots, 1)$ in F. Consequently, our $K_{1,k+1}$ has exactly one vertex in K_R, but then contains no vertices of the form $(\alpha_1, 1, 1, \ldots, 1)$ other than x. Hence, x has at most $k - 1$ other neighbors from F in this copy of $K_{1,k+1}$ from F. So G is $K_{1,k+1}$-free.

It is clear that every vertex in a copy of H (or in F) is the center of an induced $K_{1,k}$. Thus, if we add an edge between two components of G, one component must be a copy of H, and we obtain an induced $K_{1,k+1}$. Thus, it remains to consider adding an

edge within a component. Note that by the construction of H', the only possible way to add an edge is within F, which is isomorphic to H. So, it suffices to consider adding an edge to a copy of H. Suppose we add the edge uv. Without loss of generality, we may assume that u is represented by $(1, 1, \ldots, 1)$. Since u and v were not adjacent in H, their corresponding vectors must differ in at least two coordinates, say the first and second. As a consequence, v is adjacent to neither y nor w, where $y \in \{(2, 1, 1, \ldots, 1), (3, 1, 1, \ldots, 1)\}$ and $w \in \{(1, 2, 1, 1, \ldots, 1), (1, 3, 1, 1, \ldots, 1)\}$. Thus, $\{u, v, w, y\}$ is an induced $K_{1,3}$ centered at u. To this set we add vertices $\alpha^3, \alpha^4, \ldots, \alpha^k$, where α^i has all coordinates equal to 1 except that the ith coordinate is either 2 or 3. This induces $K_{1,k+1}$.

Lastly, suppose we remove an edge uv. There are three cases to consider. The first case is if uv is in a copy of H (or in F). Here, we may assume $u = (2, 1, 1, \ldots, 1)$ and $v = (3, 1, 1, \ldots, 1)$. The second case is if both u and v are in K_R. The last case is if only one endpoint, say v, is in K_R. Here, we may again assume that $u = (2, 1, 1, \ldots, 1)$. In all three cases, $(1, 1, \ldots, 1)$ together with u, v, and the vertices $\alpha^2, \ldots, \alpha^k$ defined above induce a $K_{1,k+1}$. This completes the lemma. $\qquad \square$

Corollary 4.2.3. *For fixed $k \geq 2$ and $n \geq 3^k$, $\mathrm{indsat}(n, K_{1,k+1}) = 0$.*

Theorem 4.2.4. *For $n \geq 2 \cdot 3^k$ and $k \geq 2$, there exist constants $c_1 = c_1(k)$ and $c_2 = c_2(k)$ such that $n\frac{k}{2} - c_1 \leq \mathrm{indsat}^*(n, K_{1,k+1}) \leq nk + c_2$.*

Proof. Given fixed n and k, let G and R be as defined in Construction 4.2.1.

We establish $e(G)$ by considering vertex degrees. The component H' has at most $2 \cdot 3^k$ vertices, and so (trivially) at most $\binom{2 \cdot 3^k}{2}$ edges. The remaining vertices, of which there are at most $n - 3^k$, all have degree $2k$ for a contribution of at most $(n - 3^k)k$ edges. All told, $e(G) \leq nk - k \cdot 3^k + \binom{2 \cdot 3^k}{2}$.

To show the lower bound, suppose that G is a $K_{1,k+1}$-induced-saturated graph. Let $S = \{x \in V(G)\colon \deg(x) \leq k-1\}$. We claim that $|S| \leq k$.

If $|S| > k$, then there exist $x, y \in S$ such that $xy \notin E(G)$. Let G' denote $G + xy$. As G was $K_{1,k+1}$-induced-saturated, G' must contain an induced $K_{1,k+1}$, using the edge xy with either x or y as the center of this $K_{1,k+1}$. However, as both x and y are adjacent to at most $k-1$ vertices in G, this cannot happen. So $|S| \leq k$, as claimed.

Observe:
$$e(G) \geq \frac{1}{2}\left(k(n-|S|) + \sum_{x \in S} \deg(x)\right) \geq \frac{nk}{2} - \frac{k^2}{2}.$$

This establishes the lower bound. $\qquad\square$

It is worth noting that we can extend Construction 4.2.1, as any graph formed as a Cartesian product of exactly k cliques, each of size at least three, is $K_{1,k+1}$-induced-saturated.

4.3 The claw

For sufficiently large n, Theorem 4.2.4 states that $\mathrm{indsat}^*(n, K_{1,k+1})$ is linear in n, and in particular, we know the coefficient within a factor of two. In this section, we will determine the coefficient of $\mathrm{indsat}^*(n, K_{1,3})$, which coincides with the upper bound given in Theorem 4.2.4. Additionally, we will provide better constructions than that in Construction 4.2.1, which will ultimately determine $\mathrm{indsat}^*(n, K_{1,3})$ within an additive constant of four.

Values of $\mathrm{indsat}(n, K_{1,3})$ were determined for $4 \leq n \leq 10$ by computer search[2] and are listed in Table 4.1, along with trigraphs that achieve the minimum number

---

[2]A program was written in C++ and is available at `http://www.math.unl.edu/~s-sbehren7/main/Data.html`.

of gray edges. This, together with Corollary 4.2.3 determines indsat$(n, K_{1,3})$ for all n. We now turn our attention to indsat$^*(n, K_{1,3})$.

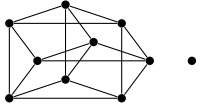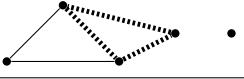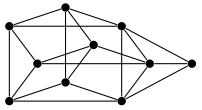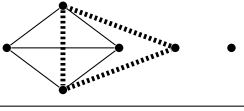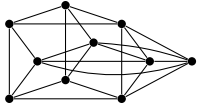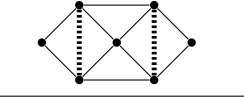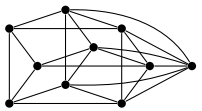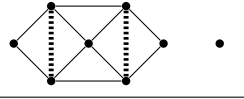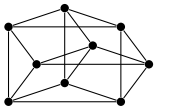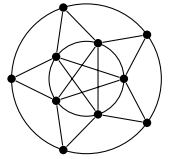|	
indsat$(4, K_{1,3}) = 3$	
indsat$(5, K_{1,3}) = 3$	
indsat$(6, K_{1,3}) = 3$	indsat$(10, K_{1,3}) = 0$
indsat$(7, K_{1,3}) = 2$	
indsat$(8, K_{1,3}) = 2$	
indsat$(9, K_{1,3}) = 0$	

Table 4.1: Values of indsat$(n, K_{1,3})$ for $4 \leq n \leq 10$ along with trigraphs realizing those values. All $K_{1,3}$-induced-saturated graphs for $n = 9$ and $n = 10$ are shown.

Theorem 4.3.1. *The following bounds hold for $n \geq 9$, $n \neq 14, 17$:*

$$\text{indsat}^*(n, K_{1,3}) = 2n \qquad \qquad \textit{if} \quad n \equiv 0 \mod 3$$

$$\text{indsat}^*(n, K_{1,3}) = 2n - 2 \qquad \quad \textit{if} \quad n \equiv 1 \mod 3$$

$$2n \leq \text{indsat}^*(n, K_{1,3}) \leq 2n + 2 \quad \textit{if} \quad n \equiv 2 \mod 3.$$

In order to prove Theorem 4.3.1, we first prove a series of lemmas that will aid in producing the lower bounds of the statement. Then we construct families of $K_{1,3}$-induced-saturated graphs that exhibit the upper bounds of Theorem 4.3.1.

The following lemma shows that $K_{1,3}$-induced-saturated graphs have few vertices of low degree.

Lemma 4.3.2. *Let G be a $K_{1,3}$-induced-saturated graph. Then G has*

1. *at most one isolated vertex,*

2. *no vertices of degree one,*

3. *at most one vertex of degree two, and*

4. *at most two vertices of degree three.*

Furthermore, if G has an isolated vertex v, then $\delta(G - v) \geq 4$. Additionally, if G has a vertex of degree three, then G does not have a vertex of degree two. If G has two vertices of degree three or a vertex of degree two, then G has a vertex of degree at least five.

Proof. Let G be a $K_{1,3}$-induced-saturated graph. Observe that if we had two isolated vertices, then adding the edge between them would not yield a $K_{1,3}$. Also, any edge of G lies in a triangle, so there are no vertices of degree one.

Suppose that u and v are vertices of degree two. Since every edge lies in a triangle the neighbors of u are adjacent, as are the neighbors of v. Thus, if u and v are not adjacent, adding the edge uv does not create an induced $K_{1,3}$. If u and v are adjacent, then $N[u] = N[v] = \{u, v, w\}$ for some w. However, removing uw does not create an induced $K_{1,3}$ as v would have to have been its center. So G has at most one vertex of degree two.

To prove (4), suppose u is a vertex of degree three with neighbors u_1, u_2, u_3. Since every edge is in a triangle, we may assume that $u_1u_2, u_2u_3 \in E(G)$. *Case 1:* $u_1u_3 \notin E(G)$. Then adding u_1u_3 creates an induced $K_{1,3}$ centered at either u_1 or u_3; say u_1. Then u_1 has two nonadjacent neighbors x and y that are distinct from u_2 and u_3. However, $\{u, u_1, x, y\}$ induces a $K_{1,3}$ in G, a contradiction. *Case 2:* $u_1u_3 \in E(G)$. In particular, every vertex of degree three in G is contained in a K_4. Let v be another vertex of degree three. By the above, $N[v]$ induces K_4. If $uv \notin E(G)$, then adding

uv does not create an induced $K_{1,3}$. Thus, u and v are adjacent, and consequently the only vertices of degree three are contained in $N[u]$.

If we remove uu_1, then an induced $K_{1,3}$ exists, centered at either u_2 or u_3. So at least one of them has degree at least four, say u_3. Similarly, removing uu_3 creates an induced $K_{1,3}$ centered at either u_1 or u_2 so that at least one of them has degree at least four. In any case, at most two vertices in $N[u]$, and as a result in G, have degree three. Thus, (4) holds.

If G has an isolate, u, and another vertex v with $\deg(v) \leq 2$, then adding uv cannot create an induced $K_{1,3}$ unless $\deg(v) = 2$. In this case, the neighbors of v cannot be adjacent, however every edge of G must be in a triangle, a contradiction.

Suppose u and v are vertices with $\deg(u) = 2$ and $\deg(v) = 3$. By previous arguments, the neighbors of u form a clique, as do the neighbors of v. Thus, if $uv \notin E(G)$, adding uv does not create an induced $K_{1,3}$. So $uv \in E(G)$, and in particular, u is in the K_4 induced by $N[v]$. However, $\deg(u) = 2$, a contradiction.

Now, suppose u and v are vertices with $\deg(u) = \deg(v) = 3$. By the above, they must be contained in the same K_4, so let u, v, x, y denote the vertices of this K_4. If we delete xy, then x and y are the leaves of a $K_{1,3}$, but this $K_{1,3}$ is not centered at u or v, so x and y have a common neighbor $z \notin \{u, v\}$. If we delete xz, the resulting $K_{1,3}$ is centered at a common neighbor of x and z. If that common neighbor is not y, then $\deg(x) \geq 5$, and if it is, then $\deg(y) \geq 5$.

Similarly, suppose $\deg(v) = 2$, with $N(v) = \{x, y\}$. Since every edge is in a triangle, $xy \in E(G)$. If we consider deleting the edge xy, we note that the $K_{1,3}$ formed does not have center v, so x and y share another neighbor z, and z has a neighbor z' nonadjacent to both x and y. Consider deleting the edge vx. The $K_{1,3}$ formed must be centered at y, so y has a neighbor nonadjacent to v or x. Then this neighbor y' is not any of the vertices already named. Similarly, x has a neighbor

$x' \notin \{v, x, y, y', z, z'\}$. Then $\{x', y'\} \subseteq N(z)$ else $G[x, x', v, z]$ or $G[y, y', v, z]$ is a $K_{1,3}$. Thus, $\deg(z) \geq 5$. $\qquad\square$

Corollary 4.3.3. *Any graph that is $K_{1,3}$-induced-saturated (on $n \geq 9$ vertices) has at least $2n-2$ edges. That is,* $\mathrm{indsat}^*(n, K_{1,3}) \geq 2n-2$ *for $n \geq 9$. Furthermore, if G is a $K_{1,3}$-induced-saturated graph that does not have an isolated vertex, then $|E(G)| \geq 2n$.*

Proof. Apply the degree-sum formula and Lemma 4.3.2. $\qquad\square$

As indicated in Corollary 4.3.3, if a graph on n vertices obtaining the minumum number of edges among $K_{1,3}$-induced-saturated graphs exists, then it is four-regular except for an isolated vertex. We provide the following structural results to show such a graph only exists if $n \equiv 1 \mod 3$.

Lemma 4.3.4. *Suppose G is a $K_{1,3}$-induced-saturated graph, and for some $v \in V(G)$, every vertex in $N[v]$ has degree precisely 4. Then $G[N(v)] \in \{2K_2, P_4\}$.*

Proof. Since we are assuming every vertex in $N[v]$ has degree 4, then we can let $N(v) = \{u, x_1, x_2, x_3\}$. Next, we show that $\Delta(G[N(v)]) \leq 2$. Suppose to the contrary that some vertex, say $u \in N(v)$, has three neighbors within $N(v)$; hence, $N(u) \cap N(v) = \{x_1, x_2, x_3\}$. By deleting ux_1, we see that u and x_1 have a common neighbor besides v. Using the symmetry of x_1, x_2, and x_3, without loss of generality $x_1x_2, x_2x_3 \in E(G)$. Now $N(x_2) = \{u, v, x_1, x_3\}$, because $\deg(x_2) = 4$. Consider deleting ux_1. The common neighbors of u and x_1 are v, x_2, and maybe x_3. Neither v nor x_2 can be the center of a $K_{1,3}$ since all of their neighbors are adjacent to u or x_1. Hence x_3 must be the center of the induced $K_{1,3}$, so $x_1x_3 \in E(G)$. But then $N(x_3) = \{v, u, x_1, x_2\}$ so the $K_{1,3}$ supposedly centered at x_3 has no third leaf.

This shows that $\Delta(G[N(v)]) \leq 2$. Because every edge is in a triangle, if $\Delta(G[N(v)]) < 2$, then $G[N(v)] = 2K_2$, so suppose $\Delta(G[N(v)]) = 2$. Then $G[N(v)]$ is either C_4 or P_4.

Suppose $x_1x_2x_3x_4 = P_4 \subseteq G[N(v)]$. If $x_1x_4 \notin E(G)$, then $G[N(v)] = P_4$, so suppose $x_1x_4 \in E(G)$. Deleting the edge x_2x_3 shows that x_2 and x_3 have a common neighbor $y \in V(G) \setminus N[v]$. Separately, consider deleting x_3x_4. The only possible common neighbors of x_3 and x_4 are v and y. Because $x_1x_4 \in E(G)$, v cannot be the center of the $K_{1,3}$ created by deleting x_3x_4, so the center is y. Then the third leaf must be some vertex $y' \notin N(x_3) \cup N(x_4)$. But we also know that $y' \notin N(x_2)$, since $\deg(x_2) = 4$, so $G[y, y', x_2, x_4]$ is an induced $K_{1,3}$, a contradiction. $\qquad \square$

For the remainder of this section, we define $R(G) := \{v \in V(G) : G[N(v)] = 2K_2\}$ and $B(G) := \{v \in V(G) : G[N(v)] = P_4\}$ for any graph G. Hence if G is a four-regular $K_{1,3}$-induced-saturated graph, then $V(G)$ is partitioned into $R(G)$ and $B(G)$. We will call the vertices in $R(G)$ *red vertices* and those in $B(G)$ *blue vertices*.

Lemma 4.3.5. *If G is a 4-regular $K_{1,3}$-induced-saturated graph, then $B(G)$ induces kK_3 for some k.*

Proof. Let $v \in B(G)$ so that $G[N(v)]$ is a path $x_1x_2x_3x_4$. Since $P_3 \subseteq G[\{v, x_1, x_3\}] \subseteq G[N(x_2)]$ and $P_3 \subseteq G[\{v, x_2, x_4\}] \subseteq G[N(x_3)]$, Lemma 4.3.4 implies that $x_2, x_3 \in B(G)$. Furthermore, as deleting x_2x_3 creates an induced $K_{1,3}$, which cannot be centered at v, then x_2 and x_3 share another common neighbor, call it y. Since $N(x_2) = \{v, x_1, x_3, y\}$, $x_1 \in B(G)$ if and only if x_1 and y are neighbors. So if $x_1y \in E(G)$, we consider adding vy to G. This creates an induced $K_{1,3}$, which must be centered at y. However, since G is 4-regular, y has at most one neighbor outside of $\{x_2, x_3, x_4, v\}$ and cannot be the center of this induced $K_{1,3}$, a contradiction. Thus, $x_1 \in R(G)$, and by symmetry, $x_4 \in R(G)$. Repeating the above

argument for x_2 instead of v shows that $y \in R(G)$. Hence, $\{x_2, x_3, v\} \subseteq B(G)$ but $N(\{x_2, x_3, v\}) = \{x_1, x_4, y\} \subseteq R(G)$ and so $\{x_2, x_3, v\}$ induces a triangle of vertices in $B(G)$. □

An example of a 4-regular $K_{1,3}$-induced-saturated graph, with $R(G), B(G) \neq \emptyset$ is shown in Figure 4.1. Observe that $B(G)$ induces $8K_3$, which is in accordance with Lemma 4.3.5.
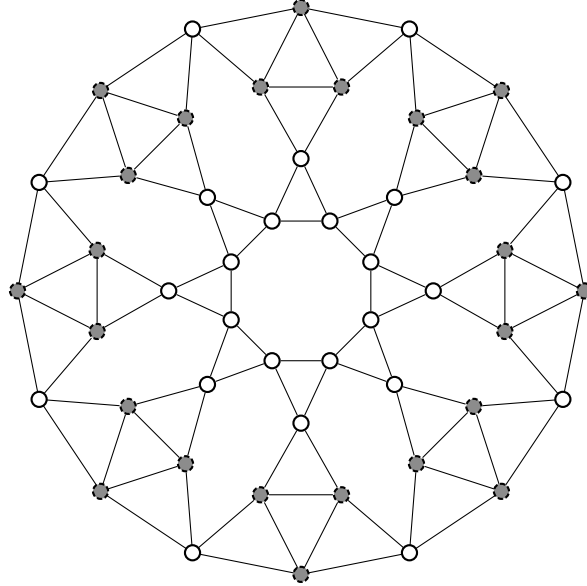
Figure 4.1: A 4-regular $K_{1,3}$-induced-saturated graph. Vertices in $R(G)$ are white, and vertices in $B(G)$ are gray.

Lemma 4.3.6. *Let G be a 4-regular $K_{1,3}$-induced-saturated graph. Every edge of G is in either one or two triangles, and there are $|B(G)|$ edges that are in two triangles.*

Proof. Recall that every edge in a $K_{1,3}$-induced-saturated graph is in at least one triangle. Suppose there exists $xy \in E(G)$ where x and y have three common neighbors u, v, w. Then $G[N(x)]$ cannot be in $\{2K_2, P_4\}$, which contradicts Lemma 4.3.4. Hence, each edge is in at most two triangles.

Let b be the number of edges that are in two triangles. Label edge xy with vertex z if xyz is a triangle, and allow for multiple labels. Thus b edges have two labels and hence

$$|E(G)| + b = \sum_{z \in V(G)} |\{e \in E(G) : e \text{ has label } z\}|.$$

Since each red vertex gives its label to two triangles and each blue vertex gives its label to three triangles, we have $\sum_{z \in V(G)} |\{e \in E(G) : e \text{ has label } z\}| = 2|R(G)| + 3|B(G)|$. Thus, since G is 4-regular,

$$\begin{aligned}
2n + b = |E(G)| + b \\
= \sum_{z \in V(G)} |\{e \in E(G) : e \text{ has label } z\}| \\
= 2|R(G)| + 3|B(G)| \\
= 2(n - |B(G)|) + 3|B(G)| \\
= 2n + |B(G)|
\end{aligned}$$

Therefore, there are precisely $|B(G)|$ edges that are in two triangles. $\square$

Proposition 4.3.7. *If G is a 4-regular, $K_{1,3}$-induced-saturated graph on n vertices, then $n \equiv 0 \mod 3$.*

Proof. Let $b = |B(G)|$. By Lemma 4.3.5, 3 divides b. By Lemma 4.3.6, $2n - b$ edges are in precisely one triangle, and b edges are in precisely two triangles. If t is the number of triangles in G, then $3t = (2n - b) + (2b) = 2n + b$. Since 3 divides b, we know 3 divides $2n$, and so 3 divides n. $\square$

The previous lemmas will be used in the proof of Theorem 4.3.1 to obtain a lower bound $\text{indsat}^*(n, K_{1,3}) \geq 2n - 1$ for $n \equiv 0 \mod 3$. The next two propositions show that certain degree sequences do not have a $K_{1,3}$-induced-saturated realization. This allows us to increase the lower bound of $\text{indsat}^*(n, K_{1,3})$ for certain values of n.

Proposition 4.3.8. *If G is a $K_{1,3}$-induced-saturated graph, then the degree sequence of G is not $(5, 5, 4, \ldots, 4)$.*

Proof. Suppose G is a counterexample to the claim, and let v be a vertex of degree 5.

Case 4.3.8.1. $\Delta(G[N(v)]) = 4$.

That is, v has a neighbor u such that $X := N(u) \cap N(v)$ is a set of order 4. If we delete vx' for some $x' \in X$, then the resulting $K_{1,3}$ is not centered at u since the neighbors of u are adjacent to v. Thus x' and v share a neighbor $x \in X$ and there is some $\in N(x) \setminus [N(x') \cup N(v)]$. Now $N(x) = \{u, v, x', y\}$ and $uy, vy, x'y \notin E(G)$, so the edge xy is in no triangle, a contradiction.

Case 4.3.8.2. $\Delta(G[N(v)]) = 3$.

That is, there exist $u \in N(v), w \notin N[u]$, and $X \subseteq N(u)$ with $|X| = 3$ so that $N(v) = \{u, w\} \cup X$. Since deleting the edge vw creates a $K_{1,3}$, there exist vertices x' and y such that x' is a common neighbor of v and w, y is adjacent to x', and y is not adjacent to w nor v. Note $x' \in X$ and $y \notin N[v]$. Then to prevent a $K_{1,3}$ in G with center x' and leaves u, w, y, we have $uy \in E(G)$. Then, u, v are the vertices of degree 5 and all other vertices have degree 4 so that $N(x') = \{u, v, y, w\}$. Since u is not the center of a $K_{1,3}$, and x' has no neighbors in X, the vertices of $X \setminus \{x\}$ (call them a and b) are adjacent.

Note that u was chosen as an arbitrary vertex of $N(v)$ with three neighbors in $N(v)$, and we showed $\deg(u) = 5$.

Now, $\deg(a) = \deg(b) = 4$ and each of a and b currently has two neighbors in $N(v)$. If a (or b) were adjacent to w, then the argument previously applied to u would guarantee that $\deg(a) = 5$ (or $\deg(b) = 5$), thus giving us at least three vertices of degree 5. Therefore a and b each have a neighbor outside of $N[v]$; due to the necessity that every edge be in a triangle, they share this neighbor, which we shall name z. It is possible that $z = y$. Suppose $z \neq y$, then deleting az should create an induced $K_{1,3}$ centered at at common neighbor of a and z. However, the only option is b, which is not the center of such a $K_{1,3}$, a contradiction. So suppose $z = y$, then $\deg_{N(u)}(a) = 3$. By the previous argument with u, we must have $\deg_G(a) = 5$, a contradiction.

Case 4.3.8.3. $\Delta(G[N(v)]) \leq 2$.

$N(v)$ has no independent set of size three, lest it be the center of a $K_{1,3}$. Then $G[N(v)] \in \{K_2 + K_3, C_5\}$.

Suppose first $G[N(v)] = K_2 + K_3$, with $\{x_1, x_2, x_3\}$ inducing K_3. We may suppose $\deg(x_1) = \deg(x_2) = 4$ since at most one of the vertices in the copy of K_3 may have degree 5. So each of x_1 and x_2 have a neighbor outside of $N[v]$, say y and z, respectively. If $y \neq z$, then since every edge is contained in a triangle, x_3 is adjacent to both y and z. However, this implies that $\deg(z) = 5$ and $\Delta(G[N(z)]) \geq 3$, as evidenced by x_1. This puts us in Case 4.3.8.2.

So $y = z$, and consequently, $N[x_1] = N[x_2]$. The only common neighbors of x_1 and y are x_2 and possibly x_3. If $x_3 \notin N(x) \cap N(y)$, removing x_1y should create an induced $K_{1,3}$ centered at x_2, but it does not. Thus, removing x_1y creates an induced $K_{1,3}$ centered at x_3, which implies that x_3 is adjacent to y, as well as another vertex y' not in $N[v] \cup \{y\}$. However, this implies that $\deg(x_3) = 5$, and $\Delta(G[N(x_3)]) \geq 3$, as evidenced by x_1. This also puts us in Case 4.3.8.2.

Suppose now $G[N(v)] = C_5$ with cycle $x_1x_2x_3x_4x_5$. We may assume that x_1, x_2, x_3, and x_4 all have degree 4. The only common neighbors of v and x_2 are x_1 and x_3. When removing vx_2 we obtain an induced claw centered at either x_1 or x_3. Without loss of generality, assume it is x_3. Since x_4 is not a leaf of a $K_{1,3}$ that features v as a leaf, x_3 has a neighbor $y \notin N(v) \cup N(x_2)$. Since $G[x_3, x_2, x_4, y]$ cannot be a $K_{1,3}$, we must have $yx_4 \in E(G)$. Similarly, if we delete vx_3, the candidates for center of the ensuing $K_{1,3}$ are x_2 and x_4; we know the neighborhood of x_4, and so see that x_2 is the center. Then there exists $y' \in N(x_2)$ such that $y' \notin N(v) \cup N(x_3)$, and as before $y'x_1 \in E(G)$. Now we know the neighborhoods of x_1, x_2, x_3, and x_4. If we add the edge x_1x_4, we find that no $K_{1,3}$ is formed, a contradiction. $\qquad\square$

Proposition 4.3.9. *Let G be a $K_{1,3}$-induced-saturated graph. Then for any $n \geq 7$, the degree sequence of G is not $(6, 4, \ldots, 4)$.*

Proof. Suppose G is a counterexample to this claim. Let v have degree six, and let $F = G[N(v)]$, so $|F| = 6$, $\Delta(F) \leq 3$, and $\alpha(F) \leq 2$ else v is the center of a $K_{1,3}$. In fact $\alpha(F) = 2$ in order for the vertices of $N(v)$ to have degree four in G. If $\delta(F) = 3$, then $N[v]$ is a component of G, and this component is $K_{1,3}$-induced-saturated. However, from the computer search, with results listed in Table 4.1, we know that there is no nontrivial $K_{1,3}$-induced-saturated graph on fewer than nine vertices. Therefore $\delta(F) \leq 2$. Indeed, we claim $\delta(F) = 2$. If $\delta(F) \leq 1$, let u be a vertex with minimum F-degree (i.e. $\deg_F(u)$ is minimum), and let $T = F \setminus N[u]$. Then T is a clique, else two nonadjacent vertices in T together with u and v form a $K_{1,3}$. Hence $|T| = 4$ and the vertices of T have no neighbors outside of $N[v]$ in G. Now, deleting the edge between v and any vertex of T does not create an induced $K_{1,3}$, so $\delta(F) = 2$.

Let u be a vertex in F with $\deg_F(u) = 2$, and let $T = F \setminus N[u]$. As before, T

is a clique, specifically a triangle. Let $N_F(u) = \{u', u''\}$. Since $\deg_F(u) = 2$, u has one neighbor w outside of $N[v]$; since every edge is in a triangle, we may assume that w is adjacent to u'. Now, the only common neighbors of v and u are in $N_F(u)$. Since $\delta(F) = 2$, u' must have another neighbor in F other than u. Thus, the only neighbor of u' not in $N[v]$ is w, and if we delete vu, the resulting induced $K_{1,3}$ cannot be centered at u'. So it must be centered at u'', which in turn has a neighbor w'' outside of $N[v] \cup \{w\}$. Since $u''w''$ is in a triangle and $\delta(F) = 2$, u'' and w'' share a neighbor t'' in F. Since $\delta(F) = 2$, no vertex in F has two neighbors outside $N[v]$. So $t'' \neq u'$, and hence $t'' \in T$. But now $\deg(t'') \geq 5$, a contradiction. $\qquad\square$

Finally, we construct graphs which we use to find an upper bound for $\mathrm{indsat}^*(n, K_{1,3})$.

Lemma 4.3.10. *If G is a graph where the neighborhood of every vertex induces $2K_2$, then G is $K_{1,3}$-induced-saturated.*

Proof. Since no vertex has three independent neighbors, G contains no induced $K_{1,3}$. Suppose we delete an edge xy. Since every edge is in a triangle, say xyz, deleting xy leaves z as the center of a $K_{1,3}$ with leaves x, y, and any other neighbor of z. If we add an edge between two vertices with no common neighbors, then we take the new edge together with two nonadjacent neighbors of one of the vertices and find a $K_{1,3}$. Therefore it suffices to consider adding an edge xy, where x and y share a neighbor. Let $N(x) = \{u_1, u_2, v_1, v_2\}$ with $u_1u_2, v_1v_2 \in E(G)$, and suppose $u_1 \in N(y)$. Then $u_2 \notin N(y)$ otherwise $N(u_2)$ would contain a P_3 and not be $2K_2$. Similarly, both v_1 and v_2 cannot be in $N(y)$. So we may assume $v_2 \notin N(y)$. Then upon adding xy, $\{x, y, u_2, v_2\}$ induces a $K_{1,3}$. $\qquad\square$

Lemma 4.3.11. *Let G be a graph with at most one isolated vertex, where each non-trivial component is one of the graphs in Figure 4.2. Then G is $K_{1,3}$-induced-saturated.*
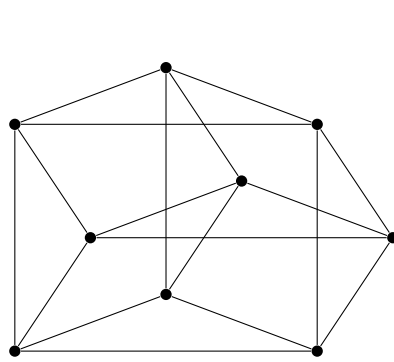
(a) $H = K_3 \square K_3$, 9 vertices
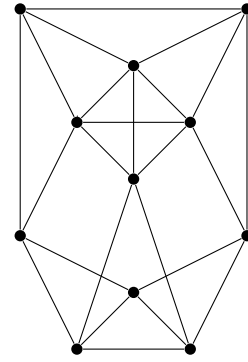
(b) Graph J on 11 vertices
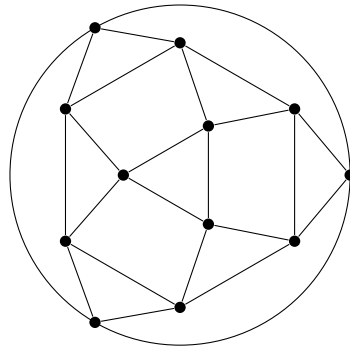
(c) Graph K on 12 vertices
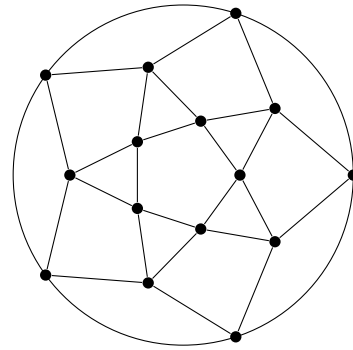
(d) Graph L on 15 vertices

Figure 4.2: These graphs are $K_{1,3}$-induced-saturated.

Proof. By inspection, the graph in Figure 4.2b is $K_{1,3}$-induced-saturated, and since the graphs in Figures 4.2a, 4.2c, and 4.2d have the property that the neighborhood of every vertex induces $2K_2$, they are $K_{1,3}$-induced-saturated by Lemma 4.3.10.

Now let G be a graph with at most one isolated vertex and each of the remaining components are one of the graphs from Figure 4.2. Since each nontrivial component of G is $K_{1,3}$-induced-saturated, we only need to consider adding an edge xy between components. When we add the edge xy, at least one of x and y must be in a nontrivial component, say x. By inspection we see every vertex in every graph of Figure 4.2 has two nonadjacent neighbors, and in particular, this holds for x. Thus, x together with these two neighbors and y induce a $K_{1,3}$. Therefore, G is $K_{1,3}$-induced-saturated. $\square$

We now can prove Theorem 4.3.1.

Proof of Theorem 4.3.1. We exhibit graphs with the desired number of edges to prove the upper bounds.

Case 4.3.11.1. $n \equiv 0 \mod 3$, $n \geq 9$

Use $\lfloor n/9 \rfloor - 1$ copies of H, together with one copy of H, K, or L, for a graph with $2n$ edges. Alternatively, we could generalize L for $n \geq 15$ by having $n/3$ vertices in the outer cycle, $n/3$ vertices in the inner cycle, and $n/3$ vertices between the two cycles.

Case 4.3.11.2. $n \equiv 1 \mod 3$, $n \geq 10$

Use an isolated vertex with a graph from Case 4.3.11.1 for a graph with $2n - 2$ edges.

Case 4.3.11.3. $n \equiv 2 \mod 3$, $n \geq 20$ or $n = 11$.

If $n = 11$, the graph J suffices. If $n \geq 20$, then take J and a construction from Case 4.3.11.1. This achieves $2n + 2$ edges.

For the lower bound, let G be any $K_{1,3}$-induced-saturated graph. Corollary 4.3.3 gives us a general lower bound of $2n - 2$. Suppose G has no isolated vertex. Then by Corollary 4.3.3, $e(G) \geq 2n$, as desired. Suppose then that G does have an isolated vertex, and $n \not\equiv 1 \mod 3$. Then $(n - 1) \not\equiv 0 \mod 3$, so by Lemmas 4.3.2 and 4.3.7, the minimum degree of the non-isolated vertices is at least 4, and $\Delta(G) \geq 5$. Then $e(G) \geq \lceil \frac{4(n-1)+1}{2} \rceil = 2n - 1$, with equality only if the degree sequence of G is $(5, 5, 4, \ldots, 4, 0)$ or $(6, 4, \ldots, 4, 0)$. Since the graph obtained by deleting the isolate is $K_{1,3}$-induced-saturated, by Propositions 4.3.8 and 4.3.9, $e(G) \geq 2n$. □

4.4   Families of graphs

In this section we extend the definition of induced saturation to families of graphs in the natural way.

Definition 4.4.1. For a family $\mathcal{F}$ of graphs, a trigraph T is $\mathcal{F}$-*induced-saturated* if no realization of T contains any member of $\mathcal{F}$ as an induced subgraph, but whenever any black or white edge of T is turned to gray, some member of $\mathcal{F}$ occurs as an induced subgraph of some realization.

The *induced saturation number* of $\mathcal{F}$ with respect to n, written $\text{indsat}(n, \mathcal{F})$, is the minimum number of gray edges in an $\mathcal{F}$-induced-saturated trigraph with n vertices.

For any family $\mathcal{F}$ containing all graphs on k vertices, $\text{indsat}(n, \mathcal{F}) = \binom{n}{2}$.

Construction 4.4.2 and Proposition 4.4.3 demonstrate that for any family $\mathcal{F}$, all of whose elements are odd cycles, even cycles with a pendant, or even cycles with a triangle chord, $\text{indsat}(n, \mathcal{F}) = 0$ for n sufficiently large. However, we could have $\text{indsat}(n, \mathcal{F}) \neq 0$ even if there is some $G \in \mathcal{F}$ such that $\text{indsat}(n, G) = 0$ as demonstrated in Proposition 4.4.4 below. One may suspect this is because of the presence of P_4, which has nonzero induced-saturation number, yet it is also possible for a family $\mathcal{F}$ to consist of graphs that each individually have induced saturation number zero, while the induced saturation number of $\mathcal{F}$ is nonzero. We provide an example of this in Proposition 4.4.5.

Construction 4.4.2 ([7])**.** For $k \geq 3$ and $n \geq (k+1)^2 + 2$, let $n = (k+1)t - s$, where $t = \left\lceil \frac{n}{k+1} \right\rceil \geq k+2$ and $0 \leq s \leq t - 3$. Let $G_{n,k}$ be formed from the Cartesian product $K_{k+1} \,\square\, K_t$ by removing s vertices from one copy of K_t.

Proposition 4.4.3 ([7])**.** *Let C'_{2k} denote a cycle of length $2k$ with a pendant vertex, and $\hat{C}_{2k}$ denote an even cycle with a chord between two vertices at distance 2 from*

each other. If $H \in \{C_{2k-1}, C'_{2k}, \hat{C}_{2k}\}$, then the graph $G_{n,k}$ in Construction 4.4.2 is H-induced-saturated.

Proposition 4.4.4. *For all n,* $\mathrm{indsat}(n, \{2K_2, P_4, C_4\}) \neq 0$.

Proof. The graphs that contain no induced $2K_2, P_4$, or C_4 are precisely the *threshold graphs* [28]. These graphs are characterized in a second way: they are constructed by iteratively adding a vertex to a graph either as an isolate or a dominating vertex. Thus, an n-vertex threshold graph can be represented as a string of n symbols from $\{-, +\}$ as follows: on the vertex set $V = \{v_1, \ldots, v_n\}$, for every $i > j$, $v_i v_j$ is an edge if and only if the ith symbol in the string is $+$.

We claim that for any threshold graph G with at least one edge, there exists $e \in E(G)$ such that $G - e$ is also threshold. Let $\pi = s_1, \ldots, s_n$ be a string of symbols from $\{-, +\}$ representing G. Suppose there exists $i \in [n-1]$ such that $s_i = -$ and $s_{i+1} = +$, and let i be minimal with this property. Then the graph $G' = G - v_i v_{i+1}$ is represented by the symbol list $\pi' = s_1 \ldots s_{i-1} s_{i+1} s_i s_{i+1} \ldots s_n$, so G' is threshold. If no such index i exists, then π is a list consisting only of $+$, so G is the complete graph K_n; however, $K_n - e$ is also threshold.

Thus, for any graph G with no induced $2K_2, P_4$, or C_4, there exists an edge $e \in G$ such that $G - e$ also has no induced $2K_2, P_4$, or C_4. It follows that $\mathrm{indsat}(n, \{2K_2, P_4, C_4\}) \neq 0$. $\qquad\square$

The family of *split graphs* is another family of graphs that can be characterized by a set of forbidden induced subgraphs. A split graph is a graph whose vertex set can be partitioned into a clique and an independent set. Földes and Hammer [38] showed that a graph is a split graph if and only if it contains no induced $2K_2$, C_4, or C_5.

Proposition 4.4.5. *For all n,* $\mathrm{indsat}(n, \{2K_2, C_4, C_5\}) \neq 0$.

Proof. Since adding or deleting an edge between the clique part and the independent set of a split graph still results in a split graph, it follows that $\mathrm{indsat}(n, \{2K_2, C_4, C_5\}) \neq 0$. $\qquad\square$

We have shown that $\mathrm{indsat}(n, 2K_2)$, $\mathrm{indsat}(n, C_4)$, and $\mathrm{indsat}(n, C_5)$ are all equal to zero for sufficiently large n. Thus, this example shows that even though every graph in a family has induced-saturation number zero, the family itself may not have induced-saturation number zero.

Other families characterized by a (not necessarily finite) family of forbidden induced subgraphs include perfect graphs [26], trivially perfect graphs [85], [42], interval graphs [62], and line graphs [8]. It would be interesting to determine $\mathrm{indsat}(n, \mathcal{F})$ and $\mathrm{indsat}^*(n, \mathcal{F})$ for these families. We suspect that doing so will be much more difficult than for threshold and split graphs, as the families of forbidden graphs are significantly more complicated.

Bibliography

[1] Nirmala Achuthan, Narasimaha R. Achuthan, and Mudin Simanihuruk. On 3-uniform hypergraphic sequences. *J. Combin. Math. Combin. Comput.*, 14:3–13, 1993.

[2] Martin Aigner and Eberhard Triesch. Realizability and uniqueness in graphs. *Discrete Math.*, 136(1-3):3–20, 1994. Trends in discrete mathematics.

[3] Kenneth Appel, Wolfgang Haken, and John Koch. Every planar map is four colorable. II. Reducibility. *Illinois J. Math.*, 21(3):491–567, 1977.

[4] Kannan Balakrishnan, Boštjan Brešar, Manoj Changat, Sandi Klavžar, Iztok Peterin, and Ajitha R. Subhamathi. Almost self-centered median and chordal graphs. *Taiwanese J. Math.*, 16(5):1911–1922, 2012.

[5] Michael Barrus, Stephen G. Hartke, Kyle Jao, and Douglas West. Length thresholds for graphic lists given fixed largest and smallest entries and bounded gaps. *Discrete Math.*, 312(9):1494–1501, 2012.

[6] Sarah Behrens, Catherine Erbes, Michael Ferrara, Stephen G. Hartke, Benjamin Reiniger, Hannah Spinoza, and Charles Tomlinson. New results on degree sequences of uniform hypergraphs. *Electron. J. Combin.*, 20(4):Paper 14, 18, 2013.

[7] Sarah Behrens, Catherine Erbes, Michael Santana, Derrek Yager, and Elyse Yeager. Graphs with induced-saturation number zero. *ArXiv e-prints*, March 2015.

[8] Lowell W. Beineke. Characterizations of derived graphs. *J. Combinatorial Theory*, 9:129–135, 1970.

[9] Annabell Berger. The connection between the number of realizations for degree sequences and majorization. *CoRR*, abs/1212.5443, 2012.

[10] Nirupama S. Bhave, Bhagyashri Y. Bam, and Charusheela M. Deshpande. A characterization of degree sequences of linear hypergraphs. *J. Indian Math. Soc. (N.S.)*, 75(1-4):151–160 (2009), 2008.

[11] Halina Bielak. Minimal realizations of graphs as central subgraphs. In *Graphs, hypergraphs and matroids (Żagań, 1985)*, pages 13–24. Higher College Engrg., Zielona Góra, 1985.

[12] David Billington. Lattices and degree sequences of uniform hypergraphs. *Ars Combin.*, 21(A):9–19, 1986. Thirteenth Australasian conference on combinatorial mathematics and computing (Sydney, 1985).

[13] David Billington. Conditions for degree sequences to be realizable by 3-uniform hypergraphs. *J. Combin. Math. Combin. Comput.*, 3:71–91, 1988.

[14] Béla Bollobás. The diameter of random graphs. *Trans. Amer. Math. Soc.*, 267(1):41–52, 1981.

[15] Fred Buckley. The central ratio of a graph. *Discrete Math.*, 38(1):17–21, 1982.

[16] Fred Buckley. Self-centered graphs. In *Graph theory and its applications: East and West (Jinan, 1986)*, volume 576 of *Ann. New York Acad. Sci.*, pages 71–78. New York Acad. Sci., New York, 1989.

[17] Fred Buckley and Frank Harary. *Distance in graphs*. Addison-Wesley Publishing Company, Advanced Book Program, Redwood City, CA, 1990.

[18] Fred Buckley, Zevi Miller, and Peter J. Slater. On graphs containing a given graph as center. *J. Graph Theory*, 5(4):427–434, 1981.

[19] Arthur Busch, Michael Ferrara, Stephen Hartke, Michael Jacobson, Hemanshu Kaul, and Douglas West. Packing of graphic n-tuples. *J. Graph Theory*, 70(1):29–39, 2012.

[20] Sergio Cabello and Günter Rote. Obnoxious centers in graphs. *SIAM J. Discrete Math.*, 24(4):1713–1730, 2010.

[21] Hatice Calik and Barbaros C. Tansel. Double bound method for solving the p-center location problem. *Computers & Operations Research*, 40(12):2991 – 2999, 2013.

[22] Gerard J. Chang. Centers of chordal graphs. *Graphs Combin.*, 7(4):305–313, 1991.

[23] Gary Chartrand, Garry L. Johns, Song Lin Tian, and Steven J. Winters. Directed distance in digraphs: centers and medians. *J. Graph Theory*, 17(4):509–521, 1993.

[24] Zhibo Chen. On the problem of the existence of a graph H with $A(H) = 3$. *J. Math. (Wuhan)*, 4(3):267–272, 1984.

[25] Sheshayya Choudum. On graphic and 3-hypergraphic sequences. *Discrete Math.*, 87(1):91–95, 1991.

[26] Maria Chudnovsky, Neil Robertson, Paul Seymour, and Robin Thomas. The strong perfect graph theorem. *Ann. of Math. (2)*, 164(1):51–229, 2006.

[27] Maria Chudnovsky and Paul Seymour. The structure of claw-free graphs. In *Surveys in combinatorics 2005*, volume 327 of *London Math. Soc. Lecture Note Ser.*, pages 153–171. Cambridge Univ. Press, Cambridge, 2005.

[28] Václav Chvátal and Peter L. Hammer. Aggregation of inequalities in integer programming. In *Studies in integer programming (Proc. Workshop, Bonn, 1975)*, pages 145–162. Ann. of Discrete Math., Vol. 1. North-Holland, Amsterdam, 1977.

[29] Charles J. Colbourn, William L. Kocay, and Douglas R. Stinson. Some NP-complete problems for hypergraph degree sequences. *Discrete Appl. Math.*, 14(3):239–254, 1986.

[30] Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein. *Introduction to algorithms.* MIT Press, Cambridge, MA, third edition, 2009.

[31] Alexander K. Dewdney. Degree sequences in complexes and hypergraphs. *Proc. Amer. Math. Soc.*, 53(2):535–540, 1975.

[32] Christoph Dürr, Flavio Guiñez, and Martín Matamala. Reconstructing 3-colored grids from horizontal and vertical projections is NP-hard. In *Algorithms—ESA 2009*, volume 5757 of *Lecture Notes in Comput. Sci.*, pages 776–787. Springer, Berlin, 2009.

[33] Paul Erdős and Tibor Gallai. Graphs with prescribed degrees of vertices (hungarian). *Mat. Lapok*, 11:264–274, 1960.

[34] Paul Erdős, András Hajnal, and John W. Moon. A problem in graph theory. *Amer. Math. Monthly*, 71:1107–1110, 1964.

[35] Péter L. Erdős, István Miklós, and Zoltán Toroczkai. A simple Havel-Hakimi type algorithm to realize graphical degree sequences of directed graphs. *Electron. J. Combin.*, 17(1):Research Paper 66, 10, 2010.

[36] Jill R. Faudree, Ralph J. Faudree, and John R. Schmitt. A survey of minimum saturated graphs. *Electron. J. Combin.*, Dynamic Survey 18, 2011.

[37] Robert W. Floyd. Algorithm 97: Shortest path. *Commun. ACM*, 5(6):345–, June 1962.

[38] Stéphanie Földes and Peter L. Hammer. On split graphs and some related questions. In *Problémes Combinatoires et Théorie Des Graphes*, volume 260 of *Colloques internationaux C.N.R.S.*, pages 138–140, Orsay, France, 1976.

[39] Delbert R. Fulkerson, Alan J. Hoffman, and Michael H. McAndrew. Some properties of graphs with multiple edges. *Canad. J. Math.*, 17:166–177, 1965.

[40] Irving Gabelman. *The functional behavior of majority (threshold) elements*. PhD thesis, Syracuse U, 1961.

[41] GLPK (GNU linear programming kit), version 4.44, 2015. `http://www.gnu.org/software/glpk`.

[42] Martin Charles Golumbic. Trivially perfect graphs. *Discrete Math.*, 24(1):105–107, 1978.

[43] Catherine S. Greenhill. The switch markov chain for sampling irregular graphs. *CoRR*, abs/1412.5249, 2014.

[44] Seifollah L. Hakimi. On realizability of a set of integers as degrees of the vertices of a linear graph. I. *J. Soc. Indust. Appl. Math.*, 10:496–506, 1962.

[45] Thomas C. Hales. A proof of the Kepler conjecture. *Ann. of Math. (2)*, 162(3):1065–1185, 2005.

[46] Frank Harary and Phillip A. Ostrand. The cutting center theorem for trees. *Discrete Math.*, 1(1):7–18, 1971/72.

[47] Václav Havel. A remark on the existence of finite graphs (czech). *Časopis Pěst. Mat.*, 80:477–480, 1955.

[48] Sandra Mitchell Hedetniemi, Stephen T. Hedetniemi, and Peter J. Slater. Centers and medians of $C_{(n)}$-trees. *Utilitas Math.*, 21:225–234, 1982.

[49] IBM ILOG CPLEX Optimizer, version 12.6.0.0. `http://www-01.ibm.com/software/integration/optimization/cplex-optimizer/`, Last 2015.

[50] Thenkarai N. Janakiraman. On self-centered graphs. *J. Ramanujan Math. Soc.*, 7(1):83–92, 1992.

[51] Mark Jerrum and Alistair Sinclair. Fast uniform generation of regular graphs. *Theoretical Computer Science*, 73(1):91 – 100, 1990.

[52] Ravi Kannan, Prasad Tetali, and Santosh Vempala. Simple Markov-chain algorithms for generating bipartite graphs and tournaments. *Random Structures Algorithms*, 14(4):293–308, 1999.

[53] Petteri Kaski and Olli Pottonen. libexact users guide, version 1.0, 2008. `http://koti.kapsi.fi/pottonen/libexact.html`.

[54] László Kászonyi and Zsolt Tuza. Saturated graphs with minimal number of edges. *J. Graph Theory*, 10(2):203–210, 1986.

[55] Sandi Klavžar, Kishori P. Narayankar, and Hanumappa B. Walikar. Almost self-centered graphs. *Acta Math. Sin. (Engl. Ser.)*, 27(12):2343–2350, 2011.

[56] Daniel J. Kleitman and Da-Lun Wang. Algorithms for constructing graphs and digraphs with given valences and factors. *Discrete Math.*, 6:79–88, 1973.

[57] William Kocay and Pak Ching Li. On 3-hypergraphs with equal degree sequences. *Ars Combin.*, 82:145–157, 2007.

[58] Georgy N. Kopylov and Eugeniy A. Timofeev. Centers and radii of graphs. *Uspehi Mat. Nauk*, 32(6(198)):226, 1977.

[59] Alexandr Kostochka, Christopher Stocker, and Peter Hamburger. A hypergraph version of a graph packing theorem by Bollobás and Eldridge. *J. Graph Theory*, 74(2):222–235, 2013.

[60] Clement W. H. Lam, Larry Thiel, and Stanley Swiercz. The nonexistence of finite projective planes of order 10. *Canad. J. Math.*, 41(6):1117–1123, 1989.

[61] Renu Laskar and Douglas Shier. On powers and centers of chordal graphs. *Discrete Appl. Math.*, 6(2):139–147, 1983.

[62] Cornelis G. Lekkerkerker and Jan Ch. Boland. Representation of a finite graph by a set of intervals on the real line. *Fund. Math.*, 51:45–64, 1962/1963.

[63] Xiang Wu Liu. A class of graphs with $\alpha(H) = 4$ that are not trees. *Natur. Sci. J. Harbin Normal Univ.*, 9(3):1–5, 1993.

[64] Xiang Wu Liu. Sufficient conditions for a graph to satisfy $a(H) = 3$. *Natur. Sci. J. Harbin Normal Univ.*, 9(2):7–12, 1993.

[65] Xiang Wu Liu. Classifications of a class of graphs H according to the parameter $a(H)$. *Heilongjiang Daxue Ziran Kexue Xuebao*, 13(4):10–12, 1996.

[66] Xiang Wu Liu. Determination of the value of $a(H)$ for a class of graphs by longitude and latitude of vertices. *Natur. Sci. J. Harbin Normal Univ.*, 14(6):1–4, 1998.

[67] Xiang Wu Liu and Yu Xin He. Two classes of graphs satisfying $A(H) = 3$. *J. Math. (Wuhan)*, 9(4):423–430, 1989.

[68] Nadimpalli V. R. Mahadev and Uri N. Peled. *Threshold graphs and related topics*, volume 56 of *Annals of Discrete Mathematics*. North-Holland Publishing Co., Amsterdam, 1995.

[69] Ryan R. Martin and Jason J. Smith. Induced saturation number. *Discrete Math.*, 312(21):3096–3106, 2012.

[70] Brendan D. McKay and Adolfo Piperno. Practical graph isomorphism, II. *J. Symbolic Comput.*, 60:94–112, 2014.

[71] Julius Petersen. Die Theorie der regulären graphs. *Acta Math.*, 15(1):193–220, 1891.

[72] Monika Pilśniak and Mariusz Woźniak. A note on packing of two copies of a hypergraph. *Discuss. Math. Graph Theory*, 27(1):45–49, 2007.

[73] Monika Pilśniak and Mariusz Woźniak. On packing of two copies of a hypergraph. *Discrete Math. Theor. Comput. Sci.*, 13(3):67–74, 2011.

[74] Ruth Pordes, Don Petravick, Bill Kramer, Doug Olson, Miron Livny, Alain Roy, Paul Avery, Kent Blackburn, Torre Wenaus, Frank Wrthwein, Ian Foster, Rob Gardner, Mike Wilde, Alan Blatecky, John McGee, and Rob Quick. The Open Science Grid. *Journal of Physics: Conference Series*, 78(1):012057, 2007.

[75] Andrzej Proskurowski. Centers of 2-trees. *Ann. Discrete Math.*, 9:1–5, 1980. Combinatorics 79 (Proc. Colloq., Univ. Montréal, Montreal, Que., 1979), Part II.

[76] Jaideep Ray, Ali Pinar, and Comandur Seshadhri. Are we there yet? When to stop a Markov chain while generating random graphs. In Anthony Bonato and Jeannette Janssen, editors, *Algorithms and Models for the Web Graph*, volume 7323 of *Lecture Notes in Computer Science*, pages 153–164. Springer Berlin Heidelberg, 2012.

[77] Neil Robertson, Daniel Sanders, Paul Seymour, and Robin Thomas. The four-colour theorem. *J. Combin. Theory Ser. B*, 70(1):2–44, 1997.

[78] Valerio Schiavoni. A fast Java implementation of Floyd-Warshall. `http://www.jroller.com/vschiavoni/entry/a_fast_java_implementation_of`, November 25, 2008. (Archived by WebCite at `http://www.webcitation.org/6Pjp76C70`).

[79] Gerard Sierksma and Han Hoogeveen. Seven criteria for integer sequences being graphic. *J. Graph Theory*, 15(2):223–231, 1991.

[80] Peter J. Slater. Some definitions of central structures. In KheeMeng Koh and HianPoh Yap, editors, *Graph Theory Singapore 1983*, volume 1073 of *Lecture Notes in Mathematics*, pages 169–178. Springer Berlin Heidelberg, 1984.

[81] William A. Stein et al. *Sage Mathematics Software (Version 5.8)*. The Sage Development Team, 2015. `http://www.sagemath.org`.

[82] Paul Turán. Eine Extremalaufgabe aus der Graphentheorie. *Mat. Fiz. Lapok*, 48:436–452, 1941.

[83] Stephen Warshall. A theorem on boolean matrices. *J. Assoc. Comput. Mach.*, 9:11–12, 1962.

[84] Douglas B. West. *Introduction to graph theory*. Prentice Hall, Inc., Upper Saddle River, NJ, 2nd edition, 2001.

[85] Elliot S. Wolk. The comparability graph of a tree. *Proc. Amer. Math. Soc.*, 13:789–795, 1962.

[86] Jian-Hua Yin and Jiong-Sheng Li. Two sufficient conditions for a graphic sequence to have a realization with prescribed clique size. *Disc. Math.*, 301:218–227, 2005.

[87] Zhen Guo Zhang, Huan Xing, and Xiang Wu Liu. An infinite class of graphs with $A(H) = 3$. *Natur. Sci. J. Harbin Normal Univ.*, 19(6):11–12, 21, 2003.