ON CONNECTEDNESS AND GRAPH POLYNOMIALS

by

Lucas Mol

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

at

Dalhousie University
Halifax, Nova Scotia
April 2016

Table of Contents

List of Tables

List of Figures

Abstract

Given a graph G whose nodes are perfectly reliable and whose edges fail independently with probability $q \in [0,1]$, the *all-terminal reliability* of G is the probability that all vertices of G can communicate with one another. The all-terminal reliability is a polynomial in q whose roots (*all-terminal reliability roots*) were conjectured to have modulus at most 1 by Brown and Colbourn. This conjecture was proven false by Sokal and Royle, but only by a slim margin. We present an upper bound on the modulus of any all-terminal reliability root in terms of the number of vertices of the graph. We find all-terminal reliability roots of greater modulus than any previously known, and we study simple graphs with all-terminal reliability roots of modulus greater than 1.

Given a graph G whose edges are perfectly reliable and whose nodes each operate independently with probability $p \in [0,1]$, the *node reliability* of G is the probability that at least one node is operational and that the operational nodes can all communicate in the subgraph that they induce. We explore analytic properties of the node reliability on the interval $[0,1]$ including monotonicity, concavity, and fixed points. Our results demonstrate a stark contrast between this model of network robustness and models that arise from coherent set systems (including all-terminal reliability).

A *connected set* of a graph G is a nonempty subset of vertices of G that induces a connected subgraph. The *connected set polynomial* of G is the generating polynomial of the collection of connected sets of G. The computational complexity, and the nature and location of the roots of the connected set polynomial are investigated. Our results have direct implications for node reliability. Further, we consider the connected set polynomials of trees – the total number of *subtrees* of a tree has recently garnered much interest and the connected set polynomial extends this notion.

List of Abbreviations and Symbols Used

Graph theory

$E(G)$ .... edge set of graph G

$V(G)$ .... vertex set of graph G

$N(v)$ .... open neighbourhood of vertex v

$N[v]$ .... closed neighbourhood of vertex v

$G[W]$ .... induced subgraph of G on the subset W of vertices

$G - W$ .... graph obtained from G by removing all vertices in W

$G - v$ .... graph obtained from G by deleting vertex v

$G \cong H$ .... graphs G and H are isomorphic

$G \cup H$ .... disjoint union of graphs G and H

$G + H$ .... join of graphs G and H

$H(u, v)$ .... graph H with special vertices u and v (also called a gadget)

$G[H(u, v)]$ .... edge substitution of the gadget $H(u, v)$ into G

$G \odot H$ .... lexicographic product of G and H

Graph polynomials

$\mathrm{Rel}(G; q)$ .... all-terminal reliability of G

$F(G; x)$ .... generating polynomial for the F-coefficients of all-terminal reliability

$H(G; x)$ .... generating polynomial for the H-coefficients of all-terminal reliability

$T(G; x, y)$ .... Tutte polynomial of G

$M_w(G; x)$ .... generating polynomial of the order ideal of monomials $\mathcal{M}_w(G)$

$\mathrm{Rel}_K(G; p)$ .... K-terminal reliability of G

$\mathrm{nRel}(G; p)$ .... node reliability of G

$\mathrm{spRel}(G; p)$ .... K-split reliability of G

$C(G;x)$ .... connected set polynomial of G

$C_v(G;x)$ .... v-rooted connected set polynomial of G

$g(G;x)$ .... the g-convexity polynomial of graph G

Named graphs

K_n .... complete graph on n vertices

O_n .... empty graph on n vertices

$K_{m,n}$ .... complete bipartite graph with bipartition sets of size m and n

C_n .... cycle on n vertices

P_n .... path on n vertices

K_n^- .... complete graph on n vertices minus an edge

$G_{m,n}^{a,b}$ .... see the discussion preceding Proposition 2.2.2

$G^{(k,n)}$ .... the graph $G_{3,3}^{k,6k}[K_n^-(u,v)]$ with u and v nonadjacent in K_n^-

$P_{n,k}$ .... path on $n-1$ vertices with a leaf added to internal vertex k

Miscellaneous

ATR .... all-terminal reliability

gcd .... greatest common divisor

$\mathrm{B}(a_1,\ldots,a_n)$ .... number of sign changes in the sequence $(a_1,\ldots,a_n)$

$\overline{a}$ .... complex conjugate of complex number a

A^* .... conjugate transpose of complex matrix A

$\mathcal{M}_w(G)$ .... order ideal of monomials for the chip-firing game on G

Acknowledgements

First of all I would like to thank my supervisor, Dr. Jason Brown. Jason's advice and guidance throughout the process was invaluable. His help in motivating and framing the research, as well as his help in honing the results, was absolutely crucial.

Next I would like to thank my committee members Dr. Jeannette Janssen and Dr. Richard Nowakowski. Their careful reading of the thesis was much appreciated. Their feedback and questions along the way have been most helpful and have made the thesis stronger.

I owe special thanks to my external examiner, Dr. David Pike, for making an extra effort to attend the external defence in person after 120km/h winds caused his scheduled flight to be cancelled. His comments and questions on the thesis and at the defence were much appreciated.

I'd like to thank everyone I interacted with on a regular basis in the Chase Building, including faculty, staff, and my fellow graduate students. In particular, I'd like to thank Ellen, Maria, Paula, and Queena in the office. I won't single out any faculty or graduate students as there are too many to list and I will inevitably miss someone. Suffice it to say that the Chase Building was a wonderful place to work largely because of the people in it. Thank you all and all the best for the future!

Finally, I'd like to thank a few people on a more personal level. My parents have been very supportive of my schooling from the start, and I appreciate that they'll be at my graduation (hopefully my sisters too – Leah, Jacoba, and Katelin). My girlfriend Rowan has been an incredible support throughout the writing of the thesis and I thank her from the bottom of my heart. I'd also like to thank Rowan's family – Mary, Heather, and Poppy – for making me feel at home in Halifax.

Chapter 1

Introduction

Networks play an increasingly important role in modern life; we rely on electrical networks, transportation networks, and social networks, to name but a few. In most networks, there are components which are not always perfectly reliable; that is, sometimes parts of the network fail. Power lines are torn down in a storm and homes lose power, a car accident occurs and closes down a road temporarily, or a Facebook user is offline and unable to receive a message.

Since the occasional failure of some components is practically unavoidable, we would like our networks to be able to function properly despite the failure of some components. One way to measure the robustness of a network is to assume that the components fail randomly, and to assign a probability of failure to each component of the network. The *reliability* of a network is then the probability that the network functions adequately. Of course, there are many different notions of "functioning adequately", each giving rise to a different notion of reliability. If we consider the power grid of a city we would like power to be delivered to all homes. In a public transportation network we may want to ensure that riders can travel between all of the major terminals. In a social network we may want to make sure that everyone who is interested in attending an event receives an invitation.

For us, a network (or *graph*) consists of a finite set of *vertices* or *nodes*, and a set of *edges* that represent links between pairs of nodes. For certain networks it is more likely that the edges will fail while the nodes remain operational – take for example an electrical network where a signal is to be passed from one node to another. Physical wires are easily damaged and may not be able to pass the signal as intended. In other cases the edges are very reliable while the nodes frequently fail – cell phone reception has become very reliable in recent years, but phone calls are often missed simply because the phones themselves are switched to silent, or because the intended receiver is unable to answer. Generally in networks where edges represent wireless links, node failure is more common than edge failure.

In this thesis we focus mainly on two different notions of reliability. For the first model we assume that nodes are perfectly reliable while edges fail with a given probability. While in general the edges could have different probabilities of failure and have failure dependencies, we make the simplifying assumption that all edges fail independently with probability $q \in [0,1]$ (and *operate* – that is, perform properly – with probability $p = 1 - q$). The *all-terminal reliability* is the probability that at least a spanning tree is operational. In other words, it is the probability that all nodes can communicate with one another. For the second model we assume that the *edges* are perfectly reliable while the *nodes* fail with a given probability. Again, we make the simplifying assumption that all nodes fail independently with probability $q \in [0,1]$ and operate with probability $p = 1 - q$. The *node reliability* is the probability that at least one node is operational and that all of the operational nodes can communicate with one another. We will see that the assumption that all components fail independently with the same probability q implies that both all-terminal reliability and node reliability are polynomials (in q or p, depending on the perspective that we take).

As a simple example, we compute the all-terminal reliability and the node reliability of the cycle C_4 on 4 vertices, pictured in Figure 1.1. For the all-terminal reliability, the edges operate independently with probability p and fail with probability $1 - p$. All nodes can communicate with one another if and only if either all 4 edges are operational or exactly 3 of the 4 edges are operational. All 4 edges are

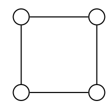


Figure 1.1: The cycle C_4.

simultaneously operational with probability p^4, while any particular set of 3 edges is operational and the other edge fails with probability $p^3(1-p)$. There are four different subsets containing 3 edges, so the all-terminal reliability of C_4 (the probability that all nodes can communicate) is given by

$$\mathrm{Rel}(C_4; p) = p^4 + 4p^3(1-p).$$

For node reliability, the *nodes* operate independently with probability p and fail with probability $1-p$. We determine the states in which at least one node is operational and all of the operational nodes can communicate. There are 4 such states in which exactly one node is operational (corresponding to the vertices), 4 such states in which exactly two nodes are operational (corresponding to the edges), 4 such states in which exactly three nodes are operational (corresponding to vertices again, as any one vertex can fail while the others remain operational), and finally 1 such state in which all four nodes are operational. The probability that a particular set of k nodes is operational while the remaining $4-k$ nodes fail is $p^k(1-p)^{4-k}$. Therefore, the node reliability of C_4 is given by

$$\mathrm{nRel}(C_4; p) = p^4 + 4p^3(1-p) + 4p^2(1-p)^2 + 4p(1-p)^3.$$

Shier [55] states that the two primary objectives of reliability theory are to assess the reliability of systems (this is the *analysis* side of reliability theory), and to design the most reliable system (if possible) from given components (this is the *synthesis* side of reliability theory). The analysis of reliability includes the study of the algorithmic complexity of computing the reliability (i.e. computing all coefficients of the polynomial), polynomial time algorithms for computing the reliability for restricted families of graphs, efficient bounding procedures, and analytic properties of the functions. Analytic properties of interest include the shape of the reliability on the interval $[0, 1]$, as well as the nature and location of the roots in the complex plane. Meanwhile, synthesis usually involves trying to find a *uniformly best graph* in a given class; that is, a graph whose reliability is the largest in the class for all $p \in [0, 1]$. The classes of all graphs on a fixed number of vertices and edges have been studied most often.

For all-terminal reliability, much work has been done in all of the topics mentioned in the previous paragraph – see [28], for example, or [11] for a more recent survey. In this thesis our focus is on the roots of all-terminal reliability polynomials. It was conjectured in [12] that the roots of all-terminal reliability polynomials lie in a disk of unit radius in the complex plane. Despite some results which confirmed the conjecture for particular families of graphs [25, 68], the conjecture was proven false in [54], but only by a slim margin. We prove the first general upper bound on the modulus of a root of an all-terminal reliability polynomial of a graph of order n (in terms of n), and we also find roots further outside the conjectured unit disk than any previously known.

Node reliability has garnered considerably less attention in the literature than all-terminal reliability. On the analysis side, it has been shown that the problem of computing the sequence of coefficients of the node reliability polynomial is #P-complete, even for the graphs that are both planar and bipartite [62]. On the other hand, polynomial time algorithms for computing the node reliabilities of certain restricted families of graphs including trees and series-parallel graphs have been found in [29]. Work on the synthesis side on the existence and identification of uniformly best graphs in particular classes has been done in [34, 48, 74]. In this thesis we undertake the first in-depth study of the analytic properties of node reliability polynomials. We study the shape of node reliability on the interval $[0, 1]$ including monotonicity, concavity, and fixed points. Our results will often demonstrate sharp contrasts between all-terminal reliability and node reliability, which are surprising given the similarity of the formulations.

We then turn our attention to the roots of node reliability, and our deepest results on the roots of node reliability are achieved through the *connected set polynomial*. A *connected set* of a graph G is a nonempty subset of vertices of G that induces a connected subgraph. The connected set polynomial of G is the generating polynomial of the collection of connected sets of G. This places the connected set polynomial in the same general context as other generating polynomials of graphs including the matching polynomial, the independence polynomial, and the domination polynomial. Much research concerning the computational complexity and the combinatorial and analytic properties of polynomials exists in the literature. It will become evident that the connected set polynomial is closely related to the node reliability polynomial, and so our results on the connected set polynomial will often have immediate implications for node reliability. We study the computational complexity and the nature and location of the roots of the connected set polynomial in general.

The connected set polynomial of a tree T counts the number of *subtrees* of T, and hence we will sometimes call it the *subtree polynomial* of T. The number of subtrees of a tree has recently received much attention in the literature; it has applications to a broad range of topics including combinatorial chemistry and phylogeny. The search for a tree in a given class with the maximum or minimum total number of subtrees has been of primary interest, and the subtree polynomial provides a natural way to extend some results in this direction. Additionally, we determine the trees for which the sequence of coefficients of the connected set polynomial is unimodal, and also those for which it is log-concave. Finally, we study the roots of connected set polynomials of trees, as their location in the complex plane appears to be very special.

Before moving on to the background material for the thesis, we make a final note about the study of roots of graph polynomials in general. The roots of various graph polynomials including the independence polynomial, the matching polynomial, the chromatic polynomial, and the domination polynomial have been studied extensively. The chromatic polynomial $P(G;\lambda)$ is a prime example. In fact, the chromatic polynomial was introduced by Birkhoff in [5] with the goal of demonstrating that 4 is not a root of $P(G;\lambda)$ for any planar graph G (which implies the Four Colour Theorem). Hence *chromatic roots* (the roots of chromatic polynomials) have been important since the inception of the chromatic polynomial. Many results on chromatic roots have since been proven. For the real roots, it is well known that $(-\infty, 0)$ and $(0,1)$ are maximal root-free intervals of the chromatic polynomial, and Jackson proved that $\left(1, \frac{32}{27}\right]$ is another maximal root-free interval [43]. Finally, Thomassen proved that any interval (λ_1, λ_2) with $\frac{32}{27} \leq \lambda_1 < \lambda_2$ contains a chromatic root [65], concluding the search for root-free intervals. There are also many significant results on complex chromatic roots. For example, Sokal achieved a proof that if λ is a chromatic root of G, then $|\lambda| < 8\Delta$ where Δ is the maximum degree of G [57], and also proved that the collection of all chromatic roots is dense in the complex plane [58]. In short, the roots of various graph polynomials have garnered considerable attention in the literature, and so our study of the roots of all-terminal reliability, node reliability, and the connected set polynomial is completely natural.

1.1 Background

Before we proceed with our new results, we provide some general background on graph theory. The material in this section can be found in any introductory text on graph theory (see [10, 70], for example).

A *(multi)graph* G is an ordered pair (V, E), where V is a set and E is a multiset whose members are unordered pairs from V. The members of V are called *vertices*, and the members of E are called *edges*. A graph is called *finite* if it has a finite number of vertices and edges. We stress that our definition of graph allows multiple edges but not loops. A *simple graph* is a graph with no multiple edges. For any graph G, we let $V(G)$ denote the set of vertices of G and we let $E(G)$ denote the (multi)set of edges of G. The *order* of G is $|V(G)|$ and the *size* of G is $|E(G)|$. An edge $\{u, v\} \in E(G)$ is said to *join* the vertices u and v. We say that u and v are the *endvertices* of the edge $\{u, v\}$. Two vertices x and y are called *adjacent* if they are joined by an edge. A vertex x is *incident* to an edge e exactly when x is an endvertex of e. Two edges are *incident* if they have an endvertex in common.

Adjacent vertices are also called *neighbours*. The (*open*) *neighbourhood* of a vertex v is the set of all neighbours of v, denoted $N_G(v)$ (or $N(v)$ when the context is clear). The *closed neighbourhood* $N_G[v]$ (or $N[v]$) of a vertex v is the set of all neighbours of v together with v itself; that is, $N[v] = N(v) \cup \{v\}$. The *degree* of vertex v is the number of edges incident to v, denoted $\deg(v)$. A vertex of degree 0 is called an *isolated vertex*, while a vertex of degree 1 is called a *leaf*. An edge incident to a leaf is called a *pendant edge*.

Let $G = (V, E)$ be a graph. We say that a graph $G' = (V', E')$ is a *subgraph* of G if $V' \subseteq V$ and $E' \subseteq E$. The *induced subgraph* of G on vertex set $V' \subseteq V$, denoted $G[V']$, is the subgraph containing *all* edges of E that join two vertices of V'. A subset C of vertices of a graph G is called a *clique* if $G[C]$ is complete.

We often construct new graphs from old graphs by removing certain vertices. For a subset $W \subseteq V$, we define $G - W = G[V \backslash W]$, the graph obtained by removing all vertices in W and all edges incident with a vertex in W. We write $G - v$ instead of $G - \{v\}$.

Let G be a graph with vertex set V. A subset $W \subseteq V$ is called *independent* if no two members of W are adjacent in G. A graph is called *bipartite* if its vertex set can be partitioned into at most two disjoint independent sets. The independent sets in such a partition are called *bipartition sets.*

Graphs G and H are said to be *isomorphic* if there is a bijection between their vertex sets which preserves the edges. We distinguish between isomorphic graphs only if the labelling of the vertices has some importance. Otherwise, we consider isomorphic graphs essentially equal. When G and H are isomorphic we write $G \cong H$, or simply $G = H$.

Many graphs have been named in the literature. The graph containing n vertices and all $\binom{n}{2}$ possible edges is called the *complete graph* on n vertices, denoted K_n. The graph containing n vertices and no edges is called the *empty graph* on n vertices, denoted O_n. For $m, n \in \mathbb{N}$, the *complete bipartite graph* on bipartition sets of size m and n has all possible edges between the bipartition sets, and is denoted $K_{m,n}$. The graph $K_{1,n-1}$ is called the *star* on n vertices. The *path* on n vertices is denoted P_n and the cycle on n vertices is denoted C_n.

For the standard definitions of *walk*, *closed walk*, *path*, *cycle*, and *length* of a walk, we refer the reader to [70]. The existence of a walk is an equivalence relation on the vertices of a graph whose equivalence classes are called *connected components.* A graph G is *connected* if it has exactly one connected component; that is, there is a walk between every pair of vertices of G. A *cut vertex* of a graph is a vertex whose removal increases the number of connected components, while a *bridge* is an edge whose removal disconnects G.

A noncomplete graph is 2-*connected* if it is connected and contains no cut vertices. A complete graph K_n is 2-connected if $n \geq 3$. More generally, a noncomplete graph is k-*connected* for some $k \geq 2$ if the removal of any set of at most $k-1$ vertices leaves the graph connected. A complete graph K_n is k-connected if $n \geq k+1$. A graph is 2-*edge connected* if it is connected and has no bridges. More generally, a graph is k-*edge-connected* if the removal of any set of at most $k-1$ edges leaves the graph connected.

A *forest* is an acyclic graph (a graph with no cycles). A *tree* is a connected forest. Any vertex of degree at least 2 in a tree is called a *central* vertex.

Let H be a graph. A graph G is called H-free if no subset of vertices of G induces the graph H. A *chordal graph* is one in which every cycle of length 4 or more has a *chord* – that is, an edge that is not part of the cycle but joins two of the vertices of the cycle. Equivalently, a chordal graph is a graph that is C_k-free for all $k \geq 4$.

A *block* or 2-*connected component* of a graph G is a maximal 2-connected subgraph of G. A *block graph* (sometimes called a *clique tree*) is a graph in which every block is a clique. Block graphs were introduced by Harary in [37]. For each graph G, Harary defined the block graph $B(G)$ of a given graph G as that graph whose vertices are the blocks $B_1, B_2, \ldots, B_N$ of G, and whose edges are determined by taking two vertices B_i and B_j as adjacent if and only if they contain a cut vertex of G in common. The main result of [37] is that a graph B is a block graph for some graph G if and only if every block of B is a clique. This characterization has become the commonly used definition for block graphs.

Given graphs $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ with disjoint vertex sets, we have the *disjoint union*

$$G_1 \cup G_2 = (V_1 \cup V_2, E_1 \cup E_2),$$

and the *join*

$$G_1 + G_2 = (V_1 \cup V_2, E_1 \cup E_2 \cup \{\{v, w\} : \ v \in V_1, w \in V_2\}).$$

When $H = K_1$ with single vertex v, we denote $G + H$ simply by $G + v$. Finally, for vertices v_1 and v_2 and G_1 and G_2, respectively, the *vertex bonding* of G_1 and G_2 at v_1 and v_2 is obtained from the disjoint union $G_1 \cup G_2$ by identifying v_1 and v_2.

Finally, a *rooted graph* is a graph with one distinguished vertex, called the *root*.

An isomorphism of rooted graphs must respect the root.

Chapter 2

All-Terminal Reliability

Let $G = (V, E)$ be a finite (multi)graph in which each edge fails independently with probability $q \in [0, 1]$ and vertices are always reliable. The *all-terminal reliability* of G, denoted $\mathrm{Rel}(G; q)$, is the probability that all vertices of G can communicate with one another; that is, the probability that at least a spanning tree is operational. All-terminal reliability is a well-studied model of network robustness, and much research has been carried out on a variety of algorithmic and theoretical issues including algorithmic complexity, polynomial time algorithms for restricted families, efficient bounding procedures, the existence of optimal graphs, and analytic properties of the functions (see [28], for example, or [11] for a more recent survey). Note that all-terminal reliability is often studied in terms of $p = 1 - q$, the probability that each edge is operational, but our results on all-terminal reliability are easier to state and prove in terms of q, so we deal exclusively in the variable q in this chapter.

The all-terminal reliability of a connected graph G with edge set E, denoted $\mathrm{Rel}(G; q)$, is indeed always a polynomial in q of degree (at most) $m = |E|$, as a subgraph with operational edges $E' \subseteq E$ arises with probability

$$(1 - q)^{|E'|} q^{|E| - |E'|}.$$

Summing this probability over all sets E' for which all vertices of G can communicate gives the all-terminal reliability of G. The polynomial turns out to have degree exactly m, as will be seen from the H-form of the polynomial, described later.

As a polynomial, it is natural to inquire about the nature and location of the roots of all-terminal reliability polynomials, called *all-terminal reliability roots* or *ATR roots* henceforth. For example, in [13], it was shown that every graph has an edge subdivision whose ATR roots are all real (this had implications on some outstanding unimodality conjectures for the coefficients of the all-terminal reliability polynomial under various expansions).

More interestingly, ATR roots were noted to have modulus at most 1 (in q) for small graphs, and it was conjectured in [12] that indeed this was the case for all graphs. This contrasts sharply with what is known for other graph polynomials, such as chromatic polynomials [58], independence polynomials [18], and domination polynomials [24] – it is known that the roots are dense in the complex plane for all of these polynomials. Despite some results and generalizations in the affirmative [25, 68], the conjecture for ATR roots was shown to be false in [54]. However,

- the ATR roots provided were only outside the unit disk by a slim margin – the largest modulus of an ATR root found was approximately 1.04;
- the simple graphs with ATR roots outside the unit disk were quite large, with the smallest having over 1500 vertices and over 3000 edges; and
- all of the simple graphs with ATR roots outside the unit disk had many vertices of degree 2, and it is unclear whether all simple graphs with ATR roots outside the unit disk have such low edge connectivity.

Finally, although ATR roots of modulus greater than 1 have been found, there is no known general upper bound on the modulus of an ATR root.

In this chapter, we continue the exploration of the location of ATR roots. In Section 2.1, we find a nontrivial (though non-constant) bound on the modulus of any ATR root of a graph G in terms of the order of the graph. In Section 2.2.1, we study graphs with ATR roots of modulus greater than 1, finding graphs with ATR roots of greater modulus than any previously known. Finally, in Section 2.2.2 we consider simple graphs with ATR roots of modulus greater than 1. We find a smaller example of a simple graph with ATR roots outside of the unit disk, and we find simple graphs that have ATR roots outside of the unit disk and have much higher edge connectivity than any previously known examples.

2.1 An Upper Bound on the Modulus of any All-Terminal Reliability Root

Our presentation of a general upper bound on the modulus of any ATR root will be a long and winding one, requiring us to draw connections to simplicial complexes, sets of monomials, and a particular game on graphs.

2.1.1 All-Terminal Reliability and Simplical Complexes

As the polynomials of degree at most m form a vector space over the real numbers (or rational numbers), there are a number of expansions of all-terminal reliability polynomials in terms of different bases for the polynomial space; one pertinent to our discussion is expressed as follows. For a connected graph G, let F_i denote the number of subsets of E of cardinality i whose removal leaves the graph connected. If G has order n and size m (i.e. n vertices and m edges), then the all-terminal reliability of G is given by

$$\mathrm{Rel}(G;q) = \sum_{i=0}^{m-n+1} F_i q^i (1-q)^{m-i}.$$

This expansion is called the *F-form* of all-terminal reliability.

The coefficients F_i arise in another context as well (see [28]). A *(simplicial) complex* $\mathcal{K}$ on a finite set X is a nonempty collection of subsets of X that is closed under containment, i.e. if $B \in \mathcal{K}$ and $A \subseteq B$ then $A \in \mathcal{K}$. The elements of $\mathcal{K}$ are called the *faces* of the complex and the maximal faces with respect to containment are called *facets* or *bases*. The *dimension* $d = d(\mathcal{K})$ of a complex $\mathcal{K}$ is the cardinality of a largest facet. Let F_i be the number of faces of $\mathcal{K}$ of cardinality i. The sequence $(F_0, F_1, \ldots, F_d)$ is called the *F-vector* of $\mathcal{K}$.

For a connected graph G of order n and size m with edge set E, the subsets $S \subseteq E$ such that $G-S$ is connected are the faces of a simplicial complex of dimension $m-n+1$ called the *cographic matroid* of G, and it is clear that the sequence of coefficients of the F-form of the all-terminal reliability of G is precisely the F-vector of the cographic matroid of G. Let

$$F(G;x) = \sum_{i=0}^{m-n+1} F_i x^i$$

denote the generating polynomial of the F-vector of the cographic matroid of G, which we will call the *F-polynomial* of G. We have

$$\begin{aligned}
\mathrm{Rel}(G;q) &= \sum_{i=0}^{m-n+1} F_i q^i (1-q)^{m-i} \\
&= (1-q)^m \sum_{i=0}^{m-n+1} F_i \left(\tfrac{q}{1-q}\right)^i \\
&= (1-q)^m F\left(G; \tfrac{q}{1-q}\right).
\end{aligned}$$

The connection to simplicial complexes leads to a second expansion of all-terminal reliability. Recall the F-form of all-terminal reliability,

$$\mathrm{Rel}(G;q) = \sum_{i=0}^{m-n+1} F_i q^i (1-q)^{m-i}.$$

Since at least a spanning tree must be operational in order for all nodes to be able to communicate with one another, we see that $F_i = 0$ for all $i < n-1$, and therefore all nonzero terms in the F-form have a factor of $(1-q)^{n-1}$. When we factor out

$(1-q)^{n-1}$ and expand the rest in terms of powers of q, we get the *H-form* of all-terminal reliability:

$$\mathrm{Rel}(G;p) = (1-q)^{n-1} \sum_{k=0}^{m-n+1} H_k q^k.$$

The sequence $(H_0, H_1, \ldots, H_{m-n+1})$ is called the *H-vector* of the cographic matroid (see [28]). Moreover, the generating polynomial

$$H(G;x) = \sum_{k=0}^{m-n+1} H_k x^k$$

of the H-vector of the cographic matroid turns out to be an evaluation of the well-known two-variable Tutte polynomial. For a subset $A \subseteq E$, let $k(A)$ denote the number of connected components of $G[A]$. Let $r(A) = n - k(A)$. The *Tutte polynomial* is given by

$$T(G;x,y) = \sum_{A \subseteq E} (x-1)^{r(E)-r(A)} (y-1)^{|A|-r(A)}.$$

It follows from a result in [8] (see [51]) that

$$T(G;1,x) = H(G;x). \tag{2.1}$$

The values H_i, while guaranteed to be rational numbers from a linear algebra perspective, turn out to be nonnegative integers – there is an *order ideal of monomials* (a finite set of monomials closed under division) such that H_i counts the number of monomials of degree i in the order ideal (see [14, 28]). This connection was derived only via a deep connection between simplical complexes and commutative algebra (see [4]). A newer and different connection between cographic matroids and order ideals of monomials has been afforded by the *chip-firing game* [50, 51], which we describe in some detail, as it is crucial to what follows.

2.1.2 The Chip-Firing Game and Order Ideals of Monomials

Let $G = (V, E)$ be a connected multigraph without loops, and let w denote a special vertex of G. A *configuration* of G is a function $\theta : V \to \mathbb{Z}$ for which $\theta(v) \geq 0$ for all $v \neq w$ and $\theta(w) = -\sum_{v \neq w} \theta(v)$. For $v \neq w$, the number $\theta(v)$ represents the number of chips on vertex v. We imagine that the special vertex w has infinitely many chips. In configuration θ, a vertex $v \neq w$ is *ready to fire* if $\theta(v) \geq \deg(v)$; vertex w is ready to fire if and only if no other vertex is ready. One can think of w as playing the role of a government, stimulating the economy when necessary. *Firing* vertex u changes the configuration from θ to θ', where

$$\theta'(u) = \theta(u) - \deg(u)$$

and for $v \neq u$

$$\theta'(v) = \theta(v) + l(u, v),$$

where $l(u, v)$ is the number of edges between u and v in G. A configuration is *stable* when $\theta(v) < \deg(v)$ for all $v \neq w$; that is, if and only if w is ready to fire.

A *firing sequence* $\Theta = (\theta_0, \theta_1, \ldots, \theta_k)$ is a sequence of configurations in which θ_i is obtained from θ_{i-1} by firing one vertex that is ready to fire for each $i \in \{1, \ldots, k\}$. It is *nontrivial* when $k > 0$. We write $\theta_0 \to \theta_k$ when some nontrivial firing sequence starting with θ_0 and ending with θ_k exists. Configuration θ is *recurrent* if $\theta \to \theta$. Stable, recurrent configurations are called *critical*. For a critical configuration θ, a *critical sequence* is a legal firing sequence of minimal length that makes θ recur.

Lemma 2.1.1 (Merino, [50])**.** *Let G be a graph and let θ be a critical configuration of G. Any critical sequence of θ consists of firing every vertex of G exactly once.*

For example, consider the complete graph K_4 and label the vertices t, u, v, and

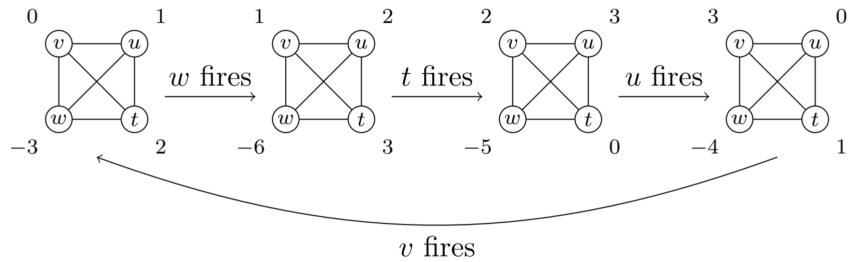


Figure 2.1: The firing sequence beginning from configuration ψ.

w. Let w be the special vertex and consider the initial configuration ψ defined by

$$\psi(t) = 2, \quad \psi(u) = 1, \quad \text{and} \quad \psi(v) = 0.$$

This configuration is stable as the vertices t, u, and v are not ready to fire, meaning that w is ready to fire. Figure 2.1 illustrates the firing sequence that results from initial configuration ψ, and demonstrates that ψ is recurrent; t will fire after w, followed by u and then v, at which point we will have returned to configuration ψ. Thus ψ is a critical configuration.

A *monomial* in the indeterminates (i.e. variables) $x_1, \ldots, x_m$ is a product of non-negative integer powers of the variables. Let $\mathrm{Mon}(x_1, \ldots, x_n)$ denote the set of all monomials in the indeterminates $x_1, \ldots, x_m$. An *order ideal of monomials* M in the indeterminates $x_1, \ldots, x_n$ is a subset of $\mathrm{Mon}(x_1, \ldots, x_n)$ that is closed under division; that is, for m_1 and m_2 in $\mathrm{Mon}(x_1, \ldots, x_n)$, if $m_1 \in M$ and $m_2 | m_1$ then $m_2 \in M$. An order ideal of monomials M is called *pure* if all maximal monomials in M (with respect to division) have the same degree. For $i \geq 0$ let a_i be the number of monomials of degree i in M, and let d be the largest integer for which $a_d \neq 0$. The *degree sequence* of M is the sequence $(a_0, a_1, \ldots, a_d)$.

Let $\mathcal{C}$ be the set of all critical configurations of G. For each $v \in V \backslash \{w\}$, let x_v be an indeterminate. For each $\theta \in \mathcal{C}$, define a monomial

$$m_\theta = \prod_{v \in V \setminus \{w\}} x_v^{\deg(v)-1-\theta(v)}.$$

For example, the critical configuration ψ discussed above has corresponding monomial

$$m_\psi = x_u x_v^2.$$

It was proven in [51] that the set

$$\mathcal{M}_w(G) = \{m_\theta :\ \theta \in \mathcal{C}\}$$

which consists of the monomials in the $n-1$ indeterminates $\{x_v :\ v \in V(G)\setminus\{w\}\}$ corresponding to the critical configurations of G is an order ideal of monomials. Essentially, given a critical configuration, we can add a chip to any vertex v with at most $\deg(v) - 2$ chips and the configuration remains critical (the extra chip need not move in the resulting firing sequence).

Let a_i be the number of monomials of degree i in $\mathcal{M}_w(G)$ for all $i \geq 0$. The generating polynomial of $\mathcal{M}_w(G)$ is given by

$$M_w(G; x) = \sum_{k=0}^{m-n+1} a_i x^i.$$

The (surprising!) key result that connects the chip-firing game to all-terminal reliability was proven in [50].

Theorem 2.1.2 (Merino [50])**.** *For a graph G and a vertex w, the generating polynomial $M_w(G; x)$ of the order ideal of monomials $\mathcal{M}_w(G)$ is an evaluation of the Tutte* *polynomial of G; in particular,*

$$T(G; 1, x) = M_w(G; x).$$

Theorem 2.1.2 implies that $M_w(G; x)$ is independent of the choice of w. Even more interestingly, we have the following corollary.

Corollary 2.1.3. *The degree sequence of $\mathcal{M}_w(G)$ is identical to the H-vector of the cographic matroid.*

Corollary 2.1.3 follows immediately from the fact that, from (2.1), the generating polynomial of the H-vector of the cographic matroid is the same evaluation of the Tutte polynomial as $M_w(G; x)$, the generating polynomial of $\mathcal{M}_w(G)$. In addition, Merino proved the following about the order ideal of monomials $\mathcal{M}_w(G)$.

Theorem 2.1.4 (Merino [51])**.** *For a graph G with special vertex w, the set $\mathcal{M}_w(G)$ is a pure order ideal of monomials.*

Equipped with links between all-terminal reliability, simplicial complexes and order ideals of monomials, we are now ready to state and prove a general bound on the modulus of any root of the all-terminal reliability polynomial of a connected graph of order n.

2.1.3 An Upper Bound on the Modulus of any All-Terminal Reliability Root

First we note that the all-terminal reliability of any disconnected graph is identically zero, so that all complex numbers are ATR roots of disconnected graphs, but in a fairly trivial manner. We also note that any connected graph on n vertices and $m < n$ edges is a tree, and

$$\mathrm{Rel}(T; q) = (1 - q)^{n-1}$$

for any tree T on n vertices. Thus, the only ATR root of any tree T on n vertices is $z = 1$ with multiplicity $n-1$. Since we know all there is to know about ATR roots of disconnected graphs and trees, the result below, which gives the first general upper bound on the modulus of any ATR root of a graph of order n, is concerned only with connected graphs of order n and size $m \geq n$.

Theorem 2.1.5. *Let G be a connected graph of order n and size $m \geq n$ with b bridges (edges whose removal disconnects G). Any root z of $\mathrm{Rel}(G; q)$ satisfies $|z| \leq n - b - 1$. Moreover, if G has a vertex w such that*

- *w has no incident multiple edges and*
- *w is incident with a non-bridge*

then $|z| \leq n - b - 2$.

Proof. Note that $b \leq n-2$ as $m \geq n$ implies that G has some cycle. Consider

$$\begin{aligned}\mathrm{Rel}(G; q) &= \sum_{i=0}^{m-n+1} F_i q^i (1-q)^{m-i} \\ &= (1-q)^{n-1} \sum_{i=0}^{m-n+1} H_i q^i.\end{aligned}$$

It is clear that $F_0 = 1$, and $F_1 = m - b$. Factoring $(1-q)^{n-1}$ from the F-form gives

$$\sum_{i=0}^{m-n+1} F_i q^i (1-q)^{m-n+1-i} = \sum_{i=0}^{m-n+1} H_i q^i.$$

The only constant term on the left-hand side is given by F_0, so that $H_0 = F_0 = 1$. Collecting all of the linear terms on the left-hand side, we find

$$H_1 = F_1 - (m-n+1)F_0 = m - b - (m-n+1) = n - b - 1.$$

Consider now the generating function

$$H(z) = \sum_{i=0}^{m-n+1} H_i z^i$$

for the H-vector of the cographic matroid; this is identical to the generating function $M_w(G; z)$ mentioned earlier for the order ideal of monomials derived from the chip-firing game. For the first statement of the theorem it suffices to show that the roots of $H(z)$ lie in the disk $|z| \leq n - b - 1$.

Huh [42] very recently answered an outstanding conjecture about H-vectors of matroids, proving that the H-vector of the cographic matroid of a connected graph is a log-concave sequence, that is,

$$H_i^2 \geq H_{i-1} H_{i+1}$$

for all $i \in \{1, \ldots, m-n\}$. It follows directly that

$$\frac{H_{i-1}}{H_i} \leq \frac{H_i}{H_{i+1}}$$

for all $i \in \{1, \ldots, m-n\}$. Thus we have

$$\frac{H_{i-1}}{H_i} \leq \frac{H_{m-n}}{H_{m-n+1}}$$

for all $i \in \{1, \ldots, m-n\}$. The well-known Eneström-Kakeya Theorem (see [1], for example) states that a polynomial $\sum_{i=0}^{d} a_i z^i$ with positive coefficients has roots in the annulus

$$\min\left(\left\{\frac{a_{i-1}}{a_i} : i = 1, \ldots, d\right\}\right) \leq |z| \leq \max\left(\left\{\frac{a_{i-1}}{a_i} : i = 1, \ldots, d\right\}\right).$$

From this and Huh's result, it follows that the roots of $H(z)$ have modulus bounded above by

$$\max\left(\left\{\frac{H_{i-1}}{H_i} : i = 1, \ldots, m-n+1\right\}\right) = \frac{H_{m-n}}{H_{m-n+1}}.$$

By Corollary 2.1.3 and Theorem 2.1.4, there is a pure order ideal of monomials $\mathcal{M} = \mathcal{M}_w(G)$ with H_i being the number of monomials of degree i in $\mathcal{M}$. Consider the set

$$S = \{(l, x)\colon\ l \in \mathcal{M} \text{ of degree } m-n+1,\ x \text{ is a variable that divides } l\}.$$

For each $(l, x) \in S$, note that l/x is a monomial of degree $m-n$ in $\mathcal{M}$, and the purity of $\mathcal{M}$ ensures that each monomial of degree $m-n$ in $\mathcal{M}$ appears at least once under this construction. It follows that $H_{m-n} \leq |S|$, and since $|S| \leq H_1 H_{m-n+1} = (n-b-1)H_{m-n+1}$, we have that

$$\frac{H_{m-n}}{H_{m-n+1}} \leq n-b-1,$$

so that any root z of H satisfies $|z| \leq n-b-1$. It follows that the roots of

$$\mathrm{Rel}(G; q) = (1-q)^{n-1} H(G; q)$$

satisfy $|z| \leq n-b-1$, and we have proven the first statement.

For the second statement, let w be any vertex of G without any incident multiple edges and with some incident edge whose removal does not disconnect the graph. We first note that the conditions on w imply that vertex w is contained in a cycle of length at least 3, and thus $b \leq n-3$. Consider the order ideal of monomials $\mathcal{M}_w(G)$ constructed via the chip-firing game with w being the special vertex. We claim that each monomial in $\mathcal{M}_w(G)$ is a product of nonzero powers of at most $n-b-2$ variables;

that is, for each monomial $l \in \mathcal{M}_w(G)$, there are $n-1-(n-b-2) = b+1$ distinct indeterminates $x_{v_1}, \ldots, x_{v_{b+1}}$ such that $x_i \nmid l$.

Every monomial in $\mathcal{M}_w(G)$ corresponds to a critical configuration of G. First of all, let B denote the set containing each vertex on the opposite end of a bridge from w. Note that any vertex v in B must have $\deg(v) - 1$ chips, as otherwise v will never fire. Thus, every vertex in B must have degree 0 in every monomial corresponding to a critical configuration. Suppose – to reach a contradiction – that we have a critical configuration in which the b vertices in B are the only vertices with degree many less 1 chips. After w fires, the only vertices that are ready to fire are the neighbours of w that are connected to w by bridges, i.e. the vertices in B. By Lemma 2.1.1, the vertex w never fires again, and this leaves the neighbours of w that are not in B (there must be at least 2) with only one more chip than they started with, which is still not enough for them to fire. This is a contradiction since the initial configuration was assumed to be critical, and thus in any critical configuration ϕ at least one neighbour u of w that is not in B must have $\deg(u) - 1$ chips. This means that the indeterminate x_u has degree 0 in the monomial corresponding to the critical configuration ϕ.

Therefore, any monomial $l \in \mathcal{M}_w(G)$ is divisible by at most $n-b-2$ indeterminates. The same pair counting argument as in the previous part of the proof shows that

$$H_{m-n} \leq (n-b-2)H_{m-n+1}$$

and we conclude via the Eneström-Kakeya Theorem that if z is a root of $H(G;x)$ then

$$|z| \leq n-b-2,$$

and hence any root z of $\mathrm{Rel}(G;q) = (1-q)^{n-1}H(G;q)$ satisfies $|z| \leq n-b-2$. $\square$

Corollary 2.1.6. *If G is a connected simple graph with b bridges then any root z of $\mathrm{Rel}(G;q)$ satisfies either $|z| \leq n-b-2$ or $z = 1$.*

Proof. Let G be a connected simple graph. Either G is a tree and has unique ATR root $z = 1$, or G has a cycle. In the latter case, G has some vertex w that is incident with a non-bridge (and w is not incident to any multiple edges as G is a simple graph). By Theorem 2.1.5, any ATR root z of G satisfies $|z| \leq n - b - 2$. $\square$

A consequence of the proof of Theorem 2.1.5 is that

$$\frac{H_{m-n}(G)}{H_{m-n+1}(G)} \leq n - 2$$

for any simple graph G of order n and size $m \geq n$. While this bound is not best possible, we are off by at most a factor of 2. We can show that

$$\frac{H_{m-n}(K_n)}{H_{m-n+1}(K_n)} = \frac{n-2}{2}$$

for all n by considering the critical configurations of the chip-firing game on K_n. Fix special vertex w of K_n. The *minimal critical configurations* (minimal in terms of the number of chips – the configurations corresponding to monomials of order $m - n + 1$) are obtained by assigning $0, 1, \ldots, n-2$ chips to the remaining vertices in any order (see [15], for example). There are $(n-1)!$ such assignments. Each critical configuration corresponding to a monomial of order $m - n$ can be obtained from exactly two distinct minimal critical configurations by adding a single chip to a particular vertex that has strictly less than $n - 2$ chips (adding a chip to a vertex with $n - 2$ chips leads to a configuration that is no longer stable). Since there are $n - 2$ vertices with strictly less than $n - 2$ chips in each minimal critical configuration of K_n, we have

$$2H_{m-n}(K_n) = (n-2)H_{m-n+1}(K_n) \quad \Longrightarrow \quad \frac{H_{m-n}(K_n)}{H_{m-n+1}(K_n)} = \frac{n-2}{2}.$$

Using a computer algebra system, we have verified for all $n \leq 9$ that if G is a simple graph on n vertices and $m \geq n$ edges then

$$\frac{H_{m-n}(G)}{H_{m-n+1}(G)} \leq \frac{n-2}{2},$$

with equality if and only if $G \cong K_n$, leading us to believe that our bound can be improved.

2.2 All-Terminal Reliability Roots outside of the Unit Disk

We now turn to providing more example of graphs with ATR roots outside the unit disk centred at the origin of the complex plane (referred to simply as the *unit disk* henceforth).

2.2.1 All-Terminal Reliability Roots of Larger Modulus

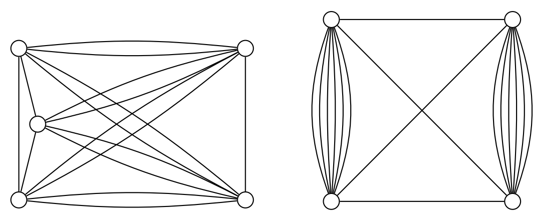
Brown and Colbourn investigated the roots of all-terminal reliability polynomials in [12] where they made the following conjecture.

Conjecture 2.2.1 (Brown-Colbourn Conjecture)**.** Let G be a connected graph. If z is a root of $\mathrm{Rel}(G;q)$ then $|z| \leq 1$. In other words, ATR roots all lie inside the unit disk.

While Wagner proved that the Brown-Colbourn conjecture is true for series-parallel graphs [68], the conjecture was proven false in general by Sokal and Royle [54]. However, the largest known modulus of an ATR root is approximately 1.04, as noted in [16]. We improve on this here, finding ATR roots that are almost three times further outside of the unit disk.

We generalize the graphs that were found to have ATR roots outside of the unit disk in [54]. For positive integers m, n, a, and b, let $G_{m,n}^{a,b}$ be the graph on $m+n$

vertices defined as follows. Take disjoint complete graphs K_m and K_n and replace every edge by a edges in parallel (i.e. replace each edge with a *bundle* of a edges), and then connect every nonadjacent pair of vertices with b parallel edges. The graphs $G^{1,2}_{3,2}$ and $G^{6,1}_{2,2}$ are shown in Figure 2.2. The graph $G^{6,1}_{2,2}$ is the smallest multigraph known to have ATR roots outside of the unit disk, and $G^{1,6}_{3,3}$ has the ATR root with the largest known modulus of approximately 1.04 (see [54]). We will see that $G^{1,6}_{n,n}$ has ATR roots of even larger modulus for $n > 3$. The following result gives us a way to compute $\mathrm{Rel}(G^{a,b}_{m,n}; q)$.



(a) The graph $G^{1,2}_{3,2}$. (b) The graph $G^{6,1}_{2,2}$.

Figure 2.2: Two examples of the graph $G^{a,b}_{m,n}$.

Proposition 2.2.2. *Let m, n, a, and b be positive integers. Then*

$$\sum_{i=1}^{m}\sum_{j=0}^{n}\binom{m-1}{i-1}\binom{n}{j}q^{a[i(m-i)+j(n-j)]+b[i(n-j)+j(m-i)]}\mathrm{Rel}\Big(G^{a,b}_{i,j}; q\Big) = 1. \tag{2.2}$$

Proof. Let G_1 and G_2 be the complete graphs of orders m and n, respectively, from which $G^{a,b}_{m,n}$ is formed. Let v be a vertex of G_1. For a particular subset C of vertices which contains v and has $i \geq 1$ vertices from G_1 and $j \geq 0$ vertices from G_2, we calculate the probability that C is a connected component in $G^{a,b}_{m,n}$. In order for C to be a connected component in $G^{a,b}_{m,n}$, all of the members of C must be able to communicate with one another and all of the members of C must be unable to communicate with any vertex outside of C. The former occurs with probability

$$\mathrm{Rel}\Big(G^{a,b}_{i,j}; q\Big)$$

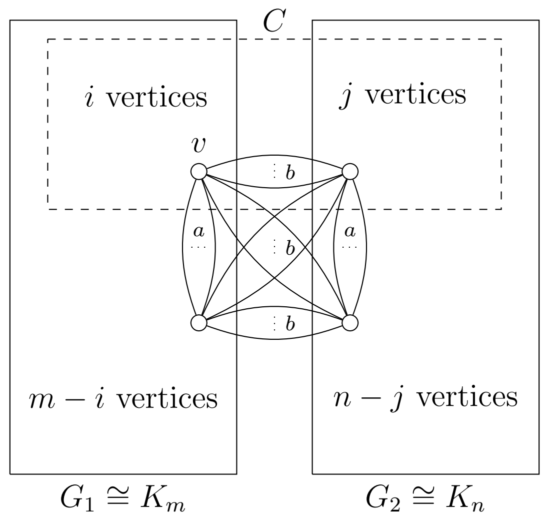


Figure 2.3: The graph $G^{a,b}_{m,n}$ and a particular subset C_v of vertices.

while the latter occurs with probability

$$q^{a[i(m-i)+j(n-j)]+b[i(n-j)+j(m-i)]},$$

as there are $a[i(m-i)+j(n-j)]+b[i(n-j)+j(m-i)]$ many edges between C and the remaining vertices of the graph. A rough sketch of $G^{a,b}_{m,n}$ is provided in Figure 2.3 to aid in the counting of edges.

There are $\binom{m-1}{i-1}\binom{n}{j}$ distinct sets of this form as we may choose any $i-1$ vertices from the remaining $m-1$ vertices of G_1 and any j vertices from the n vertices of G_2. Thus the probability that v lies in some connected component containing i vertices from G_1 and j vertices from G_2 is

$$\binom{m-1}{i-1}\binom{n}{j}q^{a[i(m-i)+j(n-j)]+b[i(n-j)+j(m-i)]}\mathrm{Rel}\Big(G^{a,b}_{i,j};p\Big)$$

If we sum over all possibilities for i and j we obtain 1 as v must be in some component. This gives (2.2). $\square$

n	ATR roots of $G^{1,6}_{n,n}$ of greatest modulus	Modulus
3	$0.6965978094 \pm 0.7739344775i$	1.0412603341
4	$0.7225077023 \pm 0.7873461471i$	1.0686118731
5	$0.7415248258 \pm 0.7932060873i$	1.0858337645
6	$0.7557913447 \pm 0.7946437701i$	1.0966673507
7	$0.7665525647 \pm 0.7937722633i$	1.1034841369
8	$0.7747703944 \pm 0.7917743649i$	1.1077796753
9	$0.7811493576 \pm 0.7892664429i$	1.1104664951
10	$0.7861847934 \pm 0.7865650322i$	1.1121020993
11	$0.7902223368 \pm 0.7838329136i$	1.1130343112
12	$0.7935054014 \pm 0.7811532818i$	1.1134860896

Table 2.1: ATR roots of $G^{1,6}_{n,n}$ of greatest modulus for small n. All values rounded to 10 decimal places.

Proposition 2.2.2 gives a recursion for $\mathrm{Rel}\big(G^{a,b}_{m,n}; q\big)$ in terms of the smaller polynomials $\mathrm{Rel}\Big(G^{a,b}_{i,j}; q\Big)$ for all $0 < i \le m$ and $0 \le j \le n$ with $i + j < m + n$. A base case is not necessary as $\mathrm{Rel}\Big(G^{a,b}_{1,0}; q\Big) = 1$ is actually given by the same equation. This allows us to compute $\mathrm{Rel}\big(G^{a,b}_{m,n}; q\big)$ efficiently for small values of m and n.

We numerically computed the ATR roots of the graphs $G^{a,b}_{m,n}$ for all small m, n, a, and b, and graphs of the form $G^{1,6}_{n,n}$ yielded the roots of largest modulus. While the ATR roots of the graphs $G^{1,6}_{1,1}$ and $G^{1,6}_{2,2}$ all lie inside the unit disk, for each $n \in \{3, 4, \ldots, 12\}$ the graph $G^{1,6}_{n,n}$ has ATR roots outside of the unit disk. Table 2.1 shows the ATR roots of $G^{1,6}_{n,n}$ of greatest modulus for $n \in \{3, 4, \ldots, 12\}$, and it is clear that the modulus is increasing with n for the values shown. The ATR roots of largest modulus of $G^{1,6}_{12,12}$ have modulus almost three times as far outside of the unit disk as the best examples from [54] – namely, those for the graph $G^{1,6}_{3,3}$.

2.2.2 Simple Graphs with All-Terminal Reliability Roots outside of the Unit Disk

Of course, for $n \geq 2$ the graphs $G^{1,6}_{n,n}$ discussed in the previous section contain multiple edges. In [54], several *simple* graphs were found that still violated the Brown-Colbourn Conjecture – the smallest example being the graph on 1512 vertices and 3016 edges obtained from $G^{11,1}_{2,2}$ (denoted differently in [54]) by replacing every edge with 58 edges in parallel, and then replacing every edge with two edges in series (i.e. with a path of length 2). The ATR roots of this graph can be obtained from the ATR roots of $G^{11,1}_{2,2}$ by transforming to a related generating polynomial and using reduction formulae for the series and parallel edge replacements (see [54]). All of the simple graphs in [54] with ATR roots outside of the unit disk are constructed in a similar manner, and thus they all have edge connectivity 2 (in fact, they have many vertices of degree 2). We improve on these results in two ways: we find a smaller simple graph that has ATR roots outside of the unit disk, and we find simple graphs with higher edge connectivity that have ATR roots outside of the unit disk.

In order to generate examples of simple graphs with higher edge connectivity that have ATR roots outside of the unit disk, we discuss a more general substitution operation on graphs. We generalize the idea from [54] of replacing every edge in a graph with either k edges in parallel or k edges in series. Essentially, our substitution operation involves replacing every edge in a given graph by *any* fixed graph of our choice.

We define a *gadget* $H(u, v)$ to be a connected graph H together with two special vertices u and v of H. Let G be a graph and let $H(u, v)$ be a gadget. An *edge substitution* of the gadget $H(u, v)$ into G, denoted $G[H(u, v)]$, is any graph formed by replacing each edge $\{x, y\} \in E(G)$ by a copy $H_{\{x,y\}}$ of H, identifying u with x and v with y. Note that in order to obtain a specific edge substitution we need to fix an

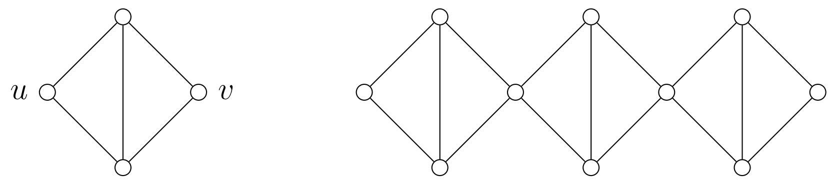


Figure 2.4: A gadget $D(u, v)$ and the edge substitution $P_4[D(u, v)]$.

orientation of G, but our results here do not depend on the orientation of G. We let $G[H(u, v)]$ denote *any* edge substitution of the gadget $H(u, v)$ into G.

We will present an expression for the all-terminal reliability of any edge substitution $G[H(u, v)]$ in terms of reliability polynomials of G and H. However, we will need more than just the all-terminal reliability of H. We introduce a new reliability polynomial which we term the $\{u, v\}$-*split reliability*. This new model of reliability will be applied to all-terminal reliability here, but there is good reason to believe that it is of interest in its own right.

Definition 2.2.1. Let G be a connected graph in which each edge fails independently with probability q, and let $\{u, v\} \subseteq V(G)$ where $u \neq v$. The $\{u, v\}$-*split reliability* of G, denoted $\mathrm{spRel}_{\{u,v\}}(G; q)$ is the probability that every vertex w in G can communicate with *exactly* one vertex from $\{u, v\}$ (i.e. every vertex in G can communicate with either u or v but not both).

For example, consider the complete graph on 4 vertices and let u and v be vertices. All 10 operational states for the $\{u, v\}$-split reliability of K_4 are pictured in Figure 2.5 – there are 8 states with two operational edges and 2 states with three operational edges. Hence, the $\{u, v\}$-split reliability of K_4 is given by

$$\mathrm{spRel}_{\{u,v\}}(K_4; q) = 8(1 - q)^2 q^4 + 2(1 - q)^3 q^3.$$

We can now present an expression for the all-terminal reliability of an edge substitution graph $G[H(u, v)]$. The key is to notice that the *internal* vertices of the gadget

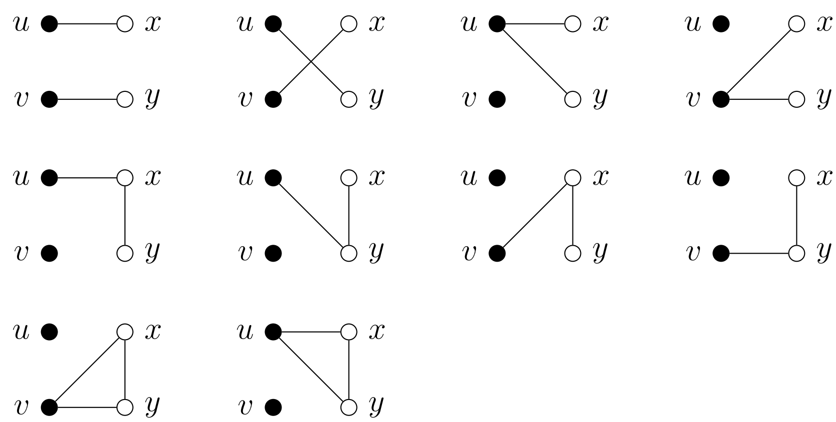


Figure 2.5: The operational states for the $\{u, v\}$-split reliability of K_4.

(that is, the vertices of H apart from u and v) can only communicate with the rest of the graph G through u and v. Thus, in any operational state of $G[H(u, v)]$ each individual copy of the gadget must either be connected, or split between u and v.

Proposition 2.2.3. *Let G be a graph on n vertices and m edges and let $H(u, v)$ be a gadget. The all-terminal reliability of any edge substitution $G[H(u, v)]$ is given by*

$$\mathrm{Rel}(G[H(u,v)];q) = [\mathrm{Rel}(H;q)]^m F\Big(G; \tfrac{\mathrm{spRel}_{\{u,v\}}(H;q)}{\mathrm{Rel}(H;q)}\Big), \tag{2.3}$$

where $F(G; x)$ is the generating polynomial of the F-vector of the cographic matroid of G.

Proof. Consider any copy of the gadget in any operational state of $G[H(u, v)]$. There are only two possibilities for the gadget if $G[H(u, v)]$ is to be operational:

(i) All of the vertices in the gadget can communicate with one another. This occurs with probability $\mathrm{Rel}(H; q)$. We say that the gadget is *operational* in this case.

(ii) Each vertex in the gadget can communicate with exactly one of the vertices u or v. This occurs with probability $\mathrm{spRel}_{\{u,v\}}(H; q)$. We say that the gadget *splits* in this case.

A gadget splitting in $G[H(u,v)]$ corresponds to an edge failing in G, while an operational gadget in $G[H(u,v)]$ corresponds to an operational edge in G. For any state ϕ of $G[H(u,v)]$, let E_{op} be the set of edges $e \in E(G)$ for which H_e is operational. The state ϕ is operational if and only if E_{op} induces a connected subgraph of G and for all $e \in E(G) \backslash E_{\text{op}}$ the gadget H_e splits. Thus, the operational states of $G[H(u,v)]$ in which $m-i$ gadgets are operational and i gadgets split correspond exactly to the operational states of G in which $m-i$ edges are operational and i edges fail, and the latter are counted by F_i (the F-coefficient of G). Thus the all-terminal reliability of any edge substitution $G[H(u,v)]$ is given by

$$\text{Rel}(G[H(u,v)];q) = \sum_{i=0}^{m-n+1} F_i [\text{Rel}(H;q)]^{m-i} \big[\text{spRel}_{\{u,v\}}(H;q)\big]^i,$$

which can be rewritten as (2.3) by factoring $[\text{Rel}(H;q)]^m$ out of the sum. □

The expression for $\text{Rel}(G[H(u,v)];q)$ given in Proposition 2.2.3 allows us to find ATR roots of $G[H(u,v)]$ by a two-step process. We first find a root r of $\text{Rel}(G;q)$ and then solve a second equation that involves the all-terminal reliability of H, the $\{u,v\}$-split reliability of H, and the root r.

Corollary 2.2.4. *Let G be a connected graph and let $H(u,v)$ be a gadget. If $r \neq 1$ is an ATR root of G, then any solution of the equation*

$$\text{spRel}_{\{u,v\}}(H;q) = \tfrac{r}{1-r} \cdot \text{Rel}(H;q) \tag{2.4}$$

is an ATR root of $G[H(u,v)]$.

Proof. Let $r \neq 1$ be a root of $\text{Rel}(G;q)$. Since we can write

$$\text{Rel}(G;q) = (1-q)^m F\Big(G; \tfrac{q}{1-q}\Big),$$

the root r of $\mathrm{Rel}(G;q)$ corresponds to the root $\frac{r}{1-r}$ of $F(G;x)$. By Proposition 2.2.3,

$$\mathrm{Rel}(G[H(u,v)];q) = [\mathrm{Rel}(H;q)]^m F\Big(G; \tfrac{\mathrm{spRel}_{\{u,v\}}(H;q)}{\mathrm{Rel}(H;q)}\Big)$$

for any gadget $H(u,v)$. Therefore, any solution of the equation

$$\frac{\mathrm{spRel}_{\{u,v\}}(H;q)}{\mathrm{Rel}(H;q)} = \frac{r}{1-r},$$

or equivalently

$$\mathrm{spRel}_{\{u,v\}}(H;q) = \tfrac{r}{1-r} \cdot \mathrm{Rel}(H;q),$$

is a root of $\mathrm{Rel}(G[H(u,v)];q)$. $\square$

Using Corollary 2.2.4 we can find ATR roots of edge substitution graphs by first finding an ATR root r of G and then solving (2.4). An inherent problem with this technique is that we can only solve for the ATR roots of a graph G *exactly* in special cases. For many graphs we can only *approximate* the ATR roots. While we can obtain very precise approximations to an ATR root r using numerical methods, we then must solve (2.4), and it is well known that the location of the roots of a polynomial can be very sensitive to small changes in the coefficients. Wilkinson's Polynomial [72] is a classic example of this phenomenon.

To get around this problem, instead of solving (2.4) numerically, we use a particular stability test due to Schur and Cohn to show that (2.4) has solutions outside of the unit disk for all values close to r, so that our numerical approximation to the ATR root r will be sufficient. The stability test that we use is described in detail in Section 11.5 of [52]. The statement of the key result below requires some new notation. The *complex conjugate* of $a \in \mathbb{C}$ is denoted $\overline{a}$ and the *conjugate transpose* of a complex matrix A is denoted A^* (this is the operation of taking the conjugate of every entry of A and then transposing). Finally, for a finite sequence $a_1, \ldots, a_n$ of nonzero real numbers, $\mathrm{B}(a_1, \ldots, a_n)$ denotes the number of sign changes in the sequence (i.e. the number of indices $k \in \{2, \ldots, n\}$ for which $a_{k-1}a_k < 0$). For example, $\mathrm{B}(-1, 1, 2, -4, -2) = 2$.

Theorem 2.2.5 (Schur-Cohn, [52], Cor. 11.5.14)**.** *Let* $f(z) = \sum_{k=0}^{n} a_k z^k$ *be a polynomial of degree* n*. Define the upper triangular matrices*

$$A_k = \begin{bmatrix} a_0 & a_1 & \ldots & a_{k-1} \\ & a_0 & \ldots & a_{k-2} \\ & & \ddots & \vdots \\ & & & a_0 \end{bmatrix} \quad \text{and} \quad B_k = \begin{bmatrix} \overline{a}_n & \overline{a}_{n-1} & \ldots & \overline{a}_{n-k+1} \\ & \overline{a}_n & \ldots & \overline{a}_{n-k+2} \\ & & \ddots & \vdots \\ & & & \overline{a}_n \end{bmatrix},$$

where the zero entries have been left blank. Suppose that for $k \in \{1, \ldots, n\}$*, the determinants*

$$M_k = \begin{vmatrix} B_k^* & A_k \\ A_k^* & B_k \end{vmatrix}$$

are all different from zero. Then f *has no root on the unit circle,*

$$\beta = \mathrm{B}(1, M_1, M_2, \ldots, M_n)$$

roots outside of the unit circle, and $\alpha = n - \beta$ *roots inside it.* □

We now have all the theory that we need to prove our desired results, and we only need to build up the particular examples. One of our goals was to find simple graphs with high edge connectivity that have ATR roots outside of the unit disk, so we will need gadgets with high edge connectivity. While the complete graph is an obvious candidate, we have found that using the complete graph minus an edge is more effective. For each $n \geq 3$, let K_n^- denote the graph obtained from K_n by deleting an edge, and suppose that u and v are nonadjacent in K_n^-. Clearly K_n^- is $(n-2)$-edge connected. The following result is straightforward.

Lemma 2.2.6. *Let G be a 2-edge connected graph. For any $n \geq 3$, the graph $G[K_n^-(u,v)]$ is $(n-1)$-edge connected.* □

In order to find ATR roots of an edge substitution $G[K_n^-(u,v)]$ using Corollary 2.2.4, we will require formulae for $\mathrm{Rel}(K_n^-;q)$ and $\mathrm{spRel}_{\{u,v\}}(K_n^-;q)$. We deal with $\mathrm{Rel}(K_n^-;q)$ first. We find a recursion for $\mathrm{Rel}(K_n^-;q)$ that is similar to the well-known recursion for $\mathrm{Rel}(K_n;q)$ (see [28], for example):

$$\mathrm{Rel}(K_n;q) = 1 - \sum_{i=1}^{n-1} \binom{n-1}{i-1} q^{i(n-i)} \mathrm{Rel}(K_i;q),$$

with the base case $\mathrm{Rel}(K_1;q) = 1$.

Proposition 2.2.7. *For any $n \geq 2$,*

$$\begin{aligned}\mathrm{Rel}(K_n^-;q) = 1 &- \sum_{i=1}^{n-1} \binom{n-2}{i-1} q^{i(n-i)-1} \mathrm{Rel}(K_i;q) \\ &- \sum_{i=3}^{n-1} \binom{n-2}{i-2} q^{i(n-i)} \mathrm{Rel}(K_i^-;q).\end{aligned}$$

Proof. Let u and v be the nonadjacent vertices in K_n^-. We find the probability that the vertex u can communicate with exactly i vertices of K_n^- (including itself) for some $i \in \{1, \ldots, n-1\}$. There are two cases:

(i) u cannot communicate with v

The probability that u can communicate with some particular subset U of i vertices from $V(K_n) - \{v\}$ is given by $\mathrm{Rel}(K_i;q)$. If u is to communicate with *only* these i vertices, then all $i(n-i)-1$ edges between U and the remaining vertices of K_n^- must be down. Since there are $\binom{n-2}{i-1}$ ways to choose the vertices of U (remember that U must contain u and must not contain v), the probability that u can communicate with exactly i vertices not including v is

$$\binom{n-2}{i-1} q^{i(n-i)-1} \mathrm{Rel}(K_i; q).$$

(ii) u can communicate with v

Note that in this case i must be at least 3. The probability that u can communicate with some particular set U of i vertices of K_n^- (necessarily containing v) is given by $\mathrm{Rel}(K_i^-; q)$. If u is to communicate with *only* these i vertices, then all $i(n-i)$ edges between U and the remaining vertices of K_n^- must be down. Since there are $\binom{n-2}{i-2}$ ways to choose the vertices of U (remember that U must contain u and v), the probability that u can communicate with exactly $i \geq 3$ vertices including v is

$$\binom{n-2}{i-2} q^{i(n-i)} \mathrm{Rel}(K_i^-; q).$$

Summing the probabilities over all possible i for each case gives the probability that K_n^- is not operational – thus $\mathrm{Rel}(K_n^-; q)$ is 1 minus these sums. □

Now that we can find $\mathrm{Rel}(K_n^-; q)$ efficiently, we consider $\mathrm{spRel}_{\{u,v\}}(K_n^-; q)$. The polynomial $\mathrm{spRel}_{\{u,v\}}(K_n^-; q)$ can be expressed in terms of the polynomials $\mathrm{Rel}(K_i; q)$ for $i \in \{1, \ldots, n-1\}$, as we demonstrate below.

Proposition 2.2.8. *For any $n \geq 2$,*

$$\mathrm{spRel}_{\{u,v\}}(K_n^-; q) = \sum_{i=1}^{n-1} \binom{n-2}{i-1} q^{i(n-i)-1} \mathrm{Rel}(K_i; q) \mathrm{Rel}(K_{n-i}; q).$$

Proof. In any operational state for the $\{u, v\}$-split reliability of K_n^-, the vertex u

must be able to communicate with exactly i vertices (including itself) for some $i \in \{1, \ldots, n-1\}$, while the vertex v must be able to communicate with all of the remaining $n-i$ vertices (including itself). Given a set V_u of i vertices such that $u \in V_u$ and $v \notin V_u$, let $V_v = V(K_n^-)\backslash V_u$, and note that $v \in V_v$. The probability that all the vertices of V_u can communicate is given by $\mathrm{Rel}(K_i; q)$ while the probability that all the vertices of V_v can communicate is given by $\mathrm{Rel}(K_{n-i}; q)$. Since no vertex of V_u can communicate with a vertex of V_v, all $i(n-i)-1$ edges that connect V_u to V_v must be down, and this occurs with probability $q^{i(n-i)-1}$. Finally, there are $\binom{n-2}{i-1}$ ways to choose the vertices of V_u – remember that V_u must contain u and must not contain v. Therefore, the probability that u can communicate with exactly i vertices while v can communicate with the remaining $n-i$ vertices is given by

$$\binom{n-2}{i-1} q^{i(n-i)-1} \mathrm{Rel}(K_i; q) \mathrm{Rel}(K_{n-i}; q).$$

Summing over $i \in \{1, \ldots, n-1\}$ gives the desired expression for $\mathrm{spRel}_{\{u,v\}}(K_n^-; q)$. $\square$

Our examples of simple graphs with roots outside the unit disk are all of the form

$$G^{(k,n)} = G_{3,3}^{k,6k}[K_n^-(u,v)]$$

for $k \geq 1$ and $n \geq 3$. We start with the base graph $G_{3,3}^{1,6}$, replace every edge with a bundle of k edges, and then substitute the gadget $K_n^-(u,v)$ for every edge. Before looking at particular examples we outline our general procedure for demonstrating that some $G^{(k,n)}$ has an ATR root outside of the unit disk.

We have found numerically that a particular ATR root R of the graph $G_{3,3}^{1,6}$ satisfies

$$0.69659 \leq \mathrm{Re}(R) \leq 0.69660 \tag{2.5}$$

and

$$0.77393 \leq \mathrm{Im}(R) \leq 0.77394. \tag{2.6}$$

This is one of the ATR roots of $G_{3,3}^{1,6}$ of greatest modulus (its conjugate $\overline{R}$ would work just as well). By Proposition 2.2.3, the all-terminal reliability of the graph $G_{3,3}^{k,6k}$ obtained from $G_{3,3}^{1,6}$ by replacing each edge with a bundle of $k \geq 1$ edges is given by

$$\mathrm{Rel}\Big(G_{3,3}^{k,6k}; q\Big) = \mathrm{Rel}\big(G_{3,3}^{1,6}; q^k\big),$$

as the all-terminal reliability of a bundle of k edges is $1 - q^k$ and the split reliability is q^k. This means that $\sqrt[k]{R}$ is an ATR root of $G_{3,3}^{k,6k}$.

Now by Corollary 2.2.4, the ATR roots of $G^{(k,n)} = G_{3,3}^{k,6k}[K_n^-(u,v)]$ include the solutions of the equation

$$\mathrm{spRel}_{\{u,v\}}(K_n^-; q) = \frac{\sqrt[k]{R}}{1 - \sqrt[k]{R}} \cdot \mathrm{Rel}(K_n^-; q). \tag{2.7}$$

Note that both $\mathrm{spRel}_{\{u,v\}}(K_n^-; q)$ and $\mathrm{Rel}(K_n^-; q)$ have a factor of $(1-q)^{n-2}$, as at least $n-2$ edges must be operational for $\{u,v\}$-split reliability of a graph on n vertices, and at least $n-1$ edges must be operational for all-terminal reliability of a graph on n vertices. So we may consider the equation

$$\frac{\mathrm{spRel}_{\{u,v\}}(K_n^-; q)}{(1-q)^{n-2}} = \frac{\sqrt[k]{R}}{1 - \sqrt[k]{R}} \cdot \frac{\mathrm{Rel}(K_n^-; q)}{(1-q)^{n-2}} \tag{2.8}$$

instead. The bounds (2.5) and (2.6) on the real and imaginary parts of the original root R of $G_{3,3}^{1,6}$ translate to bounds on the real and imaginary parts of $\frac{\sqrt[k]{R}}{1-\sqrt[k]{R}}$. For any real numbers a and b satisfying these bounds (respectively), we apply Theorem 2.2.5

to the polynomial

$$f_n(z) = \frac{\mathrm{spRel}_{\{u,v\}}(K_n^-;q)}{(1-q)^{n-2}} - (a+bi) \cdot \frac{\mathrm{Rel}(K_n^-;q)}{(1-q)^{n-2}}.$$

We are able to determine the sign of all of the required determinants using only the bounds on a and b.

In the case that $n = 3$, note that $K_3^- \cong P_3$ so that the graph

$$G^{(k,3)} = G_{3,3}^{k,6k}[K_3^-(u,v)]$$

is obtained from $G_{3,3}^{1,6}$ by only parallel and series substitutions. Hence this graph is constructed in a very similar manner to the smallest simple graph found to have an ATR root outside of the unit disk in [54]. The main difference is the choice of the base graph $G_{3,3}^{1,6}$ here (as opposed to the use of $G_{2,2}^{11,1}$ in [54]). For $k \leq 8$, we have found that the ATR roots of the graph $G^{(k,3)}$ all lie inside the unit disk. However, when $k = 9$, an ATR root is pushed outside of the unit disk, as we prove in the proposition below.

Proposition 2.2.9. *The simple graph* $G^{(9,3)} = G_{3,3}^{9,54}[K_3^-(u,v)]$ *on* 546 *vertices and* 1080 *edges has an ATR root outside of the unit disk.*

Proof. Recall that a particular ATR root R of the graph $G_{3,3}^{1,6}$ satisfies

$$0.69659 \leq \mathrm{Re}(R) \leq 0.69660 \text{ and } 0.77393 \leq \mathrm{Im}(R) \leq 0.77394.$$

From these bounds we are able to obtain

$$-1.01749 \leq \mathrm{Re}\left(\frac{\sqrt[9]{R}}{1-\sqrt[9]{R}}\right) \leq -1.01731 \text{ and } 10.70762 \leq \mathrm{Im}\left(\frac{\sqrt[9]{R}}{1-\sqrt[9]{R}}\right) \leq 10.70814.$$

It suffices to show that the polynomial

$$f_3(q) = \frac{\mathrm{spRel}_{\{u,v\}}(K_3^-;q)}{1-q} - (a+bi)\cdot\frac{\mathrm{Rel}(K_3^-;q)}{1-q}.$$

has a root outside of the unit disk for *any* real numbers a and b satisfying

$$-1.01749 \leq a \leq -1.01731 \quad \text{and} \quad 10.70762 \leq b \leq 10.70814. \tag{2.9}$$

Working either directly from the definitions or using the formulae of Proposition 2.2.7 and 2.2.8, we find that $\mathrm{Rel}(K_3^-;q) = (1-q)^2$ and $\mathrm{spRel}_{\{u,v\}}(K_3^-;q) = 2q(1-q)$. Substituting these polynomials into the equation for $f_3(q)$, we obtain

$$\begin{aligned} f_3(q) &= 2q - (a+bi)(1-q) \\ &= (2+a+bi)q - (a+bi). \end{aligned}$$

Applying the test of Theorem 2.2.5 to $f_3(q)$ with a and b as parameters, we get the single determinant

$$M_1 = 4a + 4.$$

In particular, from our bounds we know that $a < -1$, and hence $M_1 < 0$. Therefore, $\mathrm{B}(1, M_1) = 1$ (recall that $\mathrm{B}(a_1, \ldots, a_n)$ is the number of sign changes in the sequence $a_1, \ldots, a_n$), and we conclude that $f_3(q)$ has a root outside the unit disk for any a and b satisfying (2.9). $\square$

Since the graph $G_{3,3}^{9,54}[K_3^-(u,v)]$ has 546 vertices and 1080 edges, it is just over one third of the size of the smallest previously known simple graph with ATR roots outside of the unit disk (the smallest such graph found in [54] has 1512 vertices and 3016 edges). We stress that the only real difference between our graph and the graph from [54] is the choice of the graph that we start from before performing the edge substitutions. We tested many different base graphs of the form $G_{m,n}^{a,b}$ and the base graph $G_{3,3}^{1,6}$ produced the smallest simple graph with ATR roots outside of the unit disk.

It may seem as though our use of Theorem 2.2.5 in Proposition 2.2.9 is a little bit heavy-handed, as $f_3(q)$ turned out to be a linear function, and we could have verified that its single root was outside of the unit disk directly. However, Theorem 2.2.5 plays a much more important role in the proof of the following proposition. While we have observed that all of the ATR roots of the graph $G^{(k,4)} = G_{3,3}^{k,6k}[K_4^-(u,v)]$ are inside the unit disk for $k \le 6$, when $k = 7$ an ATR root is pushed outside.

Proposition 2.2.10. *The* 3*-edge connected simple graph* $G^{(7,4)} = G_{3,3}^{7,42}[K_4^-(u,v)]$ *has an ATR root outside of the unit disk.*

Proof. Recall that a particular ATR root R of the graph $G_{3,3}^{1,6}$ satisfies

$$0.69659 \le \mathrm{Re}(R) \le 0.69660 \text{ and } 0.77393 \le \mathrm{Im}(R) \le 0.77394.$$

From these bounds we obtain

$$-0.90269 \le \mathrm{Re}\left(\frac{\sqrt[7]{R}}{1-\sqrt[7]{R}}\right) \le -0.90254 \text{ and } 8.32420 \le \mathrm{Im}\left(\frac{\sqrt[7]{R}}{1-\sqrt[7]{R}}\right) \le 8.32462.$$

It suffices to show that the polynomial

$$f_4(q) = \frac{\mathrm{spRel}_{\{u,v\}}(K_4^-;q)}{(1-q)^2} - (a+bi)\cdot\frac{\mathrm{Rel}(K_4^-;q)}{(1-q)^2}.$$

has a root outside of the unit disk for any real numbers a and b satisfying

$$-0.90269 \le a \le -0.90254 \quad \text{and} \quad 8.32420 \le b \le 8.32462. \tag{2.10}$$

Using Proposition 2.2.7 we find

$$\mathrm{Rel}(K_4^-; q) = (1-q)^3(4q^2+3q+1),$$

and Proposition 2.2.8 gives

$$\mathrm{spRel}(K_4^-; q) = 2(1-q)^2(3q+1)q^2.$$

Substituting these expressions into $f_4(q)$ yields

$$f_4(q) = 2(3q+1)q^2 - (a+bi)(1-q)(4q^2+3q+1).$$

Applying Theorem 2.2.5 to $f_4(q)$ with a and b as parameters, we obtain the three determinants

$$\begin{aligned}
M_1 &= 15a^2 + 15b^2 + 48a + 36,\\
M_2 &= 144a^4 + 288a^2b^2 + 144b^4 + 1260a^3\\
&\qquad + 1260ab^2 + 3284a^2 + 980b^2 + 3456a + 1296, \text{ and}\\
M_3 &= 18432a^5 + 36864a^3b^2 + 18432ab^4 + 107520a^4 + 123648a^2b^2 + 16128b^4\\
&\qquad + 252032a^3 + 137344ab^2 + 296576a^2 + 50816b^2 + 175104a + 41472.
\end{aligned}$$

Since the determinants M_1, M_2, and M_3 have all positive coefficients, it is straightforward to bound the determinants using the bounds on a and b from (2.10). We did so using a computer algebra system and found $M_1, M_2 > 0$ and $M_3 < 0$, so that $\mathrm{B}(1, M_1, M_2, M_3) = 1$, and therefore $f_4(q)$ has exactly one solution outside of the unit disk. $\square$

Using the same procedure as in the proof of Proposition 2.2.10, we can demonstrate

n	Smallest value k for which $G^{(k,n)}$ has an ATR root outside of the unit disk	Edge connectivity of $G^{(k,n)}$	Number of vertices of $G^{(k,n)}$	Number of edges of $G^{(k,n)}$
3	9	2	546	1,080
4	7	3	846	2,100
5	6	4	1,086	3,240
6	6	5	1,446	5,040

Table 2.2: Simple graphs with ATR roots outside of the unit disk.

that there are 4-edge connected and 5-edge connected graphs with ATR roots outside of the unit disk. When we substitute the gadgets $K_5^-(u,v)$ and $K_6^-(u,v)$, we find that replacing each edge of $G_{3,3}^{1,6}$ with a bundle of 6 edges is sufficient to push an ATR root outside of the unit disk.

Proposition 2.2.11. *Both the 4-edge connected simple graph $G^{(6,5)}$ and the 5-edge connected simple graph $G^{(6,6)}$ have an ATR root outside of the unit disk.* □

All of the important information about the simple graphs we have found with ATR roots outside of the unit disk is collected in Table 2.2. While we suspect that our technique could be used to prove that there are simple graphs with yet higher edge connectivity with ATR roots outside of the unit disk, the time required to apply Theorem 2.2.5 to the polynomial $f_n(q)$ grows large very quickly; after all, the degree of $f_n(q)$ is $\binom{n-1}{2}$.

Finally, we mention that our technique for finding ATR roots of simple graphs with high edge connectivity has another important application. Let $P = \{z \in \mathbb{C}\colon\ \mathrm{Re}(z) \geq -1/2\}$. It was proven in [12] that roots of the F-polynomial are dense in P. Hence, for any $r \in P$ and any gadget $H(u,v)$, the solutions of

$$\mathrm{spRel}_{\{u,v\}}(H;q) = r \cdot \mathrm{Rel}(H;q) \tag{2.11}$$

are limits of ATR roots by Proposition 2.2.3, as the roots of a polynomial are a continuous function of its coefficients, and there are graphs whose F-polynomials have a root arbitrarily close to r. In particular, when $r = 0$ we have

$$\mathrm{spRel}_{\{u,v\}}(H; q) = 0,$$

which means that the roots of the $\{u, v\}$-split reliability of any gadget are limits of ATR roots. For small n, the polynomial $\mathrm{spRel}_{\{u,v\}}(G^{1,6}_{n,n}; q)$ has roots outside of the unit disk for any choice of u and v. While none of these roots have modulus larger than the roots of the corresponding polynomial $\mathrm{Rel}(G^{1,6}_{n,n}; q)$, this demonstrates the fact that roots of the $\{u, v\}$-split reliability polynomial might be useful if we are interested in finding ATR roots of large modulus.

If we are more creative with our choice of r, we can sometimes find roots of (2.11) that are larger in modulus than any ATR root of H alone. For example, let $r = -1/2 + 3i$, let $H = G^{1,6}_{12,12}$, and let u and v be any two vertices from the same K_{12} in the construction of H. Then (2.11) has roots of modulus slightly larger than any roots of $\mathrm{Rel}(H; q)$, meaning that there are ATR roots of even larger modulus than those we found in Section 2.2.1!

Chapter 3

Node Reliability

In this chapter we study the analytic properties of node reliability, with a particular focus on the shape of the node reliability curve on the interval $[0, 1]$. In Section 3.1 we explore the question of monotonicity of node reliability, while in Section 3.2 we consider the concavity of the curve and the number of points of inflection. In Section 3.3 we study the fixed points of node reliability in $(0, 1)$. Finally, we begin the analysis of node reliability roots in Section 3.4 – this work is continued in Chapters 4 and 5. In order to lay the ground for our results on the shape of node reliability, we first discuss the shape of all-terminal reliability (and the shape of *coherent* reliability polynomials in general) on the interval $[0, 1]$. Our findings on node reliability will demonstrate many interesting contrasts between the models of node reliability and all-terminal reliability.

The all-terminal reliability studied in the previous chapter is based on the premise that the nodes are always operational, but edges are independently operational with probability p. The all-terminal reliability model generalizes to *K-terminal reliability*, which asks the probability that all vertices in some particular subset K can communicate with one another (we call the vertices in K the *target nodes*, with the target nodes ranging from two particular vertices in the well-studied *two-terminal reliability* to the entire vertex set for all-terminal reliability). An excellent survey of these measures can be found in [28].

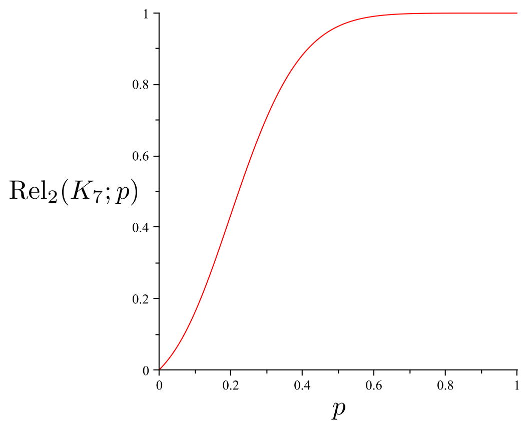
When investigating K-terminal reliability in general, one is struck by how little the size of K affects the structure of the reliability function. For example, consider the graph K_7, the complete graph of order 7. Plots of the all-terminal reliability of K_7, the two-terminal reliability of K_7 (with any two distinguished nodes) and K-terminal reliability of K_7 (with K any vertex subset of size 4) are shown in Figure 3.1. The overall shapes of the all-terminal, two-terminal and K-terminal reliabilities of K_7 on the interval $[0, 1]$ are not so different.

All of these models of reliability fit under the umbrella of *coherence*. Let X be a finite ground set; a *coherent set system* $\mathcal{S}$ on X is a nonempty subset of $\mathcal{P}(X)$ that is closed under taking supersets and does not contain the empty set (this last condition is to ensure nontriviality). The *order* of $\mathcal{S}$ is the cardinality of the ground set X. We think of the elements of X as components of a system that either operate or fail, and so we call the sets in $\mathcal{S}$ the *operational states*. Coherence is then the natural property that if we start with an operational state and make any number of failed components operational it can only improve matters (that is, will not result in a failed state). Let X have cardinality n and suppose that each element of X is independently operational with probability $p \in (0, 1)$. The *reliability* of coherent set system $\mathcal{S}$ on X, denoted $\mathrm{Rel}(\mathcal{S}; p)$, is the probability that the set of operational elements of X is in $\mathcal{S}$; that is,

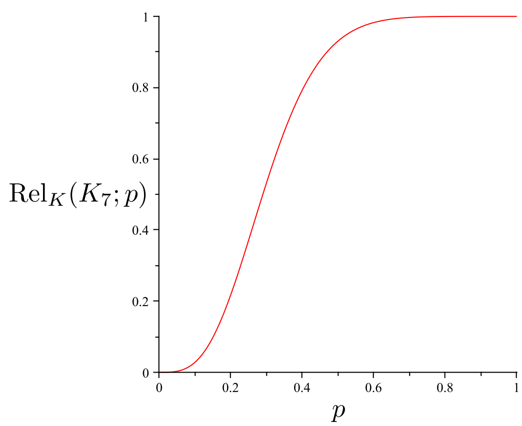
$$\mathrm{Rel}(\mathcal{S}; p) = \sum_{S \in \mathcal{S}} p^{|S|}(1-p)^{n-|S|} \tag{3.1}$$

$$= \sum_{i=1}^{n} N_i p^i (1-p)^{n-i} \tag{3.2}$$

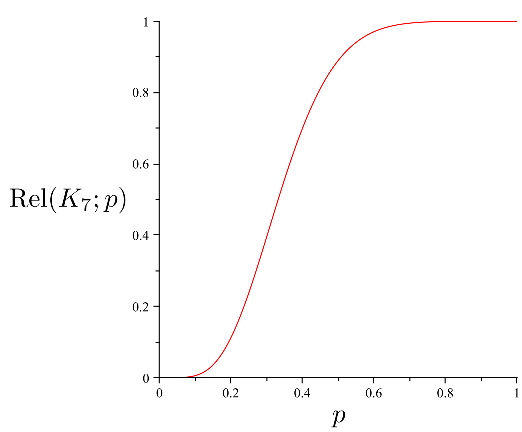
where N_i is the number of operational states of order i for each $i \in \{1, \dots, n\}$. There are obvious relevant coherent set systems underlying each of the network models introduced earlier, all on the edge set of the graph – in general for K-terminal reliability, the operational states are those edge subsets that connect all vertices of K. We note that the graph K_1 is the only connected graph whose underlying set system



(a) Two-terminal reliability of K_7.



(b) K-terminal reliability of K_7 for $|K| = 4$.



(c) All-terminal reliability of K_7.

Figure 3.1: Plots of two-terminal, K-terminal ($|K| = 4$) and all-terminal reliability of the graph K_7.

for K-terminal reliability is not coherent, as it contains the empty set.

Birnbaum, Esary, and Saunders achieved a beautiful result in [6] that describes the general shape of any *coherent reliability polynomial* (that is, the reliability of any coherent set system) on the interval $[0,1]$; it states that any coherent reliability polynomial is strictly increasing on $(0,1)$ and has at most one fixed point in $(0,1)$ (a *fixed point* of a function f is a value x for which $f(x)=x$). Moreover, when written in the form (3.2), the reliability of any coherent set system $\mathcal{S}$ with $N_1=0$ and $N_{n-1}=n$ has a *sigmoid shape* (or *S-shape*) on $[0,1]$; that is,

- $\mathrm{Rel}(\mathcal{S};0)=0$ and $\mathrm{Rel}(\mathcal{S};1)=1$,
- $\mathrm{Rel}'(\mathcal{S};p)>0$ for $p\in(0,1)$,
- $\mathrm{Rel}(\mathcal{S};p)$ has a unique fixed point $\hat{p}\in(0,1)$, and
- $\mathrm{Rel}(\mathcal{S};p)<p$ for $p\in(0,\hat{p})$ and $\mathrm{Rel}(\mathcal{S};p)>p$ for $p\in(\hat{p},1)$.

A typical S-shaped curve (the all-terminal reliability polynomial of the cycle C_4) is shown in Figure 3.2. The conditions $N_1=0$ and $N_{n-1}=n$ mean simply that the system fails whenever at most one component is operational, and that the system is operational whenever at most one component fails, respectively. For all-terminal reliability, these conditions are satisfied if and only if the graph lies on at least 3 vertices and is 2-edge connected (i.e. has no bridges).

Returning to the foundation of the network model, in some situations it is more realistic to assume that the edges are perfectly reliable and the nodes each operate independently with a given probability. Social networking websites seem to be an obvious example of this situation. In such a network, the edges represent friendship, which is (nearly) perfectly reliable. On the other hand, the vertices represent the users, who may be online or offline. In order to communicate effectively, at any given time we would like all of the operational nodes to be able to communicate with one

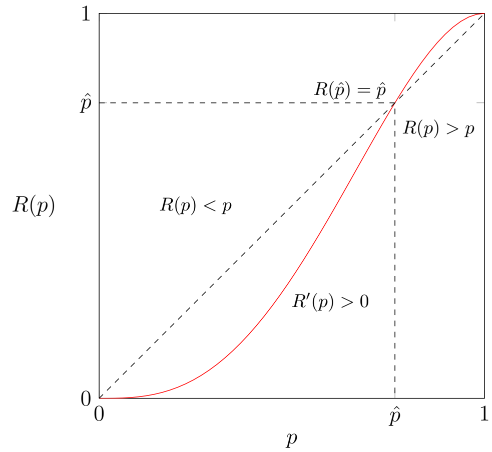


Figure 3.2: A plot of an S-shaped curve $R(p)$.

another. We refer to this model as *node reliability*, condensing the term *residual node connectedness reliability* used in [29, 61, 62], for example.

Definition 3.0.1. Consider a network G consisting of n nodes each operating independently with probability $p \in [0, 1]$. The *node reliability* of G, denoted $\mathrm{nRel}(G; p)$, is the probability that at least one node is operational and that the operating nodes can all communicate in the induced subgraph that they generate.

Like the other measures of reliability we have discussed, the node reliability of a graph is always a polynomial in p, as

$$\mathrm{nRel}(G; p) = \sum_{C \in \mathcal{C}} p^{|C|}(1-p)^{n-|C|}, \tag{3.3}$$

where $\mathcal{C}$ is the collection of all nonempty vertex subsets that induce connected subgraphs of G. We call these sets *connected sets* and refer to $\mathcal{C}$ as the *system of connected sets* of G. As a simple example, the node reliability of the complete graph K_n is

$$\mathrm{nRel}(K_n; p) = 1 - (1-p)^n,$$

as any nonempty subset of vertices induces a connected subgraph (indeed, a complete subgraph). The node reliability of K_n is equivalent to the all-terminal reliability of a bundle of n edges (the multigraph on two vertices with n edges between them).

Like that for all-terminal reliability, much of the existing work on node reliability has concerned itself with finding optimal networks, should they exist, given constraints on the number of vertices and edges allowed (see [34, 48, 61, 74]). Other research concerns the complexity of computing the polynomials; Sutner et al. showed in [62] that the problem of computing the node reliability polynomial is NP-hard, while Colbourn et al. presented efficient algorithms for computing the node reliability polynomial of several restricted families of graphs [29]. Results on both of these problems for node reliability mirror those for all-terminal reliability to a large extent.

In light of the similarity of the formulations (3.1) and (3.3) for general K-terminal reliability and node reliability, and the similarity of discoveries on the two main problems (namely synthesis and computation issues) for all-terminal reliability and node reliability, one cannot help but ask whether there is anything new for node reliability. Our investigation into the analytic structure of node reliability reveals some unexpected and remarkable differences from the structure of all instances of K-terminal reliability, and coherent reliabilities in general. In this chapter, we will show that the shape of node reliability polynomials is strikingly different than that of its coherent relatives. We will also show that the nature and location of the roots of node reliability polynomials differ greatly from those for the roots of all-terminal reliability polynomials (in this chapter and the next).

The first glaring difference between node reliability and all-terminal reliability arises from disconnected graphs. The all-terminal reliability polynomial of a disconnected graph is always identically zero, but the situation is not so trivial for the node reliability polynomial. If G is not connected, then

$$\mathrm{nRel}(G;0) = 0 \quad \text{and} \quad \mathrm{nRel}(G;1) = 0,$$

but $\mathrm{nRel}(G;p) > 0$ for all $p \in (0,1)$. The node reliability of a disconnected graph is easily computed in terms of the node reliabilities of its components.

Observation 3.0.1. *Let G be a graph with connected components $G_1, \ldots, G_k$. The node reliability of G is given by*

$$\mathrm{nRel}(G;p) = \sum_{i=1}^{k} (1-p)^{n-|V(G_i)|}\mathrm{nRel}(G_i;p).$$

Proof. In order for all operational nodes of G to communicate, they must all belong to the same component of G. Thus, the operational nodes in G must all be able to communicate in some component G_i, and all nodes outside of G_i must fail. The probability that at least one node of G_i is operational and that the operational nodes in G_i can all communicate is given by $\mathrm{nRel}(G_i;p)$, and the probability that all nodes outside of G_i fail is $(1-p)^{n-|V(G_i)|}$. Therefore, the node reliability of G is given by

$$\mathrm{nRel}(G;p) = \sum_{i=1}^{k} (1-p)^{n-|V(G_i)|}\mathrm{nRel}(G_i;p). \qquad \square$$

While our focus remains on connected graphs, disconnected graphs will provide several key examples in the material to come.

3.1 Monotonicity

It was proven in [6] that any coherent reliability polynomial is strictly increasing on $(0,1)$. We include our own short proof of this fact here as it is relevant to our work. For any coherent set system $\mathcal{S}$ on a set X of cardinality n, there is an associated set system $\mathcal{C}_{\mathcal{S}}$ on X given by

$$\mathcal{C}_{\mathcal{S}} = \{X - S\colon\; S \in \mathcal{S}\}.$$

The members of $\mathcal{C}_{\mathcal{S}}$ are the sets of components whose failure leaves the graph operational. Since $\mathcal{S}$ is coherent, the set system $\mathcal{C}_{\mathcal{S}}$ is closed under *containment*, making it a simplicial complex. We may write the reliability of the coherent system $\mathcal{S}$ in its *F-form*

$$\mathrm{Rel}(\mathcal{S}; p) = \sum_{k=0}^{n} F_k (1-p)^k p^{n-k},$$

where F_k is the number of sets of cardinality k (in simplicial complex parlance, the number of *faces* of cardinality k) in the complex $\mathcal{C}_{\mathcal{S}}$. These coefficients satisfy *Sperner's bounds* [59] for complexes:

$$(k+1)F_{k+1} \leq (n-k)F_k,$$

for $k \in \{0, \ldots, n-1\}$. A straightforward computation yields

$$\mathrm{Rel}'(\mathcal{S}; p) = \sum_{k=0}^{n} [(n-k)F_k - (k+1)F_{k+1}](1-p)^k p^{n-k-1}. \tag{3.4}$$

Note that $F_n = 0$ as $\emptyset \notin \mathcal{S}$ by the definition of coherent set system. Further, note that $\mathcal{C}_{\mathcal{S}}$ is not empty as $\mathcal{S}$ is not empty (again by the definition of coherent set system). Let t be the largest integer for which $F_t > 0$. The coefficient of the term corresponding to $k = t$ in (3.4) is strictly positive as $F_t > 1$ but $F_{t+1} = 0$, and the remaining coefficients are nonnegative by Sperner's bounds. We conclude that the coherent reliability polynomial of any coherent set system is strictly increasing on $(0, 1)$. As a corollary, the all-terminal reliability of a connected graph of order at least 2 is strictly increasing.

While there might be an expectation that a similar result holds for node reliability, the issue of monotonicity is not so obvious for node reliability. It is easy to see that the system of connected sets of a graph G is coherent if and only if G is complete (any single vertex of a graph always induces a connected subgraph, but pairs of non-adjacent vertices do not). The node reliability polynomial of the complete graph K_n, given by

$$\mathrm{nRel}(K_n; p) = 1 - (1-p)^n,$$

is clearly strictly increasing on $(0, 1)$. On the other hand, the coefficients of the F-form of the node reliability of a graph G,

$$\mathrm{nRel}(G; p) = \sum_{k=0}^{n} F_i (1-p)^i p^{n-i},$$

fail to satisfy Sperner's bounds (the essential inequalities used above in the proof of monotonicity for coherent reliability polynomials) whenever G is not complete, as then $F_{n-1} = n$ and $F_{n-2} < n(n-1)/2$. However, in spite of the failure of Sperner's bounds, there are non-complete graphs whose node reliability polynomials are always increasing on $(0, 1)$. For example, the node reliability of the complete bipartite graph $K_{n,n}$ is given by

$$(1-(1-p)^n)^2 + 2np(1-p)^{2n-1} = 1 - 2(1-p)^n + 2np(1-p)^{2n-1} + (1-p)^{2n}.$$

By a straightforward computation,

$$\mathrm{nRel}'(K_{n,n}; p) = 2n(1-p)^{n-1}\left[1 - (2n-1)p(1-p)^{n-1}\right],$$

and its sign is the same as that of $1 - (2n-1)p(1-p)^{n-1}$ for all $p \in (0, 1)$. Setting

$f_n(p) = (2n-1)p(1-p)^{n-1}$, we see that

$$f_n'(p) = (2n-1)(1-p)^{n-2}(1-np),$$

so that $p = \frac{1}{n}$ is the unique critical point of f_n in $(0,1)$, and f_n is maximized there. We will demonstrate that

$$f_n\left(\tfrac{1}{n}\right) = \tfrac{2n-1}{n}\left(1 - \tfrac{1}{n}\right)^{n-1} < 1$$

for all $n \geq 2$. Clearly $\frac{2n-1}{n} < 2$, and it remains to show that $\left(1 - \frac{1}{n}\right)^{n-1} < \frac{1}{2}$. It is sufficient to show that the function

$$t(x) = \left(1 - \tfrac{1}{x}\right)^{x-1}$$

is decreasing on $[2, \infty)$ since $t(2) = \frac{1}{2}$. A straightforward computation gives

$$\tfrac{t'(x)}{t(x)} = \ln\left(1 - \tfrac{1}{x}\right) + \tfrac{1}{x}.$$

From the power series

$$\ln(1-y) = -y - \tfrac{y^2}{2} - \tfrac{y^3}{3} - \ldots$$

which converges on $(0,1)$, we see that $\ln(1-y) < -y$ for $y \in (0,1)$. Hence we have

$$\tfrac{t'(x)}{t(x)} = \ln\left(1 - \tfrac{1}{x}\right) + \tfrac{1}{x} < -\tfrac{1}{x} + \tfrac{1}{x} = 0$$

for $x \in [2, \infty)$. Since $t(x) > 0$ for $x \in [2, \infty)$ we conclude that $t'(x) < 0$ and thus $t(x)$ is decreasing on $[2, \infty)$. Hence we have $\left(1 - \frac{1}{n}\right)^{n-1} < \frac{1}{2}$ and

$$f_n\left(\tfrac{1}{n}\right) < 2 \cdot \tfrac{1}{2} = 1.$$

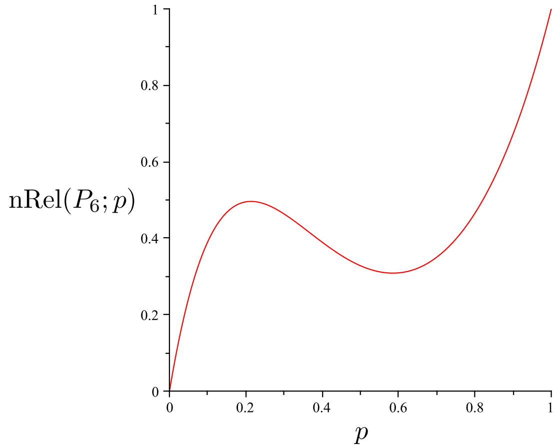


Figure 3.3: Plot of the node reliability of a path of order 6.

Therefore, $\text{nRel}'(K_{n,n}; p) > 0$ for all $p \in (0,1)$ for any $n \geq 2$, and we conclude that $\text{nRel}(K_{n,n}; p)$ is increasing on $(0,1)$.

More surprisingly, there are graphs whose node reliability polynomials are not always increasing on $(0,1)$ – it does not even appear to be very rare for the node reliability polynomial to have an interval of decrease in $(0,1)$! Figure 3.3 shows a plot of the node reliability of a path of order 6, and an interval of decrease between 0.2137 and 0.5851 is clearly evident. Using a computer algebra system, we have found that the node reliability polynomials of 37 of the 112 connected graphs of order 6 have an interval of decrease, while the node reliability polynomials of 383 of the 853 connected graphs of order 7 have an interval of decrease. We will prove in Theorem 3.1.3 that *all* graphs that are not too dense have an interval of decrease. Intuitively, the reason is that when p is close to $\frac{1}{n}$, there is a fairly high probability that exactly one vertex is operating. As p increases from $\frac{1}{n}$, we are likely to have multiple operational nodes (but still not many), and since the graph is not dense it is unlikely that such a small set of nodes will induce a connected subgraph. We now develop the theory to obtain a formal proof of this fact.

The expression for the node reliability given in (3.3) gives rise to the convenient form

$$\mathrm{nRel}(G;p) = \sum_{k=1}^{n} c_k p^k (1-p)^{n-k}, \tag{3.5}$$

where $c_k = c_k(G)$ is the number of connected sets of G of order k for each $k \in \{1, \ldots, n\}$. (Recall that a subset C of vertices is called a connected set if and only if it is nonempty and the induced subgraph $G[C]$ on C is connected.) We remark that while the problem of counting the number of connected sets in a graph has been studied in several different places in the literature [7, 56, 62, 63, 64, 73, 75, 76], very little of this work distinguishes between connected sets of different orders, which node reliability inherently does.

We refer to (3.5) as the *c-form* of the node reliability, and we refer to the coefficients of the c-form collectively as the *c-coefficients* of the node reliability polynomial. The following straightforward observation giving the exact values of certain c-coefficients was made in [61].

Observation 3.1.1. *Let G be a connected graph of order n and size m, let τ be the number of triangles of G, and let t be the number of cut vertices of G. Then*

(i) $c_1 = n$,

(ii) $c_2 = m$,

(iii) $c_3 = \left(\sum_{v \in V(G)} \binom{\deg(v)}{2} \right) - 2\tau$,

(iv) $c_{n-1} = n - t$, *and*

(v) $c_n = 1$. □

As an extension of this observation we note that if p is the order of a smallest *vertex cut* in G (a vertex cut is a set of vertices whose removal disconnects the graph), then $c_{n-k} = \binom{n}{k}$ for each $k \in \{0, \ldots, p-1\}$ and $c_{n-p} = \binom{n}{p} - t_p$ where t_p is the number of vertex cuts of order p.

We will not require explicit formulae for any of the other coefficients; we will simply bound them in terms of these known coefficients. We prove upper bounds on the c-coefficients of the node reliability polynomial in terms of lower c-coefficients. These bounds are similar in spirit to Sperner's bounds, which were used in the proof that any coherent reliability polynomial is increasing on $(0, 1)$.

Lemma 3.1.2. *For any graph G in which a largest component has order d,*

$$2c_k \leq (d-k+1)c_{k-1}$$

for all $k \in \{2, \ldots, d\}$. More generally,

$$(k-t+1)c_k \leq \binom{d-t}{k-t} c_t$$

for all $k \in \{2, \ldots, d\}$ and $t \in \{1, \ldots, k-1\}$.

Proof. We first prove that for any connected graph H of order $n \geq 1$,

$$c_k(H) \geq n-k+1$$

for each $k \in \{1, \ldots, n\}$. We proceed by induction on n. For the base case, when $n = 1$ we have $c_1 = 1 \geq 1-1+1$, and the statement is verified. Now suppose that for some $n \geq 2$, any connected graph H' of order $n-1$ satisfies $c_k(H') \geq n-k$ for all $k \in \{1, \ldots, n-1\}$. Let H be a connected graph of order n. Let v be a vertex whose removal does not disconnect H (such a vertex must exist – take a leaf of some spanning tree of H, for example). There are exactly $c_k(H-v)$ connected sets of G of order k that do not contain v, and there must be at least one connected set of H

of order k containing v by the following argument. Every connected graph H has a unique connected set of order 1 containing v (namely the singleton $\{v\}$), and given a connected set C of order $k \in \{1, \ldots, n-1\}$ in H, there must always be a vertex in $V(H) - C$ that can be added to C to form a connected set of order $k+1$, as otherwise H is not connected. Thus we have

$$c_k(H) \geq c_k(H - v) + 1.$$

Now by the induction hypothesis applied to $H - v$,

$$c_k(H) \geq c_k(H - v) + 1 \geq n - k + 1.$$

Now we are ready to prove the statement of the lemma. Let G be a graph in which a largest component has order d. For each $k \in \{1, \ldots, d\}$, let C_k be the collection of connected sets of G of order k. For any $k \geq 2$, consider a member S of C_k. The induced subgraph $G[S]$ contains at least $k - t + 1$ connected sets of order t by the argument in the previous paragraph. Clearly, any connected set of $G[S]$ must also be a connected set of G. Therefore, every member of C_k can be written in the form $W \cup X$ in at least $k - t + 1$ distinct ways, where $W \in C_t$ and X is some subset of $k - t$ vertices chosen from among the remaining vertices of the component containing W. The total number of such pairs (W, X) is at most

$$\binom{d-t}{k-t} c_t,$$

as any component contains at most d vertices. Since each member of C_k arises from at least $k - t + 1$ of these pairs, we have

$$(k - t + 1) c_k \leq \binom{d-t}{k-t} c_t \qquad \square$$

While we used Sperner's Bounds to show that the reliability polynomial of any coherent set system is increasing on $(0, 1)$, we will use the bounds of Lemma 3.1.2 to show that the node reliability of any graphs with few enough edges has an interval of decrease in the interval $(0, 1)$.

Theorem 3.1.3. *If G is a graph of order n and size $m \leq 0.0851n^2$, then $\mathrm{nRel}(G; p)$ has an interval of decrease in $(0, 1)$. In particular, $\mathrm{nRel}'\big(G; \frac{\hat{r}}{n}\big) < 0$, where $\hat{r} \approx$ 1.729474372.*

Proof. Let G be as in the theorem statement. A straightforward computation gives

$$\mathrm{nRel}'(G; p) = \sum_{k=1}^{n} p^{k-1}(1-p)^{n-k}[kc_k - (n-k+1)c_{k-1}], \tag{3.6}$$

where c_k is the number of connected sets of G of order k for $k \in \{0, \ldots, n\}$ (recall that the empty set is not considered to be a connected set, so $c_0 = 0$, and if a largest component of G has order $d < n$ then $c_{d+1} = \ldots = c_n = 0$). We find directly using the facts that $c_1 = n$ and $c_2 = m$ from Observation 3.1.1 that the sum of the first two terms (corresponding to $k = 1$ and $k = 2$) of the sum in (3.6) is given by:

$$(1-p)^{n-2}[n(1-np) + 2mp] \tag{3.7}$$

We now bound the remaining terms in the sum from (3.6) for any $p \in (0, 1)$. For ease of reading we let

$$\sigma = \sum_{k=3}^{n} p^{k-1}(1-p)^{n-k}[kc_k - (n-k+1)c_{k-1}], \tag{3.8}$$

so that

$$\mathrm{nRel}'(G; p) = (1-p)^{n-2}[n(1-np) + 2mp] + \sigma. \tag{3.9}$$

We claim that $\sigma \leq m\Big[p - \frac{1}{n-1} + \frac{(1-p)^{n-1}}{n-1}\Big]$. We first use the fact that

$$(n-k+1)c_{k-1} \geq 2c_k$$

for all $k \in \{2, \ldots, n\}$ by Lemma 3.1.2. This means that

$$kc_k - (n-k+1)c_{k-1} \leq (k-2)c_k,$$

so that from (3.8) we obtain

$$\sigma \leq \sum_{k=3}^{n} p^{k-1}(1-p)^{n-k}(k-2)c_k. \tag{3.10}$$

Using the more general version of Lemma 3.1.2 now for $t = 2$, we have

$$(k-1)c_k \leq \binom{n-2}{k-2}c_2 = m\binom{n-2}{k-2}$$

for all $k \in \{3, \ldots, n\}$. From (3.10), we obtain

$$\sigma \leq m \cdot \sum_{k=3}^{n} p^{k-1}(1-p)^{n-k}\left(\frac{k-2}{k-1}\right)\binom{n-2}{k-2}. \tag{3.11}$$

The sum in (3.11) can be evaluated using binomial identities to obtain

$$\sigma \leq m\Big[p - \tfrac{1}{n-1} + \tfrac{(1-p)^{n-1}}{n-1}\Big], \tag{3.12}$$

as claimed.

Substituting the upper bound for σ from (3.12) into (3.9), we get

$$\begin{aligned} \mathrm{nRel}'(G;p) &\le (1-p)^{n-2}[n(1-np)+2mp] + m\Big[p - \tfrac{1}{n-1} + \tfrac{(1-p)^{n-1}}{n-1}\Big] \qquad (3.13) \\ &= (1-p)^{n-2}\big[n(1-np) + m\big(2p + \tfrac{1-p}{n-1}\big)\big] + m\big[p - \tfrac{1}{n-1}\big]. \qquad (3.14) \end{aligned}$$

We substitute $p = \frac{r}{n}$ into (3.14) for some fixed $r \in (1,2)$; the exact value of r will be determined shortly.

$$\begin{aligned} \mathrm{nRel}'\big(G;\tfrac{r}{n}\big) &\le \big(1-\tfrac{r}{n}\big)^{n-2}\Big[n(1-r) + m\Big(\tfrac{2r}{n} + \tfrac{n-r}{n(n-1)}\Big)\Big] + m\big[\tfrac{r}{n} - \tfrac{1}{n-1}\big] \\ &< \big(1-\tfrac{r}{n}\big)^{n-2}\big[n(1-r) + \tfrac{m}{n}(2r+1)\big] + \tfrac{m}{n}(r-1) \\ &= \big(1-\tfrac{r}{n}\big)^{n-2}\Big[n(1-r) + \tfrac{m}{n}(2r+1) + \tfrac{m}{n}(r-1)\big(1-\tfrac{r}{n}\big)^{-(n-2)}\Big] \end{aligned}$$

We now show that $\big(1-\frac{r}{n}\big)^{-(n-2)} < e^r$. We set $f(x) = \big(1-\frac{r}{x}\big)^{-(x-2)}$ for $x \in (r,\infty)$, and find easily that

$$f'(x) = \big(1-\tfrac{r}{x}\big)^{-(x-2)}\left(-\ln\Big(1-\frac{r}{x}\Big) + \frac{r(2-x)}{x^2\big(1-\frac{r}{x}\big)}\right).$$

Using the fact that $-\ln(1-y) > y$ for $y \in (0,1)$, we see that

$$-\ln\Big(1-\frac{r}{x}\Big) + \frac{r(2-x)}{x^2\big(1-\frac{r}{x}\big)} > \frac{r}{x} + \frac{r(2-x)}{x^2\big(1-\frac{r}{x}\big)} = \frac{r(2-r)}{x(x-r)},$$

and hence

$$f'(x) > \big(1-\tfrac{r}{x}\big)^{-(x-2)}\left(\frac{r(2-r)}{x(x-r)}\right) > 0$$

since we fixed $r \in (1,2)$ and $x > r$. Thus $\big(1-\frac{r}{n}\big)^{-(n-2)}$ is increasing, and as

$\lim a_n = e^r$, we have that $\left(1 - \frac{r}{n}\right)^{-(n-2)} < e^r$. Thus,

$$\mathrm{nRel}'(G; \tfrac{r}{n}) < \left(1 - \tfrac{r}{n}\right)^{n-2}\left[n(1-r) + \tfrac{m}{n}(2r + 1 + (r-1)e^r)\right].$$

We find that if

$$n(1-r) + \tfrac{m}{n}(2r + 1 + (r-1)e^r) \leq 0 \tag{3.15}$$

then $\mathrm{nRel}'(G; \frac{r}{n}) < 0$ and $\mathrm{nRel}(G; p)$ is decreasing at $p = \frac{r}{n}$. We rearrange (3.15) to obtain the sufficient condition

$$m \leq \left[\tfrac{r-1}{2r+1+(r-1)e^r}\right] n^2. \tag{3.16}$$

Using a computer algebra system, we find that the function

$$f(r) = \tfrac{r-1}{2r+1+(r-1)e^r}$$

reaches a maximum on $(1, \infty)$ of

$$f(\hat{r}) \approx 0.08510464442$$

where $\hat{r} \approx 1.729474372$. As we assumed that $m \leq 0.0851n^2$, we conclude that $\mathrm{nRel}(G; p)$ has an interval of decrease in $(0, 1)$. $\square$

Remark 3.1.1. One can solve for $\hat{r}$ of Theorem 3.1.3 exactly in terms of the well-known Lambert W function (see [30], for example). The Lambert W function is actually a set of functions; it consists of the branches of the inverse relation of the function

$$f(x) = xe^x.$$

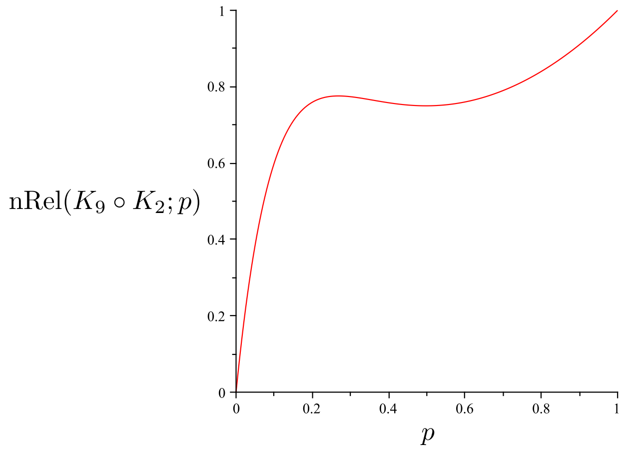


Figure 3.4: The node reliability of $K_9 \circ K_2$.

As a relation on real numbers the Lambert W function is defined only for $x \geq -\frac{1}{e}$. It is double-valued on $(-\frac{1}{e}, 0)$ and single-valued on $[0, \infty)$. The restriction $W \geq -1$ yields a single-valued function $W_0(x)$ on $(-\frac{1}{e}, \infty)$, while the lower branch on $(-\frac{1}{e}, 0)$ is denoted $W_{-1}(x)$. A computer algebra system gives

$$\hat{r} = 2W_0\left(\frac{1}{2}\sqrt{\frac{3}{e}}\right) + 1.$$

Given that graphs which are far from dense all have an interval of decrease in $(0, 1)$, one might ask how dense a graph needs to be to ensure that its node reliability polynomial is increasing on the entire interval $(0, 1)$, and indeed, it must be very dense. Consider the graph formed from the complete graph K_{n-1} by adding a single pendant edge. Let us denote this graph by $K_{n-1} \circ K_2$ (as it is indeed a *vertex bonding* of K_{n-1} and K_2). Note that $K_{n-1} \circ K_2$ has n vertices and only $n - 2$ nonedges.

The reader can verify that

$$\mathrm{nRel}(K_{n-1} \circ K_2; p) \quad = \quad 1 - p(1-p) + p(1-p)^{n-1} - (1-p)^n \qquad (3.17)$$

for all $n \geq 2$. We find that

$$\mathrm{nRel}'(K_{n-1} \circ K_2; p) \;=\; 2p - 1 + (1-p)^{n-2}(n - 2np + 1), \tag{3.18}$$

and evaluating at $p = \frac{2}{5}$ gives

$$-\tfrac{1}{5} + \left(\tfrac{3}{5}\right)^{n-2}\left(\tfrac{1}{5}n + 1\right).$$

We set $g(x) = -\frac{1}{5} + \left(\frac{3}{5}\right)^{x-2}\left(\frac{1}{5}x + 1\right)$ and find that

$$g'(x) = \left(\tfrac{3}{5}\right)^{x-2}\left[\ln\left(\tfrac{3}{5}\right)\left(\tfrac{1}{5}x + 1\right) + \tfrac{1}{5}\right],$$

which is negative for all $x \geq 0$. Further, we find that $g(7) < 0$, so that $g(n) < 0$ for all $n \geq 7$. Since $g(n) = \mathrm{nRel}'\left(K_{n-1} \circ K_2; \frac{2}{5}\right)$ for all integers $n \geq 2$ we conclude that $\mathrm{nRel}(K_{n-1} \circ K_2; p)$ is decreasing at $p = \frac{2}{5}$ for all $n \geq 7$. Figure 3.1 shows a plot of $\mathrm{nRel}(K_9 \circ K_2; p)$ which has a clearly evident interval of decrease.

We close this section with a brief look at the intervals of decrease in $(0, 1)$ of the node reliabilities of disconnected graphs. The node reliability of any disconnected graph G is decreasing on some interval $(1 - \varepsilon, 1)$, as $\mathrm{nRel}(G; p) > 0$ for all $p \in (0, 1)$ and $\mathrm{nRel}(G; 1) = 0$. More surprisingly, there are disconnected graphs whose node reliability polynomials have two distinct maximal intervals of decrease in $(0, 1)$. For example, consider the graph formed from the disjoint union of a single vertex and the star $K_{1,n-1}$. A plot of $\mathrm{nRel}(K_{1,19} \cup K_1; p)$ is shown in Figure 3.5, and one can see two separate intervals of decrease. In fact, we can prove that $\mathrm{nRel}(K_{1,n-1} \cup K_1; p)$ has two distinct maximal intervals of decrease in $(0, 1)$ for $n \geq 12$.

Proposition 3.1.4. *For any $n \geq 12$, the node reliability of $K_{1,n-1} \cup K_1$ has at least two distinct maximal intervals of decrease in* $(0, 1)$.

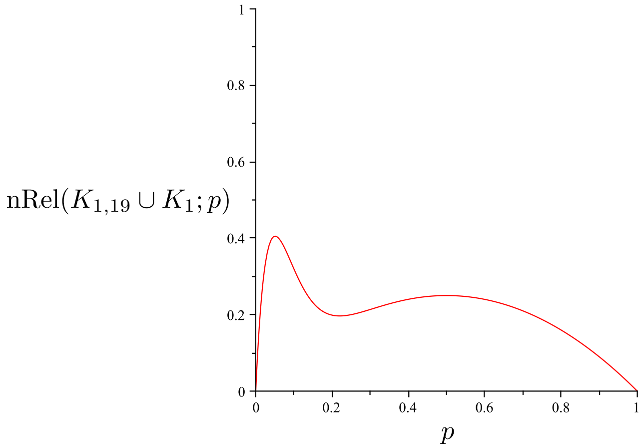


Figure 3.5: The node reliability of $K_{1,19} \cup K_1$.

Proof. First note that

$$\mathrm{nRel}(K_{1,n-1}; p) = p + (n-1)p(1-p)^{n-1},$$

as either the central vertex is operational (this occurs with probability p) or some leaf is operational and all other vertices fail (this occurs with probability $p(1-p)^{n-1}$ for each leaf). By Observation 3.0.1 we have

$$\begin{aligned}\mathrm{nRel}(K_{1,n-1} \cup K_1; p) &= (1-p)\mathrm{nRel}(K_{1,n-1}; p) + p(1-p)^n \\ &= p(1-p) + np(1-p)^n.\end{aligned}$$

By a straightforward computation,

$$\mathrm{nRel}'(K_{1,n-1} \cup K_1; p) = 1 - 2p + n(1-p)^{n-1}(1-(n+1)p).$$

We substitute $p = \frac{2}{5}$ to obtain

$$\mathrm{nRel}'\left(K_{1,n-1} \cup K_1; \tfrac{2}{5}\right) = \tfrac{1}{5} + n\left(\tfrac{3}{5}\right)^{n-1}\left(\tfrac{3-2n}{5}\right).$$

One can verify that the function

$$f(x) = \tfrac{1}{5} + x\left(\tfrac{3}{5}\right)^{x-1}\left(\tfrac{3-2x}{5}\right)$$

is increasing for all $x > 5$ and that $f(12) > 0$, which implies that $f(x) > 0$ for all $x \geq 12$. This means that $\mathrm{nRel}'\big(K_{1,n-1} \cup K_1; \frac{2}{5}\big) > 0$ for $n \geq 12$.

This gives us a point near the middle of the interval $(0, 1)$ at which the node reliability of $K_{1,n-1} \cup K_1$ is increasing. Now it suffices to show that there is a point on either side of $p = \frac{2}{5}$ at which the node reliability is decreasing. By Theorem 3.1.3, when $n \geq 10$ we have $\mathrm{nRel}'\big(K_{1,n-1} \cup K_1; \frac{\hat{r}}{n}\big) < 0$, where $\hat{r} \approx 1.729474372$ (so in particular, $\frac{\hat{r}}{n} < \frac{2}{5}$ when $n \geq 10$). Further, since $\mathrm{nRel}(K_{1,n-1} \cup K_1; 1) = 0$ and $\mathrm{nRel}\big(K_{1,n-1} \cup K_1; \frac{2}{5}\big) > 0$, there must be some point $p \in \big(\frac{2}{5}, 1\big)$ such that $\mathrm{nRel}'(K_{1,n-1} \cup K_1; p) < 0$ by the Mean Value Theorem. For any $n \geq 12$, we have shown that $\mathrm{nRel}(K_{1,n-1} \cup K_1; p)$ is decreasing at $\frac{\hat{r}}{n} < \frac{2}{5}$, increasing at $\frac{2}{5}$, and decreasing at some point in $\big(\frac{2}{5}, 1\big)$. Therefore, for any $n \geq 12$, $\mathrm{nRel}(K_{1,n-1} \cup K_1; p)$ has two distinct maximal intervals of decrease in $(0, 1)$. □

We would be very interested to know if there are any *connected* graphs whose node reliabilities have two distinct maximal intervals of decrease in $(0, 1)$. Using a computer algebra system, we have verified that no such examples exist among all connected graphs of order at most 8.

3.2 Concavity and Inflection Points

We now turn to the question of concavity and points of inflection. In [20], Brown, Koç, and Kooij proved that the all-terminal reliability of almost every simple graph has an inflection point in $(0, 1)$. The arguments there can be extended easily to show that for any coherent set system $\mathcal{S}$ with $N_1 = 0$ and $N_{n-1} = n$, the coherent reliability polynomial $\mathrm{Rel}(\mathcal{S}; p)$ is concave up near $p = 0$ and concave down near $p = 1$.

Hence, under these weak conditions, a coherent reliability polynomial has at least one point of inflection in $(0, 1)$. In [35], Graves demonstrated that coherent reliability polynomials can have two inflection points in $(0, 1)$. Later, several families of all-terminal reliability polynomials having two inflection points in $(0, 1)$ were presented in [20]. Finally, in [36], Graves and Milan proved that all-terminal reliability polynomials of multigraphs can have arbitrarily many inflection points in the interval $(0, 1)$. The existence of simple graphs whose all-terminal reliability polynomials have more than two inflection points is still an open problem. For all-terminal reliability, or more generally for coherent reliability, the families which are known to have more than one point of inflection in $(0, 1)$ are rather thin – very few examples of any particular order n are known (see [20, 35, 36]).

What is the case for node reliability? It is not difficult to see that for any $n \geq 2$ the complete graph on n vertices is concave down on the entire interval $(0, 1)$, as

$$\mathrm{nRel}(K_n; p) = 1 - (1 - p)^n,$$

so that

$$\mathrm{nRel}''(K - n; p) = -n(n - 1)(1 - p)^{n-2} < 0$$

for $n \geq 2$ and $p \in (0, 1)$. Thus there are graphs whose node reliability has no inflection points in $(0, 1)$. The remainder of this section concerns finding node reliability polynomials with one or more inflection point.

By a straightforward computation, the second derivative of the node reliability of G is given by

$$\mathrm{nRel}''(G; p) = \sum_{k=1}^{n-1} d_k p^{k-1}(1 - p)^{n-k-1}, \tag{3.19}$$

with

$$d_k = (k+1)kc_{k+1} - 2k(n-k)c_k + (n-k+1)(n-k)c_{k-1}, \tag{3.20}$$

where c_k is the number of connected sets of G on k vertices for $k \in \{0, \ldots, n\}$. We use this notation for d_k throughout the remainder of this chapter.

Lemma 3.2.1. *Let G be a graph on $n \geq 2$ vertices. The node reliability of G is concave down near $p = 0$.*

Proof. Consider $\mathrm{nRel}''(G;p)$ in the form given in (3.19). When p is sufficiently close to 0, the sign of $\mathrm{nRel}''(G;p)$ will be the same as the sign of d_1, the coefficient of the $(1-p)^{n-2}$ term. This coefficient is given by

$$d_1 = 2c_2 - 2(n-1)c_1,$$

since $c_0 = 0$ (recall that the empty set is not connected). By Observation 3.1.1, we have

$$d_1 = 2m - 2n(n-1) \leq 2\binom{n}{2} - 2n(n-1) = -n(n-1).$$

Therefore, $\mathrm{nRel}''(G;p) < 0$ for p sufficiently close to 0, and we conclude that the node reliability of G is concave down near $p = 0$. □

Lemma 3.2.1 demonstrates another major difference between the shape of the node reliability polynomial and the shape of the all-terminal reliability polynomial – while the node reliability of any graph on $n \geq 2$ vertices is concave *down* near $p = 0$ by Lemma 3.2.1, the all-terminal reliabiliy of any graph on $n \geq 3$ vertices is concave *up* near $p = 0$ [20].

Near $p = 1$, the node reliability may be concave up or concave down. In the next theorem we prove that the node reliability of any tree T on at least 4 vertices is concave up near $p = 1$ in order to reach the conclusion that the node reliability of T has at least one inflection point in $(0, 1)$. This is again very different from the case for K-terminal reliability (including two-terminal and all-terminal reliability); the K-terminal reliability of a tree T is equal to p^k (with k being the number of edges in a minimum subtree containing all vertices of K), and hence has no inflection points in $(0, 1)$.

Theorem 3.2.2. *Let T be a tree on $n \geq 4$ vertices. The node reliability polynomial of T has at least one point of inflection in* $(0, 1)$.

Proof. First suppose that $T \cong K_{1,n-1}$ for some $n \geq 4$. We have

$$\mathrm{nRel}(K_{1,n-1}; p) = p + (n-1)p(1-p)^{n-1},$$

so straightforward computation gives

$$\mathrm{nRel}''(K_{1,n-1}; p) = (n-1)^2(np-2)(1-p)^{n-3}.$$

We see that $\mathrm{nRel}(K_{1,n-1}; p)$ is concave down on $\left(0, \frac{2}{n}\right)$ and concave up on $\left(\frac{2}{n}, 1\right)$, so that the intended conclusion holds.

Now let T be a tree on $n \geq 4$ vertices that is not isomorphic to $K_{1,n-1}$. Consider $\mathrm{nRel}''(T; p)$ in the form given in (3.19). When p is sufficiently close to 1, the sign of $\mathrm{nRel}''(T; p)$ will be the same as the sign of the coefficient d_{n-1} of the p^{n-2} term (as long as this coefficient is nonzero). The coefficient d_{n-1} is given by

$$d_{n-1} = n(n-1)c_n - 2(n-1)c_{n-1} + 2c_{n-2}. \tag{3.21}$$

By Observation 3.1.1, $c_n = 1$ and $c_{n-1} = n - t$, where t is the number of cut vertices of T. Since T is a tree, we can write $n - t = r$ where r is the number of leaves of T, so that $c_{n-1} = r$. Further, $c_{n-2} = \binom{r}{2} + s$, where s is the number of leaves adjacent to a vertex of degree 2 in G, as the connected sets of order $n-2$ consist of either all vertices but a pair of leaves or all vertices but a leaf and an adjacent vertex of degree 2. Substituting these values into (3.21), we obtain

$$\begin{aligned} d_{n-1} &= n(n-1) - 2(n-1)r + 2\left[\binom{r}{2} + s\right] \\ &= n(n-1) - 2(n-1)r + r(r-1) + 2s \\ &= 2s + n(n-1) - r(2n-r-1). \end{aligned}$$

Since $T \ncong K_{1,n-1}$, we have $r < n-1$ and thus

$$r(2n-r-1) = n(n-1) - (n-r)(n-1-r) < n(n-1).$$

Thus we have

$$d_{n-1} = 2s + n(n-1) - r(2n-r-1) > 2s > 0.$$

We conclude that $\mathrm{nRel}''(T;p)$ is positive for p sufficiently close to 1, and therefore that $\mathrm{nRel}(T;p)$ is concave up near $p=1$. Recall from Lemma 3.2.1 that $\mathrm{nRel}(T;p)$ is concave down near 0. We conclude that $\mathrm{nRel}(T;p)$ has at least one inflection point in $(0,1)$ by the Intermediate Value Theorem applied to $\mathrm{nRel}''(T;p)$. □

In the proof of Theorem 3.2.2 we saw that the node reliability polynomial of the star $K_{1,n-1}$ has exactly one inflection point in $(0,1)$ for any $n \geq 4$, and we conjecture that all trees of order at least 4 have exactly one inflection point in $(0,1)$. While many of the typical 'S-shaped' all-terminal reliability polynomials also appear to have a single point of inflection in $(0,1)$, the node reliability of any tree on $n \geq 4$ vertices appears to have an 'N-shape' on $(0,1)$ as opposed to the 'S-shape' of the all-terminal reliability polynomials. Figure 3.6 provides a plot showing the node

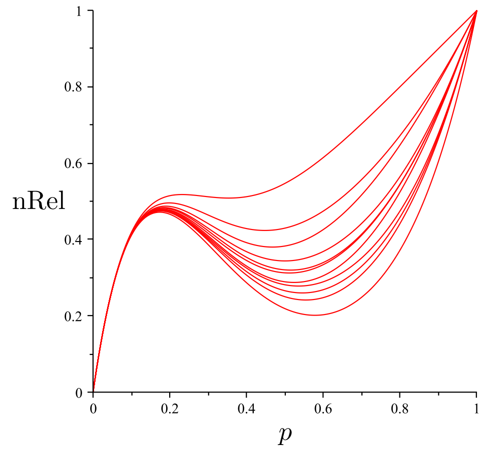


Figure 3.6: Node reliability polynomials of all trees on 7 vertices.

reliability polynomials of all trees on 7 vertices.

We next present a family of graphs whose node reliability polynomials each have at least *two* inflection points in $(0, 1)$. Unlike the examples for coherent and all-terminal reliability, our family provides numerous examples of each order n. We will require the following lemma.

Lemma 3.2.3. *Let G be a 2-connected graph. The node reliability of G is concave down near $p = 1$.*

Proof. Again we consider $\mathrm{nRel}''(G; p)$ in the form given in (3.19). Let t be the order of a smallest vertex cut-set in G (note that $t \geq 2$ as G is 2-connected by assumption). We must have

$$c_k = \binom{n}{k} \text{ for all } k > n - t$$

and

$$c_{n-t} < \binom{n}{t}.$$

Therefore, for any $k > n - t + 1$, the coefficient d_k of $p^{k-1}(1-p)^{n-k-1}$ in $\mathrm{nRel}''(G; p)$

is given by

$$\begin{aligned}
d_k &= (k+1)kc_{k+1} - 2k(n-k)c_k + (n-k+1)(n-k)c_{k-1} \\
&= (k+1)k\binom{n}{k+1} - 2k(n-k)\binom{n}{k} + (n-k+1)(n-k)\binom{n}{k-1} \\
&= \frac{n!}{(k-1)!(n-k-1)!} - 2\frac{n!}{(k-1)!(n-k-1)!} + \frac{n!}{(k-1)!(n-k-1)!} \\
&= 0
\end{aligned}$$

Thus the $p^{n-t}(1-p)^{t-2}$ term is the leading term of $\mathrm{nRel}''(G;p)$ near $p=1$, and the coefficient d_{n-t+1} of this term is given by

$$\begin{aligned}
d_{n-t+1} &= (n-t+2)(n-t+1)c_{n-t+2} \\
&\qquad - 2(n-t+1)(t-1)c_{n-t+1} + t(t-1)c_{n-t} \\
&< (n-t+2)(n-t+1)\binom{n}{n-t+2} \\
&\qquad - 2(n-t+1)(t-1)\binom{n}{n-t+1} + t(t-1)\binom{n}{n-t} \\
&= \frac{n!}{(n-t)!(t-2)!} - 2\frac{n!}{(n-t)!(t-2)!} + \frac{n!}{(n-t)!(t-2)!} \\
&= 0,
\end{aligned}$$

where we used the fact that $c_{n-t} < \binom{n}{n-t}$. Therefore, for p sufficiently close to 1, the coefficient of the leading term of $\mathrm{nRel}''(G;p)$ is negative. We conclude that the node reliability of G is concave down near $p=1$. $\square$

Theorem 3.2.4. *Let G be a graph of order n and size m. If $m \leq 0.0851n^2$ and G is 2-connected then $\mathrm{nRel}(G;p)$ has at least two distinct points of inflection in $(0,1)$.*

Proof. Lemma 3.2.1 tells us that $\mathrm{nRel}(G;p)$ is concave down near 0, and Lemma 3.2.3 tells us that $\mathrm{nRel}(G;p)$ is concave down near 1. By Theorem 3.1.3, $\mathrm{nRel}(G;p)$ contains an interval of decrease in $(0,1)$. In fact, $\mathrm{nRel}'\big(G;\frac{\hat{r}}{n}\big) < 0$, where $\hat{r} \approx 1.729474372$

(see Theorem 3.1.3 and Remark 3.1.1). Since $\mathrm{nRel}(G;p) < 1$ for all $p \in (0,1)$ and $\mathrm{nRel}(G;1) = 1$, $\mathrm{nRel}(G;p)$ must be increasing on some neighbourhood $(\tilde{p},1)$. Let $\hat{p} \in (\tilde{p},1)$ so that $\hat{p} > \frac{\hat{r}}{n}$ and $\mathrm{nRel}'(G;\hat{p}) > 0$. By the Mean Value Theorem, there is some point $c \in \left(\frac{\hat{r}}{n}, \hat{p}\right)$ such that

$$\mathrm{nRel}''(G;c) = \frac{\mathrm{nRel}'(G;\hat{p}) - \mathrm{nRel}'\left(G;\frac{\hat{r}}{n}\right)}{\hat{p} - \frac{\hat{r}}{n}} > 0.$$

Therefore, $\mathrm{nRel}(G;p)$ is concave down on some neighbourhood of 0 and some neighbourhood of 1, and concave up at some point c inside the interval. We conclude that $\mathrm{nRel}(G;p)$ has at least two points of inflection in $(0,1)$ by the Intermediate Value Theorem applied to $\mathrm{nRel}''(G;p)$. $\square$

Theorem 3.2.4 demonstrates that it is not so rare for a node reliabiliy polynomial to have two (or more) points of inflection in $(0,1)$. By comparison, the families of graphs presented in [20] whose all-terminal reliability polynomials have 2 points of inflection contain far fewer graphs of order n.

Are there graphs whose node reliabilities have three (or more) inflection points in $(0,1)$? Indeed, we have found many graphs of small order with three inflection points in $(0,1)$. The graph shown in Figure 3.7 is the unique graph on at most 7 vertices satisfying this property. We have also found that the node reliabilities of 84 of the 11117 nonisomorphic connected graphs on 8 vertices have three points of inflection in $(0,1)$. Is there an infinite family of graphs whose node reliabilities have three points of inflection in $(0,1)$? All of the small graphs that we have found whose node reliabilities have three points of inflection in $(0,1)$ have exactly one leaf and exactly one cut vertex. Is this true for every graph whose node reliability has three points of inflection in $(0,1)$?

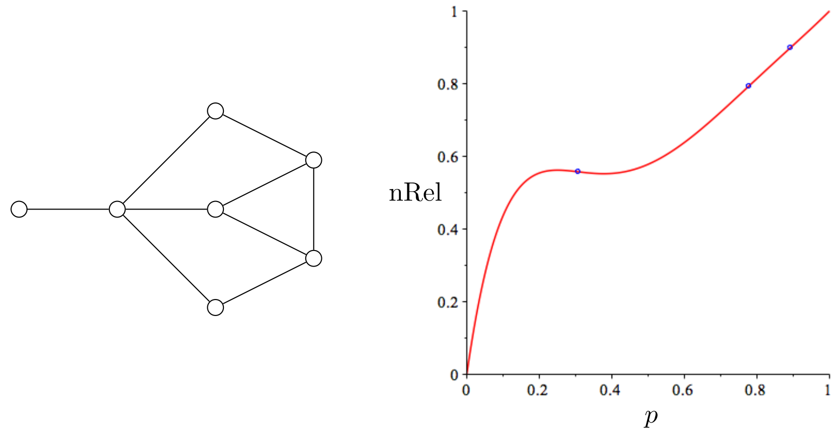


Figure 3.7: The unique graph of order at most 7 whose node reliability has three points of inflection in $(0,1)$. The node reliability is shown to the right, with the inflection points labelled in blue.

3.3 Fixed Points

A key result proven in [6] is that the reliability polynomial of any coherent set system of order at least 2 has at most one fixed point in $(0,1)$. As a corollary, the all-terminal reliability of any connected graph with at least 2 edges has at most one fixed point in $(0,1)$. There are node reliability polynomials with no fixed points in $(0,1)$ (e.g. complete graphs, stars), node reliability polynomials with exactly one fixed point in $(0,1)$ (based on calculations for all graphs of small order it appears that the node reliability of any tree not isomorphic to a star has exactly one fixed point in this interval), and of course, exactly one node reliability polynomial (for the graph K_1) with all $p \in (0,1)$ being fixed points. Surprisingly, there are many node reliability polynomials with two or more distinct fixed points in $(0,1)$. We will prove that the node reliability of any sufficiently large 2-connected graph of bounded degree has at least two fixed points in $(0,1)$. We will require the following lemma.

Lemma 3.3.1. *If G is a connected graph on n vertices with t cut vertices, then*

$$\mathrm{nRel}'(G;0) = n$$

and

$$\mathrm{nRel}'(G;1) = t$$

Proof. By a straightforward computation,

$$\mathrm{nRel}'(G;p) = \sum_{k=1}^{n} p^{k-1}(1-p)^{n-k}[kc_k - (n-k+1)c_{k-1}], \tag{3.22}$$

where c_k is the number of connected sets of G of order k for $k \in \{0, \ldots, n\}$ (recall that $c_0 = 0$). Substituting into (3.22) and using Observation 3.1.1 yields

$$\mathrm{nRel}'(G;0) = c_1 = n$$

and

$$\mathrm{nRel}'(G;1) = nc_n - c_{n-1} = n - (n-t) = t. \qquad \square$$

The following result follows almost immediately from Lemma 3.22.

Corollary 3.3.2. *Let G be a connected graph on $n \geq 2$ vertices having $t \geq 2$ cut vertices. Then* $\mathrm{nRel}(G;p)$ *has at least one fixed point in* $(0,1)$.

Proof. We will show that $\mathrm{nRel}(G;p) > p$ on some interval $(0, p_0)$ and $\mathrm{nRel}(G;p) < p$ on some interval $(p_1, 1)$ by considering the function

$$f(p) = \mathrm{nRel}(G;p) - p.$$

It is clear that $f(0) = f(1) = 0$.

By Lemma 3.22, we have $\mathrm{nRel}'(G;0) = n \geq 2$ and hence $f'(0) = n - 1 \geq 1$. This implies that $f'(p) > 0$ on some interval $(0, p_0)$ since f' is continuous. Suppose towards a contradiction that $f(\hat{p}) \leq 0$ at some point $\hat{p} \in (0, p_0)$. By the Mean Value Theorem, there is some value $c \in (0, \hat{p})$ such that

$$f'(c) = \frac{f(\hat{p}) - f(0)}{\hat{p} - 0} = \frac{f(\hat{p})}{\hat{p}} \leq 0,$$

a contradiction as $f'(p) > 0$ on $(0, p_0)$. Hence $f(p) > 0$ for any $p \in (0, p_0)$, or equivalently $\mathrm{nRel}(G; p) > p$ on $(0, p_0)$.

Similarly, by Lemma 3.22, we have $\mathrm{nRel}'(G; 1) = t \geq 2$ and hence $f'(1) = \mathrm{nRel}'(G; 1) - 1 \geq 1$. This implies that $f'(p) > 0$ on some interval $(p_1, 1)$ since f' is continuous. Suppose towards a contradiction that $f(\hat{p}) \geq 0$ at some point $\hat{p} \in (p_1, 1)$. By the Mean value Theorem, there is some value $c \in (\hat{p}, 1)$ such that

$$f'(c) = \frac{f(1) - f(\hat{p})}{1 - \hat{p}} = \frac{-f(\hat{p})}{1 - \hat{p}} \leq 0,$$

a contradiction as $f'(p) > 0$ on $(p_1, 1)$. Hence $f(p) < 0$ for any $p \in (p_1, 1)$, or equivalently $\mathrm{nRel}(G; p) < p$ on $(p_1, 1)$.

Let $p_0^- \in (0, p_0)$ and $p_1^+ \in (p_1, 1)$, so that $\mathrm{nRel}(G; p_0^-) > p$ and $\mathrm{nRel}(G; p_1^+) < p$. By the Intermediate Value Theorem, $\mathrm{nRel}(G; p) = p$ for some $p \in (p_0^-, p_1^+)$. □

We now prove the main result of this section which demonstrates that there are infinitely many graphs whose node reliability polynomials each have at least two fixed points in $(0, 1)$. We reiterate that this is very different from the case for coherent reliability polynomials, which have at most one fixed point in $(0, 1)$.

Theorem 3.3.3. *Let G be a 2-connected graph on n vertices and let G have maximum degree Δ. For fixed Δ, if n is sufficiently large then $\mathrm{nRel}(G; p)$ has at least two fixed points in $(0, 1)$.*

Proof. By the same argument as in the proof of Corollary 3.3.2, $\mathrm{nRel}(G; p) > p$ on some interval $(0, p_0)$. We now show that $\mathrm{nRel}(G; p) > p$ on some interval $(p_1, 1)$ using a very similar argument. As in the proof of Corollary 3.3.2, let

$$f(p) = \mathrm{nRel}(G;p) - p.$$

By Lemma 3.22, we have $\mathrm{nRel}'(G;1) = 0$ since G is 2-connected and hence has no cut vertices. Thus we have $f'(1) = -1$, and so $f'(p) < 0$ on some interval $(p_1, 1)$ since f' is continuous. Suppose towards a contradiction that $f(p) \leq 0$ at some point $\hat{p} \in (p_1, 1)$. By the Mean Value Theorem, there is some value $c \in (\hat{p}, 1)$ such that

$$f'(c) = \frac{f(1) - f(\hat{p})}{1 - \hat{p}} = \frac{-f(\hat{p})}{1 - \hat{p}} \geq 0,$$

a contradiction as $f'(p) < 0$ for all $p \in (p_1, 1)$. Hence $f(p) > 0$ for any $p \in (p_1, 1)$, or equivalently $\mathrm{nRel}(G;p) > p$ on $(p_1, 1)$.

Now it is sufficient to prove that $\mathrm{nRel}'(G;p) < p$ for some $p \in (0,1)$, as the conclusion will follow from the Intermediate Value Theorem.

We claim that $\mathrm{nRel}\big(G; \frac{1}{\Delta^2}\big) < \frac{1}{\Delta^2}$ for n sufficiently large. The node reliability polynomial of G is given by

$$\mathrm{nRel}(G;p) = np(1-p)^{n-1} + mp^2(1-p)^{n-2} + \sum_{k=3}^{n} c_k p^k (1-p)^k,$$

where c_k is the number of connected sets of G of order k for each $k \in \{3, \ldots, n\}$.

By Observation 3.1.1,

$$c_3 \leq n\binom{\Delta}{2}$$

and so by Lemma 3.1.2,

$$c_k \leq \frac{c_3}{k-2}\binom{n-3}{k-3} \leq \frac{n}{n-2}\binom{\Delta}{2}\binom{n-2}{k-2}$$

for each $k \geq 3$. Thus we have for $p \in (0,1)$ that

$$\begin{aligned}
\sum_{k=3}^{n} c_k p^k (1-p)^k &\leq \sum_{k=3}^{n} \frac{n}{n-2} \binom{\Delta}{2} \binom{n-2}{k-2} p^k (1-p)^{n-k} \\
&= \frac{n}{n-2} \binom{\Delta}{2} \sum_{k=3}^{n} \binom{n-2}{k-2} p^k (1-p)^{n-k} \\
&= \frac{n}{n-2} \binom{\Delta}{2} p^2 \sum_{k=1}^{n-2} \binom{n-2}{k} p^k (1-p)^{n-k-2} \\
&= \frac{n}{n-2} \binom{\Delta}{2} p^2 \left[1-(1-p)^{n-2}\right] \\
&< \frac{n}{n-2} \binom{\Delta}{2} p^2.
\end{aligned}$$

Using this bound on $\displaystyle\sum_{k=3}^{n} c_k p^k (1-p)^k$ and the elementary bound $m \leq \frac{n\Delta}{2}$, we have

$$\begin{aligned}
\mathrm{nRel}(G;p) &< np(1-p)^{n-1} + \frac{n\Delta}{2} p^2 (1-p)^{n-2} + \frac{n}{n-2} \binom{\Delta}{2} p^2 \\
&= np(1-p)^{n-2} \left[(1-p) + \frac{\Delta}{2} p\right] + \frac{n}{n-2} \binom{\Delta}{2} p^2.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\mathrm{nRel}\big(G; \tfrac{1}{\Delta^2}\big) &< \tfrac{n}{\Delta^2} \big(1 - \tfrac{1}{\Delta^2}\big)^{n-2} \big(1 - \tfrac{1}{\Delta^2} + \tfrac{1}{2\Delta}\big) + \tfrac{n}{n-2} \tbinom{\Delta}{2} \tfrac{1}{\Delta^4} \\
&= \tfrac{n}{\Delta^2} \big(1 - \tfrac{1}{\Delta^2}\big)^{n-2} \big(1 - \tfrac{1}{\Delta^2} + \tfrac{1}{2\Delta}\big) + \tfrac{1}{2} \big(\tfrac{n}{n-2}\big) \big(\tfrac{\Delta-1}{\Delta}\big) \tfrac{1}{\Delta^2}
\end{aligned}$$

For $n \geq 2\Delta$ we have

$$\big(\tfrac{n}{n-2}\big) \big(\tfrac{\Delta-1}{\Delta}\big) \leq 1.$$

Therefore,

$$\mathrm{nRel}\big(G; \tfrac{1}{\Delta^2}\big) < n\big(1 - \tfrac{1}{\Delta^2}\big)^{n-2} \big(1 - \tfrac{1}{\Delta^2} + \tfrac{1}{2\Delta}\big) \tfrac{1}{\Delta^2} + \tfrac{1}{2\Delta^2}.$$

Δ	n
2	14
3	40
4	82
5	142
6	220
7	318
8	435
9	573
10	732

Table 3.1: Sufficiently large order n for a 2-connected graph of maximum degree Δ to have node reliability with two fixed points in $(0,1)$.

It is clear that for n sufficiently large we will have

$$n\left(1-\tfrac{1}{\Delta^2}\right)^{n-2}\left(1-\tfrac{1}{\Delta^2}+\tfrac{1}{2\Delta}\right) \leq \tfrac{1}{2} \tag{3.23}$$

as

$$\lim_{n\to\infty} n\left(1-\tfrac{1}{\Delta^2}\right)^{n-2} = 0.$$

We conclude that for n sufficiently large,

$$\mathrm{nRel}\left(G;\tfrac{1}{\Delta^2}\right) < \tfrac{1}{2\Delta^2}+\tfrac{1}{2\Delta^2} = \tfrac{1}{\Delta^2}$$

as claimed, and so the node reliability has at least two fixed points for n sufficiently large. □

Remark 3.3.1. We note that the inequality (3.23) can be solved exactly in terms of the Lambert W function (see Remark 3.1.1) in order to determine just how large n

must be in terms of Δ. Using a computer algebra system, we find

$$n \geq \frac{1}{\ln\left(1-\frac{1}{\Delta^2}\right)} W_{-1}\left(\frac{\left(1-\frac{1}{\Delta^2}\right)^2 \ln\left(1-\frac{1}{\Delta^2}\right)}{2\Delta^2+\Delta-2}\right),$$

where W_{-1} is the lower branch of the Lambert W function. Table 3.1 shows the smallest value n for which (3.23) is satisfied for $\Delta \in \{2, \ldots, 10\}$.

3.4 The Roots of Node Reliability

Along with the shape of the node reliability curve on $(0, 1)$, the roots of the node reliability polynomial have not received very much attention in the literature. Given that the roots of all-terminal reliability have been studied extensively (along with the roots of other models of reliability including strongly-connected reliability [16]), it seems natural to ask questions about the roots of node reliability. Questions involving the realness of the roots, bounding of the roots, and determining the closure of the roots are compelling. We begin our discussion of the roots of node reliability in this section but our deepest results are achieved through the *connected set polynomial*, introduced formally in the next chapter.

A root of the node reliability polynomial of a graph G is called a *node reliability root* of G. It is fairly easy to see that no graph on $n \geq 1$ vertices has a real node reliability root in the interval $(0, 1)$. For any $p \in (0, 1)$, consider

$$\mathrm{nRel}(G; p) = \sum_{k=1}^{n} c_k p^k (1-p)^{n-k}.$$

The coefficient $c_1 = n$, all other coefficients are nonnegative, $p > 0$, and $1 - p > 0$, so that

$$\mathrm{nRel}(G; p) > 0$$

for all $p \in (0, 1)$.

However, the node reliability polynomial can have real roots outside of the interval $(0, 1)$. In fact, we demonstrate that the node reliability polynomial has arbitrarily large real roots. This contrasts the situation for all-terminal reliability polynomials – it was proven in [12] that the real ATR roots of connected graphs are contained in $\{0\} \cup (1, 2]$ in the variable $p = 1 - q$ (this corresponds to the set $[-1, 0) \cup \{1\}$ in the variable q).

Theorem 3.4.1. *For n sufficiently large, the polynomial* $\mathrm{nRel}(C_{2n+1}; p)$ *has a real root in the interval* $(2n^2 - 1, 2n^2)$.

Proof. We begin by finding a convenient closed form for the node reliability of C_{2n+1}. Note first of all that the cycle C_n satisfies $c_k = n$ for all $k \in \{1, \ldots, n - 1\}$, as the connected sets of order k are exactly the n sets of k consecutive vertices. And of course $c_n = 1$ for any connected graph on n vertices. Thus we have

$$\begin{aligned}\mathrm{nRel}(C_{2n+1}; p) &= p^{2n+1} + (2n + 1)\sum_{k=1}^{2n} p^k(1 - p)^{2n+1-k}\\ &= p^{2n+1} + (2n + 1)(1 - p)^{2n+1}\sum_{k=1}^{2n}\left(\tfrac{p}{1-p}\right)^k.\end{aligned}$$

Using the basic sum identity

$$\sum_{k=1}^{n} x^k = \frac{x^{n+1} - x}{x - 1},$$

we obtain

$$\begin{aligned}\mathrm{nRel}(C_{2n+1}; p) &= p^{2n+1} + (2n + 1)(1 - p)^{2n+1}\frac{\left(\tfrac{p}{1-p}\right)^{2n+1} - \tfrac{p}{1-p}}{\tfrac{p}{1-p} - 1}\\ &= p^{2n+1} + (2n + 1)\left[p^{2n+1} - p(1 - p)^{2n}\right] \cdot \tfrac{1-p}{2p-1}\\ &= \tfrac{p}{2p-1}\left[(2p - 1)p^{2n} + (2n + 1)(1 - p)\left(p^{2n} - (p - 1)^{2n}\right)\right].\end{aligned}$$

n	$2n^2-1$	Largest real root of $\mathrm{nRel}(C_{2n+1};p)$	$2n^2$
2	7	$7.623283978\ldots$	8
3	17	$17.72122653\ldots$	18
4	31	$31.75632217\ldots$	32
5	49	$49.77458407\ldots$	50
6	71	$71.78581587\ldots$	72
7	97	$97.79343113\ldots$	98
8	127	$127.7989376\ldots$	128
9	161	$161.8031061\ldots$	162
10	199	$199.8063720\ldots$	200

Table 3.2: The largest real root of $\mathrm{nRel}(C_{2n+1};p)$ for $n \in \{2,\ldots,10\}$.

We have verified using a computer algebra system that

$$\lim_{n\to\infty} \mathrm{nRel}(C_{2n+1}; 2n^2-1) = -\infty$$

and

$$\lim_{n\to\infty} \mathrm{nRel}(C_{2n+1}; 2n^2) = \infty.$$

Thus for n sufficiently large, $\mathrm{nRel}(C_{2n+1}; 2n^2-1) < 0$ and $\mathrm{nRel}(C_{2n+1}; 2n^2) > 0$. We conclude by the Intermediate Value Theorem that $\mathrm{nRel}(C_{2n+1};p)$ has a root in the interval

$$\left(2n^2-1, 2n^2\right)$$

when n is sufficiently large. $\square$

Remark 3.4.1. It appears that "sufficiently large" in the statement of Theorem 3.4.1 is not very large. The largest real root of $\mathrm{nRel}(C_{2n+1};p)$ (approximated using a computer algebra system) is shown in Table 3.2 for $n \in \{2,\ldots,10\}$ and the root lies in the interval $(2n^2-1, 2n^2)$ for all $n \in \{2,\ldots,10\}$.

In the next chapter we shift the focus of our study to the *connected set polynomial*, which is the generating polynomial of the c-coefficients for node reliability. The connected set polynomial is easier to analyse in some ways, and our work will have direct consequences pertaining to node reliability. In particular, we will learn much more about node reliability roots from our study of the connected set polynomial. Of course, the connected set polynomial is of interest in its own right, as well! It is defined in the same general context as the well-studied independence polynomial and domination polynomial, among others.

Chapter 4

The Connected Set Polynomial

Recall that the node reliability of a graph G on n vertices can be written

$$\sum_{k=1}^{n} c_k p^k (1-p)^{n-k},$$

where $c_k = c_k(G)$ is the number of connected sets of order k in G for each $k \in \{1, \ldots, n\}$. In this chapter we study the related generating polynomial for the collection of connected sets,

$$C(G;x) = \sum_{k=1}^{n} c_k x^k,$$

which we call the *connected set polynomial* of G. Given the node reliability of a graph G, the connected set polynomial of G is easy to find, and vice versa. Explicitly, we have

$$\begin{aligned} \mathrm{nRel}(G;p) &= \sum_{k=1}^{n} c_k p^k (1-p)^{n-k} \\ &= (1-p)^n \cdot \sum_{k=1}^{n} c_k \left(\frac{p}{1-p}\right)^k \\ &= (1-p)^n \cdot C\Big(G; \tfrac{p}{1-p}\Big) \end{aligned}$$

and

$$\begin{aligned} C(G;x) = \sum_{k=1}^{n} c_k x^k &= (1+x)^n \cdot \sum_{k=1}^{n} c_k \left(\frac{x}{1+x}\right)^k \cdot \left(\frac{1}{1+x}\right)^{n-k} \\ &= (1+x)^n \cdot \sum_{k=1}^{n} c_k \left(\frac{x}{1+x}\right)^k \cdot \left(1-\frac{x}{1+x}\right)^{n-k} \\ &= (1+x)^n \cdot \mathrm{nRel}\big(G; \tfrac{x}{1+x}\big). \end{aligned}$$

Hence our study of the connected set polynomial will have direct implications for node reliability.

More generally, whenever we are interested in counting the number of sets of each order in a given collection, it is natural to study the generating polynomial of the collection. The analytic and algebraic properties of the generating polynomial are of interest and can actually give us information about the related counting problem. Many graph polynomials arise from this general setting, including the independence polynomial (see [47] for a survey) and the domination polynomial (introduced in [2]). From this perspective the connected set polynomial is interesting in its own right as well.

We note that the problem of counting the number of connected sets in a graph has been studied in [7, 56, 63, 64, 73, 75, 76], although this work is mostly concerned with the *total* number of connected sets. For example, Björklund et al. recently proved an upper bound on the number of connected sets in graphs with bounded degree [7]. Aside from [7], the focus has been on trees, and in particular on finding the tree in a given class with the most (or least) connected sets. We study the connected set polynomials of trees in Chapter 5.

In Section 4.1, we study the complexity of evaluating the connected set polynomial at a fixed complex number z, demonstrating that the problem is #P-hard except possibly at several specific choices of z. In Section 4.2, we study the *connected set* *roots* (the roots of connected set polynomials). Our main results include a bound on the modulus of the connected set roots of a graph G in terms of the order of G, and a proof that the collection of all connected set roots is dense in the complex plane. A corollary of this density result is that the collection of all node reliability roots is dense in the complex plane, revealing another potential difference between the node reliability and the all-terminal reliability – it is unknown whether all-terminal reliability roots are dense in the complex plane, but it is suspected that they are not.

Before we proceed with the main results of this section, we state the formal definition of the connected set polynomial and present some basic results and examples.

Definition 4.0.1. The *connected set polynomial* of a graph G, denoted $C(G;x)$, is the generating polynomial of the collection of connected sets of G; that is,

$$C(G;x) = \sum_{W \in \mathcal{C}(G)} x^{|W|},$$

where $\mathcal{C}(G)$ is the collection of connected sets of G. Alternatively, we can write

$$C(G;x) = \sum_{k=1}^{n} c_k x^k,$$

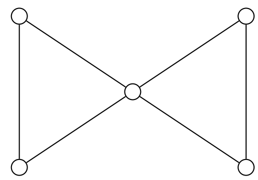
where c_k is the number of connected sets of G of order k for each $k \in \{1, \ldots, n\}$.

Example 4.0.1. The nonisomorphic graphs G_1 and G_2 pictured in Figure 4.1 have the same connected set polynomial;

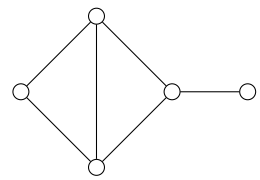
$$C(G_1;x) = C(G_2;x) = x^5 + 4x^4 + 6x^3 + 6x^2 + 5x,$$

which the reader can verify by counting the connected sets of G_1 and G_2 of each order. $\diamondsuit$

Note that the connected set polynomial is well-defined for disconnected graphs –



(a) The graph G_1.



(b) The graph G_2.

Figure 4.1: A pair of graphs whose connected set polynomials are equal.

the connected set polynomial of a disconnected graph is given simply by the sum of the connected set polynomials of its components.

Observation 4.0.1. *Let G be a graph with connected components $G_1, G_2, \ldots, G_k$. The connected set polynomial of G is given by*

$$C(G;x) = \sum_{i=1}^{k} C(G_i;x).$$

In particular, for disjoint graphs G and H,

$$C(G \cup H;x) = C(G;x) + C(H;x).$$

Proof. Every connected set of G is contained entirely in exactly one connected component of G. □

For any (possibly even disconnected) graph G of order n and size m, it is clear that $c_1 = n$ and $c_2 = m$. In addition, G is connected if and only if $c_n = 1$. More generally, G is k-connected if and only if

$$c_{n-i} = \binom{n}{n-i}$$

for all $i \in \{1, \ldots, k-1\}$. The degree of $C(G;x)$ is the order of a largest component in G, and $C(G;1) = \sum_{k=1}^{n} c_k$ is the total number of connected sets of G. Thus the connected set polynomial holds a lot of information about the connectedness of a graph.

4.1 Complexity

Many straightforward algorithms for computing the connected set polynomial of a graph G exist. The most direct approach requires running through each subset $X \subseteq V(G)$ and determining whether the induced subgraph $G[X]$ is connected. This algorithm is clearly exponential in the order of the graph, and although we can improve on this algorithm, finding a polynomial algorithm is likely to be hard.

It was proven in [62] that computing the number of connected sets of a graph is #P-complete, even for split graphs, and this is equivalent to computing $C(G;1)$. This tells us immediately that the problem of computing the connected set polynomial is NP-hard, but this answer doesn't quite tell the whole story about the complexity of evaluating the connected set polynomial at specific values. We can easily evaluate the connected set polynomial at 0 (we have $C(G;0) = 0$ for any graph G), and one cannot help but wonder if there are any other points in the complex plane at which the connected set polynomial is easy to evaluate.

For this reason, we would like a result that tells us more about the complexity of evaluating $C(G;z)$ at various fixed values of z. Vertigan and Welsh demonstrated such a result for the well-known Tutte polynomial in [67]. They found that the exact evaluation of the Tutte polynomial is #P-hard at all but a few special points and two special hyperbolae, even for the class of planar bipartite graphs. We demonstrate here that evaluating the connected set polynomial exactly is #P-hard at all but a countable set of points in the complex plane (only three of which are real numbers). Our proof will require the use of the *lexicographic product* of graphs.

Definition 4.1.1. The *lexicographic product* or *graph composition* $G \odot H$ of graphs

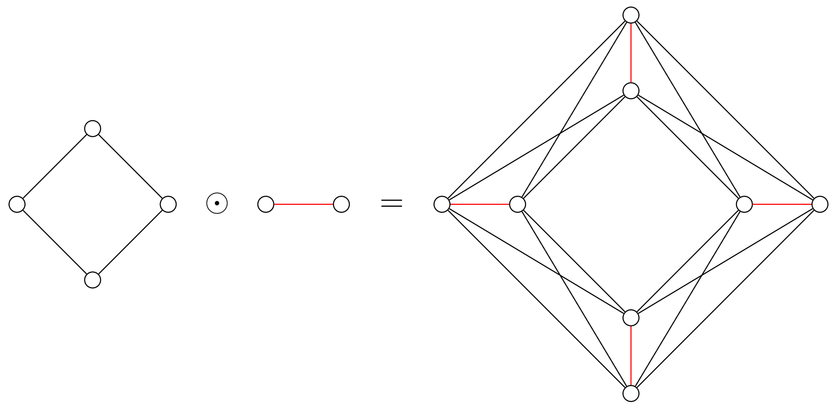


Figure 4.2: The lexicographic product $C_4 \odot K_2$. The red edges denote the copies of K_2 that are replacing vertices of C_4.

G and H is the graph on vertex set $V(G) \times V(H)$ such that vertices (u, x) and (v, y) are adjacent if and only if either

- u is adjacent to v in G; or
- $u = v$ and x is adjacent to y in H.

Intuitively, the lexicographic product $G \odot H$ is the operation of replacing every vertex of G with a copy of H. Figure 4.2 shows an example of a lexicographic product graph, namely $C_4 \odot K_2$.

The next lemma gives a formula for the connected set polynomial $C(G \odot H; x)$ in terms of $C(G; x)$ and $C(H; x)$. The special case of this formula for $H = K_n$ will be important to several proofs in this chapter.

Lemma 4.1.1. *Let G be a graph of order n_G and let H be a graph of order n_H. The connected set polynomial of the lexicographic product $G \odot H$ is given by*

$$C(G \odot H; x) = C(G; (x+1)^{n_H} - 1) + n_G[C(H; x) - (x+1)^{n_H} + 1].$$

In particular,

$$C(G \odot K_n; x) = C(G; (x+1)^n - 1).$$

Proof. Let $C_\odot$ be a subset of $V(G \odot H)$. Define

$$C_G = \{v \in V(G) \mid (v,x) \in C_\odot \text{ for some } x \in V(H)\}$$

and

$$C_v = \{x \mid (v,x) \in C_\odot\} \text{ for each } v \in C_G.$$

We see that $C_\odot$ is a connected set if and only if either

(i) C_G is a connected set of G of order at least two (in this case, for each $v \in C_G$, the set C_v can be any nonempty subset of vertices of H); or

(ii) $C_G = \{v\}$ for some v and C_v is a connected set of H.

The connected sets of G of order at least two are enumerated by

$$C(G;x) - n_G x,$$

and hence the connected sets of $G \odot H$ corresponding to case (i) are enumerated by

$$C(G;(x+1)^{n_H} - 1) - n_G[(1+x)^{n_H} - 1].$$

Meanwhile, the connected sets of $G \odot H$ corresponding to case (ii) are enumerated by

$$n_G C(H;x).$$

We conclude that

$$C(G \odot H;x) = C(G;(x+1)^{n_H} - 1) + n_G[C(H;x) - (x+1)^{n_H} + 1]. \qquad \square$$

We are now ready to prove that for a fixed $z \in \mathbb{C}$, evaluating the connected set polynomial at z is #P-hard in all but possibly a few specific cases.

Theorem 4.1.2. *Fix a complex number $z \neq -1$ such that $z+1$ is not a root of unity. Then the problem:*

INPUT: graph G

OUTPUT: the evaluation $C(G; z)$

is #P-hard. In particular, the problem is #P-hard for any real number r except 0, *and possibly* -1 *and* -2.

Proof. Suppose that we had a polynomial time algorithm (polynomial in terms of the order n of the graph) for evaluating the connected set polynomial at some fixed complex number $z \neq -1$ such that $z+1$ is not a root of unity. Then for any graph G on n vertices and each $k \in \{1, \ldots, n+1\}$ we could compute $C(G \odot K_k; z)$ – since $G \odot K_k$ has kn vertices and $k \leq n+1$, the time required to compute $C(G \odot K_k; z)$ for all $k \in \{1, \ldots, n+1\}$ is still polynomial in n. By Lemma 4.1.1,

$$C(G \odot K_k; z) = C(G; (z+1)^k - 1).$$

This means that our evaluations of $C(G \odot K_k; z)$ give a system of $n+1$ linear equations in the coefficients of $C(G; x)$, namely

$$C(G \odot K_k; z) = c_1\big[(z+1)^k - 1\big] + c_2\big[(z+1)^k - 1\big]^2 + \ldots + c_n\big[(z+1)^k - 1\big]^n$$

for $k \in \{1, \ldots, n+1\}$. We can solve for the coefficients in polynomial time using Gauss-Jordan elimination, provided that there is a unique solution. The system of equations has a unique solution if and only if its corresponding matrix is invertible.

The matrix for the system of equations can be written

$$\begin{bmatrix} 1 & z & z^2 & \dots & z^n \\ 1 & (z+1)^2-1 & ((z+1)^2-1)^2 & \dots & ((z+1)^2-1)^n \\ 1 & (z+1)^3-1 & ((z+1)^3-1)^2 & \dots & ((z+1)^3-1)^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & (z+1)^{n+1}-1 & ((z+1)^{n+1}-1)^2 & \dots & ((z+1)^{n+1}-1)^n \end{bmatrix}$$

This is a *Vandermonde matrix* (see [40], for example) for the function $C(G;x)$ at the $n+1$ points

$$\begin{aligned} \alpha_1 &= (z+1)^1 - 1 = z \\ \alpha_2 &= (z+1)^2 - 1 \\ &\vdots \\ \alpha_{n+1} &= (z+1)^{n+1} - 1 \end{aligned}$$

The determinant of a Vandermonde matrix is known to be nonzero if and only if all of the α_i are distinct (see [40]). For our matrix, this condition is satisfied by the assumption that $z \neq -1$ and $z+1$ is not a root of unity. Therefore, the matrix is invertible and we can use Gauss-Jordan elimination to solve for the coefficients of $C(G;x)$ in polynomial time. In particular, we can find $C(G;1)$ in polynomial time; a contradiction as $C(G;1)$ is #P-complete [62]. We conclude that evaluating the connected set polynomial at a fixed complex number $z \neq 1$ such that $z+1$ is not a root of unity is #P-hard. $\square$

By Theorem 4.1.2, the only real numbers at which we are unsure of the complexity of evaluating the connected set polynomial are $r = -2$ and $r = -1$, as we know that

$C(G;0) = 0$ for all graphs G.

Recall that

$$\mathrm{nRel}(G;p) = (1-p)^n \cdot C\Big(G; \tfrac{p}{1-p}\Big),$$

so we can say something about the complexity of computing $\mathrm{nRel}(G;p)$ for fixed p. In particular, we have the following result.

Corollary 4.1.3. *Fix a real number $p \notin \{0,1,2\}$. Then the problem:*

INPUT: graph G

OUTPUT: the evaluation $\mathrm{nRel}(G;p)$

is #P-hard.

Proof. Note that since $p \neq 1$ the number $\frac{p}{1-p}$ is well-defined and real. Further, note that

$$\tfrac{p}{1-p} = 0 \iff p = 0$$

and

$$\tfrac{p}{1-p} = -2 \iff p = 2,$$

while $\frac{p}{1-p} = -1$ has no solutions. Hence by the assumption that $p \notin \{0,1,2\}$, we have $\frac{p}{1-p} \notin \{0,-1,-2\}$. The result now follows immediately by the fact that

$$\mathrm{nRel}(G;p) = (1-p)^n \cdot C\Big(G; \tfrac{p}{1-p}\Big)$$

and the problem of evaluating $C(G; \frac{p}{1-p})$ is #P-hard by Theorem 4.1.2. □

In particular, Corollary 4.1.3 tells us that evaluating the node reliability polynomial is #P-hard for any fixed $p \in (0,1)$, the interval of primary interest for any measure of reliability. Of course, node reliability is easy to evaluate at $p = 0$ (it is 0 for all graphs) and at $p = 1$ (it is 1 for connected graphs and 0 for disconnected graphs), so $p = 2$ is the only real number at which the complexity of evaluating the node reliability is unknown; this is an interesting open question.

Since evaluating the node reliability exactly is #P-hard for any fixed $p \in (0,1)$, it would be useful to have a method for bounding $\mathrm{nRel}(G;p)$ on this interval. We demonstrate such a method here, considering connected graphs first. Recall that the node reliability of a connected graph G on n vertices is given by

$$\mathrm{nRel}(G;p) = \sum_{k=1}^{n} c_k p^k (1-p)^k,$$

where c_k is the number of connected sets of G of order k. While the computation of $\mathrm{nRel}(G;p)$ is #P-hard for any $p \in (0,1)$, we can compute some of the c_k coefficients in polynomial time. We can compute any finite number of coefficients at the bottom ($c_1, c_2, \ldots, c_s$ for some fixed s) in polynomial time by simply running through all possible subsets of appropriate order and checking whether they induce connected graphs (there are only $\binom{n}{k}$ vertex subsets of order k to run through, which is polynomial in n as $k \leq s$). Similarly we can find any finite number of coefficients at the top ($c_t, \ldots, c_n$ for some fixed value of $n - t$) in polynomial time.

Similar observations have been made about all-terminal reliability (see Chapter 5 of [28], for example). Briefly, several coefficients at either end of the all-terminal reliability polynomial can be computed in polynomial time, and then Sperner's bounds [59] can be used to find upper and lower bounds on the remaining coefficients in terms of the known coefficients. While many methods have been used to bound the all-terminal reliability on $(0,1)$, this method is straightforward and most other methods are simply improvements of this basic idea.

While Sperner's Bounds do not hold for the c-coefficients of node reliability (which count the number of connected sets of each order), we do have the bounds of Lemma 3.1.2:

$$(k-t+1)c_k \leq \binom{n-t}{k-t} c_t$$

for all $k \in \{2, \ldots, n\}$ and $t \in \{1, \ldots, k-1\}$. In particular, if we know the exact values of the coefficients $c_1, \ldots, c_s$, and $c_t, \ldots, c_n$, then we have

$$c_k \leq \binom{n-s}{k-s} \frac{c_s}{k-s+1}$$

and

$$c_k \geq \frac{(t-k+1)c_t}{\binom{n-k}{t-k}}$$

for $k \in \{s+1, \ldots, t-1\}$. In addition to the bounds of Lemma 3.1.2, we have another similar set of bounds, presented in the lemma below.

Lemma 4.1.4. *For any connected graph G of order n,*

$$kc_k \geq c_{k-1}$$

for all $k \in \{2, \ldots, n\}$. More generally,

$$\tbinom{k}{t} c_k \geq c_t$$

for all $k \in \{2, \ldots, n\}$ and $t \in \{1, \ldots, k-1\}$.

Proof. Let G be a connected graph of order n, let $k \in \{2, \ldots, n\}$ and let $t \in \{1, \ldots, k-1\}$. Every connected set of order t is contained in some connected set of order k (since G is connected, we can always add a vertex to a connected set of order less than n that will preserve connectedness). Therefore, the collection of all subsets of order t of all connected sets of order k will contain all connected sets of order t. We have c_k connected sets of order k, and $\binom{k}{t}$ subsets of order t for each of these c_k connected sets, giving at most $\binom{k}{t} c_k$ connected sets of order t. □

In particular, if we know the values of $c_1, \ldots, c_s$ and $c_t, \ldots, c_n$, then by Lemma

4.1.4 we have

$$c_k \geq \frac{c_s}{\binom{k}{s}}$$

and

$$c_k \leq \tbinom{t}{k} c_t$$

for all $k \in \{s+1, \ldots, t-1\}$. Putting our two bounds together gives the following lemma.

Lemma 4.1.5. *Let G be a connected graph on n vertices, and suppose that we know $c_1, \ldots, c_s$ and $c_t, \ldots, c_n$ with $s < t$. For $k \in \{s+1, \ldots, t-1\}$, we have*

$$\underset{\sim}{c}_k \leq c_k \leq \tilde{c}_k,$$

where

$$\underset{\sim}{c}_k = \max\left\{ \frac{(t-k+1)c_t}{\binom{n-k}{t-k}}, \frac{c_s}{\binom{k}{s}} \right\}$$

and

$$\tilde{c}_k = \min\left\{ \tbinom{n-s}{k-s} \tfrac{c_s}{k-s+1}, \tbinom{t}{k} c_t \right\}. \qquad \square$$

From here it is easy to bound the node reliability on the interval $(0,1)$. Suppose that we know $c_1, \ldots, c_s$ and $c_t, \ldots, c_n$. The product

$$p^k (1-p)^{n-k}$$

is positive on $(0,1)$ for all $k \in \{1, \ldots, n\}$, meaning that

$$\sum_{k=1}^{n} \underset{\sim}{c}_k p^k (1-p)^{n-k} \leq \mathrm{nRel}(G;p) \leq \sum_{k=1}^{n} \tilde{c}_k p^k (1-p)^{n-k},$$

where $\underset{\sim}{c}_k$ and $\tilde{c}_k$ are defined as in Lemma 4.1.5 for $k \in \{s+1, \ldots, t-1\}$ and $\underset{\sim}{c}_k = \tilde{c}_k = c_k$

otherwise. The upper bound $\sum_{k=1}^{n} \tilde{c}_k p^k (1-p)^{n-k}$ can be larger than 1 in practice, but obviously the node reliability is always less than 1 on the interval $(0,1)$ as it represents a probability. Therefore, we have the following theorem.

Theorem 4.1.6. *Let G be a connected graph of order n. Suppose that we know the coefficients $c_1, \ldots, c_s$ and $c_t, \ldots, c_n$. Then*

$$\sum_{k=1}^{n} \underset{\sim}{c}_k p^k (1-p)^{n-k} \leq \mathrm{nRel}(G;p) \leq \max\left\{1, \sum_{k=1}^{n} \tilde{c}_k p^k (1-p)^{n-k}\right\},$$

where

$$\underset{\sim}{c}_k = \begin{cases} \max\left\{ \frac{(t-k+1)c_t}{\binom{n-k}{t-k}}, \frac{c_s}{\binom{k}{s}} \right\} & \text{if } k \in \{s+1, \ldots, t-1\}, \text{ and} \\ c_k & \text{otherwise;} \end{cases}$$

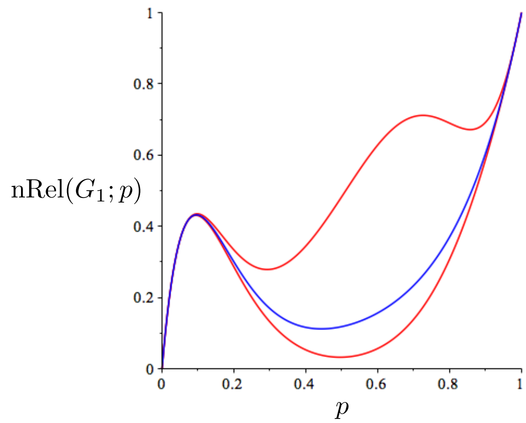
and

$$\tilde{c}_k = \begin{cases} \min\left\{ \binom{n-s}{k-s} \frac{c_s}{k-s+1}, \binom{t}{k} c_t \right\} & \text{if } k \in \{s+1, \ldots, t-1\}, \text{ and} \\ c_k & \text{otherwise.} \end{cases}$$

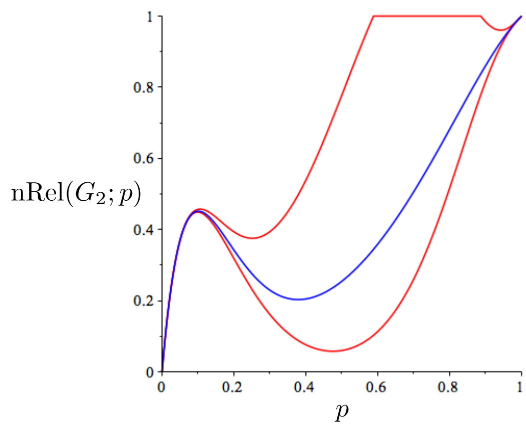
□

To demonstrate the use of these bounds on node reliability, we randomly generated some graphs on 12 vertices, including each possible edge independently at random with various fixed probabilities, and discarding any disconnected graphs that resulted from the process. A plot of the node reliability of the resulting graphs, along with the lower and upper bounds (using $s = 3$ and $t = 10$) are shown in Figure 4.3. We have found that the bounds often appear to be tighter for graphs with fewer edges.

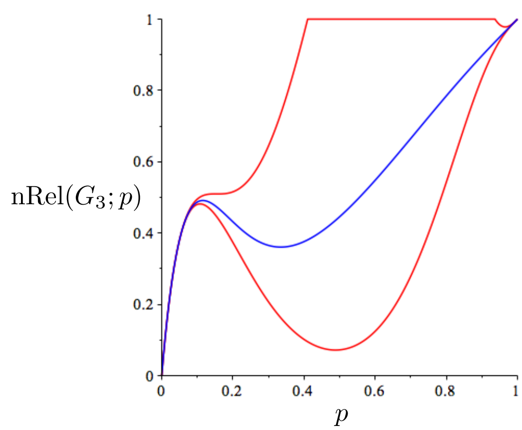
Finally, we mention that the bounds of Lemma 4.1.5 on the c-coefficients of connected graphs also allow us to bound the node reliability of disconnected graphs. First, we can find the connected components of a disconnected graph G in polynomial time. Then we can bound the c-coefficients of each connected component of G separately using Lemma 4.1.5, and sum the lower and upper bounds over all components to get bounds on the c-coefficients of G. The bounds on the c-coefficients of G



(a) Node reliability of a graph G_1 on 12 vertices and 13 edges



(b) Node reliability of a graph G_2 on 12 vertices and 17 edges



(c) Node reliability of a graph G_3 on 12 vertices and 23 edges

Figure 4.3: Bounds (red) on the node reliabilities (blue) of various randomly produced graphs.

transfer directly to bounds on $\mathrm{nRel}(G;p)$.

4.2 Roots of the Connected Set Polynomial

As we have stated several times throughout this thesis, the roots of many graph polynomials including the matching polynomial, the independence polynomial, the domination polynomial, and the chromatic polynomial have been studied extensively. This means that it is natural to ask questions about the roots of the connected set polynomial. For any graph G, we call a root of the connected set polynomial of G a *connected set root* of G. Most of our results have direct implications for the roots of node reliability, and we present these as corollaries when they are of interest.

4.2.1 Realness and Connected Set Roots

In this section we focus on the nature of connected set roots, primarily with regard to their realness. A wide variety of results concerning the realness of the roots of graph polynomials exists in the literature. The question of whether a polynomial has all real roots is often of primary interest, as having all real roots implies that the sequence of coefficients of the polynomial is log-concave (and hence unimodal if the coefficients are positive) by a result due to Newton (see [60]). For example, it was shown in [38] that all roots of the matching (generating) polynomial are real. A generalization of this result was recently proven by Chudnovsky and Seymour in [26], namely that the independence polynomial of any claw-free graph has all real roots (the matching polynomial of G is an evaluation of the independence polynomial of the line graph of G, and line graphs are claw-free).

Here we settle the problem of whether the connected set roots of a connected graph G are all real; we demonstrate that no connected graph on 3 or more vertices has all real connected set roots! Since $C(K_1;x) = x$ and $C(K_2;x) = 2x + x^2$, this implies that the graphs K_1 and K_2 are the only connected graphs whose connected set roots are all real. This is in stark contrast to the situation for the matching polynomial and the even the independence polynomial, as the relatively large family of all claw-free graphs have all real independence roots.

In order to prove that every graph of order at least 3 has a nonreal connected set root, we will need the following straightforward result, demonstrated on page 265 of [66]. Since all nonzero real roots of the connected set polynomial must be negative, we will actually apply the corollary that follows immediately.

Theorem 4.2.1. *Let* $f(z) = \sum_{k=0}^{n} a_k z^k$ *with* $a_n \neq 0$. *If all zeros of* f *are real and positive then*

$$\frac{a_{n-1}}{a_n} \cdot \frac{a_1}{a_0} \geq n^2. \qquad \square$$

Corollary 4.2.2. *Let* $f(z) = \sum_{k=0}^{n} a_k z^k$ *with* $a_n \neq 0$. *If all zeros of* f *are real and negative then*

$$\frac{a_{n-1}}{a_n} \cdot \frac{a_1}{a_0} \geq n^2.$$

Proof. Suppose all zeros of $f(z)$ are real and negative. Then all zeros of $f(-z)$ are real and positive. Thus by Theorem 4.2.1, we must have

$$\frac{(-1)^{n-1}a_{n-1}}{(-1)^n a_n} \cdot \frac{-a_1}{a_0} \geq n^2.$$

Cancelling signs gives the desired result. $\square$

Now we are ready to prove that every connected graph of order at least 3 has a nonreal connected set root.

Theorem 4.2.3. *Let* G *be a connected graph of order* n. *If* $n \geq 3$ *then* $C(G;x)$ *has a nonreal root.*

Proof. Let $C(G;x) = \sum_{k=1}^{n} c_k x^k = x\sum_{k=1}^{n} c_k x^{k-1}$. Suppose – to reach a contradiction – that $C(G;x)$ has all real roots. Since $c_k > 0$ for all $k \in \{1,\ldots,n\}$, all roots of

$C(G;x)/x = \sum_{k=1}^{n} c_k x^{k-1}$ must be strictly negative. Thus by Corollary 4.2.2, we must have

$$\frac{c_{n-1}}{c_n} \cdot \frac{c_2}{c_1} \geq (n-1)^2.$$

However, from Observation 3.1.1,

$$\frac{c_{n-1}}{c_n} \cdot \frac{c_2}{c_1} = (n-t)\frac{m}{n},$$

where m is the number of edges and t is the number of cut vertices. Since $m \leq \binom{n}{2}$ and $n \geq 3$,

$$(n-t)\frac{m}{n} \leq (n-t)\frac{(n-1)}{2} \leq n\frac{(n-1)}{2} = \frac{n}{2}(n-1) < (n-1)^2,$$

a contradiction. Therefore, $C(G;x)/x$ (and hence $C(G;x)$ itself) must have a nonreal root. $\square$

Theorem 4.2.3 implies almost immediately that any connected graph of order at least 3 has a nonreal node reliability root. The details are below.

Corollary 4.2.4. *Let G be a connected graph of order n and size m. If $n \geq 3$ then* $\mathrm{nRel}(G;p)$ *has a nonreal root.*

Proof. By Theorem 4.2.3, the connected set polynomial $C(G;x)$ has some nonreal root ζ. Let $p_\zeta = \frac{\zeta}{1+\zeta}$, which is well-defined as $\zeta \neq -1$. The reader can verify that p_ζ

is real if and only if ζ is real, so p_ζ is nonreal. We evaluate

$$\begin{aligned} \mathrm{nRel}(G;p_\zeta) &= (1-p_\zeta)^n C(G;\tfrac{p_\zeta}{1-p_\zeta}) \\ &= (1-p_\zeta)^n C\left(G;\frac{\frac{\zeta}{1+\zeta}}{1-\frac{\zeta}{1+\zeta}}\right) \\ &= (1-p_\zeta)^n C(G;\zeta) \\ &= 0, \end{aligned}$$

which shows that p_ζ is a nonreal root of $\mathrm{nRel}(G;p)$. $\square$

Since nonreal roots of polynomials with real coefficients come in conjugate pairs, every connected graph on 3 or more vertices has at least *two* nonreal connected set roots. Using a computer algebra system we have found that there is a connected graph on n vertices with *exactly* two nonreal connected set roots for each $n \in \{3,4,5,6\}$. However, there is no connected graph on 7 vertices with exactly two nonreal connected set roots.

It is obvious that every graph has a real connected set root, namely 0 (corresponding to the node reliability root at 0). We can prove that any connected graph G of even order n must also have a negative real connected set root. Close to 0, the dominant term of $C(G;x)$ is nx, so that $C(G;x) < 0$ on some interval $(-\varepsilon, 0)$. Further, $\lim_{n\to-\infty} C(G;x) = \infty$ since $C(G;x)$ has even degree n, so it follows by the Intermediate Value Theorem that $C(G;r) = 0$ for some real number $r < 0$.

On the other hand, there are many connected graphs of odd order for which 0 is the only real connected set root. The smallest such examples are P_3 and K_3 (their connected set polynomials have degree 3 and by Theorem 4.2.3 they each have a pair of nonreal connected set roots). It is easy to see that

$$C(K_{2n+1};x) = (x+1)^{2n+1} - 1$$

has no nonzero real roots for all $n \geq 1$ – the roots of $C(K_{2n+1}; x)$ are the $(2n+1)^{\text{st}}$ roots of unity shifted to the left 1.

We have also found that there are purely imaginary connected set roots. This contrasts sharply with the chromatic polynomial, where it is widely suspected that there are *no* purely imaginary roots (see [9]). The graph $K_{1,4}$ has

$$\begin{aligned} C(K_{1,4}; x) &= x^5 + 4x^4 + 6x^3 + 4x^2 + 5x \\ &= x(x^2 + 4x + 5)(x^2 + 1), \end{aligned}$$

so that $\pm i$ are connected set roots of $K_{1,4}$. In fact, we can find an infinite number of purely imaginary connected set roots by considering the graph $G_t = P_4 \cup tK_2 \cup 2K_1$ for each $t \geq 0$; that is, the disjoint union of P_4, t copies of K_2, and 2 copies of K_1. We find

$$\begin{aligned} C(P_4 \cup tK_2 \cup 2K_1; x) &= C(P_4; x) + tC(K_2; x) + 2C(K_1; x) \\ &= (x^4 + 2x^3 + 3x^2 + 4x) + t(x^2 + 2x) + 2x \\ &= x^4 + 2x^3 + (3 + t)x^2 + (6 + 2t)x \\ &= x(x + 2)(x^2 + t + 3), \end{aligned}$$

and hence for all $t \geq 0$ the graph G_t has connected set roots at $\pm i\sqrt{t+3}$.

The fact that the graph G_t has connected set roots at $\pm i\sqrt{t+3}$ also demonstrates that the collection of all connected set roots is unbounded in modulus. We will see more examples of families with connected set roots tending to infinity in the next section.

Before moving on to the next section we discuss the realness of connected set roots of disconnected graphs in general. Theorem 4.2.3 implies that the connected set polynomial of a *connected* graph has all real roots if and only if it has degree at most 2. However, we have found *disconnected* graphs whose connected set polynomials have degree 3 and still have all real roots. Consider the graph $K_3 \cup kK_2$, i.e. the disjoint union of K_3 and k copies of K_2. We have

$$\begin{aligned}
C(K_3 \cup kK_2; x) &= C(K_3; x) + kC(K_2; x) \\
&= x^3 + 3x^2 + 3x + k(x^2 + 2x) \\
&= x^3 + (3+k)x^2 + (3+2k)x \\
&= x(x^2 + (3+k)x + (3+2k))
\end{aligned}$$

The discriminant of the quadratic factor in the above expression factors to

$$(k-3)(k+1).$$

Therefore, $C(K_3 \cup kK_2; x)$ has all real roots for $k \geq 3$. One can show similarly that $C(P_3 \cup kK_2; x)$ has all real roots for $k \geq 6$. The question of whether or not there are disconnected graphs whose connected set polynomials have degree greater than 3 (or possibly even arbitrarily high degree) and still have all real roots remains open.

4.2.2 Bounding the Connected Set Roots

While there is no constant bound on the modulus of an arbitrary connected set root, the main result of this section is that the modulus of any connected set root of a graph G is bounded above by the order of the graph G. The proof is similar to the proof of Theorem 2.1.5, which gives a bound on the modulus of any all-terminal reliability root of the graph G in terms of the order of G. Similar results exist for many graph polynomials including the independence polynomial [21] and the chromatic polynomial [57]. A tool commonly used in the proof of such results is the well-known Eneström-Kakeya Theorem, which we restate here for ease of reading.

We also include a characterization from [1] of the polynomials for which the bounds of the Eneström-Kakeya Theorem are tight.

Theorem 4.2.5 (Eneström-Kakeya (c.f. [1]))**.** *If $f(x) = a_0 + a_1x + ... + a_nx^n$ has positive real coefficients, then all complex roots of f lie in the annulus*

$$r \leq |z| \leq R,$$

where

$$r = \min\Big\{\tfrac{a_i}{a_{i+1}} : 0 \leq i \leq n-1\Big\} \text{ and } \quad R = \max\Big\{\tfrac{a_i}{a_{i+1}} : 0 \leq i \leq n-1\Big\}.$$

Further, f has a root of modulus r if and only if

$$\gcd\{j \in \{1, \ldots, n+1\} \colon \; a_{j-1} > ra_j\} > 1,$$

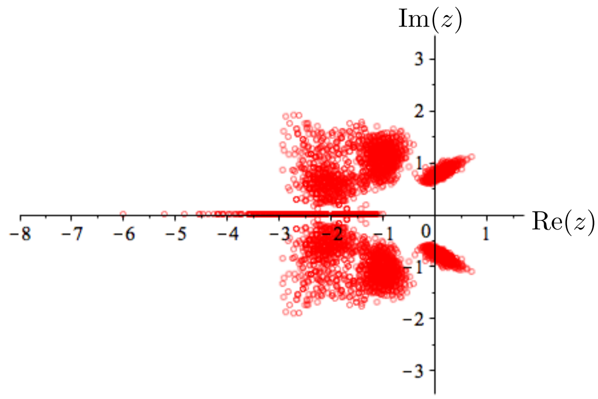
where $a_{n+1} = 0$, and f has a root of modulus R if and only if

$$\gcd\{j \in \{1, \ldots, n+1\} \colon \; a_{d-j} < Ra_{d+1-j}\} > 1,$$

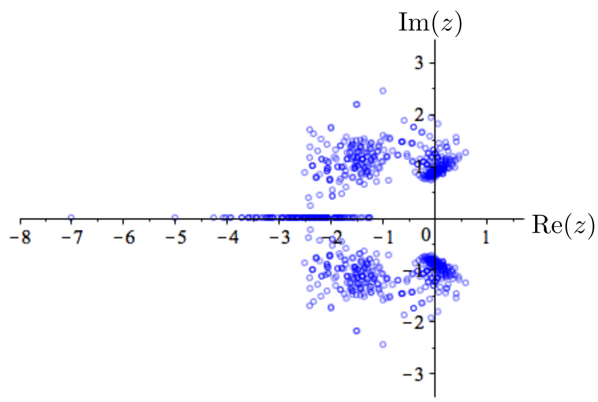
where $a_{-1} = 0$. □

The nonzero connected set roots of all graphs on 7 vertices are shown in Figure 4.4. It is evident that there are no connected set roots close to 0 and also that there are no connected set roots of modulus greater than 7 – it does appear that there is a connected graph with connected set root at or near -6 and a disconnected graph with a connected set root at or near -7. We will confirm all of these observations in this section.

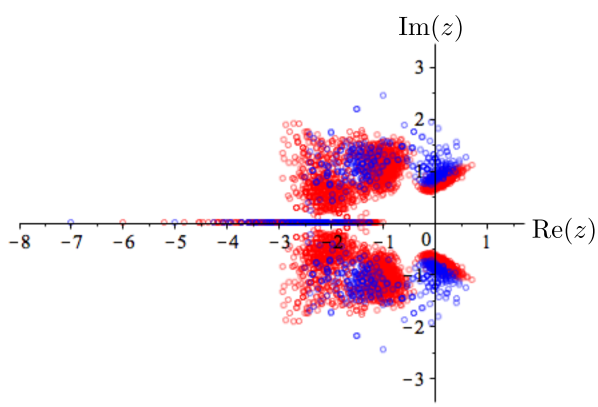
Theorem 4.2.6. *Let G be a connected graph on $n \geq 2$ vertices. If z is a nonzero*



(a) Connected set roots of all connected graphs on 7 vertices.



(b) Connected set roots of all disconnected graphs on 7 vertices.



(c) Connected set roots of all graphs on 7 vertices.

Figure 4.4: The nonzero connected set roots of all graphs on 7 vertices. The roots are partially transparent so that the viewer gets an impression of the density of the roots.

root of $C(G;x)$*, then*

$$\tfrac{2}{n-1} \leq |z| \leq n.$$

Proof. Let $C(G;x) = \displaystyle\sum_{k=1}^{n} c_k x^k$. By Lemma 3.1.2, we have

$$2c_k \leq (n-k+1)c_{k-1} \Longrightarrow \frac{c_{k-1}}{c_k} \geq \frac{2}{n-k+1} \geq \frac{2}{n-1}$$

for all $k \in \{2, \ldots, n\}$. Moreover, by Lemma 4.1.4 we have

$$kc_k \geq c_{k-1} \quad \Longrightarrow \quad \frac{c_{k-1}}{c_k} \leq k \leq n$$

for all $k \in \{2, \ldots, n\}$. Therefore, by the Eneström-Kakeya Theorem (Theorem 4.2.5) applied to $C(G;x)/x$, if z is a nonzero root of $C(G;x)$ then $\tfrac{2}{n-1} \leq |z| \leq n$. □

We can extend Theorem 4.2.6 to all graphs (including disconnected graphs) as follows.

Corollary 4.2.7. *Let* G *be a nonempty, though possibly disconnected graph on* n *vertices. Let* d *be the order of a largest component in* G*, and for each* $k \in \{1, \ldots, d-1\}$ *let* α_k *be the number of components of order* k *in* G*. Let* $\alpha = \max\{\alpha_1, \ldots, \alpha_{d-1}\}$*. If* z *is a nonzero root of* $C(G;x)$*, then*

$$\tfrac{2}{d-1} \leq |z| \leq d + \alpha \leq n.$$

Proof. Note that $d \geq 2$ as G is nonempty and thus contains at least one edge. We first prove the upper bound. For each $k \in \{2, \ldots, d\}$, let β_k be the number of components of G of order *at least* k. Without loss of generality let $G_1, \ldots, G_{\beta_k}$ be the components of G of order at least k. By Lemma 4.0.1 and the bounds of Lemma 4.1.4 applied to the components $G_1, \ldots, G_{\beta_k}$, we have

$$\begin{aligned} kc_k(G) &= k[c_k(G_1) + \ldots + c_k(G_{\beta_k})] \\ &\geq c_{k-1}(G_1) + \ldots c_{k-1}(G_{\beta_k}) \\ &= c_{k-1}(G_1 \cup \ldots \cup G_{\beta_k}) \end{aligned}$$

There are at most α connected components of G of order $k-1$, and every component of order $k-1$ has exactly one connected set of order $k-1$. Therefore,

$$c_{k-1}(G) \leq c_{k-1}(G_1 \cup \ldots \cup G_{\beta_k}) + \alpha \leq kc_k(G) + \alpha.$$

Thus we have

$$\frac{c_{k-1}(G)}{c_k(G)} \leq k + \frac{\alpha}{c_k(G)} \leq k + \alpha \leq d + \alpha.$$

Therefore, by the Eneström-Kakeya Theorem, any root z of $C(G;x)$ satisfies $|z| \leq d+\alpha$. The inequality $d+\alpha \leq n$ follows almost immediately from the definitions of d and α. We know that $\alpha = \alpha_j$ for some particular $j \in \{1, \ldots, d-1\}$, so G has α components of order j. Since G also has at least one component of order d, we have

$$n \geq d + j\alpha \geq d + \alpha.$$

For the lower bound, by Lemma 3.1.2 we have

$$\frac{c_{k-1}(G)}{c_k(G)} \geq \frac{2}{d-k+1} \geq \frac{2}{d-1},$$

and we conclude by the Eneström-Kakeya Theorem that any nonzero root z of $C(G;x)$ satisfies $|z| \geq \frac{2}{d-1}$. □

We have shown that any connected set root z of an arbitrary graph G of order n

satisfies $|z| \leq n$, and an obvious question is whether or not this bound is tight. We prove that the bound is tight for disconnected graphs by finding a disconnected graph of order n with a connected set root of modulus n for each $n \geq 2$. In fact, we prove that this is the *only* graph on n vertices with a connected set root of modulus n. Further, for n sufficiently large we find a connected graph on n vertices with a connected set root of modulus close to $n-1$. Thus, while the upper bound of Theorem 4.2.7 may not quite be tight for connected graphs, it is within 1.

Proposition 4.2.8. *The unique graph (up to isomorphism) on n vertices with a connected set root of modulus n is the graph $K_2 \cup O_{n-2}$; that is, the graph on n vertices with exactly one edge.*

Proof. By inspection,

$$C(K_2 \cup O_{n-2}; x) = x^2 + nx = x(x+n),$$

and hence $-n$ is a connected set root of $K_2 \cup O_{n-2}$ for all $n \geq 2$.

In order to see that $K_2 \cup O_{n-2}$ is the unique graph on n vertices with a connected set root of modulus n, first note that the empty graph O_n has only a single root at 0. Now we may assume that G is a graph on n vertices and $m \geq 2$ edges. Let

$$C(G; x) = c_1 x + c_2 x^2 + c_3 x^3 + \ldots + c_d x^d,$$

where d is the order of a largest component of G. By Theorem 4.2.5, the polynomial

$$C(G; x)/x = c_1 + c_2 x + \ldots + c_d x^{d-1}$$

has some of its roots on the boundary

$$|z| = R = \max\left\{\tfrac{c_i}{c_{i+1}} : \ 1 \leq i \leq d-1\right\}$$

only if the greatest common denominator of the set

$$S = \{i \in \{1, 2, \ldots, d\}\colon\ c_{d-i} < Rc_{d+1-i}\}$$

is greater than 1, where $c_0 = 0$. We way assume that $R = n$ as otherwise any root z of $C(G; x)/x$ satisfies $|z| \leq R < n$ and we are done. Using the facts that $c_1 = n$ and $c_2 = m \geq 2$, we have

$$c_0 = 0 < n \cdot n = nc_1$$

so that $d \in S$ and

$$c_1 = n < n \cdot 2 \leq n \cdot m = nc_2$$

so that $d - 1 \in S$. Since the consecutive integers $d - 1$ and d are both in S, we must have $\gcd(S) > 1$. We conclude by Theorem 4.2.5 that $C(G; x)/x$ has no root of modulus n. Therefore, $K_2 \cup O_{n-2}$ is the only graph on n vertices with a connected set root on the boundary $|z| = n$. □

Proposition 4.2.9. *Let $\varepsilon \in (0, 1)$. For n sufficiently large, there is a connected graph of order n with a connected set root of modulus within ε of $n - 1$.*

Proof. Consider the cycle C_n for some $n \geq 3$. From the discussion of the connected sets of the cycle in Theorem 3.4.1,

$$C(C_n; x) = x^n + n \cdot \sum_{k=1}^{n-1} x^k,$$

and hence

$$C(C_n; x) = x^n + n \cdot \frac{x^n - x}{x - 1}.$$

Substituting $x = 1 - n$, we obtain

$$\begin{aligned} C(C_n; 1-n) &= (1-n)^n + n \cdot \frac{(1-n)^n - (1-n)}{-n} \\ &= (1-n)^n - (1-n)^n + (1-n) \\ &= 1 - n < 0. \end{aligned}$$

Let $\varepsilon \in (0,1)$. We will show that for n sufficiently large,

$$C(C_n; 1-n+\varepsilon) > 0$$

when n is odd and

$$C(C_n; 1-n-\varepsilon) > 0$$

when n is even. It will then follow by the Intermediate Value Theorem that C_n must have a connected set root in $(1-n, 1-n+\varepsilon)$ if n is odd, and a connected set root in $(1-n-\varepsilon, 1-n)$ if n is even.

We have

$$\begin{aligned} C(C_n; 1-n+\varepsilon) &= (1-n+\varepsilon)^n + n \cdot \frac{(1-n+\varepsilon)^n - (1-n+\varepsilon)}{-n+\varepsilon} \\ &= (1-n+\varepsilon)^n(1 - \tfrac{n}{n-\varepsilon}) + \tfrac{n}{n-\varepsilon}(1-n+\varepsilon), \end{aligned}$$

which is clearly dominated by the first term for n sufficiently large. Since $1 - \frac{n}{n-\varepsilon} < 0$, this term is positive if and only if n is odd, so that $C(C_n; 1-n+\varepsilon) > 0$ for sufficiently large odd n. Similarly, we have

$$\begin{aligned} C(C_n; 1-n-\varepsilon) &= (1-n-\varepsilon)^n + n \cdot \frac{(1-n-\varepsilon)^n - (1-n-\varepsilon)}{-n-\varepsilon} \\ &= (1-n-\varepsilon)^n(1 - \tfrac{n}{n+\varepsilon}) + \tfrac{n}{n+\varepsilon}(1-n-\varepsilon), \end{aligned}$$

Graph	Nonzero connected set roots of smallest modulus	Modulus
$K_{2,1}$	$-1.0000000000 \pm 1.4142135624i$	1.7320508076
$K_{2,2}$	$-0.4348022826 \pm 1.0434274359i$	1.1303954348
$K_{3,2}$	$-0.1878126933 \pm 0.8572041254i$	0.8775377601
$K_{3,3}$	$-0.1036608393 \pm 0.6865199499i$	0.6943019597
$K_{4,3}$	$-0.0441420165 \pm 0.5823193347i$	0.5839900043
$K_{4,4}$	$-0.0220825810 \pm 0.4969667341i$	0.4974571090

Table 4.1: The nonzero connected set roots of smallest modulus of all graphs of order n for small n. All values rounded to 10 decimal places.

which once again is dominated by the first term for n sufficiently large. Since $1-\frac{n}{n+\varepsilon} > 0$, this term is positive if and only if n is even, so that $C(C_n; 1-n-\varepsilon) > 0$ for sufficiently large even n. □

The lower bound of Theorem 4.2.7 appears to be less tight. We have found using a computer algebra system that for $n \in \{3, \ldots, 8\}$, the graph on n vertices with the nonzero connected set root of smallest modulus is the (nearly) balanced complete bipartite graph $K_{\lceil \frac{n}{2} \rceil, \lfloor \frac{n}{2} \rfloor}$. The nonzero connected set roots of $K_{\lceil \frac{n}{2} \rceil, \lfloor \frac{n}{2} \rfloor}$ of smallest modulus are nonreal for all $n \in \{3, \ldots, 8\}$. The modulus of a smallest root for $n \in \{3, \ldots, 8\}$ is shown in Table 4.1. At least for these small n, the smallest root is relatively far from our lower bound of $\frac{2}{n-1}$.

We can show that the lower bound of Theorem 4.2.7 is never sharp for any graph with a component of order at least 3. Let G be a graph with a component of order at least 3 and let d be the order of a largest component in G. By Theorem 4.2.5, the polynomial

$$C(G; x)/x = c_1 + c_2 x + \ldots + c_d x^{d-1}$$

has a root on the boundary

$$|z| = r = \min\left\{\tfrac{c_i}{c_{i+1}} : 1 \leq i \leq d-1\right\}$$

if and only if the greatest common divisor of the set

$$S = \{j \in \{2,3,\ldots,d+1\}\colon\ c_{j-1} > rc_j\}$$

is strictly greater than 1, where $c_{d+1} = 0$. We way assume that $r = \frac{2}{d-1}$ as otherwise any root z of $C(G;x)/x$ satisfies $|z| \geq r > \frac{2}{d-1}$ and we are done. First of all, note that $d+1 \in S$ as

$$c_d > 0 = c_{d+1}.$$

From the proof of Theorem 4.2.7, we have

$$\frac{c_{j-1}}{c_j} \geq \frac{2}{d-j+1}$$

for all $j \in \{2,\ldots,d\}$. Thus for all $j \in \{3,\ldots,d\}$ we have

$$c_{j-1} \geq \frac{2}{d-j+1}c_j > \frac{2}{d-1}c_j.$$

Therefore, we have $\{3,\ldots,d+1\} \subseteq S$, meaning that $\gcd(S) > 1$. Finally, this implies that $C(G;x)/x$ (and hence $C(G;x)$, as well) has no roots on the boundary $|z| = \frac{2}{d-1}$.

In a slightly different direction, we can show that any connected graph G on n vertices has a connected set root with modulus at least $\frac{n-t}{n}$, where t is the number of cut vertices of G. We will actually demonstrate the truth of the slightly stronger statement that $C(G;x)$ has a connected set root z with $\mathrm{Re}(z) \leq -\frac{n-t}{n}$. We will make use of the Gauss-Lucas Theorem (See [49], for example) that gives a geometrical relation between the roots of a polynomial f and the roots of f'.

Theorem 4.2.10 (Gauss-Lucas)**.** *If f is a non-constant polynomial with complex coefficients, then all roots of f' lie in the convex hull of the set of roots of f.* □

Using the Gauss-Lucas Theorem we can gain information about the roots of

$C(G;x)$ by differentiating until we reach a linear function whose unique root is easy to determine. We use this procedure to prove the next result.

Theorem 4.2.11. *If G is a connected graph on n vertices with t cut vertices, then the connected set polynomial of G has a root z with* $\mathrm{Re}(z) \leq -\frac{n-t}{n}$.

Proof. By Observation 3.1.1, $c_n = 1$ and $c_{n-1} = n - t$, so that the $(n-1)^{\text{st}}$ derivative of $C(G;x)$ is given by

$$C^{(n-1)}(G;x) = n!x + (n-1)!(n-t).$$

Therefore, the unique root of $C^{(n-1)}(G;x)$ is $-\frac{n-t}{n}$. Thus, by the Gauss-Lucas Theorem (4.2.10), the convex hull of the set of roots of $g(G;x)$ must contain the point $-\frac{n-t}{n}$. We conclude that $C(G;x)$ has a root z with $\mathrm{Re}(z) \leq -\frac{n-t}{n}$. □

In particular, if G is 2-connected (i.e. if $t = 0$), then $C(G;z)$ has a root with $\mathrm{Re}(z) \leq -1$. This implies that any 2-connected graph for which $z = -1$ is not a root has a connected set root outside of the disk $|z| \leq 1$.

In the next section, we prove that the collection of all connected set roots is dense in the complex plane. We use an approach similar to the one used to demonstrate that the roots of independence polynomials are dense in the complex plane [19] and that the roots of domination polynomials are dense in the complex plane [24]. The approach relies on the use of lexicographic products to "fan out" connected set roots.

Before we proceed to our density result, we first show that for any fixed graph G, the connected set roots of the family $\{G \odot K_n \colon\ n \in \mathbb{N}\}$ are bounded. By Lemma 4.1.1,

$$C(G \odot K_n; x) = C(G; (1+x)^n - 1)$$

so that x is a root of $C(G \odot K_n; x)$ if and only if

$$(x+1)^n - 1 = r$$

for some root r of $C(G; x)$. Essentially, as n increases, all of the connected set roots of $G \odot K_n$ are fanned out evenly around the point -1, but they are also pulled towards the unit circle centred at -1.

Theorem 4.2.12. *Fix a graph G. Any connected set root x of $G \odot K_n$ satisfies*

$$|x+1| \leq \max\{1, |r_* + 1|\},$$

where r_* is a connected set root of G furthest from -1.

Proof. Let r_* be a root of $C(G; x)$ at greatest distance from 1. From Lemma 4.1.1, any root x of $C(G \odot K_n; x)$ satisfies

$$(x+1)^n = r + 1$$

for some root r of $C(G; x)$. Since r_* is a root of $C(G; x)$ at greatest distance from 1, we have

$$|x+1|^n = |r+1| \leq |r_* + 1|,$$

and hence

$$|x+1| \leq |r_* + 1|^{1/n}.$$

We conlcude that if $|r_* + 1| \leq 1$ then $|x+1| \leq 1$, and if $|r_* + 1| > 1$ then $|x+1| \leq |r_* + 1|$. $\square$

This tells us that if we hope to find a family of graphs of the form

$$\{H \odot K_n\colon\ H \in \mathcal{H}, n \in \mathbb{N}\}$$

whose connected set roots are dense in the complex plane, the collection of connected set roots of all graphs in $\mathcal{H}$ must at least be unbounded – hence $\mathcal{H}$ must contain infinitely many graphs.

4.2.3 The Closure of the Collection of Connected Set Roots

We now build up the necessary tools to prove that the collection of all connected set roots is dense in the complex plane. Results on the closure of the roots of many other graph polynomials have been obtained. For example, the roots of the following polynomials have been shown to be dense in the complex plane: chromatic polynomials [58], domination polynomials [24], independence polynomials [19], and strongly-connected reliability polynomials [16]. On the other hand, while it is known that all-terminal reliability roots are dense in the disk $|z| \leq 1$ (in the variable q), it is unknown whether they are dense in the entire complex plane, and it is suspected that they are not.

Our proof that the collection of all connected set roots is dense in the complex plane involves several steps, so we give a brief summary of our method here. Essentially, in order to find a connected set root close to an arbitrary complex number, we use the lexicographic product with a complete graph to “fan out” a connected set root at appropriate distance from -1 evenly around the point $z = -1$. We first prove that if for every positive real number $r > 0$ there is a connected set root whose distance from -1 is arbitrarily close to r, then the collection of all connected set roots is dense in the complex plane. Our proof follows that given for Theorem 11 in [24] that the roots of domination polynomials are dense in the complex plane. We then present some background on the Beraha-Kahane-Weiss Theorem for finding limits of roots of families of polynomials, which allows us to find a limiting curve of connected set roots that extends from the point $z = -1$ to infinity. This gives us a limit of connected set roots at any distance from -1, which completes the proof, as we can then find a connected set root whose distance from -1 is arbitrarily close to any positive real number r.

Lemma 4.2.13 (adapted from [24])**.** *Suppose that for any $r > 0$ and $\varepsilon > 0$ there is a connected set root z satisfying*

$$||z+1| - r| < \varepsilon;$$

that is, for any real number r there is a connected set root z whose distance from -1 is arbitrarily close to r. Then the collection of all connected set roots is dense in the complex plane.

Proof. Let $r > 0$ and $\theta \in [0, 2\pi)$. It suffices to show that for any $\varepsilon > 0$ there is a root x of a connected set polynomial such that $x + 1$ has modulus within ε of r and argument within ε of θ. We may assume that $\varepsilon < r$, so that $r - \varepsilon > 0$. We can choose m large enough so that $\frac{\pi}{m} < \varepsilon$, and hence for any complex number $w \neq 0$, there is an m^{th} root of w whose argument is within ε of θ.

By the supposition in the theorem statement, there is a connected set root z of some graph G that satisfies

$$(r-\varepsilon)^m < |z+1| < (r+\varepsilon)^m.$$

Consider the graph $G \odot K_m$. By Lemma 4.1.1, any complex number x satisfying

$$(x+1)^m - 1 = z$$

is a root of $C(G \odot K_m; x)$. For any such x, we have

$$|x+1|^m = |z+1|,$$

so that

$$r - \varepsilon < |x+1| < r + \varepsilon.$$

Further, for at least one such x, the argument of $x+1$ is within ε of θ by our choice of m. $\square$

In order to show that connected set roots are dense in the complex plane, by Lemma 4.2.13 it now suffices to show that there is a limiting curve of connected set roots that extends from the point -1 to infinity. We will need a precise definition for a limit of roots.

Definition 4.2.1. If $\{f_n(x)\colon\ n \in \mathbb{N}\}$ is a family of (complex) polynomials, we say that a number $z \in \mathbb{C}$ is a *limit of roots* of $\{f_n(x)\colon\ n \in \mathbb{N}\}$ if there is a sequence $\{z_n\colon\ n \in \mathbb{N}\}$ such that $f_n(z_n) = 0$ and $z_n \to z$ as $n \to \infty$.

Under certain nondegeneracy conditions given in [3], z is a limit of roots of $\{f_n(z)\colon n \in \mathbb{N}\}$ if and only if either $f_n(z) = 0$ for all sufficiently large n, or z is a limit point of the set of all roots of the family. The main result in [3] concerns limits of roots of certain recursively defined families of polynomials. The solution of the recursion

$$P_{n+k}(z) = -\sum_{i=1}^{k} f_i(z) P_{n+k-i}(z)$$

depends on the roots of the characteristic equation

$$Q_z(\lambda) = \lambda^k + \sum_{i=1}^{k} f_i(z)\lambda^{k-i} = 0.$$

Let these roots be $\lambda_1(z), \lambda_2(z), \ldots, \lambda_k(z)$, with possible repetitions. If the $\lambda_i(z)$ are distinct for a particular z, then

$$P_n(z) = \sum_{i=1}^{k} \alpha_i(z)\lambda_i(z)^n, \tag{4.1}$$

where $\alpha_i(z)$ are fixed polynomials determined by solving the system of equations that arises from letting $n = 0, \ldots, k-1$ in (4.1). If there are repeated root values at z, the solution is modified in the usual way (see [3]). For example, if $\lambda_i(z) = \lambda_j(z)$, the term $\alpha_i\lambda_i^n + \alpha_j\lambda_j^n$ is replaced by $\alpha_{i_1}\lambda_i^n + n\alpha_{i_2}\lambda_i^{n-1}$.

Beraha, Kahane, and Weiss characterized the limits of roots of such a recursive family in [3], and Brown and Hickman made the observation that any family of polynomials of the form (4.1) satisfies such a recursion [17]. This gives the following important theorem, which we refer to as the *Beraha-Kahane-Weiss Theorem.*

Theorem 4.2.14 (Beraha-Kahane-Weiss Theorem, cf. [17])**.** *Let*

$$f_n(x) = \alpha_1(x)\lambda_1(x)^n + \alpha_2(x)\lambda_2(x)^n + \ldots + \alpha_k(x)\lambda_k(x)^n,$$

where the $\alpha_i(x)$ and $\lambda_i(x)$ are fixed non-zero polynomials such that for no pair $i \neq j$ is $\lambda_i(x) = \omega\lambda_j(x)$ for some $\omega \in \mathbb{C}$ of unit modulus. Then $z \in \mathbb{C}$ is a limit of roots of the family $\{f_n(x) \colon\ n \in \mathbb{N}\}$ if and only if either

(i) *two or more of the $\lambda_i(z)$ are of equal modulus, and strictly greater (in modulus) than the others; or*

(ii) *for some j, $\lambda_j(z)$ has modulus strictly greater than all the other $\lambda_i(z)$ have, and $\alpha_j(z) = 0$.*

The same characterization holds when the characteristic equation of the associated recursion has repeated roots. In particular, if the term $\alpha_i\lambda_i^n + \alpha_j\lambda_j^n$ in $f_n(x)$ is replaced by $\alpha_{i_1}\lambda_i^n + n\alpha_{i_2}\lambda_i^{n-1}$, the same conclusion holds. In this case part (ii) *needs to be* *reworded slightly to: for some j, $\lambda_j(z)$ has modulus strictly greater than all the other $\lambda_i(z)$ have, and $\alpha_{j_k}(z) = 0$ for some k.*

We now present a simple decomposition of the connected set polynomial of an arbitrary graph G which will be useful in computing the connected set polynomials of the family of graphs used in the proof of our main result. The idea behind the decomposition is to pick a vertex v of G and to count the connected sets of G that contain v separately from the connected sets that do not contain v. An analogous decomposition holds for many generating polynomials related to graphs, and perhaps the best known example of such a decomposition is for the independence polynomial $I(G; x)$ (see [39], for example). The polynomial $xI(G - N[v]; x)$ enumerates the independent sets of G that contain v (recall that $N[v]$ denotes the closed neighbourhood of v in G) while the polynomial $I(G - v; x)$ enumerates the independent sets of G that do not contain v. This gives

$$I(G; x) = I(G - v; x) + xI(G - N[v]; x).$$

For the connected set polynomial, the connected sets of G that do not contain v are counted by the polynomial $C(G - v; x)$, and we introduce the *rooted connected set polynomial* to count the connected sets of G that contain v.

Definition 4.2.2. Let G be a rooted graph on n vertices with root v. Let $\mathcal{C}_v(G) \subseteq \mathcal{C}(G)$ be the collection of all connected sets of G containing vertex v. The *rooted connected set polynomial* of G at v, denoted $C_v(G; x)$, is the generating polynomial of the collection $\mathcal{C}_v(G)$; that is,

$$C_v(G; x) = \sum_{C \in \mathcal{C}_v(G)} x^{|C|}.$$

Alternatively, we can write

$$C_v(G;x) = \sum_{k=1}^{n} r_k x^k,$$

where r_k is the number of connected sets of G of order k that contain v.

Observation 4.2.15. *Let G be a graph of order $n \geq 1$ and let v be any vertex of G. Then*

$$C(G;x) = C(G-v;x) + C_v(G;x).$$

Proof. We can partition the connected sets of G into those which contain v and those which do not. The polynomial $C_v(G;x)$ enumerates the former while the polynomial $C(G-v;x)$ enumerates the latter. □

First of all, we can use Observation 4.2.15 to compute $C(P_n;x)$. Let $i \in \{1,\ldots,n\}$ and let v_i be a leaf of the graph P_i. It is clear that

$$C_{v_i}(P_i;x) = \sum_{k=1}^{i} x^k,$$

as there is a unique connected set of P_i of order k containing the leaf v_i for each $k \in \{1,\ldots,i\}$. By Observation 4.2.15, we have

$$C(P_n;x) = C(P_{n-1};x) + C_{v_n}(P_n;x)$$

Applying Observation 4.2.15 recursively leads to

$$\begin{aligned} C(P_n;x) &= \sum_{i=1}^{n} C_{v_i}(P_i;x) \\ &= \sum_{i=1}^{n}\sum_{k=1}^{i} x^k \\ &= \sum_{k=1}^{n} (n-k+1)x^k \end{aligned}$$

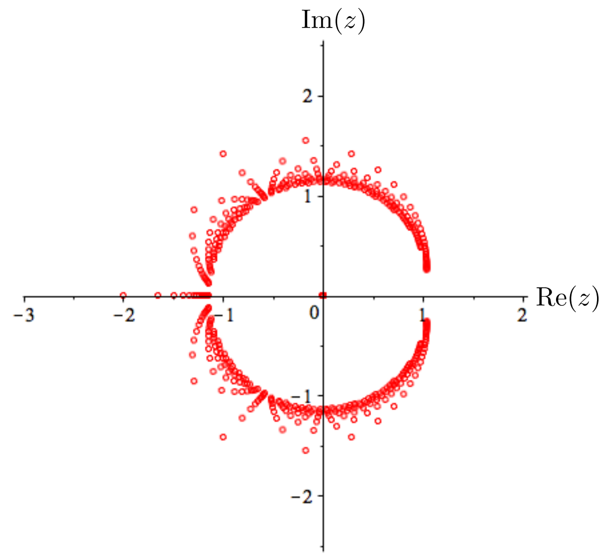


Figure 4.5: The connected set roots of P_n for $n \in \{1, \ldots, 30\}$.

Hence $c_k(P_n) = n - k + 1$ for all $k \in \{1, \ldots, n\}$. This can be verified directly as a connected set of order k in P_n must consist of k consecutive vertices, and there are $n - k + 1$ sets of k consecutive vertices in P_n.

We will need a closed form for $C(P_n; x)$, and this is easily obtained. We have

$$\begin{aligned}(x-1)C(P_n;x) &= xC(P_n;x) - C(P_n;x)\\ &= x^{n+1} + x^n + \ldots + x^2 - nx\\ &= \frac{x^{n+2} - x^2}{x-1} - nx,\end{aligned}$$

so that

$$C(P_n;x) = \frac{x^{n+2} - x^2}{(x-1)^2} - \frac{nx}{x-1}.$$

The connected set roots of the path P_n for $n \leq 30$ are shown in Figure 4.5. In Section 5.4 we will prove a result that implies the modulus of any connected set root of P_n (for any n) is at most 2. However, we see below that joining a single vertex to the path P_n has a drastic effect on the connected set roots.

Lemma 4.2.16. *Let G be a graph on n vertices. The connected set polynomial of the* *join $G+v$ is given by*

$$C(G+v;x) = C(G;x) + x(x+1)^n.$$

Proof. By Observation 4.2.15,

$$C(G+v;x) = C(G;x) + C_v(G+v;x).$$

The connected sets of $G+v$ containing v correspond simply to the vertex subsets of G, since v is adjacent to every vertex of G. Explicitly, $U \subseteq V(G)$ corresponds to the connected set $U \cup \{v\}$ of $G+v$. Hence the connected sets of $G+v$ containing v are enumerated by $x(x+1)^n$. This gives

$$C(G+v;x) = C(G;x) + x(x+1)^n. \qquad \square$$

The connected set roots of $P_n + v$ for $n \leq 50$ are shown in Figure 4.6, and one can see that there are roots that appear to grow large in several directions. We are now ready to prove our main result – the proof entails showing that the limiting curve of the connected set roots of the family of graphs $\{P_n + v\colon\ n \in \mathbb{N}\}$ extends from -1 to infinity. The result then follows immediately by Lemma 4.2.13.

Theorem 4.2.17. *The collection of all connected set roots is dense in the complex plane, even if we restrict to connected graphs.*

Proof. It suffices to show that the supposition of Lemma 4.2.13 is true. We do so by proving that the limits of roots of the family $\{C(P_n+v;x)\colon\ n \geq 1\}$ include the points on the line $\mathrm{Re}(z) = -\frac{1}{2}$ of modulus at least one, the points on the circle $|z+1| = 1$ with $\mathrm{Re}(z) \geq -\frac{1}{2}$, and the points on the circle $|z| = 1$ with $\mathrm{Re}(z) \leq -\frac{1}{2}$. See Figure 4.7 for an illustration of the limiting curve, which clearly extends from -1 to ∞. For

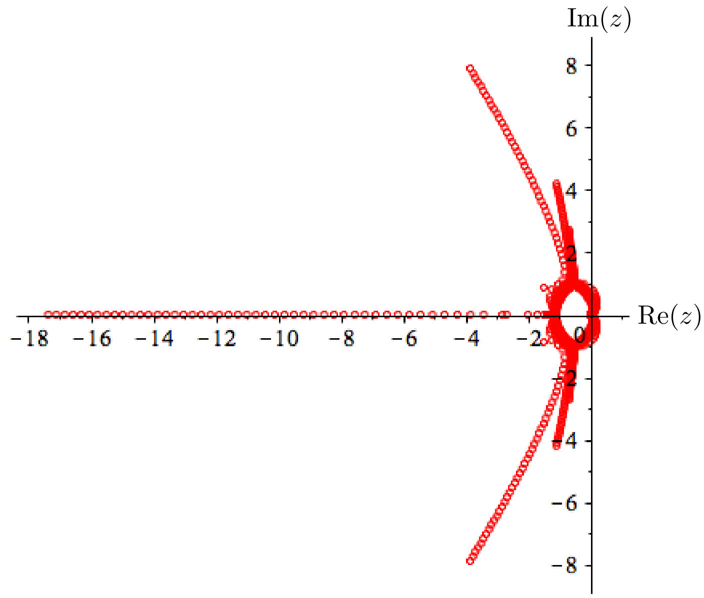


Figure 4.6: The connected set roots of $P_n + v$ for $n \in \{1, \ldots, 50\}$.

$n \geq 1$, the connected set polynomial of $P_n + v$ is given by

$$\begin{aligned} C(P_n + v; x) &= C(P_n; x) + x(x+1)^n \\ &= \frac{x^{n+2} - x^2}{(x-1)^2} - \frac{nx}{x-1} + x(x+1)^n \end{aligned}$$

Consider the polynomial $f_n(x) = (x-1)^2 C(P_n + v; x)$. We multiply by $(x-1)^2$ to clear the denominators of the rational terms and this adds only a simple root at $x = 1$. We rewrite $f_n(x)$ as follows:

$$\begin{aligned} f_n(x) &= x^{n+2} - x^2 - nx(x-1) + (x-1)^2 x(x+1)^n \\ &= x(x-1)^2(x+1)^n + x^2 \cdot x^n - x^2 - nx(x-1) \\ &= \alpha_1 \lambda_1^n + \alpha_2 \lambda_2^n + \alpha_{3_1} \lambda_3^n + n\alpha_{3_2} \lambda_3^{n-1}, \end{aligned}$$

where

$$\alpha_1 = x(x-1)^2, \;\; \alpha_2 = x^2, \;\; \alpha_{3_1} = -x^2, \;\; \text{and } \alpha_{3_2} = -x(x-1);$$

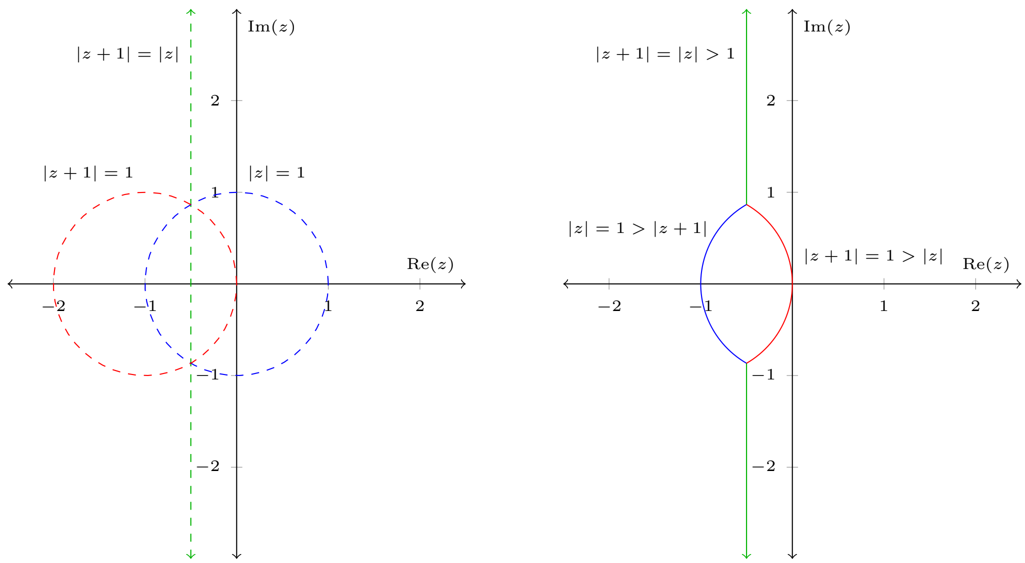


Figure 4.7: The curves $|z+1|=|z|$, $|z+1|=1$, and $|z|=1$ (left), and the limiting curve for the connected set roots of P_n+v (right).

and

$$\lambda_1 = x+1,\ \lambda_2 = x,\ \text{and } \lambda_3 = 1.$$

Clearly no α_i is identically zero and no $\lambda_i = \omega\lambda_j$ for $i \neq j$ and some complex number ω of unit modulus, so the nondegeneracy conditions of Theorem 4.2.14 are satisfied. Applying part (i) of Theorem 4.2.14 involves three cases. Figure 4.7 is provided to aid the reader in seeing the characterizations below.

Case (i): $|\lambda_1| = |\lambda_2| \geq |\lambda_3|$

The condition $|z+1| = |z|$ is true if and only if z is equidistant from -1 and 0; that is, if and only if $\mathrm{Re}(z) = -\frac{1}{2}$. Further, when $|z+1| = |z|$ we have $|z+1| \geq 1$ and $|z| \geq 1$ if and only if z has modulus at least one. Hence, we have $|z+1| = |z| \geq 1$ if and only if z lies on the line $\mathrm{Re}(z) = -\frac{1}{2}$ and z has modulus at least one.

Case (ii): $|\lambda_1| = |\lambda_3| \geq |\lambda_2|$

The condition $|z+1| = 1$ is true if and only if z lies on the circle of radius 1 centred at the point -1. Further, when $|z+1| = 1$ we have $|z+1| \geq |z|$ and $1 \geq |z|$ if and only if $\mathrm{Re}(z) \geq -\frac{1}{2}$. Hence, we have $|z+1| = 1 \geq |z|$ if and only if z lies on the circle of radius 1 centred at -1 and $\mathrm{Re}(z) \geq -\frac{1}{2}$.

Case (iii): $|\lambda_2| = |\lambda_3| \geq |\lambda_1|$

The condition $|z| = 1$ is true if and only if z lies on the circle of radius 1 centred at the point 0. Further, when $|z| = 1$ we have $|z| \geq |z+1|$ and $1 \geq |z+1|$ if and only if $\mathrm{Re}(z) \leq -\frac{1}{2}$. Hence, we have $|z| = 1 \geq |z+1|$ if and only if z lies on the circle of radius 1 centred at 0 and $\mathrm{Re}(z) \leq -\frac{1}{2}$.

Finally, it is straight forward to verify that part (ii) of Theorem 4.2.14 yields only a single limit of roots at $x = 1$. We suspect that this limit of roots is only the result of multiplying by $(x-1)^2$.

Since the limiting curve of the connected set roots of the graphs $P_n + v$ extends continuously from the point -1 to infinity and we can find a connected set root arbitrarily close to any point on this curve, the supposition of Lemma 4.2.13 is satisfied. We conclude that connected set roots are dense in the complex plane. The result holds even if we restrict to connected graphs as we have only used the connected set roots of the connected graphs $(P_n + v) \odot K_m$ for $n, m \in \mathbb{N}$. $\square$

We can also make a similar statement to Theorem 4.2.17 for *disconnected* graphs, which is surprising as asymptotically almost all graphs of order n are connected. The proof is very similar to that of Theorem 4.2.17, but we use the disconnected graphs $P_n \cup K_n$ in place of the connected graphs $P_n + v$. The connected set roots of $P_n \cup K_n$ for $n \leq 30$ are shown in Figure 4.8, and appear to have the same limiting curve as the connected set roots of the family $P_n + v$ shown in Figure 4.7. We confirm that this is true in the proof of the next theorem.

Theorem 4.2.18. *The collection of connected set roots of disconnected graphs is dense in the complex plane.*

Proof. It suffices to show that the supposition of Lemma 4.2.13 is true even for disconnected graphs. We will show that the limiting curve of the connected set roots of

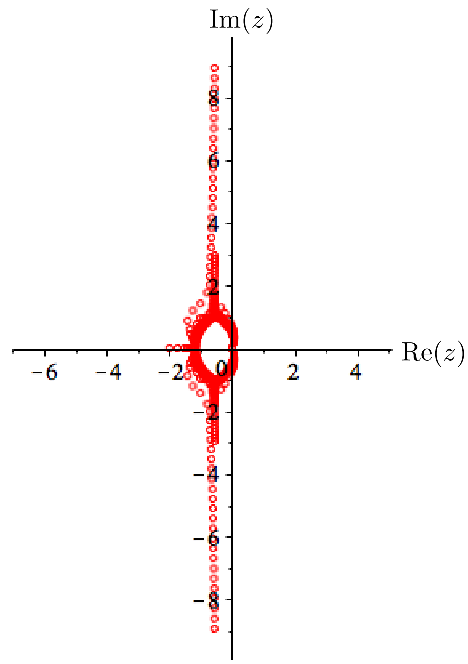


Figure 4.8: The connected set roots of $P_n \cup K_n$ for $n \in \{1, \ldots, 30\}$.

the family of graphs $\{P_n \cup K_n \colon\ n \geq 1\}$ extends from -1 to infinity. By Lemma 4.0.1 we have

$$\begin{aligned} C(P_n \cup K_n; x) &= C(P_n; x) + C(K_n; x) \\ &= \frac{x^{n+2} - x^2}{(x-1)^2} - \frac{nx}{x-1} + (x+1)^n - 1 \end{aligned}$$

As in the proof of Theorem 4.2.17, consider the polynomials

$$g_n(x) = (x-1)^2 C(P_n \cup K_n; x).$$

We rewrite $g_n(x)$ as follows:

$$\begin{aligned} g_n(x) &= x^{n+2} - x^2 - nx(x-1) + (x-1)^2(x+1)^n - (x-1)^2 \\ &= (x-1)^2(x+1)^n + x^2 \cdot x^n - x^2 - (x-1)^2 - nx(x-1) \\ &= \alpha_1 \lambda_1^n + \alpha_2 \lambda_2^n + \alpha_{3_1} \lambda_3^n + n\alpha_{3_2} \lambda_3^{n-1}, \end{aligned}$$

where

$$\alpha_1 = x(x-1)^2,\ \alpha_2 = x^2,\ \alpha_{3_1} = -x^2 - (x-1)^2,\ \text{and}\ \alpha_{3_2} = -x(x-1);$$

and

$$\lambda_1 = x+1,\ \lambda_2 = x,\ \text{and}\ \lambda_3 = 1.$$

From here we see that the limiting curve of the connected set roots of $P_n \cup K_n$ is the same as the limiting curve of the connected set roots of $P_n + v$ shown in Figure 4.2.17. Since the limiting curve extends from -1 to infinity, we conclude by Lemma 4.2.13 that the connected set roots of disconnected graphs are dense in the complex plane. □

Finally, as a corollary to Theorem 4.2.17 and Theorem 4.2.18 we can prove that the collection of all node reliability roots is also dense in the complex plane, whether we restrict to connected graphs or disconnected graphs. This is potentially another significant difference between node reliability and all-terminal reliability. While all-terminal reliability roots were shown to be dense in the disk $|z| \leq 1$ (in the variable q) in [12], the largest known modulus of an all-terminal reliability root is 1.1134860896, found in Chapter 2 of this thesis! It seems as though all-terminal reliability roots are bounded in modulus by some constant, which would make them far from dense in the complex plane. If this is indeed the case then we have found another striking difference between all-terminal reliability and node reliability.

Corollary 4.2.19. *The collection of all node reliability roots is dense in the complex plane, even if we restrict to connected graphs (or disconnected graphs).*

Proof. Let $z \in \mathbb{C}$ and let $\varepsilon > 0$. We will find a complex number $\tilde{z}$ and a graph G such that $|z-\tilde{z}|<\varepsilon$ and $\mathrm{nRel}(G;\tilde{z})=0$. Recall that

$$C(G;x)=(1+x)^n\cdot\mathrm{nRel}\big(G;\tfrac{x}{1+x}\big),$$

so that any connected set root $x\neq-1$ of the graph G yields a node reliability root $\frac{x}{1+x}$ of G. The complex function $f(x)=\frac{x}{1+x}$ is a Möbius transformation and hence it is one-to-one and continuous on its domain. Let $x=f^{-1}(z)=\frac{z}{1-z}$ (we may assume that $z\neq 1$). Since f is continuous, we can find $\delta>0$ such that $|x-\tilde{x}|<\delta$ implies $|f(x)-f(\tilde{x})|=|z-f(\tilde{x})|<\varepsilon$. By Theorem 4.2.17, the connected set roots of connected graphs are dense in the complex plane, and hence there is some connected graph G with connected set root $\tilde{x}\neq-1$ satisfying $|x-\tilde{x}|<\delta$. The complex number $f(\tilde{x})=\frac{\tilde{x}}{1+\tilde{x}}$ satisfies $|z-f(\tilde{x})|<\varepsilon$ and $\mathrm{nRel}(G;f(\tilde{x}))=0$.

The result for disconnected graphs follows analogously from Theorem 4.2.18. $\square$

Chapter 5

The Subtree Polynomial

A *subtree* of a tree T is a connected induced subgraph of T with at least one vertex. Recently, there has been much interest in enumerating the subtrees of a tree. The problem of finding the tree in a given class that maximizes (minimizes) the total number of subtrees has been of primary interest. It is not hard to demonstrate that among the class of all trees on n vertices, the path P_n has the least subtrees, while the star $K_{1,n-1}$ has the most (see [63]). In [64], Székely and Wang describe the binary tree with the maximum (minimum) number of subtrees among all binary trees with n leaves. In [73], Yan and Yeh describe the tree on n vertices with diameter at least d which has the maximum number of subtrees, and the tree on n vertices with maximum degree at least Δ which has the minimum number of subtrees. In [76], the authors generalize a result from [44] in characterizing the tree with given degree sequence that has the largest number of subtrees. On the other hand, the tree with given degree sequence that has the minimum number of subtrees is not always unique; this problem was studied in [56, 75].

The number of subtrees of a tree has a strong connection to combinatorial chemistry, where *topological indices* are used to describe the structural properties of graphs. One such index, the *Wiener index* of a graph (sometimes called the *path number*), is the sum of the distances between all pairs of vertices. The Wiener index has been studied extensively as a molecular descriptor (see [31, 71], for example). There is an interesting negative correlation between the number of subtrees of a tree and its Wiener index, first pointed out in [63]: among certain classes of trees, the tree that maximizes the number of subtrees minimizes the Wiener index, and the tree that minimizes the number of subtrees maximizes the Wiener index. In fact, the negative correlation between the Wiener index and the number of subtrees has been shown to be stronger than the correlation between the Wiener index and several other common topological indices used in combinatorial chemistry (see [69]).

Apart from this connection to the Wiener index, the study of the number of subtrees of a tree is of interest in its own right in fields including graph theory, number theory, and computer science. The number of subtrees of a tree also has applications to the study of phylogenetics (see [45], for example).

In [73], the authors studied a weighted variant of the problem. The *weight* of a subtree S of a tree T is the product of the weights on the vertices and edges of S. The authors found a linear-time algorithm to count the sum of the weights of all subtrees of a tree T. When both weight functions map identically to 1, this sum is exactly the number of subtrees of T. Alternatively, if the weight function on the vertex set maps identically to x while the weight function on the edge set maps identically to 1, the sum of the weights is a polynomial in x whose coefficients count the number of subtrees of T of each order. Note that the subtrees of T are in one-to-one correspondence with their vertex sets, which are exactly the connected sets of T. Thus, when the weight function on the vertex set maps identically to x while the weight function on the edge set maps identically to 1, the sum of the weights of all subtrees is exactly $C(T;x)$. This polynomial is only briefly mentioned in [73], but we commit this chapter to its study.

In Section 5.1 we prove that the graphs for which the connected sets are the same as the nonempty *g-convex* sets are exactly the block graphs (this result is included here as trees are block graphs). The characterization is of interest as the *g-convexity polynomial* – the generating polynomial of the collection of g-convex sets – has been studied in [23, 22]. In particular, it means that most of the results proven about the connected set polynomials of trees in this chapter apply immediately to the g-convexity polynomials of trees as well (some results on the roots of connected set polynomials of trees will require small adjustments because the constant term of the connected set polynomial is always 0 while the constant term of the g-convexity polynomial is 1). In Section 5.2 we prove that the path (star) on n vertices has the coefficient-wise smallest (largest) connected set polynomial among the class of all trees on n vertices. This strengthens the result that the path (star) on n vertices has the fewest (greatest) total number of subtrees among all trees on n vertices (see [63]). In Section 5.3, we characterize the trees that have unimodal connected set polynomial and the trees that have log-concave connected set polynomial. Finally, in Section 5.4 we study the roots of connected set polynomials of trees. While we saw in Section 4.2.3 that the collection of connected set roots of all connected graphs is dense in the complex plane, it appears that the collection of connected set roots of trees is bounded. We present a bound on the modulus of a connected set root of a tree in terms of the number of leaves, which improves on the bound of Theorem 4.2.7 for connected graphs in general.

5.1 Connected Sets and Convexity

Although we are interested here in connected sets, there has been considerable interest in the literature in *convex sets*, which are connected (at least in a connected graph) but also satisfy a stronger property. Several notions of convexity exist for vertex subsets of graphs (see [32]), all of which try to capture the essential properties of convexity for a subset of Euclidean space. The notion of convexity that we study here is called *geodesic convexity* or *g-convexity.*

Definition 5.1.1 (c.f. [32])**.** A nonempty subset X of vertices of a graph G is called *geodesically convex (g-convex)* if $G[X]$ is connected and whenever u and v belong to

X, all vertices on shortest paths between u and v also lie in X. Additionally, the empty set is g-convex.

Note that many definitions of g-convex set do not require that the set induces a connected graph. We have chosen to include this stipulation because it more closely mirrors convexity in Euclidean space, where there is always at least one shortest path between every pair of points, and thus every convex set is connected. The requirement that a g-convex set must induce a connected graph implies that every nonempty convex set must be a connected set, which is not true in disconnected graphs otherwise.

The generating polynomial of the collection of g-convex sets of a graph was studied by Brown and Oellermann in [23, 22]. We call this polynomial the *g-convexity polynomial*.

Definition 5.1.2. Let $\mathcal{X}(G)$ be the collection of g-convex sets of G. The *g-convexity polynomial* of G, denoted $g(G;x)$, is the generating polynomial of the collection of convex sets of G; that is,

$$g(G;x) = \sum_{X \in \mathcal{X}(G)} x^{|X|}.$$

Alternatively, we can write

$$g(G;x) = \sum_{k=0}^{n} g_k x^k,$$

where g_k is the number of g-convex sets of G of order k for each $k \in \{0, \ldots, n\}$.

We briefly summarize the work that has been done to date on g-convexity polynomials. In any graph G, the empty set, the singletons, the edges, and the entire vertex set are necessarily g-convex sets. The graphs for which these are the only convex sets, called *g-minimal graphs*, were studied in [22]. It is clear that g-minimal graphs have the fewest convex sets among the class of all graphs on the same number of vertices and edges. While it was found that the collection of all roots of g-convexity polynomials is unbounded, it was proven that no root of a g-convexity polynomial of a g-minimal graph has modulus exceeding $\beta \approx 2.1475$.

The main purpose of this section is to describe the graphs for which the connected set polynomial and the g-convexity polynomial are the same up to the constant term. Several relationships between the coefficients c_k of the connected set polynomial of a connected graph G and the coefficients g_k of the g-convexity polynomial of G are obvious. Note that $c_k = g_k$ for $k \in \{1, 2, n\}$ as for any connected graph on n vertices the singletons, the edges, and the entire vertex set are both connected and g-convex. Further, every nonempty convex set of a graph G is also a connected set, so we must have

$$g_k \leq c_k$$

for all $k \in \{1, \ldots, n\}$. Here we characterize the connected graphs for which equality holds for all $k \in \{1, \ldots, n\}$. That is, we determine all graphs G for which

$$W \text{ is a connected set} \quad \Rightarrow \quad W \text{ is a convex set}$$

for any $W \subseteq V(G)$. We introduce the term *simply convex* to describe the graphs for which this property holds. For all such graphs, the connected set polynomial and the g-convexity polynomial are identical up to the constant term (remember that the empty set is considered a convex set but not a connected set). This characterization is of interest because the study of the connected set polynomial is very closely linked to the study of the g-convexity polynomial for simply convex graphs.

Definition 5.1.3. A connected graph G is called *simply convex* if for any nonempty subset $W \subseteq V(G)$, W is connected if and only if W is g-convex.

We first show that simply convex graphs do not have any induced cycles of order 4 or more or any induced *diamonds*. The *diamond* $D = K_4 - e$ is the unique simple

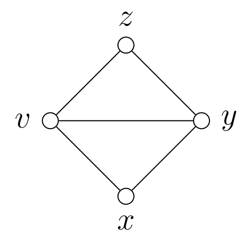


Figure 5.1: The diamond graph D.

graph (up to isomorphism) on 4 vertices and 5 edges.

Lemma 5.1.1. *If G is simply convex then G is a diamond-free chordal graph.*

Proof. Let G be simply convex and suppose G contains an induced cycle of length $k \geq 4$. Label the vertices $v_1, v_2, \ldots, v_k$ in cyclic order. The subset $\{v_1, v_2, \ldots, v_{k-1}\}$ is connected but not convex, as it does not contain the shortest path (v_{k-1}, v_k, v_1) between v_{k-1} and v_1.

Suppose now that G contains a diamond graph as an induced subgraph. Without loss of generality let the diamond be labelled as in Figure 5.1. Then the set $\{x, y, z\}$ is connected but not convex in G, as it does not contain the shortest path (x, v, z) between x and z. $\square$

Lemma 5.1.1 leads us to a nice characterization of simply convex graphs. Recall that a *block graph* is a connected graph in which every block is a clique. We make use of several known characterizations of block graphs to demonstrate that a graph G is simply convex if and only if it is a block graph.

Theorem 5.1.2. *Let G be a graph. The following are equivalent:*

(i) *G is a block graph;*

(ii) *G is a diamond-free chordal graph;*

(iii) *Between every two vertices in G there is exactly one induced chordless path;*

(iv) *G is simply convex.*

Proof. The equivalences (i) $\Leftrightarrow$ (ii) and (ii) $\Leftrightarrow$ (iii) are stated in Theorem 1.1 of [46]. Lemma 5.1.1 covers the implication (iv) $\Rightarrow$ (ii). Lastly, we show (iii) $\Rightarrow$ (iv). Suppose that between every two vertices in G there is exactly one induced chordless path. This means that there is a unique shortest path between every two vertices in G and that this shortest path is contained in any connected induced subgraph of G containing this pair of vertices. Therefore, G is simply convex. $\square$

Theorem 5.1.2 has some important implications for the remainder of the material in this chapter. Since trees are block graphs, for any tree T we have

$$C(T;x) + 1 = g(T;x).$$

The differing constant term has little to no effect on most of the results that follow in the remainder of this chapter. In particular, the main results of Section 5.2 and Section 5.3 (Theorem 5.2.1 and Theorem 5.3.6, respectively) apply equally well to g-convexity polynomials. The roots of g-convexity polynomials of trees differ slightly from the roots of the corresponding connected set polynomials due to the different constant term, but overall we have found that they are still very similar in nature. Our results on the roots of connected set polynomials of trees can be reworked slightly into very similar results that hold for the roots of g-convexity polynomials of trees.

5.2 Paths and Stars

Since the focus of much of the existing research on subtrees has been on finding the tree with the maximum (minimum) total number of subtrees in a given class (if such a tree exists), it is natural to ask the following question: is there a tree in a given class that maximizes (minimizes) *all* of the coefficients of the connected set polynomial?

Of course, such a tree would necessarily have the maximum (minimum) total number of subtrees among all trees in the class, but satisfies a stronger property in that it has more (fewer) subtrees *of each order* than any other tree. We answer this question for the class of all trees on n vertices. Recall that the star on n vertices has the most subtrees in this class while the path on n vertices has the least [63]. It is obvious that all trees of order n satisfy $c_1 = n$, $c_2 = n-1$, and $c_n = 1$, so our result is concerned only with the coefficients c_k for $k \in \{3, \ldots, n-1\}$.

Theorem 5.2.1. *Let T be a tree of order $n \geq 4$. For any $k \in \{3, \ldots, n-1\}$,*

$$n - k + 1 \leq c_k \leq \binom{n-1}{k-1},$$

with equality on the left if and only if $T = P_n$ and equality on the right if and only if $T = K_{1,n-1}$.

The rest of this section is essentially devoted to proving Theorem 5.2.1. We found in Section 4.2.3 that the path P_n has connected set polynomial

$$C(P_n; x) = \sum_{k=1}^{n} (n-k+1)x^k,$$

whose coefficients match the lower bound of Theorem 5.2.1.

The connected set polynomial of the star $K_{1,n-1}$ can be computed easily using Lemma 4.2.16, as $K_{1,n-1} \cong O_{n-1} + v$, where O_{n-1} is the empty graph on $n-1$ vertices. We have

$$\begin{aligned} C(K_{1,n-1}; x) &= C(O_{n-1}; x) + x(x+1)^{n-1} \\ &= (n-1)x + \sum_{k=1}^{n} \binom{n-1}{k-1} x^k \\ &= nx + \sum_{k=2}^{n} \binom{n-1}{k-1} x^k. \end{aligned}$$

Thus we see that the path and star of order n meet the lower and upper bounds, respectively, of Theorem 5.2.1. The next lemma will show that any tree on n vertices not isomorphic to the path P_n must have strictly more connected sets of each order $k \in \{3, \ldots, n-1\}$ than the path P_n.

Lemma 5.2.2. *Let T be a tree on n vertices that is not isomorphic to P_n. If $k \in \{3, \ldots, n-1\}$ then*

$$c_k(T) > n - k + 1.$$

Proof. We proceed by mathematical induction on n. For the base case, suppose that T is a tree of order 4 and that $T \ncong P_4$. Then $T \cong K_{1,3}$ and

$$C(T; x) = 1 + 4x + 3x^2 + 3x^3 + x^4.$$

Thus, $c_3(T) = 3 > 2 = n - 3 + 1$, and the base case is satisfied.

Suppose now that for some $n \geq 4$, every tree T of order n not isomorphic to P_n satisfies $c_k > n + k - 1$ for all $k \in \{3, \ldots, n-1\}$. Let T be a tree of order $n+1$ with $T \ncong P_{n+1}$. Let v be a leaf of T such that $T - v \ncong P_n$. Such a leaf v is guaranteed to exist by the conditions $T \ncong P_{n+1}$ and $n \geq 4$. We note that the graph $T - v$ is connected since v is a leaf. By Observation 4.2.15,

$$C(T; x) = C(T - v; x) + C_v(T; x).$$

Let $C_v(T; x) = \sum_{k=1}^{n+1} r_k x^k$. Then we have

$$c_k(T) = c_k(T - v) + r_k(T).$$

Since $T - v \ncong P_n$ we have $c_k(T - v) > n - k + 1$ for $k \in \{3, \ldots, n-1\}$ by the induction hypothesis. Note also that $c_n(T - v) = 1$ as $T - v$ is connected. Since T is connected, there must be some connected set of order k containing v for each $k \in \{3, \ldots, n\}$, and hence $r_k(T) \geq 1$. Putting all of this together, we have

$$c_k(T) > (n-k+1) + 1 = (n+1) - k + 1$$

for all $k \in \{3, \ldots, n\}$. Therefore, by mathematical induction, any tree T of order $n \geq 4$ not isomorphic to P_n satisfies $c_k(T) > n - k + 1$ for all $k \in \{3, \ldots, n-1\}$. □

We now turn our attention to the upper bound of Theorem 5.2.1. We will demonstrate that any tree on n vertices not isomorphic to the star $K_{1,n-1}$ must have strictly fewer connected sets of each order $k \in \{3, \ldots, n-1\}$ than $K_{1,n-1}$.

Lemma 5.2.3. *Let T be a tree on n vertices that is not isomorphic to $K_{1,n-1}$. If $k \in \{3, \ldots, n-1\}$ then*

$$c_k(T) < \binom{n-1}{k-1}.$$

Proof. The connected sets of order $k \geq 3$ of a tree T correspond to the subtrees of T of order k, which correspond to their set of $k-1$ edges. In other words, each connected set of order $k \geq 3$ in a tree corresponds to a set of $k-1$ edges of the tree. For the star, every subset of edges induces a subtree, meaning that

$$c_k(K_{1,n-1}) = \binom{n-1}{k-1}.$$

Let $T \ncong K_{1,n-1}$ be a graph of order n. We will show that for each $k \in \{3, \ldots, n-1\}$, the tree T has a set of $k-1$ edges that induces a disconnected graph, so that $c_k(T) < \binom{n-1}{k-1}$ for each $k \in \{3, \ldots, n-1\}$.

Recall that a pendant edge in a graph G is an edge incident to a leaf. Since $T \ncong K_{1,n-1}$, there must be some non-pendant edge $e = \{u, v\}$ in $E(T)$. Pick edges e_u and e_v which are incident to u and v, respectively, but distinct from e. Any set of edges containing e_u and e_v but not e induces a subgraph with at least two components (i.e. not a subtree). Let the edges in $E(T)\backslash\{e, e_u, e_v\}$ be $e_1, e_2, \ldots, e_{n-3}$. For each $k \in \{3, \ldots, n-1\}$, the set

$$E_k := \begin{cases} \{e_u, e_v\} & \text{if } k = 3 \\ \{e_u, e_v, e_1, \ldots, e_{k-3}\} & \text{if } k \in \{4, ..., n-1\} \end{cases}$$

contains $k-1$ edges and induces a disconnected subgraph. Therefore not every set of $k-1$ edges induces a subtree of T and hence $c_k < \binom{n-1}{k-1}$ for all $k \in \{3, \ldots, n-1\}$. □

Lemma 5.2.3 is the last piece in the proof of Theorem 5.2.1. In summary, we have improved the result from [63] that among all trees on $n \geq 4$ vertices, the path has the least subtrees while the star has the most. We can now say that among all trees on $n \geq 4$ vertices, for all $k \in \{3, \ldots, n-1\}$ the path has the least subtrees of order k while the star has the most subtrees of order k. It would be interesting to know whether other results such as those from [64, 73, 76] that give the tree in a given class with the largest (or smallest) total number of subtrees can be strengthened in this way. We close this section with a corollary to Theorem 5.2.1 which lends yet more weight to the result, and gives more motivation to determine whether there is a tree in a given class with the coefficient-wise greatest or least subtree polynomial. We say that a graph G has *uniformly best* node reliability among the collection of graphs $\mathcal{H}$ if $\mathrm{nRel}(G;p) \geq \mathrm{nRel}(H;p)$ on $[0,1]$ for every $H \in \mathcal{H}$. *Uniformly worst* is defined analogously.

Corollary 5.2.4. *Among the class of all trees on n vertices, the path P_n has the uniformly worst node reliability while the star $K_{1,n-1}$ has the uniformly best node reliability.*

Proof. Let T be a tree on n vertices. By Theorem 5.2.1,

$$c_k(P_n) \leq c_k(T) \leq c_k(K_{1,n-1})$$

for all $k \in \{1, \ldots, n\}$, and hence

$$\mathrm{nRel}(P_n; p) \leq \mathrm{nRel}(T; p) \leq \mathrm{nRel}(K_{1,n-1}; p). \qquad \square$$

We note that the star on n vertices was known to be uniformly best among the class of all trees on n vertices (see [61]), but the result that the path on n vertices is uniformly worst in this class is new to us.

5.3 Unimodality and Log-Concavity

A sequence $a_0, a_1, \ldots, a_n$ is said to be *unimodal* if there is some k for which

$$a_0 \leq a_1 \leq \ldots \leq a_k \geq a_{k+1} \geq \ldots a_n.$$

In this case k is called a *mode* of the sequence (note that a mode is not necessarily unique). The sequence $a_0, a_1, \ldots, a_n$ is called *log-concave* if

$$a_k^2 \geq a_{k-1}a_{k+1} \text{ for all } k \in \{1, \ldots, n-1\}.$$

It is not hard to see that a log-concave sequence of positive terms is unimodal. A polynomial is called unimodal (log-concave) if its sequence of coefficients is unimodal (log-concave). A common question in the study of graph polynomials is whether or not a given polynomial is unimodal or log-concave. For example, a long-standing conjecture that the sequence of absolute values of the coefficients of the chromatic polynomial is log-concave has recently been proven in [41]. The following result that gives a sufficient condition for a real polynomial with positive coefficients to have log-concave polynomial is due to Newton (see [60]).

Theorem 5.3.1. *Let*

$$f(x) = \sum_{k=0}^{n} a_k x^k$$

be a real polynomial with positive coefficients. If f has all real roots then

$$\frac{a_k^2}{\binom{n}{k}^2} \geq \frac{a_{k-1}}{\binom{n}{k-1}} \frac{a_{k+1}}{\binom{n}{k+1}} \text{ for all } k \in \{1, \ldots, n-1\}.$$

Since $\binom{n}{k}^2 \geq \binom{n}{k-1}\binom{n}{k+1}$, this implies that $a_k^2 \geq a_{k-1}a_{k+1}$ for all $k \in \{1, \ldots, n-1\}$, so that the sequence $a_0, \ldots, a_n$ is log-concave, and hence unimodal as all terms are positive.

Theorem 5.3.1 has been used to demonstrate the log-concavity of certain graph polynomials. For example, it was proven in [26] that independence polynomials of claw-free graphs have all real roots and hence are log-concave.

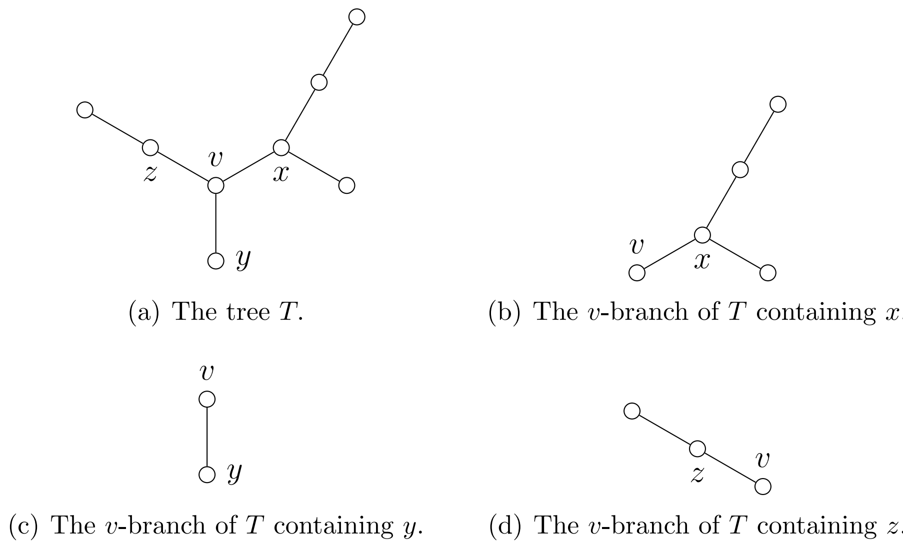
In this section we characterize the trees that have unimodal connected set polynomial and the trees that have log-concave connected set polynomial. Unfortunately, Theorem 5.3.1 will not be useful for us – Theorem 4.2.17 implies that no tree of order 3 or more has all real roots. We approach the problem directly instead.

We will see that most trees have neither unimodal nor log-concave connected set polynomial. Consider a tree T on n vertices with connected set polynomial

$$C(T;x) = \sum_{k=1}^{n} c_k x^k.$$

Since $c_1 = n > n-1 = c_2$ by Observation 3.1.1, our technique for demonstrating that a tree T has a non-unimodal subtree polynomial is to show that either c_3 or c_4 is at least n. As we will shortly see, almost all trees satisfy either $c_3 \geq n$ or $c_4 \geq n$.

Once again, the path to our main result requires plenty of building blocks. The following definition makes several proofs in this section much easier to write.



(a) The tree T.

(b) The v-branch of T containing x.

(c) The v-branch of T containing y.

(d) The v-branch of T containing z.

Figure 5.2: A tree T and its v-branches.

Definition 5.3.1. Let T be a tree and let v be a vertex of T. A *v-branch of T* is a maximal subtree of T having v as a leaf. The set of v-branches of T is denoted $\mathcal{B}(T, v)$. The v-brances of a particular tree are illustrated in Figure 5.2.

Observation 5.3.2. *Let T be a tree. If $v \in V(T)$ then T has $\deg(v)$ distinct v-branches.*

Proof. The vertex v is incident to $\deg(v)$ distinct edges. Each v-branch of T contains exactly one of these edges. Further, the maximality condition in the definition of v-branch implies that choosing an edge incident to v completely determines a v-branch. $\square$

We now move on to some results that will narrow down our search for trees with unimodal connected set polynomial greatly. Essentially, we find a lower bound on the coefficient c_3 of the cubic term in the connected set polynomial of a tree in terms of the order of the tree and the degree of any particular vertex.

Lemma 5.3.3. *Let T be a tree on n vertices. If T has a vertex v of degree $d \geq 3$, then $c_3(T) \geq n + \binom{d}{2} - d - 1$.*

Proof. Each subtree of T of order 3 either has central vertex v or is contained entirely in some v-branch of T. There are $\binom{d}{2}$ subtrees of T of order 3 with central vertex v. Each v-branch $B \in \mathcal{B}(T, v)$ contains exactly $c_3(B)$ subtrees of order 3. Therefore,

$$c_3(T) = \binom{d}{2} + \sum_{B \in \mathcal{B}(T,v)} c_3(B)$$

For each $B \in \mathcal{B}(T, v)$, if $V(B) \geq 3$ we have

$$c_3(B) \geq |V(B)| - 3 + 1 = |V(B)| - 2$$

by Theorem 5.2.1. If $V(B) \leq 2$ then $c_3(B) = 0$ and $c_3(B) \geq |V(B)| - 2$ still holds. By Observation 5.3.2 we know that T must have d distinct v-branches. Therefore,

$$\begin{aligned}
c_3(T) &\geq \binom{d}{2} + \sum_{B \in \mathcal{B}(T,v)} (|V(B)| - 2) \\
&= \binom{d}{2} + \sum_{B \in \mathcal{B}(T,v)} |V(B)| - 2d \\
&= \binom{d}{2} + (n + d - 1) - 2d \\
&= n + \binom{d}{2} - d - 1,
\end{aligned}$$

which completes the proof. □

Lemma 5.3.3 has several fairly straight forward corollaries – the lower bound we find for c_3 is at least n for every tree with a vertex of degree 4 or more, and every tree with two or more vertices of degree 3.

Corollary 5.3.4. *Let T be a tree on n vertices. If T has four or more leaves, then $C(T; x)$ is not unimodal.*

Proof. We first demonstrate that if T has four or more leaves then T either has a vertex of degree four or more, or at least two vertices of degree 3. Let the degree sequence of t be

$$(d_1, d_2, \ldots, d_{n-4}, 1, 1, 1, 1).$$

By the handshaking lemma (which says that the sum of the degrees of all vertices in a graph is equal to twice the number of edges) we know that

$$\sum_{k=1}^{n-4} d_k = 2(n-1) - 4 = 2(n-4) + 2,$$

and this $2(n-4)+2$ must be distributed among $n-4$ locations in the degree sequence. By the generalized pigeonhole principle, either some degree is at least 4 or at least two degrees are 3.

First suppose that T has a vertex v of degree 4. By Lemma 5.3.3, we have

$$c_3(T) \geq n + \binom{d}{2} - d - 1.$$

Since $d \geq 4$ we have $\binom{d}{2} - d = \frac{d(d-3)}{2} \geq 2$ and thus

$$c_3(T) \geq n + 1.$$

Therefore, $c_3(T) > c_1(T) > c_2(T)$ and we conclude that $C(T; x)$ is not unimodal.

Suppose instead that T has at least two vertices of degree 3. Let u and v be vertices of T of degree 3. Let B_1, B_2, and B_3 be the u-branches of T and without loss of generality let $v \in V(B_1)$. Then

$$c_3(T) = 3 + c_3(B_1) + c_3(B_2) + c_3(B_3)$$

by the same argument as in the proof of Lemma 5.3.3. Applying Lemma 5.3.3 to B_1

at vertex v yields

$$c_3(B_1) \geq |V(B_1)| - 1,$$

and the lower bound of Theorem 5.2.1 applied to B_i for $i \in \{2,3\}$ gives

$$c_3(B_i) \geq |V(B_i)| - 2$$

for $i \in \{2,3\}$. Thus we have

$$\begin{aligned} c_3(T) &\geq 3 + (|V(B_1)| - 1) + (|V(B_2)| - 2) + (|V(B_3)| - 2) \\ &= (|V(B_1)| + |V(B_2)| + |V(B_3)|) - 2 \\ &= n \end{aligned}$$

Therefore, $c_3(T) \geq c_1(T) > c_2(T)$ and we conclude that $C(T;x)$ is not unimodal. $\square$

Corollary 5.3.4 gives a very restrictive necessary condition for a tree to have a unimodal connected set polynomial – it says that any tree with a unimodal connected set polynomial has at most 3 leaves. Thus, we can tell already that very few trees have a unimodal connected set polynomial. From here, our approach involves computing the connected set polynomials of certain trees with 3 leaves and verifying their unimodality directly. The following general result will be needed.

Theorem 5.3.5. *Let T be a tree with vertex v. Let $u_1, u_2, \ldots, u_{\deg(v)}$ be the neighbours of v. Then*

$$C_v(T;x) = x \cdot \prod_{i=1}^{\deg(v)} [1 + C_{u_i}(B_i - v; x)],$$

where B_i is the v-branch of T containing u_i.

Proof. We proceed by induction on $\deg(v)$. In the base case, suppose $\deg(v) = 1$. We know that T has a unique connected set of order 1 containing v. Consider the set of connected subgraphs of T of order $k \geq 2$ containing v. Removing v from each subgraph in this set gives the set of connected subgraphs of $B - v$ of order $k - 1$ containing u, the unique neighbour of v. Further, this operation of removing v is clearly a one-to-one correspondence. Thus we have

$$C_v(T;x) = x + x \cdot C_u(T - v;x) = x \cdot [1 + C_u(B - v;x)],$$

which matches the theorem statement for $\deg(v) = 1$.

Now suppose that for some $d \geq 1$, the statement holds. Let T be a tree with vertex v of degree $d+1$. Let $u_1, \ldots, u_{d+1}$ be the neighbours of v and let $B_1, \ldots, B_{d+1}$ be the corresponding v-branches of T. By the inductive hypothesis, the tree $T' = T - V(B_{d+1} - v)$ satisfies

$$C_v(T';x) = x \prod_{i=1}^{d} [1 + C_{u_i}(B_i - v;x)].$$

The connected sets of T' containing v are also connected sets of T containing v, and all connected sets of T containing v but not u_{d+1} are contained in T'. Further, the union of any connected set of $B_{d+1} - v$ of order k_1 containing the vertex u_{d+1} and any connected set of T' of order k_2 containing v is a connected set of T of order $k_1 + k_2$ containing v. Furthermore, all connected sets of T containing both v and u_{d+1} arise in this way. Therefore,

$$\begin{aligned} C_v(T;x) &= C_v(T';x) \cdot \left[1 + C_{u_{d+1}}(B_{d+1} - v;x)\right] \\ &= x \cdot \prod_{i=1}^{d+1} [1 + C_{u_i}(B_i - v;x)] \end{aligned}$$

By the principle of mathematical induction, the statement holds. $\square$

We are now ready to prove the first main result of this section, which characterizes the trees having unimodal connected set polynomial.

Theorem 5.3.6. *Let T be a tree of order n. The connected set polynomial $C(T;x)$ is unimodal if and only if T has an induced P_{n-1}.*

Proof. ($\Leftarrow$) Suppose T has an induced P_{n-1}. Then either $T \cong P_n$ or $T \cong P_{n,k}$ for some $k \in \{2, \ldots, n-2\}$, where $P_{n,k}$ denotes the tree on vertices $v_1, \ldots, v_n$ with edges

$$\{v_1, v_2\}, \{v_2, v_3\}, \ldots, \{v_{n-2}, v_{n-1}\}, \{v_k, v_n\};$$

i.e. the $n-1$ vertices $v_1, \ldots, v_{n-1}$ make up an ordered path and v_n has unique neighbour v_k in the interior of this path. The graph $P_{n,k}$ is illustrated in Figure 5.3.

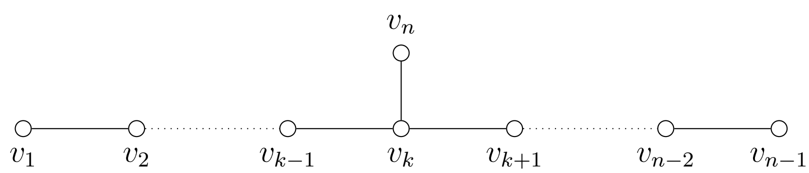


Figure 5.3: The tree $P_{n,k}$.

Suppose first that $T \cong P_n$. We saw in Section 4.2.3 that

$$C(P_n; x) = \sum_{k=1}^{n} (n+1-k)x^k,$$

which is clearly unimodal (in fact, the sequence of coefficients is monotone decreasing).

Suppose instead that $T \cong P_{n,k}$ for some $k \in \{2, \ldots, n-2\}$. We demonstrate that $C(P_{n,k}; x)$ is unimodal by finding an explicit formula for $C(P_{n,k}; x)$. By Lemma 4.2.15, we have

$$C(P_{n,k}; x) = C(P_{n-1}; x) + C_{v_n}(P_{n,k}; x). \tag{5.1}$$

By Theorem 5.3.5, we have

$$C_{v_n}(P_{n,k};x) = x \cdot [1 + C_{v_k}(P_{n-1};x)], \tag{5.2}$$

and applying Theorem 5.3.5 once more, we have

$$\begin{aligned} C_{v_k}(P_{n-1};x) &= x \cdot [1 + C_u(P_{k-1};x)] \cdot [1 + C_w(P_{n-1-k};x)] && (5.3)\\ &= x \cdot [1 + x + \ldots + x^{k-1}] \cdot [1 + x + \ldots + x^{n-k-1}], && (5.4) \end{aligned}$$

where u is a leaf of P_{k-1} and w is a leaf of P_{n-1-k}. Finally, we substitute (5.4) into (5.2) to obtain

$$C_{v_n}(P_{n,k};x) = x + x^2 \cdot [1 + x + \ldots + x^{k-1}] \cdot [1 + x + \ldots + x^{n-k-1}]. \tag{5.5}$$

Now we wish to express $C_{v_n}(P_{n,k};x)$ in such a way that the coefficient of each power is easier to read off. Continuing from (5.5) and assuming without loss of generality that $k - 1 \leq n - 1 - k$, we have

$$\begin{aligned} C_{v_n}(P_{n,k};x) &= x + x^2 \cdot \left[\sum_{i=0}^{k-1} (i+1)x^i + \sum_{i=k}^{n-k-2} kx^i + \sum_{i=n-k-1}^{n-2} (n-1-i)x^i \right] \\ &= x + \sum_{i=2}^{k+1} (i-1)x^i + \sum_{i=k+2}^{n-k} kx^i + \sum_{i=n-k+1}^{n} (n+1-i)x^i. \end{aligned}$$

Subsituting this expression for $C_{v_n}(P_{n,k};x)$ and the known formula for $C(P_{n-1};x)$ into (5.1) gives

$$\begin{aligned} C(P_{n,k};x) &= C(P_{n-1};x) + C_{v_n}(P_{n,k};x) \\ &= \sum_{i=1}^{n-1} (n-i)x^i + x + \sum_{i=2}^{k+1} (i-1)x^i + \sum_{i=k+2}^{n-k} kx^i + \sum_{i=n-k+1}^{n} (n+1-i)x^i, \end{aligned}$$

and regrouping yields

$$C(P_{n,k};x) = nx + \sum_{i=2}^{k+1}(n-1)x^i + \sum_{i=k+2}^{n-k}(n+k-i)x^i + \sum_{i=n-k+1}^{n}(2(n-i)+1)x^i.$$

Since

$$\sum_{i=k+2}^{n-k}(n+k-i)x^i = (n-2)x^{k+2} + (n-3)x^{k+3} + \ldots + (2k)x^{n-k}$$

and

$$\sum_{i=n-k+1}^{n}(2(n-i)+1)x^i = (2k-1)x^{n-k+1} + (2k-3)x^{n-k+2} + \ldots + x^n,$$

we can see directly that the sequence of coefficients of $C(P_{n,k};x)$ is nonincreasing, and hence $C(P_{n,k};x)$ is unimodal and we are done with this direction.

($\Rightarrow$) Suppose T has no induced P_{n-1}. Then $n \geq 5$ (every tree on $n \leq 4$ vertices has an induced P_{n-1}) and T has at least three leaves (if T has less than 3 leaves then $T \cong P_n$). By Corollary 5.3.4, if T has four or more leaves then it has non-unimodal connected set polynomial, so we may assume that T has exactly three leaves. In this case, T has only one vertex v of degree greater than 2, $\deg(v) = 3$, and v is not adjacent to a leaf (otherwise T would have an induced P_{n-1}). This means that the three v-branches of T are all paths on at least 3 vertices each. Let the v-branches of T be called B_1, B_2, and B_3. A sketch of T is given in Figure 5.4.

Each subtree of T of order 4 either has v as a central vertex or is contained entirely in a single v-branch of T. By inspection, there are 7 subtrees of T of order 4 which contain v as a central vertex (one where v has degree 3, and six where v has degree 2). Recall that $c_4(P_n) = n-3$ for $n \geq 3$ and that B_1, B_2, and B_3 are paths on 3 or

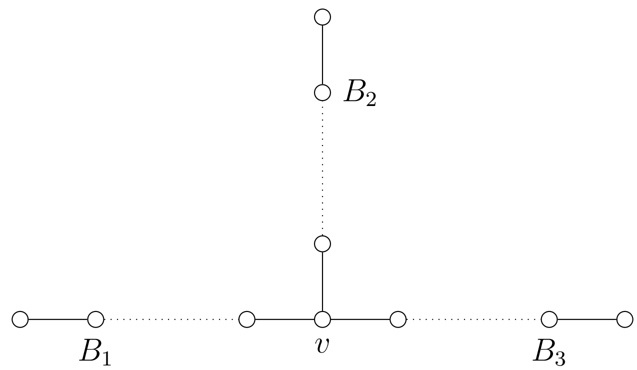


Figure 5.4: A sketch of T.

more vertices. Therefore,

$$\begin{aligned} c_4(T) &= 7 + c_4(B_1) + c_4(B_2) + c_4(B_3) \\ &= 7 + (|V(B_1)| - 3) + (|V(B_2)| - 3) + (|V(B_3)| - 3) \\ &= (|V(B_1)| + |V(B_2)| + |V(B_3)|) - 2 \\ &= n \end{aligned}$$

Thus $c_4(T) = c_1(T) > c_2(T)$ and $C(T; x)$ is not unimodal. $\square$

The number of nonisomorphic trees of order n containing an induced P_{n-1} is exactly $\lceil \frac{n-1}{2} \rceil$. It is well known that the total number of nonisomorphic trees on n vertices grows exponentially (see [27], for example). Therefore, almost all (unlabelled) trees have non-unimodal subtree polynomial.

We conclude this section with a characterization of the trees with log-concave connected set polynomial. Since log-concavity of a sequence of positive numbers implies unimodality, we know by Theorem 5.3.6 that any tree T with log-concave connected set polynomial must be isomorphic to either P_n or $P_{n,k}$. We have formulae for both $C(P_n; x)$ and $C(P_{n,k}; x)$, so characterizing the trees with log-concave connected set polynomial is very straightforward.

Theorem 5.3.7. *Let T be a tree of order n. The connected set polynomial $C(T;x)$ is log-concave if and only if $T = P_n$.*

Proof. We know that log-concavity of a polynomial with positive coefficients implies unimodality. Thus by Theorem 5.3.6, any tree of order n that has log-concave connected set polynomial must be isomorphic to either P_n or $P_{n,k}$ for some $k \in \{2, \ldots, n-2\}$.

First of all, $c_i(P_n) = n - i + 1$ for $i \in \{2, \ldots, n-1\}$, and hence

$$c_i(P_n)^2 = (n-i+1)^2 > (n-i)(n-i+2) = c_{i-1}(P_n)c_{i+1}(P_n)$$

for $i \in \{2, \ldots, n-1\}$. Thus we have shown directly that $C(P_n;x)$ is strictly log-concave.

On the other hand, for any $k \in \{2, \ldots, n-2\}$ we have

$$c_2(P_{n,k})^2 = (n-1)^2 < n(n-1) = c_1(P_{n,k})c_3(P_{n,k}),$$

and thus $C(P_{n,k};x)$ is not log-concave.

We conclude that for a tree T of order n, $C(T;x)$ is unimodal if and only if $T \cong P_n$. $\square$

While we have shown that there are very few trees that have unimodal connected set polynomial, this is largely due to the fact that $c_1 = n$ and $c_2 = n-1$ for all trees, and hence the sequence of coefficients must in fact be nonincreasing (a stronger condition than unimodality). Intuitively, it seems much more likely that the connected set polynomial of a graph will be unimodal if it has at least as many edges as it has vertices, so that $c_2 \geq c_1$. In this situation, the sequence of coefficients of the connected set polynomial may first rise and then fall.

We have verified that every connected graph on $n \leq 8$ vertices and $m \geq n$ edges has unimodal connected set polynomial. On the other hand, while all disconnected graphs on $n \leq 7$ vertices and $m \geq n$ edges have unimodal connected set polynomial, there are 4 distinct disconnected graphs on 8 vertices and at least 8 edges whose connected set polynomials are not unimodal.

As for log-concavity, there are many graphs of small order with at least as many edges as vertices whose connected set polynomials are not log-concave. For example, there are 50 connected graphs (and 5 disconnected graphs) on 6 vertices and at least 6 edges whose connected set polynomials are not log-concave.

5.4 Connected Set Roots of Trees

In this section we continue the study of connected set roots begun in Section 4.2, but here we focus exclusively on trees. In general, if z is a nonzero connected set root of a connected graph G on n vertices then

$$\tfrac{2}{n-1} \leq |z| \leq n$$

by Theorem 4.2.6. However, plotting the connected set roots of all trees of a given small order makes it seem as though this bound may be far from tight for trees. For example, the nonzero connected set roots of all trees on 10 vertices are pictured in Figure 5.5.

While we demonstrated in Proposition 4.2.9 and Proposition 4.2.8 that the cycle C_n has a connected set root close to $-(n-1)$ for $n \geq 3$ and that the disconnected graph $K_2 \cup O_{n-2}$ has a connected set root at $-n$ for $n \geq 2$, respectively, there are no trees of order 10 with a connected set root of modulus anywhere close to 10. In fact, we have numerically solved for the connected set roots of all trees of order at most 12, and the connected set root of largest modulus among all of those trees belongs to

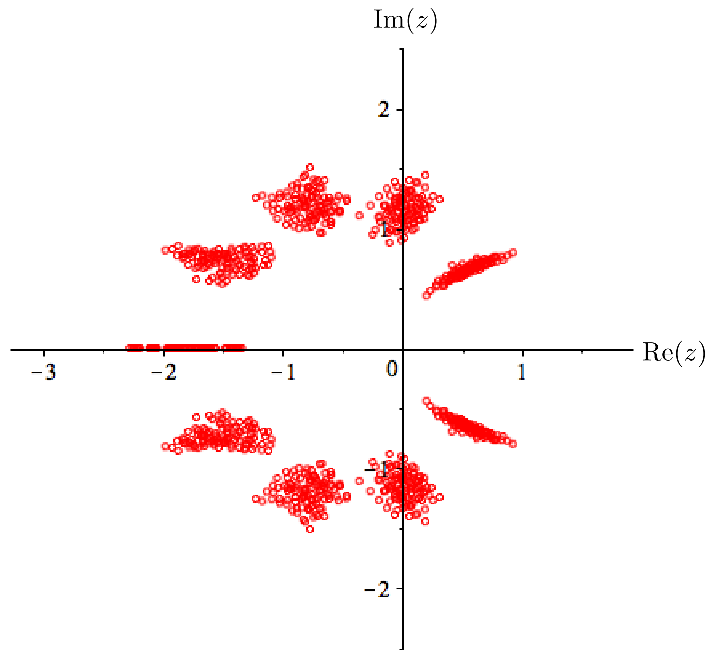


Figure 5.5: The nonzero connected set roots of all trees on 10 vertices.

the star on only 4 vertices! We have

$$C(K_{1,3};x) = x^4 + 3x^3 + 3x^2 + 4x,$$

so we can solve for the root in question exactly – it is

$$-1 - \sqrt[3]{3} \approx -2.44225.$$

If we restrict to trees, we can improve on the general upper bound on the modulus of a connected set root found in Theorem 4.2.6. We will show that the modulus of a connected set root of a tree T cannot exceed the number of leaves of T. While this bound still seems far from best possible based on our computational evidence, it is a step in the right direction. The following lemma will be used in our proof of the improved upper bound on the modulus of a connected set root of a tree. We rely on the fact that connected sets of a tree T are in one-to-one correspondence with the subtrees of T.

Lemma 5.4.1. *Let S be a subtree of T. If S has s leaves and T has t leaves then $s \leq t$.*

Proof. We can obtain S from T by recursively deleting leaf vertices not contained in S. Therefore, it suffices to show that deleting a leaf from a tree cannot increase the number of leaves. Suppose R is a tree with r leaves and that v is a leaf vertex of R. Let u be the unique neighbour of v. The number of leaves in $R - v$ is given by

$$\begin{cases} r-1 & \text{if } \deg(u) \geq 3 \\ r & \text{if } \deg(u) = 2 \\ 0 & \text{if } \deg(u) = 1 \end{cases}$$

and therefore $R - v$ has at most r leaves. □

Theorem 5.4.2. *Suppose T is a tree on $n \geq 2$ vertices with r leaves. If z is a root of $C(T;x)$ then $\frac{2}{n-1} \leq |z| \leq r$.*

Proof. Let T be a tree of order $n \geq 2$, and let

$$C(T;x) = \sum_{k=0}^{n} c_k x^k.$$

For the upper bound it suffices to show that $\frac{c_k}{c_{k+1}} \leq r$ for each $k \in \{1, \ldots, n-1\}$ by the Eneström-Kakeya Theorem. Let T_k be the set of subtrees of order k for each $k \in \{1, \ldots, n\}$. For each $1 \leq k \leq n-1$, every member of T_k is the result of deleting some leaf from a member of T_{k+1}. Thus the number of subtrees of order k is at most the sum of the number of leaves over all members of T_{k+1}. We have c_{k+1} subtrees in T_{k+1} and each such subtree has at most r leaves by Lemma 5.4.1. Therefore,

$$c_k \leq r c_{k+1},$$

as desired. The lower bound in the theorem statement was proven in general for connected set polynomials of connected graphs in Theorem 4.2.6, so we are done. □

We can indeed demonstrate that the bound of Theorem 5.4.2 is far from tight for at least one particular choice of r. We show below that the modulus of a root of $C(K_{1,n-1};x)$ cannot exceed $1+\sqrt[3]{3}$ for any $n \in \mathbb{N}$. This means that there is no tree on n vertices with $n-1$ leaves having a subtree root close to $n-1$ in modulus, the upper bound given by Theorem 5.4.2.

Proposition 5.4.3. *Let $n \geq 2$. If z is a root of $C(K_{1,n-1};x)$ then $|z| \leq 1+\sqrt[3]{3}$.*

Proof. We demonstrate the stronger result that if z is a root of $C(K_{1,n-1};x)$ then $|z+1| \leq \sqrt[3]{3}$. Recall that

$$C(K_{1,n-1};x) = x(x+1)^{n-1} + (n-1)x = x\big[(x+1)^{n-1} + (n-1)\big],$$

and hence z is a root of $C(K_{1,n-1};x)$ if and only if either $z=0$ or $z+1$ is an $(n-1)^{\text{st}}$ root of $-(n-1)$. The modulus of any $(n-1)^{\text{st}}$ root of $-(n-1)$ is given by

$$f(n) = (n-1)^{\frac{1}{n-1}}$$

Considering f as a function of a real variable $x>1$, a straightforward computation yields

$$\tfrac{f'(x)}{f(x)} = \tfrac{1}{(x-1)^2}[1-\ln(x-1)].$$

Since $f(x)>0$ for $x>1$, we see that $f'(x)>0$ when $x<e+1$ and $f'(x)<0$ when $x>e+1$. Therefore, the maximum of $f(n)$ for $n \geq 2$ falls at either $n=3$ or $n=4$. By inspection, we find that the maximum is $f(4)=\sqrt[3]{3}$.

We conclude that any root z of $C(K_{1,n-1};x)$ for $n \geq 2$ satisfies $|z+1| \leq \sqrt[3]{3}$, and hence $|z| \leq 1+\sqrt[3]{3}$. □

The following is a straightforward consequence of Theorem 5.4.2 and Proposition 5.4.3. It gives an upper bound on the modulus of any connected set root of a tree, but appears to be far from tight.

Corollary 5.4.4. *Let T be a tree of order $n \geq 5$. If z is a root of $C(T;x)$ then $|z| \leq n-2$.*

Proof. By Proposition 5.4.3, if $T = K_{1,n-1}$ then any root z of $C(T;x)$ satisfies $|z| < 1 + \sqrt[3]{3} \leq n-2$ as $n \geq 5$. Otherwise, T has at most $n-2$ leaves and if z is a root of $C(T;x)$ then $|z| \leq n-2$ by Theorem 5.4.2. $\square$

While we do not have a proof that there is a constant bound on the modulus of any connected set root of a tree T, we can prove that the roots of the *rooted* connected set polynomial have modulus at most 2.

Theorem 5.4.5. *Let T be a tree with vertex v. Any root z of the rooted connected set polynomial $C_v(T;z)$ satisfies $|z| \leq 2$.*

Proof. For any graph G, let $\widetilde{C}_v(G;z) = 1 + C_v(G;z)$. We will first demonstrate that for any tree T, we have $\left|\widetilde{C}_v(T;z)\right| \geq |z| - 1$ for any z satisfying $|z| > 2$. We proceed by induction on the order of the tree.

In the base case, the tree K_1 has

$$\left|\widetilde{C}_v(K_1;z)\right| = |z+1| \geq |z| - 1$$

for all $z \in \mathbb{C}$, but for $|z| > 2$ in particular. Now suppose that for some $n \geq 2$, any tree S of order strictly less than n satisfies

$$\left|\widetilde{C}_v(S;z)\right| \geq |z| - 1$$

for $|z| > 2$. Let T be a tree of order n, and let v be a vertex of T. Let $\deg(v) = d$ and let the neighbours of v be $u_1, \ldots, u_d$. By Theorem 5.3.5,

$$\widetilde{C}_v(T;z) = 1 + z \cdot \prod_{i=1}^{d} \left[\widetilde{C}_{u_i}(B_i - v; z)\right], \tag{5.6}$$

where B_i is the v-branch of T containing u_i. Applying the reverse triangle inequality and then the induction hypothesis to (5.6) yields

$$\begin{aligned} \left|\widetilde{C}_v(T;z)\right| &\geq |z| \cdot \prod_{i=1}^{d} \left|\widetilde{C}_{u_i}(B_i - v; z)\right| - 1 \\ &\geq |z|(|z|-1)^d - 1. \end{aligned}$$

Finally, by the hypothesis that $|z| > 2$ we have $(|z|-1)^d > 1$, which gives

$$\left|\widetilde{C}_v(T;z)\right| > |z| - 1.$$

We have shown that any tree T satisfies $\left|\widetilde{C}_v(T;z)\right| \geq |z| - 1$ for any z with $|z| > 2$. Since $C_v(T;z) = \widetilde{C}_v(T;z) - 1$ by definition, we conclude that

$$|C_v(T;z)| \geq \left|\widetilde{C}_v(T;z)\right| - 1 \geq (|z|-1) - 1 > 0$$

for $|z| > 2$. □

We don't see any immediate consequence of Theorem 5.4.5 that gives a constant bound on the moduli of connected set roots of trees (i.e. the roots of the connected set polynomial $C(T;x)$ as opposed of the roots of the *rooted* connected set polynomial $C_v(T;x)$). However, the hope is that Theorem 5.4.5 may eventually aid in the proof of such a bound on the connected set roots of trees.

We conclude this section with a very interesting observation. Among all trees on n vertices for $n \in \{3, \ldots, 12\}$, we have found that the star $K_{1,n-1}$ has the connected

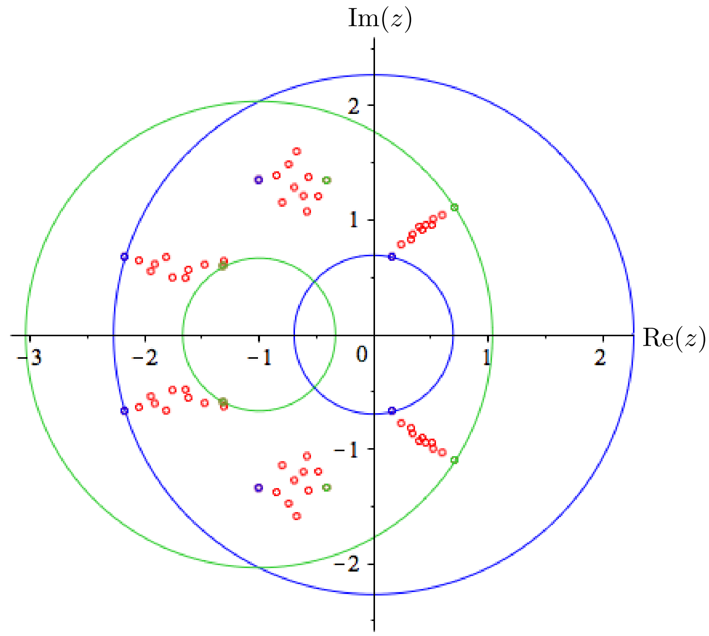


Figure 5.6: The nonzero connected set roots of all trees on 7 vertices. The connected set roots of $K_{1,n-1}$ are shown in blue while the connected set roots of P_n are shown in green. All other roots are shown in red.

set root of greatest modulus *and* the connected set root of smallest modulus. On the other hand, the path P_n has the connected set root at greatest distance from -1 *and* the connected set root at smallest distance from -1.

The nonzero connected set roots of all trees on 7 vertices are shown in Figure 5.6. In particular, the connected set roots of $K_{1,6}$ are shown in blue while the connected set roots of P_7 are shown in green. All other roots are shown in red. We have also included the circles (blue) centred at 0 of largest and smallest radius that intersect such a connected set root (they intersect the roots of the star $K_{1,6}$) and the circles (green) centred at -1 of largest and smallest radius that intersect such a connected set root (they intersect the connected set roots of the path P_7).

Recall that $K_{1,n-1}$ is the tree on n vertices with the most subtrees and P_n is the tree on n vertices with the least subtrees (in fact, they have the coefficient-wise largest and smallest subtree polynomials among all trees on n vertices, by Theorem 5.2.1). We can't help but wonder if the location of the roots of subtree polynomials is somehow related to the total number of subtrees. This would be interesting as a strong correlation between the number of subtrees of a tree and the Wiener index has been confirmed in [69]. In any case, it appears that the connected set roots of trees are worthy of more study.

Chapter 6

Conclusion

This thesis was concerned primarily with studying the analytic properties of two different reliability functions, namely all-terminal reliability and node reliability. We introduced the connected set polynomial which allowed us to achieve certain results on the computational complexity and the roots of node reliability more easily, and our results on the connected set polynomial are of interest in their own right. In particular, the connected set polynomials of trees provide a means for extending some recent results concerning the tree in a given class with the most (or least) total number of subtrees.

6.1 All-Terminal Reliability

We have proven a nonconstant bound on the modulus of any ATR root, and we have found ATR roots with larger modulus than any previously known (though they are still relatively small in modulus). We found the ATR roots of the graphs $G^{1,6}_{n,n}$ for $n \leq 12$, producing ATR roots of modulus greater than 1.11. We suspect that for $n \geq 13$ the graphs $G^{1,6}_{n,n}$ will have ATR roots of even larger modulus. What is the limiting behaviour of the modulus of the ATR root of largest modulus of the graph $G^{1,6}_{n,n}$? We are not even certain that the sequence of moduli need be increasing.

It still seems as though all-terminal reliability roots are bounded in modulus by some constant. We note, however, that data from small graphs and special families of graphs does not necessarily rule the day. For example, it was originally conjectured by Farrell that chromatic roots lie in the right half-plane [33], and this was proven false by a slim margin by Read and Royle in [53] before Sokal finally proved that chromatic roots are in fact dense in the entire complex plane [58]! The question of whether ATR roots are bounded in modulus by some constant is a tantalizing open problem.

Our study of simple graphs with ATR roots outside of the unit disk produced a smaller example than any previously known, although it is still rather large (it has 546 vertices and 1080 edges). What is the smallest simple graph with ATR roots outside of the unit disk?

In addition to finding a smaller simple graph with ATR roots outside of the unit disk, we found simple graphs with edge connectivity as high as 5 that have ATR roots outside of the unit disk. This is notable as all previously known examples have edge connectivity 2, with many vertices of degree 2. We also have good candidates for simple graphs of even higher edge connectivity that have ATR roots outside of the unit disk, although the computations required to prove that the roots are outside of the unit disk become increasingly large. Are there simple graphs with arbitrarily high edge connectivity that have ATR roots outside of the unit disk?

Finally, while we produced examples of simple graphs that have ATR roots outside of the unit disk and have *edge* connectivity higher than 2, the *vertex* connectivity of all of our examples is still only 2. Are there simple graphs with vertex connectivity greater than 2 that have ATR roots outside of the unit disk? Every graph formed from an edge substitution by a gadget on 3 or more vertices has vertex connectivity at most 2, so the theory we developed in Section 2.2.2 does not lend itself well to solving this problem.

6.2 Split Reliability

All-terminal reliability and $\{u, v\}$-split reliability are simultaneously generalized by the following notion: Let G be a graph in which each edge fails independently with probability q and let $K \subseteq V(G)$. The K*-split reliability* of G, denoted $\mathrm{spRel}_K(G;q)$, is the probability that every vertex of G can communicate with *exactly* one vertex from K. When $|K| = 1$ this is all-terminal reliability and when $|K| = 2$ this is $\{u, v\}$-split reliability.

The $\{u, v\}$-split reliability has proven itself useful in the study of all-terminal reliability, but we believe that K-split reliability could have several applications outside of all-terminal reliability as well and is worthy of study in its own right. For example, consider a network with a set K of *leader nodes* which give orders or instructions, where we would like all of the other vertices to receive orders from exactly one of the leader nodes. The condition for K-split reliability ensures that every node receives orders from exactly one leader node, and thus conflicting orders cannot be given. Chain of command structures seem to be an obvious application of this concept.

We also note that K-split reliability gives a measure of the reliability of disconnected graphs. Let G be a graph with components $G_1, \ldots, G_k$ and let $K = \{v_1, \ldots, v_k\}$ where $v_i \in V(G_i)$ for $i \in \{1, \ldots, k\}$. Then

$$\mathrm{spRel}_K(G;q) = \prod_{i=1}^{k} \mathrm{Rel}(G_i;q).$$

This seems to be a natural measure of the reliability of a disconnected network with edge failures.

6.3 Node Reliability

What is striking about node reliability is that on the surface its definition is analogous to that of other well-known forms of reliability (such as all-terminal, two-terminal and K-terminal), but its shape and analytic properties can be so different. The frequent lack of monotonicity, the contrasting concavity near 0, the frequency of points of inflection, the multiplicity of fixed points, and the nature and location of the roots all illustrate that node reliability is quite different from the other models of probabilistic robustness on graphs (or even coherent systems), and merits further attention.

In Section 3.1 we proved that the node reliability of any graph of order n and size $m \leq 0.0851n^2$ is decreasing at the point $\frac{\hat{r}}{n}$, where $\hat{r} \approx 1.729474372$. Could it be true that the node reliabilities of almost all graphs of order n have an interval of decrease in $(0,1)$? We have also found examples of disconnected graphs with two disjoint maximal intervals of decrease in $(0,1)$. Are there connected graphs with two such intervals of decrease?

For those graphs whose node reliability polynomials have an interval of decrease in $(0,1)$, a natural question to ask is how long the interval of decrease can be. For any $n \geq 2$, the node reliability polynomial of the empty graph O_n on n vertices is given by $\mathrm{nRel}(O_n; p) = p(1-p)^{n-1}$ which can easily be seen to be decreasing on the interval $(\frac{1}{n}, 1)$. This means that the interval of decrease can have length arbitrarily close to 1 for disconnected graphs, but for connected graphs we conjecture that the length is at most $\frac{1}{2}$. We can demonstrate that the length of the interval of decrease can be arbitrarily close to $\frac{1}{2}$ for connected graphs, and we give a brief sketch of this result. Let f_n be the node reliability polynomial of $K_{n-1} \circ K_2$ for each $n \geq 2$ (recall that $K_{n-1} \circ K_2$ is the complete graph K_{n-1} with a single pendant edge added, also called a *vertex bonding* of K_{n-1} and K_2). From the expression for f_n' given in (3.18), we find that for $p \in \left(0, \frac{1}{2}\right)$,

$$f_n'(p) < g_n(p) = 2p - 1 + (n+1)(1-p)^{n-2}.$$

We find that

$$\lim_{n \to \infty} \left\{ g_n\left(\tfrac{1}{\ln n}\right) \right\} = -1$$

and

$$\lim_{n \to \infty} \left\{ n g_n\left(\tfrac{1}{2} - \tfrac{1}{n}\right) \right\} = -2$$

so that both $g_n\left(\frac{1}{\ln n}\right) < 0$ and $g_n\left(\frac{1}{2} - \frac{1}{n}\right) < 0$ for n sufficiently large. Also,

$$g_n'(p) = 2 - (n+1)(n-2)(1-p)^{n-3}$$

has a unique real root

$$q_n = 1 - \left(\frac{2}{(n+1)(n-2)}\right)^{1/(n-3)}.$$

As $\lim_{n\to\infty}\{q_n \ln n\} = 0$, it follows that for n sufficiently large, q_n is to the left of $\frac{1}{\ln n}$, and so g_n, and hence f_n', is negative on $\left(\frac{1}{\ln n}, \frac{1}{2} - \frac{1}{n}\right)$, which has length tending to 1/2.

In Section 3.2 we found a large family of graphs whose node reliabilities have 2 inflection points in $(0,1)$. We also found a finite number of graphs whose node reliabilities have 3 points of inflection in $(0,1)$. This leads to two open questions: Are there infinitely many graphs whose node reliabilities have 3 inflection points in $(0,1)$? Can the node reliability have arbitrarily many inflection points in $(0,1)$, as has been shown for all-terminal reliability [36]?

In Section 3.3 we found a large family of graphs whose node reliabilities have two fixed points in $(0,1)$. This is very different from the case for coherent reliability polynomials, which were shown in [6] to have at most one fixed point in $(0,1)$. Are there graphs whose node reliabilities have three or more fixed points in $(0,1)$? We have verified that the node reliability of any connected graph on at most 8 vertices has at most two fixed points in $(0,1)$.

Finally, in Section 3.4 and also in Chapter 4 we discovered much about the nature and location of the roots of node reliability. We first found that the real roots of node reliability can be arbitrarily large in modulus – this is very different from all-terminal reliability where the real roots are contained in $\{0\} \cup (1,2]$ (ignoring disconnected graphs for which the all-terminal reliability is identically zero). In Section 4.2.3 we proved that node reliability roots are dense in the entire complex plane, and note that this is not suspected to be true for all-terminal reliability.

6.4 The Connected Set Polynomial

In Section 4.1 we demonstrated that the problem of evaluating the connected set polynomial is #P-hard at any complex number $z \neq -1$ such that $z+1$ is not a complex root of unity. In turn, this tells us that the problem of evaluating the node reliability polynomial is #P-hard for any real number $p \notin \{0, 1, 2\}$. While $\mathrm{nRel}(G; 0) = 0$ for any graph G, and $\mathrm{nRel}(G; 1)$ is 1 if G is connected and 0 otherwise (and hence can be computed quickly), we do not know the complexity of evaluating the node reliability polynomial at 2. Can $\mathrm{nRel}(G; 2)$ be found in polynomial time? More generally, can we evaluate the connected set polynomial at z in polyomial time if $z + 1$ is a root of unity?

In Section 4.2 we studied the roots of connected set polynomials. We found that every connected graph of order at least 3 has a nonreal connected set root. We proved that any connected set root of a graph of order n has modulus at most n, and demonstrated that this bound is tight for disconnected graphs and nearly tight for connected graphs. Finally, we found that the closure of the collection of connected set roots is the entire complex plane.

This last result leads to another open question. What is the closure in $\mathbb{R}$ of the collection of all *real* connected set roots? By Theorem 4.2.17, we can find a connected set root arbitrarily close to any real number x, but not necessarily a *real* connected set root. Since the connected set polynomial has all nonnegative coefficients, the real connected set roots must be nonpositive. Is the collection of real connected set roots dense in $(-\infty, 0]$?

We demonstrate here that any rational number $x \leq -2$ is a connected set root of some graph. Let $x = -\frac{p}{q}$ where $p, q \in \mathbb{N}$ and assume that $x \leq -2$, so that $p \geq 2q$.

Consider the graph $qK_2 \cup O_{p-2q}$; that is, the disjoint union of q copies of K_2 and $p-2q$ isolated vertices. The connected set polynomial of this graph is given by

$$C(qK_2 \cup O_{p-2q}; x) = qx^2 + px = qx\Big(x + \tfrac{p}{q}\Big),$$

which has a root at $-\frac{p}{q}$.

From the above paragraph we can conclude that real connected set roots are at least dense in $(-\infty, -2]$, but we are unsure whether this density extends to the interval $(-2, 0)$. We have found many small graphs with connected set roots at -1 – a graph has a connected set root at -1 if and only if it has the same number of even and odd connected sets. We have also found many small graphs with a connected set root in $(-2, -1)$, and several small graphs with a connected set root in $(-1, 0)$. What is the closure of the real connected set roots in $(-2, 0)$?

6.5 The Subtree Polynomial

In Section 5.2 we proved that the path P_n and the star $K_{1,n-1}$ have the coefficient-wise least and greatest connected set polynomials, respectively, among all trees on n vertices. This generalizes a result of [63] that the path and star have the least and greatest total number of subtrees. It would be interesting to know if similar results on the tree in a given class with the least or greatest total number of subtrees can be generalized. For example, in [73], Yan and Yeh describe the tree on n vertices with diameter at least d which has the maximum number of subtrees, and the tree on n vertices with maximum degree at least Δ which has the minimum number of subtrees. Do these results extend to all coefficients of the connected set polynomial? Other results of this type can be found in [64, 76]. Any extensions of these results to all coefficients of the connected set polynomial would have immediate applications for node reliability, as the graph in a given class with the coefficient-wise greatest (least) connected set polynomial necessarily has the uniformly best (worst) node reliability among all graphs in the class.

In Section 5.3, we characterized the trees that have unimodal connected set polynomial and the trees that have log-concave connected set polynomial. The question appears to be more difficult for graphs with at least as many edges as vertices. Is the connected set polynomial of every connected graph on n vertices and $m \geq n$ edges unimodal? We have verified that this is true for $n \leq 8$.

Finally, in Section 5.4 we studied the roots of connected set polynomials of trees. While the collection of connected set roots of all graphs is dense in the complex plane, it appears that the connected set roots of trees are bounded in modulus by some constant. The root of largest known modulus belongs to the tree $K_{1,3}$ and has modulus $1 + \sqrt[3]{3} \approx 2.44225$. Is this the largest that the modulus of a connected set root of a tree can be? While we have proven that the roots of the *rooted* connected set polynomial $C_v(T; x)$ of a tree T are contained in the disk of radius 2 centred at the origin, the best general upper bound that we have found on the modulus of a root of the connected set polynomial $C(T; x)$ is given by the number of leaves of T.

Bibliography

[1] N. Anderson, E. B. Saff, and R. S. Varga. On the Eneström-Kakeya theorem and its sharpness. *Linear Algebra Appl.*, 28:5–16, 1979.

[2] J. L. Arocha and B. Llano. Mean value for the matching and dominating polynomial. *Discuss. Math. Graph Theory*, 20:57–69, 2000.

[3] S. Beraha, J. Kahane, and N. J. Weiss. Limits of zeros of recursively defined families of polynomials. In G. Rota, editor, *Studies in foundations and combinatorics*, volume 1 of *Advances in mathematics, Supplementary studies*, pages 213–232. Academic Press, New York, 1978.

[4] L. J. Billera. Polyhedral theory and commutative algebra. In M. Grötschel A. Bachem, B. Korte, editor, *Mathematical Programming: The State of the Art*, pages 57–77. Springer Berlin Heidelberg, 1983.

[5] G. D. Birkhoff. A determinant formula for the number of ways of coloring a map. *Annal. Math.*, 14(2):42–46, 1912.

[6] Z. W. Birnbaum, J. D. Esary, and S. C. Saunders. Multi-component systems and structures and their reliability. *Technometrics*, 3(1):55–77, 1961.

[7] A. Björklund, T. Husdelft, P. Kaski, and M. Koivisto. The traveling salesman problem in bounded degree graphs. *ACM Trans. Algorithms*, 8(2):Art. 18, 13pp, 2012.

[8] A. Björner, L. Lovász, and P. W. Shor. Homology and shellability of matroids and geometric lattices. In N. White, editor, *Matroid Applications*, Encyclopedia of Mathematics and its Applications, pages 226–283. Cambridge Univ. Press, 1992.

[9] A. Bohn. A dense set of chromatic roots which is closed under multiplication by positive integers. *Discrete Math.*, 321:45–52, 2014.

[10] B. Bollobás. *Graph theory: an introductory course*, volume 63 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1979.

[11] J. I. Brown and C. J. Colbourn. Network reliability. in Handbook of the Tutte Polynomial, forthcoming.

[12] J. I. Brown and C. J. Colbourn. Roots of the reliability polynomial. *SIAM J. Discrete Math.*, 5:571–585, 1992.

[13] J. I. Brown and C. J. Colbourn. On the log concavity of reliability and matroidal sequences. *Adv. in Appl. Math.*, 15(1):114–127, 1994.

[14] J. I. Brown, C. J. Colbourn, and R. J. Nowakowski. Cohen-Macaulay rings in network reliability. *SIAM J. Discrete Math.*, 9(3):377–392, 1996.

[15] J. I. Brown, C. J. Colbourn, and R. J. Nowakowski. Chip firing and all-terminal network reliability bounds. *Discrete Optim.*, 6:436–445, 2009.

[16] J. I. Brown and D. Cox. The closure of the set of roots of strongly connected reliability polynomials is the entire complex plane. *Discrete Math.*, 309(16):5043–5047, 2009.

[17] J. I. Brown and C. A. Hickman. On chromatic roots of large subdivisions of graphs. *Discrete Math.*, 242(1-3):17–30, 2002.

[18] J. I. Brown, C. A. Hickman, and R. J. Nowakowski. On the location of roots of independence polynomials. *J. Algebraic Combin.*, 19(3):273–282, 2004.

[19] J. I. Brown, C. A. Hickman, and R. J. Nowakowski. On the location of roots of independence polynomials. *J. Algebraic Combin.*, 19(3):273–282, 2004.

[20] J. I. Brown, Y. Koç, and R. E. Kooij. Inflection points for network reliability. *Telecommunication Systems*, 56(1):79–84, May 2014.

[21] J. I. Brown and R. J. Nowakowski. Bounding the roots of independence polynomials. *Ars. Combin.*, 58:113–120, 2001.

[22] J. I. Brown and O. R. Oellermann. Graphs with a minimal number of convex sets. *Graphs Combin.*, 30(6):1383–1397, 2013.

[23] J. I. Brown and O. R. Oellermann. On the spectrum and number of convex sets in graphs. *Discrete Math.*, 338(7):1144–1153, 2015.

[24] J. I. Brown and J. Tufts. On the roots of domination polynomials. *Graphs Combin.*, 30:527–547, 2014.

[25] Y.-B. Choe, J. G. Oxley, A. D. Sokal, and D. G. Wagner. Homogeneous multivariate polynomials with the half-plane property. *Adv. in Appl. Math.*, 32(1–2):88–187, 2004.

[26] M. Chudnovsky and P. Seymour. The roots of the independence polynomial of a clawfree graph. *J. Combin. Theory Ser. B*, 97(3):350–357, 2007.

[27] D. Cieslik. *Counting graphs: an introduction with specific interest in phylogeny*, volume 20 of *Biometrie und medizinische Informatik*. Shaker, 2012.

[28] C. J. Colbourn. *The Combinatorics of Network Reliability*. Oxford University Press, Inc., New York, NY, USA, 1987.

[29] C. J. Colbourn, A. Satyanarayana, C. Suffel, and K. Sutner. Computing residual connectedness reliability for restricted networks. *Discrete Appl. Math.*, 44:221–232, 1993.

[30] R. M. Corless, G. H. Gonnet, D. E. G. Hare, D. J. Jeffrey, and D. E. Knuth. On the Lambert W function. *Adv. Comput. Math.*, 5:329–359, 1996.

[31] A. A. Dobrynin, R. Entringer, and I. Gutman. Wiener index of trees: Theory and applications. *Acta Appl. Math.*, 66:211–249, 2001.

[32] M. Farber and R. E. Jamison. Convexity in graphs and hypergraphs. *SIAM J. Algebraic Discrete Methods*, 7(3):433–444, 1986.

[33] E. J. Farrell. Chromatic roots – some observations and conjectures. *Discrete Math.*, 29(2):161–167, 1980.

[34] O. Goldschmidt, P. Jaillet, and R. LaSota. On reliability of graphs with node failures. *Networks*, 24:251–259, 1994.

[35] C. Graves. Inflection points of coherent reliability polynomials. *Australas. J. Combin.*, 49:111–126, 2011.

[36] C. Graves and D. Milan. Reliability polynomials having arbitrarily many inflection points. *Networks*, 64(1):1–5, 2014.

[37] F. Harary. A characterization of block-graphs. *Canad. Math. Bull.*, 6:1–6, 1963.

[38] O. J. Heilmann and E. H. Lieb. Theory of monomer-dimer systems. *Comm. Math. Phys.*, 25(3):190–232, 1972.

[39] C. Hoede and X. Li. Clique polynomials and independent set polynomials of graphs. *Discrete Math.*, 125:219–228, 1994.

[40] R. A. Horn and C. R. Johnson. *Topics in matrix analysis*. Cambridge Univ. Press, 1991.

[41] J. Huh. Milnor numbers of projective hypersurfaces and the chromatic polynomial of graphs. *J. Amer. Math. Soc.*, 25(3):907–927, July 2012.

[42] J. Huh. h-vectors of matroids and logarithmic concavity. *Adv. Math.*, 270:49–59, 2015.

[43] B. Jackson. A zero-free interval for chromatic polynomials of graphs. *Combin. Probab. Comput.*, 2:325–336, 1993.

[44] R. Kirk and H. Wang. Largest number of subtrees of trees with a given maximum degree. *SIAM J. Discrete Math.*, 22:985–995, 2008.

[45] B. Knudsen. *Optimal multiple parsimony alignment with affine gap cost using a phylogenetic tree*, volume 2812, pages 433–446. Springer, Berlin, 2003.

[46] V. B. Le and N. N. Tuy. The square of a block graph. *Discrete Math.*, 310:734–741, 2010.

[47] V. E. Levit and E. Mandrescu. The independence polynomial of a graph - a survey. In S. Bozapalidis, A. Kalampakas, and G. Rahonis, editors, *Proceedings of the 1st International Conference on Algebraic Informatics*, pages 233–254. Aristotle University of Thessaloniki, Department of Mathematics, Thessaloniki, 2005.

[48] S. Liu, K. Cheng, and X. Liu. Network reliability with node failures. *Networks*, 35(2):109–117, 2000.

[49] M. Marden. *Geometry of Polynomials.* Number 3 in Mathematical Surveys. American Mathematical Society, 2nd edition, 1966.

[50] C. Merino. Chip-firing and the Tutte polynomial. *Ann. Comb.*, 1(3):253–259, 1997.

[51] C. Merino. The chip firing game and matroid complexes. In *Discrete models: combinatorics, computation, and geometry (Paris, 2001)*, Discrete Math. Theor. Comput. Sci. Proc., AA, pages 245–255 (electronic). Maison Inform. Math. Discrèt. (MIMD), Paris, 2001.

[52] Q. I. Rahman and G. Schmeisser. *Analytic Theory of Polynomials.* Clarendon Press, Oxford, 2002.

[53] R. C. Read and G. F. Royle. Chromatic roots of families of graphs. *Graph Theory, Combinatorics, and Applications*, 2:1009–1029, 1991.

[54] G. Royle and A. D. Sokal. The Brown-Colbourn conjecture on zeros of reliability polynomials is false. *J. Combin. Theory Ser. B*, 91:345–360, 2004.

[55] D. R. Shier. *Network Reliability and Algebraic Structures.* Clarendon Press, Oxford, 1991.

[56] A. V. Sills and H. Wang. The minimal number of subtrees of a tree. *Graphs Combin.*, 2015(1):255–264, 31.

[57] A. D. Sokal. Bounds on the complex zeros of (di)chromatic polynomials and Potts model partition functions. *Combin. Probab. Comput.*, 10:41–77, 2001.

[58] A. D. Sokal. Chromatic roots are dense in the whole complex plane. *Combin. Probab. Comput.*, 13(2):221–261, 2004.

[59] E. Sperner. Ein Sats über Untermengen einer endlichen Menge. *Mathematische Zeitschrift*, 27:544–548, 1928.

[60] R. P. Stanley. Log-concave and unimodal sequences in algebra, combinatorics, and geometry. In *Graph theory and its applications: East and West (Jinan, 1986)*, volume 576 of *Ann. New York Acad. Sci.*, pages 500–535. New York Acad. Sci., New York, 1989.

[61] C. Stivaros. *On the residual node connectedness network reliability model.* PhD thesis, Dept. of Electrical Engineering and Computer Science, Stevens Institute of Technology, 1990.

[62] K. Sutner, A. Satyanarayana, and C. Suffel. The complexity of the residual node connectedness reliability problem. *SIAM J. Comput.*, 20(1):149–155, 1991.

[63] L. A. Székely and H. Wang. On subtrees of trees. *Adv. in Appl. Math.*, 34(1):138–155, 2005.

[64] L. A. Székely and H. Wang. Binary trees with the largest number of subtrees. *Discrete Appl. Math.*, 155(3):374–385, 2007.

[65] C. Thomassen. The zero-free intervals for chromatic polynomials of graphs. *Combin. Probab. Comput.*, 6:497–506, 1997.

[66] P. van Mieghem. *Graph Spectra for Complex Networks.* Cambridge University Press, 2010.

[67] D. L. Vertigan and D. J. A. Welsh. The computational complexity of the Tutte plane: the bipartite case. *Probab. Combin. Comput. Sci.*, 1:181–187, 1992.

[68] D. G. Wagner. Zeros of reliability polynomials and f-vectors of matroids. *Combin. Probab. Comput.*, 9:167–190, 2000.

[69] S. G. Wagner. Correlation of graph-theoretical indices. *SIAM J. Discrete Math.*, 21(1):33–46, 2007.

[70] D. B. West. *Introduction to graph theory.* Prentice-Hall, second edition, 2001.

[71] H. Wiener. Structural determination of paraffin boiling points. *J. Amer. Chem. Soc.*, 69:17–20, 1947.

[72] J. H. Wilkinson. The evaluation of the zeros of ill-conditioned polynomials. *Numer. Math.*, 1:150–180, 1959.

[73] W.-G Yan and Y.-N Yeh. Enumeration of subtrees of trees. *Theoret. Comp. Sci.*, 369(1–3):256–268, 2006.

[74] S. Yu, F.-M. Shao, and H. Meng. Uniformly optimal graphs in some classes of graphs with node failures. *Discrete Math.*, 310(1):159–166, 2010.

[75] X.-M. Zhang and X.-D. Zhang. The minimal number of subtrees with a given degree sequence. *Graphs Combin.*, 31(1):309–318, 2015.

[76] X.-M. Zhang, X.-D. Zhang, D. Gray, and H. Wang. The number of subtrees of trees with given degree sequence. *J. Graph Theory*, 73(3):280–295, 2013.