Towards New Bounds for the 2-Edge Connected Spanning Subgraph Problem

Philippe Legault

Thesis submitted to the Faculty of Graduate and Postdoctoral Studies
in partial fulfillment of the requirements for the degree of
Master of Computer Science

Ottawa-Carleton Institute for Computer Science
University of Ottawa

Abstract

Given a complete graph $K_n = (V, E)$ with non-negative edge costs $c \in \mathbb{R}^E$, the problem multi-2EC$_{\text{cost}}$ is that of finding a 2-edge connected spanning multi-subgraph of K_n with minimum cost. It is believed that there are no efficient ways to solve the problem exactly, as it is NP-hard. Methods such as approximation algorithms, which rely on lower bounds like the linear programming relaxation multi-2EC$_{\text{cost}}^{\text{LP}}$ of multi-2EC$_{\text{cost}}$, thus become vital to obtain solutions guaranteed to be close to the optimal in a fast manner.

In this thesis, we focus on the integrality gap αmulti-2EC$_{\text{cost}}$ of multi-2EC$_{\text{cost}}^{\text{LP}}$, which is a measure of the quality of multi-2EC$_{\text{cost}}^{\text{LP}}$ as a lower bound. Although we currently only know that $\frac{6}{5} \leq \alpha$multi-2EC$_{\text{cost}} \leq \frac{3}{2}$, the integrality gap for multi-2EC$_{\text{cost}}$ has been conjectured to be $\frac{6}{5}$. We explore the idea of using the structure of solutions for αmulti-2EC$_{\text{cost}}$ and the concept of convex combination to obtain improved bounds for αmulti-2EC$_{\text{cost}}$. We focus our efforts on a family J of half-integer solutions that appear to give the largest integrality gap for multi-2EC$_{\text{cost}}$. We successfully show that the conjecture αmulti-2EC$_{\text{cost}} = \frac{6}{5}$ is true for any cost functions optimized by some $x^* \in J$.

We also study the related problem 2EC$_{\text{size}}$, which consists of finding the minimum size 2-edge connected spanning subgraph of a 2-edge connected graph. The problem is NP-hard even at its simplest, when restricted to cubic 3-edge connected graphs. We study that case in the hope of finding a more general method, and we show that every 3-edge connected cubic graph $G = (V', E')$, with $n = |V'|$ allows a 2EC$_{\text{size}}$ solution for G of size at most $\frac{7n}{6}$. This improves upon Boyd, Iwata and Takazawa's guarantee of $\frac{6n}{5}$ and extend Takazawa's $\frac{7n}{6}$ guarantee for bipartite cubic 3-edge connected graphs to all cubic 3-edge connected graphs.

Acknowledgements

First, I would like to express my sincere gratitude to my advisor, Dr. Sylvia Boyd, for her continuous support during my studies, and for sharing her passion of combinatorial optimization. Her guidance was critical during my research and writing of this thesis. I could not have imagined having a better mentor for my Master's study.

I would also like to thank the University of Ottawa and the Natural Sciences and Engineering Research Council of Canada for their financial support in the form of scholarships.

I am also grateful to my family and friends who have encouraged me along the way.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

Reliability is one of the key attributes of most networks, and is often achieved with redundancy, which infers extra cost. In the case of data centers physically linked with fiber optic cables, this reliability requirement might translate into the addition of extra cables linking some locations, so that an alternate path is available should a cable fail, i.e. the network remains connected in the catastrophic event of a link failure. The layout of dependable networks at minimum cost is an important problem in network design, and has practical, cost-saving applications such as the planning of power lines, roads, railways and communication networks, all of which can survive the loss of a link. The cost is not always financial and depends on the nature of the network, e.g. minimizing network latency in the case of a data network.

The problem statement can be modelled by an undirected multi-graph $G = (V, E)$, whose vertices represent entities, and whose edges with non-negative weight linking the vertices stand for possible links and their projected cost. Multi-edges are allowed, as it is sometimes more practical to double a link than to build two separate ones: for instance, sensors in nuclear power plants must each be connected by three wires to the main system, and it is much cheaper to install the wires alongside one another than to route them through different paths. Moreover, the input graph typically has a edge between each pair of vertex and is thus labelled *complete graph*; should this not be the case, then missing edges can be added with a sufficiently high cost. The output to the problem is a subgraph which allows multi-edges, which we call a *multi-subgraph* and includes all the vertices, which we call *spanning*. Moreover, said subgraph stays connected in the event that an edge is removed, which we call *2-edge connected*, and minimizes the sum of the cost of the edges it contains. More formally, given a complete graph G, the *Minimum Cost 2-Edge Connected Spanning Subgraph Problem*, henceforth multi-2EC$_{\text{cost}}$, consists of finding a 2-edge connected spanning multi-subgraph of G of minimum cost.

This task becomes increasingly difficult as the input graph grows bigger and it is believed that there exists no efficient way to find the optimal solution to multi-2EC$_{\text{cost}}$ for graphs of arbitrary size, as the problem is NP-hard ([8, 13]) and MAX SNP-hard[1]. This is the case even when restricting the problem to simpler cases [8]. As many real-world applications depend on finding a solution in a reasonnable amount of time, methods named *heuristics* have been devised to trade the optimality of the solution found for an improved runtime. In other words, these methods generally run faster, but likely output a solution that is not optimal. When a heuristic runs in polynomial-time and provides a guarantee that the cost of the solutions it produces will be within a constant factor $\alpha > 1$ of the cost

[1] Refer to [13] for details about complexity classes.

of the optimal solution, it is called an *α-approximation algorithm*.

A strong guarantee for an α-approximation algorithm for multi-2EC$_{\text{cost}}$ is heavily reliant on the quality of the lower bound used for the problem —a lower bound is obtained via a polynomial-time process that outputs a value always less than or equal to the cost of the optimal for multi-2EC$_{\text{cost}}$: the closer to the optimal comes that value, the better guarantee can then be obtained for approximation algorithms depending on the lower bound.

Naturally, the study of lower bounds becomes crucial to the development of improved approximation algorithms. One such lower bound, the *Linear Programming Relaxation Bound*, denoted by multi-2EC$^{\text{LP}}_{\text{cost}}$, uses a mathematical model named linear programming (LP) to weaken some constraints from the original problem of multi-2EC$_{\text{cost}}$. More specifically, the LP relaxation of multi-2EC$_{\text{cost}}$ allows the variables representing the occurence of edges in the solution to have a fractional value; the concept is described in more depth in Section 2.2. In this thesis, we focus almost exclusively on the LP relaxation of multi-2EC$_{\text{cost}}$ and seek to better understand its quality as a lower bound to multi-2EC$_{\text{cost}}$, which is crucial to applications such as branch and bound search and the cutting-plane method. We study the *integrality gap* which is the worst-case ratio between the cost of the optimal multi-2EC$_{\text{cost}}$ solution for a graph G, and the cost of its linear relaxation solution on the same graph G and over all cost functions [1]. A polynomial-time proof of an upper bound of k on the integrality gap also serves as a k-approximation algorithm. More information is provided in Section 2.2.2.

Before continuing further, we must distinguish between different variants of multi-2EC$_{\text{cost}}$ and other related problems, as our work extends to more than was discussed so far. The next subsection explains the differences between the different problems, and the notation we use; Subsection 1.1.3 details the way to which these problems relate to the integrality gap.

1.1 multi-2EC$_{\text{cost}}$

There are different variants of multi-2EC$_{\text{cost}}$, which we label and define here. The variants discussed here have weights, or costs, on the edges: we refer to a vector $c \in \mathbb{R}^{E}_{\geq 0}$ mapping the edges to their non-negative weight as the *cost function* for multi-2EC$_{\text{cost}}$. Variants that are not defined are equivalent to one of the versions mentioned in the list below. Note that while multi-2EC$_{\text{cost}}$ can also be defined for digraphs, we focus solely on the undirected version of the problem.

For a complete weighted graph $K_n = (V, E)$ with cost function $c : E \mapsto \mathbb{R}_{\geq 0}$, we say that the costs are *metric* if and only if they satisfy the *triangle inequality*, i.e.

$$c_{ab} + c_{bc} \geq c_{ac} \quad \text{for all } a, b, c \in V.$$

The two main multi-2EC$_{\text{cost}}$ variants are listed below. The version of the problem referred to in the previous section is (metric) multi-2EC$_{\text{cost}}$ and is a superset of the other variants. For a variant A of multi-2EC$_{\text{cost}}$, we denote by $\mathrm{OPT}(A(G))$ the cost of the optimal solution to A for a graph G.

1.1.1 (metric) multi-2EC$_{\text{cost}}$

The problem is that of finding the minimum cost 2-edge connected spanning multi-subgraph of a complete graph $K_n = (V, E)$ on n vertices with edge costs $c \in \mathbb{R}^E_{\geq 0}$.

We may assume without loss of generality that the edge costs are metric, as multiple copies of an edge are allowed [5]. More specifically, the cost of an edge $uv \in E$ that does not respect the triangle inequality is replaced by the cost of the shortest path between u and v. That path is selected (or doubled) if uv appears in the solution. Note that if the graph is not complete, we can make it into a complete graph by adding in all the "missing" edges and assigning an appropriately large cost to each of them.

1.1.2 2EC$_{\text{cost}}$

The problem 2EC$_{\text{cost}}$ is that of finding the minimum cost 2-edge connected spanning subgraph of a complete graph $K_n = (V, E)$ on n vertices with edge costs $c \in \mathbb{R}^E_{\geq 0}$. Note that multiple copies of an edge are not allowed in the subgraphs and the costs may not always be metric. However, if the costs are metric, then OPT(metric 2EC$_{\text{cost}}(K_n)$) = OPT(metric multi-2EC$_{\text{cost}}(K_n)$). Lastly, OPT(multi-2EC$_{\text{cost}}(K_n)$) $\leq$ OPT(2EC$_{\text{cost}}(K_n)$), as the former is a relaxation of the latter.

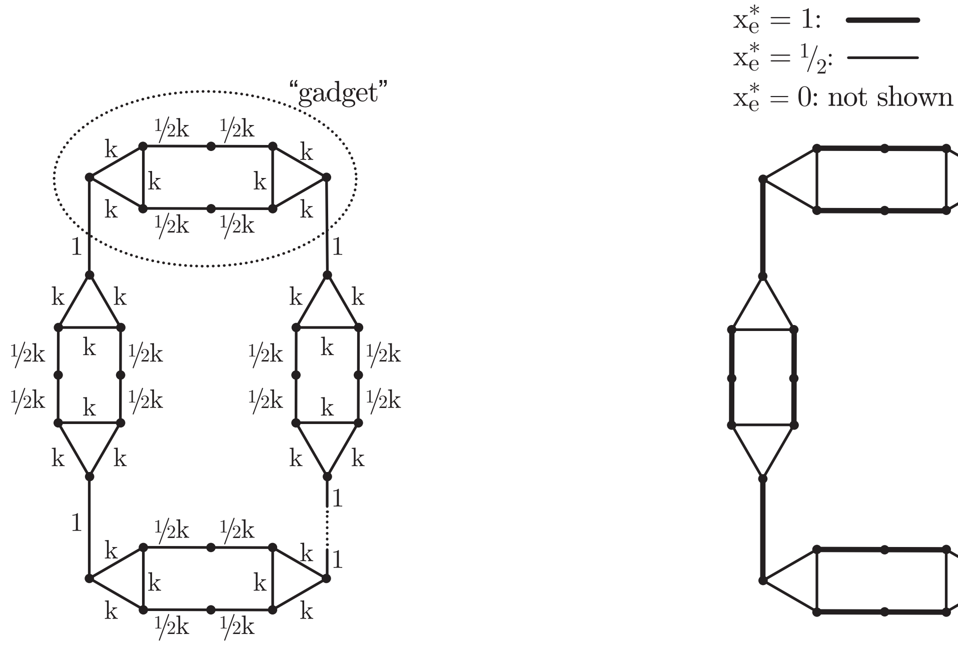
1.1.3 The Integrality Gap of multi-2EC$_{\text{cost}}$

Even though multi-2EC$_{\text{cost}}$ has been intensively studied, little is known about the integrality gap αmulti-2EC$_{\text{cost}}$, except that $\frac{6}{5} \leq \alpha$multi-2EC$_{\text{cost}} \leq \frac{3}{2}$ [1]. In [5], Carr and Ravi study αmulti-2EC$_{\text{cost}}$ and conjecture that αmulti-2EC$_{\text{cost}} = \frac{4}{3}$; however no examples are known for which the integrality gap ratio comes close to $\frac{4}{3}$. In [1], Alexander, Boyd and Elliott-Magwood also study αmulti-2EC$_{\text{cost}}$ and make the following stronger conjecture based on their findings:

Conjecture 1. *[1] The integrality gap αmulti-2EC$_{cost}$ for multi-2EC$^{LP}_{cost}$ is $\frac{6}{5}$.*

To investigate αmulti-2EC$_{\text{cost}}$ further, a natural next step is to study αmulti-2EC$_{\text{cost}}$ for some interesting class of cost functions. Given any feasible solution $x^* \in \mathbb{R}^E_{\geq 0}$ for multi-2EC$^{\text{LP}}_{\text{cost}}$, we denote the occurence of an edge e in x^* by x^*_e. The *support graph* G_{x^*} of x^* is defined to be the subgraph of K_n obtained by taking all edges $e \in E$ for which $x^*_e > 0$. A feasible solution x^* for multi-2EC$^{\text{LP}}_{\text{cost}}$ is called a *half-integer solution* if $x^*_e \in \{0, \frac{1}{2}, 1\}$ for all $e \in E$, and it is called *degree-tight* if $\sum_{uv}(x^*_{uv} : u \in V) = 2$ for all $v \in V$. Finally, a degree-tight half-integer solution is called a *half-triangle solution* if the edges in the support graph G_{x^*} corresponding to $x^*_e = \frac{1}{2}$ (called *half-edges*) form disjoint 3-cycles (called *half-triangles*) joined by paths of edges of value 1 (called *1-paths*). The support graph of a half-triangle solution is called a *half-triangle graph*.

The half-integer solutions are of interest as there is evidence that $\frac{\text{OPT(multi-2EC}_{\text{cost}})}{\text{OPT(multi-2EC}^{\text{LP}}_{\text{cost}})}$ is greatest for cost functions optimized at such solutions [1, 5]. For example, the largest such ratio known is asymptotically $\frac{6}{5}$ [1], and comes from the infinite family of multi-2EC$^{\text{LP}}_{\text{cost}}$ problems shown in Figure 1.1a, where the numbers shown are the edge costs, edges uv not shown have cost equal to the minimum cost uv path, and the "gadget" pattern is repeated k times. This family is optimized for multi-2EC$^{\text{LP}}_{\text{cost}}$ by the half-triangle solution x^* shown



(a) The edge costs. (b) The half-triangle optimal solution x^*.

Figure 1.1: An example for which αmulti-2EC$_{\text{cost}} = \frac{6}{5}$[1].

in Figure 1.1b[2].Also, in a computational study which found αmulti-2EC$_{\text{cost}}$ exactly for all K_n up to $n = 10$, αmulti-2EC$_{\text{cost}}$ was given by a half-triangle solution for all values of n [1].

[2]To compute the $\frac{6}{5}$ ratio, we will examine a gadget independently of the rest of the graph. First notice that a minimum cost 2-edge connected spanning subgraph of a gadget must use nine edges, four of which have a cost of k, four with a cost of $\frac{1}{2}k$ and one with a cost of 1, for a total cost of $6k + 1$. For every gagdet in the half-triangle solution of Figure 1.1b however, there are four edges with an occurence of 1 and a weight of $\frac{k}{2}$, six edges with an occurence of $\frac{1}{2}$ and a weight of k, and one edge with an occurence and weight of 1. The cost of the LP solution for the gadget is thus $5k + 1$. As the number of gadgets k approaches infinity, the integrality gap $\frac{6k+1}{5k+1}$ for the example approaches $\frac{6}{5}$.

1.2 2EC_{size}

A related problem to multi-2EC_{cost} which we denote by 2EC_{size} consists of finding a 2-edge connected spanning subgraph with a minimum number of edges, for a given 2-edge connected multi-graph $G = (V, E)$. Note that multi-2EC_{cost} and 2EC_{size} are fundamentally different, as the former requires the input graph to be K_n. Furthermore, 2EC_{size} does not have costs on the edges and multiple copies of an edge are not necessary [6].

Like multi-2EC_{cost}, the problem has important applications in network design, when the cost of links is uniform, or unimportant. However, 2EC_{size} has been shown to be NP-hard [18], and also MAX SNP-hard, even in the simpler case of cubic graphs (graphs with all vertices of degree 3) [8]. Since 2EC_{size} on cubic graphs is the simplest among all 2EC_{size} problems which remain NP-hard, we concentrate on solving 2EC_{size} for those graphs in the hope that successful methods can be generalized.

The concept of integrality gap is defined over all cost functions for multi-2EC_{cost}: as 2EC_{size} is defined on graphs with a unit cost function, we define a concept similar to the integrality gap to assess the quality of $2\text{EC}_{\text{size}}^{\text{LP}}$ as a lower bound. We use $\alpha 2\text{EC}_{\text{size}}$ to denote the worst case ratio between $\text{OPT}(2\text{EC}_{\text{size}}(G))$ and $\text{OPT}(2\text{EC}_{\text{size}}^{\text{LP}}(G))$ over all input graphs G. Currently, it is known that, $\frac{8}{7} \leq \alpha 2\text{EC}_{\text{size}} \leq \frac{4}{3}$, for 2EC_{size} [3, 27].

1.3 Thesis Contributions

The top three contributions of this thesis are listed below.

1. We show that the integrality gap of the LP relaxation for multi-$2EC_{\text{cost}}$ equals $\frac{6}{5}$ when restricted to cost functions optimized at half-triangle solutions, a family of cost functions which has been shown to provide the worst case by empirical testing up to $n = 10$ [1]. The family also includes the worse case known for the integrality gap. This supports Conjecture 1 which states that the integrality gap is $\frac{6}{5}$ for the LP relaxation for all cost functions; it also improves on Carr and Ravi's upper bound of $\frac{4}{3}$ for the integrality gap of a family of cost function including the one we study [5].

2. The problem $2EC_{\text{size}}$ restricted to cubic 3-edge connected graphs is the simplest form of the problem known to remain NP-hard [8] and thus, warrants further study in the hopes of generalizing successful methods. We show that every 3-edge connected cubic graph $G = (V, E)$, with $n = |V|$ allows a $2EC_{\text{size}}$ solution for G of size at most $\frac{7n}{6}$, which improves upon Boyd, Iwata and Takazawa's guarantee of $\frac{6n}{5}$ [4]. Our methods are not polynomial and thus, do not result in an approximation algorithm. Nevertheless, they give hope that a $\frac{7}{6}$-approximation algorithm exists, which would improve on the existing methods which gives a $\frac{6}{5}$-approximation [4]. It would also extend Takazawa's $\frac{7}{6}$-approximation for bipartite cubic 3-edge connected graphs to all cubic 3-edge connected graphs [28].

3. We define a new reduction operation for 3-edge connected cubic graphs with specific connectivity requirement, i.e. to be essentially 4-edge connected (see Section 2.1 for a definition). Our operation is useful for any inductive proof on essentially 3-edge connected cubic simple graphs, as it exploits all the properties of the graph to permit a more complex reduction.

1.4 Thesis Outline

We finish this chapter with a literature review on the different variants of multi-2EC$_{\text{cost}}$and 2EC$_{\text{size}}$. Following this, we structure the remainder of this thesis as follows:

Chapter 2 introduces the different definitions and notation for concepts used in the thesis, and delves deeper in some areas of graph theory and integer and linear programming.

Chapter 3 is dedicated to 2EC$_{\text{size}}$ and shows that every 3-edge connected cubic graph $G = (V, E)$, with $n = |V|$ allows a 2-edge connected spanning subgraph of size $\frac{7n}{6}$. This improves on the bound of $\frac{6n}{5}$ in [4] and implies that a $\frac{7}{6}$-approximation algorithm exists. Such an approximation would improve on Boyd, Iawata and Takazawa's $\frac{6}{5}$-approximation algorithm for 3-edge connected cubic graphs, and would also generalize Takazawa's $\frac{7}{6}$-approximation algorithm for 3-edge connected cubic bipartite graphs.

Chapter 4 uses results from Chapter 3 and demonstrates that the integrality gap of the LP relaxation of multi-2EC$_{\text{cost}}$ equals $\frac{6}{5}$ when restricted to the family of cost functions optimized at half-triangle solutions. This family has been identified to give the worst case in a computational study which found the integrality gap for multi-2EC$_{\text{cost}}$ exactly for all K_n up to $n = 10$. Moreover, it is a cost function belonging to that family which gives the current worse case for the integrality gap.

Chapter 5 is a survey of seven known lower bounds to 2EC$_{\text{size}}$; we also compare the bounds and find two which together, outperform all the others.

In Chapter 6, we make some concluding remarks and suggestions for future work.

1.5 Literature Review

$\mathbf{2EC_{cost}}$

For general weights, Frederickson and Ja'Ja' are the first to approximate a solution for the problem of augmenting a given graph to be 2-connected, and obtain an approximation factor of 3 [9]. Their approximation factor is later improved to 2 by Jain in [17] and in the same paper, it is noted that the algorithm has the same approximation guarantee for multi-$2EC_{cost}$. Much effort was since invested in specific cases of $2EC_{cost}$, and parent problems, such as the Steiner Tree Problem in [14, 15, 19, 24]: these problems either comprise more requirements than $2EC_{cost}$ or focus on graphs with a specific structure, like planar graphs.

multi-$\mathbf{2EC_{cost}}$

Frederickson and JaJa' [10] obtained a $\frac{3}{2}$-approximation algorithm in 1982, by modifying an algorithm named Christofides' algorithm for a very similar problem —the Travelling Salesman Problem [7].

$\mathbf{2EC_{size}}$

The first to improve the 2-approximation ratio for $2EC_{size}$were Khuller and Vishkin [20], in 1994, with their $\frac{3}{2}$-approximation algorithm based on depth-first search trees and a method called "tree-carving". The $\frac{3}{2}$-approximation by Khuller and Vishkin [20] was refined in 1998 to $\frac{17}{12}$ by Cheriyan, Sebő and Szigeti [6] via ear decompositions, then to $\frac{4}{3}$ by Sebő and Vygen [27], in 2014. Krysta and Kumar [21] improved the ratio to $\frac{4}{3} - \epsilon$ based on a charging scheme. Other results have been claimed, but have either been shown to be false, or are without conclusive proofs [22].

In parallel, much development also occured for special cases, especially cubic and subcubic graphs (i.e. graphs with degree less or equal to 3 everywhere). Krysta and Kumar [21] designed a $\frac{21}{16}$-approximation algorithm for cubic graphs. For subcubic graphs, in 2002, Csaba, Karpinski and Krysta [8] attained the approximation ratio of $\frac{5}{4} + \epsilon$, which was then improved to $\frac{5}{4}$ by Boyd, Fu and Sun [3] in 2014, using circulations. When $2\mathrm{EC}_{\mathrm{size}}$ is further restricted to 3-edge connected cubic graphs, we denote the work of Huh [23], and of Boyd, Iwata and Takazawa [4], who achieved ratios of $\frac{5}{4}$, and $\frac{6}{5}$, respectively. Very recently, Takazawa [28] obtained a $\frac{7}{6}$-approximation algorithm for bipartite cubic 3-edge connected graphs, inspired from the work of Boyd, Iwata and Takazawa, by using 2-factors covering specified cuts.

In the general case, Sebő and Vygen in [27] proved that the unit integrality gap $\alpha 2\mathrm{EC}_{\mathrm{size}}$ is bounded above by $\frac{4}{3}$, and Boyd, Fu and Sun in [3] showed that is is bounded below by $\frac{8}{7}$. The integrality gap $2\mathrm{EC}_{\mathrm{size}}$ for cubic 3-edge connected graphs was shown in 2013 by Boyd, Iwata and Takazawa [4] to be bounded above by $\frac{6}{5}$, and it is known that they are bounded below by the Petersen graph, at $\frac{11}{10}$, as shown in Figure 1.2 (where bold lines represent edges in the subgraph and dotted edges stand for edges in the Petersen graph, but not in the subgraph).

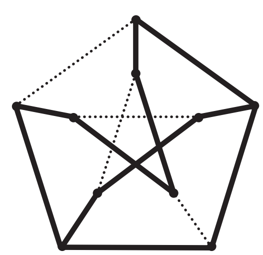


Figure 1.2: The minimum size 2-edge connected spanning subgraph for the Petersen graph.

Chapter 2

Preliminaries

Here we introduce definitions, notation and concepts in graph theory and on integer programming. We also delve into the integrality gap of multi-2EC$_{\text{cost}}$ and the unit integrality gap for 2EC$_{\text{size}}$, as well as include the integer programming formulation and the linear programming relaxation for the two problems. These notions will be used in subsequent chapters.

2.1 Graph Theory

Here we define some notation and concepts related to graph theory that will be used in the thesis. We refer the reader to Bondy and Murty's work in [2] for more details.

Let $G = (V, E)$ be a simple undirected graph with vertex set V and edge set E. All graphs in this thesis can be assumed undirected. The *order* of G is $|V|$, also referred to as n.

An edge of G consists of a pair of nodes, called *endpoints*. When more than one edge e has the same endpoints, G is considered a *multi-graph*, e is a *multi-edge* and E is a multi-set, as it may contain an edge $uv \in E$ more than once. It is common to denote an edge $e \in E$ with endpoints $u, v \in V$ as $e = uv$. When there exists an edge between every pair of vertices, G can also be denoted by K_n. The *degree* of a vertex $v \in V$ is the number of edges with v as an endpoint. A graph is *cubic* if all vertices have exactly 3 incident edges, and is *subcubic* if all vertices have 3 incident edges or less. A *subgraph* of G is a graph $H = (V_H, E_H)$ where $V_H \subseteq V$ and $E_H \subseteq E$, such that if edge $uv \in E_H$, then $u, v \in V_H$. Subgraph H is said to be *spanning* if $V_H = V$. When H contains a multi-edge, we call it a *multi-subgraph* and note that its edge set is a multi-set. For any multi-subgraph H of G we sometimes use the notation $E(H)$ and $V(H)$ to denote the edge set and the vertex set for H, respectively, and we use $\chi^{E(H)} \in \mathbb{R}^E$ to denote the incidence vector of subgraph H (i.e. $\chi_e^{E(H)}$ is the number of copies of edge e in H). A *component* of G is a subgraph of G for which any two vertices are connected to each other by paths, and which is not connected to any other vertices in G.

For any subset $A \subseteq V$, the *complement* of A is $\overline{A} = V \setminus A$. For any two non-overlapping subsets $A, B \subseteq V$, the edge set between A and B is denoted by $E[A : B]$, i.e. $E[A : B] = \{ij \in E \mid i \in A, j \in B\}$. For any $A \subset V$ we define $\delta(A) = E[A : V \setminus A]$. Note that the subset $F = \delta(A)$ of E for some $A \subset V$ is called a *cut-set* or a *cut* as its removal disconnects G. Graph G is considered k-edge connected if and only if all cuts of G have size $\geq k$, $k \in \mathbb{N}$. An edge cut that contains k edges is a *k-edge cut*. An edge cut F of G is *essential* if $G' = (V, E \setminus F)$ has at least two components each containing more than one vertex. Each of those components is called a *shore* of the cut. For a $k \in \mathbb{N}$, G is said to be *essentially k-edge connected* if and only if G does not have an essential edge cut F

with $|F| < k$. If G is cubic, simple and essentially 4-edge connected, then it is also 3-edge connected. If the removal of a single vertex $v \in V$, or a single edge $e \in E$, disconnects G, then that vertex is called a *cut-vertex* and that edge is called a *bridge*.

A *2-matching* of G is a subgraph of G with all vertices of degree 2 or less. A 2-matching M is *maximum* if and only if no other matching contains more edges than M and is *maximal* if it is not the proper subset of another 2-matching. An *independent set* of G is a set $T \subseteq V$ such that for any $u, v \in T$, $uv \notin E$. A maximum independent set contains at least as many vertices as any other independent set for the graph; a maximal independent set is not the proper subset of another independent set. A *Depth-First Search* is a graph traversal technique that starts at a vertex and explores as far as possible along each branch, never visiting a vertex twice, and backtracks when necessary. A *reduction* operation of G is a transformation of the graph into a smaller graph, typically via the removal of edges or vertices, such that the graph keeps a set of invariants. A *walk* in a simple graph such as G is a sequence $v_0, v_1, \ldots, v_n$ where v_i are vertices $v_{i-1}v_i \in E$ for $i = 1, \ldots, n$. A walk for which $v_0 = v_n$ is called a tour; when a tour visits every vertex of a graph exactly once, it is called a *Hamiltonian tour*.

2.2 Integer and Linear Programming

For more details about integer programming, we refer the reader to Schrijver's book [25].

Given a vector $x \in \mathbb{R}^p$ of *decision variables* and costs $c \in \mathbb{R}^p$, a linear program, or LP, is an optimization problem consisting of a linear *objective function* $c^\intercal x$ subject to a finite number m of linear equality and linear inequality constraints on x. We can use matrix $A \in \mathbb{R}^{m \times p}$ and vector $b \in \mathbb{R}^m$ to represent these inequalities when the LP is in standard form, i.e.

$$\max c^\intercal x \text{ s.t. } Ax \leq b, x \geq 0 \quad \text{or} \quad \min c^\intercal x \text{ s.t. } Ax \leq b, x \geq 0.$$

The vector x is called a *feasible solution* when it satisfies all constraints of the LP. If it minimizes the LP of a minimization problem, or if it maximizes the LP of a maximization problem, it is an *optimal solution*. LPs can be solved in polynomial time in the number of constraints using the Interior Point Method.

If instead the decision variables must be integer-valued (a requirement called *integrality constraint*, i.e. $x \in \mathbb{Z}^p$) the optimization problem is known as an *integer linear program* (ILP). Finding the optimum of general ILPs is NP-hard but removing the integrality constraint of an ILP gives us its *linear relaxation*. In the case of multi-2EC$_{\text{cost}}$, the number of constaints of the LP is exponential on the size of the input graph, which would make it impractical to solve. However, as it is possible to find in polynomial time a number of inequalities that bring us closer to the solution, a problem known as the *separation problem*, not all constraints of the LP have to be generated and as such, the LP relaxation of multi-2EC$_{\text{cost}}$ can be solved in polynomial time using the Ellipsoid Method [25]. Therefore, the linear relaxation of multi-2EC$_{\text{cost}}$ serves as a lower bound. LP relaxations also become useful in techniques such as branch and bound.

The concept of convex combination is used extensively: in general, a *convex combination* is a linear combination of vectors or scalars where all the coefficients are non-negative and sum to 1. In the context of this thesis, as we use convex combinations exclusively with vectors related to graph structures, we say that a vector $y \in \mathbb{R}^E$ is a *2EC convex combination* if there exist 2-edge connected spanning multi-subgraphs H_i with multipliers $\lambda_i \in \mathbb{R}_{\geq 0}$, $i = 1, 2, \ldots, j$ such that $y = \sum_{i=1}^{j} \lambda_i \chi^{E(H_i)}$ and $\sum_{i=1}^{j} \lambda_i = 1$. A 2EC convex combination can be thought of as a discrete probability distribution where the outcomes are 2-edge connected spanning multi-subgraphs and the probabilities play the same role as the λ_i's. If for a graph $G = (V, E)$, y is a 2EC convex combination of G that select every edge exactly k times overall, with $0 < k < 1$, then it follows that at least one H_i has at most $k|E|$ edges. Similarly, one H_i will have a cost of at most $k \sum_{e \in E} c_e$ for cost function $c \in \mathbb{R}^E_{\geq 0}$.

2.2.1 Duality Theory

Every minimization LP, known as a *primal* problem has a *dual* problem. Given decision variable $y \in \mathbb{R}^m$ the following relationship between primal and dual can be established.

$$
\begin{aligned}
\text{primal:} \quad & \min\{c^\intercal x : Ax \geq b, x \geq 0\} \\
\text{dual:} \quad & \max\{b^\intercal y : A^\intercal y \leq c, y \geq 0\}
\end{aligned}
\tag{2.1}
$$

Note that the dual of the dual of any LP is the original LP. A feasible solution of the dual of a minimization problem provides a lower bound to the primal's objective value. This concept is called the *Weak Duality Theorem* and is expressed below.

Theorem 1 ([25])**.** *Given a feasible solution $x \in \mathbb{R}^p$ of an LP, and a feasible solution $y \in \mathbb{R}^m$ of its dual, as well as the coefficients of their objective functions $c \in \mathbb{R}^p$ and*

$b \in \mathbb{R}^m$, *respectively,*

$$\sum_{i=1}^{m} b_i y_i \leq \sum_{i=1}^{p} c_j x_j.$$

2.2.2 Integrality Gap

Let $G = (V, E)$ be a graph with cost function $c \in \mathbb{R}^E$, and let $\mathrm{OPT}(G)$ and $\mathrm{OPT}_{\mathrm{LP}}(G)$ denote the optimal value of an ILP associated with G and its LP relaxation, respectively. The optimal value for the LP relaxation of a maximization ILP provides an upper bound to the optimal value of the objective function. When the ILP is a minimization problem, the LP relaxation optimal value provides a lower bound to the optimal value of the objective function. The quality of the bound is vital for applications such as the branch and bound method and the ratio of $\mathrm{OPT}(G)$ over $\mathrm{OPT}_{\mathrm{LP}}(G)$ is used to measure it. The closer this ratio comes close to 1, the better the bound. However, we would like to consider the quality of the bound for all input graphs and cost functions and as such, the concept of *integrality gap* is used, which is the worst case ratio between $\mathrm{OPT}_{\mathrm{LP}}(G)$ and $\mathrm{OPT}(G)$ over all input graphs G and all cost functions. Additionally, a polynomial-time constructive proof that the integrality gap equals $k \in \mathbb{R}$ would provide a k-approximation algorithm for the ILP. To see this, consider that such a proof lists a polynomial number of steps that result in a feasible ILP solution of cost always less than or equal to k times the cost of $\mathrm{OPT}({}_{\mathrm{LP}}G)$, and $\mathrm{OPT}_{\mathrm{LP}}(G) \leq \mathrm{OPT}(G)$.

2.2.3 ILP Formulation of multi-2EC$_{\text{cost}}$

The ILP of multi-2EC$_{\text{cost}}(K_n)$ for graph K_n is given as follows. Let x_e represent the number of copies of edge e used in the solution. Given a set $F \subseteq E$, let $x(F) = \sum_{e \in F} x_e$.

$$
\begin{aligned}
\text{Minimize} \quad & c^\intercal x & & \\
\text{Subject to} \quad & x(\delta(S)) \geq 2 & & \text{for all } \emptyset \neq S \subset V, \\
& x_e \geq 0, \text{ and integer} & & \text{for all } e \in E.
\end{aligned} \tag{2.2}
$$

The LP relaxation of multi-2EC$_{\text{cost}}$ is obtained by removing the integer constraint from (2.2), and is denoted by multi-2EC$^{\text{LP}}_{\text{cost}}$(G), i.e.

$$
\begin{aligned}
\text{Minimize} \quad & c^\intercal x & & \\
\text{Subject to} \quad & x(\delta(S)) \geq 2 & & \text{for all } \emptyset \neq S \subset V, \\
& x_e \geq 0 & & \text{for all } e \in E.
\end{aligned} \tag{2.3}
$$

The integrality gap of multi-2EC$^{\text{LP}}_{\text{cost}}$ is denoted by αmulti-2EC$_{\text{cost}}$.

2.2.4 ILP Formulation of $2EC_{\text{size}}$

For a graph $G = (V, E)$, let x_e be 1 if edge e is included in the $2EC_{\text{size}}$ solution, and 0 otherwise. Denoted by $2EC^{\text{ILP}}_{\text{size}}(\text{G})$, the ILP of $2EC_{\text{size}}(\text{G})$ for graph G is given as follows:

$$
\begin{aligned}
\text{Minimize} \quad & \sum_{e \in E} x_e \\
\text{Subject to} \quad & x(\delta(S)) \geq 2 && \text{for all } \emptyset \neq S \subset V, \\
& 0 \leq x_e \leq 1, \text{ and integer} && \text{for all } e \in E.
\end{aligned} \tag{2.4}
$$

The LP relaxation of $2EC^{\text{ILP}}_{\text{size}}$ is obtained by removing the integer constraint from (2.4), and is denoted by $2EC^{\text{LP}}_{\text{size}}(\text{G})$, i.e.

$$
\begin{aligned}
\text{Minimize} \quad & \sum_{e \in E} x_e \\
\text{Subject to} \quad & x(\delta(S)) \geq 2 && \text{for all } \emptyset \neq S \subset V, \\
& 0 \leq x_e \leq 1 && \text{for all } e \in E.
\end{aligned} \tag{2.5}
$$

By associating a variable d_S with every $\emptyset \neq S \subset V$, we obtain the dual LP of (2.5).

$$
\begin{aligned}
\text{Maximize} \quad & \sum_{\emptyset \neq S \subset V} 2d_S \\
\text{Subject to} \quad & \sum_{e \in \delta(S)} d_S \leq 1 && \text{for all } e \in E, \\
& d_S \geq 0 && \text{for all } \emptyset \neq S \subset V.
\end{aligned} \tag{2.6}
$$

Dual 2.6 will prove useful in the completion of certain proofs. Let $\mathbb{S} = \{S : \emptyset \neq S \subset V\}$

and note that by the Weak Duality Theorem (Theorem 1), we have the following result.

Theorem 2. *Any feasible solution $x \in \mathbb{R}^E$ to (2.5) has an objective value greater than or equal to the objective value of any feasible solution $d \in \mathbb{R}^{\mathbb{S}}$ to (2.6), i.e. $\sum_{e \in E} x_e \geq \sum_{S \in \mathbb{S}} d_S$.*

Chapter 3

A $\frac{7}{6}$ Bound for the Unit Integrality Gap for 2EC_{size}

In this chapter, we show that the unit integrality gap for 2EC_{size} is at most $\frac{7}{6}$ when restricted to cubic 3-edge connected graphs. To obtain this for such a graph $G = (V, E)$, we demonstrate in Section 3.2 that $\frac{7}{9}\chi^{E(G)}$ is a convex combination of incidence vectors of 2-edge connected spanning subgraphs of G. As a warm-up for the proof, we first show a lesser result in Section 3.1, namely that $\frac{4}{5}\chi^{E(G)}$ is a convex combination of incidence vectors of 2-edge connected spanning subgraphs of G. The result —also of independent interest, as will be seen in Subsection 3.1.1— allows for a strengthening of the proof to $\frac{7}{9}$ when combined with a new reduction operation that is specific to essentially 4-edge connected cubic simple graphs. Finally, Section 3.3 ties the previous results of the chapter to 2EC_{size}.

3.1 $\frac{4}{5}\chi^{E(G)}$ as a 2EC$_{\text{size}}$ Convex Combination

In this section, we show that for a cubic 3-edge connected graph $G = (V, E)$, $\frac{4}{5}\chi^{E(G)}$ is a 2EC$_{\text{size}}$ convex combination. We then portray the importance of the findings with regards to multi-2EC$_{\text{cost}}$ and to 2EC$_{\text{size}}$. Finally, we generalize our result and show that for a 3-edge connected graph $G = (V, E)$, even when not cubic, $\frac{4}{5}\chi^{E(G)}$ is still a 2EC$_{\text{size}}$ convex combination.

Lemma 1. *Let $G = (V, E)$ be a cubic 3-edge connected multi-graph. Then $\frac{4}{5}\chi^{E(G)}$ is a 2EC$_{size}$ convex combination.*

Besides being useful in the proof of our main result in Chapter 3, Lemma 1 is included as a warm-up to the proof for $\frac{7}{9}\chi^{E(G)}$ as a 2EC$_{\text{size}}$ convex combination; the lemma shares a similar proof structure to it, albeit simpler.

3.1.1 Relation to 2EC$_{\text{size}}$ and multi-2EC$_{\text{cost}}$

Lemma 1 is of independent interest, as it implies that for any 3-edge connected cubic graph G with edge costs $c \in \mathbb{R}^E$, there exists a 2-edge connected spanning subgraph of cost at most $\frac{4}{5}\sum(c_e : e \in E)$ (see Section 2.2 for a detailed explanation of this property). Cubic graphs have $|E| = \frac{3}{2}|V|$. Hence, for $c_e = 1$ for all $e \in E$ this shows that there exists a 2-edge connected spanning subgraph of G with at most $\frac{4}{5}|E| = \frac{4}{5}(\frac{3}{2}|V|) = \frac{6}{5}|V|$ edges, a result previously shown in [4].

Lemma 1 also independently lends support to Conjecture 1. To see this, we first show that the vector $x \in \mathbb{R}^E$ with $x_e = \frac{2}{3}$ for all $e \in E$ is a feasible solution for multi-2EC$^{\text{LP}}_{\text{cost}}$. It is easy to show that x belongs in the polytope of multi-2EC$^{\text{LP}}_{\text{cost}}$, as all x_e's are non-negative, and the minimum size cut in G has cardinality 3, which implies that the sum of the x_e's in that cut is at least 2 (see Section 2.2.3 for the ILP formulation of multi-2EC$^{\text{LP}}_{\text{cost}}$). As such, Conjecture 1 being true would imply that $\frac{6}{5}x = \frac{4}{5}\chi^{E(G)}$ must be a 2EC$_{\text{size}}$ convex combination, which Lemma 1 shows. However, to show similar support for the famous conjecture that states that the integrality gap for the Travelling Salesman Problem is $\frac{4}{3}$ is still an open problem (see [26]).

3.1.2 Proof of $\frac{4}{5}\chi^{E(G)}$ as a 2EC$_{\text{size}}$ Convex Combination

In this section, we provide a proof of Lemma 1, namely that for a cubic 3-edge connected multi-graph $G = (V, E)$, $\frac{4}{5}\chi^{E(G)}$ is a 2EC$_{\text{size}}$ convex combination.

Suppose the contrary, and let G be the counter-example with the smallest number of vertices for which the lemma does not hold. Because G is cubic and 3-edge connected, G is always simple except when $|V| = 2$, where there exists a multi-edge (top of Figure 3.1). Should this not be the case and G would contain a multi-edge or a loop with one endpoint at vertex $v \in V$, then at least one edge incident to v would be a bridge, which would contradict the 3-edge connectivity of G. The lemma can be shown to be true directly for the unique graphs G with $|V| = 2$ and $|V| = 4$ as shown in Figure 3.1 (note that cubic graphs always have an even number of vertices, so $|V| \neq 3$). In the figure, the subgraphs H_i and the corresponding λ_i values for the convex combination that gives $\frac{4}{5}\chi^{E(G)}$ are shown, where bold lines indicate edges in H_i and dotted lines indicate edges in G not in H_i. Thus, without loss of generality, we can assume $|V| > 4$. We consider two cases, based on the

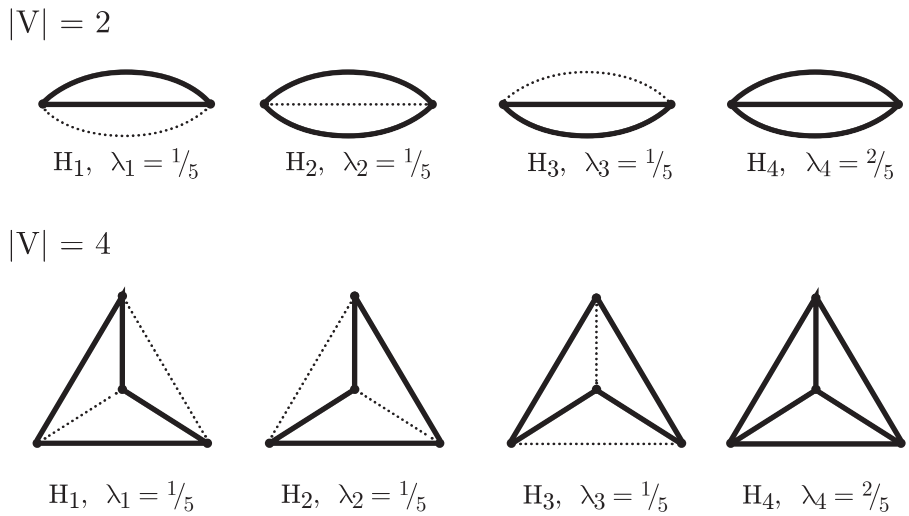


Figure 3.1: Proof of Lemma 1 for $G = (V, E)$, when $|V| \leq 4$.

existence of an essential 3-edge cut.

Case 1. *G has no essential 3-edge cut.*

For any edge $uv \in E$, let the unlabeled adjacent vertices at u be a and b, and the unlabeled adjacent vertices at v be c and d. Since G is 3-edge connected, has no essential 3-edge cut and $|V| > 4$, it follows that a, b, c and d are all distinct. If this were not the case, then there would exist a cycle of length 3 (e.g. cycle auv if $a = d$) which would imply that three vertices (in this case a, u and v) would form an essential 3-edge cut. The vertices neighbour to u and v are illustrated on the left of Figure 3.2, where some incident edges are not shown for vertices a, b, c and d.

We will now proceed by contradiction, and apply a transformation to G to obtain a smaller graph. Removing u and v and their incident edges, and adding edges ab and cd

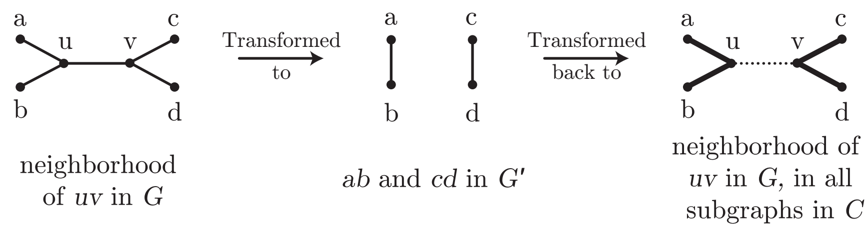


Figure 3.2: Inductive step for Case 1.

yield a new cubic graph $G' = (V', E')$ with fewer vertices than G, as is illustrated in the middle of Figure 3.2, with some edges incident to a, b, c and d omitted. As G did not have any essential 3-edge cut, the removal of a single edge cannot diminish the edge connectivity of the graph, so G' is at least 3-edge connected. Furthermore, G' is smaller than graph G, so the lemma holds for G' by contradiction, i.e. there exists a set of 2-edge connected spanning subgraphs H_i with multipliers λ_i, $i = 1, 2, \ldots, k$ such that an edge e occurs $\frac{4}{5}$ times overall in the 2EC_{size} convex combination, which we label C.

We now extend the 2EC_{size} convex combination C in G' to G. For each 2-edge connected spanning subgraph of G' in C, we extend it by removing edges ab and cd if they are in the subgraph, and adding vertices u and v and edges au, bu, cv and vd. The newly formed subgraphs for G are spanning, because $V(G) = V(G') \cup \{u, v\}$ and vertices u and v are included in all the subgraphs. The subgraphs are also 2-edge connected, because there did not exist a 1-edge cut for the subgraphs in G', and none could have been created by adding 1-paths aub and cvd. The transformation from a subgraph in C to one suitable for G is illustrated on the right of Figure 3.2.

In the newly formed convex combination of incidence vectors of 2-edge connected spanning subgraphs for G, the occurrence of an edge $e \in E$ is

$$z_e = \begin{cases} 0 & \text{if } e = uv, \\ 1 & \text{if } e \in \{ua, ub, vc, vd\}, \\ \frac{4}{5} & \text{otherwise.} \end{cases}$$

Taking in turn all edges in E as edge uv means that we have $m = |E|$ 2EC$_{\text{size}}$ convex combinations, which we will refer to as $\mathbb{M}_e$ for each $e \in E$. An edge $f \in E$ does not occur in $\mathbb{M}_f$, occurs 1 time in $\mathbb{M}_e$ for each of the four edges e adjacent to f, and occurs $\frac{4}{5}$ in the rest of the 2EC$_{\text{size}}$ convex combinations $\mathbb{M}_e$. We now take a convex combination of the m 2EC$_{\text{size}}$ convex combinations $\mathbb{M}_e$, $e \in E$, by multiplying every multiplier λ_i used in these 2EC$_{\text{size}}$ convex combinations by $\frac{1}{m}$. Summing the occurrence of every edge in this new convex combination gives

$$\frac{1}{m}\left(4 + \frac{4}{5}(m-5)\right) = \frac{4}{5}. \tag{3.1}$$

Equation 3.1 states that any edge in the new 2EC$_{\text{size}}$ convex combination has occurence $\frac{4}{5}$, i.e. $\frac{4}{5}\chi^{E(G)}$ for G; hence the lemma is true for G, contradiction.

Case 2. *G has an essential 3-edge cut.*

Notice that the ends of the three edges in the cut must be distinct since G is 3-edge connected and cubic. In this case we contract each shore of the cut to a single vertex, to obtain graphs $G_1 = (V_1, E_1)$ with pseudo-vertex v_1 and $G_2 = (V_2, E_2)$ with pseudo-vertex v_2 (as shown in Figure 3.3). Both G_1 and G_2 are cubic, 3-edge connected and smaller than

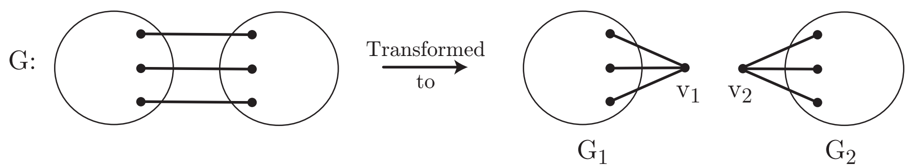


Figure 3.3: Contracting both sides of an essential 3-edge cut of G.

G, so the lemma is true for G_1 and G_2 by hypothesis. Moreover, as each subgraph used in the convex combination is 2-edge connected, pseudo-vertices v_1 and v_2 must always be connected by at least 2 edges to the rest of their respective subgraphs. There are exactly four patterns formed by the occurrence of the edges incident to v_1 and v_2: one pattern with all three edges incident with v_1 (or v_2) present, and then for each edge incident with v_1 (or v_2) one pattern where this edge is missing and the other two present. Furthermore, by hypothesis, each edge occurs $\frac{4}{5}$ times overall in the 2EC convex combination (i.e. the sum of the lambdas for subgraphs containing said edge is $\frac{4}{5}$), which means the sum of the lambdas for a pattern containing only two of the same edges incident to v_1 (or v_2) must be $\frac{1}{5}$. The first pattern occurs $\frac{2}{5}$ of the time in the 2EC convex combination and each of the other patterns occurs $\frac{1}{5}$ of the time. These constant patterns allow us to "glue" (reconnect the edges as they were before the inductive step) the subgraphs for G_1 and G_2 together, in such a way that identical patterns at v_1 and v_2 are matched. The process to transform the 2-edge connected spanning subgraphs in the convex combination for G_1 and G_2 is illustrated in Figure 3.4, where the occurence of a pattern for v_1 and v_2 is indicated on the left. The gluing process results in a 2EC convex combination for $\frac{4}{5}\chi^{E(G)}$ that shows the lemma is true for G, which gives a contradiction. □

We can also easily extend the results of Lemma 1 to non-cubic 3-edge connected graphs.

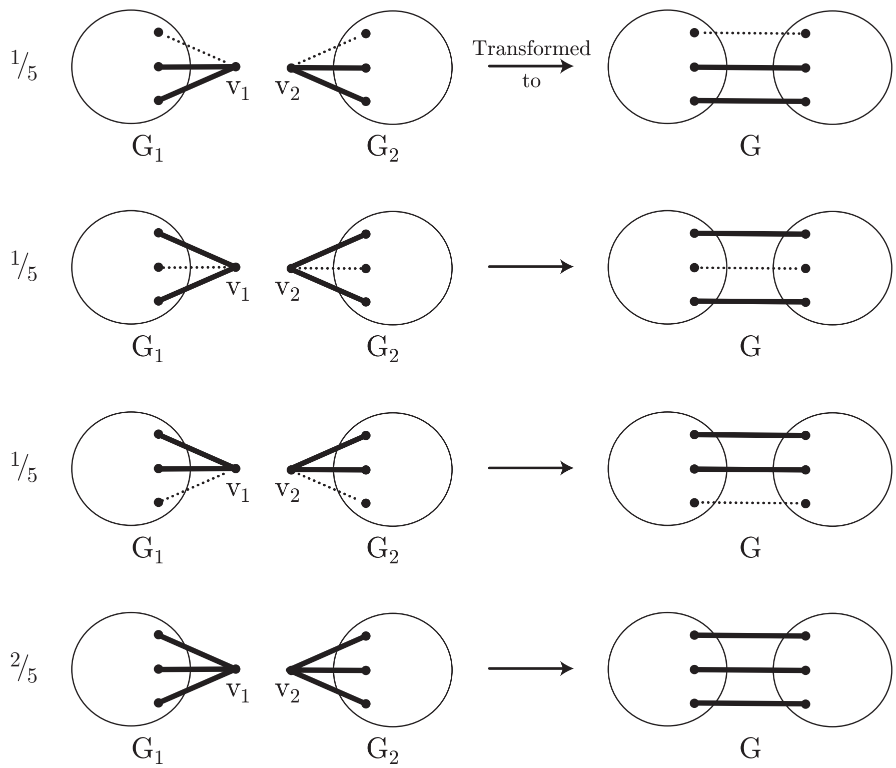


Figure 3.4: “Gluing” the 2-edge connected spanning subgraphs in the $2EC_{\text{size}}$ convex combinations for G_1 and $G1$.

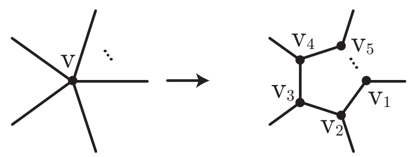


Figure 3.5: Replacing a vertex of degree 5 with a 5-cycle.

Theorem 3. *Let $G = (V, E)$ be a 3-edge connected graph. Then $\frac{4}{5}\chi^{E(G)}$ is a $2EC_{size}$ convex combination.*

Proof. Without loss of generality, we assume that there are no cut vertices, as two or more components induced by a cut-vertex can always be solved separately and still yield a 2EC_{size} convex combination for G.

Because we are assuming that G is 3-edge connected, we know that all vertices have degree at least 3. Replace all vertices of degree $k > 3$ by a k-cycle, as shown in Figure 3.5 for a vertex v of degree 5. No vertex is a cut-vertex in G, which implies that the resulting graph, labelled G', is still 3-edge connected, and is also cubic. Therefore, the theorem is true for G' by Lemma 1: this means that there exists a convex combination of incidence vectors of 2-edge connected spanning subgraphs of G' that select each edge exactly $\frac{4}{5}$ times. We now use this 2EC_{size} convex combination to go back to G and conclude the proof. Shrinking the k-cycles introduced before into the vertices of graph G in each subgraph gives a convex combination of incidence vectors of 2-edge connected spanning subgraphs with the same properties. Thus the theorem is true for G. $\square$

3.2 $\frac{7}{9}\chi^{E(G)}$ as a 2EC_{size} Convex Combination

In this section, we demonstrate that for a cubic 3-edge connected graph $G = (V, E)$, $\frac{7}{9}\chi^{E(G)}$ is a 2EC_{size} convex combination. This strengthening of Lemma 1 is due to the following result, which leads to a new reduction operation specific to essentially 4-edge connected cubic simple graphs.

Lemma 2. *Given an essentially 4-edge connected cubic simple graph $G = (V, E)$ with $|V| > 6$ and edges $au, uv, vc, vd \in E$, no two subsets $S, S' \subset V$ different than $\{u, v\}$ exist such that $au, vd \in \delta(S)$, $au, vc \in \delta(S')$, and $|\delta(S)| = |\delta(S')| = 4$.*

Lemma 2 implies that an essentially 4-edge connected cubic simple graph can be transformed into a smaller graph with the same properties, and is used in Subsection 3.2.2.

3.2.1 A Reduction Operation for Essentially 4-Edge Connected Cubic Simple Graphs

Here we provide a proof of Lemma 2. Suppose that $G = (V, E)$ with $|V| > 6$, edges $au, uv, vc, vd \in E$ and subsets $S, S' \subset V$ is the smallest counter-example to the lemma.

Let edges incident to u and v be as illustrated in Figure 3.6, with the unlabelled vertex neighbour to u called b. Some edges incident to the unlabelled vertices are not shown.

G is simple and essentially 4-edge connected, and $|V| > 6$, which implies that labelled vertices in Figure 3.6 are distinct, as there would otherwise exist an essential 3-edge cut. Without loss of generality, assume that $u \in S$ and $u \in S'$ (if not, take the complement

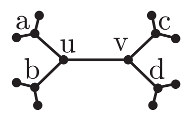


Figure 3.6: Edges incident to u and v in G.

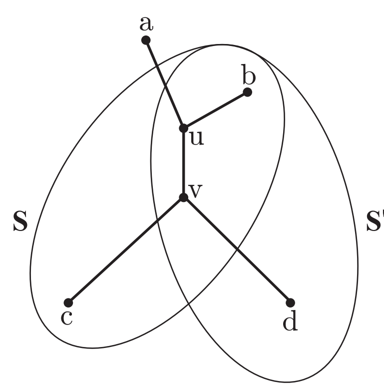


Figure 3.7: Disposition of the vertices adjacent to u and v in S and S'.

of the set). The cuts $\delta(S)$ and $\delta(S')$ are minimal and G is cubic, which implies that two adjacent edges may not be in the same cut. Therefore $v, b, c \in S$ and $v, b, d \in S'$. We already know that $a, d \notin S$ and $a, c \notin S'$. Figure 3.7 shows this disposition of the vertices in the cut.

Throughout this proof, we exploit the symmetric submodularity property of the function $|\delta(\cdot)|$, which states that

$$\begin{aligned} |\delta(S)| + |\delta(S')| &\geq |\delta(S \cup S')| + |\delta(S \cap S')|, \\ 8 &\geq |\delta(S \cup S')| + |\delta(S \cap S')|. \end{aligned} \tag{3.2}$$

Because $u, v \in S \cap S'$, then $|S \cap S'| > 1$; similarly because $a \notin S \cup S'$, then $|\overline{S \cup S'}| > 0$. It follows that $|\delta(S \cup S')| \geq 3$ and $|\delta(S \cap S')| \geq 4$ (since G has no essential 3-edge cut). There exist two cases, when $|\delta(S \cup S')| = 3$ and $|\delta(S \cap S')| = 4$ or 5 and where $|\delta(S \cup S')| = 4$, because the cardinality of the cuts are restricted by (3.2).

Case 1. $|\delta(S \cup S')| = 3$ *and* $|\delta(S \cap S')| = 4$ *or* 5*.*

Graph G is simple and essentially 4-edge connected, which implies that one of the shores of $\delta(S \cup S')$ must consist of a single vertex, as the cardinality of this cut is 3: vertices u and v belong to $S \cup S'$, so $\overline{S \cup S'} = \{a\}$.

Any partition of a cubic graph has an even number of odd parts, because cubic graphs have an even number of vertices. We use this parity to show that $|\delta(S \cap S')| = 5$. The partition $\{S \setminus S', S' \setminus S, S \cap S', \overline{S \cup S'}\}$ (displayed in Figure 3.8) has $|\overline{S \cup S'}|$ odd, which implies that at least one of $S \setminus S'$, $S' \setminus S$ and $S \cap S'$ also has an odd number of vertices. The graph G is cubic, which implies that for any subset $S \subseteq V$ of vertices, $\delta(S)$ is odd if and only if $|S|$ is odd. Given that $|\delta(S)| = |\delta(S')| = 4$, then $|S|$ and $|S'|$ must be even. This implies that the parity of $|S' \setminus S|$ and $|S \cap S'|$ must be the same; similarly, the parity of $|S \setminus S'|$ and $|S \cap S'|$ must also be the same. Therefore, either $|S \setminus S'|$, $|S' \setminus S|$ and $|S \cap S'|$ are all odd, or all even, to respect parity. But from the above, they cannot all be even. Thus, $|S \cap S'|$ is odd, which means that $|\delta(S \cap S')|$ must be odd as well. Therefore,

$$|\delta(S \cap S')| = 5. \tag{3.3}$$

All edges at u and v are known: let us examine an edge bx, for a vertex $x \in V$. We

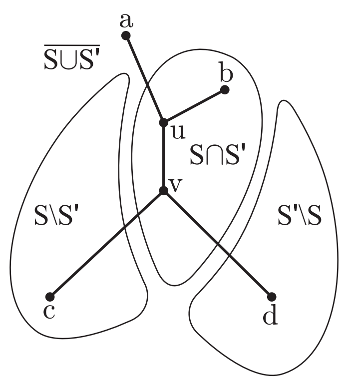


Figure 3.8: Partition of G to show the parity of parts, in Case 1.

already know of 3 edges in $\delta(S \cap S')$, which means that there are exactly 2 others edges e and f in the cut. Assume that $x \in S \cap S'$: this means that bu, e and f are in an essential 3-edge cut, because such a cut contains b and x. Such a cut contradicts the fact that G is essentially 4-edge connected. Therefore, $x \notin S \cap S'$, which immediately implies that $|S \cap S'| = 3$. We have that

$$4 = |\delta(S)| = |\delta(S \setminus S')| + |\delta(S \cap S')| - 2|E[S \setminus S' : S \cap S'])|, \tag{3.4}$$

$$4 = |\delta(S')| = |\delta(S' \setminus S)| + |\delta(S \cap S')| - 2|E[S' \setminus S : S \cap S']|. \tag{3.5}$$

We also have that $ba \notin E$ or else there would be an essential 3-edge cut $\delta(\{u, b, a\})$, which means that from the five edges in $\delta(S \cap S')$, four link to either $S \setminus S'$ or $S' \setminus S$:

$$4 = |E[S' \setminus S : S \cap S']| + |E[S \setminus S' : S \cap S']|. \tag{3.6}$$

From the premise $|V| > 6$, it is implied that one of $S \setminus S'$ and $S' \setminus S$ contains more than one vertex. Without loss of generality, let it be $S \setminus S'$. Because G has no essential 3-edge cut and $|S \setminus S'| > 1$, we infer that.

$$4 \leq |\delta(S \setminus S')|, \tag{3.7}$$

$$3 \leq |\delta(S' \setminus S)|. \tag{3.8}$$

The symmetric submodularity property ensures that

$$\begin{aligned} |\delta(S)| + |\delta(S')| &\geq |\delta(S \setminus S')| + |\delta(S' \setminus S)| \\ 8 &\geq |\delta(S \setminus S')| + |\delta(S' \setminus S)|. \end{aligned} \tag{3.9}$$

We will now use algebraic manipulations to show a contradiction. We subtract (3.5) from (3.4) and add (3.6) twice:

$$8 = |\delta(S \setminus S')| - |\delta(S' \setminus S)| + 4|E[S' \setminus S : S \cap S']|. \tag{3.10}$$

In equation (3.10), $|\delta(S \setminus S')| - |\delta(S' \setminus S)|$ must be a multiple of 4. Because inequalities (3.7), (3.8) and (3.9) restrict the values of $|\delta(S \setminus S')|$ and $|\delta(S' \setminus S)|$, their difference must be zero, i.e.

$$|\delta(S \setminus S')| = 4. \tag{3.11}$$

Equation (3.10) is simplified to

$$2 = |E[S \setminus S' : S \cap S']|. \tag{3.12}$$

We now conclude by substituting (3.3), (3.11) and (3.12) in (3.4):

$$\begin{aligned} 4 &= |\delta(S \setminus S')| + |\delta(S \cup S')| - 2|E[S \setminus S' : S \cap S']| \\ &= 4 + 5 - 2 \times 2 \\ &= 5, \end{aligned}$$

which gives a contradiction.

Case 2. *$|\delta(S \cup S')| = 4$ and $|\delta(S \cap S')| = 4$.*

Since $|\delta(S \cap S')| = 4$, it follows that $|S \cap S'|$ is even, because G is cubic. So there is at least one more vertex w in $S \cap S'$. Since G is 3-edge connected, at least one edge incident with w is in $\delta(S \cap S')$. However, $vc, vd, au \in \delta(S \cap S')$ already, so we only have one other edge e in the cut (as shown in Figure 3.9). Therefore w is one end of e. This means that $\{ub, e\}$ is a 2-edge cut in G, which is a contradiction. $\square$

3.2.2 Proof of $\frac{7}{9}\chi^{E(G)}$ as a $2\mathrm{EC}_{\mathbf{size}}$ Convex Combination

This result is a strenghtening of Lemma 1 and both share similar proofs, starting with the existence of two cases which depend on the presence of an essential 3-edge cut. We will

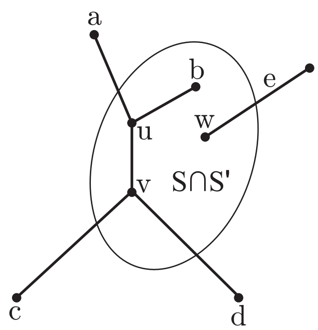


Figure 3.9: Disposition of the vertices adjacent to u and v when $|\delta(S \setminus S')| = 4$.

now modify the case when there are no essential 3-edge cut by making use of Lemma 2. By doing so, a more effective transformation can be used which results in an economy of edges.

Theorem 4. *Given a 3-edge connected cubic graph $G = (V, E)$, the vector $y \in \mathbb{R}^E$ defined by $y_e = \frac{7}{9}$ for all $e \in E$ is a convex combination of incidence vectors of 2-edge connected spanning subgraphs H_i, $i = 1, 2, \ldots, k$.*

Proof

Let $G = (V, E)$ be the smallest counter-example for which the theorem does not hold. There are only three 3-edge connected essentially 4-edge connected graphs with 6 vertices or less, on which the theorem can be shown to be true directly, as demonstrated in Figure 3.10, where bold lines represent edges in the subgraph, and dotted lines represent edges in G not in the subgraph. In the figure, for each such graph G, the subgraphs H_i and the corresponding λ_i values for the required convex combination are shown. The smallest cubic 3-edge connected simple graph which is not essentially 4-edge connected has $|V| = 6$, which means that either

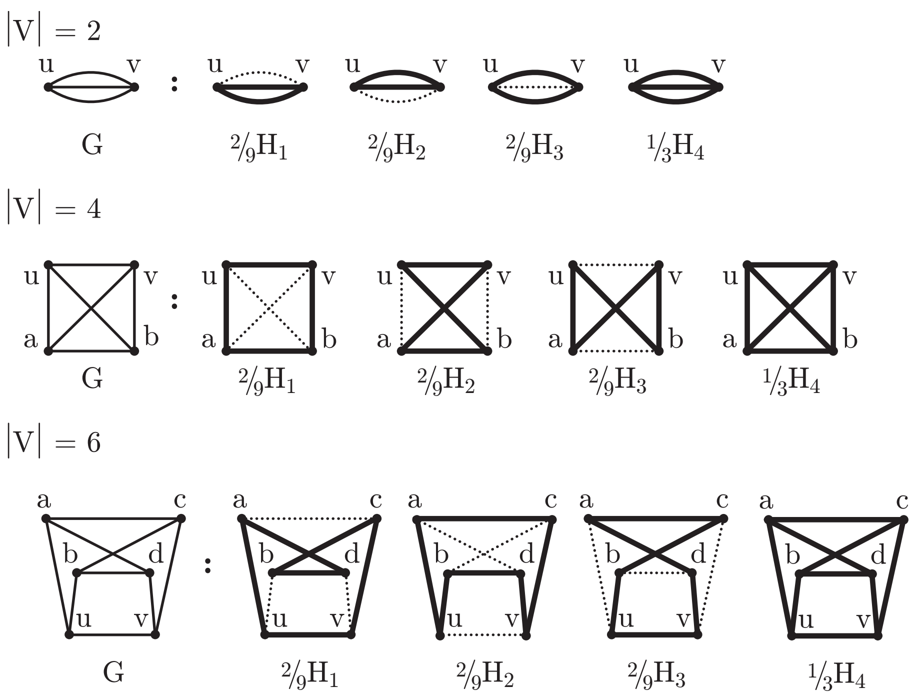


Figure 3.10: Base cases for Theorem 4. In the subgraphs H_i, bold lines indicate the edges in the subgraph and dotted lines indicate the edge omitted in the subgraph.

1. G is essentially 4-edge connected and $|V| > 6$,

2. G has an essential 3-edge cut and $|V| \geq 6$.

Case 1. *G is essentially 4-edge connected and $|V| > 6$.*

We pick an arbitrary edge $uv \in E$, and we label the other adjacent vertices of u as a and b, and the other adjacent vertices at v as c and d. We will use this edge to create the vector $z \in \mathbb{R}^E$ as a convex combination of incidence vectors of 2-edge connected spanning subgraphs, where for all $e \in E$,

$$z_e = \begin{cases} 1 & \text{if } e = uv, \\ \frac{1}{2} & \text{if } e \in \{ua, ub, vc, vd\}, \\ \frac{8}{9} & \text{if } e \neq uv \text{ and } e \text{ adjacent to one of } ua, ub, vc \text{ or } vd, \\ 1 & \text{if } e \neq uv \text{ and } e \text{ adjacent to two of } ua, ub, vc \text{ or } vd, \\ \frac{7}{9} & \text{otherwise.} \end{cases}$$

Graph G has no essential 3-edge cut and $|V| > 6$, which implies that a, b, c and d are distinct. Lemma 2 states that there does not exist two subsets $A, B \subset V$ different from $\{u, v\}$ such that $au, vc \in \delta(A)$, $au, vd \in \delta(B)$ and $|\delta(A)| = |\delta(B)| = 4$. Thus, assuming that $\delta(B)$ is an essential 4-edge cut for some B automatically means that $\delta(A)$ is not a 4-edge cut for any such A, and vice-versa. Without loss of generality, assume that edges au and vc are not in an essential 4-edge cut together, other than $\delta(\{u, v\})$, i.e. $\delta(A) > 4$.

By Lemma 2 again, there does not exist two subsets $C, D \subset V$ different from $\{u, v\}$ such that $bu, vd \in \delta(C)$, $bu, vc \in \delta(D)$ and $\delta(C) = \delta(D) = 4$. Knowing that $|\delta(A)| > 4$ for all A such that $au, vc \in \delta(A)$ and $A \neq \{u, v\}$ implies that $|\delta(C)| > 4$ for all C such that $bu, vd \in \delta(C)$ and $C \neq \{u, v\}$, which means that edges bu and vd are not in an essential 4-edge cut together.

Armed with these facts, we create graph G_1 by removing edges au and vc, and G_2 by removing edges bu and vd. In both graphs, any vertex s of degree 2 with incident edges st and sr is removed, and the edge sr is added. Because $\delta(A) > 4$, $bd, ac \notin E$, no multi-edge are created. This situation is illustrated in the first part of Figure 3.11; the second part of the figure illustrates the same situation, but when $|\delta(A)| = 4$ and $|\delta(B)| > 4$, for any B such that $au, vd \in \delta(B)$ and $B \neq \{u, v\}$.

The astute reader may ask what happends if an edge links one of a, b, c or d: assume that such an edge $rs \in E$ exists, for $\{r, s\} \subset \{a, b, c, d\}$. It is immediately apparent that r and s cannot both be neighbor to the same vertex, either u or v, otherwise there would exist an essential 3-edge cut, e.g. $\{rsu\}$. Without loss of generality, let r be adjacent to u and let s be adjacent to v. For simplicity, let $r = a$ and $s = c$. Cut $\delta(\{a, c\})$ is an essential 4-edge cut and $ua, vc \in \delta(\{a, c\})$. By Lemma 2 and because $|V| > 6$, edges au and vd are not in an essential 4-edge cut together, and edges bu and vc are not in an essential 4-edge cut together, which means that the transformation is valid.

Both G_1 and G_2 are 3-edge connected and cubic, because G was simple and essentially 4-edge connected and the two edges removed in G were not in an essential 4-edge cut together.

The theorem holds for G_1 and G_2, as they are both smaller than G, and gives two convex combinations of incidence vectors of 2-edge connected spanning subgraphs C_1 and

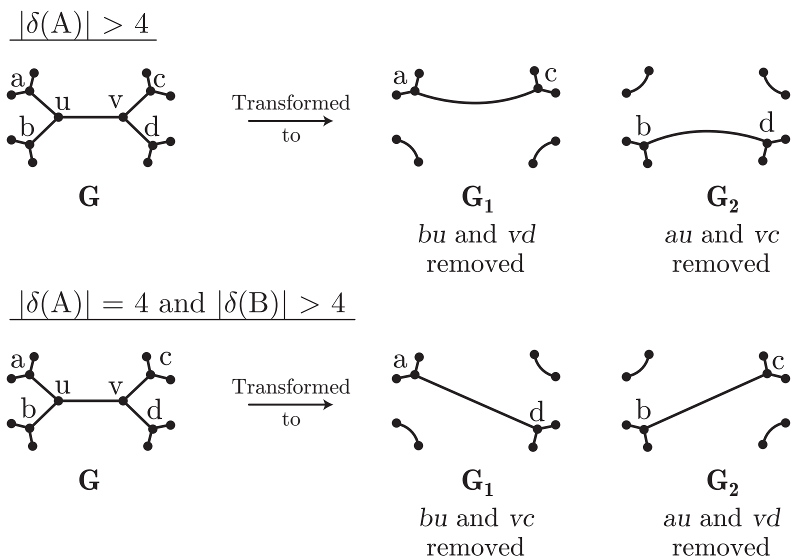


Figure 3.11: Partial transformation of G into G_1 or G_2 in Case 2 of Theorem 4.

C_2 for G_1 and G_2, respectively. Note that an edge in the convex combination is present $\frac{7}{9}$ times overall in subgraphs. We will now modify the subgraphs in C_1 and C_2 to form sets C_1' and C_2' of subgraphs for G in the following way: assume that edges s and t were the two edges that we removed to transform G into G_i, then any 2-edge connected spanning subgraph for G_i is 2-edge connected and spanning for G with the same edge selection, save for edges adjacent to s and t which we always select, and s and t which we always omit. This is illustrated on the left side of Figure 3.12, where bold lines represent edges in the subgraph, dotted line represent edges omitted in the subgraph and dashed lines represent edges which may or may not be in the subgraph. In the figure, edges incident to a, b, c and d, but not adjacent to uv are shown as distinct, for simplicity.

We now take $\frac{1}{2}C_1' + \frac{1}{2}C_2'$ to obtain a 2EC convex combination for G. The occurence of each edge in this convex combination is shown in the rightmost part of Figure 3.12. Note that this is the occurrence for the case that no edges with both endpoints in the set $\{a, b, c, d\}$ exist. In the event such edges exist, then these edges have an occurrence of 1, which does not affect the end result, as is shown later.

The convex combination $\frac{1}{2}C_1' + \frac{1}{2}C_2'$ is symmetrical (see the rightmost part of Figure 3.12): therefore, repeating all the steps from Case 1, taking every edge of G in turn as edge uv gives $|E|$ convex combinations which we label F_i, for $i = 1, \ldots, |E|$. We build a new convex combination by setting the lambda value $\lambda_{F_i} = \frac{1}{|E|}$ for each F_i, and conclude that it selects every edge $\leq \frac{7}{9}$ times overall. We shall show this property in more detail here for an edge $yz \in E$:

- edge yz is treated as edge uv in exactly one of the $|E|$ convex combinations. It then has an occurrence of 1.
- there are no doubled edges, which means that yz is adjacent to uv in exactly 4 of the $|E|$ convex combinations. It then has an occurrence of $\frac{1}{2}$.
- Because G is cubic, there are exactly 8 ways to find an edge yz exactly one edge away from uv, which are handled differently if yz and uv are in a 4-cycle together or not.

  Assume that yz and uv are not in a 4-cycle r times out of $|E|$, where $r \leq 8, r \in \mathbb{N}$ (the graph is cubic and simple, so there can be at most 8 such distinct edges for yz). It then has an occurrence of $\frac{8}{9}$.

  If yz and uv are in a 4-cycle (i.e. $yu, zu \in E$), then yz is one edge away from uv both by yu and zu. This happens t times over the $|E|$ convex combinations and yz

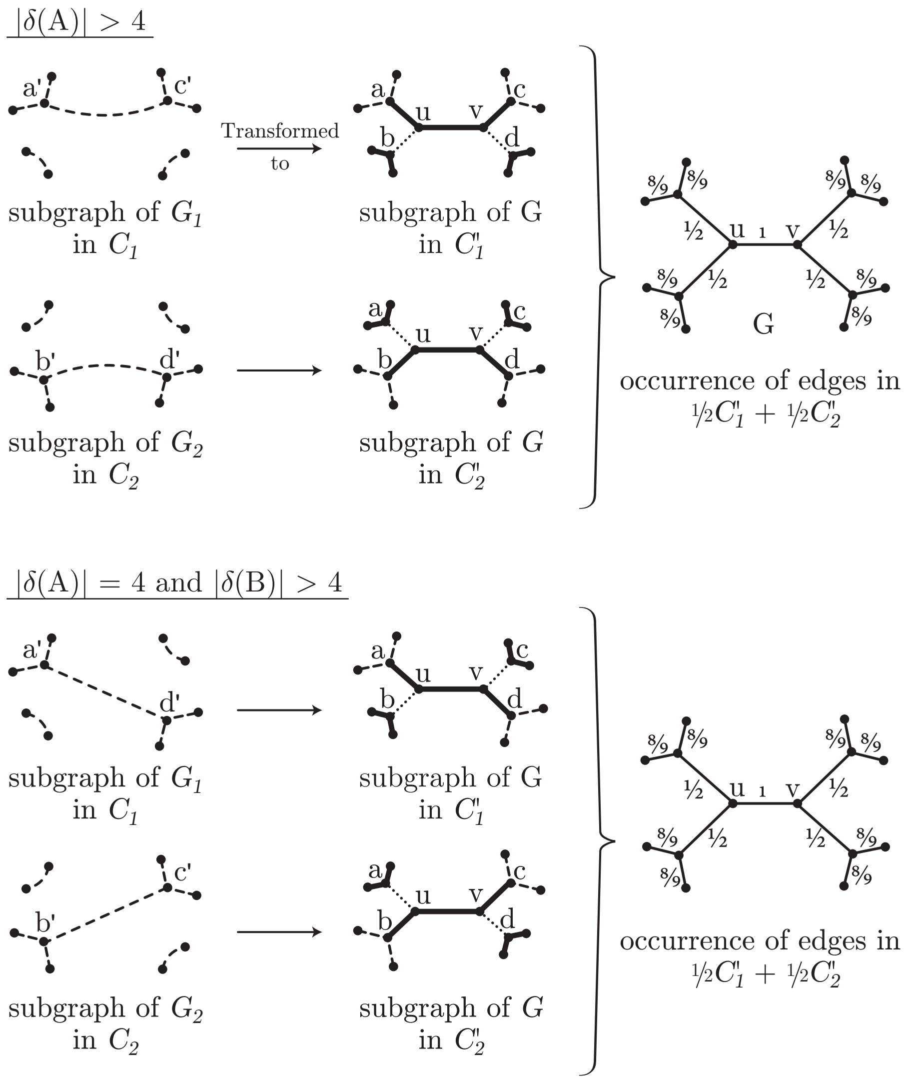


Figure 3.12: Transformation of the 2-edge connected spanning subgraphs for G_1 and G_2 into 2-edge connected spanning subgraphs for G.

then has an occurrence of 1. Because there are 8 ways to be one edge away from uv, and any 4-cycle with yz and uv uses two of those ways,

$$2t + r = 8, \text{ for } t \in \mathbb{N}. \tag{3.13}$$

- The occurrence of yz in the other convex combinations is $\frac{7}{9}$.

The average occurrence of an edge over the $|E|$ convex combinations is

$$\begin{aligned}
&\leq \frac{1}{|E|}\Big(1 + t + 4 \times \frac{1}{2} + r \times \frac{8}{9} + (|E| - r - t - 5) \times \frac{7}{9}\Big) \\
&= \frac{1}{|E|}\Big(\frac{2t}{9} + \frac{r}{9} - \frac{8}{9} + \frac{7|E|}{9}\Big) \\
&= \frac{7}{9} + \frac{1}{9|E|}(2t + r - 8). \qquad (3.14)
\end{aligned}$$

Since $2t+r = 8$, by (3.13), it follows from (3.14) that the average edge occurrence over the $|E|$ convex combinations is at most $\frac{7}{9}$. The convex combination which results from giving each convex combination a weight of $\frac{1}{|E|}$ selects every edge $\frac{7}{9}$ times. If an edge is selected less than $\frac{7}{9}$ of the time, we add it back arbitrarily to have it selected exactly $\frac{7}{9}$ of the time. The theorem holds.

Case 2. *G has an essential 3-edge cut C and $|V| \geq 6$.*

Notice that the ends of the three edges in C must be distinct because G is 3-edge connected. In this case we contract each shore of C to a single pseudo-vertex, to obtain graphs $G_1 = (V_1, E_1)$ with pseudo-vertex v_1 and $G_2 = (V_2, E_2)$ with pseudo-vertex v_2 (as

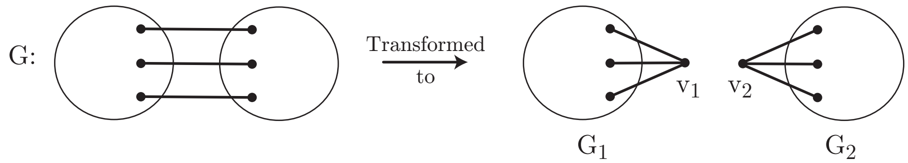


Figure 3.13: Contracting both sides of an essential 3-edge cut of G.

shown in Figure 3.13). Both G_1 and G_2 are smaller than G so the theorem holds for G_1 and G_2. Moreover the patterns formed by the occurrence of the edges incident to v_1 and v_2 are unique and identical in the subgraphs in the corresponding 2EC convex combination. More specifically, as a minimum of two edges incident to v_1 (or v_2) must be included each 2-edge connected spanning subgraph, there are four possible outcomes, or patterns. The first is that all three edges incident to v_1 (or v_2) are included in the subgraph. The second, third and fourth, which occur exactly $\frac{2}{9}$ of the time each, is that one of the incident edges to v_1 (or v_2) will not be in the subgraph, on both sides of the cut, and this is true for each of the three incident edges. These constant patterns allow us to glue (reconnect the edges as there were before the inductive step) the subgraphs for G_1 and G_2 together, in such a way that identical patterns at v_1 and v_2 are matched and the resulting subgraphs are 2-edge connected. Figure 3.14 displays the patterns formed by the occurrence of the edges incident to v_1 and v_2 and their occurence, on the left. The figure also illustrates the process of gluing the 2-edge connected spanning subgraph in the convex combinations for G_1 and G_2 .

This results in a convex combination that shows that the theorem holds for G, which gives a contradiction. □

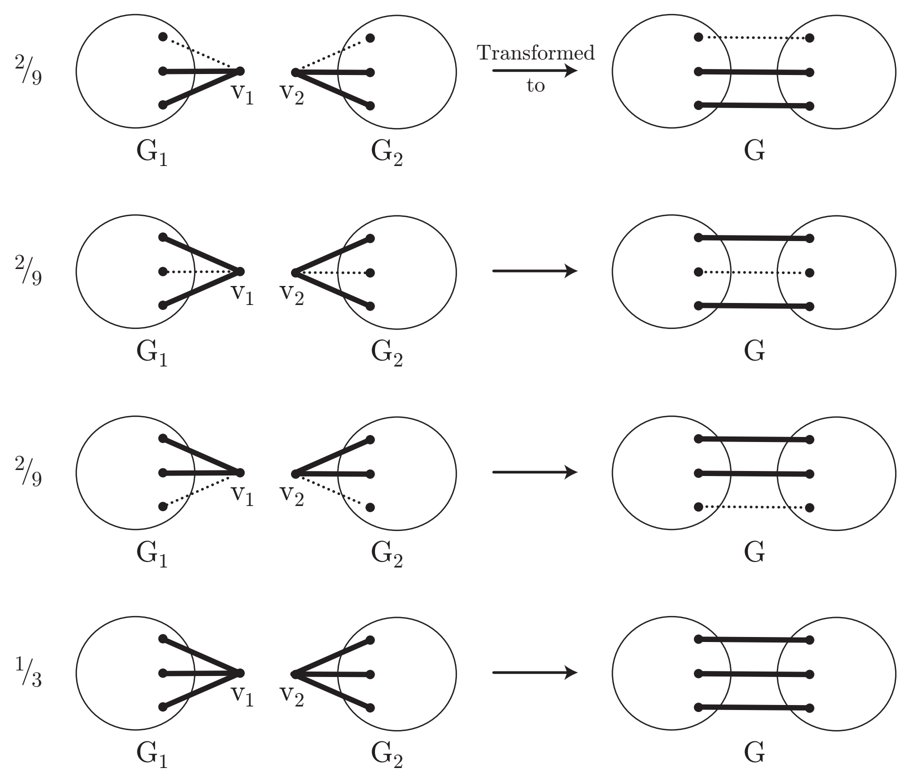


Figure 3.14: "Gluing" the 2-edge connected spanning subgraphs in the $2EC_{\text{size}}$ convex combinations for G_1 and $G1$.

3.3 A New Bound on the Unit Integrality Gap

The use of a convex combination in Theorem 4 gives the following result.

Corollary 1. *Given a 3-edge connected cubic graph $G = (V, E)$ with $n = |V|$, there exists a 2-edge connected spanning subgraph with at most $\frac{7}{6}n$ edges.*

In other words, we show that the unit integrality gap for 2EC is bounded above by $\frac{7}{6}$ for 3-edge connected cubic graphs, which improves upon Boyd, Iwata and Takazawa's upper bound of $\frac{6}{5}$ [4]. Our methods have an exponential number of steps and are not polynomial and thus, do not result in an approximation algorithm. Nevertheless, they give hope that a $\frac{7}{6}$-approximation algorithm exists, which would improve upon the $\frac{6n}{5}$-approximation algorithm by Boyd, Iwata and Takazawa [4], and the $\frac{7n}{6}$-approximation algorithm when G is bipartite, by Takazawa [28].

Chapter 4

Improved Bounds for αmulti-2EC$_{\text{cost}}$

This chapter is dedicated to the idea of using the structure of solutions for multi-2EC$_{\text{cost}}^{\text{LP}}$ and the concept of convex combination to obtain improved bounds for αmulti-2EC$_{\text{cost}}$. We focus our efforts on the family of cost functions introduced in Subsection 1.1.3, which was characterized by having $x_e \in \{0, \frac{1}{2}, 1\}$ and half-edges forming disjoint 3-cycles, linked by 1-paths, as that family appears to give the largest integrality gap for multi-2EC$_{\text{cost}}^{\text{LP}}$. Section 4.1 elaborates on the choice of half-triangle solution, and on its impact. In Section 4.2, we successfully show that the conjecture αmulti-2EC$_{\text{cost}}^{\text{LP}} = \frac{6}{5}$ is true for any cost functions optimized at half-triangle solutions.

4.1 Studying Half-Triangle Solutions

In this chapter, we show that Conjecture 1 is true for any cost function optimized at a half-triangle solution. More specifically, we show that for any half-triangle solution x^* and any cost function $c \geq 0$, there exists a solution of multi-2EC$_{\text{cost}}$ of cost at most $\frac{6}{5}cx^*$, which implies that αmulti-2EC$_{\text{cost}} = \frac{6}{5}$ for any cost function optimized at a half-triangle solution. Note that previously, $\frac{4}{3}$ was known, as Carr and Ravi [5] showed that for any degree-tight half-integer solution x^* and any cost function $c \geq 0$, there exists a solution of multi-2EC$_{\text{cost}}$ of cost at most $\frac{4}{3}cx^*$.

4.2 $\frac{6}{5}x^*$ is a multi-2EC$_{\text{cost}}$ Convex Combination

Our main theorem for the chapter is the following.

Theorem 5. *Let $x^* \in \mathbb{R}^E$ be a half-triangle solution. Then $\frac{6}{5}x^*$ is a multi-2EC$_{cost}$ convex combination.*

This theorem implies that for any non-negative cost vector $c \in \mathbb{R}^E$ optimized at a half-triangle solution x^*, we have $\frac{6}{5}cx^* = \sum_{i=1}^{j} \lambda_i c\chi^{E(H_i)}$ for 2-edge connected spanning multi-graphs H_i, with $\lambda_i \in \mathbb{R}_{\geq 0}$, $i = 1, 2, \ldots, j$, $\sum_{i=1}^{j} \lambda_i = 1$. This implies that for at least one of the H_i, $c\chi^{E(H_i)} \leq \frac{6}{5}cx^* = \frac{6}{5}\text{OPT}(\text{multi-2EC}^{\text{LP}}_{\text{cost}})$ (see Section 2.2 for details of this property of convex combinations). Since $\text{OPT}(\text{multi-2EC}_{\text{cost}}) \leq c\chi^{E(H_i)}$, it follows that $\frac{\text{OPT}(\text{multi-2EC}_{\text{cost}})}{\text{OPT}(\text{multi-2EC}^{\text{LP}}_{\text{cost}})} \leq \frac{6}{5}$ for such cost functions. As there exists a family of half-triangle solutions which show αmulti-2EC$_{\text{cost}} \geq \frac{6}{5}$ asymptotically, as shown in Figure 1.1a [1], we obtain the following corollary to Theorem 5.

Corollary 2. *The integrality gap αmulti-2EC$_{cost}$ is $\frac{6}{5}$ when restricted to cost functions optimized at half-triangle solutions.*

Proof Preliminaries for Theorem 5

Note that in order to prove Theorem 5, we need to have subgraphs in the multi-2EC$_{\text{cost}}$ convex combination in which the edges of the half-triangles are used quite infrequently; in fact at least $\frac{1}{5}$ of the time, we need to use at most one edge of each triangle in the subgraphs while maintaining 2-edge connectivity, a difficult task to achieve. To guide us in this task, we found it useful to first find a 2EC$_{\text{size}}$ convex combination for the simpler graph which is obtained from the half-triangle graph G by shrinking the half-triangles and replacing the 1-paths by single edges. We refer to a graph obtained from G in this manner as G's corresponding cubic graph. Notice that such graphs are 2-edge connected and cubic, as they are obtained from the supportg graphs of multi-2EC$_{\text{cost}}$ solutions.

In this section we prove Theorem 5, namely that for any half-triangle solution x^*, $\frac{6}{5}x^*$ can be expressed as a multi-2EC$_{\text{cost}}$ convex combination. Lemma 1 will be instrumental here, as it guides us in the task of obtaining the correct occurrences of half-triangle edges and 1-paths in the multi-2EC$_{\text{cost}}$ convex combination for $\frac{6}{5}x^*$. Unfortunately, this lemma only holds for 3-edge connected graphs, and thus it cannot be used in the way we would like in the situation where our half-triangle graph G has a 2-edge cut across two 1-paths.

One possible idea that could be used to overcome this problem, for which we only sketch the details here, would be to proceed inductively. We could split G across the 2-edge cut to obtain two half-triangle graphs by removing the two 1-paths in the 2-edge cut, and replacing them with a 1-edge on each side of the cut. In this way we would obtain two smaller half-triangle solutions x' and x''. We could then find the multi-2EC$_{\text{cost}}$ convex combination for $\frac{6}{5}x'$ and $\frac{6}{5}x''$ inductively, and then try to glue these convex combinations together. This almost works, except that under some circumstances we may not be able to guarantee 2-edge connectivity in the resulting glued subgraphs.

To deal with this problem, we begin this section by proving a stronger result than what we need. We prove this result for the special case of half-triangle graphs in which all the 1-paths consist of a single edge, which we call *cubic half-triangle graphs*. Note that this stronger result is only true in this simpler cubic case, which is why we will prove it separately as a lemma before our proof of the main theorem. In this stronger result, we specify a special 1-edge p which will occur less than the other 1-edges in the multi-2EC$_{\text{cost}}$ convex combination. This special edge p will be instrumental in guaranteeing that the subgraphs obtained in the gluing process described above for 2-edge cuts remain 2-edge connected.

Definition 1. *A cubic half-triangle graph $G = (V, E)$ and a specified 1-edge $p \in E$ have property Q (denoted $Q(G, p)$) if the vector $z^* \in \mathbb{R}^E$ defined by*

$$z^*_e = \begin{cases} \frac{3}{5} & \text{if } e \text{ is a half-edge of } G, \\ \frac{4}{5} & \text{if } e = p, \\ \frac{6}{5} & \text{otherwise,} \end{cases}$$

is a multi-2EC$_{cost}$ convex combination in which all of the 2-edge connected spanning multi-subgraphs use one or zero copies of the edge p and the half-edges, and either one or two copies of all other 1-edges.

Lemma 3. *Property $Q(G, p)$ holds for all cubic half-triangle graphs $G = (V, E)$ and any 1-edge $p \in E$ not in a 2-edge cut in G.*

Proof

Any half-triangle graph is 2-edge connected, by definition: as a half-triangle graph is the support graph of a multi-2EC$_{\text{cost}}$ solution, which is constrained by the inequalities of the LP (2.3) for multi-2EC$_{\text{cost}}$, then every edge cut to that graph must have cardinality greater than or equal to 2. We consider two cases, depending on the existence of a 2-edge cut in the graph.

Case 1. *Graph G has no 2-edge cut.*

Let $G' = (V', E')$ be the graph obtained from G by shrinking each half-triangle to a pseudo-vertex. Graph G' is cubic and 3-edge connected, therefore Lemma 1 holds, and yields a 2EC$_{\text{size}}$ convex combination for G' with an edge occurrence of $\frac{4}{5}$ for all edges. Let the subgraphs in this convex combination be H'_i, $i = 1, 2, \ldots, k$ with multipliers $\lambda'_1, \lambda'_2, \ldots, \lambda'_k$.

For each subgraph H'_i in the 2EC$_{\text{size}}$ convex combination for G', the half-triangles (previously contracted to pseudo-vertices) will now be expanded and omitted 1-edges e, $e \neq p$, will be doubled. We will add half-edges to each expanded H'_i in such a way that we create a multi-2EC$_{\text{cost}}$ convex combination for the original half-triangle graph G that gives the occurrence for each edge required by the lemma. To accomplish this, for each triangle T in each subgraph, we will add half-edges in patterns, where each pattern is used a fraction of the time (either $\frac{1}{2}$ or $\frac{1}{3}$). To facilitate this, we simply assume that we start with a new 2EC$_{\text{size}}$ convex combination of G' which contains six copies of each subgraph H'_i, where each copy has a coefficient of $\frac{\lambda'_i}{6}$.

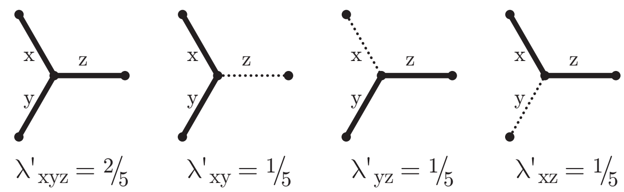


Figure 4.1: The occurence of edges x, y and z in G'.

Now consider any triangle T in G and let its incident edges be x, y and z. In the 2EC_{size} convex combination created for G', each subgraph H_i' will have all three or just two of these edges because each 2-edge connected spanning subgraph of G' must use at least two edges at every vertex of G'. Let

$$\begin{aligned}
\lambda'_{xyz} &= \sum(\lambda'_i : \{x, y, z\} \subseteq E(H'_i)),\\
\lambda'_{xy} &= \sum(\lambda'_i : \{x, y\} \subseteq E(H'_i) \text{ and } z \notin E(H'_i)),\\
\lambda'_{yz} &= \sum(\lambda'_i : \{y, z\} \subseteq E(H'_i) \text{ and } x \notin E(H'_i)) \text{ and}\\
\lambda'_{xz} &= \sum(\lambda'_i : \{x, z\} \subseteq E(H'_i) \text{ and } y \notin E(H'_i)).
\end{aligned}$$

Note that $\lambda'_{xyz} + \lambda'_{xy} + \lambda'_{yz} + \lambda'_{xz} = 1$, and each of the 1-edges x, y and z occur $\frac{4}{5}$ of the time, thus each is missing exactly $\frac{1}{5}$ of the time. A visual for the occurence of edges x, y and z is included in Figure 4.1, as well as the values for λ'_{xyz}, λ'_{xy}, λ'_{yz}, and λ'_{xz}. Thus

$$\lambda'_{xyz} = \frac{2}{5} \text{ and } \lambda'_{xy} = \lambda'_{yz} = \lambda'_{xz} = \frac{1}{5}. \tag{4.1}$$

First we consider any expanded triangle T which is not incident with edge p. For each subgraph H'_i in which all three edges x, y and z occur, we include two of the three edges in T $\frac{1}{3}\lambda'_{xyz}$ of the time. These patterns and their corresponding occurrences are illustrated in Figure 4.2, and result in an occurrence of $\frac{2}{3}\lambda'_{xyz} = \frac{4}{15}$ for each edge of T overall, by (4.1). Note that using each pattern one third of the time can be accomplished by using the patterns of Figure 4.2 for T for two of the six copies of each H'_i where x, y and z occur. Then for each subgraph H'_i in which z is omitted and x and y occur, we consider both triangle T and the other triangle T' incident with z. In this case we include the edges in T incident with z $\frac{1}{2}\lambda'_{xy}$ of the time, and the other edge in T $\frac{1}{2}\lambda'_{xy}$ of the time, and do the opposite in triangle T'. In all cases we also include two copies of edge z. The patterns are illustrated in Figure 4.3 and result in an occurrence of $\frac{1}{2}\lambda'_{xy} = \frac{1}{10}$ for each edge in T. Note that using each pattern half of the time can be accomplished by using each of the two patterns shown in Figure 4.3 for T (and T') in three of the six copies of each H'_i in which z is omitted. We do the same for the cases where x or y are omitted in H'_i. The total occurrence of each half-edge in T is

$$\frac{2}{3}\lambda'_{xyz} + \frac{1}{2}\lambda'_{xy} + \frac{1}{2}\lambda'_{yz} + \frac{1}{2}\lambda'_{xz},$$

which by (4.1) is $\frac{17}{30} < \frac{3}{5}$. We can arbitrarily add back half-edges in the multi-2EC$_{\text{cost}}$ convex combinations to obtain an occurrence of exactly $\frac{3}{5}$ for these edges.

Note that each 1-edge which is not p is now doubled whenever it was previously omitted in a subgraph, and thus occurs $\frac{6}{5}$ of the time ($\frac{4}{5}$ of the time as a single edge and $\frac{1}{5}$ of the time as a doubled edge). Also note that all patterns used in the expansion of the half-triangles ensure that the new multi-subgraphs created from the subgraphs H'_i for G' are also 2-edge connected and spanning in G, as required.

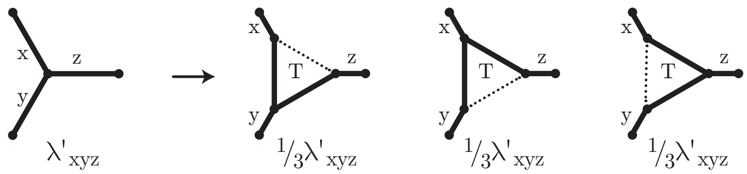


Figure 4.2: Patterns used for triangle expansion for subgraphs containing x, y and z, $p \notin \{x, y, z\}$.

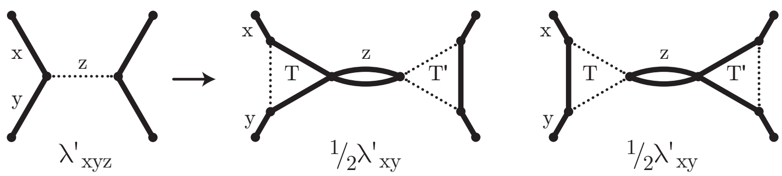


Figure 4.3: Patterns used for triangle expansion for an omitted edge z, $z \neq p$.

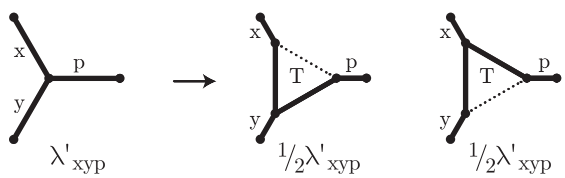


Figure 4.4: Patterns used for triangle expansion for subgraphs containing x, y and p.

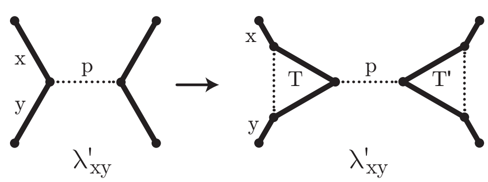


Figure 4.5: Pattern used for triangle expansion for omitted edge p.

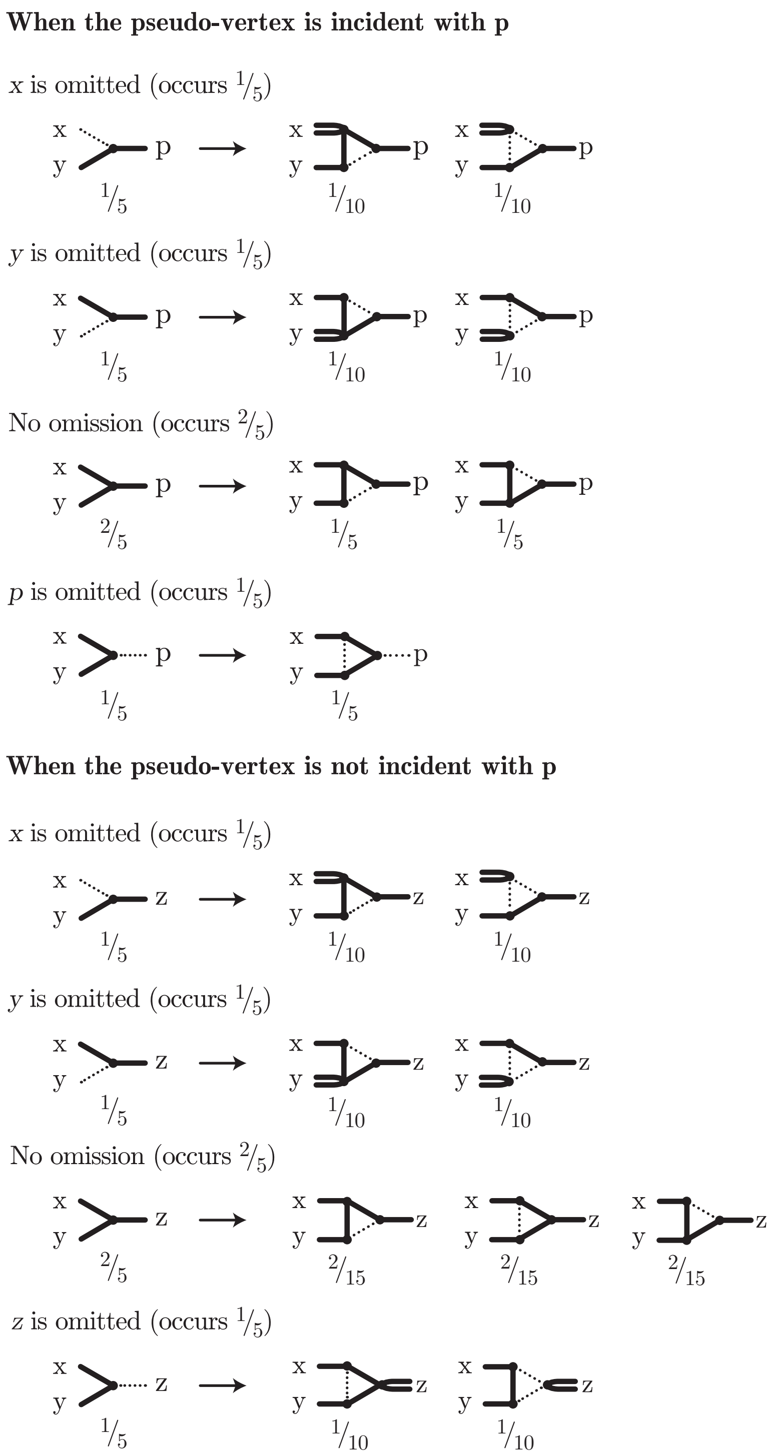


Figure 4.6: Summary of pattern occurrences upon expanding the pseudo-vertices of G'.

Next we consider any expanded triangle T which is incident with edge p, and without loss of generality let $p = z$. For each subgraph H_i in which all three edges x, y and p occur, we include two of the three edges in T $\frac{1}{2}\lambda'_{xyp} = \frac{1}{5}$ of the time, using the two patterns illustrated in Figure 4.4 (so we use each pattern in three of the six copies of H'_i). Then for each subgraph H'_i in which p is omitted and x and y occur, we include the two edges of T incident with p, as shown in Figure 4.5. Remember that this occurs $\lambda'_{xy} = \frac{1}{5}$ of the time, by (4.1). Triangles T incident with edge p otherwise have the same patterns as triangles not incident to p. Note that we do not double edge p. The total occurrence of each edge of T is exactly $\frac{3}{5}$, and p occurs exactly $\frac{4}{5}$ of the time (see Figure 4.6 for a complete illustration of these operations and the pattern occurrences).

We now have that, over all cases, the half-edge occurrence is $\frac{3}{5}$, p occurs $\frac{4}{5}$ of the time, and the occurrence of the other 1-edges is $\frac{6}{5}$. Furthermore, none of the 2-edge connected spanning multi-subgraphs use more than one copy of a half-edge or the edge p, and all of them use either one or two copies of a 1-edge. Thus $Q(G, p)$ holds.

Case 2. *G has a 2-edge cut $C = \{hi, jk\}$.*

Suppose the contrary, and let G be the smallest counter-example for which $Q(G, p)$ does not hold. Recall that p is specified to be an edge not in a 2-edge cut. Let G_1, G_2 be the two shores of the cut C in G, with h and j in G_1 and i and k in G_2, and without loss of generality choose C such that $G_1 + hj$ is 3-edge connected and does not contain p. Such a cut C always exists, since for any 2-edge cut $\delta(S)$, $S \subset V$, we can choose S such that edge p is in the subgraph induced by $V \setminus S$, and then we can choose C to be $\delta(Y)$ for the smallest subset Y of S such that $\delta(Y)$ is a 2-edge cut. Note that the vertices h, i, j, k are

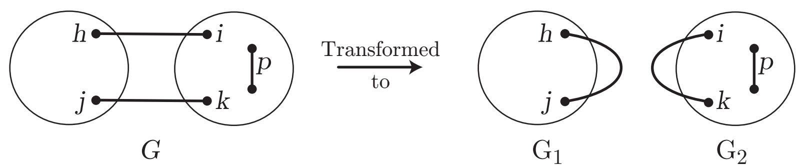


Figure 4.7: Splitting G in G_1 and G_2 when there exists a 2-edge cut in G across two 1-edges.

distinct, because G is a half-triangle graph, and $hi \neq p$ and $jk \neq p$ by assumption. Figure 4.7 illustrates the transformation from G to $G_1 + hj$ and $G_2 + ik$. By smaller example and Case 1, $Q(G_1 + hj, hj)$ and $Q(G_2 + ik, p)$ hold. We now "glue" together in the obvious way, the multi-subgraphs in the 2EC convex combination for $G_1 + hj$ where hj is omitted with the subgraphs in the multi-2EC convex combination for $G_2 + ik$ which have ik doubled (both patterns occur $\frac{1}{5}$ of the time) by removing the double edge ik and adding two copies of edges hi and jk. Similarly, we glue the subgraphs for $G_1 + hj$ and $G_2 + ik$ where hj and ik occur as single edges in the subgraphs (both patterns occur $\frac{4}{5}$ of the time) by removing hj and ik and adding edges hi and jk. Both cases are shown in Figure 4.8, where bold lines represent edges in the subgraph, dotted lines represent edges not in the subgraph, and dashed lines represent edges which may or may not be in the subgraph. We obtain $Q(G, p)$, contradiction. $\square$

Using the result of Lemma 3, we can now easily prove our main theorem that says that $\frac{6}{5}x^*$ is a multi-2EC convex combination for any half-triangle solution x^*.

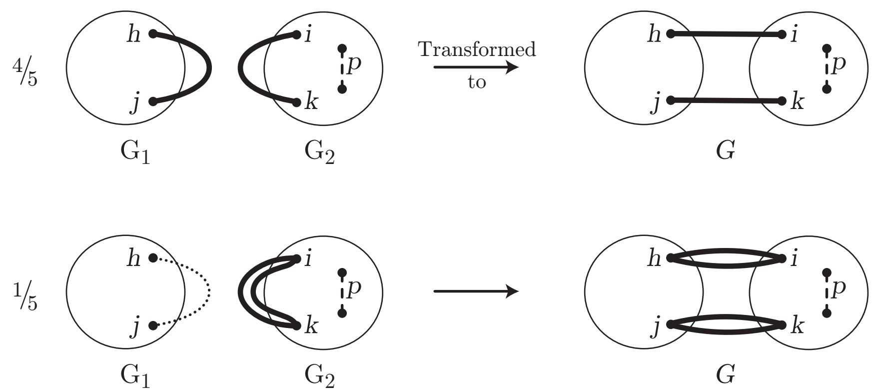


Figure 4.8: Splitting G in G_1 and G_2 when there exists a 2-edge cut in G across two 1-paths.

Proof of Theorem 5

Let $x^* \in \mathbb{R}^E$ be any half-triangle solution and let $G = (V, E)$ be the corresponding half-triangle graph. Without loss of generality, we can assume that G is not a cycle of 1-edges, as the theorem trivially holds in this case. Given the result of Lemma 3, basically all that is left to complete the proof is to handle 1-paths that do not consist of a single edge in the half-triangle graph G. To do this, we take G and replace each 1-path L by a single edge e_L to obtain a cubic half-triangle graph $G' = (V', E')$. Let p be any 1-edge in E' not in a 2-edge cut of G'. Since we are assuming that the original support graph G contains half-triangles (i.e. is not a cycle of 1-edges), such an edge p must exist. Then, by Lemma 3, we have that $Q(G', p)$ holds. In the 2-edge connected multi-subgraphs H'_i, $i = 1, 2, \ldots, j$ in the 2EC_{size} convex combination for $Q(G', p)$, we know edge p occurs exactly $\frac{4}{5}$ of the time as a single edge. By adding two copies of edge p to every subgraph H'_i where it does not appear, we have every 1-edge in G' (including edge p) appears $\frac{6}{5}$ of the time in the convex combination, and appears in every subgraph either singly or doubled. We can then replace each 1-edge e_L whenever it appears in a subgraph by its corresponding path L to obtain the required multi-2EC$_{\text{cost}}$ convex combination for the original half-triangle solution x^*. $\square$

As a final remark, we note that the methods we use to obtain the multi-2EC$_{\text{cost}}$ convex combinations for Theorem 5 are constructive, but not polynomial-time (in particular the methods of Lemma 3). Thus it is still an open question to find a $\frac{6}{5}$-approximation algorithm for multi-2EC$_{\text{cost}}$ for the special cost functions $c \in \mathbb{R}^E$ optimized at half-triangle solutions.

Chapter 5

A Comparison of the Lower Bounds for $2EC_{size}$

The quality of lower bounds directly affects the guarantee of approximation algorithms for $2EC_{size}$ and as such, is a vital area of study. In this chapter, we provide a survey of seven known lower bounds for $2EC_{size}$, compare them to each other and identify two that perform better than the others. The analysis is complicated by the fact that lower bounds often cannot be compared directly. Section 5.1 states the assumptions we make to simplify our analysis; Section 5.2 gives a brief review of the bounds and formalizes them; and Section 5.3 contains the results of the analysis.

5.1 Assumptions

The lower bounds analyzed here apply to undirected connected multi-graphs. For such a graph $G = (V, E)$, assume without loss of generality that G is 2-edge connected, as any solution to $2\mathrm{EC}_{\mathrm{size}}$ would double all the bridges of G. Furthermore, assume the absence of cut-vertices: if there exists a cut-vertex $v \in V$ for which $G - v$ has k components $G_i = (V_i, E_i)$, with $i = 1, \ldots, k$ and $k > 1$, we let G'_i be the graph obtained by adding v and edges $va \in E$ for $a \in V_i$, to G_i for $i = 1, \ldots, k$, then

$$\mathrm{OPT}(2\mathrm{EC}_{\mathrm{size}}(G)) = \sum_{i=1}^{k} \mathrm{OPT}(2\mathrm{EC}_{\mathrm{size}}(G_i)).$$

5.2 A Survey of the Bounds

In this section, we describe seven known lower bounds for $2\mathrm{EC}_{\mathrm{size}}$. All bounds can be applied to any 2-edge connected multi-graph $G = (V, E)$ with no cut-vertex. In fact, the bound of Subsection 5.2.6 specifically requires that G have no cut-vertex, and several other bounds require G to be at least 2-edge connected.

5.2.1 The Vertex Count Bound

Since any solution $2\mathrm{EC}_{\mathrm{size}}$ must include at least two edges at every vertex of the given graph G, it follows that $n = |V|$ is a natural lower bound to $2\mathrm{EC}_{\mathrm{size}}$. In fact, if graph G has a Hamiltonian tour, that solution serves as the optimal for $2\mathrm{EC}_{\mathrm{size}}$. We refer to this bound as the *Vertex Count Bound*.

5.2.2 The LP Relaxation Bound

The LP relaxation of $2\mathrm{EC}_{\mathrm{size}}$, denoted by $2\mathrm{EC}^{\mathrm{LP}}_{\mathrm{size}}$, forgoes the constraint that edge occurrences must have an integer value, as shown in (2.4) and (2.5). The LP relaxation provides a lower bound for $2\mathrm{EC}_{\mathrm{size}}$ because $\mathrm{OPT}(2\mathrm{EC}^{\mathrm{LP}}_{\mathrm{size}}(G)) \leq \mathrm{OPT}(2\mathrm{EC}_{\mathrm{size}}(G))$ for any graph G. We refer to this bound as the *LP Relaxation Bound*.

5.2.3 The Independent Set Bound

Garg, Santosh and Singla remark in [12] that any 2-edge connected spanning subgraph of G has at least two edges incident to each vertex of an independent set of G, i.e.

Lemma 4 ([12]). *If a graph has an independent set of size I then the minimum 2-vertex connected spanning subgraph has at least $2I$ edges.*

Note that all 2-vertex connected subgraphs are also 2-edge connected, by definition. Furthermore, while maximal independent sets can easily be constructed using a greedy approach, finding a maximum independent set is NP-hard and as such, the quality of this bound is highly variable on the size of the independent set found. We refer to this bound as the *Independent Set Bound*. Note that this bound differs from the Vertex Count Bound by the fact that an edge is not counted twice; as a result, the Independent Set Bound has a coefficient of 2.

5.2.4 The Tree-Carving Bound

Khuller and Vishkin introduce in [20] a useful data structure to find lower bounds for $2\mathrm{EC}_{\mathrm{size}}$, which they name the *tree-carving* of a graph and which they formalize as follows.

Definition 2 ([20]). *A tree-carving of a graph $G = (V, E)$ is a partition of the vertex set V into subsets $V_1, \ldots, V_i, \ldots, V_k$ with the following properties. Each subset constitutes a node of a tree Γ. For every vertex $v \in V_i$, all the neighbors of v in G belong either to V_i itself, or to V_j where V_j is adjacent to V_i in the tree Γ. The size of the tree-carving is k.*

In order to find a tree-carving, a Depth-First Search approach is used, and the resulting tree is modified to comply with Definition 2. While this approach does not find the largest possible tree-carving, it has the advantage of being polynomial in the number of edges. Khuller and Vishkin then prove the following lower bound.

Lemma 5 ([20]). *If the graph $G = (V, E)$ has a tree-carving of size k, then a lower bound on the number of edges of any 2-edge connected spanning subgraph in G is $2(k-1)$.*

As $2(k-1)$ can be smaller then n, they conclude by establishing the stronger composite lower bound $max(n, 2(k-1))$ for a tree-carving of size k. For the purpose of this paper, we use $2(k-1)$ as the bound and refer to it as the *Tree-Carving Bound*. A tree-carving can be found in polynomial time, but it is not known if a maximum size tree-carving can also be found efficiently.

5.2.5 The Components Bound

Garg, Santosh and Singla strenghten the concept of tree-carving used by Khuller and Vishkin (Subsection 5.2.4) by observing the following.

Lemma 6 ([12])**.** *Let S be a subset of vertices, $|S| = k$, such that its removal disconnects the graph into d components ($d > k$). Then any 2-edge connected spanning subgraph of G contains at least $n + d - k$ edges.*

Garg, Santosh and Singla further state that the bound of Lemma 6 is always as good as, or better than the Tree-Carving Bound [12]. Furthermore, they generalize their result to obtain a better lower bound, as seen below.

Lemma 7 ([12])**.** *Let $S_1, \ldots, S_i, \ldots, S_p$ be mutually disjoint sets of vertices, $|S_i| = k_i$, such that the removal of the vertices in S_i breaks the graph into d_i components ($d_i > k_i$). Then any 2-edge connected spanning subgraph of G contains at least $n + \sum_{i=1}^{p}(d_i - k_i)$ edges.*

In this paper, we use the lower bound described in Lemma 7, and refer to it as the *Components Bound*. While any disjoint sets of vertices $S_1, \ldots, S_i, \ldots, S_p$, $p \geq 1$, with desired properties can be found in polynomial time, it is unknown whether the optimal set of such S_i's that would maximize $n + \sum_{i=1}^{p}(d_i - k_i)$ can be found efficiently.

5.2.6 The 2-Matching Bound

Jothi, Raghavachari, and Varadarajan use the concept of maximum 2-matching in [18] to establish a novel lower bound for $2\mathrm{EC}_{\mathrm{size}}$, as is shown below. In their paper, they confuse the terms maximal matching and maximum matching: since it is clear from their approach that they use the concept of a maximum matching, we use the correct term in this thesis.

Lemma 8 ([18]). *Let $G = (V, E)$ be a graph defined on $|V| \geq 4$ vertices, with no cut-vertex. Let H be a maximum 3-cycle-free 2-matching of G consisting of a collection of cycles C and a set of paths P. Then $OPT(2EC_{size}(G)) \geq n + |P|$.*

Jothi, Raghavachari, and Varadarajan state in [18] that finding such a 2-matching can be done in polynomial time [16]. We refer to this lower bound as the *2-Matching Bound*.

5.2.7 The Obliged Edges Bound

In [3], Boyd, Fu and Sun use the concept of obliged edges —edges that must be included in the solution— to obtain a new lower bound. Let $\Omega \subseteq E$ contain all obliged edges, and let $F \subseteq E$ contain all edges in a 2-edge cut in G. Note that these edges must be in any solution of $2\mathrm{EC}_{\mathrm{size}}(G)$. Then, a lower bound for $2\mathrm{EC}_{\mathrm{size}}$ such that the obliged edges in Ω are included in the solution is

$$\frac{1}{2}\sum_{v \in V} max(2, |\delta(v) \cap (F \cup \Omega)|).$$

In the context of this paper, we do not consider obliged edges (i.e. $\Omega = \emptyset$), and thus simplify the lower bound to

$$\frac{1}{2}\sum_{v \in V} max(2, |\delta(v) \cap F|).$$

The lower bound can be computed in polynomial time by identifying all edges of G in 2-edge cuts; we refer to the bound as the *Obliged Edges Bound*.

5.3 The Relationship Between the Bounds

Three of the lower bounds, namely the Independent Set Bound, the Tree-Carving Bound and the Components Bound are dependent on the quality of the independent set, the tree-carving and the sets S_i, respectively. For instance, the empty set is an independent set, but would yield a terrible lower bound of $2|\emptyset| = 0$. As such, we always compare the best result possible via the lower bound.

Lemma 9. *The Vertex Count Bound does not outperform the LP Relaxation Bound for any 2-edge connected multi-graph $G = (V, E)$.*

Proof. We use the dual (2.6) of the LP Relaxation Bound defined by (2.5) to prove this relationship. Let $\mathbb{S} = \{S : \emptyset \neq S \subset V\}$, and let $d \in \mathbb{R}^{\mathbb{S}}$ such that

$$d_S = \begin{cases} \frac{1}{2} & \text{if } |S| = 1, \\ 0 & \text{otherwise.} \end{cases}$$

Note that d is a feasible solution to (2.6), and has objective value

$$\sum_{\emptyset \neq S \subset V} 2d_S = \sum_{v \in V} 2d_{\{v\}} = |V|.$$

Thus by the Weak Duality Theorem (see Theorem 1), the objective value of the primal (2.5) must be at least that of the dual (2.6), which implies that $|V| \leq 2\mathrm{EC}_{\mathrm{size}}^{\mathrm{LP}}(G)$. $\square$

Lemma 10. *The Independent Set Bound does not outperform the LP Relaxation Bound for any 2-edge connected multi-graph $G = (V, E)$.*

Proof. We prove this relationship with the dual (2.6) of the LP Relaxation Bound defined by (2.5). Let $\mathbb{S} = \{S : \emptyset \neq S \subset V\}$. Given an independent set $T \subseteq V$ of G, with $|T| = I$, let $d \in \mathbb{R}^{\mathbb{S}}$ such that

$$d_S = \begin{cases} 1 & \text{if } S = \{v\} \text{ for } v \in T, \\ 0 & \text{otherwise.} \end{cases}$$

Variable d is a feasible solution to (2.6) because d_S is non-negative for all $S \in \mathbb{S}$ and $\sum_{e \in \delta(S)} d_S \leq 1$ for all $e \in E$ (an edge uv cannot have both endpoints in T because T is an independent set).

The objective value of (2.6) with d is $2I$. By the Weak Duality Theorem (see Theorem 1), it follows that $2I \leq \mathrm{OPT}(2\mathrm{EC}_{\mathrm{size}}^{\mathrm{LP}}(G))$. $\square$

Lemma 11. *The Components Bound does not outperform the LP Relaxation Bound for a 2-edge connected multi-graph $G = (V, E)$.*

Proof. We model this proof on that of Garg, Santosh and Singla in [12]. Let $S_1, \ldots, S_i, \ldots, S_p$ be mutually disjoint sets of vertices, $|S_i| = k_i$, such that the removal of the vertices in S_i breaks the graph into d_i components ($d_i > k_i$). Let G' be the graph formed by coalescing the vertices of one S_i into a single vertex v. Since any $2\mathrm{EC}_{\mathrm{size}}^{\mathrm{LP}}(G)$ solution also provides a

$2\mathrm{EC}^{\mathrm{LP}}_{\mathrm{size}}(G')$ solution,

$$\mathrm{OPT}(2\mathrm{EC}^{\mathrm{LP}}_{\mathrm{size}}(G')) \leq \mathrm{OPT}(2\mathrm{EC}^{\mathrm{LP}}_{\mathrm{size}}(G)).$$

The removal of vertex v forms d components in G'. Let G_i denote the i^{th} component along with vertex v. Then,

$$\mathrm{OPT}(2\mathrm{EC}^{\mathrm{LP}}_{\mathrm{size}}(G')) = \sum_{i=1}^{d} \mathrm{OPT}(2\mathrm{EC}^{\mathrm{LP}}_{\mathrm{size}}(G_i)).$$

If n_i is the number of vertices in the graph G_i then $\mathrm{OPT}(2\mathrm{EC}^{\mathrm{LP}}_{\mathrm{size}}(G_i)) \geq n_i$. Also, $\sum_{i=1}^{d} n_i = n + d - k$ and therefore

$$\mathrm{OPT}(2\mathrm{EC}^{\mathrm{LP}}_{\mathrm{size}}(G)) \geq \mathrm{OPT}(2\mathrm{EC}^{\mathrm{LP}}_{\mathrm{size}}(G')) = \sum_{i=1}^{d} \mathrm{OPT}(2\mathrm{EC}^{\mathrm{LP}}_{\mathrm{size}}(G_i)) \geq n + d - k.$$

This can be generalized to take into account the other S_i's, and gives $\mathrm{OPT}(2\mathrm{EC}^{\mathrm{LP}}_{\mathrm{size}}(G)) \geq n + \sum_{i=1}^{p}(d_i - k_i)$. □

Lemma 12. *The Obliged Edges Bound does not outperform the LP Relaxation Bound for any 2-edge connected multi-graph $G = (V, E)$.*

Proof. Consider any 2-edge cut C in G, where $\{e, f\} = \delta(C)$; the second constraint of (2.5) applied to C states that

$$x(\delta(C)) \geq 2,$$
$$\text{i.e. } x_e + x_f \geq 2.$$

Therefore, $x_e = x_f = 1$ (as there always exists a solution where non-bridge edges are not doubled) and as such, any edges $e \in F$ always appears in the $2\text{EC}^{\text{LP}}_{\text{size}}$ solution with value $x_e = 1$. Thus $x(\delta(v)) \geq \max(2, \delta(v) \cup F)$ for all $v \in V$. Summing this constraint for all $v \in V$ gives $\frac{1}{2}\sum_{v \in V} max(2, |\delta(v) \cap F|) \leq \text{OPT}(2\text{EC}^{\text{LP}}_{\text{size}}(G))$. □

Lemma 13. *The Independent Set Bound does not outperform the maximum value of the Tree-Carving Bound for any 2-edge connected multi-graph $G = (V, E)$.*

Proof. Given an independent set T with $|T| = I$, let $S = V \setminus T$. Then the vertices in T along with the set S form a tree-carving of size $k = I + 1$ and give the Tree-Carving Bound a value of $2(k-1) = 2I$. □

Lemma 14. *The Independent Set Bound does not outperform the maximum value of the Components Bound for any 2-edge connected multi-graph $G = (V, E)$.*

Proof. Garg, Santosh and Singla state in [12] that the Tree-Carving Bound never outperforms the Components Bound. Since Lemma 13 shows that the Independent Set Bound never outperforms the Tree-Carving Bound, we can infer that the Independent Set Bound never outperforms the Components Bound either. □

Lemma 15. *The Tree-Carving Bound does not outperform the LP Relaxation Bound for any 2-edge connected multi-graph $G = (V, E)$.*

Proof. The Tree-Carving Bound never outperforms the Components Bound [12]. Since Lemma 11 shows that the Components Bound never outperforms the LP Relaxation Bound, we can infer that the Tree-Carving Bound too never outperforms the LP Relaxation Bound.

□

Lemma 16. *The Independent Set Bound does not outperform the 2-Matching Bound for any 2-edge connected multi-graph $G = (V, E)$.*

Proof. Let T be an independent set of G of size I, and let H be a maximal 3-cycle-free 2-matching of G consisting of a collection of cycles C and a set of paths P.

If $2I \leq n$, then the result follows. So assume $2I > n$, which means that more than half of the vertices of G are contained in T. Knowing that cycles are only possible if we alternate between vertices in T and vertices not in T, at best, it follows that there are at least $|V| - 2(|V| - I)$ paths, which simplifies to $-|V| + 2I$. Therefore, $n + |P| \geq n - |V| + 2I = 2I$. □

5.3.1 Analysis

Two bounds emerge as the best over all graphs, namely the LP Relaxation Bound and the 2-Matching Bound. While not directly comparable, the combination of the two (i.e. $\max(n + |P|, 2\text{EC}^{\text{LP}}_{\text{size}})$) would provide the best known lower bound to 2EC_{size} currently known. In order to reach this conclusion, we examine the relationship between every distinct pair of bounds. Should one bound perform equally to, or better than the other, we prove this in one of Lemma 9, 10, 11, 12, 13, 14, 15 and 16. Should the bounds not be directly comparable, we show examples where both bounds outperform the other in Figures 5.2, 5.3, 5.4, 5.6, 5.5, 5.7, and 5.8. Tables 5.1, 5.2, 5.3, 5.4, 5.5, 5.6 and 5.7 provide a justification of the relationship between all pairs of distinct bounds. Finally, Figure 5.1 provides a visual of the relationship between all bounds.

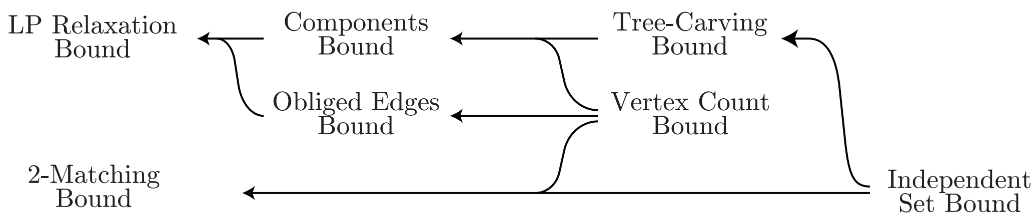


Figure 5.1: A comparison of the lower bounds: an arrow from bound A to bound B shows that for any 2-edge connected multi-graph G without cut-vertex, $A(G) \leq B(G)$.

Table 5.1: Relationship between the Vertex Count Bound and other bounds

Bounds	Relationship to the Vertex Count Bound
LP Relaxation Bound	$\geq$, by Lemma 9
Independent Set Bound	incomparable, as seen in graphs $K_{4,3}$ and B of Figures 5.2 and 5.4
Tree-Carving Bound	incomparable, as seen in graphs $K_{4,3}$ and B of Figures 5.2 and 5.5
Components Bound	$\geq$, trivially
2-Matching Bound	$\geq$, trivially
Obliged Edges Bound	$\geq$, trivially, as $\frac{1}{2}\sum_{v \in V} max(2, \lvert\delta(v) \cap \emptyset\rvert) = n$

Note that the value of the bounds in the referenced figures has been verified by computer to ensure optimality, using the linear programming solver Gurobi [11] with LPs and ILPs generated by a custom Python program.

Table 5.2: Relationship between the LP Relaxation Bound and other bounds

Bounds	Relationship to the LP Relaxation Bound
Vertex Count Bound	$\leq$, by Lemma 9
Independent Set Bound	$\leq$, by Lemma 10
Tree-Carving Bound	$\leq$, by Lemma 15
Components Bound	$\leq$, by Lemma 11
2-Matching Bound	incomparable, as seen in graphs A and B of Figures 5.3 and 5.7
Obliged Edges Bound	$\leq$ by Lemma 12

Table 5.3: Relationship between the Independent Set Bound and other bounds

Bounds	Relationship to the Independent Set Bound
Vertex Count Bound	incomparable, as seen in graphs $K_{4,3}$ and B of Figures 5.2 and 5.4
LP Relaxation Bound	$\geq$, by Lemma 10
Tree-Carving Bound	$\geq$, by Lemma 13
Components Bound	$\leq$, by Lemma 14
2-Matching Bound	$\geq$, by Lemma 16
Obliged Edges Bound	incomparable, as seen in graphs $K_{4,3}$ and A of Figures 5.4 and 5.8

Table 5.4: Relationship between the Tree-Carving Bound and other bounds

Bounds	Relationship to the Tree-Carving Bound
Vertex Count Bound	incomparable, as seen in graphs $K_{4,3}$ and B of Figures 5.2 and 5.5
LP Relaxation Bound	$\geq$, by Lemma 15
Independent Set Bound	$\leq$, by Lemma 13
Components Bound	$\geq$, as stated in [12]
2-Matching Bound	incomparable, as seen in graphs A and B of Figures 5.5 and 5.7
Obliged Edges Bound	incomparable, as seen in graphs $K_{4,3}$ and B of Figures 5.5 and 5.8

Table 5.5: Relationship between the Components Bound and other bounds

Bounds	Relationship to the Components Bound
Vertex Count Bound	$\leq$, trivially
LP Relaxation Bound	$\geq$, by Lemma 11
Independent Set Bound	$\leq$, by Lemma 14
Tree-Carving Bound	$\leq$, as stated in [12]
2-Matching Bound	incomparable, as seen in graphs A and B of Figures 5.7 and 5.6
Obliged Edges Bound	incomparable, as seen in graphs $K_{4,3}$ and C of Figures 5.8 and 5.6

Table 5.6: Relationship between the 2-Matching Bound and other bounds

Bounds	Relationship to the 2-Matching Bound
Vertex Count Bound	$\leq$, trivially
LP Relaxation Bound	incomparable, as seen in graphs A and B of Figures 5.3 and 5.7
Independent Set Bound	$\leq$, by Lemma 16
Tree-Carving Bound	incomparable, as seen in graphs A and B of Figures 5.5 and 5.7
Components Bound	incomparable, as seen in graphs A and B of Figures 5.7 and 5.6
Obliged Edges Bound	incomparable, as seen in graphs $K_{4,3}$ and A of Figures 5.7 and 5.8

Table 5.7: Relationship between the Obliged Edges Bound and other bounds

Bounds	Relationship to the Obliged Edges Bound
Vertex Count Bound	$\leq$, trivially, as $\frac{1}{2}\sum_{v\in V} max(2, \|\delta(v)\cap\emptyset\|) = n$
LP Relaxation Bound	$\geq$ by Lemma 12
Independent Set Bound	incomparable, as seen in graphs $K_{4,3}$ and A of Figures 5.4 and 5.8
Tree-Carving Bound	incomparable, as seen in graphs $K_{4,3}$ and B of Figures 5.5 and 5.8
Components Bound	incomparable, as seen in graphs $K_{4,3}$ and C of Figures 5.8 and 5.6
2-Matching Bound	incomparable, as seen in graphs $K_{4,3}$ and A of Figures 5.7 and 5.8

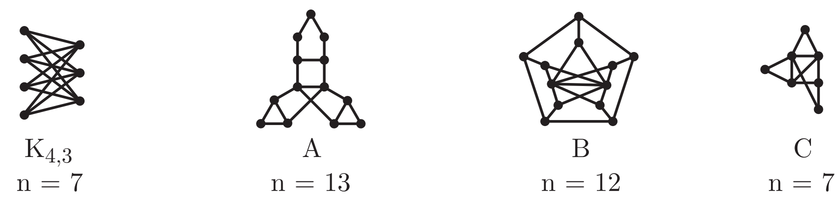


Figure 5.2: Value for the Vertex Count Bound for select graphs.

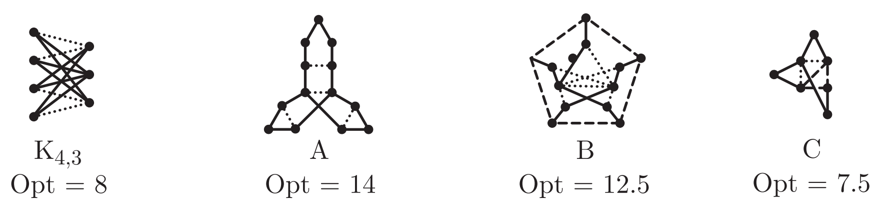


Figure 5.3: Value for the LP Relaxation Bound: displayed is the solution to $2EC^{LP}_{size}$ for select graphs, where dotted lines represent edges omitted in the solution, dashed lines stand for edges with occurrence $\frac{1}{2}$ and solid lines represent edges with an occurrence of 1.

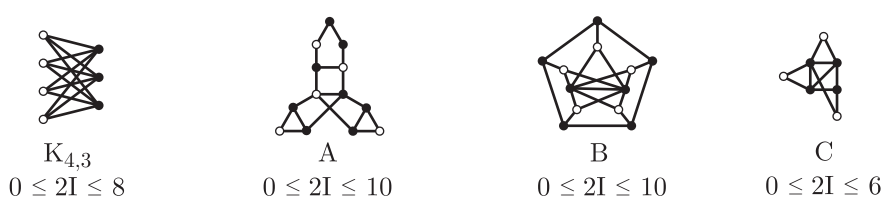


Figure 5.4: Range of values for the Independent Set Bound: displayed is a maximum size independent set T with $I = |T|$ for select graphs, where white outlined circles represent vertices included in T and solid black circles represent vertices not in T.

$K_{4,3}$ $\quad$ A $\quad$ B $\quad$ C

$0 \leq 2(k - 1) \leq 8$ $\quad$ $0 \leq 2(k - 1) \leq 14$ $\quad$ $0 \leq 2(k - 1) \leq 10$ $\quad$ $0 \leq 2(k - 1) \leq 6$

Figure 5.5: Range of values for the Tree-Carving Bound: displayed is a maximum size tree-carving of size k for select graphs, where dashed ellipses represent nodes of the tree.

$K_{4,3}$ $\quad$ A $\quad$ B $\quad$ C

$7 \leq \text{Opt} \leq 8$ $\quad$ $13 \leq \text{Opt} \leq 14$ $\quad$ $\text{Opt} = 12$ $\quad$ $\text{Opt} = 7$

Figure 5.6: Range of values *Opt* for the Components Bound. In the figure, removable vertices in the set S_1 are outlined, while others are filled.

$K_{4,3}$ $\quad$ A $\quad$ B $\quad$ C

$n + |P| = 8$ $\quad$ $n + |P| = 13$ $\quad$ $n + |P| = 13$ $\quad$ $n + |P| = 8$

Figure 5.7: Value for the 2-Matching Bound: displayed is a maximum 2-matching M consisting of cycles C and paths P, where dotted lines represent edges omitted in M and solid lines represent edges included in M.

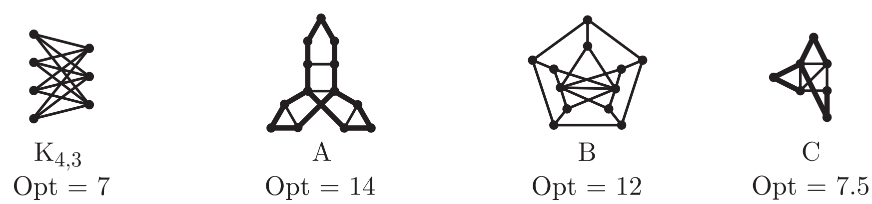


Figure 5.8: Value *Opt* for the Obliged Edges Bound: displayed in bold are edges in 2-edge cuts.

Chapter 6

Conclusion and Future Work

In this thesis, we show that the integrality gap of the LP relaxation to multi-2EC$_{\text{cost}}$ equals $\frac{6}{5}$ when restricted to half-triangle solutions, a family of cost functions which have been shown to provide the worst case by empirical testing up to $n = 10$ [1]. This supports a conjecture by Alexander, Boyd and Elliott-Magwood [1], as well as improves on Carr and Ravi's upper bound of $\frac{4}{3}$ for half-triangle solutions. We then investigate the related problem of 2EC$_{\text{size}}$: the simplest version of the problem that remains NP-hard takes cubic 3-edge connected graphs as input. We show that every such graph $G = (V, E)$ with $n = |V|$ allows a 2EC$_{\text{size}}$ solution of size at most $\frac{7n}{6}$, which improves upon Boyd, Iwata and Takazawa's result of $\frac{6n}{5}$.

We propose the following directions for future work.

1. Carr and Ravi in [5] identify half-integer, or $\frac{1}{2}$-integral, solutions as the simplest fractional solutions to the LP relaxation to multi-2EC$_{\text{cost}}$ and prove the integrality

gap for multi-2EC$_{\text{cost}}$ to be at most $\frac{4}{3}$ for the family of cost functions optimized at half-integers. Further studies by Alexander, Boyd and Elliott-Magwood show that it is half-integer solutions, more specifically half-triangle solutions, that generate the worst case. A natural next step would be to try to show that αmulti-2EC$_{\text{cost}} \leq \frac{4}{3}$ for cost functions optimized at $\frac{1}{3}$-integral solutions. Alternately, one could try to show that the result of Chapter 4 could be extended to show αmulti-2EC$_{\text{cost}} \leq \frac{6}{5}$ for all cost functions optimized at half-integer solutions.

2. We prove in Chapter 3 that every cubic 3-edge connected graph allows a 2-edge connected spanning subgraph of size at most $\frac{7n}{6}$. The proof is constructive but not polynomial, and gives hope that a polynomial time algorithm to find such a subgraph exists, which would improve upon [4, 28]. Furthermore, we know of no example that approaches the ratio of $\frac{7}{6}$ (the largest is the Petersen graph, at $\frac{11}{10}$) and we therefore believe that 2-edge connected spanning subgraphs with less edges exist.

3. Given a graph $G = (V, E)$ and a maximal 2-matching M of G, let $n = |V|$ and P be the set of paths in M. Chapter 5 shows that 2EC$^{\text{LP}}_{\text{size}}$ and $n + |P|$, also called the 2-Matching Bound, provide the best lower bounds for 2EC$_{\text{size}}$. Yet the use of these bounds for approximation algorithms is not abundant in the literature and no approximation algorithm makes use of the two lower bounds at once, i.e. $\max(\text{OPT}(2\text{EC}^{\text{LP}}_{\text{size}}(G)), n + |P|)$. We believe that this improved lower bound could be used to strengthen existing approximation algorithms or devise new ones.

References

[1] Alexander, A., Boyd, S., Elliott-Magwood, P., On the integrality gap of the 2-edge connected subgraph problem. *Technical Report TR-2006-04*, SITE, University of Ottawa, Ottawa, Canada, 2006.

[2] Bondy, A., Murty, U., Graph Theory (Graduate Texts in Mathematics), *Graduate Texts in Mathematics*, 44, Springer, New York, 2008.

[3] Boyd, S., Fu, Y., Sun, Y., A 5/4 approximation for subcubic 2EC using circulations and obliged edges, *Discrete Appl. Math.*, Elsevier Science Publishers, 209 Issue C, 48-58, 2016.

[4] Boyd, S., Iwata, S., Takazawa, K., Finding 2-factors closer to TSP tours in cubic graphs, *SIAM Journal on Discrete Mathematics*, 27, 918-939, 2013.

[5] Carr R., Ravi R., A new bound for the 2-edge connected subgraph problem, *Proceedings of Integer Programming and Combinatorial Optimization (IPCO)*, Lecture Notes in Computer Science, Springer, 112-125, 1998.

[6] Cheriyan, J., Sebő, A., Szigeti, Z., An Improved Approximation Algorithm for Minimum Size 2-Edge Connected Spanning Subgraphs, *Proc. 6th IPCO*, SIAM, LNCS, 1412, 1998.

[7] Christofides, N., Worst-case analysis of a new heuristic for the travelling salesman problem, *Report 388*, Graduate School of Industrial Administration, Carnegie Mellon University, Pittsburgh, United States, 1976.

[8] Csaba, B., Karpinski, M., Krysta, P., Approximability of dense and sparse instances of minimum 2-connectivity, tsp and path problems, *D. Eppstein (Ed.), SODA*, ACM/SIAM, 74–83, 2002.

[9] Frederickson, G. N., Ja'Ja', J. Approximation algorithms for several graph augmentation problems, *SIAM J. Comput.*, 10, 2, 270-283, 1981.

[10] Frederickson, G. N., Ja'Ja', J. On the relationship between the biconnectivity augmentation and travelling salesman problems, *Theoretical Computer Science*, 19, 189–201, 1982.

[11] G. Optimization. Gurobi 5.6.3. http://www.gurobi.com, 2014.

[12] Garg, N., Santosh, V., Singla, A., Improved Approximation Algorithms for Biconnected Subgraphs via Better Lower Bounding Techniques, *Proceedings of the Fourth Annual ACM-SIAM Symposium on Discrete Algorithms* SODA '93, 103–111, 1993.

[13] Garey, M., Johnson, D. Computers and Intractability: A Guide to the Theory of NP-completeness. W.H. Freeman and Company, New York, New York, 1979.

[14] Goemans, M. X., Bertsimas , D. J. Survivable Networks, Linear Programming Relaxations and the Parsimonious Property, *Math. Program.*, Springer-Verlag New York, Inc., 145–166, 1993.

[15] Goemans, M. X., Goldberg, A. V., Plotkin, S., Shmoys, D. B., Tardos, É., Williamson, D. P. Improved Approximation Algorithms for Network Design Problems, *Proceedings of the Fifth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '94, 223–232, 1994.

[16] Hartvigsen, D., Extensions of Matching Theory, *Ph.D. Thesis*, Carnegie-Mellon University, 1984.

[17] Jain, K., A Factor 2 Approximation Algorithm for the Generalized Steiner Network Problem, *Combinatorica*, 21, 1, 39–60, 2001.

[18] Jothi, R., Raghavachari, B., Varadarajan, S., A 5/4-Approximation Algorithm for Minimum 2-Edge-Connectivity, *Proceedings of the Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '03, Society for Industrial and Applied Mathematics, 262–273, 2003.

[19] Khuller, S., Thurimella, R., Approximation Algorithms for Graph Augmentation, *UMIACS-TR-132*, University of Maryland at College Park, College Park, United States, 1991.

[20] Khuller, S., Vishkin, U., Biconnectivity approximations and graph carvings, *Journal of the Association for Computing Machinery*, 41, 214-235, 1994.

[21] Krysta, P., Kumar, V. S. A., Approximation Algorithms for Minimum Size 2-Connectivity Problems, *STACS 2001: 18th Annual Symposium on Theoretical Aspects of Computer Science Dresden, Germany, February 15–17, 2001 Proceedings*, 431–442, 2001.

[22] Heeger, K., Vygen, J., Two-connected spanning subgraphs with at most 10/7 OPT edges, `arXiv:1609.00147 [math.CO]`, 2016.

[23] Huh, W.T., Finding 2-edge connected spanning subgraphs, *Operations Research Letters*, Elsevier, 32, 212-216, 2004.

[24] Ravi, R., Klein, P., When Cycles Collapse: A General Approximation Technique for Constrained Two-Connectivity Problems, *Technical Report*, SITE, Brown University, Providence, United States, 1992.

[25] Schrijver, A., Combinatorial Optimization - Polyhedra and Efficiency, Springer, 2003.

[26] Sebő, A., Benchetrit, Y., Stehlik, M., Problems about uniform covers, with tours and detours, *Matematisches Forschungsinstitut Oberwolfach* No. 51/2014, DOI: 10.4171/OWR/2014/51, 2912-2915, 2015.

[27] Sebő, A., Vygen, J., Shorter Tours by Nicer Ears: 7/5-Approximation for the graph-TSP, 3/2 for the Path Version, and 4/3 for Two-edge-connected Subgraphs, *Combinatorica*, Springer-Verlag New York, Inc., 34 Issue 5, 597–629, 2014.

[28] Takawaza, K., A 7/6-approximation algorithm for the minimum 2-edge connected subgraph problem in bipartite cubic graphs, *Information Processing Letters*, Elsevier, 116, 9, 550–553, 2016.

Index