

Linearizable Instances of the Quadratic Minimum Spanning Tree Problem on 3-Connected Graphs

Lucas Waddell* Dante Mancino†

Abstract

An instance of the quadratic minimum spanning tree problem (QMSTP) is called *linearizable* if it can be rewritten as an instance of the *linear* minimum spanning tree problem in such a way that the objective function value is preserved at all feasible solutions. Previous work has shown that a sufficient condition for linearizability is that the (symmetric) matrix of cost coefficients is a weak sum matrix. This condition has been shown to also be necessary for QMSTP instances defined on complete graphs and complete bipartite graphs whose two parts each have at least three vertices, but *not* necessary for QMSTP instances defined on certain other types of graphs. In this paper, we exploit a recent polyhedral framework for studying linearizable quadratic combinatorial optimization programs to prove the following: among simple connected graphs with at least four edges, the graphs for which every linearizable QMSTP instance has a weak sum symmetric cost matrix are exactly the 3-connected graphs.

Acknowledgments: D. Mancino's work on this project was supported by Bucknell University's Hoover Fund for Undergraduate Math Research.

Key Words: nonlinear combinatorial optimization, linearizable, quadratic minimum spanning tree

Declarations of Interest: none

1 Introduction

Consider a graph $G = (V, E)$ with $|V| = n$ vertices and $|E| = m$ edges (all of the graphs in this paper will be simple, connected, and undirected, unless otherwise stated). The quadratic minimum spanning tree problem (QMSTP), first introduced by Assad and Xu in 1992 [1], is an NP-hard combinatorial optimization program that can be formulated as

$$\text{QMSTP: minimize } \left\{ \sum_{e \in E} c_e x_e + \sum_{e \in E} \sum_{\substack{f \in E \\ f > e}} q_{ef} x_e x_f : x \in \mathbf{X}(G), x \text{ binary} \right\},$$

*Department of Mathematics & Statistics, Bucknell University, Lewisburg, PA 17837, lucas.waddell@bucknell.edu

†Department of Mathematics & Statistics, Bucknell University, Lewisburg, PA 17837, dcm023@bucknell.edu

where

$$\mathbf{X}(G) \equiv \left\{ \begin{array}{l} \mathbf{x} \in \mathbb{R}^m : \sum_{e \in E} x_e = n - 1, \\ \sum_{e \in E(S)} x_e \leq |S| - 1 \quad \forall S \subset V, |S| \geq 2, \\ \mathbf{0} \leq \mathbf{x} \leq \mathbf{1} \end{array} \right\}, \quad (1)$$

with $E(S)$ defined as the set of all edges having both endpoints in S . Note that the objective function assumes an ordering of the edges in E . The problem's name arises from the fact that the objective function is quadratic in the m binary variables $\mathbf{x}$, and the constraints of $\mathbf{X}(G)$, together with the binary restrictions, define the set of all 0-1 incidence vectors of spanning trees of G . Recall that a spanning tree of an undirected, connected graph G is a connected subgraph that contains no cycles and covers all of the vertices of G . The first constraint of $\mathbf{X}(G)$ enforces that the spanning tree contains the correct number of edges, which is always one less than the number of vertices of G . The second set of constraints prevents the chosen edges from forming any cycles. A detailed description of the set $\mathbf{X}(G)$, known as the *spanning tree polytope*, can be found in [5, Section 6.3, p. 144], for example. Often, it is convenient to equivalently rewrite the objective function of QMSTP as

$$\sum_{e \in E} \sum_{f \in E} w_{ef} x_e x_f, \quad (2)$$

where the symmetric cost matrix $\mathbf{W} \in \mathbb{R}^{m \times m}$ is defined by

$$w_{ef} = \begin{cases} \frac{1}{2} q_{ef}, & \text{if } e < f \\ \frac{1}{2} q_{fe}, & \text{if } e > f \\ c_e, & \text{if } e = f \end{cases}. \quad (3)$$

The number of feasible solutions to QMSTP can grow rapidly; Cayley's formula, for example, tells us that a complete graph with n vertices has n^{n-2} spanning trees. In fact, earlier exact methods reported solving benchmark instances on complete graphs having up to 50 vertices [8, 9]. As a result, a recent research direction involves the investigation of QMSTP instances that are in fact readily solvable [3, 4, 12], which includes the study of so-called *linearizable* instances. An instance of QMSTP on a graph $G = (V, E)$ with objective coefficients $\mathbf{c} \in \mathbb{R}^m$ and $\mathbf{q} \in \mathbb{R}^{m(m-1)/2}$ is said to be linearizable if there exist a vector $\hat{\mathbf{c}} \in \mathbb{R}^m$ and a scalar κ so that

$$\sum_{e \in E} c_e x_e + \sum_{e \in E} \sum_{\substack{f \in E \\ f > e}} q_{ef} x_e x_f = \kappa + \sum_{e \in E} \hat{c}_e x_e \quad (4)$$

for all $\mathbf{x} \in \mathbf{X}(G)$, $\mathbf{x}$ binary, with $\mathbf{X}(G)$ given by (1). Once a linearization is available for a QMSTP instance, it can be readily solved by applying a polynomial-time minimum spanning tree algorithm such as Kruskal's algorithm or Prim's algorithm (see, for example, [5, Section 6.1, pp. 136–137]).

The first work to study linearizable QMSTP instances [3] gives the following sufficient condition for linearizability.

Theorem 1.1 ([3]). *Consider an instance of QMSTP defined on a graph $G = (V, E)$ with objective function coefficients $\mathbf{c} \in \mathbb{R}^m$ and $\mathbf{q} \in \mathbb{R}^{m(m-1)/2}$. If the symmetric cost matrix $\mathbf{W} \in \mathbb{R}^{m \times m}$ as defined in (3) is a weak sum matrix (meaning that there exists a vector $\mathbf{a} \in \mathbb{R}^m$ such that $w_{ef} = a_e + a_f$ for all $e \neq f$), then the instance is linearizable. Specifically, the linearized objective of (4) has $\hat{c}_e = c_e + 2(n-2)a_e$ for all $e \in E$, and $\kappa = 0$.*

The paper [12] builds upon the work of [3] by providing and exploiting an alternative polyhedral characterization of linearizable instances. In particular, given an instance of QMSTP, they define the set

$$P(G) \equiv \left\{ (\mathbf{x}, \mathbf{y}) \in \mathbb{R}^m \times \mathbb{R}^{m(m-1)/2} : \mathbf{x} \in \mathbf{X}(G), \mathbf{x} \text{ binary}, y_{ef} = x_e x_f \quad \forall e < f \right\}, \quad (5)$$

and show that the m equations

$$\sum_{\substack{e \in E \\ e < f}} y_{ef} + \sum_{\substack{e \in E \\ e > f}} y_{fe} = (n-2)x_f \quad \forall f \in E \quad (6)$$

are valid for $\text{aff}\{P(G)\}$, where $\text{aff}\{\bullet\}$ denotes the affine hull of the set $\bullet$. In other words, the equations (6) hold for all $(\mathbf{x}, \mathbf{y}) \in P(G)$. This leads to the following sufficient condition for QMSTP linearizability.

Theorem 1.2 ([12]). *Consider an instance of QMSTP defined on a graph $G = (V, E)$ with objective function coefficients $\mathbf{c} \in \mathbb{R}^m$ and $\mathbf{q} \in \mathbb{R}^{m(m-1)/2}$. If there exists a vector $\bar{\mathbf{c}} \in \mathbb{R}^m$ such that the equation*

$$\sum_{e \in E} \bar{c}_e x_e + \sum_{e \in E} \sum_{\substack{f \in E \\ f > e}} q_{ef} y_{ef} = 0 \quad (7)$$

can be expressed as a linear combination of the equations (6), then the instance is linearizable.

The result of Theorem 1.2 is intuitive, because if an equation such as (7) exists, it can be directly used to substitute all of the quadratic terms out of the objective function. As it turns out, the sufficient condition for QMSTP linearizability of Theorem 1.2 is equivalent to the sufficient condition of Theorem 1.1, due to the following result of [12].

Theorem 1.3 ([12]). *Given an instance of QMSTP defined on a graph $G = (V, E)$ with objective function coefficients $\mathbf{c} \in \mathbb{R}^m$ and $\mathbf{q} \in \mathbb{R}^{m(m-1)/2}$, the matrix $\mathbf{W} \in \mathbb{R}^{m \times m}$ as defined in (3) is a weak sum matrix if and*

only if there exists a vector $\bar{c} \in \mathbb{R}^m$ such that (7) can be expressed as a linear combination of the equations (6).

A question first posed in [3] is to determine exactly which graph structures cause the sufficient condition for QMSTP linearizability of Theorem 1.1 (or, equivalently, the sufficient condition of Theorem 1.2) to also be necessary. They show that the condition is necessary for any QMSTP instance defined on a complete graph K_n or a complete bipartite graph K_{n_1, n_2} having $\min\{n_1, n_2\} \geq 3$. However, they also show that for every biconnected (i.e., 2-vertex-connected) graph $G = (V, E)$ with $|V| \geq 4$ and at least one vertex of degree 2, there exists a linearizable QMSTP instance on G whose symmetric cost matrix is not weak sum. In their concluding remarks, the authors ask whether it would be enough for G to be biconnected with a minimum vertex degree of at least 3 in order to guarantee that the linearizability condition of Theorem 1.1 is both necessary and sufficient.

More recently, [12] makes further progress on the open problem by providing a framework, stated below as Theorem 1.4, for determining when the sufficient linearizability condition of Theorem 1.2 is also necessary.

Theorem 1.4 ([12]). *Consider a simple connected graph $G = (V, E)$. The sufficient linearizability condition of Theorem 1.2 (or, equivalently, the sufficient condition of Theorem 1.1) is also necessary for every QMSTP instance on G if and only if the structure of G enforces that for any equation of the form*

$$\sum_{e \in E} \alpha_e x_e + \sum_{e \in E} \sum_{\substack{f \in E \\ f > e}} \beta_{ef} y_{ef} = \kappa \quad (8)$$

that is valid for $\text{aff}\{P(G)\}$, there exists a vector $\bar{\alpha} \in \mathbb{R}^m$ such that

$$\sum_{e \in E} \bar{\alpha}_e x_e + \sum_{e \in E} \sum_{\substack{f \in E \\ f > e}} \beta_{ef} y_{ef} = 0 \quad (9)$$

can be expressed as a linear combination of the equations (6).

Using Theorem 1.4, [12] answers the question in the concluding remarks of [3] negatively; for biconnected graphs, the necessity of the condition of Theorem 1.1 is not determined by the graph's minimum vertex degree. In particular, it is possible to construct a biconnected graph with arbitrarily large minimum vertex degree such that there exists an equation of the form (8) that is valid for $\text{aff}\{P(G)\}$ for which it is impossible to find a vector $\bar{\alpha} \in \mathbb{R}^m$ so that (9) is expressible as a linear combination of the equations (6).

The main contribution of this paper, given in Theorem 3.3, is to extend the results of [3] and [12] by showing that among simple connected graphs with at least four edges, the graphs for which every lineariz-

able QMSTP instance has a weak sum symmetric cost matrix are exactly the 3-connected graphs, thereby completely resolving the open question of [3]. Recall the following definition of a 3-connected graph.

Definition 1.1. *A graph $G = (V, E)$ is called **3-connected** if $|V| \geq 4$ and deleting any set of fewer than three vertices, together with all incident edges, leaves a connected graph.*

The remainder of this paper is organized as follows. Section 2 focuses on the *only if* direction of our main result, showing that the sufficient condition for QMSTP linearizability of Theorems 1.1 and 1.2 is not also necessary for non-3-connected graphs. Section 3, on the other hand, focuses on the *if* direction, showing that the sufficient condition for QMSTP linearizability of Theorems 1.1 and 1.2 is also necessary for 3-connected graphs. Section 4 gives concluding remarks.

2 Linearizable QMSTPs on Non-3-Connected Graphs

In this section, we consider graphs that are not 3-connected. A consequence of Theorem 1.1 is that any QMSTP instance defined on a graph with fewer than four edges will be linearizable, because every symmetric square matrix of order at most three is weak sum. Therefore, we focus our attention on graphs having at least four edges.

Theorem 2.1. *Let $G = (V, E)$ be a simple connected graph with $|E| \geq 4$. If G is not 3-connected, then there exists a linearizable QMSTP instance on G whose symmetric cost matrix W , as defined in (3), is not weak sum.*

Proof. By Theorem 1.4, it is sufficient to find an equation of the form (8) that is valid for $\text{aff}\{P(G)\}$, but for which there exists no vector $\bar{\alpha} \in \mathbb{R}^m$ such that (9) can be expressed as a linear combination of the equations (6).

Since the graph is not 3-connected, we know that there exist two vertices, say u and v , whose removal (along with any incident edges) would disconnect the graph. Let ℓ be the edge connecting vertices u and v , if it exists, and define $L = E(\{u, v\})$, so that either $L = \{\ell\}$ or $L = \emptyset$. There must exist two subgraphs of G , say $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ such that $V_1 \cup V_2 = V$, $V_1 \cap V_2 = \{u, v\}$, $V_1 \setminus \{u, v\}$ and $V_2 \setminus \{u, v\}$ are both nonempty, $E_1 \cup E_2 = E$, $E_1 \cap E_2 = L$, and no edge connects a vertex in $V_1 \setminus \{u, v\}$ to a vertex in $V_2 \setminus \{u, v\}$. See Figure 1 for an illustration. Without loss of generality, we assume the edge labels are such that $e < f$ for all $e \in E_1 \setminus L$ and $f \in E_2 \setminus L$.

There are two cases to consider. First, suppose that either $E_1 \setminus L$ or $E_2 \setminus L$ contains exactly one edge. Without loss of generality, we will assume that $E_1 \setminus L$ contains exactly one edge, call it e^* , and therefore V_1

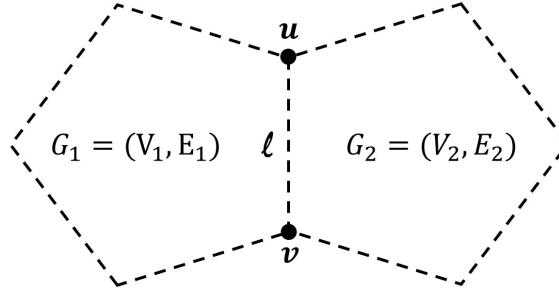


Figure 1: The general form of a non-3-connected graph used in the proof of Theorem 2.1.

contains exactly one vertex in addition to u and v . Any spanning tree T of G must contain edge e^* and $|V| - 2$ additional edges from E_2 . This implies that, for any edge $f \in E_2$, the equation

$$\sum_{e \in E_2 \setminus \{f\}} x_e x_f = (|V| - 3)x_f \quad (10)$$

is true for any spanning tree. Therefore, for any edge $f \in E_2$, the equation

$$\sum_{\substack{e \in E_2 \\ e < f}} y_{ef} + \sum_{\substack{e \in E_2 \\ e > f}} y_{fe} - (|V| - 3)x_f = 0 \quad (11)$$

holds for all $(x, y) \in P(G)$, meaning that (11) is valid for $\text{aff}\{P(G)\}$. Note that (11) is an equation of the form (8).

All that remains to show is that there exists no vector $\bar{\alpha} \in \mathbb{R}^m$ such that the equation

$$\sum_{e \in E} \bar{\alpha}_e x_e + \sum_{\substack{e \in E_2 \\ e < f}} y_{ef} + \sum_{\substack{e \in E_2 \\ e > f}} y_{fe} = 0 \quad (12)$$

can be expressed as a linear combination of the m equations (6). Suppose that there *did* exist such a linear combination, and let $\lambda \in \mathbb{R}^m$ be the multipliers of that linear combination. Arbitrarily choose two edges $i, j \in E_2 \setminus \{f\}$, which must exist because $|E| \geq 4$ (for notational convenience, assume the edges are labeled such that $e^* < f < i < j$). Since the coefficient of y_{ij} in (12) is 0, we must have that $\lambda_i + \lambda_j = 0$, so it cannot be the case that both λ_i and λ_j are positive. Without loss of generality, assume that $\lambda_i \leq 0$. Since the coefficient of y_{fi} in (12) is 1, it must be true that $\lambda_f = 1 - \lambda_i$, so $\lambda_f > 0$. Furthermore, since the coefficient of y_{e^*f} in (12) is zero, it must be true that $\lambda_{e^*} = -\lambda_f$, so $\lambda_{e^*} < 0$. Hence, $\lambda_{e^*} + \lambda_i < 0$, but the coefficient of y_{e^*i} in (12) is zero, a contradiction.

For the second case, we assume that both $|E_1 \setminus L| \geq 2$ and $|E_2 \setminus L| \geq 2$. Consider any spanning tree T of G ,

which must contain $|V| - 1 = |V_1| + |V_2| - 3$ edges. Note that T can contain at most $|V_1| - 1$ edges from E_1 , and at most $|V_2| - 1$ edges from E_2 , otherwise a cycle would be formed. There are only three possibilities: (i) T consists of $|V_1| - 2$ edges from $E_1 \setminus L$, $|V_2| - 2$ edges from $E_2 \setminus L$, and also edge ℓ , (ii) T consists of $|V_1| - 1$ edges from $E_1 \setminus L$ and $|V_2| - 2$ edges from $E_2 \setminus L$, or (iii) T consists of $|V_1| - 2$ edges from $E_1 \setminus L$ and $|V_2| - 1$ edges from $E_2 \setminus L$. For any of these possibilities, the equation

$$\left(\sum_{e \in E_1 \setminus L} x_e - (|V_1| - 2) \right) \left(\sum_{e \in E_2 \setminus L} x_e - (|V_2| - 2) \right) = 0 \quad (13)$$

is true. Therefore, the equation

$$\sum_{e \in E_1 \setminus L} \sum_{f \in E_2 \setminus L} y_{ef} - \left[(|V_1| - 2) \sum_{e \in E_2 \setminus L} x_e \right] - \left[(|V_2| - 2) \sum_{e \in E_1 \setminus L} x_e \right] = -(|V_1| - 2)(|V_2| - 2) \quad (14)$$

holds for all $(\mathbf{x}, \mathbf{y}) \in P(G)$, meaning that (14) is valid for $\text{aff}\{P(G)\}$. Note that (14) is an equation of the form (8).

All that remains to show is that there exists no vector $\bar{\alpha} \in \mathbb{R}^m$ such that the equation

$$\sum_{e \in E} \bar{\alpha}_e x_e + \sum_{e \in E_1 \setminus L} \sum_{f \in E_2 \setminus L} y_{ef} = 0 \quad (15)$$

can be expressed as a linear combination of the m equations (6). Suppose that there *did* exist such a linear combination, and let $\lambda \in \mathbb{R}^m$ be the multipliers of that linear combination. Arbitrarily choose edges $i \in E_1 \setminus L$ and $j \in E_2 \setminus L$. Since the coefficient of y_{ij} in (15) is 1, we must have that $\lambda_i + \lambda_j = 1$, so at least one of λ_i and λ_j must be positive. Without loss of generality, assume that $\lambda_i > 0$. Now choose any additional two edges $g \in E_1 \setminus L$ and $h \in E_2 \setminus L$ (for notational convenience, assume the edges are labeled such that $i < g < j < h$). Since the coefficient of y_{ig} in (15) is zero, it must be true that $\lambda_g = -\lambda_i$, so $\lambda_g < 0$. Furthermore, since the coefficient of y_{gh} in (15) is 1, it must be true that $\lambda_h = 1 - \lambda_g$, so $\lambda_h > 1$. Hence, $\lambda_i + \lambda_h > 1$, but the coefficient of y_{ih} in (15) is 1, a contradiction. $\square$

3 Linearizable QMSTPs on 3-Connected Graphs

In this section, we show that the sufficient condition for QMSTP linearizability of Theorem 1.2 (or, equivalently, the sufficient condition of Theorem 1.1) is also necessary if the QMSTP instance is defined on a 3-connected graph. The proof relies on machinery involving edge contractions within graphs, so we first

define some necessary notation and terminology.

Given a graph $G = (V, E)$, suppose that we wish to form a new graph, say $G' = (V', E')$, by contracting some edge $\ell \in E$. Let the vertices u and v be the two endpoints of ℓ , and let

$$C(\ell) = \{(e, f) \in E \times E : e < f \text{ and the edges } \{e, f, \ell\} \text{ form a triangle}\}. \quad (16)$$

For example, in the graph G pictured in Figure 2, $C(1) = \{(2, 3)\}$ and $C(3) = \{(1, 2), (4, 5)\}$.

The contraction operation consists of two steps. First, we remove the edge ℓ and merge the vertices u and v into one new vertex, say w , meaning that

$$V' = (V \setminus \{u, v\}) \cup \{w\}. \quad (17)$$

During this step, any edge that was incident to either u or v in G becomes incident to w instead, which results in every pair of edges in $C(\ell)$ becoming a double edge. The second step removes exactly one edge from each pair in $C(\ell)$ to ensure that G' is a simple graph. Call this set of removed edges Γ , noting that there are $2^{|C(\ell)|}$ possible sets Γ . In this way, we have that $E' = E \setminus (\{\ell\} \cup \Gamma)$. Since the contraction operation depends on both ℓ and Γ , we will use the notation $G' = G[\ell, \Gamma]$. Figure 2 gives an example of a graph G along with two different contractions. Notice that the edges in the contractions retain their labels from the original graph.

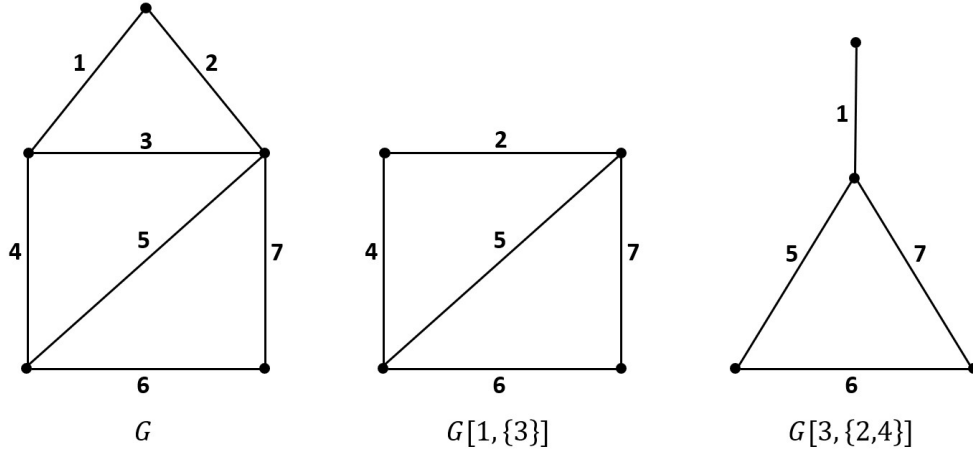


Figure 2: An example of a graph G along with the two contractions $G[1, \{3\}]$ and $G[3, \{2, 4\}]$.

Lemma 3.1. *Given a graph $G = (V, E)$, an edge $\ell \in E$, and a corresponding set Γ as defined above, there is a 1-1 correspondence between spanning trees of G that contain ℓ but no edges in Γ , and spanning trees of $G' = G[\ell, \Gamma]$.*

Proof. Consider any spanning tree $T = (V, E_T)$ of G such that $\ell \in E_T$ and $E_T \cap \Gamma = \emptyset$. Then $T' = (V', E_T \setminus \{\ell\})$ is a spanning tree of G' , with V' defined as in (17). Conversely, consider any spanning tree $T' = (V', E_{T'})$ of G' . Then $T = (V, E_{T'} \cup \{\ell\})$ is a spanning tree of G that contains ℓ and no edges in Γ . $\square$

To illustrate the result of Lemma 3.1, consider the graph G and the contraction $G[3, \{2, 4\}]$ pictured in Figure 2. There are three spanning trees of G that contain the edge 3, but do not contain either of the edges 2 or 4. These spanning trees are defined by the edge sets $\{1, 3, 5, 6\}$, $\{1, 3, 5, 7\}$, and $\{1, 3, 6, 7\}$. Likewise, there are three spanning trees of the contraction $G[3, \{2, 4\}]$, defined by the edge sets $\{1, 5, 6\}$, $\{1, 5, 7\}$, and $\{1, 6, 7\}$.

Definition 3.1. Given a 3-connected graph $G = (V, E)$, we say that an edge $\ell \in E$ is **3-contractible** if $G' = G[\ell, \Gamma]$ is also 3-connected for any choice of the set Γ . (Note that if G' is 3-connected for one choice of Γ , then it is 3-connected for any choice of Γ .)

To illustrate Definition 3.1, consider the 3-connected graph G pictured in Figure 3. The edge e_1 is not 3-contractible, because the contraction $G[e_1, \Gamma]$ is not 3-connected for any choice of the set Γ . The edge e_2 is 3-contractible, however, because the contraction $G[e_2, \Gamma]$ is 3-connected for any choice of the set Γ .

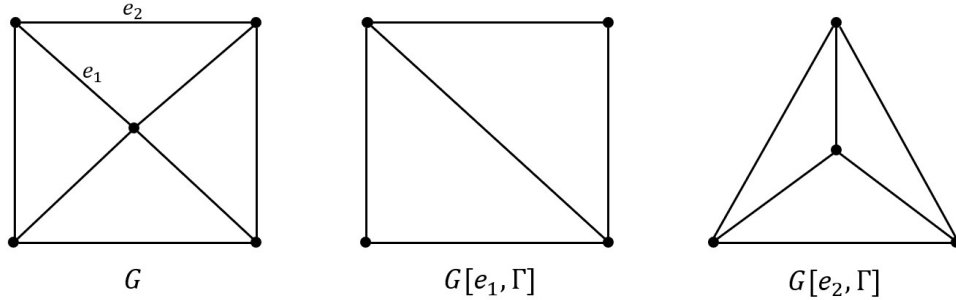


Figure 3: For the 3-connected graph G , the edge e_1 is not 3-contractible, but the edge e_2 is.

Theorem 3.2. Let $G = (V, E)$ be a simple 3-connected graph with $|V| = n$ and $|E| = m$. Then every linearizable QMSTP instance defined on G has a weak sum symmetric cost matrix W , as defined in (3).

Proof. By Theorem 1.4, it is sufficient to show that given any equation of the form (8) that is valid for $\text{aff}\{P(G)\}$, there exists a vector $\bar{\alpha} \in \mathbb{R}^m$ such that (9) can be expressed as a linear combination of the equations (6). In particular, consider some arbitrary equation of the form (8), say

$$\sum_{e \in E} \hat{\alpha}_e x_e + \sum_{e \in E} \sum_{\substack{f \in E \\ f > e}} \hat{\beta}_{ef} y_{ef} = \hat{\kappa}, \quad (18)$$

that is valid for $\text{aff}\{P(G)\}$.

We will now define two sets, given below in (20) and (21), that form a partition of the edge pairs that define the variables y_{ef} in (18). Because elements of these sets require edge indices to appear in increasing order, it is often convenient for us to adopt the notation

$$\langle j, k \rangle = \begin{cases} (j, k) & \text{if } j < k \\ (k, j) & \text{if } j > k, \end{cases} \quad (19)$$

to ensure that indices are ordered correctly. Given any three distinct edges $i, j, k \in E$, we define

$$S_{i,j,k}^E \equiv \{(e, f) \in E \times E : e < f \text{ and either } e = i, f = i, \text{ or } (e, f) = \langle j, k \rangle\}, \quad (20)$$

and

$$R_{i,j,k}^E \equiv \{(e, f) \in E \times E : e < f\} \setminus S_{i,j,k}^E. \quad (21)$$

Note that $|S_{i,j,k}^E| = m$ and $|R_{i,j,k}^E| = \frac{m^2 - 3m}{2}$.

Notice that within the left-hand-side coefficient matrix of the m equations (6), the columns corresponding to those variables y_{ef} having $(e, f) \in S_{i,j,k}^E$ form an $m \times m$ submatrix of zeros and ones. This submatrix is the unsigned vertex-edge incidence matrix of a graph with vertex set E and edge set $\{\{i, e\} : e \in E \setminus \{i\}\} \cup \{\{j, k\}\}$. Specifically, this graph is a star centered at i , together with the additional edge $\{j, k\}$, so it is connected and contains a triangle formed by edges $\{i, j\}, \{i, k\}$ and $\{j, k\}$. Therefore, the submatrix has full rank [11]. This means that we can take a linear combination of the equations (6) so that the resulting coefficients on the variables y_{ef} having $(e, f) \in S_{i,j,k}^E$ are any values that we want them to be. Specifically, there must exist a linear combination of the equations (6) that yields an equation of the form

$$\sum_{e \in E} \bar{\alpha}_e x_e + \sum_{(e,f) \in S_{i,j,k}^E} \hat{\beta}_{ef} y_{ef} + \sum_{(e,f) \in R_{i,j,k}^E} \bar{\beta}_{ef} y_{ef} = 0, \quad (22)$$

where the coefficients $\hat{\beta}_{ef}$ having $(e, f) \in S_{i,j,k}^E$ are the same as those found in (18). Since the equations (6) are valid for $\text{aff}\{P(G)\}$, we know that (22) is valid for $\text{aff}\{P(G)\}$ as well.

The final step of the proof is to show that given any equation of the form (8) that is valid for $\text{aff}\{P(G)\}$ and any distinct edges $i, j, k \in E$, each β_{ef} having $(e, f) \in R_{i,j,k}^E$ is completely determined by the β_{ef} having $(e, f) \in S_{i,j,k}^E$. If we can prove this final step, recognizing that (18) and (22) are both of the form (8) and are both valid for $\text{aff}\{P(G)\}$, then the coefficients $\hat{\beta}_{ef}$ of (18) and $\bar{\beta}_{ef}$ of (22) would then satisfy $\hat{\beta}_{ef} = \bar{\beta}_{ef}$ for

all $(e, f) \in R_{i,j,k}^E$, so that equation (22) would be expressible as

$$\sum_{e \in E} \bar{\alpha}_e x_e + \sum_{e \in E} \sum_{\substack{f \in E \\ f > e}} \hat{\beta}_{ef} y_{ef} = 0,$$

meaning that (9) can be expressed as a linear combination of the equations (6), as desired. We will proceed by proving the stronger result that each β_{ef} having $(e, f) \in R_{i,j,k}^E$ can be defined in terms of only those β_{st} having $(s, t) \in S_{i,j,k}^E$ in exactly the following manner:

$$\beta_{ef} = \beta_{(i,e)} + \beta_{(i,f)} - \beta_{(i,j)} - \beta_{(i,k)} + \beta_{(j,k)}. \quad (23)$$

We will use the shorthand notation $\mathbb{E}_{i,j,k}^{(e,f)}$ to represent the single equation (23), and the notation $\mathbb{E}_{i,j,k}^E$ to represent the entire system of $|R_{i,j,k}^E| = \frac{m^2 - 3m}{2}$ such equations. Thus, to prove the final step, we simply need to show that the equations of the system $\mathbb{E}_{i,j,k}^E$ in the variables β from (8) hold true.

The proof is inductive on the number of vertices n in the graph G . For the $n = 4$ base case, we consider the complete graph on four vertices, which is the smallest possible 3-connected graph, and the only 3-connected graph having four vertices. The proofs for both the $n = 4$ and $n = 5$ base cases are found in Appendix B.

For the inductive step, let $n \geq 6$, and assume that, for every 3-connected graph $\tilde{G} = (\tilde{V}, \tilde{E})$ having $|\tilde{V}| = n - 1$, the coefficients of every equation of the form (8) valid for $\text{aff}\{P(\tilde{G})\}$ satisfy $\mathbb{E}_{i,j,k}^{\tilde{E}}$ for every choice of distinct edges $i, j, k \in \tilde{E}$. Now consider a 3-connected graph $G = (V, E)$ with $|V| = n$, together with an arbitrary equation of the form (8) that is valid for $\text{aff}\{P(G)\}$. The key to the inductive argument is that we can identify six distinct edges within G – three 3-contractible edges which we will call ℓ_1, ℓ_2 , and ℓ_3 , and three additional edges which we will call q, r , and s – such that for any pair of edges $(e, f) \in R_{q,r,s}^E$, there exists an edge $\ell \in L = \{\ell_1, \ell_2, \ell_3\}$ and a Γ such that the edges e, f, q, r and s all remain in the contracted graph $G[\ell, \Gamma]$ (that is, $\{e, f, q, r, s\} \subseteq E \setminus (\{\ell\} \cup \Gamma)$). For clarity of the argument, we refer to this step that we must still establish as the *six-edge identification step*.

By the inductive hypothesis and Theorem 5.2, completion of the six-edge identification step will imply that the equation $\mathbb{E}_{q,r,s}^{(e,f)}$ in the variables β from (8) holds for any $(e, f) \in R_{q,r,s}^E$. Therefore, Theorem 5.1 from Appendix A will give that the equations of the system $\mathbb{E}_{i,j,k}^E$ in the variables β from (8) hold true for any distinct edges $i, j, k \in E$, and the proof will be complete.

We first consider the six-edge identification step for the case where $|V| = 6$. In this case, G must contain one of the non-isomorphic minimally 3-connected graphs pictured in Figure 4 as a spanning subgraph [2, 6], and we can identify the six edges as shown in the figure. (A 3-connected graph is *minimally* 3-connected if

the removal of any edge results in a graph that is no longer 3-connected.)

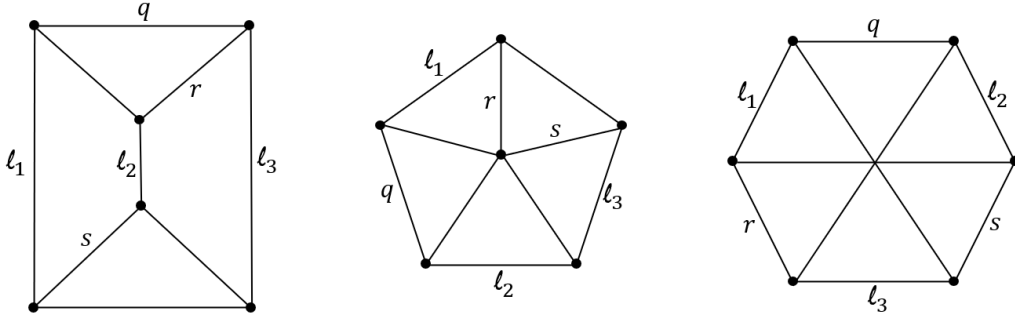


Figure 4: The minimally 3-connected graphs on 6 vertices.

In each graph pictured in Figure 4, contracting any $\ell \in L = \{\ell_1, \ell_2, \ell_3\}$ results in a wheel on five vertices. The corresponding contraction of G therefore contains a wheel as a spanning subgraph, meaning that it is 3-connected. Thus, ℓ_1, ℓ_2 , and ℓ_3 are 3-contractible. Observe that, in each configuration in Figure 4, the six selected edges form a connected spanning subgraph containing a cycle C that includes q, r , and s . Consider any two distinct edges $e, f \in E$, and define $H = (V, E_H)$, with $E_H = \{e, f, q, r, s\}$. Note that H could contain fewer than five edges depending on the selection of e and f . For the sake of contradiction, suppose that none of ℓ_1, ℓ_2 , or ℓ_3 admits a choice of Γ for which every edge of H is retained in the contracted graph.

Then for each $\ell \in L$, either $\ell \in E_H$, or two edges of H form a triangle with ℓ . (If neither condition held, ℓ could be contracted while retaining every edge of H .) Consequently, the endpoints of ℓ are joined by a path of length at most two in H . Adding the edges in L to H therefore cannot join different connected components; call this new graph H^+ . Since H^+ contains the connected spanning subgraph formed by the six selected edges, we know that H^+ is connected and spans V . Hence H is also connected and spans V , since H has the same number of connected components as H^+ . Since H has six vertices and at most five edges, it must be a spanning tree.

Let k denote the length of C . Replace each edge of C by the unique path in H joining its endpoints. This produces a closed walk in H visiting all k vertices of C . The edges q, r , and s each contribute one step to the walk, while each of the remaining $k - 3$ edges belongs to L and therefore contributes at most two steps. Thus the walk has length at most $3 + 2(k - 3) = 2k - 3$.

On the other hand, the edges traversed by this walk form a connected subgraph of the tree H containing at least k vertices, and hence at least $k - 1$ edges. Since a closed walk in a tree traverses every edge it uses at least twice, its length is at least $2(k - 1) = 2k - 2$, a contradiction.

Therefore, some $\ell \in L$ admits a choice of Γ such that $E_H \subseteq E \setminus (\{\ell\} \cup \Gamma)$. Since e and f were arbitrary,

this establishes the six-edge identification step for $|V| = 6$.

It remains to complete the six-edge identification step when $|V| \geq 7$. By [7, 10], the set of 3-contractible edges cannot be covered by two vertices. Hence there must exist two vertex-disjoint 3-contractible edges, call them ℓ_1 and ℓ_2 , as well as a third distinct 3-contractible edge ℓ_3 . Since the three edges in $L = \{\ell_1, \ell_2, \ell_3\}$ do not form a cycle (because ℓ_1 and ℓ_2 are disjoint) and G is connected, we know that there exists a spanning tree $T = (V, E_T)$ containing all three edges in L . Moreover, $|E_T \setminus L| = n - 4 \geq 3$, so we may choose three distinct edges $q, r, s \in E_T \setminus L$. The subgraph $F = (V, E_F)$, where $E_F = L \cup \{q, r, s\}$, is therefore a forest with six edges and exactly $n - 6$ connected components, counting isolated vertices (see, for example, [5, Section 2.2, p. 20]).

Now consider any two distinct edges $e, f \in E$, and define $H = (V, E_H)$, with $E_H = \{e, f, q, r, s\}$. Since H has at most five edges, it has at least $n - 5$ connected components. For the sake of contradiction, suppose that none of ℓ_1, ℓ_2 , or ℓ_3 admits a choice of Γ for which every edge of H is retained in the contracted graph.

As in the $|V| = 6$ case, the endpoints of each $\ell \in L$ are then joined by a path of length at most two in H . Consequently, adding the edges in L to H does not join different connected components, so the resulting graph (again, call it H^+) still has at least $n - 5$ components. We know that H^+ contains F as a spanning subgraph and therefore has at most $n - 6$ components, a contradiction.

Thus some $\ell \in L$ admits a choice of Γ such that $E_H \subseteq E \setminus (\{\ell\} \cup \Gamma)$. Since e and f were arbitrary, this completes the six-edge identification step for $|V| \geq 7$. $\square$

The main result of the paper, which completely resolves the open question studied by [3] and [12], follows as a direct result of Theorems 2.1 and 3.2.

Theorem 3.3. *Let $G = (V, E)$ be a simple connected graph with $|E| \geq 4$. Then every linearizable QMSTP instance on G has a weak sum symmetric cost matrix W , as defined in (3), if and only if G is 3-connected.*

4 Conclusions

This paper uses the polyhedral framework of [12] to resolve an open question of [3] regarding linearizable QMSTP instances. In particular, we show that among simple connected graphs with at least four edges, the graphs for which every linearizable QMSTP instance has a weak sum symmetric cost matrix are exactly the 3-connected graphs. A remaining open problem is to obtain a general characterization of linearizable QMSTP instances on graphs that are not 3-connected.

5 Appendix A

Consider an instance of QMSTP defined on a graph $G = (V, E)$ along with an equation of the form (8) that is valid for $\text{aff}\{P(G)\}$. This appendix proves a result about equations defined by (23) in the variables β from (8), given in Theorem 5.1, that is used in the proof of Theorem 3.2. Then Theorem 5.2 gives a result that is critical to the inductive argument in the proof of Theorem 3.2.

Theorem 5.1. *If the equations of the system $\mathbb{E}_{i,j,k}^E$ defined by (23) hold for some distinct edges $i, j, k \in E$, then the equations of the system $\mathbb{E}_{r,s,t}^E$ hold for every choice of distinct edges $r, s, t \in E$.*

Proof. Suppose that the equations of $\mathbb{E}_{i,j,k}^E$ hold. We first construct a vector $\lambda \in \mathbb{R}^m$ such that

$$\beta_{\langle e,f \rangle} = \lambda_e + \lambda_f \quad (24)$$

for all distinct $e, f \in E$. Define

$$\lambda_i = \frac{\beta_{\langle i,j \rangle} + \beta_{\langle i,k \rangle} - \beta_{\langle j,k \rangle}}{2}, \text{ and } \lambda_q = \beta_{\langle i,q \rangle} - \lambda_i \text{ for all } q \in E \setminus \{i\}.$$

By construction, (24) holds for every pair of edges involving i . It also holds for the pair $\langle j, k \rangle$, since

$$\lambda_j + \lambda_k = \beta_{\langle i,j \rangle} + \beta_{\langle i,k \rangle} - 2\lambda_i = \beta_{\langle j,k \rangle}.$$

Thus (24) holds for all pairs of edges in $S_{i,j,k}^E$. For every remaining pair $(e, f) \in R_{i,j,k}^E$, the corresponding equation of $\mathbb{E}_{i,j,k}^E$ gives

$$\begin{aligned} \lambda_e + \lambda_f &= \beta_{\langle i,e \rangle} + \beta_{\langle i,f \rangle} - 2\lambda_i \\ &= \beta_{\langle i,e \rangle} + \beta_{\langle i,f \rangle} - \beta_{\langle i,j \rangle} - \beta_{\langle i,k \rangle} + \beta_{\langle j,k \rangle} \\ &= \beta_{ef}. \end{aligned}$$

Therefore, (24) holds for all pairs of distinct edges $e, f \in E$.

Now consider any distinct $r, s, t \in E$ and any $(e, f) \in R_{r,s,t}^E$. Then we have that

$$\begin{aligned} &\beta_{\langle r,e \rangle} + \beta_{\langle r,f \rangle} - \beta_{\langle r,s \rangle} - \beta_{\langle r,t \rangle} + \beta_{\langle s,t \rangle} \\ &= (\lambda_r + \lambda_e) + (\lambda_r + \lambda_f) - (\lambda_r + \lambda_s) - (\lambda_r + \lambda_t) + (\lambda_s + \lambda_t) \\ &= \lambda_e + \lambda_f = \beta_{ef}. \end{aligned}$$

This is the desired equation of $\mathbb{E}_{r,s,t}^E$. Since $(e, f) \in R_{r,s,t}^E$ and the distinct edges $r, s, t \in E$ were arbitrary, the conclusion follows. $\square$

Theorem 5.2. *Consider an instance of QMSTP defined on a 3-connected graph $G = (V, E)$ with $|V| \geq 6$, along with an equation (EQ1) of the form (8) that is valid for $\text{aff}\{P(G)\}$. Furthermore, suppose we know that for any instance of QMSTP defined on a 3-connected graph $\tilde{G} = (\tilde{V}, \tilde{E})$ having $|\tilde{V}| = |V| - 1$, along with any equation (EQ2) of the form (8) that is valid for $\text{aff}\{P(\tilde{G})\}$, the equations of the system $\mathbb{E}_{i,j,k}^{\tilde{E}}$ in the variables β from (EQ2) hold true for any three distinct edges $i, j, k \in \tilde{E}$.*

Let $\ell \in E$ be a 3-contractible edge, and let $q, r, s \in E \setminus \{\ell\}$ be distinct. If $(e, f) \in R_{q,r,s}^E$, and there exists a contraction $G[\ell, \Gamma]$ such that $\{e, f, q, r, s\} \subseteq E \setminus (\{\ell\} \cup \Gamma)$, then the equation $\mathbb{E}_{q,r,s}^{(e,f)}$ holds for the coefficients β from (EQ1).

Proof. Suppose that (EQ1) is given by

$$\sum_{e \in E} \hat{\alpha}_e x_e + \sum_{e \in E} \sum_{\substack{f \in E \\ f > e}} \hat{\beta}_{ef} y_{ef} = \hat{\kappa}, \quad (25)$$

and arbitrarily choose any $(e, f) \in R_{q,r,s}^E$ that satisfies the conditions given in the statement of the theorem. That is, there must exist a contraction $G' = (V', E') = G[\ell, \Gamma]$ for which $e, f, q, r, s \in E'$, meaning that $e, f, q, r, s \notin \Gamma$.

The fact that (25) is valid for $\text{aff}\{P(G)\}$ means that it holds for all binary $\mathbf{x} \in \mathbf{X}(G)$, with $y_{ef} = x_e x_f$. Isolating only those spanning trees of G that contain edge ℓ and no edge from the set Γ , we have that

$$\sum_{e \in E'} (\hat{\alpha}_e + \hat{\beta}_{(e,\ell)}) x_e + \sum_{e \in E'} \sum_{\substack{f \in E' \\ f > e}} \hat{\beta}_{ef} y_{ef} = \hat{\kappa} - \hat{\alpha}_\ell \quad (26)$$

holds for all $(\mathbf{x}, \mathbf{y})$ corresponding to such spanning trees. Lemma 3.1 then gives us that (26) is an equation of the form (8) that is valid for $\text{aff}\{P(G')\}$. Since ℓ is 3-contractible, G' is 3-connected. Moreover, $|V'| = |V| - 1$, meaning that the hypothesis applies to (26) and therefore the equations of the system $\mathbb{E}_{q,r,s}^{E'}$ must hold true, which includes the equation $\mathbb{E}_{q,r,s}^{(e,f)}$. $\square$

6 Appendix B

Theorem 3.2: $n = 4$ base case. Consider an instance of QMSTP defined on the complete graph with four vertices, $G = (V, E) = K_4$, along with an equation of the form (8) that is valid for $\text{aff}\{P(G)\}$. The equations

of the system $\mathbb{E}_{i,j,k}^E$ defined by (23) in the variables β from (8) hold true for any distinct edges $i, j, k \in E$.

Proof. Note that portions of this argument appear in [12], but are included here for completeness. Consider an instance of QMSTP defined on the complete graph $G = (V, E) = K_4$ with $n = 4$ vertices and $m = 6$ edges, as shown in Figure 5, along with some arbitrary equation of the form (8) that is valid for $\text{aff}\{P(G)\}$.

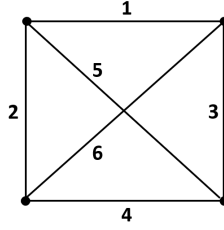


Figure 5: The complete graph with $n = 4$ vertices and $m = 6$ edges.

There are $4^{4-2} = 16$ solutions to $x \in \mathbf{X}(G)$ of (1), x binary, meaning that $|P(G)| = 16$, with each of the 16 corresponding spanning trees containing 3 edges. Notationally, we distinguish each of these solutions with a row vector $\theta \in \{1, 2, 3, 4, 5, 6\}^3$ so that x_θ denotes that binary solution having $x_{\theta(i)} = 1$ for all $i \in \{1, 2, 3\}$, and all other $x_e = 0$. For example, $x_{(1,2,4)}$ has $x_1 = x_2 = x_4 = 1$ and $x_3 = x_5 = x_6 = 0$. Define y_θ appropriately so that $(x_\theta, y_\theta) \in P(G)$. Since (8) is satisfied by all $(x, y) \in P(G)$, the 16 equations listed below must hold, where $L(x_\theta)$ denotes that linear equation obtained by setting $x = x_\theta$ and $y = y_\theta$ within (8).

$$(A1) \quad L(x_{(1,2,3)}) : \alpha_1 + \alpha_2 + \alpha_3 + \beta_{12} + \beta_{13} + \beta_{23} = \kappa$$

$$(A2) \quad L(x_{(1,2,4)}) : \alpha_1 + \alpha_2 + \alpha_4 + \beta_{12} + \beta_{14} + \beta_{24} = \kappa$$

$$(A3) \quad L(x_{(1,2,5)}) : \alpha_1 + \alpha_2 + \alpha_5 + \beta_{12} + \beta_{15} + \beta_{25} = \kappa$$

$$(A4) \quad L(x_{(1,3,4)}) : \alpha_1 + \alpha_3 + \alpha_4 + \beta_{13} + \beta_{14} + \beta_{34} = \kappa$$

$$(A5) \quad L(x_{(1,3,6)}) : \alpha_1 + \alpha_3 + \alpha_6 + \beta_{13} + \beta_{16} + \beta_{36} = \kappa$$

$$(A6) \quad L(x_{(1,4,5)}) : \alpha_1 + \alpha_4 + \alpha_5 + \beta_{14} + \beta_{15} + \beta_{45} = \kappa$$

$$(A7) \quad L(x_{(1,4,6)}) : \alpha_1 + \alpha_4 + \alpha_6 + \beta_{14} + \beta_{16} + \beta_{46} = \kappa$$

$$(A8) \quad L(x_{(1,5,6)}) : \alpha_1 + \alpha_5 + \alpha_6 + \beta_{15} + \beta_{16} + \beta_{56} = \kappa$$

$$(A9) \quad L(x_{(2,3,4)}) : \alpha_2 + \alpha_3 + \alpha_4 + \beta_{23} + \beta_{24} + \beta_{34} = \kappa$$

$$(A10) \quad L(x_{(2,3,5)}) : \alpha_2 + \alpha_3 + \alpha_5 + \beta_{23} + \beta_{25} + \beta_{35} = \kappa$$

$$(A11) \quad L(x_{(2,3,6)}) : \alpha_2 + \alpha_3 + \alpha_6 + \beta_{23} + \beta_{26} + \beta_{36} = \kappa$$

$$(A12) \quad L(x_{(2,4,6)}) : \alpha_2 + \alpha_4 + \alpha_6 + \beta_{24} + \beta_{26} + \beta_{46} = \kappa$$

$$(A13) \quad L(\mathbf{x}_{(2,5,6)}) : \alpha_2 + \alpha_5 + \alpha_6 + \beta_{25} + \beta_{26} + \beta_{56} = \kappa$$

$$(A14) \quad L(\mathbf{x}_{(3,4,5)}) : \alpha_3 + \alpha_4 + \alpha_5 + \beta_{34} + \beta_{35} + \beta_{45} = \kappa$$

$$(A15) \quad L(\mathbf{x}_{(3,5,6)}) : \alpha_3 + \alpha_5 + \alpha_6 + \beta_{35} + \beta_{36} + \beta_{56} = \kappa$$

$$(A16) \quad L(\mathbf{x}_{(4,5,6)}) : \alpha_4 + \alpha_5 + \alpha_6 + \beta_{45} + \beta_{46} + \beta_{56} = \kappa$$

In accordance with (20) and (21), we have that $S_{1,5,6}^E = \{(1,2), (1,3), (1,4), (1,5), (1,6), (5,6)\}$, and $R_{1,5,6}^E = \{(2,3), (2,4), (2,5), (2,6), (3,4), (3,5), (3,6), (4,5), (4,6)\}$. The following linear combinations of equations (A1)–(A16) above recursively express the nine coefficients β_{ef} having $(e, f) \in R_{1,5,6}^E$ in terms of the six coefficients β_{ef} having $(e, f) \in S_{1,5,6}^E$.

$$(A2) - (A3) - (A12) + (A13) : \beta_{46} = \beta_{14} - \beta_{15} + \beta_{56}$$

$$(A1) - (A2) - (A11) + (A12) : \beta_{36} = \beta_{13} - \beta_{14} + \beta_{46} = \beta_{13} - \beta_{14} + (\beta_{14} - \beta_{15} + \beta_{56}) \\ = \beta_{13} - \beta_{15} + \beta_{56}$$

$$(A4) - (A7) - (A14) + (A16) : \beta_{35} = \beta_{13} - \beta_{16} + \beta_{56}$$

$$(A2) - (A3) - (A9) + (A10) : \beta_{34} = \beta_{14} - \beta_{15} + \beta_{35} = \beta_{14} - \beta_{15} + (\beta_{13} - \beta_{16} + \beta_{56}) \\ = \beta_{13} + \beta_{14} - \beta_{15} - \beta_{16} + \beta_{56}$$

$$(A4) - (A5) - (A14) + (A15) : \beta_{45} = \beta_{14} - \beta_{16} + \beta_{56}$$

$$(A6) - (A3) - (A16) + (A13) : \beta_{26} = \beta_{12} - \beta_{14} + \beta_{46} = \beta_{12} - \beta_{14} + (\beta_{14} - \beta_{15} + \beta_{56}) \\ = \beta_{12} - \beta_{15} + \beta_{56}$$

$$(A6) - (A7) - (A10) + (A11) : \beta_{25} = \beta_{15} - \beta_{16} + \beta_{26} - \beta_{35} + \beta_{36} + \beta_{45} - \beta_{46} \\ = \beta_{15} - \beta_{16} + (\beta_{12} - \beta_{15} + \beta_{56}) - (\beta_{13} - \beta_{16} + \beta_{56}) + (\beta_{13} - \beta_{15} + \beta_{56}) \\ + (\beta_{14} - \beta_{16} + \beta_{56}) - (\beta_{14} - \beta_{15} + \beta_{56}) \\ = \beta_{12} - \beta_{16} + \beta_{56}$$

$$(A7) - (A8) - (A12) + (A13) : \beta_{24} = \beta_{14} - \beta_{15} + \beta_{25} = \beta_{14} - \beta_{15} + (\beta_{12} - \beta_{16} + \beta_{56}) \\ = \beta_{12} + \beta_{14} - \beta_{15} - \beta_{16} + \beta_{56}$$

$$(A11) - (A12) - (A15) + (A16) : \beta_{23} = \beta_{24} + \beta_{35} - \beta_{45} \\ = (\beta_{12} + \beta_{14} - \beta_{15} - \beta_{16} + \beta_{56}) + (\beta_{13} - \beta_{16} + \beta_{56}) - (\beta_{14} - \beta_{16} + \beta_{56}) \\ = \beta_{12} + \beta_{13} - \beta_{15} - \beta_{16} + \beta_{56}$$

Notice that these equations can be written as the following system, which is exactly $\mathbb{E}_{1,5,6}^E$ defined by (23).

$$\begin{aligned} \mathbb{E}_{1,5,6}^E : \quad & \beta_{23} = \beta_{12} + \beta_{13} - \beta_{15} - \beta_{16} + \beta_{56} \\ & \beta_{24} = \beta_{12} + \beta_{14} - \beta_{15} - \beta_{16} + \beta_{56} \\ & \beta_{25} = \beta_{12} - \beta_{16} + \beta_{56} \\ & \beta_{26} = \beta_{12} - \beta_{15} + \beta_{56} \\ & \beta_{34} = \beta_{13} + \beta_{14} - \beta_{15} - \beta_{16} + \beta_{56} \\ & \beta_{35} = \beta_{13} - \beta_{16} + \beta_{56} \\ & \beta_{36} = \beta_{13} - \beta_{15} + \beta_{56} \\ & \beta_{45} = \beta_{14} - \beta_{16} + \beta_{56} \\ & \beta_{46} = \beta_{14} - \beta_{15} + \beta_{56} \end{aligned}$$

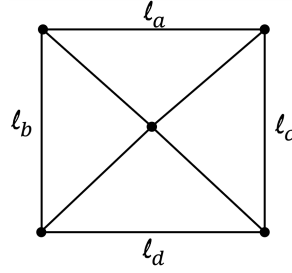
Since the equations of the system $\mathbb{E}_{1,5,6}^E$ hold true, we know by Theorem 5.1 found in Appendix A that the equations in the system $\mathbb{E}_{i,j,k}^E$ hold true for any choice of distinct edges $i, j, k \in E$. $\square$

Theorem 3.2: n = 5 base case. Consider an instance of QMSTP defined on a 3-connected graph $G = (V, E)$ with $|V| = 5$, along with an equation of the form (8) that is valid for $\text{aff}\{P(G)\}$. The equations of the system $\mathbb{E}_{i,j,k}^E$ defined by (23) in the variables β from (8) hold true for any distinct edges $i, j, k \in E$.

Proof. Consider an instance of QMSTP defined on a 3-connected graph $G = (V, E)$ with $|V| = 5$, along with some arbitrary equation of the form (8) that is valid for $\text{aff}\{P(G)\}$, say

$$\sum_{e \in E} \alpha_e x_e + \sum_{e \in E} \sum_{\substack{f \in E \\ f > e}} \beta_{ef} y_{ef} = \kappa. \quad (27)$$

The fact that (27) is valid for $\text{aff}\{P(G)\}$ means that it holds for all binary $x \in \mathbf{X}(G)$, with $y_{ef} = x_e x_f$. Note that G must contain the minimally 3-connected graph pictured in Figure 6 as a spanning subgraph [2, 6]. (A 3-connected graph is *minimally* 3-connected if the removal of any edge results in a graph that is no longer 3-connected.) The four “rim edges” in Figure 6, labeled ℓ_a, ℓ_b, ℓ_c , and ℓ_d , are 3-contractible.


 Figure 6: The minimally 3-connected graph with $n = 5$ vertices.

Let $L = \{\ell_a, \ell_b, \ell_c, \ell_d\}$. We will proceed by showing that there exists a vector $\lambda \in \mathbb{R}^m$ such that $\beta_{\langle e, f \rangle} = \lambda_e + \lambda_f$ for all distinct edges $e, f \in E$. Substituting this identity into (23) will establish the desired result.

Let $\mathcal{E}$ be the collection of all retained edge sets of the form $E' = E \setminus (\{\ell\} \cup \Gamma)$, where ℓ ranges over L and Γ ranges over all admissible choices for the contraction $G' = G[\ell, \Gamma]$. Since contracting any edge $\ell \in L$ will produce K_4 (even when G has additional edges not pictured in Figure 6), every $E' \in \mathcal{E}$ must consist of six edges.

For any such contraction $G' = (V', E')$, isolating only those spanning trees of G that contain edge ℓ and no edge from the set Γ , we have that

$$\sum_{e \in E'} (\alpha_e + \beta_{\langle e, \ell \rangle}) x_e + \sum_{e \in E'} \sum_{\substack{f \in E' \\ f > e}} \beta_{ef} y_{ef} = \kappa - \alpha_\ell \quad (28)$$

holds for all (x, y) corresponding to such spanning trees. Lemma 3.1 then gives us that (28) is an equation of the form (8) that is valid for $\text{aff}\{P(G')\}$. Applying the $n = 4$ base case, the equations of $\mathbb{E}_{p,q,r}^{E'}$ hold for any distinct $p, q, r \in E'$. Define

$$\lambda_p^{E'} = \frac{\beta_{\langle p, q \rangle} + \beta_{\langle p, r \rangle} - \beta_{\langle q, r \rangle}}{2}, \text{ and } \lambda_t^{E'} = \beta_{\langle p, t \rangle} - \lambda_p^{E'} \text{ for all } t \in E' \setminus \{p\}.$$

Using the same argument that was used to establish (24) in the proof of Theorem 5.1 in Appendix A, these definitions together with the equations of $\mathbb{E}_{p,q,r}^{E'}$ give

$$\lambda_e^{E'} + \lambda_f^{E'} = \beta_{\langle e, f \rangle} \quad (29)$$

for all distinct edges $e, f \in E'$. Moreover, for any three distinct edges $e, f, g \in E'$, (29) implies

$$\frac{\beta_{\langle e, f \rangle} + \beta_{\langle e, g \rangle} - \beta_{\langle f, g \rangle}}{2} = \frac{(\lambda_e^{E'} + \lambda_f^{E'}) + (\lambda_e^{E'} + \lambda_g^{E'}) - (\lambda_f^{E'} + \lambda_g^{E'})}{2} = \lambda_e^{E'}.$$

Thus $\lambda^{E'}$ is uniquely determined. In particular, if $E'_1, E'_2 \in \mathcal{E}$ have at least three edges in common, then $\lambda_e^{E'_1} = \lambda_e^{E'_2}$ for every $e \in E'_1 \cap E'_2$.

We next show that this agreement holds for two arbitrary sets $E'_1, E'_2 \in \mathcal{E}$. Suppose that E'_1 and E'_2 arise from contracting $\ell_1 \in L$ and $\ell_2 \in L$, respectively. If $\ell_1 = \ell_2$, then $E'_1 \cup E'_2$ excludes ℓ_1 , so $|E'_1 \cup E'_2| \leq 9$ (since $|E| \leq 10$). If ℓ_1 and ℓ_2 are distinct and share a vertex, write $\ell_1 = \{u, v\}$ and $\ell_2 = \{v, w\}$ (meaning, for example, that the edge ℓ_1 is incident to vertices u and v). Each of E'_1 and E'_2 contains at most one of the three possible edges $\{u, v\}, \{v, w\}, \{u, w\}$. Hence their union omits at least one of the ten possible edges on the vertex set V , again giving $|E'_1 \cup E'_2| \leq 9$. In either case,

$$|E'_1 \cap E'_2| = |E'_1| + |E'_2| - |E'_1 \cup E'_2| \geq 6 + 6 - 9 = 3,$$

so $\lambda_e^{E'_1} = \lambda_e^{E'_2}$ for every $e \in E'_1 \cap E'_2$.

It remains to consider the case where ℓ_1 and ℓ_2 are opposite rim edges (i.e., they do not share a vertex). Fix $e \in E'_1 \cap E'_2$. Each of the other two rim edges shares a vertex with both ℓ_1 and ℓ_2 , and at least one of them, call it ℓ_3 , differs from e . Since $\ell_3 \neq e$, an admissible choice of Γ allows us to contract ℓ_3 while retaining e . Let $E'_3 \in \mathcal{E}$ be the resulting retained edge set. By the preceding argument, E'_3 shares at least three edges with each of E'_1 and E'_2 . Since e belongs to all three sets, we have that $\lambda_e^{E'_1} = \lambda_e^{E'_3} = \lambda_e^{E'_2}$. Therefore, because $e \in E'_1 \cap E'_2$ was chosen arbitrarily, we have $\lambda_e^{E'_1} = \lambda_e^{E'_2}$ for every $e \in E'_1 \cap E'_2$.

We next show that every pair of distinct edges lies in some $E' \in \mathcal{E}$. Consider any distinct edges $e, f \in E$. If an edge $\ell \in L$ is different from both e and f , and the edges $\{e, f, \ell\}$ do not form a triangle, then an admissible choice of Γ allows us to contract ℓ while retaining both e and f . At most three edges fail these conditions: e, f , and the edge completing a triangle with them, if such an edge exists. Since L contains four edges, some $\ell \in L$ must admit a contraction retaining both e and f . Thus every pair of distinct edges lies in some $E' \in \mathcal{E}$.

In particular, every edge itself belongs to at least one member of $\mathcal{E}$. We may therefore define $\lambda_e = \lambda_e^{E'}$ using any E' containing e , and the argument above makes this definition independent of the choice of E' . For any distinct edges $e, f \in E$, choosing a set E' containing both edges gives

$$\beta_{\{e, f\}} = \lambda_e^{E'} + \lambda_f^{E'} = \lambda_e + \lambda_f.$$

Now fix any distinct $i, j, k \in E$ and any $(e, f) \in R_{i,j,k}^E$. We have

$$\begin{aligned} & \beta_{\langle i,e \rangle} + \beta_{\langle i,f \rangle} - \beta_{\langle i,j \rangle} - \beta_{\langle i,k \rangle} + \beta_{\langle j,k \rangle} \\ &= (\lambda_i + \lambda_e) + (\lambda_i + \lambda_f) - (\lambda_i + \lambda_j) - (\lambda_i + \lambda_k) + (\lambda_j + \lambda_k) \\ &= \lambda_e + \lambda_f = \beta_{ef}. \end{aligned}$$

This is precisely (23), so the equations of $\mathbb{E}_{i,j,k}^E$ hold for every choice of distinct $i, j, k \in E$.

□

References

- [1] A. Assad and W. Xu. The quadratic minimum spanning tree problem. *Naval Research Logistics*, 39:399–417, 1992.
- [2] J.P. Costalonga, R.J. Kingan, and S.R. Kingan. Constructing minimally 3-connected graphs. *Algorithms*, 14(1):9, 2021.
- [3] A. Ćustić and A. Punnen. A characterization of linearizable instances of the quadratic minimum spanning tree problem. *Journal of Combinatorial Optimization*, 35:436–453, 2018.
- [4] A. Ćustić, R. Zhang, and A. Punnen. The quadratic minimum spanning tree problem and its variations. *Discrete Optimization*, 27:73–87, 2018.
- [5] B. Korte and J. Vygen. *Combinatorial Optimization: Theory and Algorithms*. Springer Berlin, Heidelberg, 6th edition, 2018.
- [6] OEIS Foundation Inc. The On-Line Encyclopedia of Integer Sequences, 2026. Published electronically at <http://oeis.org/A199676>.
- [7] K. Ota and A. Saito. Non-separating induced cycles in 3-connected graphs. *Scientia Series A*, 2:101–105, 1988.
- [8] D.L. Pereira, M. Gendreau, and A.S. da Cunha. Lower bounds and exact algorithms for the quadratic minimum spanning tree problem. *Computers & Operations Research*, 63:149–160, 2015.
- [9] B. Rostami and F. Malucelli. Lower bounds for the quadratic minimum spanning tree problem based on reduced cost computation. *Computers & Operations Research*, 64:178–188, 2015.

- [10] A. Saito. Covering contractible edges in 3-connected graphs. I: covers of size three are cutsets. *Journal of Graph Theory*, 14(6):635–643, 1990.
- [11] C. Van Nuffelen. On the incidence matrix of a graph. *IEEE Transactions on Circuits and Systems*, 23(9):572, 1976.
- [12] L. Waddell. A polyhedral characterization of linearizable quadratic combinatorial optimization problems. *Discrete Applied Mathematics*, 376:281–293, 2025.