Centro de Investigación y de Estudios Avanzados
del Instituto Politécnico Nacional
Unidad Zacatenco
Departamento de Matemáticas

Ideales críticos de una gráfica y reducción de dimensión en el espacio de árboles

TESIS QUE PRESENTA

Carlos Alejandro Alfaro Montúfar

PARA OBTENER EL GRADO DE
DOCTOR EN CIENCIAS
EN LA ESPECIALIDAD EN MATEMÁTICAS

Director de tesis: Dr. Carlos Enrique Valencia Oleta

México, D.F. Agosto de 2014.

Center for Research and Advanced Studies
of the National Polytechnic Institute
Campus Zacatenco
Department of Mathematics

Critical ideals of a graph and dimension reduction in tree space

A Dissertation presented by:

Carlos Alejandro Alfaro Montúfar

to obtain the degree of

Doctor in Science

in the speciality of Mathematics

Thesis Advisor: Ph.D. Carlos Enrique Valencia Oleta

México, D.F. August 2014.

Resumen

Esta tesis se divide en dos partes: La primera parte concierne al estudio de los ideales críticos de una gráfica, y la segunda al análisis estadístico de datos con estructura de árbol.

En esta tesis damos conjuntos de subgráficas *prohibidas minimales* para $\Gamma_{\leq 1}$, $\Gamma_{\leq 2}$ y $\Gamma_{\leq 3}$. Y usamos estas gráficas prohibidas para obtener una clasificación de las gráficas en $\Gamma_{\leq 1}$ y en $\Gamma_{\leq 2}$, además damos una descripción parcial de las gráficas en $\Gamma_{\leq 3}$. Como consecuencia, damos una clasificación completa para las gráficas cuyo grupo crítico tiene 2 factores invariantes iguales a 1. Cabe señalar que éste fue una problema que no tuvo respuesta en por lo menos diez años.

Por otro lado, decimos que dos vértices son *gemelos* si tienen la misma vecindad. De esta forma tenemos dos tipos de vértices gemelos: los que son adyacentes entre sí y los que no. Esdudiamos los ideales críticos de una gráfica que tienen vértices gemelos. Específicamente, obtenemos relaciones entre los ideales críticos y sus evaluaciones. Como consecuencia damos una cota superior para el co-rango algebraico para las gráficas con vértices gemelos. Después usamos esta cota para caracterizar las gráficas con a lo más 3 ideales críticos y número de cliqué igual a 2 y a 3.

Finalizamos esta parte con los cálculos de los ideales críticos de todas las gráficas simples y conexas con a los más 9 vértices. Estos cálculos fueron usados en las secciones anteriores; por ejemplo en el cálculo de las gráficas prohibidas. Además estos datos nos ayudan para conjeturar nuevos resultados. Y damos un nuevo enfoque a las gráficas co-espectrales al ver el polinomio característico como una evaluación de los ideales críticos.

La segunda parte de esta tesis concierne al *análisis estadístico de datos con estructura de árbol*, la cual es un área nueva de la estadística con bastas áreas de aplicación. Algunas ideas del *Análisis de Componentes Principales* (PCA) han sido desarrolladas previamente para árboles binarios. Extendemos estas ideas a un espacio más general de *árboles ordenados con raíz*. Conceptos como *línea-árbol* y *componente principal hacia delante* son redefinidos para este espacio más general. Además se generaliza un algoritmo que las cálcula en tiempo polinomial. Desarrollaremos una técnica análoga a la *clásica reducción de dimensión* en PCA. Para hacer esto, definiremos las *componentes principales hacia atrás*, estas componentes son las que llevan menos cantidad de información sobre los datos. Presentaremos además, un algoritmo que las encuentra. Más aún, la relación de éstas con las componentes principales hacia delante son investigadas, y la propiedad de *independencia de caminos* entre las técnicas hacia delante y hacia atrás es demostrada. Estos métodos son aplicados a un conjunto de datos de arterias cerebrales de 98 sujetos. Usando estas técnicas, investigaremos los efectos de envejecimiento de las estructuras de las arterias cerebrales de mujeres y hombres. Un segundo conjunto de datos de la estructura organizacional de una compañía grande en los Estados Unidos es también analizada, y las diferencias estructurales a través de diferentes tipos de departamantos dentro de la compañía son explorados.

La investigación hecha en la primera parte de la tesis se ve reflejada en los artículos: [**1, 2, 3, 4, 5, 6, 7**] de la bibliografía de la primera parte. Y la de la segunda parte en el artículo [**1**] y la patente [**2**] de la segunda bibliografía.

Abstract

This thesis is divided in two parts. The first part is devoted to the study of critical ideals of a graph, and the second part to the statistical analysis of tree structured data.

In this thesis we provide the sets of minimal forbidden subgraphs for $\Gamma_{\leq 1}$, $\Gamma_{\leq 2}$ and $\Gamma_{\leq 3}$. And we use these forbidden subgraphs to get a complete classification of the graphs in $\Gamma_{\leq 1}$ and $\Gamma_{\leq 2}$, and a partial classification of the graphs in $\Gamma_{\leq 3}$. As a consequence we give a complete classification of the simple graphs whose critical group has two invariant factors equal to one, which was an unanswered question for at least the last ten years.

We say that two vertices of a graph are twins if they have the same neighbors. There are two types of twins depending on whether the twins are connected or not. We also study the critical ideals of a graph having twin vertices. Specifically, we obtain relations between some evaluations of the critical ideals of a graph G and the critical ideals of G with some vertices arbitrarily cloned. As a consequence, we get an upper bound for the algebraic co-rank for a graph with twin vertices. After we use this bound to characterize the graphs with at most three trivial critical ideals and clique number equal to 2 and 3.

We finish this part with the computation of the critical ideals of the simple connected graphs with at most 9 vertices. This data was used in previous sections; for instance in the computation of the forbidden graphs. This data also has help us to conjecture new results. Finally, we give a new interpretation to the co-spectral graphs by considering the characteristic polynomial as an evaluation of the critical ideals.

The second part of this thesis concerns to the statistical analysis of tree structured data, which is a new topic in statistics with wide application areas. Some Principal Component Analysis (PCA) ideas have been previously developed for binary tree spaces. Here, these ideas are extended to the more general space of rooted and ordered trees. Concepts such as tree-line and *forward* principal component tree-line are redefined for this more general space, and the optimal algorithm that finds them is generalized. An analog of classical dimension reduction technique in PCA for tree spaces is developed. To do this, *backward* principal components, the components that carry the least amount of information on tree data set, are defined. An optimal algorithm to find them is presented. Furthermore, the relationship of these to the forward principal components is investigated, and a path-independence property between the forward and backward techniques is proven. These methods are applied to a brain artery data set of 98 subjects. Using these techniques, the effects of aging to the brain artery structure of males and females is investigated. A second data set of the organization structure of a large US company is also analyzed and the structural differences across different types of departments within the company are explored.

The research done in the first part is reflected in the articles and preprints: [**1, 2, 3, 4, 5, 6, 7**] of the bibliography of the first part. And the second in the article [**1**] and the patent application [**2**] of the second bibliography.

Agradecimientos

Primero que nada, agradezco a mi asesor Carlos Valencia por sus consejos y su incondicional apoyo. Esta tesis no hubiera sido posible sin la libertad e inspiración que él me concedío. También quiero agradecer a Burcu Aydın por sus consejos y apoyo para la realización de la segunda parte de la tesis.

Agradezco a mis coautores su generosidad para dejarme usar aquí sus ideas y documentos, incluidos en las referencias citadas.

Agradezco a mi novia, Chelito, por brindarme apoyo emocional e intelectual sin ella ésta tesis no hubiera sido posible. Agradezco a mis amigos del doctorado Carmen Lozano, César Guadarrama, Hugo Corrales y Marcos Vargas por los cafés, las risas, charlas y las experiencias que aún compartimos.

Agradezco a los profesores del Departamento de Matemáticas por los comentarios y sugerencias que me brindaron. Agradezco a Pano Santos, Lyle Ramshaw, Bob Tarjan y Julie Ward por las facilidades y conocimientos que me brindaron durante mi estancia en los HP Labs.

Agradezco a las secretarias Adriana Aranda y Roxana Martínez por la toda ayuda que me brindaron.

Agradezco al Consejo Nacional de Ciencia y Tecnología (CONACyT) y al proyecto SEP-CONACyT 166059 por el apoyo económico que me brindaron durante la realización de este proyecto.

A mis padres,
con gratitud,
admiración
y amor.

Le savant n'étudie pas la nature parce que cela est utile; il l'étudie parce qu'il y prend plaisir et il y prend plaisir parce qu'elle est belle. Si la nature n'était pas belle, elle ne vaudrait pas la peine d'être connue, la vie ne vaudrait pas la peine d'être vécue. Je ne parle pas ici, bien entendu, de cette beauté qui frappe les sens, de la beauté des qualités et des apparences; non que j'en fasse fi, loin de là, mais elle n'a rien à faire avec la science; je veux parler de cette beauté plus intime qui vient de l'ordre harmonieux des parties, et qu'une intelligence pure peut saisir

Henri Poincaré

Los científicos estudian la naturaleza no porque sea útil, sino porque encuentran placer en ello, y encuentran placer porque es hermosa. Si no lo fuera, no merecería la pena conocerla, y si la naturaleza no mereciera la pena, la vida tampoco. No me refiero, claro está, a la belleza que estimula los sentidos, la de las cualidades y las apariencias; no es que menosprecie tal belleza, nada más lejos de mi intención, más ésta nada tiene que ver con la ciencia; me refiero a esa hermosura más profunda que emana del orden armonioso de las partes, susceptible de ser captada por una inteligencia pura.

Henri Poincaré

Contents

Part 1

Critical Ideals of a graph

CHAPTER 1

Introduction

In this thesis, graphs allow multiple edges and no loops. The critical group $K(G)$ of a graph G, also known as sandpile group, is a topic that has been widely developed in the last years. The interest in studying the critical group is that it is in the intersection of several areas of the mathematics, physics and computer science. For instance, the critical group had been studied in [**9, 14, 17, 18, 19, 20, 21, 23, 24, 29, 31, 33, 34, 35, 37, 38, 44, 45**].

It is known [**33**, Theorem 1.4] that the critical group of a connected graph G with n vertices can be described as follows:

$$K(G) \cong \mathbb{Z}_{d_1} \oplus \mathbb{Z}_{d_2} \oplus \cdots \oplus \mathbb{Z}_{d_{n-1}},$$

where $d_1, d_2, ..., d_{n-1}$ are positive integers with $d_i \mid d_j$ for all $i \leq j$. These integers are called *invariant factors* of the Laplacian matrix of G. Besides, if $\Delta_i(G)$ is the *greatest common divisor* of the i-minors of the Laplacian matrix $L(G)$ of G, then the *i-th* invariant factor d_i is equal to $\Delta_i(G)/\Delta_{i-1}(G)$, where $\Delta_0(G) = 1$ (for details see [**30**, Theorem 3.9]).

The computation of the invariant factors of the Laplacian matrix is an important technique used in the understanding of $K(G)$. For instance, several researchers addressed the question of how often the critical group is cyclic, that is, if $f_1(G)$ denote the number of invariant factors equal to 1, then the question is how often $f_1(G)$ is equal to $n-2$ or $n-1$. In this way, it is desirable to understand the combinatorial properties of $f_1(G)$ and the family of graphs $\mathcal{G}_i$ of simple connected graphs with $f_1(G) = i$.

Superficially, the critical group has three components: algebraic, combinatorial, and arithmetic. The methodology of these studies rely on the separation of the combinatorial and algebraic information from most of the arithmetic component by means of the introduction of a new invariant: the critical ideals. Critical ideals were defined in [**22**] as a generalization of the critical group and have been studied in [**1, 7, 22**]. The effect of avoiding the arithmetic information is the behavior of the critical ideals is easier to observe and to describe. Thus critical ideals provide a new perspective to understand the critical group theory. Until now, the major research in the area is based in the computation the critical group of small families of graphs, but few deep relations between the critical group and the combinatorial properties of the graph have been found. The effect of avoiding most of the arithmetic information is that the patterns of the behavior of the critical ideals and the graphs are easier to observe and to describe.

In Chapter 2 we explore the definition of *algebraic co-rank* of a graph, which is the number of trivial critical ideals associated to the given graph. A crucial result linking critical groups and critical ideals is [**22**, Theorem 3.6], which states that if D_G is the degree vector of G, and $d_1 \mid \cdots \mid d_{n-1}$ are the invariant factors of $K(G)$, then

$$I_i(G, D_G) = \left\langle \prod_{j=1}^{i} d_j \right\rangle = \langle \Delta_i(G) \rangle \text{ for all } 1 \leq i \leq n-1.$$

Thus if the critical ideal $I_i(G, X_G)$ is trivial, then $\Delta_i(G)$ and d_i are equal to 1. Equivalently, if $\Delta_i(G)$ and d_i are not equal to 1, then the critical ideal $I_i(G, X_G)$ is not trivial. The main result of Chapter 2 is the classification of the graphs with two trivial critical ideals (see Theorems 2.10 and Theorem 2.11). This result is then employed to fully characterize the graphs with two invariant factors equal to 1 (see Theorem 2.22). Which was an unsolved problem for at least 10 years. In turn, we develop new concepts like the *algebraic co-rank* of a graph, which is the number of trivial critical ideals associated to the given graph. We finish this chapter with the classification of the graphs with three trivial critical ideals with clique number at most 3 (see Theorem 2.33).

In Chapter 3, we recall the concepts of duplication and replications of vertices. The purpose of this chapter is to study of the critical ideals of signed multidigraphs having twin vertices. Several graph families have twin vertices. For instance, the complete multipartite graphs, the threshold graphs, the quasi-threshold graphs, or the cographs. Therefore, the description of critical ideals of graphs with twins is an important step in the development on the theory of critical ideals and critical group. Here, we will obtain relations between some evaluations of the critical ideals of a signed multidigraph G and the critical ideals of $G^{\mathbf{d}}$, where $\mathbf{d} \in \mathcal{P}^{V(G)}$. As a consequence of this partial description of the critical ideals, we get an upper bound for the algebraic co-rank of graphs with twins. This upper bound is important for instance in the classification of the graphs that have algebraic co-rank less than or equal to an integer k (see [**2**, Section 2]). We will also state three conjectures which lead into a wide and interesting panorama of the critical ideals. Also, we give a description of the critical ideals of the *k-th* duplication $d^k(G, v)$ of vertex v and *k-th* replication $r^k(G, v)$ of vertex v in terms of some of the critical ideals of G.

Chapter 4 aims at discussing some numerical experiments providing information about the critical ideals of small graphs. Much of this information have being used to conjecture, and some of it had end up in new mathematical results. These computations also lead into a wide panorama of the critical ideals.

The first part of this thesis has been written jointly with Carlos Valencia and Hugo Corrales. Each chapter corresponds to a paper or a preprint.

1. Background

1.1. Laplacian matrix. In this section we introduce the *Laplacian matrix*, or also called *Kirchhoff's matrix*, of a graph and few basic properties that will be frequently used through this thesis. Laplacian matrices had been extensively studied in different context in the last 50 years. A general account of the Laplacian matrices can be found in [**10**, Chapter 4], [**11**, Chapter 4], [**13**] or [**26**, Chapter 13].

The *Laplacian matrix* $L(G)$ of a graph G is the matrix with rows and columns indexed by the vertices of G, such that the uv-entry is the negative of the number of edges (also called *multiplicity*) between the different vertices u and v, and the degree of u otherwise.

The Laplacian matrix is closely related to the adjacency and incidence matrices. Many properties of the Laplacian matrix are inherited from these matrices. The *adjacency matrix* $A(G)$ of a graph $G = (V, E)$ is the matrix with rows and columns indexed by V whose uv-entry is the multiplicity between the vertices u and v. The *degree matrix* $Deg(G)$ of a graph G is the diagonal matrix with rows and columns indexed by V whose uu-entry is the degree of vertex u. Therefore, $L(G) = Deg(G) - A(G)$. An *orientation* σ of a graph G is the assignment of a direction of each edge, *i.e.*, we declare one end of the edge to be the *head*, and the other end to be the *tail*. Formally, if the edge uv is oriented from the tail u to the head v, then $\sigma(u, v) = 1$ and $\sigma(v, u) = -1$. The graph G together with an orientation σ is called an *oriented graph* and is denoted by G^σ. The *incidence matrix* $D(G^\sigma)$ of an oriented graph G^σ is the $\{0,1,-1\}$-matrix whose rows and columns are indexed by the vertices and the edges of G, respectively, such that the ue-entry is equal to 1 if the vertex u is the head of the edge e, -1 if u is the tail of e, and 0 otherwise. The incidence matrix of an orientation is related to the Laplacian in the following way.

LEMMA 1.1. [**26**] *If σ is an orientation of the graph G, and D is the incidence matrix of G^σ, then $L(G) = DD^T$.*

Despite that there are many different ways to give an orientation to a graph, many of the results about oriented graphs are independent of the choice of orientation. An example of this is the following lemma.

LEMMA 1.2. [**26**] *Let G be a graph with n vertices and c components. If σ is an orientation of G and D is the incidence matrix of G^σ, then the rank of D is equal to $n - c$.*

Since the rank of $D = D(G^\sigma)$, for an arbitrary orientation σ, is equal to the rank of $L(G) = DD^T$, then the rank of $L(G)$ is $n - c$. A direct proof of this lemma can be found in [**44**, Lemma 3.1].

LEMMA 1.3. [**26**] *If G is a graph with n vertices and c components, then the rank of $L(G)$ is equal to $n - c$.*

We finish this section with the Kirchhoff matrix-tree theorem. This relates the minors of size $n-1$ and spanning trees, however there are more general results. For instance, a result for any principal minor can be found in [**10**, Theorem 4.7], and a result for any minor can be found [**16**].

DEFINITION 1.4. *A spanning tree is a graph that is a tree and a spanning subgraph at the same time.*

The number of spanning trees of G is denoted by $\tau(G)$. Let $L[u,v]$ be the submatrix of the matrix L obtained by removing from L the row and column corresponding to the vertices u and v, respectively. And let $L[u] = L[u,u]$. The following celebrated result due to Kirchhoff [**32**] in an 1847 paper concerned with electrical networks.

THEOREM 1.5 (Matrix-Tree Theorem). *Let G be a graph with Laplacian matrix L. If u is a vertex of G, then* $\det L[u] = \tau(G)$.

1.2. Critical group. Let assume that G is a connected graph with n vertices: $v_1, ..., v_n$. By considering the Laplacian matrix $L(G)$ as a linear map $L(G) : \mathbb{Z}^V \to \mathbb{Z}^V$, the *cokernel* of $L(G)$ is the quotient module $\mathbb{Z}^V / \operatorname{Im} L(G)$.

DEFINITION 1.6. *The critical group $K(G)$ of G is the torsion part of the cokernel of $L(G)$.*

The critical group has been studied intensively on several contexts over the last 30 years. For example, the *group of components* [**34, 35**], the *Picard group* [**9, 14**], the *Jacobian group* [**9, 14**], the *sandpile group* [**21**], *chip-firing game* [**14, 37**], or *Laplacian unimodular equivalence* [**29, 38**].

In the rest of this section we present some basic properties of the critical group. The following theorem implies that the critical group is a finer invariant than the number of spanning trees of a connected graph G.

THEOREM 1.7. [**14**] *If G is a connected graph, then $K(G)$ has order the number $\tau(G)$ of spanning trees of G.*

For example, Cayley's celebrated formula $\tau(K_n) = n^{n-2}$ is suggestive of the structure of the critical group for the complete graph K_n, since $K(K_n) \simeq \mathbb{Z}_n^{n-2}$. On the other hand, since the number of spanning trees of a planar graph G is equal to that of its dual G^*, it follows that the sandpile groups of G and G^* have equal order.

Theorem 1.8. [**9, 21**] *For a planar graph G and any of its duals G^*, the groups $K(G)$ and $K(G^*)$ are isomorphic.*

Also, in [**44**] it is proved that if for two connected graphs G and H, the graphic matroids $M(G)$ and $M(H)$ are isomorphic, then $K(G) \simeq K(H)$.

Recall that a *block* is a maximal connected subgraph without a cutvertex. Let $G_1, ..., G_k$ be the blocks of a graph G. The following result was noticed by several authors [**21, 34, 44**].

Proposition 1.9. *The group $K(G)$ is isomorphic to $K(G_1) \oplus \cdots \oplus K(G_k)$.*

A classical result (see [**30**, Section 3.7]) asserts that the reduced Laplacian matrix is equivalent to a integer diagonal matrix with entries $d_1, d_2, ..., d_{n-1}$, where $d_i > 0$ and $d_i \mid d_j$ if $i \leq j$. The integers $d_1, \ldots, d_{n-1}$ are unique and are called *invariant factors*. With this in mind, the critical group of a connected graph with n vertices can be expressed in terms of the invariant factors as follows [**33**, Theorem 1.4]:

$$K(G) \cong \mathbb{Z}_{d_1} \oplus \mathbb{Z}_{d_2} \oplus \cdots \oplus \mathbb{Z}_{d_{n-1}}.$$

Definition 1.10. *For every integer k, let $f_k(G)$ be the number of invariant factors of the Laplacian matrix of G equal to k.*

Definition 1.11. *Let $\mathcal{G}_i = \{G : G$ is a simple connected graph with $f_1(G) = i\}$.*

The study of the graphs in $\mathcal{G}_i$ is of great interest. In particular, some results and conjectures for the graphs concerning cyclic critical group can be found in [**35**, Section 4] and [**44**, Conjectures 4.3 and 4.4]. An interesting result proved in [**19**] is that for any given connected simple graph, there is a homeomorphic graph with cyclic critical group. In [**29**] was noticed that if G is a graph of diameter d, then $f_1(G) \geq d$. Also, if $H = G - e$, then $|f_1(G) - f_1(H)| \leq 1$. Besides, it is easy to see [**21, 34, 38**] that $\mathcal{G}_1$ consists only of the complete graphs. In this sense, several researchers [**37**] expressed interest on the characterization of $\mathcal{G}_2$ and $\mathcal{G}_3$. The advances in this matter are the following. For instance, in [**40**] it was characterized the graphs in $\mathcal{G}_2$ whose third invariant factor is equal to n, $n-1$, $n-2$, or $n-3$. In [**17**], the characterizations of the graphs in $\mathcal{G}_2$ with a cutvertex, and the graphs in $\mathcal{G}_2$ whose number of independent cycles equals $n-2$ are given.

1.3. The Sandpile group. A vertex $s \in V$ will be distinguished and called *sink*. The non-sink vertex set is denoted by $\tilde{V}$. And $\mathbb{N}$ denote the non-negative integers.

1.3.1. *Stable configurations.* Let $G = (V, E)$ be a graph with vertices $V = \{v_1, ..., v_n = s\}$. A *configuration* of (G, s) is an element $c \in \mathbb{N}^n$. A non-sink vertex v is called *stable* with respect to c if $deg_G(v) > c_v$, and *unstable* otherwise. A configuration is called *stable* if every non-sink vertex $v \in \tilde{V}$ is stable. For instance, in Figure 1.a the vertices v_1 and v_2 are stable and the rest are unstable.

Let c be a configuration such that the vertex v_i is unstable. The *toppling* operation (or *firing*) is performed on a vertex v_i by decreasing c_i by the degree of v_i, and adding to the entry c_j (associated to vertex v_j adjacent with v_i) the multiplicity m_{v_i,v_j}. Note that toppling v_i is done by subtracting the column vector v_i of the Laplacian matrix of G to the configuration c. Let $L(G)_i$ denote the i-th column of the Laplacian matrix.

LEMMA 1.12. *Let u and v be different vertices of a graph G. If u and v are unstable in a configuration c, then the same configuration results as firing u then v as firing v then u.*

PROOF. First note that the vertices u and v can be fired in either order, since firing u cannot make v stable and vice-versa. The rest follows from the commutativity of vector subtraction

$$h - L(G)_i - L(G)_j = h - L(G)_j - L(G)_i.$$

□

A sequence of vertex firing is *valid* if at each step only unstable vertices are fired. A general consequence of the previous lemma asserts that starting from a given configuration, any two valid sequences of vertex firings that finish in the same configuration are a rearrangement of each other.

PROPOSITION 1.13. [**41**] *From configuration h, let $v_{i_1}, ..., v_{i_m} \in \widetilde{V}$ and $v_{j_1}, ..., v_{j_n} \in \widetilde{V}$ be valid sequences of fired vertices, resulting in the configurations h' and h'', respectively.*

(1) If h'' is stable, then $m \leq n$, and no vertex appears more times in $v_{i_1}, ..., v_{i_m}$ than in $v_{j_1}, ..., v_{j_n}$.
(2) If h' and h'' are both stable, then $m = n$, $h' = h''$, and the sequences are rearrangements of each other.

The following example shows how a stable configuration is reached from an unstable configuration by toppling the unstable vertices.

EXAMPLE 1.14. *The Laplacian matrix of the cycle with 5 vertices is*

$$L(\mathcal{C}_5) = \begin{bmatrix} 2 & -1 & 0 & 0 & -1 \\ -1 & 2 & -1 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & -1 & 2 & -1 \\ -1 & 0 & 0 & -1 & 2 \end{bmatrix}$$

Consider the configuration $c = (1, 0, 2, 2, 2)$. By subtracting the third row of $L(C_5)$, the configuration $(1, 1, 0, 3, 2)$ is obtained. And by subtracting it the fourth row of $L(C_5)$, the stable configuration $(1, 1, 1, 1, 3)$ is obtained. See Figure 1.

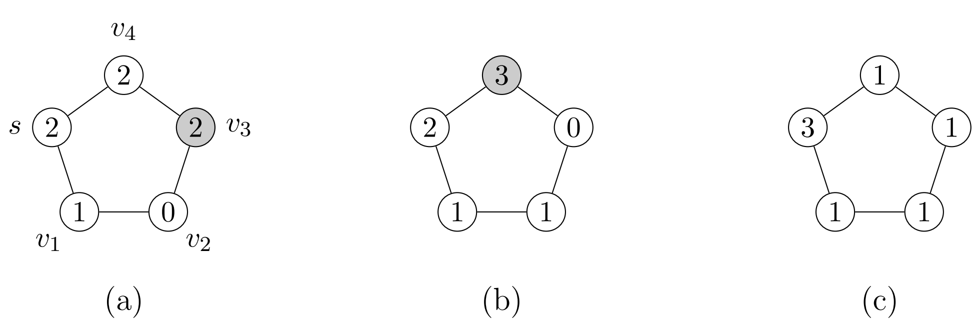


FIGURE 1. (a) An unstable configuration in C_5. (b) A new configuration obtained by toppling vertex v_3. (c) A stable configuration obtained by toppling vertex v_4.

For any configuration, a unique stable configuration is obtained by a finite sequence of topplings. The stable configuration obtained from c is denoted by $s(c)$. Thus, $s(c) = c - L(G)^t b$ for a vector $b \in \mathbb{N}^n$.

PROPOSITION 1.15. *For any non-zero configuration c, there exists a unique stable configuration $s(c)$ obtained by toppling the unstable vertices.*

PROOF. Existence. Let $V_k = \{v \in V(G) \mid d_G(v_n, v) = k\}$, where $d_G(u,v)$ denotes the distance between u and v. Let d be the greatest distance from the sink v_n. To any configuration c we associate the $(d+1)$-tuple $\mu(c) = (\mu_0(c), \mu_1(c), ..., \mu_d(c))$ given by

$$\mu_i(c) = \sum_{v \in V_i} c_v.$$

We consider the following lexicographic order $\prec$ on these d-tuples:

$$\begin{aligned} \mu(c) \prec \mu(c') \quad \Leftrightarrow \quad & \text{exists } 0 \leq k \leq d, \text{ such that} \\ & \mu_0(c) = \mu_0(c'), ..., \mu_{k-1}(c) = \mu_{k-1}(c'), \mu_k(c) < \mu_k(c'). \end{aligned}$$

Since the sum of the entries of the i-*th* row of the Laplacian matrix is equal to zero, then if $u, v \in \mathbb{N}^n$ with $u = v + L(G)^t b$ we have that $|u| = \sum_{i=0}^{d} \mu(u)_i$ is equal to $|v| = \sum_{i=0}^{d} \mu(v)_i$. Hence, there exists a finite number of configurations with the same entries sum. Thus, there exists a finite ascending chain for $\prec$ with the same entries sum.

Uniqueness. It is consequence of the commutativity of the toppling operator; $(u - L(G)_i) - L(G)_j = (u - L(G)_j) - L(G)_i$. □

1.3.2. *Recurrent configurations.* In this section, we define a special set of configurations which have the property of being an abelian group.

The sum of two configurations c, c' in (G, s) is taken entry-by-entry. That is, $c + c' := (c_1 + c'_1, ..., c_n + c'_n)$.

PROPOSITION 1.16. *If c and d are configurations in (G, s), then $s(c + d) = s(s(c) + s(d))$.*

PROOF. Since $s(c) = c'$ and $s(d) = d'$, then there exist $a, b \in \mathbb{N}^n$ such that $c' = c - aL(G)$ and $d' = d - bL(G)$ are stable configurations. Thus, $c' + d' = c + d - (a + b)L(G)$. Since $c' + d'$ is not necessary an stable configuration, then we have that $s(u + v) = s(c' + d')$. □

Since we cannot topple the sink v_n, then we say that two configurations c and c' are *equivalent* if $c_v = c'_v$ for all non-sink vertices $v \in \tilde{V}$. Thus most of the time we omit the value of the sink in the configuration. The *support* of a configuration c is the set $supp(c) = \{v \in \widetilde{V} \mid c_v \neq 0\}$.

REMARK 1.17. *Let $\sigma_{max} = (d_G(v_1)) - 1, ..., d_G(v_n) - 1)$, and $\delta = \sigma_{max} + \mathbf{1}$. Then*

i. *σ_{max} is stable.*
ii. *δ is unstable.*
iii. *$(\delta - s(\delta))_i > 0$ for all $1 \leq i \leq n - 1$.*
iv. *$\delta - s(\delta) = L(G)b$ for a vector $b \in \mathbb{N}^n$.*

DEFINITION 1.18. *The configuration c is recurrent if there exists a non-zero configuration d such that $s(c + d) = c$.*

It is easy to see that the configuration $(1,1,1,1)$ in C_5 with v_5 as sink is a recurrent configuration. One may think that if we take any unstable configuration the stabilization will produce a recurrent configuration. However, this is not true in general, for instance consider the following example.

EXAMPLE 1.19. *Consider the configuration* $(0,1,0)$ *in* C_4 *with* v_4 *as sink. It is the stabilization of the non-stable configuration* $(2,0,0)$. *However it is not recurrent.*

There exist several other definitions of recurrent configuration. In the following we show that in fact they are equivalent.

DEFINITION 1.20. *A configuration* $\beta \geq \mathbf{1}$ *of* (G,s) *is called a burning configuration if*

- $\beta = \mathbf{z}^t L(G,s)$ *for some* $\mathbf{z} \in \mathbb{Z}^{V(G)\setminus s}$,
- *for all* $v \in V(G) \setminus s$, *there exists a path to* v *from some vertex of* $\mathrm{supp}(\beta)$.

THEOREM 1.21. [**43**] *Let* β *be a burning configuration of* (G,s), *then a configuration* $\mathbf{c} \in \mathbb{N}^{V(G)\setminus s}$ *of* (G,s) *is recurrent if and only if*

$$s(\mathbf{c}+\beta) = \mathbf{c} \text{ with firing vector equal to } \beta^t L(G,s)^{-1}.$$

THEOREM 1.22. [**43**] *There exists a unique burning configuration* $\beta_{\min}$ *such that*

$$\beta_{\min}^t L(G,s)^{-1} \leq \beta'^t L(G,s)^{-1} \text{ for all } \beta' \text{ a burning configuration.}$$

Moreover, $\beta_{\min}^t L(G,s)^{-1} \geq \mathbf{1}$ *with equality if and only if* G *has no vertex* $v \in V(G) \setminus s$ *with* $\deg_G^+(v) < \deg_G^-(v)$.

THEOREM 1.23. *The following statements are equivalent:*

i. [**18**] *For all* $v \in \mathbb{N}^{\widetilde{V}}$ *there exists* $\omega \in \mathbb{N}^{\widetilde{V}}$ *such that* $s(v+\omega) = c$.
ii. [**43**] *There exists a configuration* e *such that* $s(\sigma_{max}+e) = c$.
iii. [**21**] *There exists* $\omega \in \mathbb{N}^{\widetilde{V}}$ *such that* $s(c+\omega) = c$.
iv. c *is stable and* $c + L(G)x$ *is not stable, for all* $x \neq 0 \in \mathbb{N}^V$.
v. $s(c+\beta) = c$.

PROOF.

$i \Rightarrow ii)$ Taking $v = \sigma_{max}$ we have that there exists a configuration $\omega = e$ such that $s(\sigma_{max}+e) = c$.

$ii \Rightarrow i)$ We have two cases: either $v \leq \sigma_{max}$ or $v > \sigma_{max}$. If $v \leq \sigma_{max}$, we finish. Suppose that $v > \sigma_{max}$. Then $s(v) \leq \sigma_{max}$, and $\sigma_{max} = s(v) + r$. Finally, taking $\omega = r + e$, $s(v+\omega) = s(s(v)+(r+e)) = s((s(v)+r)+e) = s(\sigma_{max}+e) = c$.

$ii \Rightarrow iii)$ Clearly, c is stable, then $c \leq \sigma_{max}$. Thus $\sigma_{max} = c+r$ and $s(c+(r+e)) = s(\sigma_{max}+e) = c$.

$iii \Rightarrow iv)$ Suppose there exists $x \neq 0 \in \mathbb{N}^V$ such that $c + L(G)x$ is stable. Then $s(c+\omega) = c$ and

$$s(c+L(G)x+\omega) = s(s(c+\omega)+L(G)x) = s(c+L(G)x) = c+L(G)x.$$

Thus $c + xL(G,s)$ is recurrent. That is, there exist two different recurrent configurations in the same class of equivalence. Contradicting Proposition 1.15.

$iv \Rightarrow ii)$ We have that c is stable and $c + L(G,s)x$ is not stable, for all $x \neq 0 \in \mathbb{N}^V$. In particular, take x such that $c + L(G)_s \geq \sigma_{max}$. Since $s(c + L(G)_s x)$ exists, there is a sequence of topplings $i_1, ..., i_k$ that carries $c + L(G)_s x$ to $s(c + L(G)_s x)$. Let $y = e_{i_1} + \cdots + e_{i_k}$, thus $c + L(G)(x-y) = s(c+L(G)x)$. There is no j such that $(x-y)_j > 0$, because it means that there is a vertex unstable. So $(x-y)_j < 0$ for each $1 \leq j \leq k$. If $x - y \neq \mathbf{0}$ then $s(c+L(G)x) < c$ which contradicts Proposition 1.13. Therefore, $(x-y) = \mathbf{0}$ and

$s(c+L(G)x)=c$. Now, since $c+L(G)_s x \geq \sigma_{max}$, there exists r such that $\sigma_{max}+r = c+L(G)x$. And thus $s(\sigma_{max}+r)=s(c+L(G)x)=c$.

$v \Rightarrow iii)$ It follows by taking $\omega=\beta$.

$iii \Rightarrow v)$ Theorem 1.21.

□

The *sandpile monoid* $\mathcal{M}$ is defined as the set of stable configurations under the operation of point-wise addition and stabilization. We denote this operation by $\oplus$. Let $I_{\sigma_{max}}$ be the ideal generated by σ_{max}, and let I_{min} be the unique minimal ideal. The ideal $I_{\sigma_{max}}$ is precisely the set of configurations that satisfies (ii) in the theorem above.

PROPOSITION 1.24. *$I_{\sigma_{max}} = I_{min}$.*

PROOF. It is clear that $I_{min} \subseteq I_{\sigma_{max}}$. On the other hand, $\sigma_{max} \in I_{min}$ because if $h \in I_{min} \subset \mathcal{M}$ there exists a stable configuration h' such that $h+h'=\sigma_{max}$. Which implies that $\sigma_{max}=h' \oplus h \in I_{min}$. Then, $I_{\sigma_{max}} \subseteq I_{min}$. □

DEFINITION 1.25. *A recurrent configuration c is minimal if there is no recurrent configuration $c' \neq c$ such that $c'_v \leq c_v$ for all $v \in \widetilde{V}$.*

PROPOSITION 1.26. *If c is a recurrent configuration, then there exists a minimal recurrent configuration c_{min} such that $c_{min} \leq c \leq \sigma_{max}$. Moreover, every configuration c between c_{min} and σ_{max} is recurrent.*

PROPOSITION 1.27. [**21**] *The minimal recurrent configurations of the complete graph K_n with n vertices are permutations of $(n-2, ..., 1, 0)$.*

Given a configuration c, we define its *level* as

$$level(c) = \sum_{v \in \widetilde{V}} c_v$$

THEOREM 1.28. [**37**] *Let G be a graph with sink vertex s and c a recurrent configuration, then*

$$|E(G)| - deg_G(s) \leq level(c) \leq 2|E(G)| - deg_G(s) - |V(G)| + 1.$$

For $i \geq 0$, we take a_i as the number of recurrent configurations with level $i + |E(G)| - deg_G(s)$. We now take the generating function of the recurrent configurations, that is, the polynomial

$$P_s(G;y) = \sum_{i=0}^{|E(G)|-|V(G)|+1} a_i y^i$$

THEOREM 1.29. [**37**] *For a graph G and sink vertex s, we have that the generating function of the recurrent configurations is the Tutte polynomial of G along the line $x=1$, that is,*

$$P_s(G;y) = T(G;1,y)$$

EXAMPLE 1.30. *Consider the complete graph K_4 with 4 vertices and sink vertex s. There are 16 recurrent configurations (see Figure 2) and the Tutte polynomial of G along the line $x=1$ is*

$$P(K_4;y) = 6 + 6y + 3y^2 + y^3.$$

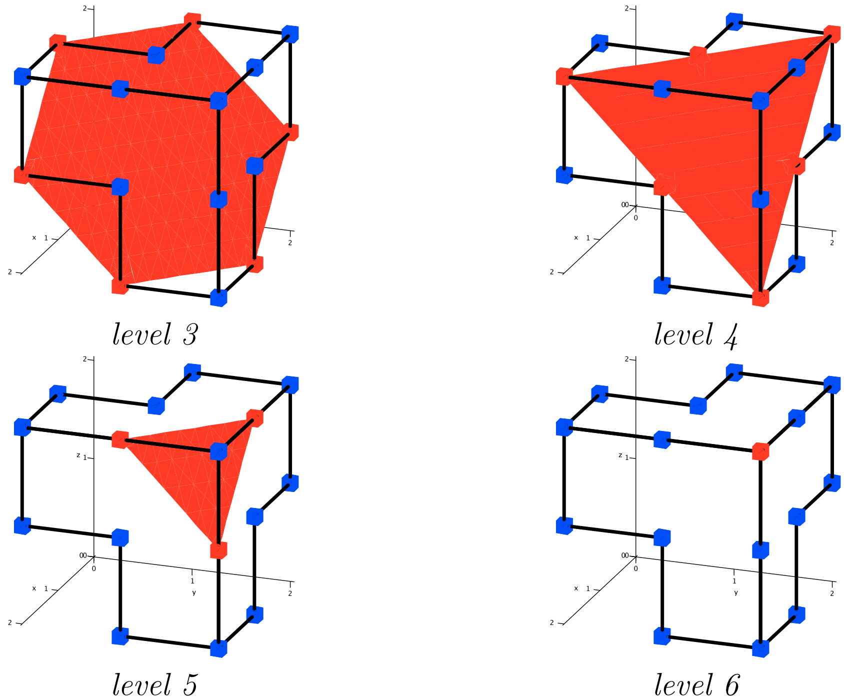


FIGURE 2. Recurrent configurations of K_4 represented in $\mathbb{R}^3$. There are 6 recurrent configurations with level 3, 6 with level 4, 3 with level 5, and 1 with level 5.

1.4. Sandpile group.

DEFINITION 1.31. *The sandpile group of G is the set of recurrent configurations and is denoted by $SP(G, s)$.*

Now we describe the sandpile group of the cycle.

PROPOSITION 1.32. *The sandpile group of the cycle is composed of the following elements:* $\mathbf{1}$, $\mathbf{1} - \mathbf{e}_{v_i}$ *for* $v_i \in \widetilde{V}$.

Let us define $u \oplus v := s(u + v)$. So we will see that the sandpile group with the operation $\oplus$ is effectively a group.

LEMMA 1.33. *Let δ be as in Remark 1.17, and $\epsilon = 2\delta - 2s(\delta)$. If u is recurrent, then $s(u+\epsilon) = u$.*

PROOF. Let u be a recurrent configuration. Hence, there exists a configuration v such that $s(u + v) = u$. Then, applying several times Remark 1.16, we have

$$\begin{aligned} s(u + v + \epsilon) &= s(u + v + 2\delta - 2s(\delta)) = s(u + v + 2s(\delta) - 2s(\delta)) \\ &= s(u + v) = u. \end{aligned}$$

On the other hand,

$$s(u + v + \epsilon) = s(u + \epsilon).$$

$\square$

PROPOSITION 1.34. *For every configuration u, there exists a unique recurrent configuration v such that $u - v \in \langle L(G)_1, ..., L(G)_{n-1} \rangle$.*

Proof. Existence. Let u be a configuration. Since $(\delta - s(\delta))_i > 0$, we can find $k > 0$ such that $u + k(\delta - s(\delta)) > \sigma_{max}$. Now, let $v = s(u + k(\delta - s(\delta)))$. Hence, v is stable. And, since $u + k(\delta - s(\delta)) > \sigma_{max}$, then there exists a configuration c such that $u + k(\delta - s(\delta)) > \sigma_{max} = v + c$. Therefore, v is recurrent.

Uniqueness. Let u and v be two recurrent configurations such that $u - v \in \langle L(G)_1, ..., L(G)_{n-1} \rangle$. Then, $u - v = \sum_{i=1}^{n-1} c_1 L(G)_i$. Take $I = \{i \mid c_i < 0\}$ and $J = \{i \mid c_i \geq 0\}$, and define

$$\beta = u + \sum_{i \in I} (-c_i) L(G)_i = v + \sum_{i \in J} c_i L(G)_i.$$

Let $k = \max_{v \in \tilde{V}} \{\mid c_v \mid d_v\}$, and $\tau = \beta + k\epsilon$. Hence

$$\tau = \beta + k\epsilon = u + \sum_{i \in I} (-c_i) L(G)_i + k\epsilon$$

Since $\tau_v \geq -c_v d_v$, we can topple $-c_v$ times each vertex of τ. Then,

$$s(\tau) = s(u + k\epsilon) = s(u + \epsilon) = u$$

On the other hand,

$$\tau = \beta + k\epsilon = v + \sum_{i \in J} (c_i) L(G)_i + k\epsilon$$

Since $\tau_v \geq c_v d_v$, we can topple c_v times each vertex of τ. Then,

$$s(\tau) = s(v + k\epsilon) = s(v + \epsilon) = v$$

Thus, by Proposition 1.15, we have that $u = v$. $\square$

Thus, each element in $\mathbb{Z}^{n-1} / \langle L(G)_1, ..., L(G)_{n-1} \rangle$ is represented by a unique recurrent configuration. Since $\mathbb{Z}^{n-1} / \langle L(G)_1, ..., L(G)_{n-1} \rangle$ is an abelian group, then it remains to prove that $SP(G, s)$ is closed under the operation $\oplus$.

Theorem 1.35. *Let G be a multigraph with distinguished vertex s. Then the sandpile group $SP(G, s)$ is an abelian group.*

Proof. It only remains to prove that $SP(G, s)$ is closed under $\oplus$. Let $u, v \in SP(G, s)$, then u, v are recurrent. Thus, there exist u' and v' such that $s(u + u') = u$ and $s(v + v') = v$. Now, $s(u + u' + v + v') = s(u + v)$ and $s(u + v + u' + v') = s(s(u + v) + u' + v')$. Hence, $s(u + v) = u \oplus v$ is recurrent. And, it turns out that $u \oplus v \in SP(G, s)$. $\square$

In fact Theorem 1.35 proves that $SP(G, s) \cong \mathbb{Z}^{n-1} / L(G, s)$, where $L(G, s)$ is the reduced Laplacian with respect to s.

Another important result is that the sandpile group does not depend on the sink vertex, that is, $SP(G, u) \cong SP(G, v)$ for $u \neq v \in V(G)$.

Theorem 1.36. *Let $G = (V, E)$ be a graph. The sandpile group of G is independent of the sink.*

Proof. Let P be the identity matrix such that the k-th row is interchanged to the n-th row. Since $P^{-1} = P$, we know that

$$L(G, v_n) = \begin{bmatrix} d_1 & -m_{1,2} & \cdots & m_{1,n-1} & -m_{1,n} \\ -m_{2,1} & d_2 & \cdots & m_{2,n-1} & -m_{2,n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ -m_{n-1,1} & -m_{n-1,2} & \cdots & d_{n-1} & -m_{n-1,n} \\ -m_{n,1} & -m_{n,2} & \cdots & -m_{n,n-1} & d_n \end{bmatrix}.$$

Now, we have the following product

$$P \cdot L(G, v_n) \cdot P = \begin{bmatrix} d_1 & -m_{1,2} & \cdots & m_{1,n} & \cdots & -m_{1,k} \\ -m_{2,1} & d_2 & \cdots & m_{2,n} & \cdots & -m_{2,k} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ -m_{n,1} & -m_{n,2} & \cdots & d_n & \cdots & -m_{n,k} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ -m_{k,1} & -m_{k,2} & \cdots & -m_{k,n-1} & \cdots & d_k \end{bmatrix} = L(G, v_k).$$

Thus both matrices are equivalent and the result follows. ∎

1.5. Integer linear programming and the sandpile group. In this section we will see how to use integer linear programming to compute the recurrent configurations of a stable configuration, the identity of the sandpile group, and the degree of a recurrent configuration.

Let G be a multigraph with vertex set $V(G) = \{v_1, \ldots, v_n, s\}$, and degree vector $\mathbf{d}_{(G,s)} = (d_G(v_1), ..., d_G(v_n))$.

Theorem 1.37. *Let G be a multigraph with sink $s \in V(G)$. Let $\mathbf{c}$ be a stable configuration of (G, s) and $\mathbf{x}^*$ be an optimal solution of the following integer linear problem:*

$$\text{(1)} \qquad \begin{array}{ll} \text{maximize} & |\mathbf{x}| \\ \text{subject to} & \mathbf{0} \leq L(G,s)^t\mathbf{x} + \mathbf{c} \leq \mathbf{d}_{(G,s)} - \mathbf{1} \\ & \mathbf{x} \geq \mathbf{0}, \end{array}$$

then $L(G, s)^t\mathbf{x}^* + \mathbf{c} \in SP(G, s)$ and $[\mathbf{c}] = [L(G, s)^t\mathbf{x}^* + \mathbf{c}]$.

Proof. Let x^* be an optimal solution of the integer linear program (1). Clearly, $\mathbf{r} = L(G, s)^t\mathbf{x}^* + \mathbf{c}$ is a stable configuration of (G, s) and $[\mathbf{c}] = [\mathbf{r}]$ in $K(G)$. Therefore, it remains to prove that $\mathbf{r}$ is a recurrent configuration of (G, s). Let $\beta_{\min}$ be the burning configuration as in Theorem 1.22, then Theorem 1.21 implies that $\mathbf{r}$ is a recurrent configuration of (G, s) if and only if $s(\mathbf{r}+\beta_{\min}) = \mathbf{r}$ with firing vector equal to $\beta_{\min}L(G, s)^{-1}$. If we assume that $\mathbf{r}$ is not a recurrent configuration of (G, s), then there exists $\mathbf{b} \in \mathbb{N}^{V(G)\setminus s}$ such that $\mathbf{0} \leq \mathbf{b} < \beta_{\min}L(G, s)^{-1}$, and $\mathbf{r} + \beta_{\min} - \mathbf{b}^tL(G, s)$ is an stable configuration of (G, s). Since $\mathbf{r} + \beta_{\min} - \mathbf{b}^tL(G, s) = \mathbf{c} + (\mathbf{x}^* + \mathbf{z} - \mathbf{b})^tL(G, s)$, where $\mathbf{z} = \beta_{\min}L(G, s)^{-1}$ and $\mathbf{z} - \mathbf{b} > \mathbf{0}$, then $\mathbf{x}' = \mathbf{x}^* + \mathbf{z} - \mathbf{b}$ is a feasible solution of the integer linear program (1) with $\mathbf{x}^* < \mathbf{x}'$; a contradiction to the optimality of $\mathbf{x}^*$. ∎

EXAMPLE 1.38. *Let $c = (0, 0, 1, 0)$ be a configuration in (C_5, v_5). The corresponding integer linear program is:*

$$\begin{array}{rl} \text{maximize} & x_1 + x_2 + x_3 + x_4 \\ \text{subject to} & \\ & \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \leq \begin{bmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \leq \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \\ & \mathbf{x} \geq \mathbf{0} \end{array}$$

By using the following code in Maple we obtain $\mathbf{x}^* = (1, 1, 1, 1)$, and $L(G, s)^t \mathbf{x}^* + \mathbf{c} = (1, 0, 1, 1)$ is a recurrent configuration.

```
with(Optimization):
LPSolve( x[1]+x[2]+x[3]+x[4],
{
-x[3]+2*x[4] <= 1,
-x[1]+2*x[2]-x[3] <= 1,
-x[2]+2*x[3]-x[4]+1 <= 1,
2*x[1]-x[2] <= 1
},
assume = {integer, nonnegative}, maximize)
```

COROLLARY 1.39. *Let G be a multigraph with sink vertex $s \in V(G)$. Let $\mathbf{x}^*$ be an optimal solution of the following integer linear problem:*

$$\begin{array}{rl} \text{maximize} & |\mathbf{x}| \\ \text{subject to} & \mathbf{0} \leq L(G, s)^t \mathbf{x} \leq \mathbf{d}_{(G,s)} - \mathbf{1} \\ & \mathbf{x} \geq \mathbf{0}, \end{array} \tag{2}$$

then $L(G, s)^t \mathbf{x}^* \in SP(G, s)$ is the identity of $K(G)$.

PROOF. It follows from Theorem 1.37 by taking $\mathbf{c} = \mathbf{0}$. □

EXAMPLE 1.40. *Now we compute the identity configuration of $SP(\mathcal{C}_3(3, 1, 1), v_1)$. The corresponding integer linear program is:*

$$\begin{array}{rl} \text{maximize} & x_1 + x_2 \\ \text{subject to} & \\ & \begin{bmatrix} 0 \\ 0 \end{bmatrix} \leq \begin{bmatrix} 4 & -3 \\ -3 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \leq \begin{bmatrix} 3 \\ 3 \end{bmatrix} \\ & \mathbf{x} \geq \mathbf{0} \end{array}$$

We use the following code in Maple to obtain $\mathbf{x}^* = (3, 3)$, and $L(G, s)^t \mathbf{x}^* = (3, 3)$ is the identity.

```
with(Optimization):
LPSolve( x[1]+x[2],
{
4*x[1]-3*x[2] <= 3,
-3*x[1]+4*x[2] <= 3,
},
assume = {integer, nonnegative}, maximize)
```

Corollary 1.41. *Let G be a connected, r-regular multigraph, then $r\mathbf{1}$ is the identity of $SP(c(G), s)$.*

Proof. The dual linear problem of (2) is given by:

$$\begin{aligned} \text{minimize} \quad & \mathbf{y}^t(\mathbf{d}_G, \mathbf{0}) \\ \text{subject to} \quad & \left(L(c(G), s), -L(c(G), s)^t\right) \mathbf{y} \geq \mathbf{1} \\ & \mathbf{y} \geq \mathbf{0}. \end{aligned}$$

Since G is a r-regular multigraph, then the vectors $\mathbf{x}^t = r\mathbf{1}$ and $\mathbf{y}^t = (\mathbf{1}, \mathbf{0})$ are feasible integral solutions of the primal and the dual linear problems, respectively, with cost equal to $r\mathbf{1} \cdot \mathbf{1} = r|V(G)| = \mathbf{1} \cdot \mathbf{d}_G$. By the weak duality theorem [**12**, Corollary 4.2], the vector $\mathbf{x}$ is an optimal solution of the integer linear problem (2). Therefore, by Corollary 1.39, $r\mathbf{1} = rL(G, s)^t\mathbf{1} = L(G, s)^t r\mathbf{1}$ is is the identity of the sandpile group of $(c(G), s)$. ◻

Corollary 1.42. *Let G be a multigraph, $s \in V(G)$, $\mathbf{c}$ be a recurrent configuration of $SP(G, s)$, and $(d, \mathbf{x})^*$ be an optimal solution of the following integer linear problem:*

$$\begin{aligned} \text{minimize} \quad & d \\ \text{subject to} \quad & d\mathbf{c} - L(G, s)^t\mathbf{x} = \mathbf{0} \\ & d \geq 1, \mathbf{x} \geq \mathbf{0}, \end{aligned} \tag{3}$$

then d is the degree of $\mathbf{c}$ in $K(G)$.

Proof. Since $\mathbf{c}$ is recurrent, then $\mathbf{c} = [\mathbf{c}]$. And $\deg_{K(G)}([\mathbf{c}]) = \min\{d \mid d[\mathbf{c}] = [\mathbf{0}]\}$. Thus

$$\deg_{K(G)}([\mathbf{c}]) = \min\{d \mid d[\mathbf{c}] = [d\mathbf{c}] = [L(G, s)^t\mathbf{x}] = [\mathbf{0}]\} = \min\{d \mid d\mathbf{c} - L(G, s)^t\mathbf{x} = \mathbf{0}, d \geq 1, \mathbf{x} \geq \mathbf{0}\}.$$

◻

Example 1.43. *Now we compute the degree of the configuration $(1, 0, 1, 1)$ of $SP(C_5, v_5)$. The corresponding integer linear program is:*

$$\begin{aligned} \text{minimize} \quad & d \\ \text{subject to} \quad & \\ & \begin{bmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} d \\ 0 \\ d \\ d \end{bmatrix} \\ & \mathbf{x} \geq \mathbf{0} \\ & d \geq 0 \end{aligned}$$

We use the following code in Maple:

```
with(Optimization):
LPSolve( d,
{
2*x[1]-x[2] = d,
-x[1]+2*x[2]-x[3] = 0,
-x[2]+2*x[3]-x[4] = d,
-x[3]+2*x[4] = d,
d >= 1,
},
assume = {integer, nonnegative})
```

Thus we obtain that $\mathbf{x}^* = (7, 9, 11, 8)$ *and the degree* $d = 5$*. In fact,* $(1, 0, 1, 1)$ *is a generator of* $SP(C_5, v_5)$*.*

CHAPTER 2

The critical ideals

The main goals of this chapter is to introduce and to study the critical ideals. We begin by recalling some basic concepts on graph theory in Section 1. Later we establish basic properties of critical ideals in Section 2. In Section 3, we will characterize the graphs with at most two trivial critical ideals by finding their minimal set of forbidden graphs. As consequence, we will get the characterization of the graphs with two invariant factors equal to one. And in Section 4 we will give two infinite families of forbidden graphs for $\Gamma_{\leq i}$. In Sections 5 and 6, we will provide a set of minimal forbidden graphs for the set of graphs with at most three trivial critical ideals. Then we use these forbidden graphs to characterize the graphs with at most three trivial critical ideals and clique number equal to 2 and 3.

1. Preliminary definitions on graphs and matrices

Given a graph $G = (V, E)$ and a subset U of V, the subgraph of G *induced* by U is denoted by $G[U]$. If u is a vertex of G, let $N_G(u)$ be the set of *neighbors* of u in G. A *maximum clique* of a graph G is a maximum complete subgraph, whose order is the *clique number* $\omega(G)$ of G. The *path* with n vertices is denoted by P_n, a *matching* with k edges by M_k, the *complete graph* with n vertices by K_n and the *trivial graph* of n vertices by T_n. The *cone* of a graph G is the graph obtained from G by adding a new vertex, called *apex*, which is adjacent with each vertex of G. The cone of a graph G is denoted by $c(G)$. Thus, the *star* S_k of $k+1$ vertices is equal to $c(T_k)$. Given two graphs $G = (V_G, E_G)$ and $H = (V_H, E_H)$, the *union* of G and H is the union of their vertex and edge sets. And it is denoted by $G \cup H$. When V_G and V_H are disjoint, their union is referred as the *disjoint union* by $G + H$. Thus the (disjoint) union of n copies of G is denoted by nG. The *join* of G and H, denoted by $G \vee H$, is the graph obtained from $G + H$ by adding all the edges between vertices of G and H. For $m, n, o \geq 1$, let $K_{m,n,o}$ be the *complete tripartite graph*. The reader can consult [**25**] for any unexplained concept of graph theory.

Let $M \in M_n(\mathbb{Z})$ be a $n \times n$ matrix with entries on $\mathbb{Z}$, $\mathcal{I} = \{i_1, \dots, i_r\} \subseteq \{1, \dots, n\}$, and $\mathcal{J} = \{j_1, \dots, j_s\} \subseteq \{1, \dots, n\}$. The submatrix of M obtained by rows $i_1, \dots, i_r$ and columns $j_1, \dots, j_s$ is denoted by $M[\mathcal{I}; \mathcal{J}]$. If $|\mathcal{I}| = |\mathcal{J}| = r$, then $M[\mathcal{I}; \mathcal{J}]$ is called *r-square submatrix* or *square submatrix* of size r of M. A *r-minor* is the determinant of a r-square submatrix. The set of i-minors of a matrix M will be denoted by $\text{minors}_i(M)$. Finally, the identity matrix of size n is denoted by I_n and the all ones $m \times n$ matrix is denoted by $J_{m,n}$. Two matrices $M, N \in M_n(\mathbb{Z})$ are *equivalent*, if there exist $P, Q \in GL_n(\mathbb{Z})$ such that $N = PMQ$. And it is denoted by $N \sim M$.

2. Graphs with few trivial critical ideals

In this section, we will introduce the critical ideals of a graph and the class of graphs with k or less trivial critical ideals, denoted by $\Gamma_{\leq k}$. After that, we define the set of minimal forbidden graphs of $\Gamma_{\leq k}$. We finish this section with the classification of $\mathcal{G}_1$, which consists of the complete graphs.

Given a connected graph $G = (V(G), E(G))$ and a set of indeterminates $X_G = \{x_u \,|\, u \in V(G)\}$, the *generalized Laplacian matrix* $L(G, X_G)$ of G is the matrix with rows and columns indexed by $V(G)$ given by

$$L(G, X_G)_{uv} = \begin{cases} x_u & \text{if } u = v, \\ -m_{uv} & \text{otherwise,} \end{cases}$$

where m_{uv} is the *multiplicity* of the edge uv, that is, the number of the edges between vertices u and v. For all $1 \leq i \leq n$, the *i-critical ideal* of G is the determinantal ideal given by

$$I_i(G, X_G) = \langle \{\det(m) \,|\, m \text{ is a square submatrix of } L(G, X_G) \text{ of size } i\} \rangle \subseteq \mathbb{Z}[X_G].$$

By convention $I_i(G, X_G) = \langle 1 \rangle$ if $i < 1$, and $I_i(G, X_G) = \langle 0 \rangle$ if $i > n$. We say that a critical ideal is *trivial* when it is equal to $\langle 1 \rangle$.

Definition 2.1. *The algebraic co-rank of G, denoted by $\gamma(G)$, is the number of critical ideals of G equal to $\langle 1 \rangle$.*

Definition 2.2. *For all $k \in \mathbb{N}$, let $\Gamma_{\leq k} = \{G \,|\, G$ is a simple connected graph with $\gamma(G) \leq k\}$ and $\Gamma_{\geq k} = \{G \,|\, G$ is a simple connected graph with $\gamma(G) \geq k\}$.*

Note that $\Gamma_{\leq k}$ and $\Gamma_{\geq k+1}$ are disjoint sets and that a characterization of one of the sets leads to a characterization of the other one. Now let us recall some basic properties about critical ideals, for more details see [**22**]. It is known that if $i \leq j$, then $I_j(G, X_G) \subseteq I_i(G, X_G)$. Moreover, if H is an induced subgraph of G, then $I_i(H, X_H) \subseteq I_i(G, X_G)$, for all $i \leq |V(H)|$, and hence $\gamma(H) \leq \gamma(G)$. This implies that $\Gamma_{\leq k}$ is closed under induced subgraphs, that is, if $G \in \Gamma_{\leq k}$ and H is an induced subgraph of G, then $H \in \Gamma_{\leq k}$.

Recall that $f_k(G)$ denote the number of invariant factors of $K(G)$ that are equal to k, and $\mathcal{G}_i = \{G : G$ is a simple connected graph with $f_1(G) = i\}$. Presumably, the set $\Gamma_{\leq k}$ behaves better than $\mathcal{G}_k$. It is not difficult to see that unlike of $\Gamma_{\leq k}$ the set $\mathcal{G}_k$ is not closed under induced subgraphs. For instance, $c(S_3)$ belongs to $\mathcal{G}_2$, but S_3 belongs to $\mathcal{G}_3$. Similarly, the graph $K_6 \setminus \{2P_2\}$ belongs to $\mathcal{G}_3$, meanwhile $K_5 \setminus \{2P_2\}$ belongs to $\mathcal{G}_2$. Moreover, if H is an induced subgraph of G, then it is not always true that $K(H) \trianglelefteq K(G)$. For example, $K(K_4) \cong \mathbb{Z}_4^2 \ntrianglelefteq K(K_5) \cong \mathbb{Z}_5^3$. Finally, Theorems 2.10 and 2.22 give us additional evidence that $\Gamma_{\leq k}$ behaves better than $\mathcal{G}_k$. Moreover, Theorem 3.6 of [**22**] implies that $\gamma(G) \leq f_1(G)$ for any graph, and thus $\mathcal{G}_k \subseteq \Gamma_{\leq k}$ for all $k \geq 0$.

Definition 2.3. *A graph G is forbidden (or an obstruction) for $\Gamma_{\leq k}$ if and only if $\gamma(G) \geq k+1$. Let $\mathbf{Forb}(\Gamma_{\leq k})$ be the set of minimal (under the induced subgraphs property) forbidden simple graphs for $\Gamma_{\leq k}$. A graph G is called γ-critical if $\gamma(G \setminus v) < \gamma(G)$ for all $v \in V(G)$.*

Thus $G \in \mathbf{Forb}(\Gamma_{\leq k})$ if and only if G is γ-critical with $\gamma(G) \geq k + 1$ and $\gamma(G - v) \leq k$ for each $v \in V(G)$.

Given a family of graphs $\mathcal{F}$, a graph G is called $\mathcal{F}$-free if no induced subgraph of G is isomorphic to a member of $\mathcal{F}$. Thus, G belongs to $\Gamma_{\leq k}$ if and only if G is $\mathbf{Forb}(\Gamma_{\leq k})$-free, or equivalently, G belongs to $\Gamma_{\geq k+1}$ if and only if G contains a graph of $\mathbf{Forb}(\Gamma_{\leq k})$ as an induced subgraph.

These ideas are useful in characterization of $\Gamma_{\leq k}$. For instance, since $\gamma(P_2) = 1$ and no one of its induced subgraphs has $\gamma \geq 1$, then $P_2 \in \mathbf{Forb}(\Gamma_{\leq 0})$. Moreover, it is easy to see that T_1 is the only connected graph P_2-free. Thus, since $I_1(T_1, \{x\}) \neq \langle 1 \rangle$, then we get that $\mathbf{Forb}(\Gamma_{\leq 0}) = \{P_2\}$, and $\Gamma_{\leq 0}$ consists of the graph with one vertex. Also, it is not difficult to prove that $\mathcal{G}_0 = \Gamma_{\leq 0}$ and that the set of non-necessarily connected graphs with algebraic co-rank equal to zero consists only of the trivial graphs. In the next section, we will use this kind of arguments in order to get

Forb$(\Gamma_{\leq k})$ and characterize $\Gamma_{\leq k}$ for k equal to 1 and 2. Now, we obtain the characterization of $\Gamma_{\leq 1}$.

THEOREM 2.4. *If G is a simple connected graph, then the following statements are equivalent:*
(i) $G \in \Gamma_{\leq 1}$,
(ii) G *is* P_3*-free,*
(iii) G *is a complete graph.*

PROOF. *(i)* $\Rightarrow$ *(ii)* Since $\gamma(P_3) = 2$, then clearly G must be P_3-free.

(ii) $\Rightarrow$ *(iii)* If G is not a complete graph, then it has two vertices not adjacent, say u and v. Let P be the smallest path between u and v. Thus, the length of P is greater or equal to 3. So, P_3 is an induced subgraph of both P and G. Therefore, G is a complete graph.

(iii) $\Rightarrow$ *(i)* It is easy to see that for any non-trivial simple connected graph, its first critical ideal is trivial. Meanwhile $I_1(K_1, \{x\}) = \langle x\rangle$. On the other hand, the 2-minors of a complete graphs are of the forms: $-1 + x_ix_j$ and $1 + x_i$. Since $-1 + x_ix_j \in \langle 1 + x_1, ..., 1 + x_n\rangle$, then

$$I_2(K_n, X_{K_n}) = \begin{cases} \langle -1 + x_1x_2\rangle & \text{if } n = 2, \text{ and,} \\ \langle 1 + x_1, ..., 1 + x_n\rangle & \text{if } n \geq 3. \end{cases} \tag{4}$$

Therefore $\gamma(K_n) \leq 1$. In fact, the set $\{1+x_1, ..., 1+x_n\}$ is a reduced Gröbner basis for $I_2(K_n, X_{K_n})$, see [**22**, Theorem 3.14]. ☐

In light of Theorem 2.4, the characterization of $\mathcal{G}_1$ is as follows: Clearly, $\mathcal{G}_1 \subseteq \Gamma_{\leq 1} \setminus \mathcal{G}_0$. Now let $G \in \Gamma_{\leq 1} \setminus \{K_1\}$, that is, $G = K_n$ with $n \geq 2$ and $f_1(G) \geq 1$. It is easy to verify from Equation 4 that the second invariant factor of $K(G)$ is equal to $I_2(K_n, X_{K_n}) \mid_{\{x_v = n-1 \,|\, v \in K_n\}}$ which is different to $\langle 1\rangle$.

COROLLARY 2.5. [**34**] *If G is a simple connected graph with $n \geq 2$ vertices, then $f_1(G) = 1$ if and only if G is a complete graph.*

A crucial fact in the proof of Theorem 2.4 was that P_3 belongs to **Forb**$(\Gamma_{\leq 1})$, and that any other connected simple graph belonging to $\Gamma_{\geq 2}$ contains P_3. This leads to the following corollary.

COROLLARY 2.6. **Forb**$(\Gamma_{\leq 1}) = \{P_3\}$.

Next corollary give us the non-connected version of Theorem 2.4.

COROLLARY 2.7. *If G is a simple non-necessary connected graph, then the following statements are equivalent:*
(i) $\gamma(G) \leq 1$,
(ii) G *is* $\{P_3, 2P_2\}$*-free,*
(iii) G *is a disjoint union of a complete graph and a trivial graph.*

We proceed with the proof of Corollary 2.7, we present a lemma that help us to calculate the critical ideal of a non-connected graph. It may be useful to recall that the *product* of the ideals I and J of a commutative ring R, which we denote by IJ, is the ideal generated by all the products ab where $a \in I$ and $b \in J$.

LEMMA 2.8. [**22**, Proposition 3.4] *If G and H are two vertex-disjoint graphs, then*

$$I_i(G + H, \{X_G, Y_H\}) = \left\langle \cup_{j=0}^{i} I_j(G, X_G) I_{i-j}(H, Y_H)\right\rangle \text{ for all } 1 \leq i \leq |V(G + H)|.$$

By this lemma we have that $\gamma(G+H)=\gamma(G)+\gamma(H)$ when G and H are vertex-disjoint.

Proof of corollary 2.7. $(i)\Rightarrow(ii)$ It follows since $\gamma(2P_2)=2$ and $\gamma(P_3)=2$.

$(ii)\Rightarrow(iii)$ Let $G_1,\ldots,G_s$ be the connected components of G. Then by Theorem 2.4 and Lemma 2.8, G_i is a complete graph for all $1\le i\le s$. Since $2P_2$ must not be an induced subgraph of G, then at most one of the G_i has order greater than 1.

$(iii)\Rightarrow(i)$ If $G=K_n+T_m$, then it is not difficult to see that $I_1(T_m,Y_{T_m})=\langle y_1,...,y_m\rangle$ and $I_2(T_m,Y_{T_m})=\left\langle\prod_{i\neq j}y_iy_j\right\rangle$. Thus by Lemma 2.8,

$$I_2(G,\{X_{K_n},Y_{T_m}\})=\langle I_2(K_n,X_{K_n}),I_1(K_n,X_{K_n})I_1(T_m,Y_{T_m}),I_2(T_m,Y_{T_m})\rangle\neq\langle 1\rangle\,.$$

□

3. Graphs with algebraic co-rank equal to two

The main goal of this section is to classify the simple graphs on $\Gamma_{\le 2}$. After, we use the contention $\mathcal{G}_2\subseteq\Gamma_{\le 2}$ to classify the simple graphs whose critical group has two invariant factors equal to 1. As in the case of $\Gamma_{\le 1}$, the characterization of $\Gamma_{\le 2}$ relies heavily in the two facts: (1) $\Gamma_{\le 2}$ is closed under induced subgraphs and (2) we have a good guessing about $\mathbf{Forb}(\Gamma_{\le 2})$. We begin by introducing a set of graphs in the $\mathbf{Forb}(\Gamma_{\le 2})$.

Proposition 2.9. *Let $\mathcal{F}_2$ be the set of graphs consisting of P_4, $K_5\setminus S_2$, $K_6\setminus M_2$,* cricket *and* dart*, see Figure 3. Then $\mathcal{F}_2\subseteq\mathbf{Forb}(\Gamma_{\le 2})$.*

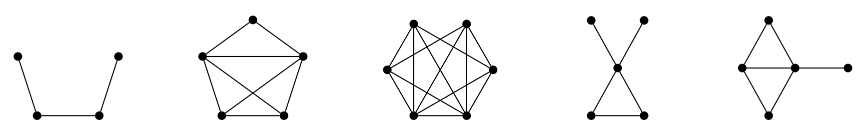


Figure 3. The set $\mathcal{F}_2$ of graphs.

Proof. It is not difficult to see that the generalized Laplacian matrix of the graphs on $\mathcal{F}_2$ are given by:

$$L(P_4)=\begin{bmatrix} x_1 & -1 & 0 & 0\\ -1 & x_2 & -1 & 0\\ 0 & -1 & x_3 & -1\\ 0 & 0 & -1 & x_4\end{bmatrix},\qquad L(K_5\setminus S_2)=\begin{bmatrix} x_1 & 0 & -1 & -1 & 0\\ 0 & x_2 & -1 & -1 & -1\\ -1 & -1 & x_3 & -1 & -1\\ -1 & -1 & -1 & x_4 & -1\\ 0 & -1 & -1 & -1 & x_5\end{bmatrix},$$

$$L(\mathsf{cricket})=\begin{bmatrix} x_1 & 0 & 0 & -1 & 0\\ 0 & x_2 & -1 & -1 & 0\\ 0 & -1 & x_3 & -1 & 0\\ -1 & -1 & -1 & x_4 & -1\\ 0 & 0 & 0 & -1 & x_5\end{bmatrix},\qquad L(\mathsf{dart})=\begin{bmatrix} x_1 & -1 & 0 & -1 & 0\\ -1 & x_2 & -1 & -1 & 0\\ 0 & -1 & x_3 & -1 & 0\\ -1 & -1 & -1 & x_4 & -1\\ 0 & 0 & 0 & -1 & x_5\end{bmatrix},$$

$$L(K_6\setminus M_2)=\begin{bmatrix} x_1 & 0 & -1 & -1 & -1 & -1\\ 0 & x_2 & -1 & -1 & -1 & -1\\ -1 & -1 & x_3 & -1 & -1 & 0\\ -1 & -1 & -1 & x_4 & -1 & -1\\ -1 & -1 & -1 & -1 & x_5 & -1\\ -1 & -1 & 0 & -1 & -1 & x_6\end{bmatrix}.$$

In these matrices, we have marked in gray some 3×3 submatrices whose determinant is equal to ± 1. Then $\gamma(G) \geq 3$ for all $G \in \mathcal{F}_2$. Finally, using any computer algebra system, like *Macaulay 2*, one can note that the graphs in $\mathcal{F}_2$ have algebraic co-rank equal to 3. Moreover, it can be checked that any of its induced subgraphs has algebraic co-rank less than or equal to 2. $\square$

One of the main results of this chapter is the following:

THEOREM 2.10. *Let G be a simple connected graph. Then, $G \in \Gamma_{\leq 2}$ if and only if G is an induced subgraph of $K_{m,n,o}$ or $T_n \vee (K_m + K_o)$ with $m, n, o \geq 0$.*

We divide the proof of Theorem 2.10 in two steps. First we classify the connected graphs that are $\mathcal{F}_2$-free. After that, we check that all these graphs have algebraic co-rank less than or equal to two.

THEOREM 2.11. *A simple connected graph is $\mathcal{F}_2$-free if and only if it is an induced subgraph of $K_{m,n,o}$ or $T_n \vee (K_m + K_o)$ with $m, n, o \geq 0$.*

PROOF. First, one implication is clear, because $K_{m,n,o}$ and $T_n \vee (K_m + K_o)$ are $\mathcal{F}_2$-free. The other part is divided in three cases by considering the clique number $\omega(G)$ of G: when $\omega(G) = 2$, $\omega(G) = 3$, and $\omega(G) \geq 4$.

The case when $\omega(G) = 2$ is very simple. Since $\omega(G) = 2$, there exist $a, b \in V(G)$ such that $ab \in E(G)$. Clearly, $N_G(a) \cap N_G(b) = \emptyset$. Moreover, if $x \in \{a, b\}$, then $uv \notin E(G)$ for all $u, v \in N_G(x)$. On the other hand, since G is P_4-free, then $uv \in E(G)$ for all $u \in N_G(a)$ and $v \in N_G(b)$. Therefore, G is the complete bipartite graph.

Now assume that $\omega(G) = 3$. Let a, b and c be vertices of G that induce a complete graph. For all $X \subseteq \{a, b, c\}$, let $V_X = \{v \in V(G) : N_G(v) \cap \{a, b, c\} = X\}$. Clearly $V_{\{a,b,c\}} = \emptyset$, because $\omega(G) = 3$. In a similar way, if $X \subseteq \{a, b, c\}$ has size two, then set V_X induce a trivial graph. Also, since G is cricket-free, V_x induces a complete graph for all $x \in \{a, b, c\}$. Thus V_x has at most two vertices.

Now, given $U, V \in V(G)$, let $E(U, V) = \{uv \in E(G) : u \in U \text{ and } v \in V\}$. Let $x \neq y \in \{a, b, c\}$ and $z \in \{a, b, c\}$ such that $\{x, y, z\} = \{a, b, c\}$. Assume that V_x, V_y and $V_{\{x,y\}}$ are not empty. Let $u \in V_x$ and $v \in V_y$. If $uv \in E(G)$, then $\{u, v, y, z\}$ induced a P_4. Therefore, $E(V_x, V_y) = \emptyset$. In a similar way, since G is P_4-free, we get $E(V_x, V_{\{x,y\}}) = \emptyset$.

CLAIM 2.12. *At least two of the sets V_a, V_b or V_c are empty. Furthermore, if $V_a \neq \emptyset$, then G is an induced subgraph of $T_l \vee (K_2 + K_2)$, where $l = |V_{\{b,c\}}| + 1$.*

PROOF. First, assume that the sets V_x and V_y are non empty. Let $u \in V_y$, $v \in V_x$. Since u and v are not adjacent, the vertices $\{u, x, y, v\}$ induce a P_4. Therefore, at least one of the sets V_x or V_y is empty.

Without loss of generality, assume that V_a is not empty. Since there is no edge between V_a and $V_{\{a,b\}}$, then $V_{\{a,b\}} = \emptyset$. Otherwise, if $u \in V_{\{a,b\}}$ and $v \in V_a$, then the vertices $\{u, v, a, b, c\}$ induces a dart. In a similar way $V_{\{a,c\}} = \emptyset$. On the other hand, if $V_{\{b,c\}}$ is not empty and there are two vertices $u \in V_{\{b,c\}}$ and $v \in V_a$ such that $uv \notin E(G)$, then the vertices $\{u, b, a, v\}$ induces a P_4. Therefore, either $E(V_a, V_{\{b,c\}}) = \{uv \,|\, u \in V_a \text{ and } v \in V_{\{b,c\}}\}$ or the set $V_{\{b,c\}}$ is empty. Finally, since V_a is a complete graph with at most two vertices, the result follows. $\square$

Now, we can assume that $V_x = \emptyset$ for all $x \in \{a, b, c\}$. Let $\{x, y, z\} = \{a, b, c\}$. If $uv \notin E(G)$ for some $u \in V_{\{x,y\}}$, $v \in V_{\{x,z\}}$, then $\{u, y, z, v\}$ induces a P_4. Therefore, $uv \in E(G)$ for all $u \in V_{\{x,y\}}$ and $u \in V_{\{x,z\}}$, and G is an induced subgraph of the complete tripartite graph.

We finish with case when $\omega(G) \geq 4$. Let $W = \{a,b,c,d\}$ be a complete subgraph of G of size four and let

$$V_i = \{v \in V(G) \setminus W \,:\, |N_G(V) \cap W| = i\} \text{ for all } i = 0,1,2,3,4.$$

Since G is $K_5 \setminus S_2$-free, then $V_2 = \emptyset$.

CLAIM 2.13. *The graph induced by the set V_1 is a complete graph.*

PROOF. Let $u, u' \in V_1$ and suppose there is no edge between u and u'. Let $x, y \in W$ be the vertices adjacent with u and u', respectively. If $x \neq y$, then $\{u, x, y, u'\}$ induces a P_4; a contradiction. On the other hand, if $x = y$, let $z \neq w \in W \setminus x$. Since u and u' are not adjacent with both z and w, then $\{x, z, w, u, u'\}$ induces a cricket; a contradiction. □

Let $v, v' \in V_3$ and assume that are adjacent. Let $x, y \in W$ such that $x \notin N_G(v)$ and $y \notin N_G(v')$. If $x \neq y$, then $\{v, v'\} \cup W$ induces a $K_6 \setminus M_2$; a contradiction. On the other hand, if $x = y$, then $\{v, v'\} \cup W$ contains a $K_5 \setminus S_2$ as induced graph; a contradiction. Therefore, V_3 induces a trivial graph.

Now let $u \in V_1$, $v \in V_3$, $x, y \in W$ such that $xu \in E(G)$, $yv \notin E(G)$. Assume that $uv \notin E(G)$. Let $z \in W \setminus \{x, y\}$. If $x = y$, then $\{v, z, x, u\}$ induces a P_4; a contradiction. On the other hand, if $x \neq y$, then G must contains a dart as induced subgraph; a contradiction. Therefore $E(V_1, V_3)$ contains all the possible edges. Since $uv \in E(G)$, then $x = y$. Otherwise, if $x \neq y$, then $\{y, z, v, u\}$ induces a P_4; a contradiction. Therefore we can assume without loss of generality that $\{a\} = N_G(V_1) \cap W = (N_G(V_3) \cap W)^c$.

Now, let $w \in V_4$, $u \in V_1$, and $v \in V_3$. If $uw \in E(G)$, then $\{u, w, a, b, c\}$ induces a $K_5 \setminus S_2$. Therefore, $E(V_1, V_4) = \emptyset$. In a similar way, if $vw \notin E(G)$, then $\{v, a, w, b, c\}$ induces a $K_5 \setminus S_2$. Therefore, $E(V_3, V_4) = \{vw \,|\, v \in V_3 \text{ and } w \in V_4\}$.

Since G is $\{K_5 \setminus S_2, K_6 \setminus M_2\}$, then is not difficult to see that the graph induced by V_4 is $\{K_2 + T_1, C_4\}$-free. Thus V_4 induces either a trivial graph, a complete graph, or a complete graph minus an edge. Moreover, if $ww' \notin E(G)$ for some $w \neq w' \in V_4$, then $\{w, w', a, v, b, c\}$ induces a $K_6 \setminus M_2$. Thus, if $V_3 \neq \emptyset$, then V_4 induces a complete graph. Therefore, if $V_1, V_3, V_4 = \emptyset$, then G is a complete graph.

CLAIM 2.14. *If $V_1, V_3 = \emptyset$ and $V_4 \neq \emptyset$, then G is an induced subgraph of $T_1 \vee (K_m + K_n)$ for some $m, n \in \mathbb{N}$.*

PROOF. If $|V_| = |V_4| = 1$, then the result is clear. Therefore, we can assume that either $|V_4| \geq 2$ or $|V_0| \geq 2$. Moreover, we need to consider three cases for V_4, when it induces a trivial graph, a complete graph, or a complete graph minus an edge. Assume that V_4 induces a trivial graph. If $|V_4| \geq 2$, let $o \in V_0$ and $w, w' \in V_4$. If $ow \in E(G)$ and $ow' \in E(G)$, then $\{o, w, w', a\}$ induces a P_4; a contradiction. Thus, either $E(o, V_4) = \{ow \,|\, w \in V_4\}$ or $E(o, V_4)$ is empty. Therefore, since G is connected, we get the result when $|V_0| = 1$.

Now, assume that $|V_0| \geq 2$. Since G is connected, there exist $o \in V_0$ such that $ow \in E(G)$ for some $w \in V_4$. Let $o' \in V_0$ such that $E(o', V_4)$ is empty. Since G is connected, there exist a path from o' to o. Let P be a minimum path between o' and o. In this case, $\{V(P), w, a\}$ induces a path with more than four vertices; a contradiction. Therefore, $E(V_0, V_4) = \{ow \,|\, o \in V_0 \text{ and } w \in V_4\}$. Moreover, since G is $K_6 \setminus M_2$-free, then V_0 induces a trivial graph and we get the result.

Now, assume that V_4 induces a complete graph. Since G is $K_5 \setminus S_2$-free, o is adjacent with at most one vertex in V_4. Moreover, all the vertices in V_0 are adjacent with a unique vertex in V_4. Otherwise, let $o, o' \in V_0$ and $w, w' \in V_4$ such that $ow, o'w' \in E(G)$ and $ow', o'w \notin E(G)$. If

$oo' \in E(G)$, then $\{a, w, o, o'\}$ induces a P_4; a contradiction. Also, if $oo' \notin E(G)$ and $ww' \in E(G)$, then $\{w, w', o, o'\}$ induces a P_4; a contradiction. Let $w \in V_4$ such that all the vertices in V_0 are adjacent with w. Then V_0 induces a complete graph. Otherwise, $\{a, b, w, o, o'\}$ induces a cricket; a contradiction. Therefore G is an induced subgraph of $T_1 \vee (K_m + K_n)$ for some $m, n \in \mathbb{N}$.

Finally, when V_4 induces a complete graph minus an edge, following similar arguments to those given in the case when V_4 induces a complete graph we get that G is an induced subgraph of $T_2 \vee (K_m + K_n)$ for some $m, n \in \mathbb{N}$. $\square$

Therefore we can assume that $V_1 \cup V_3 \neq \emptyset$. Let $u \in V_1 \cup V_3$, $o \in V_0$, and $x \neq y \in W$ such that $x \notin N_G(u)$ and $y \in N_G(u)$. If $uo \notin E(G)$, then $\{x, y, u, o\}$ induces a P_4; a contradiction. Thus, there are no edges between the vertex sets V_0 and $V_1 \cup V_3$. Moreover, let $w \in V_4$. If $ow \in E(G)$, then $\{a, b, u, w, o\}$ induces a dart when $u \in V_3$ and $\{u, a, w, o\}$ induces a P_4 when $u \in V_1$. Therefore, there are no edges between V_0 and V_4. Since G is connected, $V_0 = \emptyset$ and therefore G is an induced subgraph of $T_n \vee (K_m + K_o)$. $\square$

To finish the proof of Theorem 2.10 we need to prove that the third critical ideal of the graphs $K_{m,n,o}$ and $T_n \vee (K_m + K_o)$ is not trivial. If $m + n + o \leq 2$, then the third critical ideal is equal to zero. Also, if $m + n + o = 3$, then the third critical ideal is equal to the determinant of the generalized Laplacian matrix. Moreover, [**22**, Theorem 3.16] proves that the algebraic co-rank of the complete graph is equal to 1.

THEOREM 2.15. *If $K_{m,n,o}$ is connected with $m \geq n \geq o$ and $m + n + o \geq 4$, then*

$$I_3(K_{m,n,o}, \{X, Y, Z\}) = \begin{cases} \langle 2, \bigcup_{i=1}^m x_i, \bigcup_{i=1}^n y_i, \bigcup_{i=1}^o z_i \rangle & \text{if } m, n, o \geq 2, \\ \langle \bigcup_{i=1}^m x_i, \bigcup_{i=1}^n y_i, z_1 + 2 \rangle & \text{if } m \geq 2, n \geq 2, o = 1, \\ \langle \bigcup_{i=1}^m x_i, y_1 + z_1 + 2 \rangle & \text{if } m \geq 3, n = 1, o = 1, \\ \langle x_1x_2 + x_1 + x_2, x_1z_1 + x_1, x_2z_1 + x_2, y_1 + z_1 + 2 \rangle & \text{if } m = 2, n = 1, o = 1, \\ \langle \bigcup_{i=1}^m x_i, \bigcup_{i=1}^n y_i \rangle & \text{if } m \geq 3, n \geq 3, o = 0, \\ \langle \bigcup_{i=1}^m x_i, y_1 + y_2 \rangle & \text{if } m \geq 3, n = 2, o = 0, \\ \langle x_2y_2, x_1 + x_2, y_1 + y_2 \rangle & \text{if } m = 2, n = 2, o = 0, \\ \langle \bigcup_{i=1}^m x_i \rangle & \text{if } m \geq 3, n = 1, o = 0. \end{cases} \tag{5}$$

THEOREM 2.16. *If $T_n \vee (K_m + K_o)$ is connected with $m \geq o$, $m + n + o \geq 4$ such that $T_n \vee (K_m + K_o)$ is not the complete graph or the complete bipartite graph, then*

$$I_3(T_n \vee (K_m + K_o), \{X, Y, Z\}) = \begin{cases} \langle 2, \bigcup_{i=1}^m (x_i + 1), \bigcup_{i=1}^n y_i, \bigcup_{i=1}^o (z_i + 1) \rangle & \text{if } m, n, o \geq 2, \\ \langle \bigcup_{i=1}^m (x_i + 1), y_1 + 2, \bigcup_{i=1}^o (z_i + 1) \rangle & \text{if } m \geq 2, n = 1, o \geq 2, \\ \langle \bigcup_{i=1}^m (x_i + 1), \bigcup_{i=1}^n y_i, z_1 - 1 \rangle & \text{if } m \geq 2, n \geq 2, o = 1, \\ \langle x_1 + z_1, \bigcup_{i=1}^n y_i \rangle & \text{if } m = 1, n \geq 3, o = 1, \\ \langle x_1 + z_1, y_1 + y_2, y_2z_1, \rangle & \text{if } m = 1, n = 2, o = 1, \\ \langle \bigcup_{i=1}^m (x_i + 1), z_1y_1 + z_1 - 1 \rangle & \text{if } m \geq 2, n = 1, o = 1, \\ \langle \bigcup_{i=1}^m (x_i + 1), \bigcup_{i=1}^n y_i \rangle & \text{if } m \geq 3, n \geq 3, o = 0, \\ \langle x_1 + x_2 + 2, \bigcup_{i=1}^n y_i \rangle & \text{if } m = 2, n \geq 3, o = 0, \\ \langle \bigcup_{i=1}^m (x_i + 1), y_1y_2 + y_1 + y_2 \rangle & \text{if } m \geq 3, n = 2, o = 0, \\ \langle x_1 + x_2 + 2, x_2y_1 + y_1, x_2y_2 + y_2, y_1y_2 + y_1 + y_2 \rangle & \text{if } m = 2, n = 2, o = 0, \end{cases} \tag{6}$$

The proofs of Theorems 2.15 and 2.16 relies on the description of the 3-minors of the generalized Laplacian matrices of $K_{m,n,o}$ and $T_n \vee (K_m + K_o)$.

PROOF OF THEOREM 2.15. In order to simplify the arguments in the proof we separate it in two parts. We begin by finding the 3-minors of the generalized Laplacian matrix of the complete bipartite graph and using it to calculate their third critical ideals. An after that, we do the same for the general case of the complete tripartite graph.

LEMMA 2.17. *For $m, n \geq 1$, let $L_{m,n}$ be the generalized Laplacian matrix of the complete bipartite graph $K_{m,n}$. That is,*

$$L_{m,n} = L(K_{m,n}, \{X_{T_m}, Y_{T_n}\}) = \begin{bmatrix} L(T_m, X_{T_m}) & -J_{m,n} \\ -J_{n,m} & L(T_n, Y_{T_n}) \end{bmatrix}.$$

Then every 3-minor of $L_{m,n}$ is equal, up to sign, to one of the following polynomials listed below:

- y_{j_1}, $y_{j_1}y_{j_2}$, *and* $y_{j_1}y_{j_2}y_{j_3}$ *when* $n \geq 3$,
- $y_{j_1}y_{j_2}x_{i_1} - y_{j_1} - y_{j_2}$ *when* $n \geq 2$,
- $x_{i_1} + x_{i_2}$, $y_{j_1} + y_{j_2}$ *and* $x_{i_1}y_{j_1}$ *when* $m \geq 2$ *and* $n \geq 2$,
- x_{i_1}, $x_{i_1}x_{i_2}$, *and* $x_{i_1}x_{i_2}x_{i_3}$ *when* $m \geq 3$,
- $x_{i_1}x_{i_2}y_{j_1} - x_{i_1} - x_{i_2}$ *when* $m \geq 2$,

where $1 \leq i_1 < i_2 < i_3 \leq n$ and $1 \leq j_1 < j_2 < j_3 \leq n$.

PROOF. Before to proceed with the proof we establish some notation corresponding to row and column indices. Let $\mathcal{I} = \{i_1, i_2, i_3\}$ such that $1 \leq i_1 < i_2 < i_3 \leq m+n$, and $\mathcal{J} = \{j_1, j_2, j_3\}$ such that $1 \leq j_1 < j_2 < j_3 \leq m+n$. Let $\mathcal{I}_1 = \mathcal{I} \cap [m]$, $\mathcal{I}_2 = \mathcal{I}_1^c$, $\mathcal{J}_1 = \mathcal{J} \cap [m]$, and $\mathcal{J}_2 = \mathcal{J}_1^c$. Also in the following $i'_t = i_t - m$ and $j'_t = j_t - m$, for all $1 \leq t \leq 3$.

In order to find all the 3-minors of $L_{m,n}$ we need to calculate the determinants of all non-singular matrices of the form $L_{m,n}[\mathcal{I}, \mathcal{J}]$. Since the generalized Laplacian matrix is symmetric, we can assume without loss of generalization that $|\mathcal{I}_2| \leq |\mathcal{J}_2|$. Let $L = L_{m,n}[\mathcal{I}; \mathcal{J}]$ be non-singular. First, consider the case when $\mathcal{I}_2$ is empty. Since the determinant of L is equal to zero when $|\mathcal{J}_2| \geq 2$, only remains to consider the cases when $|\mathcal{J}_2| = 0$ or $|\mathcal{J}_2| = 1$. If $|\mathcal{J}_2| = 0$, then $m \geq 3$, L is a submatrix of $L(T_m, X_{T_m})$, and the determinant of L is equal to $x_{i_1}x_{i_2}x_{i_3}$. In a similar way, if $|\mathcal{J}_2| = 1$, then $m \geq 3$, $n \geq 1$, and L is equal to (up to row permutation)

$$\begin{bmatrix} x_{j_1} & 0 & -1 \\ 0 & x_{j_2} & -1 \\ 0 & 0 & -1 \end{bmatrix}$$

whose determinant is equal to $-x_{j_1}x_{j_2}$.

Now, consider the case when $|\mathcal{I}_2| = 1$. In a similar way, L has determinant different from zero when $|\mathcal{J}_2| = 1$ or $|\mathcal{J}_2| = 2$. If $|\mathcal{J}_2| = 1$, then there are essentially only four 3×3 non-singular submatrices of $L_{m,n}$:

$$\begin{bmatrix} x_{i_1} & 0 & -1 \\ 0 & A & -1 \\ -1 & -1 & B \end{bmatrix},$$

where A is equal to 0 (when $m \geq 3$) and x_{i_2}, and B is equal to 0 (when $n \geq 2$) and $y_{i'_3}$. Clearly $\det(L) = ABx_{i_1} - A - x_{i_1}$. Thus we have the following minors: $x_{i_1}x_{i_2}y_{i'_3} - x_{i_1} - x_{i_2}$, $-x_{i_1} - x_{i_2}$, $-x_{i_1}$. If $|\mathcal{J}_2| = 2$, then $m \geq 2$, $n \geq 2$, and L has determinant equal to

$$\det \begin{bmatrix} x_{j_1} & -1 & -1 \\ 0 & -1 & -1 \\ -1 & 0 & y_{i'_3} \end{bmatrix} = -x_{j_1}y_{i'_3}.$$

When $|\mathcal{I}_2| = 2$ we have two cases, when either $|\mathcal{J}_2| = 2$ or $|\mathcal{J}_2| = 3$. If $|\mathcal{J}_2| = 2$, then L is equal to:

$$\begin{bmatrix} A & -1 & -1 \\ -1 & y_{i'_2} & 0 \\ -1 & 0 & B \end{bmatrix}$$

where A is equal to 0 (when $m \geq 2$) or x_{i_1} and B is equal to 0 (when $n \geq 3$) or $y_{i'_3}$. It is easy to see that $\det(L) = ABy_{i'_2} - A - y_{i'_2}$. Thus we have the following minors: $x_{i_1}y_{i'_2}y_{i'_3} - y_{i'_2} - y_{i'_3}$, $-y_{i'_2} - y_{i'_3}$, $-y_{i'_2}$. If $|\mathcal{J}_2| = 3$, then $m \geq 1$, $n \geq 3$ and there are only one non-singular matrix whose determinant is equal to

$$\det \begin{bmatrix} -1 & -1 & -1 \\ y_{i'_2} & 0 & 0 \\ 0 & y_{i'_3} & 0 \end{bmatrix} = -y_{i'_2}y_{i'_3}.$$

Finally, if $|\mathcal{I}_2| = 3$, then $n \geq 3$, L is a submatrix of $L(T_m, Y_{T_m})$, and therefore its determinant is equal to $y_{i'_1}y_{i'_2}y_{i'_3}$. □

We can use Lemma 2.17 to get the third critical ideal of the complete bipartite graph. For instance, it is not difficult to see that $I_3(K_{m,n}, \{X, Y\}) = \langle \bigcup_{i=1}^m x_i, \bigcup_{i=1}^n y_i \rangle$ when $m \geq 3$ and $n \geq 3$. In a similar way, since $x_{i_1} + x_{i_2}, x_{i_1}y_{j_1}, y_{j_1}y_{j_2}x_{i_1} - y_{j_1} - y_{j_2}, x_{i_1}x_{i_2}, x_{i_1}x_{i_2}x_{i_3} \in \langle \bigcup_{i=1}^m x_i, y_1 + y_2 \rangle$, $I_3(K_{m,n}, \{X, Y\}) = \langle \bigcup_{i=1}^m x_i, y_1 + y_2 \rangle$ when $m \geq 3$ and $n = 2$. The other cases follow in a similar way.

Therefore in order to calculate the third critical ideal of the complete tripartite graph we need to calculate their 3-minors as below.

THEOREM 2.18. *For $m, n, o \geq 1$, let $L_{m,n,o}$ be the generalized Laplacian matrix of the tripartite complete graph $K_{m,n,o}$. That is,*

$$L_{m,n,o} = L(K_{m,n,o}, \{X_{T_m}, Y_{T_n}, Z_{T_o}\}) = \begin{bmatrix} L(T_m, X_{T_m}) & -J_{m,n} & -J_{m,o} \\ -J_{n,m} & L(T_n, Y_{T_n}) & -J_{n,o} \\ -J_{o,m} & -J_{o,n} & L(T_o, Z_{T_o}) \end{bmatrix}.$$

Then every 3-minor of $L_{m,n,o}$ is equal, up to sign, to one of the polynomials listed below:

- x_{i_1}, $x_{i_1}x_{i_2}$, *and* $x_{i_1}x_{i_2}x_{i_3}$ *when* $m \geq 3$,
- y_{j_1}, $y_{j_1}y_{j_2}$, *and* $y_{j_1}y_{j_2}y_{j_3}$ *when* $n \geq 3$,
- z_{k_1}, $z_{k_1}z_{k_2}$, *and* $z_{k_1}z_{k_2}z_{k_3}$ *when* $o \geq 3$,
- x_{i_1}, y_{j_1}, $x_{i_1} + 2$, $y_{j_1} + 2$, $x_{i_1} + x_{i_2}$, $y_{j_1} + y_{j_2}$, *and* $x_{i_1}y_{j_1}$ *when* $m \geq 2$ *and* $n \geq 2$,
- x_{i_1}, z_{k_1}, $x_{i_1} + 2$, $z_{k_1} + 2$,$x_{i_1} + x_{i_2}$, $z_{k_1} + z_{k_2}$, *and* $x_{i_1}z_{k_1}$ *when* $m \geq 2$ *and* $o \geq 2$,
- y_{j_1}, z_{k_1}, $y_{j_1} + 2$, $z_{k_1} + 2$, $y_{j_1} + y_{j_2}$, $z_{k_1} + z_{k_2}$, *and* $y_{j_1}z_{k_1}$, *when* $n \geq 2$ *and* $o \geq 2$,
- $y_{j_1} + z_{k_1} + 2$, $x_{i_1}(y_{j_1} + 1)$, $x_{i_1}(z_{k_1} + 1)$, $x_{i_1}x_{i_2} + x_{i_1} + x_{i_2}$, $x_{i_1}x_{i_2}y_{j_1} - x_{i_1} - x_{i_2}$, *and* $x_{i_1}x_{i_2}z_{k_1} - x_{i_1} - x_{i_2}$ *when* $m \geq 2$,
- $x_{i_1} + z_{k_1} + 2$, $y_{j_1}(x_{i_1} + 1)$, $y_{j_1}(z_{k_1} + 1)$, $y_{j_1}y_{j_2} + y_{j_1} + y_{j_2}$, $y_{j_1}y_{j_2}x_{i_1} - y_{j_1} - y_{j_2}$, *and* $y_{j_1}y_{j_2}z_{k_1} - y_{j_1} - y_{j_2}$ *when* $n \geq 2$,
- $x_{i_1} + y_{j_1} + 2$, $z_{k_1}(x_{i_1} + 1)$, $z_{k_1}(y_{j_1} + 1)$, $z_{k_1}z_{k_2} + z_{k_1} + z_{k_2}$, $z_{k_1}z_{k_2}x_{i_1} - z_{k_1} - z_{k_2}$, *and* $z_{k_1}z_{k_2}y_{j_1} - z_{k_1} - z_{k_2}$ *when* $o \geq 2$,
- 2 *when* $m \geq 2$, $n \geq 2$ *and* $o \geq 2$,
- $-2 - x_i - y_j - z_k + x_iy_jz_k$,

where $1 \leq i_1 < i_2 < i_3 \leq m$, $1 \leq j_1 < j_2 < j_3 \leq n$, *and* $1 \leq k_1 < k_2 < k_3 \leq o$.

PROOF. We will follow a similar proof to the proof given for Lemma 2.17. Let $\mathcal{I} = \{i_1, i_2, i_3\}$ with $1 \leq i_1 < i_2 < i_3 \leq m + n + o$ and $\mathcal{J} = \{j_1, j_2, j_3\}$ with $1 \leq j_1 < j_2 < j_3 \leq m + n + o$. Moreover, let $\mathcal{I}_1 = \mathcal{I} \cap [m]$, $\mathcal{I}_2 = \mathcal{I} \cap \{m+1, \ldots, m+n\}$, $\mathcal{I}_3 = \mathcal{I} \cap \{m+n+1, \ldots, m+n+o\}$, $\mathcal{J}_1 = \mathcal{J} \cap [m]$, $\mathcal{J}_2 = \mathcal{J} \cap \{m+1, \ldots, m+n\}$, $\mathcal{J}_3 = \mathcal{J} \cap \{m+n+1, \ldots, m+n+o\}$. Also, in the following $i'_t = i_t - m$, $i''_t = i_t - m - n$, $j'_t = j_t - m$ and $j''_t = j_t - m - n$, for $t \in [3]$.

Let $L = L_{m,n,o}[\mathcal{I};\mathcal{J}]$. First, in the same way that in the proof of Lemma 2.17 we can assume that L is non-singular. Several of the 3-minor of $L_{m,n,o}$ can be calculated using Lemma 2.17. For instance, if $\mathcal{I}_i = \mathcal{J}_i = \emptyset$ for some $i = 1,2,3$, then L is a submatrix of $L(K_{m,n}, \{X_{T_m}, Y_{T_n}\})$ and the corresponding 3-minor can be calculated using Lemma 2.17. Therefore we can assume that, if $\mathcal{I}_i = \emptyset$, then $\mathcal{J}_i \neq \emptyset$ for all $i = 1,2,3$. Moreover, if $\mathcal{I}_i = \emptyset$, then $|\mathcal{J}_i| = 1$ for all $i = 1,2,3$. Because otherwise either L will have two identical columns; a contradiction to the fact that L is non-singular. In a similar way, if $\mathcal{J}_i = \emptyset$, then $|\mathcal{I}_i| = 1$ for all $i = 1,2,3$. If $|\mathcal{I}_i| = 3$ for some $i = 1,2,3$, then L is a submatrix of the generalized Laplacian matrix of a complete bipartite graph. Therefore we can assume that $|\mathcal{I}_i| \leq 2$ and $|\mathcal{J}_i| \leq 2$ for all $i = 1,2,3$.

The first case that we need to consider is when $\mathcal{I}_i \neq \emptyset \neq \mathcal{J}_i$ for all $1 \leq i \leq 3$, that is, $|\mathcal{I}_i| = |\mathcal{J}_i| = 1$ for all $1 \leq i \leq 3$. In this case we have that

$$L = \begin{bmatrix} A & -1 & -1 \\ -1 & B & -1 \\ -1 & -1 & C \end{bmatrix},$$

where A is equal to 0 (when $m \geq 2$) or x_{i_1}, B is equal to 0 (when $n \geq 2$) or $y_{i'_2}$, and C is equal to 0 (when $n \geq 2$) or $z_{i''_3}$. Since $\det L = ABC - A - B - C - 2$ we get eight of the 3-minors of $L_{m,n,o}$. Since $|\mathcal{I}_i| \leq 2$ ($|\mathcal{J}_i| \leq 2$) for all $i = 1,2,3$, then there are no two $\mathcal{I}$'s ($\mathcal{J}$'s) empty. Therefore only remains the cases: when only one of the $\mathcal{I}$'s is empty and the case when one of the $\mathcal{I}$'s is empty and one of the $\mathcal{J}'s$ is empty.

Consider the case when only one of the sets $\mathcal{I}$'s is empty, that is, $|\mathcal{J}_i| = 1$ for all $i = 1,2,3$. Assume that $\mathcal{I}_3 = \emptyset$. Then we need to consider the following two matrices (when $|\mathcal{I}_1| = 1$ and $|\mathcal{I}_1| = 2$):

$$L_1 = \begin{bmatrix} A & -1 & -1 \\ -1 & 0 & -1 \\ -1 & B & -1 \end{bmatrix} \quad \text{and} \quad L_2 = \begin{bmatrix} A & -1 & -1 \\ 0 & -1 & -1 \\ -1 & B & -1 \end{bmatrix},$$

where A is equal to 0 (when $m \geq 2$ and $m \geq 3$, respectively) or $x_{i'_1}$ and B is equal to 0 (when $n \geq 3$ and $n \geq 2$, respectively) or $y_{i'_3}$. It is not difficult to see that $\det(L_1) = AB - B$ and $\det(L_2) = AB - A$. Thus, we get the minors $x_{i_1}y_{i'_3} - y_{i'_3}$ (when $n \geq 2$), $x_{i_1}y_{i'_3} - x_{i_1}$ (when $m \geq 2$), $-y_{i'_3}$ and $-x_{i_1}$ (when $m \geq 2$ and $n \geq 2$). We get similar 3-minors when $\mathcal{I}_2 = \emptyset$ or $\mathcal{I}_1 = \emptyset$.

Finally, consider the case when one of the $\mathcal{I}$'s is empty and one of the $\mathcal{J}'s$ is empty. Assume that $\mathcal{I}_3 = \emptyset$ and $\mathcal{J}_2 = \emptyset$. Then $|\mathcal{I}_2| = 1$ and L is equal to:

$$\begin{bmatrix} A & 0 & -1 \\ 0 & A' & -1 \\ -1 & -1 & -1 \end{bmatrix},$$

where A is equal to 0 or $x_{i'_1}$ and A' is equal to 0 or x_{i_2}. Clearly $\det L = -AA' - A - A'$. Thus we get the 3-minors $x_{i_1}x_{i_2} + x_{i_1} + x_{i_2}$ (when $m \geq 2$) and x_{i_1} and x_{i_2} (when $m \geq 3$). Similarly when $\mathcal{J}_1 = \emptyset$ and the other cases. □

Now the computation of the third critical ideal of the tripartite complete graph can easily done by using previous theorem. □

Proof of Theorem 2.16. Similarly to the proof of Theorem 2.15 we need to find the 3-minors of the generalized Laplacian matrix of $T_n \vee (K_m + K_o)$. We begin with $K_m \vee T_n$ and after that we do the same for $T_n \vee (K_m + K_o)$. We omit the proofs of Lemma 2.19 and Theorem 2.20

because are rutinary and both follows by using similar arguments to those in Lemma 2.17 and in Theorem 2.18, respectively.

LEMMA 2.19. *For $m, n \geq 1$, let $L'_{m,n}$ be the generalized Laplacian matrix of $K_m \vee T_n$. That is,*

$$L'_{m,n} = L(K_m \vee T_n, \{X_{K_n}, Y_{T_m}\}) = \begin{bmatrix} L(K_m, X_{K_m}) & -J_{m,n} \\ -J_{n,m} & L(T_n, Y_{T_n}) \end{bmatrix}.$$

Then the 3-minors (with positive leading coefficient) of $L'_{m,n}$ are the following:

- y_{j_1}, $y_{j_1}y_{j_2}$, *and* $y_{j_1}y_{j_2}y_{j_3}$ *when* $n \geq 3$,
- $x_{i_1}y_{j_1}y_{j_2} - y_{j_1} - y_{j_2}$ *when* $n \geq 2$,
- $x_{i_1}x_{i_2}y_{j_1} - x_{i_1} - x_{i_2} - y_{j_1} - 2$ *when* $m \geq 2$,
- y_{j_1} *when* $m \geq 2$ *and* $n \geq 3$,
- $x_{i_1} + 1$ *when* $m \geq 3$ *and* $n \geq 2$,
- $x_{i_1} + x_{i_2} + 2$, $x_{i_1} + y_{j_1}$, $x_{i_1}y_{j_1}y_{j_2}$ *and* $y_{j_1}y_{j_2} + y_{j_1} + y_{j_2}$ *when* $m \geq 2$ *and* $n \geq 2$,
- $(x_{i_1}+1)(x_{i_2}+1)$, $(x_{i_1}+1)(y_{i_1}+1)$, *and* $x_{i_1}x_{i_2}x_{i_3} - x_{i_1} - x_{i_2} - x_{i_3} - 2$ *when* $m \geq 3$,

where $1 \leq i_1 < i_2 < i_3 \leq m$ *and* $1 \leq j_1 < j_2 < j_3 \leq n$.

THEOREM 2.20. *For $m, n, o \geq 1$, let $L'_{m,n,o}$ be the generalized Laplacian matrix of $T_n \vee (K_m + K_o)$. That is,*

$$L'_{m,n,o} = L(T_n \vee (K_m + K_o), \{X_{K_m}, Y_{T_n}, Z_{K_o}\}) = \begin{bmatrix} L(K_m, X_{K_m}) & -J_{m,n} & \mathbf{0}_{m,o} \\ -J_{n,m} & L(T_n, Y_{T_n}) & -J_{n,o} \\ \mathbf{0}_{o,m} & -J_{o,n} & L(K_o, Z_{K_o}) \end{bmatrix}.$$

Then the 3-minors (with positive leading coefficient) of $L'_{m,n,o}$ are the following:

- $x_{i_1} + 1$ *when* $m \geq 3$ *and* ($o \geq 2$ *or* $n \geq 2$),
- $z_{k_1} + 1$ *when* $o \geq 3$ *and* ($m \geq 2$ *or* $n \geq 2$),
- y_{j_1} *when* $n \geq 3$ *and* ($m \geq 2$ *or* $o \geq 2$),
- 2 *when* $m \geq 2$, $n \geq 2$, *and* $o \geq 2$,
- y_{j_1}, $y_{j_1}y_{j_3}$, *and* $y_{j_1}y_{j_2}y_{j_3}$ *when* $n \geq 3$,
- $x_{i_1}y_{j_1}z_{k_1} - x_{i_1} - z_{k_1}$,
- $x_{i_1} + 1$, $z_{k_1}(x_{i_1} + 1)$, $(x_{i_1} + 1)(x_{i_2} + 1)$, $(x_{i_1} + 1)(y_{j_1} + 1)$, *and* $x_{i_1}x_{i_2}x_{i_3} - x_{i_1} - x_{i_2} - x_{i_3} - 2$, *when* $m \geq 3$,
- $z_{k_1} + 1$, $x_{i_1}(z_{k_1} + 1)$, $(z_{k_1} + 1)(z_{k_2} + 1)$, $(z_{k_1} + 1)(y_{j_1} + 1)$, *and* $z_{k_1}z_{k_2}z_{k_3} - z_{k_1} - z_{k_2} - z_{k_3} - 2$ *when* $o \geq 3$,
- $x_{i_1} + z_{k_1}$, $y_{j_1} + y_{j_2}$, $x_{i_1}y_{j_1}$, $y_{j_1}z_{k_1}$, $x_{i_1}y_{j_1}y_{j_2} - y_{j_1} - y_{j_2}$, *and* $y_{j_1}y_{j_2}z_{k_1} - y_{j_1} - y_{j_2}$ *when* $n \geq 2$,
- $x_{i_1} + 1$, $x_{i_1}x_{i_2} - 1$, $y_{j_1}z_{k_1} + z_{k_1} - 1$, $z_{k_1}(x_{i_1} + 1)$, $x_{i_1}x_{i_2}z_{k_1} - z_{k_1}$, *and* $x_{i_1}x_{i_2}y_{j_1} - x_{i_1} - x_{i_2} - y_{j_1} - 2$, *when* $m \geq 2$,
- $z_{k_1} + 1$, $x_{i_1}(z_{k_1}z_{k_2} - 1)$, $z_{k_1}z_{k_2} - 1$, $x_{i_1}y_{j_1} + x_{i_1} - 1$, $x_{i_1}(z_{k_1} + 1)$, *and* $z_{k_1}z_{k_2}y_{j_1} - z_{k_1} - z_{k_2} - y_{j_1} - 2$, *when* $o \geq 2$,
- $x_{i_1} + 1$, $y_{j_1} + 2$, $z_{k_1} + 1$, $x_{i_1}x_{i_2} - 1$, *and* $z_{k_1}z_{k_2} - 1$ *when* $m \geq 2$ *and* $o \geq 2$,
- $x_{i_1} + 1$, y_{j_1}, $z_{k_1} - 1$, $x_{i_1} + y_{j_1}$, $x_{i_1} + x_{i_2} + 2$, $y_{j_1}(x_{i_1} + 1)$, *and* $y_{j_1}y_{j_2} + y_{j_1} + y_{j_2}$ *when* $m \geq 2$ *and* $n \geq 2$,
- $x_{i_1} - 1$, y_{j_1}, $z_{k_1} + 1$, $z_{k_1} + y_{j_1}$, $z_{k_1} + z_{k_2} + 2$, $y_{j_1}(z_{k_1} + 1)$, *and* $y_{j_1}y_{j_2} + y_{j_1} + y_{j_2}$, *when* $n \geq 2$ *and* $o \geq 2$,

where $1 \leq i_1 < i_2 < i_3 \leq m$, $1 \leq j_1 < j_2 < j_3 \leq n$, *and* $1 \leq k_1 < k_2 < k_3 \leq o$. □

Theorems 2.15 and 2.16 implies that $\mathbf{Forb}(\Gamma_{\leq 2}) = \mathcal{F}_2$. Now, we present the non-connected version of Theorem 2.10.

COROLLARY 2.21. *A simple graph has algebraic co-rank equal to two if and only if it is the disjoint union of a trivial graph with one of the following graphs:*

- $K_{m,n,o}$, *where* $m \geq 2$, $n + o \geq 1$,
- $T_n \vee (K_m + K_o)$, *where* $m, o \geq 2$, $m, n, o \geq 1$, *or* $n \geq 2$ *and* $m + o \geq 1$.

PROOF. It is not difficult to see that in the non-connected case we need to add the graphs $P_3 + P_2$ and $3P_2$ to the set of forbidden graphs. The rest follows directly from Theorem 2.10. □

We finish this section with the classification of the graphs having critical group with 2 invariant factors equal to one.

THEOREM 2.22. *The critical group of a connected simple graph has exactly two invariant factor equal to 1 if and only if it is one of the following graphs:*

- $K_{m,n,o}$, *where* $m \geq n \geq o$ *satisfy one of the following conditions:*
  - $m, n, o \geq 2$ *with the same parity,*
  - $m, n \geq 3$, $o = 1$, *and* $\gcd(m+1, n+1) \neq 1$,
  - $m \geq 2$, $n = o = 1$,
  - $m, n \geq 2$, $o = 0$ *and* $\gcd(m, n) \neq 1$,
  - $m \geq 2$, $n = 2$, *and* $o = 0$, *or*
  - $m = 2$ *and* $n = 1$.
- $T_n \vee (K_m + K_o)$, *where* $m \geq o$ *and* n *satisfy one of the following conditions:*
  - $m, n, o \geq 2$ *with the same parity,*
  - $m, o \geq 2$, $n = 1$, *and* $\gcd(m+1, o+1) \neq 1$,
  - $m, n \geq 2$, $o = 1$, *and* $\gcd(m+1, n-1) \neq 1$,
  - $m \geq 1$, $n = o = 1$,
  - $n \geq 1$, $m = o = 1$,
  - $m, n \geq 3$, $o = 0$, *and* $\gcd(m, n) \neq 1$,
  - $m \geq 2$, $n = 2$, $o = 0$, *or*
  - $m = 2$, $n \geq 2$, $o = 0$.

PROOF. It turns out from Theorems 2.15 and 2.16. □

4. The set $\mathbf{Forb}(\Gamma_{\leq k})$.

The characterization of the γ-critical graphs with a given algebraic co-rank, $\mathbf{Forb}(\Gamma_{\leq k})$, is very important. For instance, we were able to characterize $\Gamma_{\leq k}$ for k equal to 1 and 2 because we got a finite set of γ-critical graphs with algebraic co-rank equal to $k+1$ (for k equal to 1 and 2), and after that we proved that all the graphs that do not contain a graph from this set as an induced subgraph have algebraic co-rank less than or equal to k. In this section we give two infinite families of forbidden simple graphs. This will prove that $\mathbf{Forb}(\Gamma_{\leq k})$ is not empty for all $k \geq 0$. Moreover, we conjecture that $\mathbf{Forb}(\Gamma_{\leq k})$ is finite for all $k \geq 0$. To finish we present an example of a simple graph G with algebraic co-rank equal to 5 but with no 5-minor equal to 1. That is, the 1 can be obtained uniquely from a non trivial algebraic combination of 5-minors of $L(G, X)$.

We begin by proving that the path with $n+2$ vertices is γ-critical with algebraic co-rank equal to $n+1$.

THEOREM 2.23. *If* $n \geq 0$, *then* $P_{n+2} \in \mathbf{Forb}(\Gamma_{\leq n})$.

PROOF. It is not difficult to prove $\gamma(P_{n+2}) = n+1$, see Corollary 4.10 of [**22**]. On the other hand, if $H = P_{n+2} \setminus v$ for some $v \in V(P_{n+2})$, then H is a disjoint union of at most two paths. Let $H = P_{n_1} + \cdots + P_{n_s}$ with $1 \leq s \leq 2$ and $\sum_{i=}^{s} n_i = n+1$, then by lemma 2.8 we get that

$$\gamma(H) = \sum_{i=1}^{s} \gamma(P_{n_i}) = \sum_{i=1}^{s} (n_i - 1) = \sum_{i=1}^{s} n_i - s = n + 1 - s < n + 1.$$

Therefore $P_{n+2} \in \mathbf{Forb}(\Gamma_{\leq n})$. □

Now, we present another infinite family of graph that are γ-critical. Let K_n be the complete graph with n vertices and M_k a matching of K_n with k edges. We begin by finding the critical group of $K_n \setminus M_k$.

PROPOSITION 2.24. *If K_n is the complete graph with n vertices and M_k is a matching of k edges, then*

$$K(K_n \setminus M_k) \cong \begin{cases} \mathbb{Z}_n^{n-2k-2} \oplus \mathbb{Z}_{n(n-2)}^k & \text{if } n \geq 2k+2, \\ \mathbb{Z}_{n-2} \oplus \mathbb{Z}_{n(n-2)}^{k-1} & \text{if } n = 2k+1. \end{cases}$$

PROOF. If $n = 2k+1$, then the result follows by [**31**, Theorem 1]. Therefore, we assume that $n \geq 2k+2$. Given $\mathbf{a} \in \mathbb{Z}^k$, let $N_{k+1}(\mathbf{a})$ be the matrix given by

$$\begin{bmatrix} 1 & \mathbf{a} \\ \mathbf{0}^t & I_k \end{bmatrix}.$$

If $M_k = \{v_1v_2, \ldots, v_{2k-1}v_{2k}\}$, then

$$L(K_n \setminus M_k, v_n) = \begin{bmatrix} [(n-2)I_2 + J_2] \otimes I_k - J_{2k} & -J_{2k,n-2k-1} \\ -J_{n-2k-1,2k} & nI_{n-2k-1} - J_{n-2k-1} \end{bmatrix},$$

where $\otimes$ is the tensor product of matrices. Now, since $\det(N_{n-1}(\mathbf{a})) = 1$ for all $\mathbf{a}$, then

$$\begin{aligned} L(K_n \setminus M_k, v_n) &\sim N_{n-1}(\mathbf{1})^t N_{n-1}(\mathbf{1}) L(K_n \setminus M_k, v_n) N_{n-1}(-\mathbf{1}) \\ &= I_1 \oplus nI_{n-2k-2} \bigoplus_{i=1}^{k} \begin{bmatrix} n-1 & 1 \\ 1 & n-1 \end{bmatrix}. \end{aligned}$$

On the other hand,

$$\begin{aligned} \begin{bmatrix} n-1 & 1 \\ 1 & n-1 \end{bmatrix} &\sim \begin{bmatrix} 0 & 1 \\ -1 & n-1 \end{bmatrix} \begin{bmatrix} n-1 & 1 \\ 1 & n-1 \end{bmatrix} \begin{bmatrix} 1 & -(n-1) \\ 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & n(n-2) \end{bmatrix}. \end{aligned}$$

Therefore, $L(K_n \setminus M_k, v_n) \sim I_{k+1} \oplus nI_{n-2k-2} \oplus n(n-2)I_k$. □

COROLLARY 2.25. *If $n = 2k+2$, then $K_n \setminus M_k \in \mathbf{Forb}(\Gamma_{\leq k})$.*

PROOF. By Proposition 2.24 we have that

$$\gamma(K_n \setminus M_k) \leq \begin{cases} k+1 & \text{if } n \geq 2k+2, \\ k & \text{if } n = 2k+1. \end{cases}$$

Now let $n \geq 2k+2$, $M_k = \{v_1v_2, \ldots, v_{2k-1}v_{2k}\}$, and $M = L(K_n \setminus M_k, X)[\{1, \ldots, 2k+1\}, \{2, \ldots, 2k+2\}]$ be a square submatrix of generalized Laplacian matrix of $K_n \setminus M_k$. Then

$$M = \begin{bmatrix} 0 & -1 & & -1 & -1 \\ -1 & 0 & & -1 & -1 \\ & & \ddots & -1 & -1 \\ -1 & -1 & -1 & 0 & -1 \\ -1 & -1 & -1 & -1 & -1 \end{bmatrix}.$$

By [**22**, Theorem 3.13], $\det(M) = \det(L(K_k, X_{K_k}))|_{\{x_1=0,\ldots,x_{k-1}=0,x_k=-1\}} \overset{3.13}{=} -1$ and thus $\gamma(K_n \setminus M_k) = k+1$ for all $n \geq 2k+2$. Finally, if $n = 2k+2$ and $v \in V(K_n \setminus M_k)$, then $(K_n \setminus M_k) \setminus v$ is equal to $K_{n-1} \setminus M_k$ or $K_{n-1} \setminus M_{k-1}$. Hence, $\gamma((K_n \setminus M_k) \setminus v) \leq k$ and therefore $K_n \setminus M_k \in \mathbf{Forb}(\Gamma_{\leq k})$. □

This result proves that $\mathbf{Forb}(\Gamma_{\leq k})$ is not empty for all $k \geq 0$.

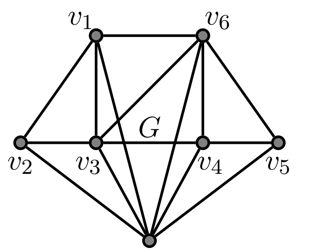
COROLLARY 2.26. *If $k \geq 0$, then* $\mathbf{Forb}(\Gamma_{\leq k})$ *is not empty.*

For $i \geq 3$, the set $\mathbf{Forb}(\Gamma_{\leq i})$ is more complex than $\mathbf{Forb}(\Gamma_{\leq 1})$ and $\mathbf{Forb}(\Gamma_{\leq 2})$. For instance, we will see that $\mathbf{Forb}(\Gamma_{\leq 3})$ has at least 49 graphs. Moreover, we conjecture that $\mathbf{Forb}(\Gamma_{\leq k})$ is finite.

CONJECTURE 2.27. *For all $k \in \mathbb{N}$ the set* $\mathbf{Forb}(\Gamma_{\leq k})$ *is finite.*

Until now, all the graphs that were presented had algebraic co-rank equal to k because its generalized Laplacian matrix has a k-minor equal to one. Next example shows a graph G with $\gamma_{\mathbb{Z}}(G) = 5$ having no a 5-minor equal to 1.

EXAMPLE 2.28. *Let G be the graph on Figure 14 and* $f_1 = \det(L(G,X)[\{1,2,3,4,5\};\{2,3,5,6,7\}]) =$



$$L(G,X) = \begin{bmatrix} x_1 & -1 & -1 & 0 & 0 & -1 & -1 \\ -1 & x_2 & -1 & 0 & 0 & 0 & -1 \\ -1 & -1 & x_3 & -1 & 0 & -1 & -1 \\ 0 & 0 & -1 & x_4 & -1 & -1 & -1 \\ 0 & 0 & 0 & -1 & x_5 & -1 & -1 \\ -1 & 0 & -1 & -1 & -1 & x_6 & -1 \\ -1 & -1 & -1 & -1 & -1 & -1 & x_7 \end{bmatrix}$$

FIGURE 4. A graph G with seven vertices and its generalized Laplacian matrix.

$x_2 + x_5 + x_2x_5$, *and* $f_2 = \det(L(G,X)[\{1,2,3,5,6\};\{2,4,5,6,7\}]) = -(1 + x_2 + x_5 + x_2x_5)$. *Then* $\langle f_1, f_2 \rangle = 1$ *and therefore* $\gamma_{\mathbb{Z}}(G) = 5$. *However, it is not difficult to check that $L(G,X)$ has no 5-minor is equal to one.*

5. Cliques, stable sets and critical ideals

Let $G = (V,E)$ a simple graph. If E' is a subset of E, the *edge-induced subgraph* $G[E']$ is the subgraph of G whose edge set is E' and whose vertex set consists of all ends of edges of E'. Let P, Q be two subsets of V, we denote by $E(P,Q)$ the set of edges of G with one end in P and the other end in Q. A *clique* of G is a subset S of V of mutually adjacent vertices, and the maximum size of a clique of G is the *clique number* $\omega(G)$ of G. A subset S of V is called an *independent set*, or *stable set*, of G if no two vertices of S are adjacent in G. The cardinality of a maximum stable set in G is called the *stability number* of G and is denoted by $\alpha(G)$.

DEFINITION 2.29. *Given a simple graph $G = (V,E)$ and a vector $\mathbf{d} \in \mathbb{Z}^V$, the graph $G^{\mathbf{d}}$ is constructed as follows. For each vertex $u \in V$ is associated a new vertex set V_u, where V_u is a clique of cardinality $-\mathbf{d}_u$ if $\mathbf{d}_u$ is negative, and V_u is a stable set of cardinality $\mathbf{d}_u$ if $\mathbf{d}_u$ is positive. And each vertex in V_u is adjacent with each vertex in V_v if and only if u and v are adjacent in G. Then the graph G is called the underlying graph of $G^{\mathbf{d}}$.*

A convenient way to visualize $G^{\mathbf{d}}$ is by means of a drawing of G, where the vertex u is colored in black if $\mathbf{d}_u$ is negative, and colored in white if $\mathbf{d}_u$ is positive. We indicate the cardinality of V_u by writing it inside the drawing of vertex u. When $|\mathbf{d}_u| = 1$, we may color u in gray and avoid

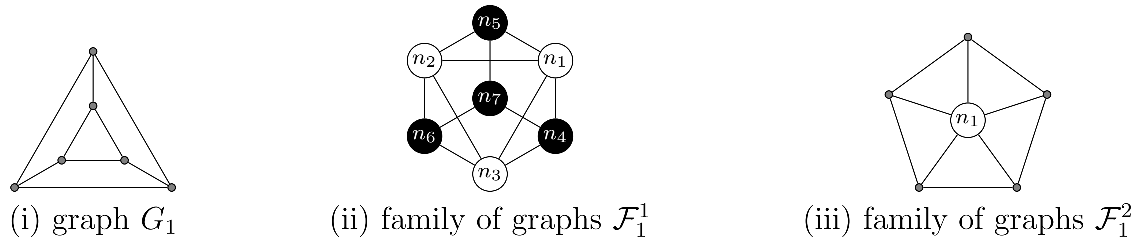


FIGURE 5. The family of graphs $\mathcal{F}_1$. A black vertex represents a clique of cardinality n_v, a white vertex represents a stable set of cardinality n_v and a gray vertex represents a single vertex. Where $n_v \geq 0$

writing the cardinality (see Figure 5). On the other hand, it will be useful to avoid writing the cardinality when $|\mathbf{d}_u| = 2$ (see Figure 6).

In general, the computation of the Gröbner bases of the critical ideals is more than complicated. However, in the rest of this section we will show a novel method, developed in [**7**], to decide for $i \leq |V(G)|$ whether the *i-th* critical ideal of $G^{\mathbf{d}}$ is trivial or not.

For $V' \subseteq V(G)$ and $\mathbf{d} \in \mathbb{Z}^{V'}$, we define $\phi(\mathbf{d})$ as follows:

$$\phi(\mathbf{d})_v = \begin{cases} 0 & \text{if } \mathbf{d}_v > 0, \\ -1 & \text{if } \mathbf{d}_v < 0. \end{cases}$$

A simpler version of Theorem 3.9 is restated as follows.

THEOREM 2.30. *Let $n \geq 2$ and G be a graph with n vertices. For $V' \subseteq V(G)$, $1 \leq j \leq n$ and $\mathbf{d} \in \mathbb{Z}^{V'}$, the critical ideal $I_j(G^{\mathbf{d}}, X_{G^{\mathbf{d}}})$ is trivial if and only if the evaluation of $I_j(G, X_G)$ at $X_G = \phi(\mathbf{d})$ is trivial.*

Thus the procedure of verifying whether a family of graphs belongs to $\Gamma_{\leq i}$ becomes in evaluation of the *i-th* critical ideal of the underlying graph of the family.

Let G_2 be the underlying graph of the family of graphs $\mathcal{F}_1^1$ (see Figure 5.ii) with vertex set $V = \{v_1, v_2, v_3, v_4, v_5, v_6, v_7\}$. Let $\mathbf{d} \in \mathbb{Z}^V$ such that $\mathbf{d}_1, \mathbf{d}_2, \mathbf{d}_3$ are positive integers and $\mathbf{d}_4, \mathbf{d}_5, \mathbf{d}_6, \mathbf{d}_7$ are negative integers. Thus $\phi(\mathbf{d}) = (0, 0, 0, -1, -1, -1, -1)$. By using a computer algebra system we can check that

$$I_4(G_2, X_{G_2}) = \langle 2, x_1, x_2, x_3, x_4 + 1, x_5 + 1, x_6 + 1, x_7 + 1 \rangle.$$

Since the evaluation $I_4(G_2, X_{G_2})$ at $X_{G_2} = \phi(\mathbf{d})$ is equal to $\langle 2 \rangle$, then by Theorem 2.30 the critical ideal $I_4(G_2^{\mathbf{d}}, X_{G_2^{\mathbf{d}}})$ is non-trivial. Therefore, each graph in this family of graphs has algebraic co-rank at most 3. Let G_3 be the underlying graph of the family of graphs $\mathcal{F}_1^2$ (see Figure 5.iii) with vertex set $V = \{v_1, v_2, v_3, v_4, v_5, v_6\}$. Let $\mathbf{d}_1$ be positive integer. By using a computer algebra system we can check that

$$I_4(G_3, X_{G_3}) = \langle x_1^2 + 5x_1 + 5, x_1 + x_2 + 3, x_1 + x_3 + 3, x_1 + x_4 + 3, x_1 + x_5 + 3, x_1 + x_6 + 3 \rangle.$$

Since the evaluation of $I_4(G_3, X_{G_3})$ at $x_1 = 0$ is non-trivial, then by Theorem 2.30 the critical ideal $I_4(G_3^{\mathbf{d}}, X_{G_3^{\mathbf{d}}})$ is non-trivial. Therefore, each graph in $\mathcal{F}_1^2$ has algebraic co-rank at most 3.

On the other hand, it can be verified that $\gamma(G_1)$, $\gamma(G_2)$ and $\gamma(G_3)$ are equal to 3. Then the graphs in $\mathcal{F}_1$ have algebraic co-rank 3, and so each induced subgraph of a graph in $\mathcal{F}_1$ has algebraic co-rank at the most 3.

PROPOSITION 2.31. *Each graph in $\mathcal{F}_1$ (see Figure 5) belongs to $\Gamma_{\leq 3}$.*

6. A description of $\Gamma_{\leq 3}$

It is possible to compute the algebraic co-rank of all connected graphs with at most 9 vertices using the software Macaulay2 [**28**] and Nauty [**36**]. The computation of the algebraic co-rank of the connected graphs with at most 8 vertices required at most 3 hours on a MacBookPro with a 2.8 GHz Intel i7 quad core processor and 16 GB RAM. Besides, the computation of the algebraic co-rank of the connected graphs with 9 vertices required 4 weeks of computation on the same computer.

Let $\mathcal{F}_3$ be the family of graphs shown in Figure 6. This family represents the graphs in $\mathbf{Forb}(\Gamma_{\leq 3})$ with at most 8 vertices. Since there exists no minimal forbidden graph with 9 vertices for $\Gamma_{\leq 3}$, then it is likely that $\mathcal{F}_3 = \mathbf{Forb}(\Gamma_{\leq 3})$.

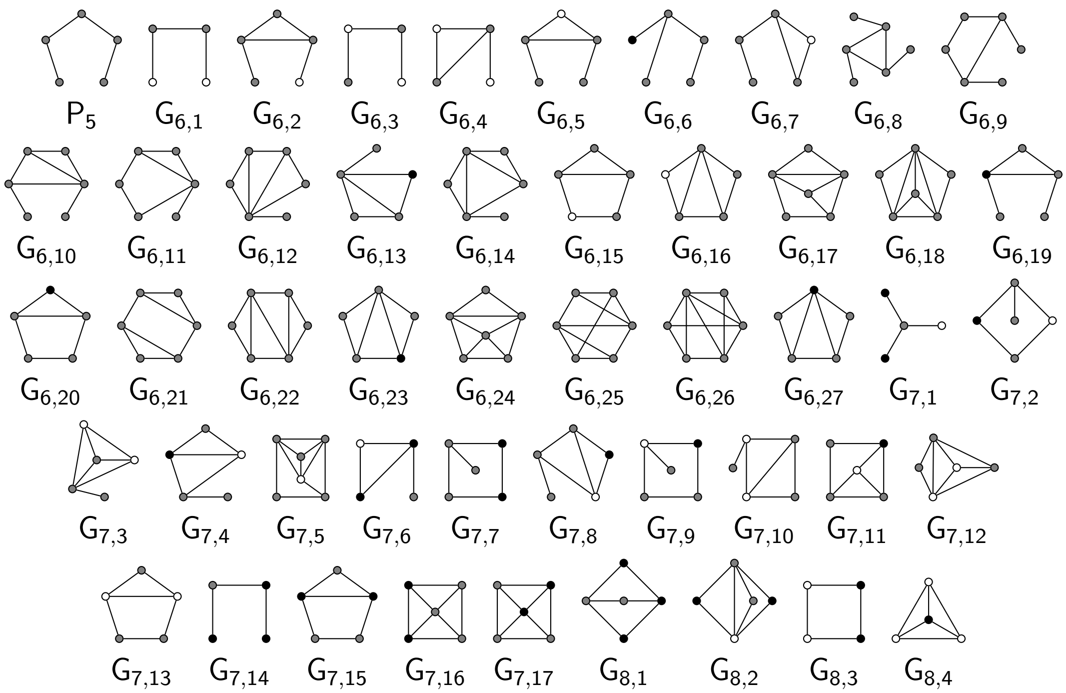


FIGURE 6. The family of graphs $\mathcal{F}_3$. A black vertex represents a clique of cardinality 2, a white vertex represents a stable set of cardinality 2 and a gray vertex represent a single vertex.

PROPOSITION 2.32. *Each graph in $\mathcal{F}_3$ belongs to* $\mathbf{Forb}(\Gamma_{\leq 3})$.

PROOF. It can easily checked, using a computer algebra system, that each graph in $\mathcal{F}_3$ is γ-critical and has algebraic co-rank equal to 4. $\square$

One of the main results of this thesis is the following:

THEOREM 2.33. *If a graph $G \in \Gamma_{\leq 3}$ has clique number at most 3, then G is an induced subgraph of a graph in $\mathcal{F}_1$.*

We divide the proof in two characterizations: the graphs in $\Gamma_{\leq 3}$ with clique number equal to 2 and 3. The converse of Theorem 2.33 is stronger, by Proposition 2.31 we have that each induced subgraph of a graph in $\mathcal{F}_1$ belongs to $\Gamma_{\leq 3}$. However, it is not difficult to recognize the graphs in $\mathcal{F}_1^1$ with clique number greater or equal than 4.

THEOREM 2.34. *Let G be a simple connected graph with $\omega(G) = 2$. Then, G is $\mathcal{F}_3$-free if and only if G is isomorphic to an induced subgraph of a graph in $\mathcal{F}_2$ (see Figure 7).*

(i) K_{n_1,n_2} *(ii)* $\mathcal{F}_2^1$ *(iii)* $\mathcal{F}_2^2$ *(iv)* $\mathcal{F}_2^3$

FIGURE 7. The family of graphs $\mathcal{F}_2$. A white vertex represents a stable set of cardinality n_v and a gray vertex represents a single vertex.

Since each graph in $\mathcal{F}_2$ is isomorphic to an induced subgraph of a graph in $\mathcal{F}_1^1$, then Proposition 2.31 implies that each graph in $\mathcal{F}_2$ belongs to $\Gamma_{\leq 3}$. Note that the graphs in $\Gamma_{\leq 1}$ and $\Gamma_{\leq 2}$ are induced subgraph of a graph in $\mathcal{F}_1^1$ (see Figure 5.ii).

THEOREM 2.35. *Let G be a simple connected graph with $\omega(G) = 3$. Then, G is $\mathcal{F}_3$-free if and only if G is isomorphic to an induced subgraph of a graph in $\mathcal{F}_1$ with clique number 3.*

Section 7 is devoted to the proof of the Theorem 2.35. Now we give the proof of Theorem 2.34.

PROOF OF THEOREM 2.34. Since each graph in $\mathcal{F}_2$ belongs to $\Gamma_{\leq 3}$, then each graph in $\mathcal{F}_2$ is $\mathcal{F}_3$-free. Thus, we get one implication.

Suppose G is $\mathcal{F}_3$-free. Let $a, b \in V(G)$ such that $ab \in E(G)$. Since $\omega(G) = 2$, then there is no vertex adjacent with a and b at the same time. For a vertex $v \in V(G)$, the neighbor set $N_G(v)$ of v in G is the set of all vertices adjacent with v. Let $\mathcal{A} = N_G(a) - b$ and $\mathcal{B} = N_G(b) - a$. Clearly, each of the sets $\mathcal{A}$ and $\mathcal{B}$ induces a trivial graph. Let us define

$$A = \{u \in \mathcal{A} : \exists v \in \mathcal{B} \text{ such that } uv \in E\}, \quad A' = \{u \in \mathcal{A} : \nexists v \in \mathcal{B} \text{ such that } uv \in E\}$$

$$B = \{u \in \mathcal{B} : \exists v \in \mathcal{A} \text{ such that } uv \in E\}, \text{ and } B' = \{u \in \mathcal{B} : \nexists v \in \mathcal{A} \text{ such that } uv \in E\}.$$

Thus we have two possible cases: when A and B are not empty and when A and B are empty.

First we consider when A and B are not empty. In this case we have the following statements:

CLAIM 2.36. *One of the sets A' or B' is empty, and the other has cardinality at most one.*

PROOF. Suppose A' and B' are not empty. Let $u \in A$, $v \in B$, $s \in A'$ and $t \in B'$. Then the vertex set $\{a, b, u, v, s, t\}$ induces a graph isomorphic to $\mathsf{G}_{6,9}$, which is impossible. Now suppose A' has cardinality more than 1. Take $w_1, w_2 \in A'$. Then $\{x, y, u, v, w_1, w_2\}$ induces a graph isomorphic to $\mathsf{G}_{6,3}$; a contradiction. Thus A' has cardinality at most one. $\square$

CLAIM 2.37. *The edge set $E(A, B)$ induces either a complete bipartite graph or a complete bipartite graph minus an edge.*

PROOF. First note that each vertex in A is incident with each vertex in B, except for at most one vertex. It is because if $u \in A$ and $v_1, v_2, v_3 \in B$ such that $uv_1 \in E(G)$ and $uv_2, uv_3 \notin E(G)$, then the induced subgraph $G[\{x, y, u, v_1, v_2, v_3\}]$ is isomorphic to $\mathsf{G}_{6,3}$; which is impossible. In a similar way, each vertex in B is incident with each vertex in A, except for at most one vertex. Thus the edge set $E(A, B)$ must be equal to the edge set $\{uv : u \in A, v \in B\}$ minus a matching. In fact, the cardinality of this matching must be at most one. Otherwise, if $u, v \in A$ and $s, t \in B$ such that $us, vt \notin E(G)$ and $ut, vs \in E(G)$, then the induced subgraph $G[\{t, u, a, v, s\}]$ is isomorphic to P_5; which is a contradiction. □

CLAIM 2.38. *It is not possible that, at the same time, the edge set $E(A, B)$ induces a complete bipartite graph minus an edge and $A' \cup B' \neq \emptyset$.*

PROOF. Suppose both situations occur at the same time. Let $u_1, u_2 \in A$ and $v_1, v_2 \in B$ such that $u_1v_1, u_1v_2, u_2v_1 \in E(G)$ and $u_2v_2 \notin E(G)$. And let $w \in A'$. Then the vertex set $\{u_1, u_2, v_1, v_2, y, w\}$ induces a graph isomorphic to $\mathsf{G}_{6,9}$; which is a contradiction. □

Thus there are three possible cases:
(a) $E(A, B) = \{uv : u \in A, v \in B\}$ and $A' \cup B' = \emptyset$,
(b) $E(A, B) = \{uv : u \in A, v \in B\}$ and $A' = T_1$, $B' = \emptyset$ or
(c) $E(A, B)$ induces a bipartite complete graph minus an edge and $A' \cup B' = \emptyset$.

Let $V_\emptyset$ denote the set of vertices that are not adjacent with a nor b. In what follows we will describe the vertex set $V_\emptyset$.
Case (a). Let $w_1, w_2 \in V_\emptyset$, $u_1, u_2 \in A$, and $v_1 \in B$. Note that it is not possible that a vertex in $V_\emptyset$ is adjacent with a vertex in A and a vertex in B at the same time, because then we get $\omega(G) \geq 3$. Moreover

CLAIM 2.39. *There exist no two vertices in $V_\emptyset$ such that one is adjacent with a vertex in A and the other one is adjacent with a vertex in B.*

PROOF. Suppose $w_1u_1, w_2v_1 \in E(G)$. There are two cases: either $w_1w_2 \in E(G)$ or $w_1w_2 \notin E(G)$. If $w_1w_2 \in E(G)$, then $G[\{w_1, w_2, v_1, b, a\}] \simeq \mathsf{P}_5$; which is impossible. And if $w_1w_2 \notin E(G)$, then $G[\{w_1, w_2, v_1, u_1, b, a\}] \simeq \mathsf{G}_{6,9}$; which is a contradiction. □

Without loss of generality, suppose $w_1 \in V_\emptyset$ is adjacent with $u_1 \in A$. Thus

CLAIM 2.40. *The vertex set $V_\emptyset$ has cardinality at most 1.*

PROOF. There are two possible cases: either each vertex in $V_\emptyset$ is adjacent with a common vertex in A or not. Suppose there exists $w_2 \in V_\emptyset$ such that u_1 is adjacent with both w_1 and w_2. Then $w_1w_2 \notin E(G)$, because otherwise w_1, w_2 and u_1 induce a K_3. Thus the vertex set $\{w_1, w_2, u_1, v_1, a, b\}$ induces a graph isomorphic to $\mathsf{G}_{6,3}$; which is forbidden. Then this case is not possible. Now suppose w_1 and w_2 are not adjacent with a common vertex in A. We have the following possible cases:

(1) $w_1u_1, w_2u_2 \in E(G)$,
(2) $w_1u_1, w_2u_2, w_1w_2 \in E(G)$, or
(3) $w_1w_2 \in E(G)$.

This yields a contradiction since in case (1) the vertex set $\{w_1, u_1, v_1, u_2, w_2\}$ induces a graph isomorphic to P_5, in case (2) the vertex set $\{w_2, w_1, u_1, v_1, b\}$ induces a graph isomorphic to P_5, and in case (3) the vertex set $\{w_2, w_1, u_1, v_1, b\}$ induces a graph isomorphic to P_5. □

If $|A| = 2$, there are two possibilities: either w_1 is adjacent with each vertex in A or w_1 is adjacent with only one vertex of A. If $|A| \geq 3$, then w_1 is adjacent with either all vertex in A or only one vertex in A. It is because if w_1 is adjacent with both $u_1, u_2 \in A$ and w is not adjacent with $u_3 \in A$, then the vertex set $\{w_1, u_1, u_2, u_3, a, b\}$ induces a graph isomorphic to $\mathsf{G}_{6,3}$. Note that when w_1 is adjacent with each vertex in A, then the graph is isomorphic to an induced subgraph of a graph in $\mathcal{F}_2^2$. Meanwhile, when w_1 is adjacent only with one vertex of A, then the graph is isomorphic to an induced subgraph of a graph in $\mathcal{F}_2^1$. Finally, when $V_\emptyset = \emptyset$, the graph is a complete bipartite graph.

Case (b). Suppose without loss of generality that $A' \neq \emptyset$. By Claims 2.39 and 2.40, the vertex set $V_\emptyset$ has cardinality at most one. Let $w \in V_\emptyset$, $u_1 \in A$, $u_2 \in A'$ and $v_1 \in B$. We have two cases: w is adjacent with either a vertex in $\mathcal{A}$ or a vertex in $\mathcal{B}$. Let us consider when w is adjacent with a vertex in $\mathcal{A}$. Here we have two possibilities: either $wu_2 \in E(G)$ or $wu_2 \notin E(G)$. However, none of the two cases is allowed, since in the first case we get that the vertex set $\{w, u_2, a, b, v_1\}$ induces a graph isomorphic to P_5, and in the second case the vertex set $\{w, u_2, a, b, u_1, v_1\}$ induces a graph isomorphic to $\mathsf{G}_{6,9}$. Thus the remaining case is that w is adjacent with a vertex in $\mathcal{B}$. In this case w must be adjacent with u_2 and each vertex in $\mathcal{B}$, because otherwise the graph P_5 appears as induced subgraph. Note that this graph is isomorphic to an induced subgraph of a graph in $\mathcal{F}_2^3$.

Case(c). Let $w \in V_\emptyset$, $u_1, u_2 \in A$, and $v_1, v_2 \in B$ such that $u_1v_2 \notin E(G)$ and $u_1v_1, u_2v_1, u_2v_2 \in E(G)$. Note that if w is adjacent with u_1 or v_2, then w is adjacent with both u_1 and v_2, because since $u_1v_2 \notin E(G)$, we would obtain P_5 as induced subgraph; which is not possible. In this case, when w is adjacent with u_1 and v_2, the graph is isomorphic to an induced subgraph of $\mathcal{F}_2^3$. Now we consider when w is adjacent with a vertex in $(A - u_1) \cup (B - v_2)$. Without loss of generality, we can suppose w is adjacent with u_2. The vertex w is not adjacent with v_1 or v_2, because otherwise a clique of cardinality 3 is obtained. On the other hand, w is not adjacent with u_1, because otherwise w would be adjacent with both vertices u_1 and v_1. Thus w is adjacent only with u_2, but the vertex set $\{w, u_1, u_2, v_2, a, b, \}$ induces a graph isomorphic to $\mathsf{G}_{6,9}$; a contradiction. Thus $V_\emptyset$ is empty, and the graph is isomorphic to an induced subgraph of $\mathcal{F}_2^2$.

Now we consider the case when A and B are empty. One of the vertex sets A' or B' has cardinality at most one, because otherwise $A' \cup B' \cup \{a, b\}$ would contain $\mathsf{G}_{6,1}$ as induced subgraph. Thus, let us assume that $A' = \{u\}$ and $|B'| > 1$. Let $V_\emptyset$ denote the vertex set whose vertices are not adjacent with both a and b.

CLAIM 2.41. *If A' and B' are not empty, and $w \in V_\emptyset$ is adjacent with a vertex in $A' \cup B'$, then w is adjacent with each vertex in $A' \cup B'$.*

PROOF. Let $v \in B'$. Suppose one of the edges uw or vw does not exist. Then $\{w, u, a, b, v\}$ induces a graph isomorphic to P_5; which is a contradiction. $\square$

Note that the vertex set $V_\emptyset$ induces a stable set, because otherwise $\omega(G) > 2$. Thus when A' and B' are not empty, the graph G is isomorphic to an induced subgraph of a graph in $\mathcal{F}_2^3$.

Now consider the case when $B = \emptyset$ and $|A| > 1$.

CLAIM 2.42. *Each vertex $w \in V_\emptyset$ is adjacent with either a unique vertex in A or each vertex in A.*

PROOF. The result is easy to check when $|A| \leq 2$. So suppose A has cardinality greater or equal than 3. Let $u_1, u_2, u_3 \in A$ such that w is adjacent with both u_1, u_2, and w is not adjacent with u_3. Since the vertex set $\{w, u_1, u_2, u_3, a, b\}$ induces a graph isomorphic to $\mathsf{G}_{6,3}$, then we get a contradiction and the result follows. $\square$

Claim 2.43. *Let $w \in V_\emptyset$ such that it is adjacent with $u \in A$. If $V_\emptyset$ has cardinality greater than 1, then each vertex in $V_\emptyset$ is adjacent either with u or with each vertex in A.*

Proof. Suppose there exists $w' \in V_\emptyset$ such that w' is not adjacent with u. Let $u' \in A$ such that w' is adjacent with w'. There are four possible cases:

- $wu', ww' \in E(G)$,
- $ww' \in E(G)$ and $wu' \notin E(G)$,
- $wu' \in E(G)$ and $ww' \notin E(G)$,
- $wu', ww' \notin E(G)$.

In the cases when $ww' \in E(G)$, the vertex set $\{w, u, a, u', w'\}$ induces a graph isomorphic to P_5; those cases do not occur. On the other hand if $wu' \in E(G)$ and $ww' \notin E(G)$, then $\{w, w', u, u', a, b\}$ induces a graph isomorphic to $\mathsf{G}_{6,9}$, and this case can not occur. Finally, if $wu', ww' \notin E(G)$, then the vertex set $\{w, u, a, u', w'\}$ induces a graph isomorphic to P_5. Which is a contradiction. Thus each pair of vertices in $V_\emptyset$ are adjacent with the same vertices in A. And the result follows. □

Thus there are two cases: when each vertex in $V_\emptyset$ is adjacent with a unique vertex u in A, and when each vertex in $V_\emptyset$ is adjacent with each vertex in A. Note that $V_\emptyset$ induces a stable set, because otherwise $\omega(G) > 2$. In the first case we have that either $|V_\emptyset| \geq 2$ or the vertex set $A - u$ is empty. It is because otherwise G would have $\mathsf{G}_{6,1}$ as induced subgraph. Therefore, in each case G is isomorphic to a graph in $\mathcal{F}_1^1$. □

7. Proof of Theorem 2.35

One implication is easy because Proposition 2.31 implies that each graph in $\mathcal{F}_1$ is $\mathcal{F}_3$-free. The other implication is much more complex.

Suppose G is $\mathcal{F}_3$-free. Let $W = \{a, b, c\}$ be a clique of cardinality 3. For each $X \subseteq \{a, b, c\}$, let $V_X = \{u \in V(G) : N_G(u) \cap \{a, b, c\} = X\}$. Note that $V_\emptyset$ denote the vertex set whose vertices are not adjacent with a vertex in $\{a, b, c\}$.

Since $\omega(G) = 3$, then the vertex set $V_{a,b,c}$ is empty, and for each pair $\{x, y\} \subset \{a, b, c\}$, the vertex set $V_{x,y}$ induces a stable set. Furthermore

Claim 2.44. *For $x \in \{a, b, c\}$, the induced subgraph $G[V_x]$ is isomorphic to either $K_{m,n}$, $2K_2$, $K_2 + K_1$, or T_n.*

Proof. First consider K_3, P_4, $K_2 + T_2$ and $P_3 + T_1$ as induced subgraphs of $G[V_x]$. Since $G[K_3 \cup x] \simeq K_4$, $G[P_4 \cup \{x, y\}] \simeq \mathsf{G}_{6,12}$, $G[K_2 + T_2 \cup \{a, b, c\}] \simeq \mathsf{G}_{7,1}$, $G[P_3 + T_1 \cup \{x, y\}] \simeq \mathsf{G}_{6,4}$ and all of them are forbidden for G, then the graphs K_3, P_4, $K_2 + T_2$ and $P_3 + T_1$ are forbidden in $G[V_x]$. Thus $\omega(G[V_x]) \leq 2$.

If $\omega(G[V_x]) = 2$, then there exist $u, v \in V_x$ such that $uv \in E(G[V_x])$. Clearly, $N_{G[V_x]}(u) \cap N_{G[V_x]}(v) = \emptyset$, and each vertex set $N_{G[V_x]}(u)$ and $N_{G[V_x]}(u)$ induces a trivial subgraph. Since $G[V_x]$ is P_4-free, then $st \in E(V_x)$ for all $s \in N_{V_x}(u)\setminus\{v\}$ and $t \in N_{V_x}(v)\setminus\{u\}$. Therefore, each component in $G[V_x]$ is a complete bipartite subgraph.

If a component in $G[V_x]$ has cardinality at least three, then $G[V_x]$ does not have another component, because the existence of another component makes that $P_3 + T_1$ appears as an induced subgraph in $G[V_x]$; which is impossible. If there is a component in $G[V_x]$ of cardinality at least two, then there is at most another component in $G[V_x]$ since $K_2 + T_2$ is forbidden in $G[V_x]$. And thus the result turns out. □

CLAIM 2.45. *If there is $x \in \{a, b, c\}$ such that the induced subgraph $G[V_x]$ has at least 2 components or is isomorphic to a complete bipartite with at least three vertices, then the vertex set V_y is empty for each $y \in \{a, b, c\} - x$.*

PROOF. Let $G[V_x]$ be as above. Then there are two vertices $u, v \in V_x$ that are not adjacent. Suppose V_y is not empty, that is, there is $w \in V_y$. There are three possibilities:

- $uw, vw \in E(G)$,
- $uw \in E(G)$ and $vw \notin E(G)$, and
- $uw, vw \notin E(G)$.

But in each case the vertex set $\{a, b, c, u, v, w\}$ induces a graph isomorphic to $\mathsf{G}_{6,15}$, $\mathsf{G}_{6,11}$, and $\mathsf{G}_{6,2}$, respectively. Since these graphs are forbidden, then we obtain a contradiction and then V_y is empty. $\square$

CLAIM 2.46. *If there is $x \in \{a, b, c\}$ such that $G[V_x]$ has at least 2 components or is isomorphic to a complete bipartite with at least three vertices, then the vertex set $V_{x,y}$ is empty for each $y \in \{a, b, c\} - x$.*

PROOF. Let $G[V_x]$ be as above. Then there are two vertices $u_1, u_2 \in V_x$ that are not adjacent. Suppose $V_{x,y} \neq \emptyset$. Let $v \in V_{x,y}$. There are three possibilities:

- $u_1v \in E(G)$ and $u_2v \in E(G)$,
- $u_1v \in E(G)$ and $u_2v \notin E(G)$, and
- $u_1v \notin E(G)$ and $u_2v \notin E(G)$.

Then in each case, $G[\{a, b, c, u_1, u_2, v\}]$ is isomorphic to $\mathsf{G}_{6,16}$, $\mathsf{G}_{6,12}$ and $\mathsf{G}_{6,4}$, respectively. Since these graphs are forbidden, we have that $V_{x,y}$ is empty. $\square$

REMARK 2.47. *Claims 2.45 and 2.46 imply that when V_x has connected 2 components or is a complete bipartite graph of at least 3 vertices, then the only non-empty vertex sets are $V_\emptyset$, V_x and $V_{y,z}$, where x, y and z are different elements of $\{a, b, c\}$ Moreover, by Claim 2.44, the vertex set V_x is one of the following vertex sets:*

- *T_n, where $n \geq 2$,*
- *complete bipartite graph with cardinality at least 3,*
- *$K_1 + K_2$ or $2K_2$.*

Next result describes the induced subgraph $G[V_a \cup V_b \cup V_c]$ when each set V_x is connected of cardinality at most 2.

CLAIM 2.48. *Suppose for each $x \in \{a, b, c\}$ the vertex set V_x is connected of cardinality at most 2. If for all $x \in \{a, b, c\}$ the set V_x is not empty, then $G[V_a \cup V_b \cup V_c]$ is isomorphic (where x, y and z are different elements of $\{a, b, c\}$) to one of the following sets:*

- *$V_x \vee (V_y + V_z)$ where $V_x = K_1$, $V_y = K_m$, $V_z = K_n$ and $m, n \in \{1, 2\}$,*
- *$V_x \vee (V_y + V_z)$ where $V_x = K_2$, $V_y = K_1$, $V_z = K_1$, or*
- *$V_x \vee (V_y \vee V_z)$ where $V_x = K_1$, $V_y = K_1$, $V_z = K_1$.*

If $V_z = \emptyset$, then $G[V_x \cup V_y]$ is isomorphic to one of the following sets:

- *$V_x + V_y$, where $V_x = K_m$, $V_y = K_n$ and $m, n \in \{1, 2\}$, or*
- *$V_x \vee V_y$, where $V_x = K_1$, $V_y = K_m$ and $m \in \{1, 2\}$.*

If $V_y = V_z = \emptyset$, then V_x is isomorphic to K_1 or K_2.

PROOF. It is not difficult to prove that either $E(V_x, V_y)$ is empty or $E(V_x, V_y)$ induces a complete bipartite graph. The result follows by checking the possibilities with a computer algebra system. $\square$

In the rest of the proof, for each case obtained in Remark 2.47 and Claim 2.48, we will analyze the remaining edges sets and the vertex set $V_\emptyset$. As before, we may refer to x, y or z as different elements of $\{a, b, c\}$. Each case will be consider in a subsection.

7.0.1. **When** $V_a \cup V_b \cup V_c = \emptyset$**.** Now we describe the induced subgraph $G[V_{a,b} \cup V_{a,c} \cup V_{b,c}]$.

CLAIM 2.49. *If the vertex sets $V_{x,y}$ and $V_{y,z}$ are not empty, and the edge set $E(V_{x,y}, V_{y,z})$ is empty, then $|V_{x,y}| = |V_{y,z}| = 1$.*

PROOF. Suppose $|V_{x,y}| \geq 2$ and $V_{y,z} \neq \emptyset$. Take $u, u' \in V_{x,y}$ and $v \in V_{y,z}$. The result follows since the vertex set $\{a, b, c, u, u', v\}$ induces a graph isomorphic to the forbidden graph $\mathsf{G}_{6,16}$. $\square$

CLAIM 2.50. *If $E(V_{x,y}, V_{y,z}) \neq \emptyset$, $E(V_{x,y}, V_{x,z}) = \emptyset$ and $E(V_{y,z}, V_{x,z}) = \emptyset$, then $V_{x,z}$ is empty.*

PROOF. Let $u \in V_{x,y}$ and $v \in V_{y,z}$ such that $uv \in E(G)$. Suppose there exists $w \in V_{x,z}$. Then $uw, vw \notin E(G)$. Since the induced subgraph $G[\{a, b, c, u, w, v\}]$ is isomorphic to $\mathsf{G}_{6,24}$, then we get a contradiction. Then $V_{x,z}$ is empty. $\square$

CLAIM 2.51. *If $E(V_{x,y}, V_{y,z}) \neq \emptyset$, then $E(V_{x,y}, V_{y,z})$ induces either a complete bipartite graph or a complete bipartite graph minus an edge.*

PROOF. Let $u \in V_{x,y}$. Suppose there exist $v, v' \in V_{y,z}$ such that $uv, uv' \notin E(G)$. Since the induced subgraph $G[\{a, b, c, u, v, v'\}]$ is isomorphic to $\mathsf{G}_{6,16}$, which is forbidden, then the vertex u is adjacent with at least all but one vertices in $V_{y,z}$. In a similar way, we have that each vertex in $V_{y,z}$ is adjacent with at at least all but one vertices in $V_{x,y}$. Therefore, the edge set $E(V_{x,y}, V_{y,z})$ induces a complete bipartite graph minus a matching. Now suppose this matching has cardinality greater or equal to 2. Then there exist $u, u' \in V_{x,y}$ and $v, v' \in V_{y,z}$ such that $uv', u'v \notin E(G)$ and $uv, u'v' \in E(G)$. Since $G[\{v, u, x, u', v'\}]$ is isomorphic to P_5, then we get a contradiction and the matching has cardinality at most 1. Therefore, $E(V_{xy}, V_{y,z})$ induces a complete bipartite graph or a complete bipartite graph minus an edge. $\square$

CLAIM 2.52. *If $E(V_{x,y}, V_{y,z}) \neq \emptyset$, $E(V_{y,z}, V_{x,z}) \neq \emptyset$ and $E(V_{x,y}, V_{x,z}) = \emptyset$, then $V_{x,y} = \{v_1\}$, $V_{x,z} = \{v_2\}$ and one of the following two statements holds:*

- *$E(v_1, V_{y,z}) = \{v_1 v : v \in V_{y,z}\}$ and $E(v_2, V_{y,z}) = \{v_2 v : v \in V_{y,z}\}$, or*
- *there exists $v_3 \in V_{y,z}$ such that $E(v_1, V_{y,z}) = \{v_1 v : v \in V_{y,z} - v_3\}$ and $E(v_2, V_{y,z}) = \{v_2 v : v \in V_{y,z} - v_3\}$.*

PROOF. Since there is no edge joining a vertex in $V_{x,y}$ and a vertex in $V_{x,z}$, then Claim 2.49 implies that $|V_{x,y}| = |V_{x,z}| = 1$. Let $V_{x,y} = \{v_1\}$ and $V_{x,z} = \{v_2\}$. By Claim 2.51, each edge set $E(V_{x,y}, V_{y,z})$ and $E(V_{y,z}, V_{x,z})$ induces either a complete bipartite graph or a complete bipartite graph minus an edge. If both sets $E(V_{x,y}, V_{y,z})$ and $E(V_{y,z}, V_{x,z})$ induce either a complete bipartite graph, then we are done. So it remains to check two cases:

- when one edge set induces a complete bipartite graph and the other one induces a complete bipartite graph minus an edge, and
- when both edge sets induce a complete bipartite graph minus an edge.

First consider the former case. Suppose there exists $v_3 \in V_{y,z}$ such that $E(v_1, V_{y,z}) = \{v_1 v : v \in V_{y,z} - v_3\}$ and $E(v_2, V_{y,z}) = \{v_2 v : v \in V_{y,z}\}$. Since $G[\{a, b, c, v_1, v_2, v_3\}]$ is isomorphic to

$\mathsf{G}_{6,24}$, then this case is not possible. Now consider second case. Let $v_3, v_4 \in V_{y,z}$. There are two possible cases: either $v_1v_3, v_2v_3 \notin E(G)$, or $v_1v_3, v_2v_4 \notin E(G)$. Suppose $v_1v_3, v_2v_4 \notin E(G)$. Since $G[\{v_4, v_1, z, v_2, v_3\}]$ is isomorphic to P_5, then this case is impossible. And therefore $v_1v_3, v_2v_3 \notin E(G)$. □

CLAIM 2.53. *If each edge set $E(V_{x,y}, V_{y,z})$ is not empty, then the induced subgraph $G[V_{a,b} \cup V_{a,c} \cup V_{b,c}]$ is isomorphic to one of the following graphs:*

- *a complete tripartite graph,*
- *a complete tripartite graph minus an edge, or*
- *a complete tripartite graph minus the edges v_1v_2, v_2v_3, v_3v_1, where $v_1 \in V_{x,y}$, $v_2 \in V_{y,z}$, $v_3 \in V_{x,z}$.*

PROOF. By Claim 2.51, the induced subgraph $G[V_{a,b} \cup V_{a,c} \cup V_{b,c}]$ is a complete tripartite graph minus at most three edges. We will analyze the cases where 2 or 3 edges have been removed.

Suppose $G[V_{a,b} \cup V_{a,c} \cup V_{b,c}]$ induces a complete tripartite graph minus 2 edges. Since $E(V_{x,y}, V_{y,z})$ induces complete bipartite graph or complete bipartite graph minus an edge, then the 2 edges cannot be removed from a unique edge set $E(V_{x,y}, V_{y,z})$. Then there are two possibilities:

- there exist $u, u' \in V_{x,y}$, $v \in V_{y,z}$ and $w \in V_{x,z}$ such that $uv, u'w \notin E(G)$, or
- there exist $u \in V_{xy}$, $v \in V_{y,z}$ and $w \in V_{x,z}$ such that $uv, uw \notin E(G)$.

In the first case the graph induced by the set $\{x, v, u, u', w, z\}$ is isomorphic to the forbidden graph $\mathsf{G}_{6,17}$; which is impossible. And in the second case, the induced subgraph $G[\{a, b, c, u, v, w\}]$ is isomorphic to $\mathsf{G}_{6,13}$, which is impossible. Thus the case where 2 edges are removed from $G[V_{a,b} \cup V_{a,c} \cup V_{b,c}]$ is not possible.

Suppose $G[V_{a,b} \cup V_{a,c} \cup V_{b,c}]$ is a complete tripartite graph minus 3 edges. Since $E(V_{x,y}, V_{y,z})$ induces a complete bipartite graph or a complete bipartite graph minus an edge, then the 3 edges cannot be removed from a unique edge set $E(V_{x,y}, V_{y,z})$. Thus there are four possible cases:

(a) $v_1v_4, v_3v_6, v_5v_2 \notin E(G)$, where $v_1, v_2 \in V_{x,y}$, $v_3, v_4 \in V_{y,z}$ and $v_5, v_6 \in V_{x,z}$
(b) $v_1v_3, v_2v_5, v_5v_4 \notin E(G)$, where $v_1, v_2 \in V_{x,y}$, $v_3, v_4 \in V_{y,z}$ and $v_5 \in V_{x,z}$
(c) $v_2v_4, v_4v_5, v_5v_2 \notin E(G)$, where $v_1, v_2 \in V_{x,y}$, $v_3, v_4 \in V_{y,z}$ and $v_5 \in V_{x,z}$ and
(d) $v_1v_3, v_3v_5, v_5v_2 \notin E(G)$, where $v_1, v_2 \in V_{x,y}$, $v_3, v_4 \in V_{y,z}$ and $v_5 \in V_{x,z}$.

Cases (a), (b) and (d) are impossible, the argument is the following. Case (a) is impossible because the induced subgraph $G[\{v_2, v_3, v_5, v_6, x, y\}]$ is isomorphic to the forbidden graph $\mathsf{G}_{6,17}$. In case (b), the induced subgraph $G[\{v_2, v_5, v_4, a, b, c\}]$ is isomorphic to the forbidden graph $\mathsf{G}_{6,24}$. And in case (d), the induced subgraph $G[\{v_2, v_3, v_5, a, b, c\}]$ is isomorphic to the forbidden graph $\mathsf{G}_{6,24}$. Thus when 3 edges are removed the only possible case is (c). □

By previous Claims we have the following cases:

(1) when the set $V_{x,y}$ is the only not empty set,
(2) when $V_{x,z} = \emptyset$, $V_{x,y} = \{v_{xy}\}$, $V_{y,z} = \{v_{yz}\}$ and $v_{xy}v_{yz} \notin E(G)$,
(3) when $V_{x,z} = \emptyset$ and $E(V_{x,y}, V_{y,z})$ induces a bipartite complete graph,
(4) when $V_{x,z} = \emptyset$ and $E(V_{x,y}, V_{y,z})$ induces a bipartite complete graph minus an edge,
(5) when $V_{a,b} = \{v_{ab}\}$, $V_{a,c} = \{v_{ac}\}$, $V_{b,c} = \{v_{bc}\}$ and $v_{ab}v_{ac}, v_{ab}v_{bc}, v_{ac}v_{bc} \notin E(G)$,
(6) when $V_{x,z} = \{v_{xz}\}$, $V_{y,z} = \{v_{yz}\}$, $v_{xz}v_{yz} \notin E(G)$, and the edge sets $E(v_{xz}, V_{x,y})$ and $E(v_{yz}, V_{x,y})$ induce a complete bipartite graph,
(7) when $V_{x,z} = \{v_{xz}\}$, $V_{y,z} = \{v_{yz}\}$, $v_{xz}v_{yz} \notin E(G)$, and there exists $v_{xy} \in V_{x,y}$ such that $E(v_{xz}, V_{x,y}) = \{v_{xz}v : v \in V_{x,y} - v_{xy}\}$ and $E(v_{yz}, V_{x,y}) = \{v_{yz}v : v \in V_{x,y} - v_{xy}\}$.

(8) when $G[V_{a,b} \cup V_{a,c} \cup V_{b,c}]$ is isomorphic to a complete tripartite graph,
(9) when $G[V_{a,b} \cup V_{a,c} \cup V_{b,c}]$ is isomorphic to a complete tripartite graph minus an edge,
(10) when $G[V_{a,b} \cup V_{a,c} \cup V_{b,c}]$ is isomorphic to a complete tripartite graph, where $v_1 \in V_{x,y}$, $v_2 \in V_{y,z}$, $v_3 \in V_{x,z}$ and $v_1v_2, v_2v_3, v_3v_1 \notin E(G)$.

Now we describe the vertex set $V_\emptyset$.

REMARK 2.54. *Let $w, w' \in V_\emptyset$. Suppose w is adjacent with a vertex in $V_{x,y}$ and with w'. Then the vertex w' is adjacent with a vertex in $V_{x,y}$, because otherwise the shortest path from w' to z would contains a graph isomorphic to P_5. Thus each vertex in $V_\emptyset$ is adjacent with a vertex in $V_{x,y}$ for some $\{x, y\} \subset \{a, b, c\}$.*

CLAIM 2.55. *If $w \in V_\emptyset$ is adjacent with $v \in V_{x,y}$, then either w is adjacent only with v and with no other vertex in $V_{x,y}$, or w is adjacent with each vertex in $V_{x,y}$. Moreover, if each vertex in $V_\emptyset$ is adjacent with a vertex in $V_{x,y}$, then either exists a vertex $v \in V_{v,x}$ such that each vertex in $V_\emptyset$ is adjacent with v, or each vertex in $V_\emptyset$ is adjacent with each vertex in $V_{x,y}$.*

PROOF. Since the first statement is easy when $V_{x,y}$ has cardinality at most 2, then we assume that $V_{x,y}$ has cardinality at least 3. Let $v', v'' \in V_{x,y}$. Suppose w is adjacent with v and v' but not adjacent with v''. Since $G[\{x, z, v, v', v'', w\}]$ is isomorphic to $\mathsf{G}_{6,3}$, then we get a contradiction. And then w is adjacent only with v or with v, v' and v''.

Let $w, w' \in V_\emptyset$. Suppose there is $v \in V_{x,y}$ such that $wv \in E(G)$ and $w'v \notin E(G)$. The vertex w is not adjacent with w', because otherwise by Remark 2.54 we get a contradiction. Let $v' \in V_{x,y}$ such that w' is adjacent with w'. Thus there are two possible cases:

- $ww', wv' \notin E(G)$, and
- $wv' \in E(G)$ and $ww' \notin E(G)$.

Since in the first case $G[\{w, v, x, v', w'\}]$ is isomorphic to P_5 and in the second case $G[\{x, z, v, v', w, w'\}]$ is isomorphic to $\mathsf{G}_{6,9}$, then we get a contradiction and thus there is no vertex in $V_{x,y}$ adjacent with a vertex in $V_\emptyset$ and not adjacent with other vertex in $V_\emptyset$. ☐

CLAIM 2.56. *Let $v \in V_{x,y}$. If each vertex in $V_\emptyset$ is adjacent with v, then $V_\emptyset$ induces either K_2 or it is a trivial graph. If furthermore there exists $v' \in V_{x,y}$ such that no vertex in $V_\emptyset$ is adjacent with v', then $V_\emptyset$ is a clique of cardinality at most 2.*

PROOF. First note that P_3 is forbidden as induced subgraph in $G[V_\emptyset]$. It is because if the vertices $w_1, w_2, w_3 \in V_\emptyset$ induce a graph isomorphic to P_3, then $G[\{x, y, v, w_1, w_2, w_3\}] \simeq \mathsf{G}_{6,7}$. Now we will see that each component in $G[V_\emptyset]$ is a clique. Let C be a component in $G[V_\emptyset]$. Suppose C is not a clique, then there are two vertices not adjacent in C, say w and w'. Let P be the smallest path contained in C between w and w'. The length of P is greater or equal to 3. So P_3 is an induced subgraph of P, and hence of C. Which is a contradiction, and therefore, C is a clique. On the other hand, the graph $K_2 + K_1$ is forbidden as induced subgraph in $G[V_\emptyset]$. It is because if $w_1, w_2, w_3 \in V_\emptyset$ such that $G[\{w_1, w_2, w_3\}] \simeq K_2 + K_1$, then $G[\{x, y, v, w_1, w_2, w_3\}] \simeq \mathsf{G}_{6,8}$. Therefore, if $G[V_\emptyset]$ has more than one component, then each component has cardinality one.

Let $v, v' \in V_{x,y}$ such that each vertex in $V_\emptyset$ is adjacent with v, and no vertex in $V_\emptyset$ is adjacent with v'. Suppose $V_\emptyset$ induces a stable set of cardinality at least 2. Take $w, w' \in V_\emptyset$. Then we get a contradiction since the induced graph $G[\{w, w', v, v', x, z\}]$ is isomorphic to $\mathsf{G}_{6,1}$. Thus $V_\emptyset$ is a clique of cardinality at most 2. ☐

Thus by Claims 2.55 and 2.56, in case (1) we have the following possible cases:

- $V_\emptyset$ is a clique of cardinality at most 2, and each vertex in $V_\emptyset$ is adjacent with only one vertex in $V_{x,y}$, and
- $V_\emptyset$ is a trivial graph and $E(V_\emptyset, V_{x,y})$ induces a complete bipartite graph.

Note that these graphs are isomorphic to an induced subgraph of a graph in $\mathcal{F}_1^1$.

Claim 2.57. *If $E(V_{x,y}, V_{y,z}) = \emptyset$ and $E(V_\emptyset, V_{x,y} \cup V_{y,z}) \neq \emptyset$, then $V_\emptyset$ is a clique of cardinality at most 2, and each vertex in $V_\emptyset$ is adjacent each vertex in $V_{x,y} \cup V_{y,z}$*

Proof. Let $v_{xy} \in V_{x,y}$, $v_{yz} \in V_{y,z}$ such that $v_{xy}v_{yz} \notin E(G)$. It is easy to see that if w is adjacent with v_{xy} or v_{yz}, then w is adjacent with both vertices, because otherwise G has an induced subgraph isomorphic to P_5. Thus each vertex in $V_\emptyset$ is adjacent with each vertex in $V_{x,y} \cup V_{y,z}$. Now suppose $w, w' \in V_\emptyset$ such that w and w' are not adjacent. Since the induced subgraph $G[\{w, w', x, y, v_{xy}, v_{yz}\}]$ is isomorphic to $\mathsf{G}_{6,15}$, then we get a contradiction and $V_\emptyset$ induces a clique of cardinality at most 2. □

By previous Claim we get that in case (2), $V_\emptyset$ is a clique of cardinality at most 2 and each vertex in $V_\emptyset$ is adjacent each vertex in $V_{x,y} \cup V_{y,z}$. This graph is isomorphic to an induced subgraph of a graph in $\mathcal{F}_1^1$.

Claim 2.58. *If the edge set $E(V_{x,y}, V_{y,z})$ induces a complete bipartite graph and $E(V_\emptyset, V_{x,y} \cup V_{y,z}) \neq \emptyset$, then $V_\emptyset$ is a clique of cardinality at most 2, and each vertex in $V_\emptyset$ is adjacent only to one vertex in $V_{x,y} \cup V_{y,z}$.*

Proof. First we prove that there each vertex in $V_\emptyset$ is adjacent with a vertex in only one of the sets $V_{x,y}$ or $V_{y,z}$. Suppose there exists a vertex $w \in V_\emptyset$ adjacent with $v \in V_{x,y}$ and $u \in V_{y,z}$. Since the induced subgraph $G[\{w, a, b, c, v, u\}]$ is isomorphic to $\mathsf{G}_{6,24}$, then we get a contradiction and each vertex in $V_\emptyset$ is adjacent only with vertices of one of the vertex sets $V_{x,y}$ or $V_{y,z}$. Suppose there are two vertices $w, w' \in V_\emptyset$ such that w is adjacent with $v \in V_{x,y}$, and w' is adjacent with $u \in V_{y,z}$. Since $G[\{w, w', u, v, x, z\}]$ is isomorphic to $\mathsf{G}_{6,9}$, then this is impossible and the vertices of $V_\emptyset$ are adjacent only to vertices in one of the vertex sets either $V_{x,y}$ or $V_{y,z}$. Now suppose that $w \in V_\emptyset$ is adjacent with two vertices in $V_{x,y}$, say v and v'. Take $u \in V_{y,z}$. So u is adjacent with both v and v'. Since the induced subgraph $G[\{w, v, v', u, a, b, c\}]$ is isomorphic to $\mathsf{G}_{7,10}$, then this cannot occur. Thus each vertex in $V_\emptyset$ is adjacent only with one vertex in $V_{x,y} \cup V_{y,z}$. Finally suppose $V_\emptyset$ induces a trivial graph of cardinality at least 2. Let $w, w' \in V_\emptyset$ adjacent with $v \in V_{x,y}$. Take $u \in V_{y,z}$, so u is adjacent with both v. Since the induced subgraph $G[\{w, w, v, u, x, z\}]$ is isomorphic to $\mathsf{G}_{6,6}$, we get a contradiction and the result follows. □

By previous Claim we have that in case (3) the vertex set $V_\emptyset$ induces a clique of cardinality at most 2 and each vertex in $V_\emptyset$ is adjacent only to one vertex in $V_{x,y} \cup V_{y,z}$. In this case, the graph is isomorphic to an induced subgraph of a graph in $\mathcal{F}_1^1$.

Claim 2.59. *Let $u \in V_{x,y}$ and $v, v' \in V_{y,z}$ such that u is adjacent with v but not with v'. If $w \in V_\emptyset$, then w is not adjacent with v.*

Proof. Suppose w is adjacent with v. Note that if w is adjacent with u or v', then w is adjacent with both u and v'. Thus there are two cases: either w is adjacent only with v, or w is adjacent with v, v' and u. Both cases are impossible because in the former case $G[\{x, z, u, v, v', w\}]$ is isomorphic to $\mathsf{G}_{6,9}$, meanwhile in the second case $G[\{y, z, u, v, v', w\}]$ is isomorphic to $\mathsf{G}_{6,17}$; which is a contradiction. □

Consider case (4). Let $u \in V_{x,y}$ and $v \in V_{y,z}$ such that u is not adjacent with v. By Claim 2.59, each vertex in $V_\emptyset$ is adjacent with u or with v. It is not difficult to see that in fact each vertex in $V_\emptyset$ is adjacent with both u and v, because otherwise G would contain P_5 as induced subgraph. By applying Claim 2.57 to the induced subgraph $G[\{u,v\} \cup V_\emptyset]$, we get that $V_\emptyset$ is a clique of cardinality at most 2. And this graph is isomorphic to an induced subgraph of a graph in $\mathcal{F}_1^1$.

Now consider case (5), by Claim 2.57, we get that $V_\emptyset$ is a clique of cardinality at most 2, and each vertex in $V_\emptyset$ is adjacent each vertex in $V_{a,b} \cup V_{b,c} \cup V_{a,c}$. And this graph is isomorphic to an induced subgraph of a graph in $\mathcal{F}_1^1$.

Claim 2.60. *Let $u_1 \in V_{x,y}$, $u_2 \in V_{y,z}$ and $u_3 \in V_{x,z}$ such that u_1 is adjacent with both u_2 and u_3, and $u_2u_3 \notin E(G)$. If $w \in V_\emptyset$, then w is not adjacent with u_1.*

Proof. Suppose w is adjacent with u_1. Note that if w is adjacent with u_2 or u_3, then w is adjacent with both u_2 and u_3. Thus there are two cases: either w is adjacent only with u_1, or w is adjacent with u_1, u_2 and u_3. Both cases are impossible because in the former case $G[\{x, y, u_1, u_2, u_3, w\}]$ is isomorphic to $\mathsf{G}_{6,12}$, meanwhile in the second case $G[\{x, z, u_1, u_2, u_3, w\}]$ is isomorphic to $\mathsf{G}_{6,25}$; which is a contradiction. □

Consider the Cases (6) and (9). Let $v_{xz} \in V_{x,z}$ and $v_{yz} \in V_{y,z}$ such that v_{xz} is not adjacent with v_{yz}. By Claims 2.57 and 2.60, each vertex in $V_\emptyset$ is adjacent only with both v_{xz} and v_{yz}. And by Claim 2.57, we get that $V_\emptyset$ is a clique of cardinality at most 2. And these graphs are isomorphic to an induced subgraph of a graph in $\mathcal{F}_1^1$.

Now consider the Case (7) and (10). Let $v_{xy} \in V_{x,y}$, $v_{xz} \in V_{x,z}$ and $v_{yz} \in V_{y,z}$ such that v_{xy} is not adjacent with v_{xz} and v_{yz}, and v_{xz} is not adjacent with v_{yz}. By Claims 2.60 and 2.57, each vertex in $V_\emptyset$ is adjacent only with v_{xy}, v_{xz} and v_{yz}. And by Claim 2.57, we get that $V_\emptyset$ is a clique of cardinality at most 2. And these graphs are isomorphic to an induced subgraph of a graph in $\mathcal{F}_1^1$.

Finally in the case (8), by Claim 2.58 we have that the vertex set $V_\emptyset$ is a clique of cardinality at most 2 and each vertex in $V_\emptyset$ is adjacent with only one vertex in $V_{a,b} \cup V_{b,c} \cup V_{a,c}$. This case corresponds to a graph that is isomorphic to an induced subgraph of a graph in $\mathcal{F}_1^1$.

7.0.2. **Case $V_x = T_n$, where $n \geq 2$.** First we will obtain that $E(V_x, V_{y,z})$ satisfies one of the following statements:

- it induces a complete bipartite graph,
- there exists a vertex $s \in V_x$, we called the apex, such that $E(V_x, V_{y,z}) = \{uv : u \in V_x - s \text{ and } v \in V_{y,z}\}$, or
- there exists a vertex $s \in V_{y,z}$, we called the apex, such that $E(V_x, V_{y,z}) = \{uv : u \in V_x \text{ and } v \in V_{y,z} - s\}$.

To do it, we will prove the following claims.

Claim 2.61. *Let $u_1, u_2 \in V_x$ and $v_1, v_2 \in V_{y,z}$ such that $u_1v_1, u_2v_2 \notin E(G)$. Then either $u_1 = u_2$ or $v_1 = v_2$.*

Proof. Suppose $u_1 \neq u_2$ and $v_1 \neq v_2$. There are three possible cases:

- $u_1v_2, u_2v_1 \notin E(G)$,
- $u_1v_2 \notin E(G)$ and $u_2v_1 \in E(G)$, or
- $u_1v_2, u_2v_1 \in E(G)$.

The first two cases are not possible since the induced subgraph $G[\{x, y, u_1, u_2, v_1, v_2\}]$ would be isomorphic to $\mathsf{G}_{6,1}$ and $\mathsf{G}_{6,9}$, respectively. In the last case, the induced subgraph $G[\{x, u_1, u_2, v_1, v_2\}]$ is isomorphic to P_5; which is impossible. Thus the result follows. □

The last claim implies that all *non-edges* in $E(V_x, V_{y,z})$ are incident to a one vertex: the apex s.

Suppose the vertex s is in V_x, and there are vertices $v_1, v_2 \in V_{y,z}$ such that $sv_1 \in E(G)$, and $sv_2 \notin E(G)$. By Claim 2.61, each vertex in $V_x - s$ is adjacent with v_1 and v_2. Then the induced subgraph $G[\{a, b, c, u_1, u_2, v_1, v_2\}] \simeq \mathsf{G}_{7,9}$, that is impossible. This implies that if the vertex $s \in V_x$ is not adjacent with a vertex in $V_{y,z}$, then s is not adjacent with all vertices in $V_{y,z}$. A similar argument yields that if the apex vertex s is in $V_{y,z}$ and s is not adjacent with a vertex in V_x, then s is not adjacent with all vertices in V_x.

Thus, we have three cases:

(a) $E(V_x, V_{y,z})$ is complete bipartite minus the edges between a vertex s (the apex) in V_x and all vertices of $V_{y,z}$,
(b) $E(V_x, V_{y,z})$ is complete bipartite minus the edges between a vertex s (the apex) in $V_{y,z}$ and all vertices in V_x, and
(c) $E(V_x, V_{y,z})$ is complete bipartite.

Claim 2.62. *If $|V_x| \geq 3$ and $v \in V_{y,z}$, then $E(v, V_x)$ either it induces a complete bipartite graph or it is empty.*

Proof. Let $u_1, u_2, u_3 \in V_x$. Suppose one of the two following possibilities happen: $vu_1 \in E(G)$ and $vu_2, vu_3 \notin E(G)$, or $vu_1, vu_2 \in E(G)$ and $vu_3 \notin E(G)$. In the first case the induced subgraph $G[\{u_1, u_2, u_3, v, x, y\}]$ is isomorphic to $\mathsf{G}_{6,3}$, meanwhile in the second case the induced subgraph $G[\{u_1, u_2, u_3, v, x, y, z\}]$ is isomorphic to $\mathsf{G}_{7,2}$. Then both cases cannot occur, and we get the result. □

Thus case (a) occur only when $|V_x| = 2$.

In what follows we describe the vertex set $V_\emptyset$, that is, the vertex set whose vertices have no edge in common with the vertex set $\{a, b, c\}$.

Remark 2.63. *Let $w \in V_\emptyset$. The vertex w is adjacent with a vertex in $V_x \cup V_{y,z}$, because otherwise the shortest path from w to $\{a, b, c\}$ would contains the graph P_5 as induced subgraph. Let $u_1, u_2 \in V_x$. If w is adjacent with u_1 or u_2, then w is adjacent with both vertices, because otherwise the induced subgraph $G[\{a, b, c, u_1, u_2, w\}]$ would be isomorphic to $\mathsf{G}_{6,6}$, which is forbidden.*

In case (a), we will see that each vertex in $V_\emptyset$ is adjacent with each vertex in $V_x \cup V_{y,z}$. Let w in $V_\emptyset$. Supppose $s \in V_x$ is the vertex that is not adjacent with any vertex in $V_{y,z}$. If $w \in V_\emptyset$ is adjacent with one of the vertices in $\{s\} \cup V_{y,z}$, then w must to be adjacent with s and each vertex in $V_{y,z}$, because otherwise let $v \in V_{y,z}$, then the induced subgraph $G[\{x, y, s, w, v\}]$ would be isomorphic to P_5. Then by Remark 2.63, w is adjacent with each vertex in $V_x \cup V_{y,z}$.

Claim 2.64. *The vertex set $V_\emptyset$ induces a stable set.*

Proof. Suppose $w_1, w_2 \in V_\emptyset$ are adjacent. Since both w_1 and w_2 are adjacent with $u \in V_x - s$ and $v \in V_{y,z}$, then the induced subgraph $G[\{u, v, w_1, w_2\}]$ is isomorphic to K_4; which is forbidden. Thus w_1 and w_2 are not adjacent, and therefore $V_\emptyset$ is a stable set. □

Thus this case corresponds to a graph that is isomorphic to an induced subgraph of a graph in $\mathcal{F}_1^1$.

Now consider case (b). Let $u_1, u_2 \in V_x$ and $s \in V_{y,z}$ such that s is not adjacent with u_1 and u_2.

Claim 2.65. *If $w \in V_\emptyset$ is adjacent with a vertex in $V_{y,z} \setminus \{s\}$, then each vertex in $V_\emptyset$ is adjacent with each vertex in $V_x \cup V_{y,z}$.*

Proof. Let $w \in V_\emptyset$. Suppose w is adjacent with $v \in V_{y,z} \setminus \{s\}$. If w is not adjacent with both vertices $u_1, u_2 \in V_x$, then we get that the induced subgraph $G[\{a, b, c, w, u_1, u_2, v\}]$ is isomorphic to $\mathsf{G}_{7,2}$, which cannot be. Thus w is adjacent with at least one vertex in V_x. Then by Remark 2.63, w is adjacent with both vertices u_1 and u_2. If there exist $v' \in V_{y,z} \setminus \{s\}$ such that w is not adjacent with v', then the vertex set $\{w, u_1, x, y, v, v'\}$ would induce a graph isomorphic to $\mathsf{G}_{6,15}$. On the other hand, if w is not adjacent with s, then the vertex set $\{w, u_1, x, y, s\}$ induces the subgraph P_5. Therefore, w is adjacent with each vertex in $V_x \cup V_{y,z}$.

Suppose there is another vertex $w' \in V_\emptyset$. By the above argument if w' is adjacent with a vertex in $V_{y,z} \setminus \{s\}$, then it must be adjacent with each vertex in $V_x \cup V_{y,z}$. Also if w' is adjacent with a vertex in $\{s\} \cup V_x$, then w' must be adjacent with each vertex in $\{s\} \cup V_x$. So suppose w' is adjacent with each vertex in $s \cup V_x$ and w' is not adjacent with each vertex in $V_{y,z} \setminus \{s\}$. Then there are two possibilities: either $ww' \notin E(G)$ or $ww' \in E(G)$. Let $v \in V_{y,z} \setminus \{s\}$. In the first case $G[\{x, y, s, v, w, w'\}] \simeq \mathsf{G}_{6,15}$ and in the second case $G[\{x, y, v, w, w'\}] \simeq \mathsf{P}_5$. Since both graphs are forbidden, then we get a contradiction and thus w' is adjacent with v. And therefore w' is adjacent with each vertex in $V_x \cup V_{y,z}$. □

Claim 2.66. *Either each vertex in $V_\emptyset$ is adjacent with each vertex in $V_x \cup \{s\}$, or each vertex in $V_\emptyset$ is adjacent with each vertex in $V_x \cup V_{y,z}$.*

Proof. Let $w \in V_\emptyset$. Clearly, if w is adjacent with a vertex in $V_x \cup \{v\}$, then w is adjacent with each vertex in $V_x \cup \{v\}$, because otherwise P_5 would be an induced subgraph. By Claim 2.65, we have that if there is a vertex $w \in V_\emptyset$ adjacent with a vertex in $V_{y,z}$, different to the apex s, then each vertex in $V_\emptyset$ is adjacent with each vertex in V_x and $V_{y,z} \setminus \{s\}$. Thus we get the result. □

Claim 2.67. *The vertex set $V_\emptyset$ induces either a clique of cardinality at most 2 or a trivial graph.*

Proof. First note that P_3 is forbidden as induced subgraph in $V_\emptyset$. It is because if $w_1, w_2, w_3 \in V_\emptyset$ induce P_3, then $G[\{x, y, u_1, w_1, w_2, w_3\}] \simeq \mathsf{G}_{6,7}$. Now we will get that each component in $G[V_\emptyset]$ is a clique. Let C be a component in $G[V_\emptyset]$. Suppose C is not a clique, then it has two vertices not adjacent, say u and v. Let P be the smallest path in C between u and v. Thus the length of P is greater or equal to 3. So P_3 is an induced subgraph of P, and hence of C. Therefore, C is a complete graph.

On the other hand, the graph $K_2 + K_1$ is forbidden for $G[V_\emptyset]$. It is because if $w_1, w_2, w_3 \in V_\emptyset$ such that $G[\{w_1, w_2, w_3\}] \simeq K_2 + K_1$, then $G[\{x, y, u_1, w_1, w_2, w_3\}] \simeq \mathsf{G}_{6,8}$; which cannot happen. Therefore, if $G[V_\emptyset]$ has more than one component, then each component has cardinality one. □

In the first case of Claim 2.66, if $|V_{y,z}| \geq 2$, then $V_\emptyset$ is either K_1 or K_2, because if $u_1 \in V_x$, $v \in V_{y,z} \setminus \{s\}$, and $w_1, w_2 \in V_\emptyset$ are adjacent, then $G[\{w_1, w_2, u_1, v, x, y\}] \simeq \mathsf{G}_{6,3}$. Otherwise if $V_{y,z} = \{s\}$, then both possibilities in Claim 2.67 are allowed. In the second case of Claim 2.66, if $|V_{y,z}| \geq 3$, we have that $V_\emptyset = \emptyset$, because if $w \in V_\emptyset$ is adjacent with $u_1, u_2 \in V_x$ and with two vertices $v_1, v_2 \in V_{y,z} \setminus \{s\}$, we get $G[\{x, y, u_1, u_2, v_1, v_2, w\}] \simeq \mathsf{G}_{7,13}$ as forbidden subgraph. If $|V_{y,z}| = 2$, then $V_\emptyset$ is trivial since $\omega(G) = 3$. And if $V_{y,z} = \{s\}$, then both possibilities in Claim 2.67 are allowed. With no much effort the reader can see that each of these cases corresponds to a graph isomorphic to an induced subgraph of a graph in $\mathcal{F}_1^1$.

Case (c). By Claim 2.66, there are two possible cases:

- either each vertex in $V_\emptyset$ is adjacent with each vertex in $V_x \cup V_{y,z}$, or
- each vertex in $V_\emptyset$ is adjacent with each vertex in V_x, and no vertex in $V_\emptyset$ is adjacent with any vertex in $V_{y,z}$.

In first case when $|V_{y,z}| \geq 2$, the vertex set $V_\emptyset$ is empty. Because if $w \in V_\emptyset$ is adjacent with the vertices $v_1, v_2 \in V_{y,z}$, and $u_1, u_2 \in V_x$, then the induced subgraph $G[\{x, y, u_1, u_2, v_1, v_2, w\}] \simeq \mathsf{G}_{7,13}$ which is forbidden. Then the vertex set $V_\emptyset$ is empty. Otherwise when $|V_{y,z}| = 1$, then the vertex set $V_\emptyset$ must be a stable set, because if there exist two adjacent vertices in $V_\emptyset$, then by taking a vertex in V_x and a vertex in $V_{y,z}$ we get K_4 that is forbidden. Finally, in the case when each vertex in $V_\emptyset$ is adjacent with each vertex in V_x, and no vertex in $V_\emptyset$ is adjacent with a vertex in $V_{y,z}$, we get that $V_\emptyset$ is a clique of cardinality at most 2. It is because if $w_1, w_2 \in V_\emptyset$ are such that $w_1w_2 \notin E(G)$, then we get $G[\{w_1, w_2, u_1, v, x, y\}] \simeq \mathsf{G}_{6,3}$ that is forbidden. And each of these graphs are isomorphic to an induced subgraph of a graph in $\mathcal{F}_1^1$.

7.0.3. **Case V_x is a complete bipartite graph of cardinality at least 3.** Assume V_x is a complete bipartite graph of cardinality at lest three with (A, B) the bipartition of V_x.

Claim 2.68. *If $v \in V_{y,z}$, then $E(v, V_x) \neq \emptyset$.*

Proof. Suppose $E(v, V_x) = \emptyset$. There exist $u_1, u_2, u_3 \in V_x$ such that $G[\{u_1, u_2, u_3\}] \simeq P_3$. And then we get contradiction since the induced subgraph $G[\{u_1, u_2, u_3, x, y, v\}] \simeq \mathsf{G}_{6,7}$; which is forbidden. $\square$

Claim 2.69. *Let $u \in V_x$ and $v \in V_{y,z}$. If u and v are adjacent, then v is adjacent with each vertex in the part (A or B) containing u.*

Proof. Suppose $u \in A$ and v is not adjacent with any vertex in $A \setminus \{u\}$. If $|A| = 1$, then the result follows, so we may assume $|A| \geq 2$. Since $G[V_x]$ is connected and has cardinality at least 3, then there exists $u' \in A$ and $u'' \in B$ such that $G[\{u, u'', u'\}] \simeq P_3$. We have u' and v are not adjacent. There are two possibilities: either $u''v \in E(G)$ or $u''v \notin E(G)$. In the first case the induced subgraph $G[\{u', u''.u, v, y\}]$ is isomorphic to P_5, and in the second case the induced subgraph $G[\{x, y, u, u', u'', v\}]$ is isomorphic to $\mathsf{G}_{6,17}$. Since both cases are forbidden, we get a contradiction. $\square$

Previous Claims suggest to divide the vertex set $V_{y,z}$ in three subsets:

- $V_{y,z}^A$, the vertices in $V_{y,z}$ that are adjacent with each vertex in A,
- $V_{y,z}^B$, the vertices in $V_{y,z}$ that are adjacent with each vertex in B, and
- $V_{y,z}^{AB}$, the vertices in $V_{y,z}$ that are adjacent with each vertex in $A \cup B$.

In what follows, we assume $|A| \geq |B|$.

Claim 2.70. *The cardinality of the sets $V_{y,z}^A$ and $V_{y,z}^B$ is no more than 1.*

Proof. Suppose there exist $v, v' \in V_{y,z}^A$. Let $u \in A$, and $u' \in B$. Since $G[\{u, u', v, v', x, y\}]$ is isomorphic to $\mathsf{G}_{6,15}$, which is forbidden, then we get a contradiction. The case $V_{y,z}^B$ is similar. $\square$

Claim 2.71. *If $|B| \geq 2$ and $V_{y,z} \neq \emptyset$, then one of the sets $V_{y,z}^B$ or $V_{y,z}^{AB}$ is empty.*

Proof. Suppose $v \in V_{y,z}^B$ and $v' \in V_{y,z}^{AB}$. Let $u, u' \in B$, and $u'' \in A$. Thus $G[\{v, v', u, u', u'', x, y\}] \simeq \mathsf{G}_{7,13}$. Which is impossible. $\square$

Since $|A| \geq 2$, then by applying previous Claim to A, one of the sets $V_{y,z}^A$ or $V_{y,z}^{AB}$ is empty.

Thus the possible cases we have are the following:

(a) $V_{y,z} = \emptyset$,
(b) $V_{y,z}^B \cup V_{y,z}^{AB} = \emptyset$ and $|V_{y,z}^A| = 1$,
(c) $V_{y,z}^A \cup V_{y,z}^{AB} = \emptyset$ and $|V_{y,z}^B| = 1$,

(d) $V_{y,z}^{AB} = \emptyset$ and $|V_{y,z}^{A}| = |V_{y,z}^{B}| = 1$,
(e) $V_{y,z}^{A} = \emptyset$, $|B| = 1$, $|V_{y,z}^{B}| = 1$ and $|V_{y,z}^{AB}| \geq 1$, and
(f) $V_{y,z}^{A} \cup V_{y,z}^{B} = \emptyset$ and $|V_{y,z}^{AB}| \geq 1$.

Now we describe $V_\emptyset$, that is, the set of vertices not adjacent with any vertex in $\{a, b, c\}$. Let $w \in V_\emptyset$. The vertex w is adjacent with a vertex in $V_x \cup V_{y,z}$, because otherwise the shortest path from w to $\{x, y\}$ would contains the graph P_5 as induced subgraph.

CLAIM 2.72. *Let $w \in V_\emptyset$. If w is adjacent with a vertex in V_x, then w is adjacent with each vertex in the parts of the partition (A, B) with cardinality greater or equal to 2.*

PROOF. Let $v \in V_x$ be a vertex adjacent with w. Suppose $v \in B$. we will prove two things: (1) if $|B| \geq 2$, then w is adjacent with each vertex in B, and (2) w is adjacent with each vertex in A.

Let us consider case when $|B| \geq 2$. We will see that w is adjacent with each vertex in B. Suppose there is a vertex $v' \in B$ not adjacent with w. Take $u \in A$. Thus there are two possibilities: either u and w are adjacent or not. The case $uw \in E(G)$ is impossible because $G[\{u, v, v', w, x, y\}] \simeq \mathsf{G}_{6,10}$, which is forbidden. Meanwhile, the case $uw \notin E(G)$ is impossible because $G[\{u, v, v', w, x, y\}] \simeq \mathsf{G}_{6,14}$, which is forbidden. Thus w is adjacent with v', and therefore w is adjacent with each vertex in B.

Now we see that w is adjacent with each vertex in A. Note that in this case $|B|$ may be equal to 1. Suppose w is not adjacent with any vertex in A. Let $u, u' \in A$. Since the induced subgraph $G[\{w, v, u, u', x, y\}]$ is isomorphic to the forbidden graph $\mathsf{G}_{6,5}$, then we get a contradiction. Thus w is adjacent with a vertex in A. Now applying the previous case (1) to A, we get that w is adjacent with each vertex in A. □

Next Claim show us what happens in the case when $|B| = 1$.

CLAIM 2.73. *If $|B| = 1$ and $E(V_\emptyset, V_x) \neq \emptyset$, then only one of the edges sets $E(V_\emptyset, V_x)$ or $E(V_\emptyset, A)$ induces a complete bipartite graph.*

PROOF. Let $w, w' \in V_\emptyset$, $u \in B$, and $u' \in A$. By Claim 2.72, vertices w and w' are adjacent with each vertex in A. Suppose w is adjacent with u, and w' is not adjacent with u. There are two possibilities: either w and w' are adjacent or not. If $ww' \in E(G)$, then the induced subgraph $G[w, w', u, x, y]$ is isomorphic to P_5; which is impossible. On the other hand, if $ww' \notin E(G)$, then $G[\{w, w', u, u', x, y\}]$ is isomorphic to $\mathsf{G}_{6,10}$, which is forbidden. So we get a contradiction, and the result follows. □

CLAIM 2.74. *Let $w \in V_\emptyset$. If w is adjacent with $v \in V_{y,z}$, then w is adjacent with each vertex in the parts of partition (A, B) with cardinality greater or equal to 2. Moreover, if $v \in V_{y,z}^{A} \cup V_{y,z}^{AB}$ and $|B| = 1$, then w is adjacent with the unique vertex in B.*

PROOF. Let $w \in V_\emptyset$ and $v \in V_{y,z}$ such that w and v are adjacent. By Claim 2.68, $E(v, V_x) \neq \emptyset$, and therefore there are three cases: $v \in V_{y,z}^{A}$, $v \in V_{y,z}^{AB}$, or $v \in V_{y,z}^{B}$.

Suppose $v \in V_{y,z}^{A}$. If $E(w, V_x) = \emptyset$, then by taking $u, u \in A$, the forbidden induced subgraph $G[W \cup \{w, v, u, u'\}] \simeq \mathsf{G}_{7,2}$ would appear and we get a contradiction. Thus $E(w, V_x) \neq \emptyset$. By Claim 2.72, w is adjacent with each vertex in the parts of V_x with cardinality greater or equal to 2. When $|B| = 1$, take $u'' \in B$. If $wu'' \notin E(G)$, then $G[\{w, v, y, x, u''\}]$ would be isomorphic to the forbidden graph P_5. Therefore, $wu'' \in E(G)$.

Suppose $v \in V_{y,z}^{AB}$. In a similar way than in previous case we get that $E(w, V_x) \neq \emptyset$, and v is adjacent with each vertex in the parts of the partition (A, B) of cardinality greater or equal to 2. In the case when $|B| = 1$, suppose $u \in B$ with $wu \notin E(G)$. Take $u' \in A$, we know that $u'w \in E(G)$. Since $G[\{w, u, u', x, y, v\}] \simeq \mathsf{G}_{6,24}$, then we get a contradiction. Thus $wu \in E(G)$.

Now suppose $v \in V_{y,z}^{B}$. It is easy to see that v is adjacent with each vertex in A, because otherwise P_5 would appear. Therefore $E(w, V_x) \neq \emptyset$, and by Claim 2.72, we have that w is adjacent with each vertex in the parts of the partition (A, B) of cardinality greater or equal to 2. $\square$

Claim 2.75. *If $E(V_\emptyset, B) \neq \emptyset$, then $E(V_\emptyset, V_{y,z}^{AB}) = \emptyset$. Moreover, if $E(V_\emptyset, B) \neq \emptyset$ and $E(V_\emptyset, V_{y,z}^{B}) \neq \emptyset$, then $V_{y,z}^{AB} = \emptyset$.*

Proof. Let $w \in V_\emptyset$ and $v \in V_{y,z}^{AB}$. Suppose w and v are adjacent. Take $u_1 \in A$ and $u_2 \in B$. Since v is in $V_{y,z}^{AB}$ and u_1 is adjacent with u_2, then $G[\{v, u_1, u_2\}]$ is isomorphic to K_3. On the other hand, by Claims 2.72 and 2.73, w is adjacent with each vertex in $A \cup B$. Thus the induced subgraph $G[\{w, v, u_1, u_2\}]$ is isomorphic to K_4 that is forbidden, and therefore w cannot be adjacent with v.

Let w is adjacent with $u_2 \in B$ and $v' \in V_{y,z}^{B}$, and suppose there exists a vertex $v \in V_{y,z}^{AB}$. Since w is not adjacent with v, then $G[\{w, v, v', u_2, x, y\}]$ is isomorphic to $\mathsf{G}_{6,15}$; which it is not possible. $\square$

Claim 2.76. *Let $w \in V_\emptyset$ and $v \in V_{y,z}$. If w and v are adjacent, then each vertex in $V_\emptyset$ is adjacent with v.*

Proof. Suppose $w' \in V_\emptyset$ such that w' is not adjacent with v. By Claims 2.72 and 2.73, both vertices w and w' are adjacent with each vertex in A. Let $u \in A$. There are four cases obtained by the combinations of the following possible cases: either $ww' \in E(G)$ or $ww' \notin E(G)$, and either $v \in V_{y,z}^{A}$ or $v \in V_{y,z}^{B}$. When $ww' \in E(G)$ and $v \in V_{y,z}^{B}$, the induced subgraph $G[\{w', w, v, y, x\}]$ is isomorphic to P_5; which is not possible. When $ww' \notin E(G)$ and $v \in V_{y,z}^{B}$, the induced subgraph $G[\{w', u, w, v, y\}]$ is isomorphic to P_5; which is not possible. When $ww' \in E(G)$ and $v \in V_{y,z}^{A}$, the induced subgraph $G[\{w', w, v, y, x\}]$ is isomorphic to P_5; which is not possible. Finally, when $ww' \notin E(G)$ and $v \in V_{y,z}^{A}$, the induced subgraph $G[\{w', w, x, y, v, u\}]$ is isomorphic to $\mathsf{G}_{6,11}$; which is not possible. Thus w' is adjacent with v. $\square$

Claim 2.77. *Let $v \in V_{y,z}^{A}$ and $v' \in V_{y,z}^{B}$. If $w \in V_\emptyset$ is adjacent with v or v', then w is adjacent with both v and v'.*

Proof. First suppose w is adjacent with v and not with v'. Let $u \in A$. By Claim 2.74, w is adjacent with a vertex $u \in A$. Thus $G[\{w, u, x, y, v'\}]$ is isomorphic to P_5 that is a contradiction. Therefore w and v' are adjacent. Now suppose w is adjacent with v' and not with v. Let $u \in B$. If u is adjacent with w, then a P_5 is obtained in a similar way than previous case. Thus assume u is not adjacent with w. Then $G[\{w, x, y, u, v, v'\}]$ is isomorphic to $\mathsf{G}_{6,9}$; which is impossible. Therefore, w and v are adjacent. $\square$

Claim 2.78. *If $V_{y,z}^{AB} \neq \emptyset$, then $E(V_\emptyset, B) = \emptyset$.*

Proof. Let $u_1, u_2 \in A$, $u_3 \in B$, $w \in V_\emptyset$ and $v \in V_{y,z}^{AB}$. By Claim 2.72, w is adjacent with u_1 and u_2. Suppose w is adjacent with u_3. By Claim 2.75, the vertices w and v are not adjacent. Thus $G[\{x, y, u_1, u_2, u_3, v, w\}]$ is isomorphic to $\mathsf{G}_{7,10}$, which is a contradiction. Therefore, $E(V_\emptyset, B) = \emptyset$. $\square$

Thus applying previous Claims to cases (a) - (f), we obtain the following possibilities:

(1) $V_{y,z} = \emptyset$, $|B| = 1$, and $E(V_\emptyset, A)$ induces a complete bipartite graph,
(2) $V_{y,z} = \emptyset$, $|B| \geq 1$, and $E(V_\emptyset, V_x)$ induces a complete bipartite graph,
(3) $V^B_{y,z} \cup V^{AB}_{y,z} = \emptyset$, $|V^A_{y,z}| = 1$, $|B| \geq 1$ and $E(V_\emptyset, V_x \cup V_{y,z})$ induces a complete bipartite graph,
(4) $V^A_{y,z} \cup V^{AB}_{y,z} = \emptyset$, $|V^B_{y,z}| = 1$, $|B| = 1$ and $E(V_\emptyset, A \cup V_{y,z})$ induces a complete bipartite graph,
(5) $V^A_{y,z} \cup V^{AB}_{y,z} = \emptyset$, $|V^B_{y,z}| = 1$, $|B| \geq 1$ and $E(V_\emptyset, V_x \cup V_{y,z})$ induces a complete bipartite graph,
(6) $V^{AB}_{y,z} = \emptyset$, $|V^A_{y,z}| = |V^B_{y,z}| = 1$, $|B| \geq 1$ and $E(V_\emptyset, V_x \cup V_{y,z})$ induces a complete bipartite graph,
(7) $V^A_{y,z} \cup V^B_{y,z} = \emptyset$, $|V^{AB}_{y,z}| \geq 1$, $|B| = 1$ and $E(V_\emptyset, A)$ induces a complete bipartite graph,
(8) $V^A_{y,z} \cup V^B_{y,z} = \emptyset$, $|V^{AB}_{y,z}| \geq 1$, $|V^B_{y,z}| = 1$, $|B| = 1$ and $E(V_\emptyset, A \cup V^B_{y,z})$ induces a complete bipartite graph.

With a similar argument as in Claim 2.67, we obtain that $V_\emptyset$ is either trivial or K_2. In cases (1), (4), (7) and (8), $V_\emptyset$ cannot be T_n with $n \geq 2$, since taking $w, w' \in V_\emptyset$, $u \in A$, $u' \in B$, then $G[\{w, w', u, u', x, y\}] \simeq \mathsf{G}_{6,2}$. On the other hand, since $\omega(G) = 3$, then $V_\emptyset$ is not isomorphic to K_2 in the cases (2), (3), (5) and (6). It is not difficult to see that in each case G is isomorphic to an induced subgraph of a graph in $\mathcal{F}^1_1$.

7.0.4. **Case when V_x induces $K_1 + K_2$ or $2K_2$.** Through this case we assume that $V_x = \{u_1, u_2, u_3, u_4\}$ such that $u_1u_2, u_3u_4 \in E(G)$. That is, $G[\{u_1, u_2, u_3, u_4\}] \simeq 2K_2$. Let $A = \{u_1, u_2\}$ and $B = \{u_3, u_4\}$. The following discussion also applies when one of the vertex set A or B has cardinality 1.

Claim 2.79. *If $v \in V_{y,z}$, then $E(v, V_x) \neq \emptyset$. Moreover, if $v \in V_{y,z}$, then v is adjacent with each vertex in one of the following sets A, B or V_x.*

Proof. Suppose there is no edge joining v and a vertex in V_x. Since $G[\{u_1, u_2, u_3, x, y, v\}] \simeq \mathsf{G}_{6,6}$ is a forbidden induced subgraph, then a contradiction is obtained. Then v is adjacent with some vertex in V_x.

Suppose v is adjacent with one of the vertices in A. We will prove that v cannot be adjacent only with one vertex of A, say u_1. Thus suppose v is adjacent with u_1, and v is not adjacent with u_2, u_3 and u_4. Since $G[\{u_1, u_2, u_3, v, x, y\}] \simeq \mathsf{G}_{6,11}$, then v is adjacent with u_2 or a vertex in B, say u_3. If v is adjacent with u_2, then we are done. So we assume v is adjacent with u_1, but v is not adjacent with u_2. This is not possible because $G[\{u_1, u_2, u_3, v, y, z\}] \simeq \mathsf{G}_{6,10}$. Therefore, either v is adjacent with both u_1 and u_2, or v is not adjacent with neither of u_1 and u_2 ☐

Claim 2.80. *If $|V_{y,z}| \geq 2$, then each vertex in $V_{y,z}$ is adjacent with each vertex of only one of the following sets A, B or V_x.*

Proof. Let $v, v' \in V_{y,z}$. Consider the following cases:

(a) v is adjacent with each vertex in A and v' is adjacent with each vertex in B,
(b) v is adjacent with each vertex in V_x and v' is adjacent with each vertex in B, and
(c) v is adjacent with each vertex in V_x and v' is adjacent with each vertex in A.

Case (a) is impossible because $G[\{u_1, v, y, v', u_3\}] \simeq P_5$, which is forbidden. On the other hand, cases (b) and (c) are not allowed because $G[\{a, b, c, v, v', u_1, u_3\}] \simeq \mathsf{G}_{7,9}$; which is forbidden. Thus, the result follows. ☐

Thus $E(V_x, V_{y,z})$ satisfies only one of the following three cases:

(1) $E(V_{y,z}, A)$ induces a complete bipartite graph,

(2) $E(V_{y,z}, B)$ induces a complete bipartite graph, or
(3) $E(V_{y,z}, V_x)$ induces a complete bipartite graph.

Now we describe $V_\emptyset$, the set of vertices not adjacent with any vertex in $\{a, b, c\}$. Let $w \in V_\emptyset$. The vertex w is adjacent with a vertex in $V_x \cup V_{y,z}$, because otherwise the shortest path from w to $\{x, y\}$ would contains the graph P_5 as induced subgraph.

Claim 2.81. *If $w \in V_\emptyset$ is adjacent with a vertex in V_x, then w is adjacent with each vertex in V_x.*

Proof. Suppose w is adjacent with u_1 and not with u_3. Since $G[W \cup \{w, u_1, u_3\}] \simeq \mathsf{G}_{7,4}$ is forbidden, w also must be adjacent with u_3. In a similar way we get the opposite case and it turns out the result. □

Consider Cases (1) and (2). The following arguments works on both cases. First note that if there exists $w \in V_\emptyset$ adjacent with $v \in V_{y,z}$, then w is adjacent with each vertex in V_x. The reason is the following. Suppose w is not adjacent with any vertex in V_x. Since v is not adjacent with a vertex in V_x, say u, and the vertices w and u are not adjacent, then we have that $G[\{w, v, y, x, u\}] \simeq \mathsf{P}_5$; which is a contradiction. Then w and u are adjacent. And by Claim 2.81, the edge set $E(w, V_x)$ induces a complete bipartite graph. On the other hand, we can prove in a similar way that if there exists $w \in V_\emptyset$ adjacent with each vertex in V_x, then w is adjacent with each vertex in $V_{y,z}$. Therefore, each vertex in $V_\emptyset$ is adjacent with each vertex in $V_x \cup V_{y,z}$. Furthermore, the set $V_\emptyset$ is a stable set. It is because if $w, w' \in V_\emptyset$ were adjacent, then by taking $u \in V_x$ and $v \in V_{y,z}$ such that u and v are adjacent, the induced subgraph $G[\{w, w', u, v, x, y\}]$ would be isomorphic to $\mathsf{G}_{6,20}$; which is impossible. These cases correspond to a graph isomorphic to an induced subgraph of a graph in $\mathcal{F}_1^1$.

Now let us consider case (3).

Claim 2.82. *If $w \in V_\emptyset$ is adjacent with a vertex in $V_{y,z}$, then each vertex in $V_\emptyset$ is adjacent with each vertex in $V_x \cup V_{y,z}$.*

Proof. Let $w \in V_\emptyset$ and $v \in V_{y,z}$ such that $wv \in E(G)$. Suppose w is not adjacent with any vertex in V_x. Take $u_2, u_3 \in V_x$ such that $u_2u_3 \in E(G)$. Since the induced subgraph $G[\{a, b, c, w, u_2, u_3, v\}]$ is isomorphic to $\mathsf{G}_{7,2}$, then we get a contradiction; and w is adjacent with a vertex in V_x. Thus by Claim 2.81, w is adjacent with each vertex in V_x.

Suppose that there exists $v' \in V_{y,z}$ such that w is not adjacent with v'. Since the vertex set $\{w, u_1, x, y, v, v'\}$ would induce $\mathsf{G}_{6,15}$, this does not occur. Therefore, w is adjacent with each vertex in $V_x \cup V_{y,z}$.

Suppose there is another vertex $w' \in V_\emptyset$. By the above argument, if w' is adjacent with a vertex in $V_{y,z}$, then it must be adjacent with each vertex in $V_x \cup V_{y,z}$. And we are done. On the other hand, by Claim 2.81, if w' is adjacent with a vertex in V_x, then w' must be adjacent with each vertex in V_x. So suppose w' is adjacent with each vertex in V_x, but not adjacent with each vertex in $V_{y,z}$. Then there are two possibilities: either $ww' \notin E(G)$ or $ww' \in E(G)$. In the first case $G[\{x, y, u_1, v, w, w'\}] \simeq \mathsf{G}_{6,11}$ and in the second case $G[\{x, y, v, w, w'\}] \simeq \mathsf{P}_5$. Since both graphs are forbidden, then w' is adjacent with v. And therefore w' is adjacent with each vertex in $V_x \cup V_{y,z}$. □

By Claims 2.81 and 2.82, we obtain that there are two possible cases: either each vertex in $V_\emptyset$ is adjacent with each vertex in $V_x \cup V_{y,z}$, or each vertex in $V_\emptyset$ is adjacent only with each vertex in V_x. Consider first case. If there exists a vertex $w \in V_\emptyset$, then w adjacent with the vertices $v \in V_{y,z}$, and $u_1, u_2 \in V_x$. Thus $G[\{w, v, u_1, u_2\}]$ is isomorphic to K_4 that is not allowed, then $V_\emptyset$ is empty. Now consider second case. Let $w \in V_\emptyset$ and $v \in V_{y,z}$. Thus w is adjacent with u_1, u_2 and u_3. Since $G[\{w, u_1, u_2, u_3, v, x, y\}]$ is isomorphic to $\mathsf{G}_{7,9}$, we get a contradiction. Then $V_\emptyset = \emptyset$. The graph in this case is isomorphic to an induced subgraph of a graph in $\mathcal{F}_1^1$.

7.0.5. **Cases when $G[V_a \cup V_b \cup V_c] = V_x \vee (V_y + V_z)$ is one of the following graphs: $K_1 \vee 2K_1$, $K_1 \vee (K_1 + K_2)$, $K_1 \vee 2K_2$, or $K_2 \vee 2K_1$.** For the sake of clarity, we suppose $E(V_a, V_b)$ and $E(V_a, V_c)$ induce a complete bipartite graph, and $E(V_b, V_c)$ is empty. Now we are going to obtain some claims that describe the edge sets joining V_x and $V_{y,z}$.

CLAIM 2.83. *Let $x, y \in \{a, b, c\}$. If $E(V_x, V_y)$ is not empty, then $E(V_x, V_{xy})$ and $E(V_y, V_{xy})$ are empty.*

PROOF. Let $v_x \in V_x$, $v_y \in V_y$ and $v_{xy} \in V_{x,y}$. Suppose v_{xy} is adjacent with both v_x and v_y. Then $G[\{a, b, c, v_{xy}, v_x, v_y\}] \simeq \mathsf{G}_{6,24}$; which is a contradiction. Now suppose v_{xy} is adjacent with v_x and not with v_y. In this case $G[\{a, b, c, v_{xy}, v_x, v_y\}] \simeq \mathsf{G}_{6,17}$; which is impossible. And therefore result turns out. $\square$

Claim 2.83 implies that $E(V_{a,b}, V_a \cup V_b) = \emptyset$ and $E(V_{a,c}, V_a \cup V_c) = \emptyset$.

CLAIM 2.84. *Let $x, y \in \{a, b, c\}$. If $V_{x,y} \neq \emptyset$ and $E(V_x, V_y) = \emptyset$, then each edge set $E(V_x, V_{x,y})$ and $E(V_y, V_{x,y})$ induces a complete bipartite graph.*

PROOF. Let $v_x \in V_x$, $v_y \in V_y$ and $v_{xy} \in V_{x,y}$. Suppose v_{xy} is not adjacent with both v_x and v_y. Then $G[\{a, b, c, v_{xy}, v_x, v_y\}] \simeq \mathsf{G}_{6,5}$; which is a contradiction. Finally, suppose v_{xy} is adjacent with v_x and not with v_y. Since $G[\{a, b, c, v_{x,y}, v_x, v_y\}] \simeq \mathsf{G}_{6,14}$ is forbidden, then we get a contradiction. And the result turns out. $\square$

Claim 2.84 implies that each vertex in $V_{b,c}$ is adjacent with each vertex in $V_b \cup V_c$.

CLAIM 2.85. *If $V_{b,c} \neq \emptyset$, then $E(V_a, V_{b,c})$ is empty.*

PROOF. Let $v_a \in V_a$, $v_b \in V_b$ and $v_{bc} \in V_{b,c}$. Suppose v_{bc} is adjacent with v_a. Then $G[\{a, b, c, v_a, v_b, v_{bc}\}]$ is isomorphic to $\mathsf{G}_{6,25}$ that is forbidden. Therefore, there is no edge joining a vertex in V_a with a vertex in $V_{b,c}$. $\square$

CLAIM 2.86. *Let $x \in \{b, c\}$. If $V_{a,x} \neq \emptyset$, then $E(V_{a,x}, V_x)$ induces a complete bipartite graph.*

PROOF. Let $v_b \in V_b$, $v_c \in V_c$ and $v_{ax} \in V_{a,x}$. Suppose v_{ax} is not adjacent with v_x. Since $G[\{a, b, c, v_{ax}, v_b, v_c\}]$ is isomorphic to $\mathsf{G}_{6,10}$ that is forbidden, then each vertex in V_x is adjacent with each vertex in $V_{a,x}$. $\square$

CLAIM 2.87. *The edge set $E(V_{a,b}, V_{a,c})$ induces a complete bipartite graph.*

PROOF. Let $v_a \in V_a$, $v_b \in V_b$, $v_{ab} \in V_{a,b}$, and $v_{ac} \in V_{a,c}$. We know that v_a is not adjacent with both v_{ab} and v_{ac}, and v_b is adjacent with v_{ac}, but v_b is not adjacent with v_{ab}. Suppose v_{ab} and v_{ac} are not adjacent. Then $G[\{b, c, v_a, v_b, v_{ab}, v_{ac}, \}] \simeq \mathsf{G}_{6,9}$; which is impossible. And therefore, each vertex in $V_{a,b}$ is adjacent with each vertex in $V_{a,c}$. $\square$

CLAIM 2.88. *Let $x \in \{b, c\}$ The edge set $E(V_{a,x}, V_{b,c})$ induces a complete bipartite graph.*

PROOF. Let $v_b \in V_b$, $v_c \in V_c$, $v_{ax} \in V_{a,x}$, and $v_{bc} \in V_{b,c}$. We know that $v_b v_{bc}, v_{bc} v_c, v_c v_{ab}, v_b v_{ac} \in E(G)$ and $v_x v_{ax}, v_b v_c \notin E(G)$. Suppose v_{ax} and v_{bc} are not adjacent. Then $G[\{b, c, v_b, v_c, v_{ax}, v_{bc}, \}] \simeq \mathsf{G}_{6,14}$; which is impossible. Therefore each vertex in $V_{a,x}$ is adjacent with each vertex in $V_{b,c}$. $\square$

Now we analyze $V_\emptyset$.

CLAIM 2.89. *If $w \in V_\emptyset$, then w cannot be adjacent with any vertex in $V_a \cup V_b \cup V_c$.*

PROOF. Let $x \in \{b,c\}$, $y \in \{b,c\} - x$, $v_a \in V_a$ and $v_x \in V_x$. Suppose $w \in V_\emptyset$ such that $E(w, V_a \cup V_b \cup V_c) \neq \emptyset$. Consider the following cases:

(a) w is adjacent only with v_a,

(b) w is adjacent only with v_x, or

(c) w is adjacent with both v_a and v_x.

Cases (a) and (b) are impossible, because G would have P_5 as induced subgraph obtained by $G[\{w, v_a, v_x, x, y\}]$ and $G[\{w, v_x, v_a, a, y\}]$, respectively. Finally in case (c), the induced subgraph $G[\{a, b, c, w, v_a, v_x\}]$ is isomorphic to $\mathsf{G}_{6,21}$; which is impossible. □

CLAIM 2.90. *There exists no vertex in $V_\emptyset$ adjacent with a vertex in $V_{b,c}$.*

PROOF. Let $w \in V_\emptyset$, $v_a \in V_a$, $v_b \in V_b$, $v_c \in V_c$, and $v_{bc} \in V_{b,c}$. Suppose w is adjacent with v_{bc}. A contradiction is obtained since $G[\{w, v_a, v_b, v_c, v_{bc}, w\}] \simeq \mathsf{G}_{6,11}$. Thus w is not adjacent with any vertex in $V_{b,c}$ □

CLAIM 2.91. *There exists no vertex in $V_\emptyset$ adjacent with a vertex in $V_{a,b} \cup V_{a,c}$.*

PROOF. Let $x \in \{b,c\}$, $y \in \{b,c\} - x$, $v_a \in V_a$, $v_b \in V_b$ and $v_{ax} \in V_{a,x}$. Suppose w is adjacent with v_{ax}. Since the induced subgraph $G[\{w, v_{ax}, v_a, v_b, v_c\}]$ is isomorphic to $\mathsf{G}_{6,5}$, then we get a contradiction. And then the result follows. □

Claims 2.89, 2.90 and 2.91 imply that no vertex in $V_\emptyset$ is adjacent with a vertex in $G \setminus V_\emptyset$, which implies that $V_\emptyset = \emptyset$. Thus the graph is isomorphic to an induced subgraph of a graph in $\mathcal{F}_1^1$.

7.0.6. **Case $G[V_a \cup V_b \cup V_c] = K_{1,1,1}$, where each vertex set $V_x = K_1$.** Let $V_a = \{v_a\}$, $V_b = \{v_b\}$, and $V_c = \{v_c\}$. By Claim 2.83, the edge sets $E(V_{x,y}, V_x)$ and $E(V_{x,y}, V_y)$ are empty for $x, y \in \{a,b,c\}$. By Claim 2.85, the edge set $E(V_{x,y}, V_z)$ is empty for $x, y, z \in \{a,b,c\}$. Now let $v_{xy} \in V_{x,y}$. Since $G[\{v_y, v_z, z, x, v_{x,y}\}]$ is isomorphic to P_5 which is a forbidden, then V_{xy} is empty for each pair $x, y \in \{a,b,c\}$. On the other hand, by Claim 2.89 the edge set $V_\emptyset$ is empty. This graph is isomorphic to G_1, see Figure 5.i.

7.0.7. **Case $V_z = \emptyset$ and $G[V_x \cup V_y] = V_x + V_y$, where $V_x = K_m$, $V_y = K_n$ and $m, n \in \{1,2\}$.** Without loss of generality, suppose $V_c = \emptyset$, and $E(V_a, V_b)$ is empty. By Claim 2.84, each vertex in $V_{a,b}$ is adjacent with each vertex in $V_a \cup V_b$.

CLAIM 2.92. *Let $x \in \{a,b\}$. If $V_{x,c} \neq \emptyset$, then $E(V_x, V_{x,c}) = \emptyset$.*

PROOF. Let $y \in \{a,b\} - x$, $v_{xc} \in V_{x,c}$, $v_a \in V_a$ and $v_b \in V_b$. Suppose v_x and v_{xc} are adjacent. There are two possible cases: either v_y and v_{xc} are adjacent or not. Since in the first case $G[\{a, b, c, v_a, v_b, v_{xc}\}] \simeq \mathsf{G}_{6,17}$ and in the second case $G[\{v_a, v_b, y, v_{xc}, c\}]$ is isomorphic to P_5, then we get a contradiction. Thus v_x and v_{xc} are not adjacent. □

By Claim 2.86, each vertex in $V_{a,c}$ is adjacent with each vertex in V_b, and each vertex in $V_{b,c}$ is adjacent with each vertex in V_a.

CLAIM 2.93. *Each vertex in $V_{a,b}$ is adjacent with each vertex in $V_{a,c} \cup V_{b,c}$.*

PROOF. Let $v_{ab} \in V_{a,b}$, $v_a \in V_a$ and $v_b \in V_b$. Suppose there exists $v_{xc} \in V_{xc}$ with $x \in \{a,b\}$ such that $v_{ab}v_{xc} \notin E(G)$. Since $G[\{v_a, v_b, v_{ab}, v_{xc}, c\}]$ is isomorphic to P_5, then we get a contradiction; and the vertices v_{ab} and v_{xc} are adjacent. And the result turns out. □

CLAIM 2.94. *Each vertex in $V_{a,c}$ is adjacent with each vertex in $V_{b,c}$.*

PROOF. Suppose there are $v_{ac} \in V_{a,c}$ and $v_{bc} \in V_{b,c}$ such that $v_{bc}v_{ac} \notin E(G)$. Let $v_a \in V_a$ and $v_b \in V_b$ Since $G[\{v_b, v_{ac}, c, v_{bc}, v_a\}] \simeq P_5$, then we get a contradiction. $\square$

Now we describe the vertex set $V_\emptyset$, that is, the set of vertices that are not adjacent with any vertex in $\{a, b, c\}$.

CLAIM 2.95. *If $w \in V_\emptyset$, then w cannot be adjacent with any vertex in $V_a \cup V_b$.*

PROOF. Let $v_a \in V_a$ and $v_b \in V_b$. Suppose $w \in V_\emptyset$ such that $E(w, V_a \cup V_b) \neq \emptyset$. Consider the following cases:
(a) w is adjacent only with v_a, or
(b) w is adjacent with both v_a and v_b.
Cases (a) and (b) are impossible, because G would have P_5 as induced subgraph obtained by $G[\{w, v_b, b, a, v_a\}]$ and $G[\{v_a, w, v_b, b, c\}]$, respectively. Thus w is not adjacent with any vertex in $V_a \cup V_b$. $\square$

CLAIM 2.96. *There is no vertex $w \in V_\emptyset$ adjacent with a vertex in $V_{a,b}$.*

PROOF. Suppose $w \in V_\emptyset$ is adjacent with $v_{ab} \in V_{a,b}$. Let $v_a \in V_a$ and $v_b \in V_b$. Since $G[\{v_a, v_b, v_{a,b}, w, a, c\}]$ is isomorphic to $\mathsf{G}_{6,2}$, then we get a contradiction and w and v_{ab} are not adjacent. $\square$

CLAIM 2.97. *There is no vertex $w \in V_\emptyset$ adjacent with a vertex in $V_{a,c} \cup V_{b,c}$.*

PROOF. Let $x \in \{a, b\}$, $v_{xc} \in V_{x,c}$, $y \in \{a, b\} - x$ and $v_y \in V_y$. Suppose $w \in V_\emptyset$ is adjacent with v_{xc}. Since $G[\{v_y, y, c, v_{xc}, w\}]$ is isomorphic to P_5, then we get a contradiction and w is not adjacent with v_{xc}. $\square$

Thus there is no edge between W and $G \setminus W$, and therefore W is empty. Therefore, the graph is isomorphic to an induced subgraph of a graph in $\mathcal{F}_1^1$.

7.0.8. **Case $V_z = \emptyset$ and $G[V_x \cup V_y] = V_x \vee V_y$, where $V_x = K_1$, $V_y = K_m$ and $m \in \{1, 2\}$.** Without loss of generality, suppose $V_c = \emptyset$, and $E(V_a, V_b)$ is complete. By Claim 2.83, $E(V_{a,b}, V_x) = \emptyset$ for $x \in \{a, b\}$.

CLAIM 2.98. *Let $x \in \{a, b\}$. Either $E(V_{x,c}, V_a)$ induces a complete bipartite graph and $E(V_{x,c}, V_b) = \emptyset$, or $E(V_{x,c}, V_a) = \emptyset$ and $E(V_{x,c}, V_b)$ induces a complete bipartite graph.*

PROOF. Let $v_a \in V_a$, $v_b \in V_b$, $v_{xc} \in V_{x,c}$ and $y \in \{a, b\} - x$. First note that v_{xc} cannot be not adjacent with v_a and v_b at the same time, because otherwise $G[\{v_a, v_b, y, c, v_{xc}\}]$ would be isomorphic to P_5. Also v_{xc} cannot be adjacent with v_a and v_b at the same time, because otherwise $G[\{a, b, c, v_a, v_b, v_{xc}\}]$ would be isomorphic to $\mathsf{G}_{6,25}$. Thus v_{xc} is adjacent only with one vertex either v_a or v_b.

Suppose v_{xc} is adjacent with v_y. Let $v'_y \in V_y$ and $v'_{xc} \in V_{x,c}$. If v_{xc} is not adjacent with v'_y, then $G[\{v_y, v'_y, x, c, v_x, v_{xc}\}]$ is isomorphic to $\mathsf{G}_{6,21}$. Thus $E(V_y, v_{xc})$ induces a complete bipartite graphs. On the other hand, if v_y and v'_{xc} are not adjacent, then $G[\{v_a, v_b, v_{xc}, c, v'_{xc}\}]$ is isomorphic to P_5. Then $E(V_{x,c}, V_y)$ induces a complete bipartite graph.

Suppose v_{xc} is adjacent with v_x. Let $v'_x \in V_x$ and $v'_{xc} \in V_{x,c}$. If v'_x and v_{xc} are not adjacent, then $G[\{v_a, v_b, v'_x, x, c, v_{xc}\}]$ is isomorphic to $\mathsf{G}_{6,22}$. Thus $E(V_x, v_{xc})$ induces a complete bipartite graph. On the other hand, if v'_{xc} and v_x are not adjacent, then $G[\{v_y, v_x, v_{xc}, c, v'_{xc}\}]$ is isomorphic to P_5. Thus $E(V_{x,c}, V_x)$ induces a complete bipartite graph. $\square$

CLAIM 2.99. *If $V_{a,c}$ and $V_{b,c}$ are not empty, then each vertex in $V_{a,c} \cup V_{b,c}$ is adjacent with each vertex in either V_a or V_b .*

PROOF. Let $v_{ac} \in V_{a,c}$, $v_{bc} \in V_{b,c}$, $v_a \in V_a$ and $v_b \in V_b$.

Suppose v_{ac} is adjacent with v_a, and v_{bc} is adjacent with v_b. Then there are two cases: either v_{ac} and v_{bc} are adjacent or not. In the first case $G[\{a, b, c, v_b, v_{ac}, v_{bc}\}]$ is isomorphic to $\mathsf{G}_{6,24}$; then this case is impossible. And in the second case $G[\{a, b, v_a, v_b, v_{ac}, v_{bc}\}]$ is isomorphic to $\mathsf{G}_{6,21}$; then this case is not possible.

Suppose v_{ac} is adjacent with v_b and v_{bc} is adjacent with v_a. Then there are two cases: either v_{ac} and v_{bc} are adjacent or not. In the first case $G[\{a, b, c, v_a, v_b, v_{ac}, v_{bc}\}]$ is isomorphic to $\mathsf{G}_{7,13}$; then this case is impossible. And in the second case $G[\{a, b, c, v_b, v_{ac}, v_{bc}\}]$ is isomorphic to $\mathsf{G}_{6,17}$; then this case is not possible.

Thus v_{ac} and v_{bc} are adjacent with the same vertex: either v_a or v_b. And the result follows. $\square$

CLAIM 2.100. *If $V_{a,c}$ and $V_{b,c}$ are not empty, then the set $E(V_{ac}, V_{bc})$ induces a complete bipartite graph.*

PROOF. Let $x \in \{a, b\}$ and $y \in \{a, b\} - x$, $v_{ac} \in V_{a,c}$, $v_{bc} \in V_{b,c}$, $v_a \in V_a$ and $v_b \in V_b$. Suppose v_{xc} and V_{yc} are not adjacent, and v_{xc} and v_{yc} are adjacent with v_x. Since $G[\{x, y, v_x, v_y, v_{x,c}, v_{y,c}\}] \simeq \mathsf{G}_{6,15}$, we get a contradiction. And then $E(V_{ac}, V_{bc})$ induces a complete bipartite graph. $\square$

CLAIM 2.101. *Let $x \in \{a, b\}$. If $V_{x,c} \neq \emptyset$, then $E(V_{x,c}, V_{a,b})$ induces a complete bipartite graph.*

PROOF. Let $y \in \{a, b\} - x$, $v_{xc} \in V_{x,c}$, $v_{ab} \in V_{a,b}$, $v_a \in V_a$ and $v_b \in V_b$. Suppose v_{xc} and v_{ab} are not adjacent. There are two cases: either v_{xc} is adjacent with v_x or v_{xc} is adjacent with v_y. If v_{xc} is adjacent with v_x, then the induced subgraph $G[\{v_{ab}, y, v_y, v_x, v_{xc}\}]$ is isomorphic to P_5; which is a contradiction. Then this case is impossible. On the other hand, if v_{xc} is adjacent with v_y, then the induced subgraph $G[\{v_{ac}, v_a, v_b, a, b, v_{xc}\}]$ is isomorphic to $\mathsf{G}_{6,9}$; which is a contradiction. Thus this case is also impossible, and therefore $E(V_{x,c}, V_{a,b}$ induces a complete bipartite graph. $\square$

Now let us describe $V_\emptyset$. By Claim 2.89, there exists no vertex in $V_\emptyset$ adjacent with a vertex in $V_a \cup V_b$.

CLAIM 2.102. *There is no vertex in $V_\emptyset$ adjacent with a vertex in $V_{a,b}$.*

PROOF. Let $v_{ab} \in V_{a,b}$ and $w \in V_\emptyset$. Suppose v_{ab} and w are adjacent. Since $G[\{w, v_{a,b}, a, v_a, v_b\}]$ is isomorphic to P_5, then we get a contradiction. And therefore v_{ab} and w are not adjacent. $\square$

CLAIM 2.103. *Let $x \in \{a, b\}$. There is no vertex in $V_\emptyset$ adjacent with a vertex in $V_{x,c}$.*

PROOF. Let $y \in \{a, b\} - x$, $v_{xc} \in V_{x,c}$, $v_{ab} \in V_{a,b}$, $v_a \in V_a$ and $v_b \in V_b$. Suppose the vertex $w \in V_\emptyset$ is adjacent with v_{xc}. There are two cases: either v_{xc} is adjacent with x or v_{xc} is adjacent with y. First case is impossible since $G[\{w, v_{xc}, v_x, v_y, y\}]$ is isomorphic to P_5; which is forbidden. And second case cannot occur because $G[\{w, y, v_{xc}, c, v_a, v_b\}] \simeq \mathsf{G}_{6,9}$. Then it follows that w is adjacent with no vertex in $V_{x,c}$. $\square$

Thus by previous Claims, the vertex set $V_\emptyset$ is empty, because there is no vertex in $V_\emptyset$ adjacent with a vertex in $G \setminus V_\emptyset$. Then G is isomorphic to an induced subgraph of a graph in $\mathcal{F}_1^1$.

7.0.9. **Case $V_y \cup V_z = \emptyset$ and V_x is K_1 or K_2.** Without loss of generality, suppose $V_b = V_c = \emptyset$ and $V_a = \{u_1, u_2\}$.

CLAIM 2.104. *Let $x \in \{b, c\}$. If $V_{a,x} \neq \emptyset$, then either $E(V_a, V_{a,x})$ induces a complete bipartite graph or $E(V_a, V_{a,x})$ is empty.*

PROOF. Let $u \in V_a$. Suppose there exist $v_1, v_2 \in V_{a,x}$ such that $uv_1 \in E(G)$ and $uv_2 \notin E(G)$. Since the induced subgraph $G[\{a, b, c, u, v_1, v_2\}]$ is isomorphic to $\mathsf{G}_{6,16}$, then we get a contradiction. Thus this case is not possible and therefore u is adjacent with either each vertex in $V_{a,x}$ or no vertex in $V_{a,x}$. Now we are going to discard the possibility that $E(u_1, V_{a,x})$ induces a complete bipartite graph and $E(u_2, V_{a,x}) = \emptyset$. Suppose there is a vertex $v \in V_{a,x}$ such that $u_1v \in E(G)$ and $u_2v \notin E(G)$. Since the induced subgraph $G[\{c, b, v, u_1, u_2\}]$ is isomorphic to P_5, then we get a contradiction. Thus either $E(V_a, V_{a,x})$ induces a complete bipartite graph or $E(V_a, V_{a,x})$ is empty. □

CLAIM 2.105. *Let $u \in V_a$. If $V_{b,c} \neq \emptyset$, then $E(u, V_{b,c})$ satisfies only one of the following:*

- *it induces a complete bipartite graph,*
- *it is an empty edge set, or*
- *it induces a complete bipartite graph minus an edge.*

PROOF. Suppose there exist $v_1, v_2, v_3 \in V_{b,c}$ such that $uv_1 \in E(G)$ and $uv_2, uv_3 \notin E(G)$. Since the induced subgraph $G[\{a, b, u, v_1, v_2, v_3\}]$ is isomorphic to $\mathsf{G}_{6,3}$, then we get a contradiction. Thus this case is not possible and the result follows. □

CLAIM 2.106. *If $V_a = \{u_1, u_2\}$ and $V_{b,c} \neq \emptyset$, then $E(V_a, V_{b,c})$ satisfies one of the following:*

- *it induces a complete bipartite graph,*
- *it is an empty edge set,*
- *it induces a complete bipartite graph minus an edge,*
- *it induces a perfect matching and $|V_a| = |V_{b,c}| = 2$, or*
- *it induces a complete bipartite graph minus two edges u_1v and u_2v, where $v \in V_{b,c}$.*

PROOF. Since cases where $|V_{b,c}| \leq 2$ can be checked easily with a computer algebra system or with similar arguments to the rest of the proof, then we asume $|V_{b,c}| \geq 3$. By Claim 2.105, we only have to check the possibilities of the edge sets $E(u_1, V_{b,c})$ and $E(u_2, V_{b,c})$. The possible cases we have to discard are the following:

- $E(u_1, V_{b,c}) = \emptyset$ and $E(u_2, V_{b,c})$ induces a complete bipartite graph,
- $E(u_1, V_{b,c}) = \emptyset$ and $E(u_2, V_{b,c})$ induces a complete bipartite graph minus an edge, and
- each edge set $E(u_1, V_{b,c})$ and $E(u_2, V_{b,c})$ induces a complete bipartite graph minus an edge and the two removed edges don't share a common vertex.

Let $v_1, v_2, v_3 \in V_{b,c}$. Suppose we are in the first case. Thus u_1 is adjacent with no vertex in $V_{b,c}$, and u_2 is adjacent with each vertex in $V_{b,c}$. Since the induced subgraph $G[\{a, b, u_1, u_2, v_1, v_2\}]$ is isomorphic to $\mathsf{G}_{6,15}$, then we get a contradiction and this case is not possible. Suppose we are in the second case. Thus u_1 is adjacent with no vertex in $V_{b,c}$, and u_2 is adjacent only with each vertex in $V_{b,c} - v_1$. Since the induced subgraph $G[\{v_1, b, v_2, u_2, u_1\}]$ is isomorphic to P_5, then we get a contradiction and this case is not possible. Finally, suppose we are in the third case. Thus u_1 is adjacent with each vertex in $V_{b,c} - v_1$, and u_2 is adjacent with each vertex in $V_{b,c} - v_2$. Since the induced subgraph $G[\{a, u_1, u_2, v_1, v_2, v_3\}]$ is isomorphic to $\mathsf{G}_{6,5}$, then we get a contradiction and this case is not possible. And the result turns out. □

Claim 2.107. *If $E(V_a, V_{a,b})$ and $E(V_a, V_{a,c})$ are empty, then $E(V_{a,b}, V_{a,c})$ induces a complete bipartite graph.*

Proof. Let $v_{ab} \in V_{a,b}$ and $v_{ac} \in V_{a,c}$. Suppose v_{ab} and v_{ac} are not adjacent. Since the induced subgraph $G[\{a, b, c, u_1, v_{ab}, v_{ac}\}]$ is isomorphic to $\mathsf{G}_{6,12}$, then we get a contradiction. And thus each vertex in $V_{a,b}$ is adjacent with each vertex in $V_{a,c}$. $\square$

Claim 2.108. *If each of the edge sets $E(V_a, V_{a,b})$ and $E(V_a, V_{a,c})$ induces a complete bipartite graph, then $|V_a| = |V_{a,b}| = |V_{a,c}| = 1$ and $E(V_{a,b}, V_{a,c}) = \emptyset$.*

Proof. First we will prove that $E(V_{a,b}, V_{a,c}) = \emptyset$. Let $v_{ab} \in V_{a,b}$ and $v_{ac} \in V_{a,c}$. Suppose v_{ab} and v_{ac} are adjacent. Since the induced subgraph $G[\{a, b, c, v_{ab}, v_{ac}, v_a\}]$ is isomorphic to $\mathsf{G}_{6,26}$, then we get a contradiction. And thus $E(V_{a,b}, V_{a,c}) = \emptyset$.

Now let $x \in \{b, c\}$. Suppose $V_{a,x}$ has cardinality at least 2. Let $y \in \{b, c\} - x$, $v_{ax}, v'_{ax} \in V_{a,x}$ and $v \in V_{a,y}$. Then $v_{ax}v, v_{ax}v \notin E(G)$. Since the induced subgraph $G[\{a, y, v_{ax}, v'_{ax}, v_a, v\}]$ is isomorphic to $\mathsf{G}_{6,16}$, then we get a contradiction. And then $|V_{a,b}| = |V_{a,c}| = 1$

Finally suppose that $|V_a| = 2$. Let $v_{ab} \in V_{a,b}$ and $v_{ac} \in V_{a,c}$. Since $G[\{v_{ab}, v_{ac}, u_1, u_2, a, x\}]$ is isomorphic to $\mathsf{G}_{6,23}$, then we get a contradiction. Thus V_a has cardinality at most 1. $\square$

Claim 2.109. *Let $x \in \{b, c\}$ and $y \in \{b, c\} - x$. If $E(V_a, V_{a,x}) = \emptyset$ and $E(V_a, V_{a,y})$ induces a complete bipartite graph, then $E(V_{a,b}, V_{a,c})$ induces a complete bipartite graph.*

Proof. Let $v_{ab} \in V_{a,b}$ and $v_{ac} \in V_{a,c}$. Suppose v_{ab} and v_{ac} are not adjacent. Since the induced subgraph $G[\{v_a, v_{ay}, y, x, v_{ax}\}]$ is isomorphic to P_5, then we get a contradiction. And thus $E(V_{a,b}, V_{a,c})$ induces a complete bipartite graph. $\square$

Claim 2.110. *Let $x \in \{b, c\}$. If $V_{a,x} \neq \emptyset$, $V_{b,c} \neq \emptyset$ and $E(V_a, V_{a,x}) = \emptyset$, then only one of the following statements is true:*

- *$E(V_a, V_{b,c}) = \emptyset$ and $E(V_{a,x}, V_{b,c})$ induces a complete graph, or*
- *each edge set $E(V_a, V_{b,c})$ and $E(V_{a,x}, V_{b,c})$ induce a complete bipartite graph.*

Proof. We will analyze the following four cases:

- when $E(V_a, V_{b,c}) = \emptyset$,
- when $E(V_a, V_{b,c})$ induces a complete bipartite graph,
- when there exist $u \in V_a$ and $v_1, v_2 \in V_{b,c}$ such that $uv_1 \in E(G)$ and $uv_2 \notin E(G)$, and
- when there exists $v \in V_{b,c}$ such that $u_1v \in E(G)$ and $u_2v \notin E(G)$, where $u_1, u_2 \in V_a$.

Consider the first case. Let $v_{ax} \in V_{a,x}$ and $v_{bc} \in V_{b,c}$. Suppose v_{ax} and v_{bc} are not adjacent. Since the induced subgraph $G[\{a, b, c, u_1, v_{ax}, v_{bc}\}]$ is isomorphic to $\mathsf{G}_{6,14}$, then we get a contradiction. And thus $E(V_{a,x}, V_{b,c})$ induces a complete bipartite graph.

Now consider the second case. Let $v_{ax} \in V_{a,x}$ and $v_{bc} \in V_{b,c}$. Suppose v_{ax} and v_{bc} are not adjacent. Since the induced subgraph $G[\{a, b, c, u_1, v_{ax}, v_{bc}\}]$ is isomorphic to $\mathsf{G}_{6,17}$, then we get a contradiction. And thus $E(V_{a,x}, V_{b,c})$ induces a complete bipartite graph.

In what follows we will prove that the last two cases are impossible which implies that the rest possibilities of Claim 2.106 are impossible.

Now consider the third case. Let $v_{ax} \in V_{a,x}$. Suppose there exist $u \in V_a$ and $v_1, v_2 \in V_{b,c}$ such that $uv_1 \in E(G)$ and $uv_2 \notin E(G)$. There are four possible cases:

- $v_{ax}v_1, v_{ax}v_2 \in E(G)$,
- $v_{ax}v_1, v_{ax}v_2 \notin E(G)$,
- $v_{ax}v_1 \in E(G)$ and $v_{ax}v_2 \notin E(G)$, or

- $v_{ax}v_1 \notin E(G)$ and $v_{ax}v_2 \in E(G)$.

Since in first case $G[\{a,b,c,u,v_{ax},v_1,v_2\}] \simeq \mathsf{G}_{7,10}$, in second case $G[\{b,c,u,v_{ax},v_1,v_2\}] \simeq \mathsf{G}_{6,10}$, in third case $G[\{b,c,u,v_{ax},v_1,v_2\}] \simeq \mathsf{G}_{6,14}$ and in fourth case $G[\{u,v_1,v_y,v_2,v_{ax}\}] \simeq \mathsf{P}_5$, then we get a contradiction.

In the last case, we get a contradiction because otherwise by first case: $v_{a,x}v \notin E(G)$, but by second case: $v_{a,x}v \in E(G)$. Thus this case is impossible. $\square$

CLAIM 2.111. *Let $x \in \{b,c\}$. If $V_{a,x} \neq \emptyset$, $V_{b,c} \neq \emptyset$, and $E(V_a, V_{a,x})$ induces a complete bipartite graph, then the edge sets $E(V_a, V_{b,c})$ and $E(V_{a,x}, V_{b,c})$ induce complete bipartite graphs.*

PROOF. Let $v_{ax} \in V_{a,x}$, $v \in V_{b,c}$ and $u \in V_a$. Suppose $uv \notin E(G)$. There are two cases: either $vv_{ax} \in E(G)$ or $vv_{ax} \notin E(G)$. Since the induced subgraph $G[\{a,b,c,u,v,v_{ax}\}]$ is isomorphic to $\mathsf{G}_{6,24}$ and $\mathsf{G}_{6,22}$, in the first case and in the second case, respectively, then we get a contradiction. Thus $uv \in E(G)$ and therefore the edge set $E(V_a, V_{b,c})$ induces a complete bipartite graph.

Now suppose $vv_{ax} \notin E(G)$. By previous result, $uv \in E(G)$. Since the induced subgraph $G[\{a,b,c,u,v,v_{ax}\}]$ is isomorphic to $\mathsf{G}_{6,25}$, then we get a contradiction and thus $vv_{ax} \in E(G)$. Therefore $E(V_{a,x}, V_{b,c})$ induces a complete bipartite graph. $\square$

By previous Claims we have the following cases:

(1) $E(V_a, V_{a,b} \cup V_{ac} \cup V_{b,c}) = \emptyset$, and each edge set $E(V_{a,b}, V_{a,c})$ and $E(V_{a,b} \cup V_{a,c}, V_{b,c})$ induces a complete bipartite graph,
(2) $E(V_a, V_{a,b} \cup V_{ac}) = \emptyset$, and each edge set $E(V_{a,b}, V_{a,c})$ and $E(V_a \cup V_{a,b} \cup V_{a,c}, V_{b,c})$ induces a complete bipartite graph,
(3) $E(V_a, V_{a,b}) = \emptyset$, and each edge set $E(V_a \cup V_{a,b}, V_{a,c})$ and $E(V_a \cup V_{a,b} \cup V_{a,c}, V_{b,c})$ induces a complete bipartite graph,
(4) $E(V_{a,b}, V_{a,c}) = \emptyset$, and each edge set $E(V_a, V_{a,b} \cup V_{a,c} \cup V_{b,c})$ and $E(V_{a,b} \cup V_{a,c}, V_{b,c})$ induces a complete bipartite graph and $|V_a| = |V_{a,b}| = |V_{a,c}| = 1$,
(5) $V_{b,c} = \emptyset$, $E(V_a, V_{a,b} \cup V_{ac}) = \emptyset$, and $E(V_{a,b}, V_{a,c})$ induces a complete bipartite graph,
(6) $V_{b,c} = \emptyset$, $E(V_a, V_{a,x}) = \emptyset$, and $E(V_a \cup V_{a,x}, V_{a,y})$ induces a complete bipartite graph, where $x \in \{a,b\}$ and $y \in \{a,b\} - x$,
(7) $V_{b,c} = \emptyset$, $E(V_a, V_{a,b} \cup V_{a,c})$ induce a complete bipartite graph, $E(V_{a,b}, V_{a,c}) = \emptyset$ and $|V_a| = |V_{a,b}| = |V_{a,c}| = 1$,
(8) $V_{a,y} = \emptyset$, $E(V_a, V_{a,x} \cup V_{b,c}) = \emptyset$, and $E(V_{a,x}, V_{b,c})$ induces a complete bipartite graph, where $x \in \{a,b\}$ and $y \in \{a,b\} - x$,
(9) $V_{a,y} = \emptyset$, $E(V_a, V_{a,x}) = \emptyset$, and $E(V_a \cup V_{a,x}, V_{b,c})$ induces a complete bipartite graph, where $x \in \{a,b\}$ and $y \in \{a,b\} - x$,
(10) $V_{a,y} = \emptyset$, $E(V_a, V_{a,x} \cup V_{b,c})$ and $E(V_{a,x}, V_{b,c})$ induce a complete bipartite graph, where $x \in \{a,b\}$ and $y \in \{a,b\} - x$,
(11) $V_{a,b} \cup V_{a,c} = \emptyset$, and $E(V_a, V_{b,c})$ induces a complete bipartite graph,
(12) $V_{a,b} \cup V_{a,c} = \emptyset$ and $E(V_a, V_{b,c}) = \emptyset$,
(13) $V_{a,b} \cup V_{a,c} = \emptyset$, and $E(V_a, V_{b,c})$ induces a complete bipartite graph minus an edge,
(14) $V_{a,b} \cup V_{a,c} = \emptyset$, $|V_{b,c}| \geq 2$ and $E(V_a, V_{b,c})$ induces a complete bipartite graph minus two edges u_1v and u_2v, where $v \in V_{b,c}$,
(15) $V_{a,b} \cup V_{a,c} = \emptyset$, $|V_a| = |V_{b,c}| = 2$ and $E(V_a, V_{b,c})$ induces a perfect matching,
(16) $V_{a,y} \cup V_{b,c} = \emptyset$ and $E(V_a, V_{a,x})$ induces a complete bipartite graph, where $x \in \{a,b\}$ and $y \in \{a,b\} - x$, and
(17) $V_{a,y} \cup V_{b,c} = \emptyset$ and $E(V_a, V_{a,x}) = \emptyset$, where $x \in \{a,b\}$ and $y \in \{a,b\} - x$.

Now we describe $V_\emptyset$, the set of vertices that are not adjacent with any vertex in $\{a, b, c\}$. Let $w \in V_\emptyset$. The vertex w is adjacent with a vertex in $V_a \cup V_{a,b} \cup V_{a,c} \cup V_{b,c}$, because otherwise the shortest path from w to $\{a, b, c\}$ would contains the graph P_5 as induced subgraph.

Claim 2.112. *If $V_{a,x} \neq \emptyset$ for some $x \in \{b, c\}$, then $E(V_\emptyset, V_a \cup V_{a,x}) = \emptyset$.*

Proof. Let $w \in V_\emptyset$, $v_{ax} \in V_{a,x}$ and $u \in V_a$. Suppose w is adjacent with u or v_{ax}. Then there are three possible cases:

(a) w is adjacent only with u,
(b) w is adjacent only with v_{ax}, or
(c) w is adjacent with both vertices u and v_{ax}.

First consider case (a). This case has two subcases: either u is adjacent with v_{ax} or not. If they are adjacent, then $G[\{u, w, v_{ax}, b, c\}]$ is isomorphic to P_5. Since it is forbidden, then v_a is not adjacent with v_{ax}. On the other hand, if they are not adjacent, then $G[\{a, b, c, w, u, v_{ax}\}]$ is isomorphic to $\mathsf{G}_{6,7}$. Therefore, case (a) is impossible. Now consider case (b). There are two possible cases: either u is adjacent with v_{ax} or not. If they are adjacent, then $G[\{a, b, c, w, u, v_{ax}\}]$ is isomorphic to $\mathsf{G}_{6,14}$; which is forbidden. Thus u is not adjacent with v_{ax}. But if they are not adjacent, then we have that $G[\{a, b, c, w, u, v_{ax}\}]$ is isomorphic to $\mathsf{G}_{6,10}$. Thus case (b) cannot occur. Finally, consider case (c). The subcases are: either u is adjacent to v_{ax} or not. If they are adjacent, then $G[\{a, b, c, w, u, v_{ax}\}]$ is isomorphic to $\mathsf{G}_{6,22}$; which is forbidden. Then u is not adjacent with v_{ax} But if they are not adjacent, then the induced subgraph $G[\{u, w, v_{ax}, b, c\}]$ is isomorphic to P_5; which is forbidden. Thus, we get that case (c) is impossible. And therefore, w is not adjacent with u or v_{ax}. From which the result follows. $\square$

Claim 2.113. *Let $x \in \{b, c\}$. If $V_{a,x} \neq \emptyset$, then $E(V_\emptyset, V_{b,c}) = \emptyset$. Moreover, if $V_{a,x} \neq \emptyset$, then $V_\emptyset = \emptyset$.*

Proof. Let $w \in V_\emptyset$, $v_{ax} \in V_{a,x}$, $v_a \in V_a$ and $v_{bc} \in V_{b,c}$. Suppose w is adjacent with v_{bc}. By Claim 2.112, we have that w is not adjacent with a vertex in $V_a \cup V_{a,x}$. By Claims 2.110 and 2.111, the vertex v_{ax} is adjacent with v_{bc}, and one of the following three cases occur:

(a) v_a is adjacent with v_{ax} and v_{bc},
(b) v_a is adjacent only with v_{bc}, and
(c) v_a is not adjacent with both v_{ax} and v_{bc}.

Case (a) is not possible because the induced subgraph $G[\{v_a, v_{ax}, b, c, v_{bc}, w\}]$ would be isomorphic to $\mathsf{G}_{6,12}$; which is impossible. In case (b), we have that $G[\{v_a, v_{ax}, b, c, v_{bc}, w\}]$ is isomorphic to $\mathsf{G}_{6,4}$. Since it is forbidden, then this case is not possible. Finally, case (c) is not possible since $G[\{v_a, a, x, v_{bc}, w\}]$ is isomorphic to P_5 that is forbidden. Thus w is not adjacent with v_{bc}. And there is no vertex in $V_\emptyset$ adjacent with a vertex in $G \setminus V_\emptyset$. Therefore $V_\emptyset$ is empty. $\square$

Thus by previous Claim, in cases (1) to (10), (16) and (17) the vertex set $V_\emptyset$ is empty. The cases when $V_{a,b} \cup V_{a,c} = \emptyset$ (cases (4) and (7)) correspond to a graph isomorphic to an induced subgraph of $\mathcal{F}_1^2$. The rest of these cases correspond to a graph isomorphic to an induced subgraph of $\mathcal{F}_1^1$.

Claim 2.114. *Let $w_a \in V_\emptyset$ such that w_a is adjacent with $u \in V_a$, and w_a is not adjacent with a vertex in $V_{b,c}$. Let $w_{bc} \in V_\emptyset$ such that w_{bc} is adjacent with $v \in V_{b,c}$ and w_{bc} is not adjacent with a vertex in V_a. Let $w \in V_\emptyset$ such that w is adjacent with $u' \in V_a$ and $v' \in V_{b,c}$. If $E(V_a, V_{b,c})$ induces a complete bipartite graph, then no two vertices of $\{w_a, w_{bc}, w\}$ exist at the same time.*

Proof. Suppose w_a and w_{bc} exist at the same time. There there are two possible cases: either w_a and w_{bc} are adjacent or not. If w_a and w_{bc} are adjacent, then $G[\{b,a,u,w_a,w_{bc}\}]$ is isomorphic to P_5; which is impossible. If w_a and w_{bc} are not adjacent, then $G[\{b,c,u,v,w_a,w_{bc}\}]$ is isomorphic to $\mathsf{G}_{6,6}$; which is impossible. Then w_a and w_{bc} do not exist at the same time.

Suppose w_a and w exist at the same time. There there are two possible cases: either $u=u'$ or $u\neq u'$. Suppose $u=u'$, then we have two possible cases: either w_a and w are adjacent or not. If w_a and w are adjacent, then $G[\{w_a,w,v',b,a\}]$ is isomorphic to P_5; which is impossible. But if w_a and w are not adjacent, then $G[\{b,c,u,v',w_a,w\}]$ is isomorphic to $\mathsf{G}_{6,11}$; which is impossible. Then $u\neq u'$. Thus suppose $u\neq u'$. Note that in the induced subgraph $G[\{a,b,c,u,v',w_a,w\}]$, the vertex w is only adjacent with v', and w_a is only adjacent with u. Then applying previous (w_a,w_{bc}) case in this induced subgraph, we get that w and w_a do not exist at the same time.

Suppose w_{bc} and w exist at the same time. There are two possible cases: either $v=v'$ or $v\neq v'$. Suppose $v=v'$, then we have two possible cases: either w_{bc} and w are adjacent or not. If w_a and w_{bc} are adjacent, then $G[\{w_{bc},w,u,a,b\}]$ is isomorphic to P_5; which is impossible. But if w_a and w_{bc} are not adjacent, then $G[\{b,c,u,v,w_{bc},w\}]$ is isomorphic to $\mathsf{G}_{6,11}$; which is impossible. Then $v\neq v'$. Suppose $v\neq v'$. Note that in the induced subgraph $G[\{a,b,c,u,v,w_a,w\}]$, the vertex w is only adjacent with u, and w_{bc} is only adjacent with v. Thus by applying first case in this induced subgraph, we get that w and w_{bc} do not exist at the same time. $\square$

Claim 2.115. *Let $w\in V_\emptyset$, $u\in V_a$ and $v\in V_{b,c}$. If $E(V_a,V_{b,c})$ induces a complete bipartite graph and w is adjacent with u and v, then $E(w,V_a\cup V_{b,c})$ induces a complete graph.*

Proof. First we see that if $v'\in V_{b,c}-v$, then w is adjacent with v'. Suppose w and v' are not adjacent. Since the induced subgraph $G[\{a,b,u,v,v',w\}]$ is isomorphic to the forbidden graph $\mathsf{G}_{6,15}$, then we get a contradiction. Thus w is adjacent with each vertex in $V_{b,c}$. Now we see that if $u'\in V_a-u$, then w is adjacent with u'. Suppose w and u' are not adjacent. Since the induced subgraph $G[\{a,b,u,u,u',w\}]$ is isomorphic to $\mathsf{G}_{6,15}$, then we get a contradiction. Therefore, w is adjacent with each vertex in $V_a\cup V_{b,c}$. $\square$

Claim 2.116. *If $|V_a|=2$ and $E(V_\emptyset,V_a)\neq\emptyset$, then either each vertex in $V_\emptyset$ is adjacent only with one vertex in V_a and $V_\emptyset$ is a clique, or each vertex in $V_\emptyset$ is adjacent with u_1 and u_2, and $V_\emptyset$ is trivial.*

Proof. Let $w,w'\in V_\emptyset$ and $i,j\in\{1,2\}$ with $i\neq j$. By proving that the following cases are not possible, it follows that the only possible cases are that either $E(\{u_1,u_2\},\{w,w'\})$ is equal to $\{wu_i,w'u_j\}$ and $ww'\in E(G)$, or $E(\{u_1,u_2\},\{w,w'\})$ is equal to $\{wu_i,wu_j,w'u_i,w'u_j\}$ and $ww'\in E(G)$. Which implies the result.

(a) $ww',wu_j,w'u_j\notin E(G)$ and $wu_i,w'u_i\in E(G)$,
(b) $ww',wu_j,w'u_i\notin E(G)$ and $wu_i,w'u_j\in E(G)$,
(c) $wu_j,w'u_i\notin E(G)$ and $ww',wu_i,w'u_j\in E(G)$,
(d) $ww',w'u_j\notin E(G)$ and $wu_i,wu_j,w'u_i\in E(G)$,
(e) $w'u_j\notin E(G)$ and $ww',wu_i,wu_j,w'u_i\in E(G)$, and
(f) $w'u_j,ww',wu_i,wu_j,w'u_i\in E(G)$.

Since in cases (a), (b) and (d) the induced subgraph $G[\{a,b,u_1,u_2,w,w'\}]$ is isomorphic to $\mathsf{G}_{6,2}$, $\mathsf{G}_{6,8}$, $\mathsf{G}_{6,10}$, respectively, then these cases are impossible. On the other hand, in case (c) the induced graph $G[\{w,w',u_j,a,b\}]$ is isomorphic to P_5: which is impossible. Also in case (e) the induced subgraph $G[\{w',w,u_j,a,b\}]$ is isomorphic to P_5; which is not possible. Finally in case (f) the induced subgraph $G[\{a,b,c,u_1,u_2,w,w'\}]$ is isomorphic to $\mathsf{G}_{7,14}$; which is not possible. $\square$

Claim 2.117. *If $E(V_a, V_{b,c})$ induces a complete bipartite graph, and each vertex in $V_\emptyset$ is adjacent only with one vertex in V_a, then there is no vertex in $V_\emptyset$ adjacent with a vertex in $V_{b,c}$.*

Proof. Let $w \in V_\emptyset$, $v \in V_{b,c}$. Suppose w is adjacent with u_1 and v, but w is not adjacent with u_2. Since the induced subgraph $G[\{a, b, w, u_1, u_2, v\}]$ is isomorphic to $\mathsf{G}_{6,17}$, then we get a contradiction and the result follows. □

Claim 2.118. *Let $w \in V_\emptyset$ and $u \in V_a$. If $E(V_a, V_{b,c})$ induces a complete bipartite graph, $E(w, V_{b,c}) \neq \emptyset$ and $E(w, V_a) = \emptyset$, then each vertex in $V_\emptyset$ is adjacent with only one vertex in $V_{b,c}$, and $V_\emptyset$ is a clique of cardinality at most 2.*

Proof. Let $v, v' \in V_{b,c}$. Suppose w is adjacent with v and v'. Since $G[\{a, b, c, u, v, v', w\}]$ is isomorphic to $\mathsf{G}_{7,9}$, then we get a contradiction. And we have that w is adjacent only with one vertex in $V_{b,c}$

Now let $w, w' \in V_\emptyset$ and $v, v' \in V_{b,c}$ such that $wv \in E(G)$ and $w'v' \notin E(G)$. Suppose $v \neq v'$. There are two cases: either w and w' are adjacent or not. If w and w' are adjacent, then the induced subgraph $G[\{b, c, v, u, w, w'\}]$ is isomorphic to $\mathsf{G}_{6,6}$, then we get a contradiction and $ww' \notin E(G)$. Now if w and w' are not adjacent, then the induced subgraph $G[\{w, v, u, v', w'\}]$ is isomorphic to P_5 and we get a contradiction. Thus $v = v'$.

Finally, let $w, w' \in V_\emptyset$ adjacent with $v \in V_{b,c}$. Suppose w and w' are not adjacent. Since the induced subgraph $G[\{a, b, v, u, w, w'\}]$ is isomorphic to $\mathsf{G}_{6,3}$, then we get a contradiction and $ww' \in E(G)$. □

In case (11), by Claims 2.114, 2.115, 2.116, 2.117 and 2.118, we have the following possible cases:

- each vertex in $V_\emptyset$ is adjacent with each vertex in $V_a \cup V_{b,c}$, and $V_\emptyset$ is a trivial graph,
- each vertex in $V_\emptyset$ is adjacent only with a vertex in V_a, and $V_\emptyset$ is a clique of cardinality at most 2, or
- each vertex in $V_\emptyset$ is adjacent only with a vertex in $V_{b,c}$, and $V_\emptyset$ is a clique of cardinality at most 2.

Each of these cases corresponds to a graph isomorphic to an induced subgraph of $\mathcal{F}_1^1$.

Remark 2.119. *Let $w \in V_\emptyset$, $u \in V_a$ and $v \in V_{b,c}$. If $uv \notin E(G)$, and w adjacent with u or v, then w is adjacent with both vertices u and v.*

In case (12), by Claim 2.116 and Remark 2.119, we have that each vertex in $V_\emptyset$ is adjacent with each vertex in $V_a \cup V_{b,c}$, and

- $V_\emptyset$ is a clique or a trivial graph when $|V_a| = 1$, or
- $V_\emptyset$ is a clique when $|V_a| = 2$.

It is not difficult to see that each of these graphs are isomorphic to an induced subgraph of $\mathcal{F}_1^1$

Claim 2.120. *Let $w \in V_\emptyset$, $u \in V_a$ and $v \in V_{b,c}$. Suppose u is adjacent with each vertex in $V_{b,c} - v$. If $E(w, V_{b,c} \cup \{u\}) \neq \emptyset$, then one of the following statements holds:*

- *each vertex in $V_\emptyset$ is adjacent with u and v,or*
- *$E(w, V_{b,c} \cup \{u\})$ induces a complete bipartite graph and $|V_{b,c} - v| = 1$.*

Proof. First case follows by Remark 2.119. Now we check when w is adjacent with a vertex in $V_{b,c} - v$. Let $v' \in V_{b,c}$. Then u is adjacent with v'. Suppose w is adjacent only with v'. Since the induced subgraph $G[\{a, b, u, v, v', w\}]$ is isomorphic to $\mathsf{G}_{6,9}$, then we have a contradiction and

w is adjacent also with u or v. Thus by Remark 2.119, w is adjacent with u, v and v'. Finally by applying Claim 2.115 to the induced subgraph $G[\{w,u\} \cup (V_{b,c}-v)]$, we get that w also is adjacent with each vertex in $V_{b,c}-v$. Thus w is adjacent with each vertex in $\{u\} \cup V_{b,c}$.

Suppose the cardinality of $V_{b,c}-v$ is at least 2. Take $v, v' \in V_{b,c}-v$. Thus w is adjacent with v, v' and v''. Since the induced subgraph $G[\{u,v,v',v'',a,w\}]$ is isomorphic to $\mathsf{G}_{6,5}$, then we get a contradiction. Thus the cardinality of $V_{b,c}-v$ is at most 1. □

Claim 2.121. *Let $w \in V_\emptyset$ and $v \in V_{b,c}$. Suppose u_1 is adjacent with each vertex in $V_{b,c}$, and u_2 is adjacent with each vertex in $V_{b,c}-v$. If $E(w,V_{b,c}) \neq \emptyset$, then either w is adjacent only with u_2 and v, or $|V_{b,c}|=1$ and $E(w,V_a \cup V_{b,c})$ induces a complete bipartite graph.*

Proof. First case follows by Remark 2.119. Now we prove that if w is adjacent with u_1, then w is adjacent with u_2 and v. Suppose w is adjacent only with u_1. Since the induced subgraph $G[\{a,b,u_1,u_2,v,w\}]$ is isomorphic to $\mathsf{G}_{6,11}$, then we have a contradiction and w is adjacent also with u_2 or v. Thus by Remark 2.119, w is adjacent with u_1, u_2 and v. Finally, suppose $|V_{b,c}| \geq 2$. Let $v' \in V_{b,c}$. By applying Claim 2.115 to the induced subgraph $G[\{w,u_1\} \cup V_{b,c}]$, we get that w also is adjacent with each vertex in $V_{b,c}$. Thus w is adjacent with each vertex in $V_a \cup V_{b,c}$. But since $G[\{a,u_1,u_2,v,v',w\}]$ is isomorphic to $\mathsf{G}_{6,18}$, then we get a contradiction, and $|V_{b,c}|=1$. □

Let $u \in V_a$ and $v \in V_{b,c}$ such that $uv \notin E(G)$. Therefore, in case (13), by Claims 2.116, 2.120 and 2.121, we have that one of the following cases holds:

- $V_\emptyset$ is a clique of cardinality at most 2, and each vertex in $V_\emptyset$ is adjacent with u and v,
- $V_a=\{u\}$, $|V_{b,c} \setminus \{v\}| \leq 1$, $V_\emptyset$ is trivial and each vertex in $V_\emptyset$ is adjacent with each vertex in $\{u\} \cup V_{b,c}$, or
- $|V_a|=2$, $V_\emptyset$ is trivial and $V_{b,c}=\{v\}$ and each vertex in $V_\emptyset$ is adjacent with u_1, u_2 and v.

It can be checked that each of these graphs is isomorphic to an induced subgraph or a graph in $\mathcal{F}_1^1$.

Claim 2.122. *Let $w \in V_\emptyset$ and $v \in V_{b,c}$ such that v is not adjacent with u_1 and u_2. Suppose $|V_{b,c}| \geq 2$ and $E(V_a, V_{b,c}-v)$ induces a complete bipartite graph. If $E(w,V_{b,c}) \neq \emptyset$, then $V_\emptyset=\{w\}$, and w is adjacent with u_1, u_2 and v.*

Proof. By Remark 2.119, we have that if w is adjacent with one vertex of $\{v,u_1,u_2\}$, then w is adjacent with v, u_1 and u_2. Now we see that w is not adjacent to any other vertex. Suppose w is adjacent with $v' \in V_{b,c}-v$. First consider the case when w is not adjacent with v, u_1 and u_2. Since $G[\{b,v,v',u_1,u_2,w\}]$ is isomorphic to $\mathsf{G}_{6,6}$, then we get a contradiction and thus w is adjacent with v, v', u_1, and u_2. But if w is adjacent with v, v', u_1, and u_2, then $G[\{b,v,v',u_1,u_2,w\}]$ is isomorphic to $\mathsf{G}_{6,20}$; which is impossible. Then w is not adjacent with any vertex in $V_{b,c}-v$.

Suppose there exist another vertex $w' \in V_\emptyset$. By previous discussion w' is adjacent with v, u_1 and u_2. Since $G[\{a,b,w,w',v,v'\}]$ is isomorphic to $\mathsf{G}_{6,1}$, then we get a contradiction. Thus $V_\emptyset$ has cardinality at most 1. □

Let $v \in V_{b,c}$ such that $u_1v, u_2v \notin E(G)$. Therefore, in case (14), by Claims 2.116 and 2.122, we have that $V_\emptyset=\{w\}$ and w is adjacent with u_1, u_2 and v. Then G is isomorphic to an induced subgraph of $\mathcal{F}_1^1$.

Claim 2.123. *Let $V_{b,c}=\{v,v'\}$ such that $vu_1, v'u_2 \in E(G)$ and $vu_2, v'u_1 \notin E(G)$. If $w \in V_\emptyset$ and $E(w,V_a \cup V_{b,c}) \neq \emptyset$, then w is adjacent with u_1, u_2, v and v'.*

Proof. Since w is adjacent with a vertex in $\{u_1, u_2, v, v'\}$, then w is adjacent with u_1 and v', or w is adjacent with u_2 and v. Suppose w is adjacent with u_1 and v', but w is not adjacent with u_2 and v. Since $G[\{a, u_1, u_2, v, v', w\}]$ is isomorphic to $\mathsf{G}_{6,11}$, then we get a contradiction, and therefore w is also adjacent with u_2 and v. $\square$

Let $V_{b,c} = \{v, v'\}$ such that $vu_1, v'u_2 \in E(G)$ and $vu_2, v'u_1 \notin E(G)$. Thus in case (15), by Claims 2.116 and 2.123, we have that $V_\emptyset$ is trivial and each vertex in $V_\emptyset$ is adjacent only with u_1, u_2, v and v'. Therefore, G is isomorphic to an induced subgraph of $\mathcal{F}_1^1$.

CHAPTER 3

Critical ideals of signed multidigraphs with twins

Two vertices of a graph are *twins* if they have the same neighbors. There are two types of twins depending on whether the twins are connected or not. Here, we study the critical ideals of a graph having twin vertices. Specifically, we obtain relations between some evaluations of the critical ideals of a graph G and the critical ideals of G with some vertices arbitrarily cloned. As a consequence, we get an upper bound for the algebraic co-rank for a graph with twin vertices.

1. Critical ideals and signed multidigraphs

A *signed multidigraph* G_σ is a pair that consists of a multidigraph G (a digraph possibly with multiple arcs) and a function σ, called the *sign*, from the edges of G into the set $\{1, -1\}$. Given the set of variables $X_G = \{x_u : u \in V(G)\}$ indexed by the vertices of G, and a principal ideal domain (PID) $\mathcal{P}$, the *generalized Laplacian matrix* $L(G_\sigma, X_G)$ of G_σ is the matrix whose entries are given by

$$L(G_\sigma, X_G)_{uv} = \begin{cases} x_u & \text{if } u = v, \\ -\sigma(uv)m_{uv}1_{\mathcal{P}} & \text{otherwise,} \end{cases}$$

where m_{uv} is the number of arcs *leaving* u and *entering* to v, and $1_{\mathcal{P}}$ is the *identity* of $\mathcal{P}$. Moreover, if $\mathcal{P}[X_G]$ is the polynomial ring over $\mathcal{P}$ in the variables X_G, then the *critical ideals* of G_σ are the determinantal ideals given by

$$I_i(G_\sigma, X_G) = \langle \{\det(m) : m \text{ is an } i \times i \text{ submatrix of } L(G_\sigma, X_G)\} \rangle \subseteq \mathcal{P}[X_G],$$

for all $1 \leq i \leq |V(G)|$. We say that a critical ideal is *trivial* when it is equal to $\langle 1 \rangle$.

Definition 3.1. *The algebraic co-rank $\gamma_{\mathcal{P}}(G_\sigma)$ of G_σ is the maximum integer i such that $I_i(G_\sigma, X_G)$ is trivial.*

Note that $\gamma_{\mathcal{P}}(G_\sigma) \leq n - 1$, since $I_n(G_\sigma, X_G) = \langle \det(L(G_\sigma, X_G)) \rangle \neq \langle 1 \rangle$. The algebraic co-rank of a graph is closely related to the combinatorial properties of the graph. For instance, if H_σ is an induced subgraph of G_σ, then $I_i(H_\sigma, X_H) \subseteq I_i(G_\sigma, X_G)$ for all $1 \leq i \leq |V(H)|$ (see [**22**, Proposition 3.3]). Therefore, $\gamma(G_\sigma) \leq \gamma(H_\sigma)$. In [**22**, Theorem 3.13] the following bounds were obtained:

$$\gamma_{\mathcal{P}}(G) \leq 2(n - \omega(G)) + 1 \text{ and } \gamma_{\mathcal{P}}(G) \leq 2(n - \alpha(G)).$$

We now introduce the concepts of duplication and replications of vertices which are key in our study. Given a digraph G and a vertex $v \in V(G)$, the *duplication* $d(G, v)$ of v is the digraph obtained from G by adding a new vertex v^1 to G and the arcs

$$\{v^1u : u \in N_+(v)\} \cup \{uv^1 : u \in N_-(v)\}.$$

In this case, v and v^1 are called *false twins*. The *replication* $r(G, v)$ of v on G is the graph obtained from $d(G, v)$ by adding the arcs vv^1 and v^1v. In this case, we say that v and v^1 are *true twins*. We say that two vertices of a digraph are *twins* if they are true or false twins. Also, let $d^k(G, v)$

and $r^k(G, v)$ denote the multidigraphs obtained from G by duplicating and replicating k times the vertex v, respectively. For signed multidigraphs, the new arcs have the same multiplicity and sign than the original arcs.

In general, given $\mathbf{d} \in \mathbb{Z}^{V(G)}$, let $G^{\mathbf{d}}$ be the graph obtained from G by duplicating $\mathbf{d}_v$ times the vertex v when $\mathbf{d}_v > 0$ and replicating $-\mathbf{d}_v$ times the vertex v when $\mathbf{d}_v < 0$. Observe that $G = G^{\mathbf{0}}$. Let $V(G^{\mathbf{d}}, v)$ denote the vertex set $\{v^1, \ldots, v^{|\mathbf{d}_v|}\}$ created by either duplicating or replicating the vertex v. The following example illustrates these concepts.

Example 3.2. *Let C_4 be the cycle with four vertices and $\mathbf{d} = (-1, 1, 1, 1)$. Thus $C_4^{\mathbf{d}}$ is the graph with eight vertices shown in Figure 8.b.*

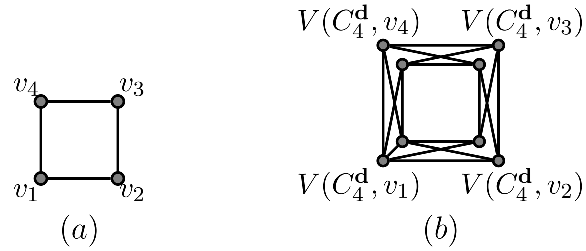


Figure 8. The cycle with four vertices and $C_4^{(-1,1,1,1)}$.

Critical ideals were firstly defined in [**22**] as an algebraic generalization of the critical group of a graph. Which is now recalled. The *Laplacian matrix* $L(G_\sigma)$ of G_σ is the evaluation of $L(G_\sigma, X_G)$ at $X_G = D_G$, where D_G is the degree vector of G. By considering $L(G_\sigma)$ as a linear map $L(G_\sigma) : \mathbb{Z}^V \to \mathbb{Z}^V$, the *cokernel* of $L(G_\sigma)$ is the quotient module $\mathbb{Z}^V / \operatorname{Im} L(G_\sigma)$. The torsion part of this module is the *critical group* $K(G_\sigma)$ of G_σ. The critical group has been studied intensively on several contexts over the last 30 years: the *group of components* [**34, 35**], the *Picard group* [**9, 14**], the *Jacobian group* [**9, 14**], the *sandpile group* [**21**], *chip-firing game* [**14, 37**], or *Laplacian unimodular equivalence* [**29, 38**]. And recently, critical ideals have played an important role in the classification and understanding of the graphs whose critical group has i invariant factors equal to 1, see [**1, 2**].

In general, the relations between the critical group and other parameters of the graph remain unknown. Actually, researches have focused on two topics. One is to determine the exact structure of $K(G)$ for some special families of graphs. The other one is to study the relationship between the critical group of a graph and that of graphs obtained from it by various constructions. There are some natural constructions on graphs which behave well with respect the critical group. For example, trivially the critical group $K(G + H)$ of *disjoint union* $G + H$ of two graphs G and H is isomorphic to $K(G) \oplus K(H)$. More interesting, in [**44**] was proved that if the graphic matroids of G and H are isomorphic, then their critical groups are isomorphic. This was proved by studying the operations of *splittings* or *mergings of one-vertex cuts* and *twistings of two-vertex cuts*.

The purpose of this chapter is to study the critical ideals of signed multidigraphs having twin vertices. Several graph families have twin vertices. For instance, the complete multipartite graphs, the threshold graphs, the quasi-threshold graphs, or the cographs. Therefore, the description of critical ideals of graphs with twins is an important step in the development on the theory of critical ideals and critical group. In Section 2, we will obtain relations between some evaluations of the critical ideals of a signed multidigraph G and the critical ideals of $G^{\mathbf{d}}$, where $\mathbf{d} \in \mathcal{P}^{V(G)}$.

As a consequence of this partial description, we will get an upper bound for the algebraic co-rank of graphs with twins. This upper bound is crucial in the classification of the graphs whose algebraic co-rank is less than or equal to an integer k (see [**2**, Section 2]). We will also state three conjectures which lead into a wide and interesting panorama of the critical ideals. In Section 3, we give a description of the critical ideals of the *k-th* duplication $d^k(G,v)$ of vertex v and *k-th* replication $r^k(G,v)$ of vertex v in terms of some of the critical ideals of G.

2. An upper bound for the algebraic co-rank of graphs with twins

This section is meant to be a first approach to the theory of critical ideals of graphs with twins. First, we present a core result, Lemma 3.3, which computes the minors of a special type of matrices, the join of matrices. In particular, we will see that almost all minors of these matrices are equal to zero. This lemma turns out to be very useful because the generalized Laplacian matrix of several multidigraphs are the join of matrices. For instance, the generalized Laplacian matrix of a graph obtained by duplicating or replicating its vertices. By using this lemma, we will get a first description for the critical ideals of the graph obtained by the duplication or replication of its vertices (Lemma 3.5 and Theorem 3.9). After, we will find that this non-accurate description of the critical ideals of graphs with twin vertices is enough to we get an upper bound for the algebraic co-rank of a graph with twins (Corollary 3.11). In fact, this bound is tight since the complete graphs satisfy it (Example 3.12). This upper bound is important in the theory of critical ideals of graphs. For instance, it can be used in the classification of the graphs whose algebraic co-rank is less than or equal to an integer k.

First we define the join of matrices and prove a lemma that will be frequently used in this chapter. Let $\mathcal{P}$ be a commutative ring with identity and let $M_n(\mathcal{P})$ denote the set of $n\times n$ matrices with entries on $\mathcal{P}$. Given two vectors $\mathbf{a}\in\mathcal{P}^{q_1}$ and $\mathbf{b}\in\mathcal{P}^{q_2}$, and two matrices $P\in M_{p_1\times p_2}(\mathcal{P})$ and $Q\in M_{q_1\times q_2}(\mathcal{P})$ such that $p_1+q_1=p_2+q_2$, then the *join* $J(P,\mathbf{a};Q,\mathbf{b})$ is the matrix

$$\begin{bmatrix} P & \mathbf{1}_{p_1}^T\mathbf{b} \\ \mathbf{a}^T\mathbf{1}_{p_2} & Q \end{bmatrix} \in M_{p_1+q_1}(\mathcal{P}).$$

Note that $L(G\vee H, X_{G\vee H}) = J(L(G,X_G),-\mathbf{1};L(H,X_H),-\mathbf{1})$. The following lemma describes the determinant of the join $J(P,\mathbf{a};Q,\mathbf{b})$.

LEMMA 3.3. *If $P\in M_{p_1\times p_2}(\mathcal{P})$, $Q\in M_{q_1\times q_2}(\mathcal{P})$ with $p_1+q_1=p_2+q_2$, $\mathbf{a}\in\mathcal{P}^{q_1}$, and $\mathbf{b}\in\mathcal{P}^{q_2}$, then*

$$\det(J(P,\mathbf{a};Q,\mathbf{b})) = \begin{cases} \det(P)\cdot\det(Q) - \det\begin{bmatrix} P & \mathbf{1}^T \\ \mathbf{1} & 0\end{bmatrix}\cdot\det\begin{bmatrix} 0 & \mathbf{b} \\ \mathbf{a}^T & Q\end{bmatrix} & \text{if } p_1=p_2, \\[2ex] \det\begin{bmatrix} P & \mathbf{1}^T\end{bmatrix}\cdot\det\begin{bmatrix}\mathbf{b} \\ Q\end{bmatrix} & \text{if } p_1=p_2+1, \\[2ex] \det\begin{bmatrix} P \\ \mathbf{1}\end{bmatrix}\cdot\det\begin{bmatrix}\mathbf{a}^T & Q\end{bmatrix} & \text{if } p_2=p_1+1, \\ 0 & \text{otherwise.}\end{cases}$$

PROOF. The proof follows by using induction on $p_1 + p_2$. Note that, if $P \in M_{1\times 0}(\mathcal{P})$, then $[\ P \quad \mathbf{1}\ ] = [1]$. Also, if $P \in M_{0\times 1}(\mathcal{P})$, then $[\ P \quad \mathbf{1}\ ]^T = [1]$. □

Note that all the square submatrices of a join of matrices are in fact join of matrices. Therefore, almost all minors of the join of matrices are equal to zero. This fact will be useful in getting a description of the critical ideals of a graph with twin vertices (see Lemmas 3.5, 3.6, 3.16 and 3.21).

Let $[n] = \{1, ..., n\}$. Given two sets $\mathcal{I} = \{i_1, \dots, i_r\} \subseteq [n]$, $\mathcal{J} = \{j_1, \dots, j_s\} \subseteq [n]$, and a matrix $M \in M_n(\mathcal{P})$, the submatrix of M obtained by the rows $i_1, \dots, i_r$ and the columns $j_1, \dots, j_s$ will be denoted by $M[\mathcal{I}; \mathcal{J}]$. On the other hand, given $\mathbf{a} \in \mathcal{P}^n$, $L \in M_n(\mathcal{P})$, and $1 \leq j \leq n$. Let

$$\text{minors}_j(L, \mathbf{a}) = \left\{ \det \begin{bmatrix} \mathbf{a}' \\ M \end{bmatrix} : \begin{bmatrix} \mathbf{a}' \\ M \end{bmatrix} \in M_j\left(\begin{bmatrix} \mathbf{a} \\ L \end{bmatrix} \right) \right\}$$

and

$$\text{minors}_j(\mathbf{a}, L) = \left\{ \det \begin{bmatrix} \mathbf{a}'^T & M \end{bmatrix} : \begin{bmatrix} \mathbf{a}'^T & M \end{bmatrix} \in M_j\left(\begin{bmatrix} \mathbf{a}^T & L \end{bmatrix} \right) \right\}.$$

CLAIM 3.4. *Let G be a signed multidigraph with $n \geq 2$ vertices and v be a vertex of G. Suppose $L(G, X_G) = J(x_v, \mathbf{a}; L(G - v, X_{G-v}), \mathbf{b})$, for some $\mathbf{a}, \mathbf{b} \in \mathcal{P}^{n-1}$. Then the critical ideal $I_j(G, X_G)$ is equal to*

$$\begin{aligned} \langle &\text{minors}_j(L(G - v, X_{G-v})), \text{minors}_j(\mathbf{a}, L(G - v, X_{G-v})), \text{minors}_j(L(G - v, X_{G-v}), \mathbf{b}), \\ &\{x_v \cdot \det(M) + \det(J(0, \mathbf{a}'; M, \mathbf{b}')) \ :\ J(x_v, \mathbf{a}'; M, \mathbf{b}') \in M_j(L(G, X_G))\}\rangle, \end{aligned}$$

when $1 \leq j \leq n - 1$, and it is equal to

$$\langle x_v \cdot \det(L(G - v, X_{G-v})) + \det(J(0, \mathbf{a}; L(G - v, X_{G-v}), \mathbf{b}))\rangle,$$

when $j = n$.

PROOF. The proof is simple, and similar to the proof of [**22**, Claim 3.12]. □

Now we give a description of the critical ideals of $d(G, v)$ in terms of the critical ideals of G. Let $X \subseteq X_G$ and $\mathbf{a} \in \mathcal{P}^{|X|}$. Through the thesis, $I(G, X_G)|_{X=\mathbf{a}}$ will denote the evaluation of $I(G, X_G)$ at $X = \mathbf{a}$.

LEMMA 3.5. *Let G be a signed multidigraph with $n \geq 2$ vertices and v be a vertex of G. Then*

$$I_j(d(G, v), X_{d(G,v)}) \subseteq \langle x_{v^0}, x_{v^1}, I_j(G, X_G)|_{x_v=0}\rangle,$$

for all $1 \leq j \leq n$. Moreover, $I_j(d(G, v), X_{d(G,v)})$ is trivial if and only if $I_j(G, X_G)|_{x_v=0}$ is trivial.

PROOF. Suppose $L(G, X_G) = J(x_v, \mathbf{a}; L(G - v, X_{G-v}), \mathbf{b})$, for some $\mathbf{a}, \mathbf{b} \in \mathcal{P}^{n-1}$. Let $\mathcal{I}, \mathcal{I}' \subseteq [n + 1]$ be two sets of size j, and $\mathcal{I}_{\{1,2\}} = \{1, 2\} \cap \mathcal{I}$ and $\mathcal{I}'_{\{1,2\}} = \{1, 2\} \cap \mathcal{I}'$. Note that $L(d(G, v), X_{d(G,v)}) = J(\text{diag}(x_{v^1}, x_{v^0}), \mathbf{a}; L(G - v, X_{G-v}), \mathbf{b})$. Let

$$m_{\mathcal{I},\mathcal{I}'} = \det(L(d(G, v), X_{d(G,v)})[\mathcal{I}, \mathcal{I}']) \in I_j(d(G, v), X_{d(G,v)}).$$

If $\mathcal{I}_{\{1,2\}} \cap \mathcal{I}'_{\{1,2\}} = \{a\}$, then Lemma 3.3 implies that for some matrix $J(x_v, \mathbf{a}'; M, \mathbf{b}') \in M_j(L(G, X_G))$, and

$$m_{\mathcal{I},\mathcal{I}'} = \det(J(x_{v^a}, \mathbf{a}'; M, \mathbf{b}')) = x_{v^a} \cdot \det(M) + \det(J(0, \mathbf{a}'; M, \mathbf{b}')).$$

If $|\mathcal{I}_{\{1,2\}}|, |\mathcal{I}'_{\{1,2\}}| = 1$ and $\mathcal{I}_{\{1,2\}} \cap \mathcal{I}'_{\{1,2\}} = \emptyset$, then $m_{\mathcal{I},\mathcal{I}'} = \det(J(0, \mathbf{a}'; M, \mathbf{b}'))$ for some $J(0, \mathbf{a}'; M, \mathbf{b}') \in M_j(L(G, X_G))$. On the other hand, since $\det(J(x, 1; 1, 0)) = \det(J(x, 0; 1, 1)) = x$, then

$$m_{\mathcal{I},\mathcal{I}'} \in \begin{cases} \{x_{v^i} \cdot \mathrm{minors}_{j-1}(\mathbf{a}, L(G-v, X_{G-v}))\}_{i=0}^{1} & \text{if } |\mathcal{I}_{\{1,2\}}| = 2, |\mathcal{I}'_{\{1,2\}}| = 1, \\ \{x_{v^i} \cdot \mathrm{minors}_{j-1}(L(G-v, X_{G-v}), \mathbf{b})\}_{i=0}^{1} & \text{if } |\mathcal{I}_{\{1,2\}}| = 1, |\mathcal{I}'_{\{1,2\}}| = 2. \end{cases}$$

Finally, since $\det(J(\mathrm{diag}(x_{v^1}, x_{v^0}), (1, 1); 0, (1, 1))) = -(x_{v^1} + x_{v^0})$, then Lemma 3.3 implies that $m_{\mathcal{I},\mathcal{I}'}$ belongs to

$$S_j(G, v) = \{x_{v^0} x_{v^1} \cdot \det(M) + (x_{v^0} + x_{v^1}) \cdot \det(J(0, \mathbf{a}'; M, \mathbf{b}')) \,:\, J(x_v, \mathbf{a}'; M, \mathbf{b}') \in M_{j-1}(L(G, X_G))\},$$

when $\mathcal{I}_{\{1,2\}}$ and $\mathcal{I}'_{\{1,2\}}$ are equal to $\{1, 2\}$.

Let $\mathrm{minors}_j(\mathbf{a}, L, \mathbf{b}) = \{\det(J(0, \mathbf{a}'; M, \mathbf{b}')) \,:\, J(0, \mathbf{a}'; M, \mathbf{b}') \in M_j\,(J(0, \mathbf{a}; L, \mathbf{b}))\}$. Therefore, for $1 \le j \le n-1$, the j-*th* critical ideal of the duplication has the following expression:

$$\begin{aligned} I_j(d(G,v), X_{d(G,v)}) \;=\;& \langle \mathrm{minors}_j(L(G-v, X_{G-v})), \{x_{v^i} \cdot \mathrm{minors}_{j-1}(L(G-v, X_{G-v}))\}_{i=0}^{1}, \\ & \mathrm{minors}_j(\mathbf{a}, L(G-v, X_{G-v})), \{x_{v^i} \cdot \mathrm{minors}_{j-1}(\mathbf{a}, L(G-v, X_{G-v}))\}_{i=0}^{1}, \\ & \mathrm{minors}_j(L(G-v, X_{G-v}), \mathbf{b}), \{x_{v^i} \cdot \mathrm{minors}_{j-1}(L(G-v, X_{G-v}), \mathbf{b})\}_{i=0}^{1}, \\ & \mathrm{minors}_j(\mathbf{a}, L(G-v, X_{G-v}), \mathbf{b}), S_j(G, v)\rangle. \end{aligned} \tag{7}$$

We assume, for the sake of clarity, that

$$\begin{aligned} \mathrm{minors}_0(\mathbf{a}, L(G-v, X_{G-v})) &= \mathrm{minors}_0(L(G-v, X_{G-v}), \mathbf{b}) = \emptyset, \\ \mathrm{minors}_1(\mathbf{a}, L(G-v, X_{G-v})) &= \{\mathbf{a}_i \,:\, 1 \le i \le n\}, \\ \mathrm{minors}_1(L(G-v, X_{G-v}), \mathbf{b}) &= \{\mathbf{b}_i \,:\, 1 \le i \le n\}, \\ \mathrm{minors}_1(\mathbf{a}, L(G-v, X_{G-v}), \mathbf{b}) &= S_1(G, v) = \emptyset, \text{ and} \\ \mathrm{minors}_2(\mathbf{a}, L(G-v, X_{G-v}), \mathbf{b}) &= S_2(G, v) = x_{v^0} x_{v^1}. \end{aligned}$$

Therefore, $I_1(d(G,v), X_{d(G,v)}) = \langle x_{v^0}, x_{v^1}, I_1(G, X_G)|_{x_v=0}\rangle$ and

$$I_2(d(G,v), X_{d(G,v)}) \subseteq \langle x_{v^0}, x_{v^1}, I_2(G, X_G)|_{x_v=0}\rangle.$$

Besides, it is not difficult to see that the ideal $I_n(d(G,v), X_{d(G,v)})$ is equal to

$$\begin{aligned} &\langle \{x_{v^i} \cdot \det(L(G-v, X_{G-v}))\}_{i=0}^{1}, \det(J(0, \mathbf{a}; L(G-v, X_{G-v}), \mathbf{b})), \\ &\{x_{v^i} \cdot \mathrm{minors}_{n-1}(\mathbf{a}, L(G-v, X_{G-v}))\}_{i=0}^{1}, \{x_{v^i} \cdot \mathrm{minors}_{n-1}(L(G-v, X_{G-v}), \mathbf{b})\}_{i=0}^{1}, S_n(G, v)\rangle. \end{aligned}$$

On the other hand, by Claim 3.4 we have that $I_j(G, X_G)|_{x_v=0}$ is equal to

$$\langle \mathrm{minors}_j(L(G-v, X_{G-v})), \mathrm{minors}_j(\mathbf{a}, L(G-v, X_{G-v})), \mathrm{minors}_j(L(G-v, X_{G-v}), \mathbf{b}), \mathrm{minors}_j(\mathbf{a}, L(G-v, X_{G-v}), \mathbf{b})\rangle,$$

for $1 \le j \le n-1$, and $I_n(G, X_G)|_{x_v=0} = \langle \det(L(G, X_G)|_{x_v=0})\rangle = \langle \det(J(0, \mathbf{a}; L(G-v, X_{G-v}), \mathbf{b}))\rangle$. Then

$$I_j(d(G,v), X_{d(G,v)}) \subseteq \langle x_{v^0}, x_{v^1}, I_j(G, X_G)|_{x_v=0}\rangle$$

for $1 \le j \le n$. Finally, it is not difficult to see, from previous equalities, that $I_j(d(G,v), X_{d(G,v)})$ is trivial if and only if $I_j(G, X_G)|_{x_v=0}$ is trivial. $\square$

Now we give a description of the critical ideals of the replication of a vertex of a signed multidigraph.

LEMMA 3.6. *Let G be a signed multidigraph with $n \geq 2$ vertices and v be a vertex of G. Then*

$$I_j(r(G,v), X_{r(G,v)}) \subseteq \langle x_{v^0}+1, x_{v^1}+1, I_j(G,X_G)|_{x_v=-1}\rangle,$$

for all $1 \leq j \leq n$. Moreover, $I_j(r(G,v), X_{r(G,v)})$ is trivial if and only if $I_j(G,X_G)|_{x_v=-1}$ is trivial.

PROOF. Suppose $L(G,X_G) = J(x_v, \mathbf{a}; L(G-v, X_{G-v}), \mathbf{b})$, for some $\mathbf{a}, \mathbf{b} \in \mathcal{P}^{n-1}$. Similarly to Lemma 3.5, we have that for $1 \leq j \leq n-1$, the j-*th* critical ideal of the replication has the following expression:

$$\begin{aligned} I_j(r(G,v), X_{r(G,v)}) \;=\; & \langle \text{minors}_j(L(G-v), X), \{(x_{v^i}+1)\cdot \text{minors}_{j-1}(L(G-v, X_{G-v}))\}_{i=0}^{1}, \\ & \text{minors}_j(\mathbf{a}, L(G-v, X_{G-v})), \{(x_{v^i}+1)\cdot \text{minors}_{j-1}(\mathbf{a}, L(G-v, X_{G-v}))\}_{i=0}^{1}, \\ & \text{minors}_j(L(G-v, X_{G-v}), \mathbf{b}), \{(x_{v^i}+1)\cdot \text{minors}_{j-1}(L(G-v, X_{G-v}), \mathbf{b})\}_{i=0}^{1}, \\ & R_j(G,v), \widetilde{S}_j(G,v)\rangle, \end{aligned} \tag{8}$$

where $R_j(G,v) = \{\det(J(-1,\mathbf{a}'; M, \mathbf{b}')) = -\det(M) + \det(J(0,\mathbf{a}'; M, \mathbf{b}')) \,:\, J(x_v, \mathbf{a}'; M, \mathbf{b}') \in M_j(L(G,X_G))\}$ and $\widetilde{S}_j(G,v) = \{(x_{v^0}+1)(x_{v^1}+1)\cdot\det(M) + ((x_{v^0}+1)+(x_{v^1}+1))\cdot\det(J(-1,\mathbf{a}'; M, \mathbf{b}')) \,:\, J(x_v, \mathbf{a}'; M, \mathbf{b}') \in M_{j-1}(L(G,X_G))\}$. Besides, the n-*th* critical ideal of the replication has the following expression:

$$\begin{aligned} I_n(r(G,v), X_{r(G,v)}) \;=\; & \langle \{(x_{v^i}+1)\cdot \det(L(G-v, X_{G-v}))\}_{i=0}^{1}, \det(J(-1,\mathbf{a}; L(G-v, X_{G-v}), \mathbf{b})), \\ & \{(x_{v^i}+1)\cdot \text{minors}_{n-1}(\mathbf{a}, L(G-v, X_{G-v}))\}_{i=0}^{1}, \\ & \{(x_{v^i}+1)\cdot \text{minors}_{n-1}(L(G-v, X_{G-v}), \mathbf{b})\}_{i=0}^{1}, \widetilde{S}_n(G,v)\rangle. \end{aligned}$$

On the other hand, by Claim 3.4 we have that $I_j(G,X_G)|_{x_v=-1}$ is equal to

$$\langle \text{minors}_j(L(G-v, X_{G-v})), \text{minors}_j(\mathbf{a}, L(G-v, X_{G-v})), \text{minors}_j(L(G-v, X_{G-v}), \mathbf{b}), R_j(G,v)\rangle,$$

for all $1 \leq j \leq n-1$, and $I_n(G,X_G)|_{x_v=-1} = \langle \det(L(G,X_G))|_{x_v=-1}\rangle = \langle \det(J(-1,\mathbf{a}; L(G-v, X_{G-v}), \mathbf{b}))\rangle$. Therefore,

$$I_j(r(G,v), X_{r(G,v)}) \subseteq \langle x_{v^0}+1, x_{v^1}+1, I_j(G,X_G)|_{x_v=-1}\rangle$$

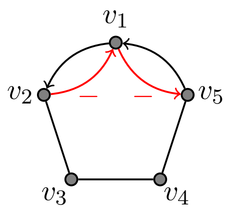
for all $1 \leq j \leq n$. Finally, it is clear that $I_j(r(G,v), X_{r(G,v)})$ is trivial if and only if $I_j(G,X_G)_{x_v=-1}$ is trivial. $\square$

REMARK 3.7. *Note that $I_j(G,X_G)|_{x_v=0}$ is trivial if and only if there exists $p \in I_j(G,X_G)$ such that $p = x_v q + 1$, and $I_j(G,X_G)_{x_v=-1}$ is trivial if and only if there exists $p \in I_j(G,X_G)$ such that $p = (x_v+1)q+1$. However, $p \in \mathcal{P}[X]$ and $a, b \in \mathcal{P}$ such that $p|_{\{x_u=a, x_v=b\}} = 1$ do not imply that $p = (x_u - a)\cdot(x_v - b)\cdot q$ for some $q \in \mathcal{P}[X \setminus \{x_u, x_v\}]$.*

Next example shows a signed digraph satisfying the equality in the ideal inclusions of Lemmas 3.5 and 3.6.

EXAMPLE 3.8. *Let G be the cycle with five vertices, where the arcs v_2v_1 and v_1v_5 have negative sign, see Figure 9. It is not difficult to check that the algebraic co-rank of the graph G is equal to 3, when $\mathcal{P} = \mathbb{Z}$. Since $I_4(G,X_G)$ is given by $\langle x_1x_2 + x_4 + 1, x_2x_3 - x_5 - 1, x_3x_4 + x_1 - 1, x_4x_5 - x_2 - 1, x_1x_5 + x_3 + 1\rangle$, then $I_4(G,X_G)|_{x_{v_1}=0} = \langle x_3+1, x_4+1, x_3x_4 - 1, x_2x_3 - x_5 - 1, x_4x_5 - x_2 - 1\rangle$ $= \langle x_3+1, x_4+1, x_2+x_5+1\rangle$ and*

$$\begin{aligned} I_4(G,X_G)|_{x_{v_1}=-1} &= \langle -x_5 + x_3 + 1, -x_2 + x_4 + 1, x_4x_5 - x_2 - 1, x_3x_4 - 2, x_2x_3 - x_5 - 1\rangle \\ &= \langle x_3 - x_5 + 1, x_2 - x_4 - 1, x_4x_5 - x_4 - 2\rangle. \end{aligned}$$



$$L(G, X_G) = \begin{bmatrix} x_1 & -1 & 0 & 0 & 1 \\ 1 & x_2 & -1 & 0 & 0 \\ 0 & -1 & x_3 & -1 & 0 \\ 0 & 0 & -1 & x_4 & -1 \\ -1 & 0 & 0 & -1 & x_5 \end{bmatrix}$$

FIGURE 9. A signed multidigraph G with five vertices and its generalized Laplacian matrix.

On the other hand, the 4-th critical ideal $I_4(d(G, v_1), X_{d(G,v_1)})$ *is equal to* $\langle x_{v_1^0}, x_{v_1^1}, x_3 + 1, x_4 + 1, x_2 + x_5 + 1\rangle$*, and the 4-th critical ideal* $I_4(r(G, v_1), X_{r(G,v_1)})$ *is equal to*

$$\langle x_{v_1^0} + 1, x_{v_1^1} + 1, x_3 - x_5 + 1, x_2 - x_4 - 1, x_4x_5 - x_4 - 2\rangle.$$

Thus this example satisfies also the opposite inclusions of Lemma 3.5 and 3.6.

Successive applications of Lemmas 3.5 and 3.6 leads to the following general result:

THEOREM 3.9. *Let G be a signed multidigraph with $n \geq 2$ vertices and $\mathbf{d} \in \mathbb{Z}^n$. Then the j-th critical ideal $I_j(G^{\mathbf{d}}, X_{G^{\mathbf{d}}})$ is included in the ideal*

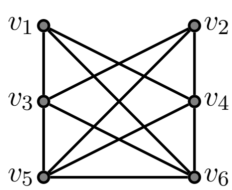
$$\left\langle \{\{x_{v^i}\}_{i=0}^{\mathbf{d}_v} : \mathbf{d}_v \geq 1\}, \{\{x_{v^i} + 1\}_{i=0}^{-\mathbf{d}_v} : \mathbf{d}_v \leq -1\}, I_j(G, X_G)|_{\{x_v=-1\,:\,\mathbf{d}_v \leq -1\} \cup \{x_v=0\,:\,\mathbf{d}_v \geq 1\}} \right\rangle,$$

for $1 \leq j \leq n$. Moreover, $I_j(G^{\mathbf{d}}, X_{G^{\mathbf{d}}})$ is trivial if and only if $I_j(G, X_G)|_{\{x_v=-1\,:\,\mathbf{d}_v \leq -1\} \cup \{x_v=0\,:\,\mathbf{d}_v \geq 1\}}$ is trivial. That is, there exists $p \in I_j(G, X_G)$ such that $p|_{\{x_v=-1\,:\,\mathbf{d}_v \leq -1\}, \{x_v=0\,:\,\mathbf{d}_v \leq 1\}} = 1$.

PROOF. The result turns out by the successive applications of Lemmas 3.5 and 3.6. $\square$

Next example illustrates Lemmas 3.5 and 3.6 and Theorem 3.9.

EXAMPLE 3.10. *Let G be the graph on Figure 10. By using a computer algebra system, we can*



$$L(G, X_G) = \begin{bmatrix} x_1 & 0 & -1 & -1 & 0 & -1 \\ 0 & x_2 & -1 & -1 & -1 & 0 \\ -1 & -1 & x_3 & 0 & -1 & -1 \\ -1 & -1 & 0 & x_4 & -1 & -1 \\ 0 & -1 & -1 & -1 & x_5 & -1 \\ -1 & 0 & -1 & -1 & -1 & x_6 \end{bmatrix}$$

FIGURE 10. A graph G with eight vertices and its generalized Laplacian matrix.

see that $\gamma_{\mathbb{Z}}(G) = 3$ *and its non-trivial critical ideals are the following:*

$$\begin{aligned}
I_4(G, X_G) &= \langle x_3, x_4, x_1x_2 + 1, (x_1 - 1)x_6 - 2, (x_2 - 1)x_5 - 2, x_1x_5 + x_5 + 2x_1, x_2x_6 + x_6 + 2x_2,\\
&\quad x_5x_6 + x_6 + x_5 + 2\rangle,\\
I_5(G, X_G) &= \langle x_2x_4x_5(x_6 + 1) - x_4x_6, x_2x_3x_4 + x_2x_3x_6 + x_2x_4x_6 + 2x_2x_3 + 2x_2x_4 + x_3x_6 + x_4x_6,\\
&\quad x_1x_3x_4 + x_1x_3x_5 + x_1x_4x_5 + 2x_1x_3 + 2x_1x_4 + x_3x_5 + x_4x_5, x_1x_4x_6(x_5 + 1) - x_4x_5,\\
&\quad x_1x_4(x_2x_6 + x_2 + x_6) - x_4, x_1x_3(x_2x_6 + x_2 + x_6) - x_3, (x_3 + x_4)(x_5x_6 + x_5 + x_6 + 2) + x_3x_4,\\
&\quad x_1x_2(x_6 + x_5) + x_5x_6(x_1 + x_2) + 2(x_1x_2 + x_1x_6 + x_2x_5) - x_5 - x_6 - 2\rangle,\\
I_6(G, X_G) &= \langle \det(L(G, X_G))\rangle.
\end{aligned}$$

From these equalities and Theorem 3.9, we can easily obtain that the critical ideals $I_4(d(G, v_i), X_{d(G,v_i)})$ *and* $I_4(r(G, v_i), X_{r(G,v_i)})$ *are trivial, for* $i \in \{1, 2\}$ *and* $j \in \{3, 4\}$*. Further, the ideals* $I_4(G^{\mathbf{e}_1-\mathbf{e}_6}, X)$, $I_4(G^{\mathbf{e}_1-\mathbf{e}_5}, X)$, $I_4(G^{\mathbf{e}_2-\mathbf{e}_5}, X)$, $I_4(G^{\mathbf{e}_2-\mathbf{e}_6}, X)$, $I_4(G^{\mathbf{e}_5-\mathbf{e}_6}, X)$, $I_4(G^{\mathbf{e}_6-\mathbf{e}_5}, X)$ *are also trivial. On the other hand,*

$$I_4(d(G, v_6), X) = \langle x_6, x_{6'}, I_4(G, X_G)|_{x_6=0}\rangle = \langle x_6, x_{6'}, 2, x_3, x_4, x_5, x_1x_2 + 1\rangle,$$

$$\begin{aligned}
I_5(d(G, v_6), X) &= \langle x_3x_6, x_3x_{6'}, x_4x_6, x_4x_{6'}, x_3x_5, x_4x_5, x_6(x_1x_2 + 1), x_{6'}(x_1x_2 + 1),\\
&\quad x_6(x_2x_5 - x_5 - 2), x_{6'}(x_2x_5 - x_5 - 2), x_6(x_1x_5 + x_5 + 2x_1), x_{6'}(x_1x_5 + x_5 + 2x_1),\\
&\quad x_6x_{6'}(x_1 - 1) - 2(x_6 + x_{6'}), x_6x_{6'}(x_2 + 1) + 2x_2(x_6 + x_{6'}),\\
&\quad (x_6x_{6'} + x_6 + x_{6'})(x_5 + 1) + (x_6 + x_{6'}), x_3x_4 + 2x_3 + 2x_4,\\
&\quad x_3(x_1x_2 - 1), x_4(x_1x_2 - 1), x_1x_2x_5 + 2x_1x_2 + 2x_2x_5 - x_5 - 2\rangle\\
&\subsetneq \langle x_6, x_{6'}, I_5(G, X_G)|_{x_6=0}\rangle, \text{ and}
\end{aligned}$$

$$\begin{aligned}
I_5(G^{\mathbf{e}_6-\mathbf{e}_5}, X) &= \langle 2(x_{5'}+1), 2(x_5+1), x_5x_{5'}-1, x_6+x_{6'}, x_6(x_1-1), x_6(x_2+1), x_6(x_5+1),\\
&\quad x_6(x_{5'}+1), x_3, x_4, x_1x_2-2x_2-1\rangle\\
&\subsetneq \langle x_5 + 1, x_{5'} + 1, x_6, x_{6'}, I_5(G, X_G)|_{\{x_6=0,x_5=-1\}}\rangle.
\end{aligned}$$

Note that, $I_5(G, X_G)|_{\{x_6=0,x_5=-1\}} = \langle x_3, x_4, x_1x_2 - 2x_2 - 1\rangle$*, and* $x_5x_{5'} - 1 = (x_5 + 1)(x_{5'} + 1) - (x_5 + 1) - (x_{5'} + 1)$.

As consequence of Theorem 3.9, we get the following bound for the algebraic co-rank of a signed multidigraph with twins.

Corollary 3.11. *Let* G *be a signed multidigraph with* n *vertices. Then* $\gamma_{\mathcal{P}}(G^{\mathbf{d}}) \leq n$*, for all* $\mathbf{d} \in \mathbb{Z}^n$*. Moreover,* $\gamma_{\mathcal{P}}(G^{\mathbf{d}}) = \gamma_{\mathcal{P}}(G^{\mathrm{supp}(\mathbf{d})})$*, where*

$$\mathrm{supp}(\mathbf{d})_v = \begin{cases} -1 & \textit{if } \mathbf{d}_v < 0,\\ 1 & \textit{if } \mathbf{d}_v > 0,\\ 0 & \textit{otherwise.}\end{cases}$$

Proof. Let $g = \gamma_{\mathcal{P}}(G^\delta)$ and $\mathbf{d} \in \mathbb{Z}^n$ such that $\mathrm{supp}(\mathbf{d}) = \delta$. By applying Theorem 3.9 to G^δ, we have that the non-trivial critical ideal $I_{g+1}(G^\delta, X_{G^\delta})$ is included in the ideal

$$\Big\langle \{x_{v^0}, x_{v^1} : \delta_v = 1\}, \{x_{v^0} + 1, x_{v^1} + 1 : \delta_v = -1\}, I_{g+1}(G, X_G)|_{\{x_v=-1\,:\,\delta_v=-1\}\cup\{x_v=0\,:\,\delta_v=1\}}\Big\rangle$$

and $I_{g+1}(G, X_G)|_{\{x_v=-1\,:\,\mathbf{d}_v\leq-1\}\cup\{x_v=0\,:\,\mathbf{d}_v\geq1\}} = I_{g+1}(G, X_G)|_{\{x_v=-1\,:\,\delta_v=-1\}\cup\{x_v=0\,:\,\delta_v=1\}} \neq \langle 1\rangle$. Since $(x_vp)|_{x_v=0} = 0$ and $((x_v + 1)p)|_{x_v=-1} = 0$ for all $p \in \mathcal{P}[X]$, then

$$I_{g+1}(G^\delta, X_{G^\delta})|_{\{x_v=-1\,:\,(\mathbf{d}-\delta)_v\leq-1\}\cup\{x_v=0\,:\,(\mathbf{d}-\delta)_v\geq1\}} \subseteq I_{g+1}(G, X_G)|_{\{x_v=-1\,:\,\mathbf{d}_v\leq-1\}\cup\{x_v=0\,:\,\mathbf{d}_v\geq1\}} \neq \langle 1\rangle.$$

Therefore, applying Theorem 3.9 to G^δ and $\mathbf{d}-\delta$, we have that

$$I_{g+1}((G^\delta)^{\mathbf{d}-\delta}, X_{(G^\delta)^{\mathbf{d}-\delta}}) = I_{g+1}(G^{\mathbf{d}}, X_{G^{\mathbf{d}}}) \neq \langle 1 \rangle$$

for all $\mathbf{d}$ with $\operatorname{supp}(\mathbf{d}) = \delta$, that is,

$$\gamma_{\mathcal{P}}(G^{\mathbf{d}}) = \gamma_{\mathcal{P}}(G^\delta).$$

On the other hand, since $I_{n+1}(G, X_G) = \langle 0 \rangle$, then

$$I_{n+1}(G^{\mathbf{d}}, X) \subseteq \langle \{x_v, \ldots, x_{v^{\mathbf{d}_v}} : \mathbf{d}_v \geq 1\}, \{x_v + 1, \ldots, x_{v^{-\mathbf{d}_v}} + 1 : \mathbf{d}_v \leq -1\} \rangle \neq \langle 1 \rangle$$

and we get the result. □

Next example show us that the upper bound given in Corollary 3.11 is tight.

EXAMPLE 3.12. *Let K_n be the complete graph with $n \geq 2$ vertices. By* [**22**, Theorem 3.15] *and* [**22**, Theorem 3.16], *we have that $\gamma_{\mathcal{P}}(K_n) = 1$ and $I_n(K_n, X_{K_n}) = \langle P \rangle$, where*

$$P = \prod_{j=1}^{n}(x_j + 1) - \sum_{i=1}^{n}\prod_{j \neq i}(x_j + 1).$$

Since the evaluation of P at $\{x_1 = 0, \cdots, x_{n-1} = 0, x_n = -1\}$ is equal to -1, then Theorem 3.9 and Corollary 3.11 imply $\gamma_{\mathcal{P}}(K_n^{\mathbf{d}}) = n$ for any $\mathbf{d} \in \mathbb{Z}^n$ such that $\mathbf{d}_i \geq 1$ when $i \in [n-1]$ and $\mathbf{d}_n \leq -1$. Also, [**22**, Theorem 3.16] *implies*

$$I_{n-1}(K_n, X) = \left\langle \left\{ \prod_{i \in \mathcal{I}}(x_i + 1) : \mathcal{I} \subseteq [n] \text{ and } |\mathcal{I}| = n-2 \right\} \right\rangle.$$

Since $I_{n-1}(K_n, X)_{\{x_i = 0 \,:\, i \in [n-1]\}} = \langle 1 \rangle$, then Theorem 3.9 and Corollary 3.11 imply $\gamma_{\mathcal{P}}(K_n^{\mathbf{d}}) = n-1$ for any $\mathbf{d} \in \mathbb{Z}^{n-1}$ such that $\mathbf{d}_i \geq 1$ when $i \in [n-1]$.

Corollary 3.11 is useful in classification and understanding of the graphs with algebraic co-rank less than or equal to an integer fixed k. A graph G is *forbidden* for the graphs with algebraic co-rank less than or equal to k if and only if $\gamma(G) \geq k+1$. One step in the classification is to prove that the graphs with no forbidden induced subgraph have algebraic co-rank at most k. This was done in [**1**] for $k = 1, 2$, by a large computation of the minors of their corresponding generalized Laplacian matrices. But now this can be easily checked by evaluations of the critical ideals of few graphs, as performed in [**2**, Section 2].

2.1. Some conjectures. In light of the previous results we conjecture the following.

CONJECTURE 3.13. *If $\gamma_{\mathcal{P}}(G - v) = \gamma_{\mathcal{P}}(G)$ for all $v \in V(G)$, then G has at least a pair of twin vertices.*

CONJECTURE 3.14. *If $\gamma_{\mathcal{P}}(G) < \frac{n}{2}$ with $n \geq 5$ vertices, then G has at least a pair of twin vertices.*

CONJECTURE 3.15. *If G is twin-free, then $\gamma_{\mathcal{P}}(G) \geq \frac{n}{2}$.*

Note that Conjecture 3.13 imply Conjecture 3.14. And Conjecture 3.14 is equivalent to Conjecture 3.15.

3. Critical ideals of graphs with twin vertices

In this section we give a description of the critical ideals of $d^k(G,v)$ and $r^k(G,v)$ in terms of some of the critical ideals of G. More precisely, if $g' = \gamma_{\mathcal{P}}(d(G,v))$ and λ is a constant that depends on G and v, then Theorem 3.19 gives a description of $I_{g'+k}(d^{i+\lambda+k}(G,v), X_{d^{i+\lambda+k}(G,v)})$ in terms of $I_{g'+k}(d^{k-1}(G,v), X_{d^{k-1}(G,v)})$. And Theorem 3.22 gives a similar description of the critical ideals of $I_{g'+k}(r^{i+\lambda+k}(G,v), X_{r^{i+\lambda+k}(G,v)})$. We begin by giving a description of the critical ideals of $d^k(G,v)$ in terms of the critical ideals of G and some minors of $G-v$. This description generalizes the description of the critical ideals of $d(G,v)$ given in Equation 7.

Before to establish the result, we introduce some notation. Given a subset S of the natural numbers, let $\binom{S}{l}$ denote the set of all subsets of S of cardinality exactly l. Moreover, if v is a vertex of a signed multidigraph, let $P_l^S(v) = \left\{\prod_{c\in C} x_{v^c} \,:\, C \in \binom{S}{l}\right\}$. We take $P_0^S(v) = \{1\}$. And for simplicity, $P_l^k(v)$ denote $P_l^{\{0\}\cup[k]}(v)$.

Lemma 3.16. *Let G be a signed multidigraph with n vertices and $v \in V(G)$. If $k, j \geq 1$ and $m = \min(k, j-1)$, then*

$$\begin{aligned} I_j(d^k(G,v), X_{d^k(G,v)}) \;=\; & \Big\langle \left\{P_l^k(v)\cdot I_{j-l}(G,X_G)|_{x_v=0}\right\}_{l=0}^{m-1}, P_m^k(v)\cdot I_{j-m}(G-v,X_{G-v}),\\ & P_m^k(v)\cdot \mathrm{minors}_{j-m}(\mathbf{a}, L(G-v,X_{G-v})),\\ & P_m^k(v)\cdot \mathrm{minors}_{j-m}(L(G-v,X_{G-v}),\mathbf{b}), S_j^k(G,v)\Big\rangle, \end{aligned}$$

where $S_j^k(G,v)$ is equal to $P_j^k(v)$ when $j \leq k+1$, and is equal to

$$\left\{\det(M)\cdot\prod_{t=0}^{k} x_{v^t} + \det(J(0,\mathbf{a}';M,\mathbf{b}'))\cdot\sum_{t=0}^{k}\prod_{s\neq t} x_{v^s} \;:\; J(x_v,\mathbf{a}';M,\mathbf{b}') \in M_{j-k}(L(G,X_G))\right\}$$

when $j \geq k+2$.

Proof. Suppose the generalized Laplacian matrix $L(d^k(G,v), X_{d^k(G,v)})$ of $d^k(G,v)$ is equal to

$$J(\mathrm{diag}(x_{v^0},\dots,x_{v^k}),\mathbf{a}; L(G-v,X_{G-v}),\mathbf{b}),$$

for some $\mathbf{a},\mathbf{b} \in \mathcal{P}^{n-1}$. Let $\mathcal{I},\mathcal{I}' \subseteq [n+k]$ be two sets of size j, $h = |\mathcal{I}\cap[k+1]|$, $h' = |\mathcal{I}'\cap[k+1]|$, and

$$m_{\mathcal{I},\mathcal{I}'} = \det(L(d^k(G,v),X_{d^k(G,v)})[\mathcal{I},\mathcal{I}']).$$

Clearly $0 \leq h, h' \leq m+1$. If $h, h' = 0$, then $m_{\mathcal{I},\mathcal{I}'} \in \mathrm{minors}_j(L(G-v,X_{G-v}))$ and $m_{\mathcal{I},\mathcal{I}'} \in I_j(G-v,X_{G-v})$. First suppose $h=0$. If $h' \geq 2$, then two columns of $L(d^k(G,v),X_{d^k(G,v)})[\mathcal{I},\mathcal{I}']$ are equal, and $m_{\mathcal{I},\mathcal{I}'} = 0$. Also, if $h'=1$, then $m_{\mathcal{I},\mathcal{I}'} \in \mathrm{minors}_j(\mathbf{a}, L(G-v,X_{G-v}))$. We can use similar arguments when $h'=0$. Thus, we assume that $h,h' \geq 1$.

Now by Lemma 3.3 we have that

$$m_{\mathcal{I},\mathcal{I}'} = \begin{cases} 0 & \text{if } |h-h'| \geq 2,\\ \det\begin{bmatrix} P & 1 \end{bmatrix}\cdot\det\begin{bmatrix} \mathbf{b}' \\ Q \end{bmatrix} & \text{if } h-h'=1,\\ \det\begin{bmatrix} P \\ \mathbf{1} \end{bmatrix}\cdot\det\begin{bmatrix} \mathbf{a}'^T & Q \end{bmatrix} & \text{if } h'-h=1,\\ \det(P)\cdot\det(Q) - \det(J(P,\mathbf{1};0,\mathbf{1}))\cdot\det(J(0,\mathbf{a}';Q,\mathbf{b}')) & \text{if } h=h', \end{cases}$$

for some submatrix P of $\text{diag}(x_{v^0}, ..., x_{v^k})$, some submatrix Q of $L(G-v, X_{G-v})$, and some subvectors $\mathbf{a}'$ of $\mathbf{a}$ and $\mathbf{b}'$ of $\mathbf{b}$. Clearly, $\det \begin{bmatrix} P & \mathbf{1} \end{bmatrix} \neq 0$ if and only if (up to row and column permutations)

$$P = \begin{bmatrix} \text{diag}(x_{v^{i_1}}, ...x_{v^{i_{h'}}}) \\ \mathbf{0} \end{bmatrix}.$$

If $h - h' = 1$, then $m_{\mathcal{I},\mathcal{I}'} \in P^k_{h'}(v) \cdot \text{minors}_{j-h'}(L(G-v, X_{G-v}), \mathbf{b}) \subsetneq P^k_{h'}(v) \cdot I_{j-h'}(G, X_G)|_{x_v=0}$, for all $1 \leq h' \leq m$. Similarly, if $h' - h = 1$, then $m_{\mathcal{I},\mathcal{I}'} \in P^k_h(v) \cdot \text{minors}_{j-h}(L(G-v, X_{G-v}), \mathbf{b}) \subsetneq P^k_h(v) \cdot I_{j-h}(G, X_G)|_{x_v=0}$, for all $1 \leq h \leq m$. On the other hand, if $h = h'$ we have the following cases:

Case I: If P has at least two zero rows, then $\det(P) = 0$, $\det(J(P, \mathbf{1}; 0, \mathbf{1})) = 0$, and thus $m_{\mathcal{I},\mathcal{I}'} = 0$.

Case II: If P has only one zero row, then $\det(P) = 0$, $\det(J(P, \mathbf{1}; 0, \mathbf{1})) = \prod_{t=1}^{h-1} x_{v^{i_t}}$, and

$$m_{\mathcal{I},\mathcal{I}'} = \prod_{t=1}^{h-1} x_{v^{i_t}} \cdot \det(J(0, \mathbf{a}; Q, \mathbf{b}')),$$

for some $(j-h+1) \times (j-h+1)$-submatrix $J(0, \mathbf{a}; Q, \mathbf{b}')$ of $L(G, X_G)|_{x_v=0}$. Thus $m_{\mathcal{I},\mathcal{I}'} \in P^k_{h-1}(v) \cdot \text{minors}_{j-h+1}(\mathbf{a}, L(G-v, X_{G-v}), \mathbf{b}) \subsetneq P^k_{h-1}(v) \cdot I_{j-h+1}(G, X_G)|_{x_v=0}$, for all $2 \leq h \leq m-1$.

Case III: If P has no zero row, then

$$m_{\mathcal{I},\mathcal{I}'} = \begin{cases} \prod_{t=1}^h x_{v^{i_t}} \cdot \det(Q) + \sum_{t=1}^h \prod_{s \neq t} x_{v^{i_s}} \cdot \det(J(0, \mathbf{a}'; Q, \mathbf{b}')) & \text{if } h < j, \\ \prod_{t=1}^h x_{v^{i_t}} & \text{if } h = j, \end{cases}$$

for some $(j-h+1) \times (j-h+1)$-submatrix $J(0, \mathbf{a}'; Q, \mathbf{b}')$ of $L(G, X_G)|_{x_v=0}$, and for all $1 \leq h \leq m$. Moreover, since

$$\sum_{t=1}^h \prod_{s \neq t} x_{v^{i_s}} \cdot \det(J(0, \mathbf{a}'; Q, \mathbf{b}')) \in P^k_{h-1}(v) \cdot \text{minors}_{j-h+1}(\mathbf{a}, L(G-v, X_{G-v}), \mathbf{b})$$

and $\prod_{t=1}^h x_{v^{i_t}} \cdot \det(Q) = m_{\mathcal{I},\mathcal{I}'} - \sum_{t=1}^h \prod_{s \neq t} x_{v^{i_s}} \cdot \det(J(0, \mathbf{a}'; Q, \mathbf{b}')) \in P^k_h(v) \cdot \text{minors}_{j-h}(L(G-v, X_{G-v})) \subsetneq P^k_h(v) \cdot I_{j-h}(G, X_G)|_{x_v=0}$ for all $0 \leq h \leq m-1$, then we get the result. $\square$

REMARK 3.17. *Note that* $I_j(G, X_G)|_{x_v=0}$ *is equal to*

$$\langle \text{minors}_j(L(G-v, X_{G-v})), \text{minors}_j(\mathbf{a}, L(G-v, X_{G-v})), \\ \text{minors}_j(L(G-v, X_{G-v}), \mathbf{b}), \text{minors}_j(\mathbf{a}, L(G-v, X_{G-v}), \mathbf{b}) \rangle$$

and the i*-th critical ideal* $I_i(T_{k+1}, X_{T_{k+1}})$ *of the graph with* $k+1$ *isolated vertices is equal to* $\langle P^k_i(v) \rangle$. *Moreover, if* $m = \min(k, j-1)$, *then*

$$I_j(d^k(G, v), X_{d^k(G,v)})|_{x_{v^0}=0} = \left\langle \left\{ P^{[k]}_l(v) \cdot I_{j-l}(G, X_G)|_{x_v=0} \right\}_{l=0}^m \right\rangle.$$

That is, $I_j(d^k(G, v), X_{d^k(G,v)})|_{x_{v^0}=0}$ *behaves almost equally as the* j*-th critical ideal of the disjoint union of* T_{k+1} *and* G.

In next example we show how to use the description of $I_j(d^k(G, v), X_{d^k(G,v)})|_{x_{v^0}=0}$.

EXAMPLE 3.18. *Let* Q_3 *be the hypercube with* $V(Q_3) = \{v_i\}_{i=1}^8$. *The reader can check that* $\gamma_{\mathbb{Z}}(Q_3) = 4$, $\gamma_{\mathbb{Z}}(d(Q_3, v_8)) = 5$, $I_7(d(Q_3, v_8), X_{d(Q_3,v_8)})|_{x_8=0} = \langle x_{v_8^1} \cdot I_6(Q_3, X_{Q_3})|_{x_8=0}, I_7(Q_3, X_{Q_3})|_{x_8=0} \rangle$, *where*

$$I_6(Q_3, X_{Q_3})_{x_8=0} = \langle x_1 - x_6, x_2 - 3x_7, x_3 - x_6, x_4 - x_7, x_5 - x_7, x_6 x_7 - 1 \rangle, \text{ and}$$

$$\begin{aligned}
I_7(Q_3, X_{Q_3})|_{x_8=0} \;=\; & \langle x_2x_4x_6 - x_4x_5x_6 - x_4x_6x_7 - x_5x_6x_7 - x_2 - x_4 + 2x_5 + 2x_7,\\
& x_2x_3x_5 - x_3x_4x_5 - x_3x_4x_7 - x_3x_5x_7 - x_2 + 2x_4 - x_5 + 2x_7,\\
& x_1x_2x_7 - x_1x_4x_5 - x_1x_4x_7 - x_1x_5x_7 - x_2 + 2x_4 + 2x_5 - x_7,\\
& x_1x_3x_7 - x_1x_4x_6 + x_3x_4x_6 - x_1x_6x_7 + x_1 - 2x_3 + x_6,\\
& x_1x_3x_5 + x_1x_4x_6 - x_3x_4x_6 - x_3x_5x_6 - 2x_1 + x_3 + x_6,\\
& x_1x_4x_5x_6 + x_1x_4x_6x_7 + x_1x_5x_6x_7 - x_1x_5 - x_5x_6 - 2x_4x_6 - 2x_1x_7 + 3,\\
& x_3x_4x_5x_6 + x_3x_4x_6x_7 + x_3x_5x_6x_7 - 2x_3x_5 - 2x_4x_6 - x_3x_7 - x_6x_7 + 3\rangle.
\end{aligned}$$

Now we are ready to give a more accurate description of some critical ideals of $d^{i+k}(G, v)$. Given $d, d' \geq 0$, let

$$\lambda(d, d') = \begin{cases} 0 & \text{if } d, d' = 0,\\ 1 & \text{otherwise.}\end{cases}$$

THEOREM 3.19. *Let G be a signed multidigraph, $v \in V(G)$, $g = \gamma_{\mathcal{P}}(G)$, $g' = \gamma_{\mathcal{P}}(d(G, v))$, $d = g - \gamma_{\mathcal{P}}(G - v)$, $d' = g' - g$, and $\lambda = \lambda(d, d')$. If $g \geq 1$, then $0 \leq d + d' \leq 2$ and*

$$I_{g'+k}(d^{k+\lambda+i}(G, v), X_{d^{k+\lambda+i}(G,v)}) = \left\langle P_k^{k+\lambda+i}(v), \left\{P_l^{k+\lambda+i}(v) \cdot I_{g'+k-l}(G, X_G)|_{x_v=0}\right\}_{l=0}^{k-1}\right\rangle$$

for all $k \geq 1$ and $i \geq 0$.

PROOF. Since $I_j(G, X_G)_{x_v=0} \subseteq I_{j-2}(G-v, X_{G-v})$, then by Lemma 3.5 we have that $0 \leq d+d' \leq 2$. On the other hand, since $m = k+\lambda$, then Lemma 3.16 implies that $I_{g'+k}(d^{k+\lambda+i}(G, v), X_{d^{k+\lambda+i}(G,v)})$ is equal to

$$\left\langle \left\{P_l^{k+\lambda+i}(v) \cdot I_{g'+k-l}(G, X_G)|_{x_v=0}\right\}_{l=0}^{k+\lambda-1}, P_{k+\lambda}^{k+\lambda+i}(v) \cdot I_{g'-\lambda}(G - v, X_{G-v}),\right.$$
$$\left. P_{k+\lambda}^{k+\lambda+i}(v) \cdot \operatorname{minors}_{g'-\lambda}(\mathbf{a}, L(G - v, X_{G-v})), P_{k+\lambda}^{k+\lambda+i}(v) \cdot \operatorname{minors}_{g'-\lambda}(L(G - v, X_{G-v}), S_{g'+k}^{k+\lambda+i}(G, v)\right\rangle.$$

If $\lambda(d, d') = 0$, then $I_{g'}(G - v, X_{G-v})$ is trivial and $\langle P_k^{k+i}(v) \cdot I_{g'}(G - v, X_{G-v})\rangle = \langle P_k^{k+i}(v)\rangle$. Therefore,

$$I_{g'+k}(d^{k+i}(G, v), X_{d^{k+i}(G,v)}) = \left\langle \left\{P_l^{k+i}(v) \cdot I_{g'+k-l}(G, X_G)|_{x_v=0}\right\}_{l=0}^{k-1}, P_k^{k+i}(v)\right\rangle.$$

Thus, assume that $\lambda(d, d') > 0$. If $d' = 0$, then $I_{g'}(G, X_G)$ is trivial, and therefore $I_{g'}(G, X_G)|_{x_v=0}$ is also trivial. Also, if $d' > 0$, then $I_{g'}(d(G, v), X_{d(G,v)})$ is trivial, and by Lemma 3.5 we have that $I_{g'}(G, X_G)|_{x_v=0}$ is trivial. Therefore,

$$I_{g'+k}(d^{k+i+1}(G, v), X_{d^{k+i+1}(G,v)}) = \left\langle \left\{P_l^{k+i+1}(v) \cdot I_{g'+k-l}(G, X_G)|_{x_v=0}\right\}_{l=0}^{k-1}, P_k^{k+i+1}(v)\right\rangle.$$

□

When $k = 1$, Theorem 3.19 can be reduced to the following simpler form

$$I_{g'+1}(d^{i+1}(G, v), X_{d^{i+1}(G,v)}) = \langle x_{v^0}, x_{v^1}, \ldots, x_{v^{i+1}}, I_{g'+1}(G, X_G)|_{x_v=0}\rangle,$$

for all $i \geq \lambda(d, d')$. Which is similar to Lemma 3.5. For a fixed integer $k' \geq \lambda + 1$, we have that Theorem 3.19 implies that

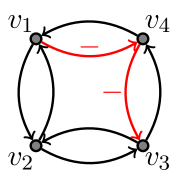
$$I_{g'+j}(d^{k'}(G, v), X_{d^{k'}(G,v)}) = \left\langle P_j^{k'}(v), \left\{P_l^{k'}(v) \cdot I_{g'+j-l}(G, X_G)|_{x_v=0}\right\}_{l=0}^{j-1}\right\rangle,$$

for all j such that $1 \leq j \leq k' - \lambda$. That is, Theorem 3.19 does not describe all the critical ideals of $d^{k'}(G, v)$. In order to obtain a description of all the critical ideals of $d^{k'}(G, v)$ it still remains to compute $d = \gamma_{\mathcal{P}}(G) - \gamma_{\mathcal{P}}(G - v)$, $d' = \gamma_{\mathcal{P}}(d(G, v)) - \gamma_{\mathcal{P}}(G)$, and $I_{g'+j}(d^k(G, v), X)$ for any $j > k' - \lambda(d, d')$. Now we present an example that may help to understood Theorem 3.19.

EXAMPLE 3.20. *Let G be the cycle with four vertices and sign σ given by*

$$\sigma(e) = \begin{cases} -1 & \textit{if } e = v_1v_4, v_4v_3, \\ 1 & \textit{otherwise.} \end{cases}$$

See Figure 11. By using a computer algebra system, we can verify that $\gamma_{\mathbb{Z}}(G) = 2$, $\gamma_{\mathbb{Z}}(G - v_1) = 2$, and $\gamma_{\mathbb{Z}}(d(G, v_1)) = 2$. Thus $d, d' = 0$ and $\lambda(d, d') = 0$. Also, it can be checked that $I_3(G, X_G) = \langle x_2 + x_4, x_1 - x_3, x_3x_4 + 2\rangle$ and $I_4(G, X_G) = \langle x_1x_2x_3x_4 + x_1x_2 + x_2x_3 - x_1x_4 - x_3x_4 - 4\rangle$.



$$L(G, X) = \begin{bmatrix} x_1 & -1 & 0 & 1 \\ -1 & x_2 & -1 & 0 \\ 0 & -1 & x_3 & -1 \\ -1 & 0 & 1 & x_4 \end{bmatrix}$$

FIGURE 11. A signed multidigraph G with four vertices and its generalized Laplacian matrix.

Since $I_3(G, X_G)|_{x_{v_1}=0} = \langle 2, x_3, x_2 + x_4\rangle$, then Theorem 3.19 implies that

$$I_3(d^{i+1}(G, v_1), X_{d^{i+1}(G,v_1)}) = \left\langle P_1^{i+1}(v_1), I_3(G, X_G)|_{x_{v_1}=0} \right\rangle = \left\langle \{x_{v_1^l}\}_{l=0}^{i+1}, 2, x_3, x_2 + x_4 \right\rangle,$$

for all $i \geq 0$. Also, since $I_4(G, X_G)|_{x_{v_1}=0} = \langle x_2x_3 - x_3x_4 - 4\rangle$, then by Theorem 3.19

$$\begin{aligned} I_4(d^{i+2}(G, v_1), X_{d^{i+2}(G,v_1)}) &= \left\langle P_2^{i+2}(v_1), P_1^{i+2}(v_1) \cdot I_3(G, X_G)|_{x_{v_1}=0}, I_4(G, X_G)|_{x_{v_1}=0} \right\rangle \\ &= \langle \{x_{v_1^l} x_{v_1^{l'}}\}_{0 \leq l < l' \leq i+2}, \{2x_{v_1^l}\}_{l=0}^{i+2}, \{x_{v_1^l} x_3\}_{l=0}^{i+2}, \{x_{v_1^l}(x_2+x_4)\}_{l=0}^{i+2}, \\ &\quad x_2x_3 - x_3x_4 - 4\rangle \end{aligned}$$

for all $i \geq 0$. Finally, since $I_j(G, X_G) = \langle 0\rangle$ for all $j \geq 5$, then

$$\begin{aligned} I_{k+2}(d^{k+i}(G, v_1), X_{d^{k+i}(G,v_1)}) &= \langle P_k^{k+i}(v_1), P_{k-1}^{k+i}(v_1) \cdot I_3(G, X_G)|_{x_{v_1}=0}, P_{k-2}^{k+i}(v_1) \cdot I_4(G, X_G)|_{x_{v_1}=0}\rangle \\ &= \langle P_k^{k+i}(v_1), \{2, x_3, x_2+x_4\} \cdot P_{k-1}^{k+i}(v_1), (x_2x_3 - x_3x_4 - 4) \cdot P_{k-2}^{k+i}(v_1)\rangle \end{aligned}$$

for all $i \geq 1, k \geq 1$. Moreover, the reader can check that

$$\begin{aligned} I_4(d(G, v), X_{d(G,v)}) &= \langle x_{v_1^0}(x_2 + x_4), x_{v_1^1}(x_2 + x_4), x_{v_1^0}(x_3x_4 + 2), x_2x_3 - x_3x_4 - 4, \\ &\quad x_{v_1^0} x_{v_1^1} x_4 + 2x_{v_1^0} + 2x_{v_1^1}, x_{v_1^0} x_3 + x_{v_1^1} x_3 - x_{v_1^0} x_{v_1^1}\rangle \\ &\neq \langle P_2^{i+1}(v), P_1^{i+1}(v) \cdot I_3(G, X_G)|_{x_{v_1}=0}, I_4(G, X_G)|_{x_{v_1}=0}\rangle. \end{aligned}$$

That is, Theorem 3.19 can not be improved.

Now we will give the description of the critical ideals of $r^k(G, v)$. This part is structured similarly than the part of the critical ideals of $d^k(G, v)$. Given a subset S of the natural numbers and a vertex $v \in V(G)$, let $\widetilde{P}_l^S(v) = \{\prod_{c \in C} x_{v^c} + 1 \,:\, C \in \binom{S}{l}\}$. We take $P_0^S(v) = \{1\}$. Also, for simplicity $\widetilde{P}_l^{\{0\} \cup [k]}(v)$ will be denoted by $\widetilde{P}_l^k(v)$. We will use similar arguments to those used in the proof of Lemma 3.16.

Lemma 3.21. *Let G be a signed multidigraph with n vertices, and $v \in V(G)$. If $k, j \geq 1$ and $m = \min(j-1, k)$, then*

$$\begin{aligned} I_j(r^k(G,v), X_{r^k(G,v)}) \; &= \; \Big\langle \{\widetilde{P}^k_l(v) \cdot I_{j-l}(G, X_G)|_{x_v=-1}\}_{l=0}^{m-1}, \widetilde{P}^k_m(v) \cdot I_{j-m}(G{-}v, X), \\ &\quad \widetilde{P}^k_m(v) \cdot \mathrm{minors}_{j-m}(\mathbf{a}, L(G{-}v, X)), \widetilde{P}^k_m(v) \cdot \mathrm{minors}_{j-m}(L(G{-}v, X), \mathbf{b}), \widetilde{S}^k_j \Big\rangle, \end{aligned}$$

where $\widetilde{S}^k_j$ is equal to

$$\left\{ \prod_{s=1}^{j} (x_{v^{l_s}}+1) - \sum_{s=1}^{j} \prod_{t \neq s} (x_{v^{l_t}}+1) : 0 \leq l_1 < \cdots < l_j \leq k \right\},$$

when $j \leq k+1$, and equal to

$$\left\{ \det(Q) \cdot \prod_{t=0}^{k} (x_{v^t}+1) + \det(J(-1, \mathbf{a}'; Q, \mathbf{b}')) \sum_{t=0}^{k} \prod_{s \neq t} (x_{v^s}+1) : J(x_v, \mathbf{a}'; Q, \mathbf{b}') \in M_{j-k}(L(G, X_G)) \right\},$$

when $j > k+1$.

Proof. Let $\mathcal{I}, \mathcal{I}' \subseteq [n+k]$be two sets of size j, $h = |\mathcal{I} \cap [k+1]|$, and $h' = |\mathcal{I}' \cap [k+1]|$. Clearly $0 \leq h, h' \leq m+1$. Suppose $L(r^k(G,v), X_{r^k(G,v)}) = J(L(K_{k+1}, X_{K_{k+1}}), \mathbf{a}; L(G-v, X_{G-v}), \mathbf{b})$ for some $\mathbf{a}, \mathbf{b} \in \mathcal{P}^{n-1}$. Let $m_{\mathcal{I},\mathcal{I}'} = \det(L(r^k(G,v), X)[\mathcal{I}, \mathcal{I}'])$.

We can use the same arguments used in the proof of Lemma 3.16 for the case when $h = 0$ or $h' = 0$. On the other hand, by Lemma 3.3

$$m_{\mathcal{I},\mathcal{I}'} = \begin{cases} 0 & \text{if } |h-h'| > 2, \\ \det \begin{bmatrix} P & \mathbf{1}^T \end{bmatrix} \det \begin{bmatrix} \mathbf{b}' \\ Q \end{bmatrix} & \text{if } h-h' = 1, \\ \det \begin{bmatrix} P \\ \mathbf{1} \end{bmatrix} \det \begin{bmatrix} \mathbf{a}'^T & Q \end{bmatrix} & \text{if } h'-h = 1, \\ \det(P)\det(Q) - \det \begin{bmatrix} P & \mathbf{1}^T \\ \mathbf{1} & 0 \end{bmatrix} \det \begin{bmatrix} 0 & \mathbf{b}' \\ \mathbf{a}'^T & Q \end{bmatrix} & \text{if } h = h', \end{cases}$$

where P is a submatrix of $L(K_{k+1}, X_{K_{k+1}})$, Q is a submatrix of $L(G-v, X_{G-v})$, $\mathbf{a}'$ is a subvector of $\mathbf{a}$ and $\mathbf{b}'$ is a subvector of $\mathbf{b}$. If $h - h' = 1$ then $\det \begin{bmatrix} P & 1 \end{bmatrix} \neq 0$ if and only if (up to row and column permutations)

$$P = \begin{pmatrix} x_{v^{i_1}} & & -1 & -1 \\ & \ddots & & \\ -1 & & x_{v^{i_{h'}}} & -1 \end{pmatrix}^T$$

for some $0 \leq l_1 < \cdots < l_{h'} \leq k$. Since $\det \begin{bmatrix} P & \mathbf{1}^T \end{bmatrix} = \prod_{s=1}^{h'} (x_{v^{i_s}} + 1)$, then $m_{\mathcal{I},\mathcal{I}'} \in \widetilde{P}^k_{h'}(v) \cdot \mathrm{minors}_{j-h'}(L(G-v, X_{G-v}), \mathbf{b}) \subsetneq \widetilde{P}^k_{h'}(v) \cdot I_{j-h'}(G, X_G)|_{x_v=-1}$. In a similar way, if $h' - h = 1$, then $m_{\mathcal{I},\mathcal{I}'} \in \widetilde{P}^k_h(v) \cdot I_{j-h}(G, X_G)|_{x_v=-1}$.

Now assume that $h = h'$. It is not difficult to see that if P has two rows equal to $-\mathbf{1}$, then $m_{\mathcal{I},\mathcal{I}'} = 0$. Let

$$R = \begin{pmatrix} x_{v^{l_1}} & & -1 \\ & \ddots & \\ -1 & & x_{v^{l_h}} \end{pmatrix}$$

where $0 \leq l_1 < \cdots < l_h \leq k$. If P has only a row equal to $-\mathbf{1}$, then P is equal to (up to row and column permutations) $R|_{x_{v^{l_h}}=-1}$. Since $\det(R|_{x_{v^{l_h}}=-1}) = -\prod_{s=1}^{h-1}(x_{v^{l_s}}+1)$ and $\det(J(R|_{x_{v^{l_h}}=-1}, \mathbf{1}; 0, \mathbf{1})) = -\prod_{s=1}^{h-1}(x_{v^{l_s}} + 1)$, then

$$m_{\mathcal{I},\mathcal{I}'} = (\det(J(0, \mathbf{a}'; Q, \mathbf{b}')) - \det(Q)) \prod_{s=1}^{h-1}(x_{v^{l_s}} + 1) = \det(J(-1, \mathbf{a}'; Q, \mathbf{b}')) \prod_{s=1}^{h-1}(x_{v^{l_s}} + 1),$$

for all $1 \leq h \leq m$. Thus $m_{\mathcal{I},\mathcal{I}'} \in \langle \widetilde{P}^k_{h-1}(v) \cdot I_{j-h+1}(G, X_G)|_{x_v=-1}\rangle$. Finally, if P has no row equal to $-\mathbf{1}$, then P is equal to (up to row and column permutations) to R. Since $\det(R) = \prod_{s=1}^{h}(x_{v^{l_s}} + 1) - \sum_{s=1}^{h}\prod_{t\neq s}(x_{v^{l_t}} + 1)$ (see [**22**, theorem 3.15]) and $\det(J(R, \mathbf{1}; 0, \mathbf{1})) = -\sum_{s=1}^{h}\prod_{t\neq s}(x_{v^{l_t}} + 1)$, then

$$\begin{aligned} m_{\mathcal{I},\mathcal{I}'} &= \det(Q) \cdot \prod_{s=1}^{h}(x_{v^{l_s}} + 1) + (\det(J(0, \mathbf{a}'; Q, \mathbf{b}')) - \det(Q)) \cdot \sum_{s=1}^{h}\prod_{t\neq s}(x_{v^{l_t}} + 1) \\ &= \det(Q) \cdot \prod_{s=1}^{h}(x_{v^{l_s}} + 1) + \det(J(-1, \mathbf{a}'; Q, \mathbf{b}')) \cdot \sum_{s=1}^{h}\prod_{t\neq s}(x_{v^{l_t}} + 1), \text{ for all } 1 \leq h \leq m. \end{aligned}$$

Since $\det(Q) \cdot \prod_{s=1}^{h}(x_{v^{l_s}} + 1) = m_{\mathcal{I},\mathcal{I}'} - \det(J(-1, \mathbf{a}'; Q, \mathbf{b}')) \cdot \sum_{s=1}^{h}\left(\prod_{t\neq s}(x_{v^{l_t}} + 1)\right) \in \widetilde{P}^k_h(v) \cdot I_{j-h}(G, X_G)|_{x_v=-1}$ we get the result. $\square$

Now we present a result similar to Theorem 3.19 for the replication of vertices.

THEOREM 3.22. *Let G be a signed multidigraph, $v \in V(G)$, $g = \gamma_{\mathcal{P}}(G)$, $g' = \gamma_{\mathcal{P}}(r(G, v))$, $d = g - \gamma_{\mathcal{P}}(G - v)$, $d' = g' - g$, and $\lambda = \lambda(d, d')$. If $g \geq 1$, then $0 \leq d + d' \leq 2$ and*

$$I_{g'+k}(r^{k+\lambda+i}(G, v), X_{r^{k+\lambda+i}(G,v)}) = \left\langle \widetilde{P}^{k+\lambda+i}_k(v), \left\{\widetilde{P}^{k+\lambda+i}_l(v) \cdot I_{g'+k-l}(G, X)|_{x_v=-1}\right\}_{l=0}^{k-1} \right\rangle,$$

for all $k \geq 1$ and $i \geq 0$.

The proof follows similar arguments of those used in Theorem 3.19.

PROOF. First since $I_j(G, X_G)|_{x_v=-1} \subseteq I_{j-2}(G - v, X_{G-v})$, then Lemma 3.6 that that $0 \leq d + d' \leq 2$. Now since $m = \min(g' + k - 1, k + \lambda + i) = k + \lambda$, then by Lemma 3.21 we have that $I_{g'+k}(r^{k+\lambda+i}(G, v), X_{r^{k+\lambda+i}(G,v)})$ is equal to

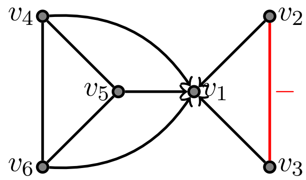
$$\left\langle \left\{\widetilde{P}^{k+\lambda+i}_l(v) \cdot I_{g'+k-l}(G, X_G)|_{x_v=-1}\right\}_{l=0}^{k+\lambda-1}, \widetilde{P}^{k+\lambda+i}_{k+\lambda}(v) \cdot I_{g'-\lambda}(G - v, X_{G-v}),\right.$$

$$\left.\widetilde{P}^{k+\lambda+i}_{k+\lambda}(v) \cdot \text{minors}_{g'-\lambda}(\mathbf{a}, L(G - v, X_{G-v})), \widetilde{P}^{k+\lambda+i}_{k+\lambda}(v) \cdot \text{minors}_{g'-\lambda}(L(G - v, X_{G-v}), \mathbf{b}), \widetilde{S}^{k+\lambda+i}_{g'+k}(G, v)\right\rangle.$$

If $\lambda(d, d') = 0$, then we use the same argument given in Theorem 3.19. Now if $\lambda(d, d') = 1$, then either $I_{g'}(G, X_G)$ is trivial or $I_{g'}(r(G, v), X_{r(G,v)})$ is trivial. In both cases we have that $I_{g'}(G, X_G)|_{x_v=-1}$, and the result turns out in a similar way than in Theorem 3.19. $\square$

Now we how an example in order to understand Theorem 3.22.

EXAMPLE 3.23. *Let G be the signed multidigraph given in Figure 12.*



$$L(G,X)=\begin{bmatrix} x_1 & 0 & 0 & 0 & 0 & 0\\ -1 & x_2 & 1 & 0 & 0 & 0\\ -1 & 1 & x_3 & 0 & 0 & 0\\ -1 & 0 & 0 & x_4 & -1 & -1\\ -1 & 0 & 0 & -1 & x_5 & -1\\ -1 & 0 & 0 & -1 & -1 & x_6 \end{bmatrix}$$

FIGURE 12. A graph G with six vertices and its generalized Laplacian matrix.

By using a computer algebra system, we can check that $\gamma_{\mathbb{Z}}(G)=\gamma_{\mathbb{Z}}(G-v_1)=2$ *and* $\gamma_{\mathbb{Z}}(r(G,v_1))=3$. *Thus* $d'=1$ *and* $\lambda(d,d')=1$. *Also, it is not difficult to check that* $I_4(G,X_G)|_{x_1=-1}=\langle x_4+1,x_5+1,x_6+1,x_2x_3-1\rangle$,
$I_5(G,X_G)|_{x_1=-1}=\langle (x_4+1)\cdot(x_2x_3-1),(x_5+1)\cdot(x_2x_3-1),(x_6+1)\cdot(x_2x_3-1),x_4x_5x_6-x_4-x_5-x_6-2\rangle$, *and* $I_6(G,X_G)|_{x_1=-1}=\langle (x_2x_3-1)\cdot(x_4x_5x_6-x_4-x_5-x_6-2)\rangle$. *Then Theorem 3.22 implies*

$$\begin{aligned} I_4(r^{i+2}(G,v_1),X_{r^{i+2}(G,v_1)}) &= \langle \{x_{v_1^l}+1\}_{0\le l\le i+2}, I_4(G,X_G)|_{x_1=-1}\rangle\\ &= \langle \{x_{v_1^l}+1\}_{0\le l\le i+2}, x_4+1,x_5+1,x_6+1,x_2x_3-1\rangle, \end{aligned}$$

for all $i\ge 0$. *Also* $I_5(r^{i+3}(G,v_1),X_{r^{i+3}(G,v_1)})$ *is equal to*

$$\left\langle \left\{(x_{v_1^l}+1)(x_{v_1^{l'}}+1)\right\}_{0\le l<l'\le i+3}, I_5(G,X_G)|_{x_1=-1}, \left\{(x_{v_1^l}+1)\cdot I_4(G,X_G)|_{x_1=-1}\right\}_{0\le l\le i+3}\right\rangle$$

for all $i\ge 0$. *Finally,* $I_{k+3}(r^{k+i+1}(G,v_1),X_{r^{k+i+1}(G,v_1)})$ *is equal to*

$$\left\langle \widetilde{P}_k^{k+i+1}(v_1), \widetilde{P}_{k-1}^{k+i+1}(v_1)\cdot I_4(G,X)|_{x_v=-1}, \widetilde{P}_{k-2}^{k+i+1}(v_1)\cdot I_5(G,X)|_{x_v=-1}, \widetilde{P}_{k-3}^{k+i+1}(v_1)\cdot I_6(G,X)|_{x_v=-1}\right\rangle,$$

for all $k\ge 3$ *and* $i\ge 0$. *On the other hand, it is not difficult to check that* $I_4(r(G,v_1),X_{r(G,v_1)})$ *is equal to*

$$\langle \{(x_{v_1^l}+1)(x_{l'}-1)\}_{0\le l\le 1, 2\le l'\le 3}, x_4+1,x_5+1,x_6+1,x_2x_3-1,x_{v_1^0}x_{v_1^1}-1\rangle,$$

which is different from $\langle x_{v_1^0}+1,x_{v_1^1}+1,x_4+1,x_5+1,x_6+1,x_2x_3-1\rangle$. *Thus Theorem 3.22 can not be improved.*

REMARK 3.24. *Note that* $I_i(K_{k+1},X_{K_{k+1}})=\langle \widetilde{P}_i^k(v)\rangle$. *Moreover, if* $m=\min(k,j-1)$, *then*

$$I_j(r^k(g,v),X_{r^k(g,v)})|_{x_v0=-1}=\left\langle \left\{\widetilde{P}_l^{[k]}(v)\cdot I_{j-l}(G,X_G)|_{x_v=-1}\right\}_{l=0}^{m}\right\rangle.$$

That is, $I_j(d^k(G,v),X)|_{x_v0=-1}$ *behaves almost equally as* $I_j(K_{k+1}+G,X_{K_{k+1}+G})$.

CHAPTER 4

Critical ideals of small graphs

The computation of the invariant factors of the Laplacian matrix is an important technique used in the understanding of $K(G)$. For instance, several researchers addressed the question of how often the critical group is cyclic, that is, if $f_1(G)$ denote the number of invariant factors equal to 1, then the question is how often $f_1(G)$ is equal to $n-2$ or $n-1$. In this way, it is desirable to understand the combinatorial properties of $f_1(G)$ and the family of graphs $\mathcal{G}_i$ of simple connected graphs with $f_1(G) = i$.

Superficially, the critical group has three components: algebraic, combinatorial, and arithmetic. The methodology of these studies rely on the separation of the combinatorial and algebraic information from most of the arithmetic component by means of the introduction of a new invariant: the critical ideals. Critical ideals were defined in [**22**] as a generalization of the critical group and have been studied in [**1, 7, 22**]. The effect of avoiding the arithmetic information is that the behavior of the critical ideals is easier to observe and to describe. Thus critical ideals provide a new perspective to understand the critical group theory.

The difficulty in critical ideals relies in that these are parameters hard to compute. However, it is possible to compute the algebraic co-rank of all simple connected graphs with at most 9 vertices using the software Macaulay2 and Nauty, see Table 1. The computation of the algebraic co-rank of the connected graphs with at most 8 vertices required at most 3 hours on a MacBookPro with a 2.8 GHz Intel i7 quad core processor and 16 GB RAM. Besides, the computation of the algebraic co-rank of the connected graphs with 9 vertices required 4 weeks of computation on the same computer. Under these conditions the time required to compute the critical ideals of the graphs with 10 vertices is at least 3.7 years.

$n\backslash\gamma$	1	2	3	4	5	6	7	8
2	1							
3	1	1						
4	1	4	1					
5	1	8	11	1				
6	1	13	52	45	1			
7	1	18	141	505	187	1		
8	1	24	315	2749	7086	941	1	
9	1	31	605	10085	93296	152365	4696	1

(a) The number of simple connected graphs with n vertices and γ trivial critical ideals.

$n\backslash f_1$	1	2	3	4	5	6	7	8
2	1							
3	1	1						
4	1	3	2					
5	1	6	11	3				
6	1	8	34	63	6			
7	1	8	53	271	509	11		
8	1	11	97	707	3226	7052	23	
9	1	15	139	1646	12822	68979	177431	47

(b) The number of simple connected graphs with n vertices and f_1 invariant factors equal to 1.

TABLE 1.

This section aims at discussing some numerical experiments using this data, which had led to numerous conjectures and results.

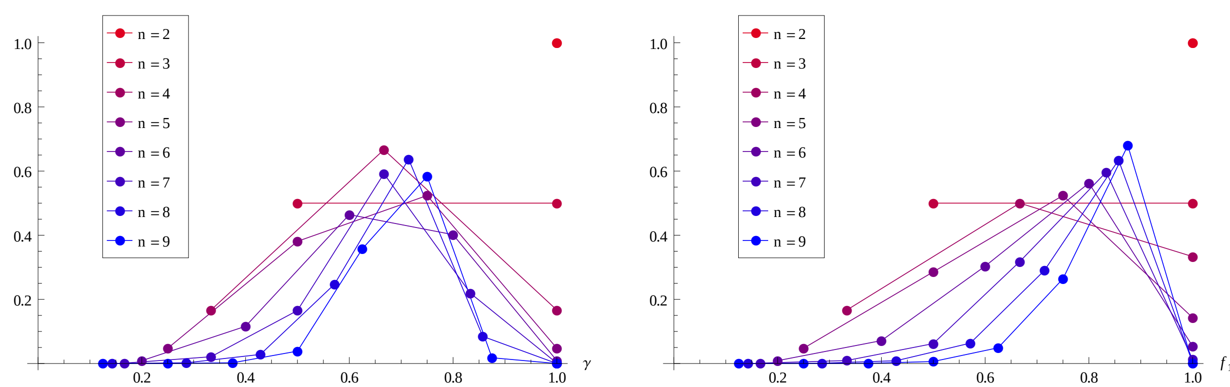
1. Numerical experiments

It follows from Kirchoff's matrix-tree theorem that the order of $K(G)$ is the number $\kappa(G)$ of spanning trees of the graph G. This is the reason the only graphs with n vertices and $n-1$ invariant factors are the trees. In the case of critical ideals, we can observe in the data that the only graph with n vertices and $n-1$ trivial critical ideals is the path ([**22**, Conjecture 4.12]). But this was completely proved until [**23**]. Contrary to the critical group, the data suggests (see Fig. 15.a) that the number of graphs with n vertices and $n-2$ trivial critical ideals is small, and there might be chances of completely classifying these graphs.

A major investigation on the behavior of the critical group is: how often the critical group is cyclic? In [**35**] and [**44**] was found that the numerical data suggests we could expect to find a substantial proportion of graphs have cyclic critical group. This behavior can be also observed in Fig. 15.b. Based on this, D. G. Wagner conjectured that almost every connected simple graph has cyclic critical group. However, a deeper study was done in [**45**] concluding that the probability that the critical group of a random graph is cyclic is asymptotically at most

$$\zeta(3)^{-1}\zeta(5)^{-1}\zeta(7)^{-1}\zeta(9)^{-1}\zeta(11)^{-1}\cdots \approx 0.7935212,$$

where ζ is the Riemann zeta function, differing from Wagner's conjecture. Besides, it is interesting [**19**] that for any given connected simple graph, there is a homeomorphic graph with cyclic critical group. See [**20, 35, 45**] for more questions and results on this topic.



(a) Scaled number of connected graphs with n vertices and γ trivial critical ideals.

(b) Scaled number of connected graphs with n vertices and f_1 invariant factors equal to 1.

FIGURE 13.

1.1. γ-critical graphs. One of the first results in the study of the critical group was [**21, 34, 38**] that $\mathcal{G}_1$ consists only of the complete graphs. In this sense, several people [**37**] posed interest on the characterization of $\mathcal{G}_2$ and $\mathcal{G}_3$. These characterizations are quite hard. In [**40**] it was characterized the graphs in $\mathcal{G}_2$ whose third invariant factor is equal to n, $n-1$, $n-2$, or $n-3$. In [**17**] the characterizations of the graphs in $\mathcal{G}_2$ with a cut vertex. Recently, a complete characterization of $\mathcal{G}_2$ was given in [**2**], and a partial description of $\mathcal{G}_3$ was given in [**3**]. There are also related results of interest to algebraic geometers. In [**17, 34**] some graph families for which the equality $f_1(G) = \beta(G)$ holds are characterized.

The major advances in this matter were provided by the properties critical ideals; since the set $\Gamma_{\leq i}$ is closed under induced subgraphs. This property allow us to define the following concepts. A graph G is *forbidden* for $\Gamma_{\leq k}$ when $\gamma(G) \geq k+1$. Let $\mathbf{Forb}(\Gamma_{\leq k})$ be the set of minimal (under induced subgraphs property) forbidden graphs for $\Gamma_{\leq k}$. Given a family $\mathcal{F}$ of graphs, a graph G is called $\mathcal{F}$-free if no induced subgraph of G is isomorphic to a member of $\mathcal{F}$. Thus $G \in \Gamma_{\leq k}$ if and only if G is $\mathbf{Forb}(\Gamma_{\leq k})$-free. Hence characterizing $\mathbf{Forb}(\Gamma_{\leq k})$ leads to a characterization of $\Gamma_{\leq k}$. And after an analysis of the *k-th* invariant factor of the Laplacian matrix of the graphs in $\Gamma_{\leq k}$, the characterization of $\mathcal{G}_k$ can be obtained. For example, in [**2**] it was proved that $\mathbf{Forb}(\Gamma_{\leq 0}) = \{P_2\}$, $\mathbf{Forb}(\Gamma_{\leq 1}) = \{P_3\}$, and $\mathbf{Forb}(\Gamma_{\leq 2}) = \{P_4, K_5 \setminus S_2, K_6 \setminus M_2, \mathsf{cricket}, \mathsf{dart}\}$. And the characterization of $\Gamma_{\leq 2}$ and $\mathcal{G}_2$ turns out.

An alternative technique of computing the elements of $\mathbf{Forb}(\Gamma_{\leq k})$ is by means of the following definition. A graph G is called *γ-critical* if $\gamma(G - v) < \gamma(G)$ for all $v \in V(G)$. Then $\mathbf{Forb}(\Gamma_{\leq k})$ is equal to the set of γ-critical graphs with $\gamma(G) \geq k + 1$ and $\gamma(G - v) \leq k$ for all $v \in V(G)$. In Table 2 it is shown the number γ-critical graphs with until 9 vertices.

$k\backslash n$	2	3	4	5	6	7	8	9
0	1							
1		1						
2			1	3	1			
3				1	27	17	4	
4					1	153	773	340
5						1	871	52333
6							1	4562
7								1

TABLE 2. The number of simple graphs with n vertices in $\mathbf{Forb}(\Gamma_{\leq k})$.

We have that P_k is the unique graph with k vertices in $\mathbf{Forb}(\Gamma_{\leq k-2})$. Therefore, $\mathbf{Forb}(\Gamma_{\leq k})$ is not empty for all $k \geq 0$. However, it is still an open question to prove that $\mathbf{Forb}(\Gamma_{\leq k})$ is finite for all $k \geq 3$. For instance, there are 49 graphs in $\mathbf{Forb}(\Gamma_{\leq 3})$ with at most 8 vertices. And since there exists no minimal forbidden graph with 9 vertices for $\Gamma_{\leq 3}$, then it is likely that there exists no graph in $\mathbf{Forb}(\Gamma_{\leq 3})$ with more than 8 vertices. A question to ask is: Is there an asymptotic behavior on the proportion of γ-critical graphs with n vertices? Is it expected to remain low?

n	2	3	4	5	6	7	8	9
% γ-critical	100	20	16.66	19.04	25.89	20.04	14.83	21.92

TABLE 3. Percentage of graphs with n vertices which are γ-critical.

1.2. Components of the critical group. The standard technique for obtaining the invariant factors of the critical group is to reduce the Laplacian matrix to its Smith normal form. Algorithmically, this is done by applying row and column operations to obtain a diagonal matrix, whose entries are the invariant factors. However the process of computing $f_1(G)$ hides several relations with the combinatorial structure of the graph. We identify three components for which $\Delta_k(G) = 1$.

- *Algebraic.* When $I_k(G, X_G) = \langle 1 \rangle$ and there is no induced subgraph of order k whose associated minor equals 1.
- *Combinatorial.* When there exists an induced subgraph of order k whose corresponding minor is equal to 1.
- *Arithmetic.* Otherwise.

Table 4 shows the number of occurrences for each type of component.

Next example shows the only graph with 7 vertices in which the algebraic component appears.

	algebraic	combinatorial	arithmetic
2	0	1	0
3	0	3	0
4	0	12	1
5	0	54	4
6	0	368	33
7	1	3420	450
8	28	53045	8672
9	10367	1445401	271641

TABLE 4.

EXAMPLE 4.1. *Let G be the graph in Fig. 14. It is not difficult to check that the algebraic co-rank of G is 5 and $L(G,X)$ has no 5-minor equal to 1.*

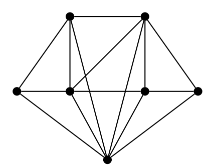


FIGURE 14. A graph G with seven vertices.

1.3. Characteristic ideals. In [**35**] D. Lorenzini showed a deep relation between the critical group and the Laplacian spectrum of a graph. For instance, Lorenzini ([**35**, proposition 3.2]) proved that if λ is an integer eigenvalue of $L(G)$ of multiplicity $\mu(\lambda)$, then $K(G)$ contains a subgroup isomorphic to $\mathbb{Z}_\lambda^{\mu(\lambda)-1}$. Another point of interest on the critical group discussed in [**14**] and [**8**, Section 13.8] is that its structure can be used to distinguish graphs in cases where other algebraic invariants, such as those derived from the spectrum, fail. In this sense critical ideals are even finer as next result shows. Let σ be a permutation on $V(G)$. Then σG is a graph on $V(G)$ such that $\{i,j\} \in E(G)$ if and only if $\{\sigma(i),\sigma(j)\} \in \sigma G$. Two graphs G and G' on the same vertex set V are called *n-cospectral* if there exists a permutation σ on V such that $I_n(G,X_G) = I_n(\sigma G', X_{\sigma G'})$.

PROPOSITION 4.2. [**27**, Propisition 1] *Let G and G' be two graphs with n vertices. Then G and G' are isomorphic if and only if they are n-cospectral.*

Let $I_i(G,t)$ denote the critical ideals where $x_i = t$ for all $1 \leq i \leq n$. Then $I_n(G,t)$ is equal to the ideal generated by the characteristic polynomial $p_G(t)$ of the adjacency matrix of G. In a similar way, if we take $x_i = d_G(i)-t$ for all $1 \leq i \leq n$, then we recover the characteristic polynomial of the Laplacian matrix of G from their critical ideals. Therefore, the ideals $I_i(G,t) \subseteq \mathcal{R}[t]$ are called *characteristic ideals*, in keeping with the terminology used in spectral graph theory.

Note that characteristic ideals depend on the base ring. We say that two graphs G and G' are *$\gamma_{\mathcal{R}}$-cospectral* if they have the same characteristic ideals on $\mathcal{R}[t]$. The term "set of $\gamma_{\mathcal{R}}$-cospectral non-ismorphic graphs" is denoted by $\gamma_{\mathcal{R}}$-SET.

EXAMPLE 4.3. *Let G_1 and G_2 be the graphs shown in Fig. 15. These graphs are cospectral, moreover they are the unique pair of $\gamma_{\mathbb{Q}}$-cospectral graphs with 6 vertices.*

$$I_i(G_1,t) = I_i(G_2,t) = \begin{cases} \langle 1 \rangle & \text{if } 1 \leq i \leq 4, \\ \langle t+1 \rangle & \text{if } i = 5, \\ \langle (t-1)\cdot(t+1)^2\cdot(t^3-t^2-5t+1) \rangle & \text{if } i = 6, \end{cases}$$

G_1 G_2

FIGURE 15.

But when the base ring is $\mathbb{Z}$*, we have that the characteristic ideals are:*

$$I_i(G_1,t) = \begin{cases} \langle 1 \rangle & \text{if } 1 \le i \le 4, \\ \langle 2(t+1), (t+1)\cdot(t^2+1) \rangle & \text{if } i = 5, \\ \langle (t-1)\cdot(t+1)^2\cdot(t^3-t^2-5t+1) \rangle & \text{if } i = 6, \end{cases}$$

and

$$I_i(G_2,t) = \begin{cases} \langle 1 \rangle & \text{if } 1 \le i \le 3, \\ \langle 2, (t+1) \rangle & \text{if } i = 4, \\ \langle 4(t+1), (t+1)\cdot(t-3) \rangle & \text{if } i = 5, \\ \langle (t-1)\cdot(t+1)^2\cdot(t^3-t^2-5t+1) \rangle & \text{if } i = 6. \end{cases}$$

In Table 5 it is shown the number of $\gamma_{\mathcal{R}}$-SETs with n vertices.

	6	7	8
$\mathbb{Q}$	1	31	660
$\mathbb{Z}$	0	3	232

TABLE 5. The number of γ-SETs with n vertices.

Bibliography

[1] C.A. Alfaro and C.E. Valencia, On the sandpile group of the cone of a graph, Linear Algebra Appl. 436 (2012) 1154–1176.
[2] C.A. Alfaro and C.E. Valencia, Graphs with two trivial critical ideals, Discrete Appl. Math.167 (2014) 33–44
[3] C.A. Alfaro and C.E. Valencia, Small clique number graphs with three trivial critical ideals, submitted.
[4] C.A. Alfaro and C.E. Valencia, Critical ideals of small graphs, in progress.
[5] C.A. Alfaro and C.E. Valencia, Digraphs with one trivial critical ideal, in progress.
[6] C.A. Alfaro and C.E. Valencia, Sandpile group and integer programming, in progress.
[7] C.A. Alfaro, H. Corrales and C.E. Valencia, Critical ideals of signed graphs with twin vertices, manuscript.
[8] A.E. Brouwer and W.H. Haemers, Spectra of Graphs, Universitext, Springer, 2012.
[9] R. Bacher, P. de la Harpe and T. Nagnibeda, The lattice of integral flows and the lattice of integral cuts on a finite graph, Bull. Soc. Math. France 125 (1997) 167–198.
[10] R.B. Bapat, Graphs and Matrices, Universitext, Springer, 2010.
[11] L.W. Beinke and R.J. Wilson, Eds., Topics in algebraic graph theory, Cambridge University Press, 2005.
[12] D. Bertsimas and J.N. Tsitsiklis, Introduction to Linear Optimization, Athena Scientific, 1997.
[13] N. Biggs, Algebraic Graph Theory, Second edition, Cambridge University Press, 1993.
[14] N. Biggs, Chip-Firing and Critical Group of a Graph, J. Algebraic Comb. 9 (1999) 25–45.
[15] A. Brandstadt, V.B. Le and J. Spinrad, Graph Classes: A Survey, SIAM Monographs on Discrete Math. Appl., Vol. 3, SIAM, 1999.
[16] S. Chaiken, A combinatorial proof of the all minors matrix tree theorem, SIAM J. Alg. Disc. Meth. 3 (1982) 319–329.
[17] W.H. Chan, Y. Hou and W.C. Shiu, Graphs whose critical groups have larger rank, Acta Math. Sinica 27 (2011) 1663–1670.
[18] W. Chen and T. Schedler, Concrete and abstract structure of the sandpile group for thick trees with loops, preprint, ArXiv:math/0701381v1.
[19] S. Chen and S.K. Ye, Critical groups for homeomorphic classes of graphs, Discrete Math. 309 (2008) 255–258.
[20] J. Clancy, N. Kaplan, T. Leake, S. Payne and M.M. Wood, On a Cohen-Lenstra heuristic for jacobians of random graphs, arxiv.
[21] R. Cori and D. Rossin, On the sandpile group of dual graphs, European J. Combin. 21 (4) (2000) 447–459.
[22] H. Corrales and C.E. Valencia, On the critical ideals of graphs, Linear Algebra Appl. 439 (2013) 3870–3892.
[23] H. Corrales, C.E. Valencia, On the critical ideals of trees, in preparation.
[24] D. Dhar, Self-organized critical state of the sandpile automathon models, Phys. Rev. Lett. 64 (14) (1990) 1613–1616.
[25] R. Diestel, Graph Theory, Fourth Edition, Springer, 2010.
[26] C. Godsil and G. Royle, Algebraic Graph Theory, GTM 207, Springer-Verlag, New York, 2001.
[27] N. Ghareghani, F. Ramezani and B. Tayfeh-Rezaie, Graphs cospectral with starlike trees, Linear Algebra Appl. 429 (2008) 2691–2701.
[28] D.R. Grayson and M.E. Stillman, *Macaulay2, a software system for research in algebraic geometry*, available at www.math.uiuc.edu/Macaulay2/.
[29] R. Grone, R. Merris and W. Watkins, Laplacian unimodular equivalence of graphs. In: R. Brualdi, S. Friedland and V. Klee (Eds.) Combinatorial and Graph-Theoretical Problems in Linear Algebra, Springer-Verlag (1993) 175-180.
[30] N. Jacobson, Basic Algebra I, Second Edition, W. H. Freeman and Company, New York, 1985.
[31] B. Jacobson, A. Niedermaier and V. Reiner, Critical groups for complete multipartite graphs and Cartesian products of complete graphs, J. Graph Theory 44 (2003) 231–250.

[32] G. Kirchhoff, Über die Auflösung der Gleichungen, aus welche man bei der Untersuchung der linearen Verteilung galvanisher Ströme gefürht wird, Ann. Phys. Chem. 72 (1847) 497–508.
[33] D.J. Lorenzini, Arithmetical Graphs, Math. Ann. 285 (1989) 481–501.
[34] D.J. Lorenzini, A finite group attached to the laplacian of a graph, Discrete Mathematics 91 (1991) 277–282.
[35] D.J. Lorenzini, Smith normal form and Laplacians, J. Combin. Theory B 98 (2008), 1271-1300.
[36] B.D. McKay, *nauty User's Guide (Version 2.4)*, available at http://cs.anu.edu.au/∼bdm/nauty/.
[37] C. Merino, The chip-firing game, Discrete Mathematics 302 (2005) 188–210.
[38] R. Merris, Unimodular Equivalence of Graphs, Linear Algebra Appl. 173 (1992),181-189
[39] C. Meyer, Matrix analysis and applied linear algebra, first edition, SIAM, 2000.
[40] Y. Pan and J. Wang, A note on the third invariant factor of the Laplacian matrix of a graph, preprint, arXiv:0912.3608v1 [math.CO]
[41] D. Perkinson, Lecture notes *http://people.reed.edu/ davidp/412/*
[42] M. Raynaud, Spécialisation du foncteur de Picard, Publications Mathématiques de l'IHÉS 38 (1970) 27-76.
[43] E.R. Speer, Asymmetric Abelian sandpiles models, Journal of statistical physics, 71 1-2 (1993) 61-74.
[44] D.G. Wagner, The critical group of a directed graph, preprint, arXiv:math/0010241v1 [math.CO]
[45] M.M. Wood, The distribution of the sandpile groups of random graphs, arxiv.

Part 2

Dimension Reduction in Tree Space

CHAPTER 5

Dimension Reduction in Tree Space

1. Introduction

In statistics, data sets that reside in high dimensional spaces are quite common. A widely used set of techniques to simplify and analyze such sets is *principal component analysis* (PCA). It was introduced by Pearson in 1901 and independently by Hotelling in 1933. A comprehensive introduction can be found in [**18**].

The main aim of PCA is to provide a smaller subspace such that the maximum amount of information is retained when the original data points are projected onto it. This smaller subspace is expressed through components. In many contexts, one dimensional subspaces are called lines, and we follow this terminology. The line that carries the most variation present in the data set is called *first principal component* (PC1). The *second principal component* (PC2) is the line such that when combined with PC1, the most variation that can be retained in a two-dimensional subspace is kept. One may repeat this procedure to find as many principal components as necessary to properly summarize the data set in a manageable sized subspace.

Another way to characterize the principal components is to consider the distances of the data points to a given subspace. The line which minimizes the sum of squared distances of data points onto it can be considered as PC1. Similarly, PC2 is the line that, when combined with PC1, the sum of squared distances of the data points to this combination is minimized. In Euclidean space, these two characterizations are equivalent.

An important topic within PCA is called *dimension reduction* (See [**20**] for dimension reduction and [**18**] pp. 144, for *backward elimination method*). The aim of dimension reduction method is to find the components such that when they are eliminated, the projection of the data onto the remaining subspace will retain the maximum amount of variation. Or alternatively, the remaining subspace will have the minimum sum of squared distances to the data points. These are the components with least influence.

We would like to note that, in the general sense, any PCA method can be regarded as a dimension reduction process. However, [**20**] reserves the term dimension reduction specifically for this method, which some other resources also refer as backward elimination, or backward PCA. We will follow [**20**]'s convention, together with "backward PCA" terminology. The original approach will be called *forward PCA*.

In Euclidean space, the choice of which technique to use depends on the needs of the end user: If only the few principal components with most variation in them are needed, then the forward approach is more suitable. If the aim is to eliminate only the few least useful components, then the backward approach would be the appropriate choice.

The historically most common space used in statistics is the Euclidean space ($\mathbb{R}^n$) and the PCA ideas were first developed in this context. In $\mathbb{R}^n$, the two definitions of PC's (maximum variation and minimum distance) are equivalent, and the components are all orthogonal to each other. In

Euclidean space, applying forward or backward PCA n times for a data set in $\mathbb{R}^n$ would provide an orthogonal basis for the whole space.

Moreover, in this context, the set of components obtained with the backward approach is the same as the one obtained by the classical forward approach, only the order of the components is reversed. This is a direct result of orthogonality properties in Euclidean space. This phenomenon can be referred as *path independence* and it is very rare in non-Euclidean spaces. In fact, to the best of the authors' knowledge, this thesis is presenting the first known example of path independence in non-Euclidean spaces.

With the advancement of technology, more and more data sets that do not fit into the Euclidean framework became available to researchers. A major source of these is biological sciences; collecting detailed images of their objects of interest using advanced imaging technologies. The need to statistically analyze such non-traditional data sets gave rise to many innovations in statistics. The type of non-traditional setting we will be focusing in is sets of trees as data. Such sets arise in many contexts, such as blood vessel trees ([**6**]), lung airways trees ([**27**]), and phylogenetic trees ([**9**]).

A first starting point in PCA analysis for trees is [**28**], who attacked the problem of analyzing the brain artery structures obtained through a set of Magnetic Resonance Angiography (MRA) images. They modeled the brain artery system of each subject as a binary tree and developed an analog of the forward PCA in the binary tree space. They provided appropriate definitions of concepts such as distance, projection and line in binary tree space. They gave formulations of first, second, etc. principal components for binary tree data sets based on these definitions. This work has been the first study in adapting classical PCA ideas from Euclidean space to the new binary tree space. [**28**]'s definitions involve a vector of attributes for each node.

The PCA formulations of [**28**] gave rise to interesting combinatorial optimization problems. [**4**] provided an algorithm to find the optimal principal components in binary tree space in linear time. This work however used the simplified versions of [**28**]'s definitions where attributes are not considered. Only topology information is included in the analysis. This development enabled a numerical analysis on a full-size data set of brain arteries, revealing a correlation between their structure and age.

In the context of PCA in non-Euclidean spaces, [**19**] gave a backward PCA interpretation in image analysis. They focus on *mildly non-Euclidean*, or manifold data, and propose the use of Principal Nested Spheres as a backward step-wise approach.

[**21**] provided a concise overview of backward and forward PCA ideas and their applications in various non-classical contexts. They also mention the possibility of backwards PCA for trees: "... The notion of backwards PCA can also generate new approaches to tree line PCA. In particular, following the backwards PCA principal in full suggests first optimizing over a number of lines together, and then iteratively reducing the number of lines." This quote essentially summarizes one of our goals in this part of the thesis.

In this work, our first goal is to define and discuss the subject of rooted ordered tree spaces. We will elaborate on the correspondence concept, which is at the heart of any numerical analysis for ordered tree data. We will also suggest some indexing methods, and provide the generalized versions of some basic definitions such as distance, projection, etc.

Secondly, we will extend the definitions and results of [**28**] and [**4**] on forward PCA from binary tree space to the more general rooted ordered tree space and proceed with providing optimal algorithms for finding forward PC's. We will provide rigorous definitions and proofs of all the generalizations. Like [**4**], we will not consider the attribute vectors of [**28**] at this time.

Next, we attack the problem of finding an analog of dimension reduction. We first provide the definition for principal components with least influence (we call these *backward principal components*) in tree space, and define the optimization problem to be solved to reach them. We then provide a linear time algorithm to solve this problem to optimality.

Furthermore, we prove that the set of backward principal components in tree space is the same as the forward set, with order reversed, just like their counterparts in the classical Euclidean space. This equivalence is significant since the same phenomenon in Euclidean space is a result of orthogonality, and the concept of orthogonality does not carry over to the tree space. This result enables the analyst to switch between the two approaches as necessary while the results remain comparable, i.e., the components and their influence do not depend on which approach is used to find them. Therefore path independence property is valid in tree space PCA as well.

Our numerical results come from two main data sets. First one is an updated version of the brain artery data set previously used by [**4**]. Using our backward PCA tool, we investigate the effect of aging in brain artery structure in male and female subjects. We define two different kinds of age effect on the artery structure: overall branchyness and location-specific effects. We report our findings on both of these effects for male and female subpopulations. Secondly, we present a statistical analysis of the organization structure of a large US company. We present evidence on the structural differences across departments focusing on finance, marketing, sales and research and development (R&D).

2. The Tree Space

An *ordered tree* is a tree where a left-to-right order among siblings in the tree is given. For example, a leftmost child of a node is distinct from, say, the rightmost child of the same node. In this work we focus on ordered trees. This order is determined using *correspondence* ideas together with an *indexing scheme* to distinguish the node siblings. Moreover, we are focusing on rooted trees. A rooted ordered tree is a tree such that there is a single node designated as a root, and each node is indexed in such a way that a correspondence structure can be established between data trees. For the rest of the thesis, we will refer to the rooted ordered tree space as the *tree space*, and the term *tree* is reserved for rooted tree graphs in which each node is distinguished from each other through indexes. A data set, $\mathcal{T}$, is an indexed finite set of n trees. Note that we are only focused on the topology of data trees, and any attribute vectors that the nodes might carry will not be part of our analysis.

In computer science, a well-studied subject is the space of labeled trees. A *labeled tree* is a tree such that each node is assigned a symbol from a finite alphabet Σ (See [**8**]). In this space, a set of labeled nodes can form different trees if they are connected in different configurations. Here, our focus is not on labeled trees but indexed ordered trees. The index of a node is only determined by the node's location in a tree.

2.1. Correspondence. The concept of correspondence is a crucial element in analyzing sets of tree shaped objects. As a result, there is a growing research interest on the issue. In binary tree context, correspondence refers to which child of each node will be assumed left or right, effectively deciding which nodes across the data trees will be considered the same. For example, in Figure 17, the nodes indexed as **2** in both trees are assumed to be the same.

A consequence of using ordered trees is that, if two nodes correspond to each other, then their parents correspond as well. We will build the indexing schemes we propose based on this limitation.

For some tree data sets, the nature of the data uniquely determines the correspondence. In other examples, ambiguities may need to be resolved when there is no obvious correspondence choice available. An example of this is the brain artery data set we use in this thesis. With this data set, there is no clear way to decide which artery trunks correspond to each other beyond the root nodes.

[**4**] proposed two schemes: descendant correspondence, where the child with the more number of descendants is assumed to be the left node, and thickness correspondence, where the child with the larger median thickness measurement is put on left. It was later understood that the descendant correspondence approximates which blood artery trunks are "main" and which are "side" branches better than the thickness correspondence. This is mostly due to the fact that the MRA imaging technique used to obtain this data is not able to measure the thickness of the arteries to a desired accuracy level.

Another option considered, but not pursued, is to decide correspondence based on the location of the arteries in the brain. The apparently arbitrary growth of the arteries covering the surface of the brain render achieving a meaningful correspondence through physical location difficult.

The correspondence decision affects the statistical results obtained from any data set. As an example, [**4**] found a connection between the branching structure of the brain with aging under descendant correspondence, but the effect was not there when thickness correspondence was used. A scheme that represents the structures within the data set incorrectly can obscure the statistical connections related to these structures.

2.2. An Indexing Scheme. The correspondence within a data set is expressed through indices. Indices are "names" given to each node in a tree, so that the nodes with the same indices across data trees correspond to each other.

For binary trees, [**28**] proposed a level-order indexing method. In this scheme the root node has index 1. For the remaining nodes, if a node has index i, then the index of its left child is $2i$ and of its right child is $2i + 1$.

In this thesis, we will propose a somewhat similar technique, called *k-way tree indexing*. A k-way tree is a rooted tree in which each node has no more than k children (See [**12**]). We call the root node 0, and the j^{th} child of the node i is called $k * i + j$. This scheme reserves a unique index for all the nodes that exist in a full k-way tree, even if they do not exist in a particular instance. Furthermore, when k is known, it is easy to deduct the location of a given node from this index.

An example tree indexed using k-way tree indexing is given in Figure 16.

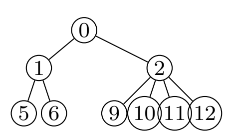


FIGURE 16. An example 4-way tree, indexed using the k-way tree indexing method.

A concept that will be useful in the following sections is the **support tree** ($Supp(\mathcal{T})$). The support tree of a data set is the smallest tree that includes all the members of the data set as sub-trees. Similarly, the **intersection tree** ($Int(\mathcal{T})$) of a data set is the largest sub-tree that is common to all of the data trees. They can be expressed as:

$$Supp(\mathcal{T}) = \cup_{i=1}^{n} t_i \text{ and } Int(\mathcal{T}) = \cap_{i=1}^{n} t_i,$$

where $\mathcal{T} = \{t_1, \ldots, t_n\}$ is a data set.

When indexing a data set of trees, the support tree should be used to determine the constant k and the indices. Constructing the support tree requires deciding a proper correspondence scheme before the indexing. This scheme will determine the order of the children of each node in every tree, and which ones correspond to which nodes in other trees.

2.3. Core Concepts. Some important concepts in the tree space include distance, median, line, and projection. These were defined previously for binary tree spaces ([**28**]). In this work we re-define these for rooted ordered tree space.

A *distance* metric between two trees is the symmetric difference of their nodes. Given two trees, t_1 and t_2, the distance between t_1 and t_2, denoted by $d(t_1, t_2)$, is

$$|t_1 \setminus t_2| + |t_2 \setminus t_1|,$$

where $|\cdot|$ is the number of nodes and $\setminus$ is the node set difference. In Figure 17, the nodes 0, 1 and 2 are common to both of the trees, so they do not contribute to the distance between them. The nodes 3, 4, 5 and 6 exist in one data tree but not in the other, therefore, the distance between the left and right trees in the figure is $|\{3, 4, 5, 6\}| = 4$.

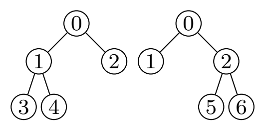


FIGURE 17. Two trees of which nodes are indexed using k-way tree indexing. The nodes 3, 4, 5 and 6 contribute to the distance.

It can easily be shown that d is a non-negative real function which has the following properties:

(1) $d(t_1, t_2) = 0$ if and only if $t_1 = t_2$,
(2) $d(t_1, t_2) = d(t_2, t_1)$, and
(3) $d(t_1, t_3) \leq d(t_1, t_2) + d(t_2, t_3)$,

where t_1, t_2 and t_3 are trees. Hence d is a metric.

Equipped with this distance metric, we can now define the *norm* or *length* of any data tree as its distance to the zero (empty tree):

$$|t| = d(t, \{\}).$$

A distance metric frequently used for labeled and unlabeled trees is the *tree edit distance*. Given a cost function and two trees, their tree edit distance is the minimal cost of a sequence of edit operations (inserting, deleting or relabeling nodes) to turn one of these trees into the other. (See [**8**], and [**16**]). In our case, we use ordered trees, and the distinction of any node from another comes from topology-based indices rather than labels. The tree edit distance for the ordered tree space only uses the insertion and deletion operations. When the cost of deletion and insertion are the same, the tree edit distance in the ordered tree space becomes equivalent to our distance definition given above. Another equivalent metric is *Robinson-Foulds distance* (see [**26**]). It is defined as the minimum number of contractions and decontractions of the edges required to transform one tree into another.

In our current setting, only the topological structure of the trees is considered as data. Any attributes of the nodes beyond their location and correspondence information is not part of the input. For this reason the distance metric given above is the best option to determine how close or far the data trees are in tree space.

In another setting where data trees or nodes carry numerical attributes, these may be incorporated into the distance measurement (see Wang and Marron (2007) for an example).

A very rich literature exists for distance metrics developed in the phylogenetic tree space. The *geodesic distance* was introduced in [**9**], and a polynomial time algorithm for computing geodesics in this space was found in [**24**]. To build the geodesic distance each tree is represented by its edge lengths as well as its topology. The distance is designed for understanding trees for which the differences between their edge lengths are very important. The emphasis on edge length is incompatible with the data sets we use in this study.

There are other distance measures that emphasize only the differences in topologies between the trees and not accounting any other attributes. For instance, the *nearest neighbor interchange (NNI) distance* was defined in [**25**] as the minimum number of *crossovers* converting a tree into another. The *subtree-prune-and-regraft (SPR) distance* was analyzed in [**15**], and the *tree bisection and reconnection (TBR) distance* was studied in [**3**]. In each of these cases these distances are defined as the minimum number of SPR or TBR operations to transform a tree to another. These distances assume that all the interior vertices of trees have degree 3.

A big limitation of distances developed for phylogenetic trees is that all trees have a fixed number of leaves. The data trees share the same leaf node set (corresponding to living or extinct species). This is a very big assumption that does not apply to our data.

In another context the *classical tree edit distance* (a variant of the tree edit distance) and *quotient Euclidean distance* are studied in [**13**]. These use the notion of "tree-shape" which is a tree embedded in $\mathbb{R}^2$ or $\mathbb{R}^3$, where the attributes describe the edges' location, size, etc. in space. These are not appropriate for a topology-only approach.

Once a notion of distance is constructed, the next step is to consider the *sample mean*, the centerpoint of a given set of data points. The fully discrete nature of the tree space means thinking in terms of sample median rather than the sample mean may be more appropriate for this space. [**28**] give a definition of the *median tree* of a tree data set. When the attribute vectors of each node is stripped, their median tree definition becomes such that the median of a tree set consisting of n trees is the tree consisting of the nodes that appear at least $n/2$ times in the data set.

The line concept in tree space is a close counterpart to the lines in Euclidean space. In the most general sense line refers to a set of points that are next to each other. These points lie in a given direction, which makes the line "one-dimensional". Due to the discrete nature of tree space, the points (trees) that are next to each other are defined as the points with distance 1, the smallest possible distance between two non-identical trees. To mimic the one-dimensional direction property, we require that every next point on the line in tree space is obtained by adding a child of most recently added node. The resulting construct is a set of trees that start from a starting tree and expand following a path away from the root, which is akin to the sense of direction in Euclidean space. A formal definition of a line in tree space is as follows:

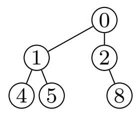
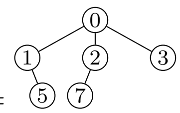
DEFINITION 5.1. *Given a data set $\mathcal{T}$, a* **tree-line**, $L = \{l_0, \ldots, l_k\}$, *is a sequence of trees where l_0 is called the starting tree, and l_i is defined from l_{i-1} by the addition of a single node $v_i \in Supp(\mathcal{T})$. In addition, each v_i is a child of v_{i-1}.*

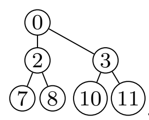
See Example 5.2 for an example tree-line.

As noted by [**28**], it is important that the starting tree l_0 should be defined minimally. This will allow the principal components of the next section to have greater flexibility to explain the variation in the data set.

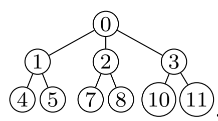
The next concept to construct is *projection* in this space. In general, the projection of a point onto an object can be defined as the closest point on the object to the projected point. This can be formalized in tree space as $P_L(t) = \arg\min_{l \in L}\{d(t,l)\}$, where $P_L(t)$ is the projection of tree t onto the object L. The projection can be regarded as the point in the object most similar to the data tree. Example 5.2 contains a small data set and a tree-line, and illustrates how the projection of each data point onto the given tree-line can be found.

EXAMPLE 5.2. *Let us consider the following data set $\mathcal{T}$ consisting of 3 data points:*

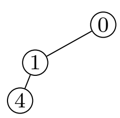
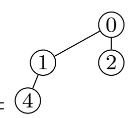
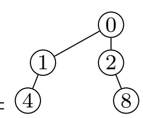
$t_1 =$  , $t_2 =$  , *and*

$t_3 =$  .

Thus the support tree $Supp(\mathcal{T})$ is

 .

Also considrer the following tree-line

$L = \Big\{$ $l_0 =$  , $l_1 =$  , $l_2 =$  $\Big\}$.

The following table gives the distance between each tree of $\mathcal{T}$ and each tree of L:

	l_0	l_1	l_2
t_1	*3*	*2*	*1*
t_2	*5*	*4*	*5*
t_3	*8*	*7*	*6*

So, we can observe that $P_L(t_1) = l_2$, $P_L(t_2) = l_1$ and $P_L(t_3) = l_2$.

The tree-lines approximate one dimensional directions in the tree space. The combination of multiple directions (in other words, “subspaces” of more than one dimension) in tree space can be represented by a *union* of tree-lines. Remember that, in Euclidean space, a point lying on the union of two lines is the combination of one point from the first line, and another point from the second line. These two points do not need to be unique. The two dimensional subspace determined by these two directions is the set of all possible combinations of the points on the two directions.

In tree space, the same concept is employed. We say that given tree-lines $L_1 = \{l_{1,0}, l_{1,1}, \ldots, l_{1,m}\}$, $L_2 = \{l_{2,0}, l_{2,1}, \ldots, l_{2,n}\}$, their union is the set of all possible unions of members of L_1 and L_2:

$$L_1 \cup L_2 \quad = \quad \{l_{1,i} \cup l_{2,j} \mid i \in \{0, \cdots, m\}, j \in \{0, \cdots, n\}\}.$$

Note that in an Euclidean setting, every point on either line is a member of the union of both lines. In the case of tree-lines however, this is only true if both lines start at the same starting point.

In light of this, the projection of a tree t onto $L_1 \cup L_2$ is the point in set $L_1 \cup L_2$ with the smallest distance to t:

$$P_{L_1 \cup L_2}(t) = \arg \min_{l \in L_1 \cup L_2} \{d(t, l)\}.$$

Note that even though the union is explained using only two directions for simplicity, the definition can easily be extended to the combination of any number of directions. Also note that the length of a projection is the norm of the projection.

Now we define the concept of "path" that will be useful later on. Given a tree-line $L = \{l_0, \cdots, l_k\}$, the **path** of L, p_L, is the unique path from the root to v_k, the last node added in L. Note that our path definition is different than the one given in [**4**], which included only the nodes added to the starting tree l_0 instead of forming a set starting from the root node.

The following lemma provides an easy-to-use formula for the projection of a data point. The proof of it can be found in the Appendix.

LEMMA 5.3. *Let t be a tree and $L = \{l_0, \cdots, l_k\}$ be a tree-line. Then*

$$P_L(t) = l_0 \cup (t \cap p_L).$$

PROOF. Since $l_i = l_{i-1} \cup v_i$, we have that

$$d(t, l_i) = \begin{cases} d(t, l_{i-1}) - 1 & \text{if } v_i \in t, \\ d(t, l_{i-1}) + 1 & \text{otherwise.} \end{cases}$$

In other words, the distance of the tree to the line decreases as we keep adding nodes of p_L that are in t, and when we step out of t, the distance begins to increase. □

It follows that projection of a tree onto a tree-line is unique.

In theory, a line extends to infinity. In this thesis, we limit our scope to the line pieces that reside within the support tree of a given data set since extending these lines outside the support tree's scope would introduce unnecessary trivialities. We also consider only the tree-lines that are not trivial: The tree-lines that consist of l_0 and at least one more point. Finally, we only consider tree-lines that are maximal, i.e., whose paths cannot be extended within $Supp(\mathcal{T})$. It can be seen that in searching for PC tree-lines, non-maximal tree-lines are dominated by maximal tree-lines that contain them.

In the light of this, we let $\mathcal{L}$ denote the set of all nontrivial maximal tree-lines with staring point l_0, contained in $Supp(\mathcal{T})$. Also we name $\mathcal{P}$ to be the set of all paths in $Supp(\mathcal{T})$ from the root to leaves that are not in l_0. It is easy to see that $\mathcal{P}$ is the set of paths of tree-lines in $\mathcal{L}$. Also note that $|\mathcal{L}| = |\mathcal{P}| = n$ and $Supp(\mathcal{T}) = l_0 \cup \bigcup_{p_L \in \mathcal{P}} p_L$.

3. Forward PCA in Tree Space

The concepts of first, second, etc. principal components were developed previously for binary tree space (See [**28**] for first principal component definition, and [**4**] for the other components.) In this section, we will present principal components for the general tree space. The proofs of this section's results are in the Appendix.

We also note that this section is devoted to the "forward PCA" approach where directions that carry the most amount of variation are sought. We will develop the "backward PCA" approach in the upcoming section.

Principal components in Euclidean space are directions where the projection of data onto it has the largest variation. They can also be defined as directions where the distance of data to it is minimum. In Euclidean space, these two formulations are equivalent, and (first few) principal components are seen as the basis of the subspace with the best **coverage** of data while keeping the dimensions low. In other words, they form the most **representative** subspace.

In tree space, however, the two formulations are not equivalent. Following [**28**] and [**4**], we chose to employ the minimum distance formulation. The principal components obtained using this formulation provides the best coverage of the data, even though they do not necessarily comply with the variation definition of [**28**], developed using the median tree notion.

The first principal component was defined as the tree-line that minimizes the sum of distances of the data points to their projections on the line. This can be viewed as the one-dimensional line that best fits the data. We will provide a similar definition below, adopted to the general tree space.

DEFINITION 5.4. *Given a starting point l_0, the* **first (forward) principal component tree-line***, PC1, is*

$$L_1^f = \arg\min_{L \in \mathcal{L}} \sum_{t \in \mathcal{T}} d(t, P_L(t)).$$

Notice that this principal component definition requires determining a starting point l_0, while its Euclidean space counterpart contains no such notion. The lines in Euclidean space can extend in both directions indefinitely. Due to the structure of the tree space, lines can indefinitely extend in only one direction (although we limit our attention to the line pieces contained in the support tree for practical purposes). Therefore, they have to be limited on the other end.

Some appropriate starting point options were discussed previously in [**28**] and [**4**]. Some suggestions are using the median tree, the root tree or the intersection tree as the starting point. The starting point tree is included in every line, and any variation that may exist within this tree cannot be detected by the tree-lines extending from it. The starting point should be selected carefully based on this information. In this thesis, the root node is selected as the starting point for both of the analyses.

The second principal component is the line which, when combined with the first principal component, minimizes the sum of the distances of the data points to this combination. The k^{th} principal component can be deduced similarly. For the tree space, this notion is formalized as follows:

DEFINITION 5.5. *Given a starting point l_0, the* **k-th (forward) principal component tree-line** *is defined recursively as*

$$L_k^f = \arg\min_{L \in \mathcal{L}} \sum_{t \in \mathcal{T}} d(t, P_{L_1^f \cup \cdots \cup L_{k-1}^f \cup L}(t)).$$

The path of the k-th principal component tree-line is p_k^f.

As we will see in Example 5.9, the definition of the principal components allows multiple solutions. A tie-breaking rule depending on the nature of the data should be established to reach consistent results in the existence of ties. Such a rule can be established by determining a total ordering over the set of all tree-lines, or equivalently, the set of all paths. When a tie breaking rule prefers path p_L over $p_{L'}$, we denote this by $p_L > p_{L'}$.

The following lemma describes a key property that will be used to interpret the projection of a tree onto a subspace defined by a set of tree-lines. The reader may refer to the Appendix for the proof.

LEMMA 5.6. *Let $L_1, L_2, \ldots,$ and L_q be tree-lines with a common starting point, and t be a tree. Then*

$$P_{L_1 \cup \cdots \cup L_q}(t) = P_{L_1}(t) \cup \cdots \cup P_{L_q}(t).$$

PROOF. For simplicity, we only prove the statement for $q = 2$. Assume that

$$L_1 = \{l_{1,0}, l_{1,1}, \ldots, l_{1,k_1}\}, L_2 = \{l_{2,0}, l_{2,1}, \ldots, l_{2,k_2}\}$$

with $l_0 = l_{1,0} = l_{2,0}$,

$$\begin{aligned} l_{1,i} &= l_{1,i-1} \cup v_{1,i} \quad \text{for } 1 \leq i \leq k_1, \text{ and} \\ l_{2,j} &= l_{2,j-1} \cup v_{2,j} \qquad \text{for } 1 \leq j \leq k_2. \end{aligned}$$

Also assume

$$P_{L_1}(t) = l_{1,r_1} \tag{9}$$

and

$$P_{L_2}(t) = l_{2,r_2}. \tag{10}$$

Let $f(i,j)$ be the distance between the trees t and $l_{1,i} \cup l_{2,j}$, for $1 \leq i \leq k_1$ and $1 \leq j \leq k_2$. Using lemma 5.3, equation (9) means

$$\begin{aligned} v_{1,i} &\in t, \quad \text{if } i \leq r_1, \text{ and} \\ v_{1,j} &\in t, \qquad \text{if } j \leq r_2. \end{aligned}$$

Hence,

$$\begin{aligned} f(i,j) &\leq f(i-1,j), \quad \text{if } i \leq r_1, \\ f(i,j) &\geq f(i-1,j), \quad \text{if } i > r_1. \end{aligned} \tag{11}$$

By symmetry, we have

$$\begin{aligned} f(i,j) &\leq f(i,j-1), \quad \text{if } j \leq r_2, \\ f(i,j) &\geq f(i,j-1), \quad \text{if } j > r_2. \end{aligned} \tag{12}$$

Overall, equations (11) and (12) imply that the function f attains its minimum at $i = r_1, j = r_2$, which is what we had to prove. $\square$

[**4**] provided a linear time algorithm to find the forward principal components in binary tree space. We will give a generalization of that algorithm in tree space, and prove that the extended version also gives the optimal PC's.

The algorithm uses a notion of "weight" of nodes in the support tree. For nodes that are covered by any of the already computed principal components or the starting point, the weight is zero. For a node that did not appear in any of the already computed principal components, the weight is the number of times it appears in the data set $\mathcal{T}$. To formalize this idea, we first need an indicator function. Let δ be an indicator function, defined as $\delta(v,t) = 1$ if $v \in t$, and 0 otherwise.

Then, given $L_1^f, \ldots,$ and L_{k-1}^f, the first $k-1$ PC tree-lines, the k-th weight of a node $v \in Supp(\mathcal{T})$ is

$$w_k(v) = \begin{cases} 0, & \text{if } v \in l_0 \cup p_1^f \cup \cdots \cup p_{k-1}^f, \\ \sum_{t \in \mathcal{T}} \delta(v,t), & \textit{otherwise.} \end{cases}$$

How will the weight $w_k(v)$ be useful in finding principal components? Remember that the k^{th} PC is the line which minimizes the sum of distances of the data points to the union of first k PC's. When selecting k^{th} PC, we consider the lines which will decrease this sum of distances most. In each candidate line, the nodes that already appear in l_0 or a previously selected PC do not contribute to this reduction. However, the nodes that have not yet appeared in l_0 or first $k-1$ PC's reduce this sum by $w_k(v)$. (This follows from Lemma 5.3.) This reasoning leads us to the following algorithm:

ALGORITHM 5.7. *Forward algorithm. Let $\mathcal{T}$ be a data set and $\mathcal{L}$ be the set of all tree-lines with the same starting point l_0.*
Input: *The first $(k-1)$-st PC tree-lines: $L_1^f, \ldots,$ and L_{k-1}^f.*
Output: *A tree-line.*
Examine the paths of the tree-lines in $\mathcal{L}$. Return the tree-line whose path maximizes the sum of w_k weights in the support tree. Break ties according to an appropriate tie-breaking rule.

The next theorem formalizes that the tree-line returned by the forward algorithm is precisely the k-th PC tree-line. The proof is in the Appendix.

THEOREM 5.8. *For a given starting point l_0, let $L_1^f, \ldots,$ and L_{k-1}^f be the first $(k-1)$-st PC tree-lines. Then, the forward algorithm returns the kth PC tree-line, L_k^f.*

PROOF. The definition of k^{th} PC tree-line in terms of paths is equivalent to the equation

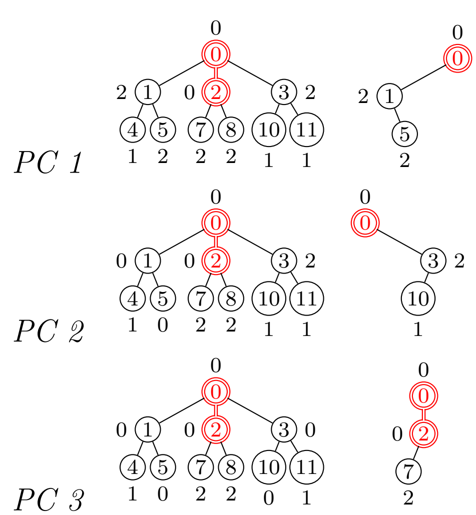
$$\begin{aligned}
p_k^f &= \arg\min_{p_L \in \mathcal{P}} \sum_{t \in \mathcal{T}} d\left(t, l_0 \cup \left(\left(\cup_{i=1\cdots k-1} p_i^f \cup p_L\right) \cap t\right)\right) \\
&= \arg\min_{p_L \in \mathcal{P}} \sum_{t \in \mathcal{T}} \left\{ \left| t \setminus \left(l_0 \cup \left(\left(\cup_{i=1\cdots k-1} p_i^f \cup p_L\right) \cap t\right)\right)\right| \right. \\
&\qquad \left. + \left| \left(l_0 \cup \left(\left(\cup_{i=1\cdots k-1} p_i^f \cup p_L\right) \cap t\right)\right) \setminus t \right| \right\} \\
&= \arg\min_{p_L \in \mathcal{P}} \sum_{t \in \mathcal{T}} \left\{ \left| t \setminus \left(l_0 \cup p_1^f \cup \cdots \cup p_{k-1}^f \cup p_L\right)\right| \right. \\
&\qquad \left. + \left| \left(l_0 \cup \left(\left(p_1^f \cup \cdots \cup p_{k-1}^f \cup p_L\right) \cap t\right)\right) \setminus t \right| \right\}
\end{aligned}$$

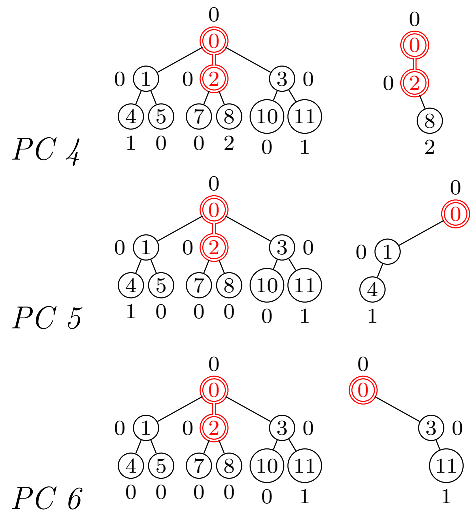
$$
\begin{aligned}
&= \arg\min_{p_L \in \mathcal{P}} \sum_{t \in \mathcal{T}} \left\{ \left| t \setminus \left(l_0 \cup p_1^f \cup \cdots \cup p_{k-1}^f \cup p_L \right) \right| + |l_0 \setminus t| \right\} \\
&= \arg\min_{p_L \in \mathcal{P}} \sum_{t \in \mathcal{T}} \left| t \setminus \left(l_0 \cup p_1^f \cup \cdots \cup p_{k-1}^f \cup p_L \right) \right| \\
&= \arg\min_{p_L \in \mathcal{P}} \sum_{t \in \mathcal{T}} \left\{ \left| t \setminus \left(l_0 \cup p_1^f \cup \cdots \cup p_{k-1}^f \right) \right| \right. \\
&\qquad \left. - \left| (t \cap p_L) \setminus \left(l_0 \cup p_1^f \cup \cdots \cup p_{k-1}^f \right) \right| \right\} \\
&= \arg\min_{p_L \in \mathcal{P}} - \sum_{t \in \mathcal{T}} \left| (t \cap p_L) \setminus \left(l_0 \cup p_1^f \cup \cdots \cup p_{k-1}^f \right) \right| \\
&= \arg\max_{p_L \in \mathcal{P}} \sum_{t \in \mathcal{T}} \left| (t \cap p_L) \setminus \left(l_0 \cup p_1^f \cup \cdots \cup p_{k-1}^f \right) \right| \\
&= \arg\max_{p_L \in \mathcal{P}} \sum_{v \in p_L} w_k(v).
\end{aligned}
$$

The last equation correspond to the path with maximum sum of w_k weights in the support tree. □

To better explain how the algorithm works, we will apply the forward algorithm to the toy data set given in Example 5.2.

EXAMPLE 5.9. *In this example, we select the tree-line with the leftmost path in the event of a tie. We take the intersection tree as the starting point (illustrated in red below). The table given below summarizes iterations of the algorithm, where each row corresponds to one iteration. At each of the iterations, the name of the principal component obtained at that iteration is given in left column. The support tree with updated weights ($w_i'(.)$) is given in the middle column. The paths of selected PC tree-lines according to these weights is given in right column.*





The forward algorithm finds the k-th PC tree-line in linear time. We finish this section with the proof of this fact.

THEOREM 5.10. *For a data set with support tree of m nodes, the running time of computing the k-th PC tree-line is $O(m)$.*

The sketch of the proof goes as follows. Forward algorithm adds the weights of the nodes in the path of each tree-line on the support tree, and returns the tree-line with the maximum sum of weights. To do this, the algorithm starts from the root node and moves down level by level, going over each node once. At each node v, a total weight $\widehat{w}(v)$ is computed. The $\widehat{w}(v)$ is the sum of weight $w_k(v)$ (as defined above) and the weight $\widehat{w}$ of the parent of v. The initial weight value $\widehat{w}(r)$ of the nodes that belong to the starting point is zero. The algorithm returns the path whose leaf-node has the maximum weight $\widehat{w}$ as the k-th principal component. Since the algorithm goes over every node in support tree once, the running time is $O(m)$.

4. Dimension Reduction in Tree Space

In this section, we will define *backward principal component tree-lines*. This structure is the tree space equivalent of the backward principal component in the classical dimension reduction setting. They represent the directions that carry the least information about the data set and thus can be taken out. Our definition describes backward principal components as directions such that when eliminated, the remaining subspace will have the minimum sum of squared distances to the data points. These are considered to be the components with least influence.

As stated before, given a data set $\mathcal{T}$, the number of distinct tree-lines that do not include each other is n. These n tree-lines are able to express the support tree of $\mathcal{T}$ fully. In our setting, we denote the tree-line with least influence as $BPCn$, the n^{th} backward principal component tree-line. The next backward PC with second least influence is $BPC(n-1)$, and so on. The indexing is selected this way to ensure compatibility with the forward components.

DEFINITION 5.11. *Given a starting point l_0, the* $\mathbf{n^{th}}$ **backward principal component tree-line***,* BPCn*, is*

$$L_n^b = \arg\min_{L \in \mathcal{L}} \sum_{t \in \mathcal{T}} d(t, P_{\bigcup L' \in \mathcal{L} \setminus \{L\}}(t)).$$

The $(\mathbf{n-k})^{\mathbf{th}}$ **backward principal component tree-line** *is defined recursively as*

$$L^b_{n-k} = \quad \arg\min_{L\in\mathcal{L}\setminus\{L^b_n,\cdots,L^b_{n-k+1}\}} \sum_{t\in\mathcal{T}} d(t, P_{\bigcup L'\in\mathcal{L}\setminus\{L^b_n,\cdots,L^b_{n-k+1},L\}}(t)). \tag{13}$$

The path associated to the $(n-k)$-th backward principal component tree-line will be denoted by p^b_{n-k}.

Like in the forward approach, finding the backward PC's require a way of measuring the effect of each node on the objective function at each step. Consider the situation where we are searching the $(n-k)^{th}$ BPC. At this point, k least influent components are already found and taken out, and the remaining subspace consists of $n-k$ components. To do this, we will need a modified version of the weight concept. The nodes that are not part of the remaining subspace are at this point irrelevant, so their *backward weights* are zero. Similarly, nodes of the starting point l_0 have weight zero. For nodes that are in the subspace (but not in the starting point), there are two possibilities. If a node is covered by only one path, and this node appears in the data trees w times, then selecting this path as BPC_{n-k} will mean this node will reduce the sum of distances by w. If this node is covered by multiple paths in the remaining subspace, then selecting any of those paths as the BPC_{n-k} will not result in removing this node from the remaining subspace, since it is already covered by other paths as well. Therefore the sum of distances will not be reduced due to this node, and it's weight is zero.

Our methodology to find the backward components relies on finding the paths whose removal from the current subspace will result in the least amount of increase in the sum of distances. At each step, we go over all the candidate paths, and sum the backward weights of all of its nodes. This sum is equal to the amount of increase in the sum of distances of the data points to the remaining subspace as a result of selecting that path as the next backward component. We select the path with the least sum of backward weights.

The formal definition of the backward weights and the backward algorithm is given as follows:

DEFINITION 5.12. *Given a starting point l_0, let $L^b_n, \ldots,$ and L^b_{n-k+1} be the last k BPC tree-lines and $\boldsymbol{B} = \mathcal{P}\setminus\{p^b_n,\ldots,p^b_{n-k+1}\}$. For $v \in Supp(\mathbf{B})$, the $(n-k)$-th backward weight of the node v is*

$$w'_{n-k}(v) = \begin{cases} 0 & \text{If } v\in l_0 \text{ or } v \text{ belongs to at} \\ & \text{least two different paths of } \mathbf{B}, \\ \sum_{t\in\mathcal{T}}\delta(v,t) & \text{Otherwise.}\end{cases}$$

ALGORITHM 5.13. *Backward Algorithm. Let $\mathcal{T}$ be a data set and $\mathcal{L}$ be the set of all tree-lines with the same starting point l_0.*
Input: *The last k BPC tree-lines: $L^b_n, \ldots,$ and L^b_{n-k+1}.*
Output: *The $(n-k)^{th}$ BPC tree-line: L^b_{n-k}.*
Let $\boldsymbol{B} = \mathcal{P}\setminus\{p^b_n,\ldots,p^b_{n-k+1}\}$.
Examine the paths of the tree-lines in $\mathcal{L}$. Return the tree-line L^b_{n-k} whose path minimizes the sum of w'_k weights in the support tree $Supp(\mathbf{B})$. If there is more than one candidate, select the BPC according to an appropriate tie-breaking rule.

The key theoretical result of the section, the optimality of the backward algorithm, is summarized as follows:

THEOREM 5.14. *For a starting point l_0, let $L^b_n, \ldots,$ and L^b_{n-k+1} be the last k BPC tree-lines. Then, the backward algorithm returns the optimum $(n-k)^{th}$ BPC tree-line, L^b_{n-k}.*

PROOF. The definition of k^{th} BPC tree-line (see Equation 13) in terms of paths is equivalent to the equation

$$
\begin{aligned}
p^b_{n-k} \quad = \quad & \arg\min_{p_L \in \mathbf{B}} \sum_{t \in \mathcal{T}} d\left(t, l_0 \cup \left(\left(\cup_{p \in \mathbf{B} \setminus \{p_L\}} p\right) \cap t\right)\right), \\
& \text{where } \mathbf{B} = \mathcal{P} \setminus \{p^b_n, \ldots, p^b_{n-k+1}\}
\end{aligned}
$$

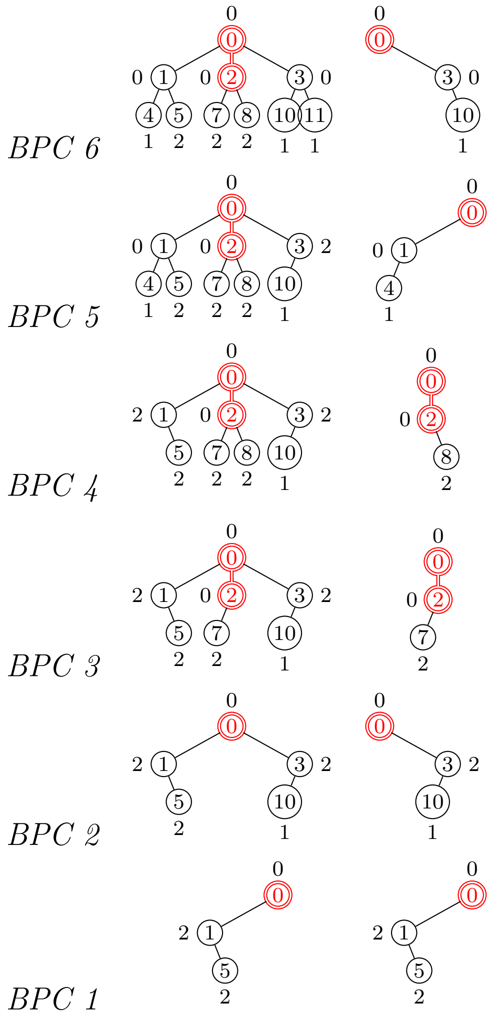
$$
\begin{aligned}
= \quad & \arg\min_{p_L \in \mathbf{B}} \sum_{t \in \mathcal{T}} \left\{ \left| t \setminus l_0 \cup \left(\left(\cup_{p \in \mathbf{B} \setminus \{p_L\}} p\right) \cap t\right) \right| \right. \\
& \left. + \left| l_0 \cup \left(\left(\cup_{p \in \mathbf{B} \setminus \{p_L\}} p\right) \cap t\right) \setminus t \right| \right\} \\
= \quad & \arg\min_{p_L \in \mathbf{B}} \sum_{t \in \mathcal{T}} \left\{ \left| t \setminus l_0 \cup \left(\left(\cup_{p \in \mathbf{B} \setminus \{p_L\}} p\right) \cap t\right) \right| \right. \\
& \left. + \left| (l_0 \setminus t) \cup \left(\left(\left(\cup_{p \in \mathbf{B} \setminus \{p_L\}} p\right) \cap t\right) \setminus t\right) \right| \right\} \\
= \quad & \arg\min_{p_L \in \mathbf{B}} \sum_{t \in \mathcal{T}} \left\{ \left| t \setminus l_0 \cup \left(\left(\cup_{p \in \mathbf{B} \setminus \{p_L\}} p\right) \cap t\right) \right| + |l_0 \setminus t| \right\} \\
= \quad & \arg\min_{p_L \in \mathbf{B}} \sum_{t \in \mathcal{T}} \left| t \setminus l_0 \cup \left(\left(\cup_{p \in \mathbf{B} \setminus \{p_L\}} p\right) \cap t\right) \right| \\
= \quad & \arg\min_{p_L \in \mathbf{B}} \sum_{t \in \mathcal{T}} \left| t \setminus l_0 \cup \left(\cup_{p \in \mathbf{B} \setminus \{p_L\}} p\right) \right| \\
= \quad & \arg\min_{p_L \in \mathbf{B}} \sum_{t \in \mathcal{T}} \left| (t \cap p_L) \setminus \left(l_0 \cup \left(\cup_{p \in \mathbf{B} \setminus \{p_L\}} p\right)\right) \right| \\
& + \sum_{t \in \mathcal{T}} \left| \left(t \cap \left(\cup_{p \in \mathcal{P} \setminus \mathbf{B}} p\right)\right) \setminus (l_0 \cup (\cup_{p \in \mathbf{B}} p)) \right| \\
= \quad & \arg\min_{p_L \in \mathbf{B}} \sum_{t \in \mathcal{T}} \left| (t \cap p_L) \setminus \left(l_0 \cup \left(\cup_{p \in \mathbf{B} \setminus \{p_L\}} p\right)\right) \right| \\
= \quad & \arg\min_{p_L \in \mathbf{B}} \sum_{t \in \mathcal{T}} \sum_{v \in (t \cap p_L) \setminus \left(l_0 \cup \left(\cup_{p \in \mathbf{B} \setminus \{p_L\}} p\right)\right)} 1 \\
= \quad & \arg\min_{p_L \in \mathbf{B}} \sum_{v \in p_L} w'_k(v).
\end{aligned}
$$

From the last equation the result follows. □

The proof of this theorem is in the Appendix.

Next, we provide an example illustrating the steps of the backward algorithm. We will apply the backward algorithm to the toy data set given in Example 5.2 and use the same starting point. Furthermore, we use the opposite tie-breaking rule we used in the forward algorithm, in this case is to select the rightmost tree-line.

EXAMPLE 5.15. *The table given below summarizes iterations of the algorithm, where each row corresponds to one iteration. At each of the iterations, the name of the backward principal* *component obtained at that iteration is given in left column. The pruned support tree with updated weights ($w'_i(.)$) is given in the middle column. The paths of selected PC tree-lines according to these weights is given in right column.*



Like the forward algorithm, the backward algorithm also finds the optimal solution in linear time:

THEOREM 5.16. *For a data set with support tree of m nodes, the running time of computing the k-th BPC tree-line is $O(m)$.*

The proof of this theorem is very similar to that of Theorem 5.10, so it will be omitted to save space.

5. Equivalence of PCA and BPCA in Tree Space

A very important aspect of tree space is that, the notion of orthogonality does not exist. In the Euclidean space equivalent of backward PCA, the orthogonality property ensures that the components do not depend on the method used to find them, i.e., the most informative principal component is the same when forward or backward approaches are used. This powerful property of path-independence brings various advantages to the analyst.

In this section, we will prove that, under appropriate tie-breaking rules, the forward and backward approaches are equivalent in the tree space as well when tree-lines are used. This is a surprising result given the lack of any notion of orthogonality. In practice, this result will ensure that the components of backward and forward approaches in binary tree space are comparable.

We will show this equivalence by proving that, for each $1 \leq k \leq n$, the k^{th} PC tree-line and the k^{th} BPC tree-line are equal. An equivalent statement is that their paths are equal: $p_k^f = p_k^b$. Without loss of generality, we will assume that a consistent tie-breaking method is established for both methods in choosing principal components whenever candidate tree-lines have the same sum of weights.

All the proofs can be found in the Appendix.

PROPOSITION 5.17. *Given an integer $1 \leq k \leq n$, let $p_1^f, ...,$ and p_k^f be the paths of the first k principal components yielded by the forward algorithm and let $p_n^b, ...,$ and p_{k+1}^b be the paths of the last $n-k$ principal components yielded by the backward algorithm, then there exist no i and j such that $1 \leq i \leq k < j \leq n$ and $p_i^f = p_j^b$.*

PROOF. Suppose there exist i and j with $1 \leq i \leq k < j \leq n$ and $p_i^f = p_j^b$. Without loss of generality, suppose that j is the largest index where the assumption holds. Let p_L denote the path $p_i^f = p_j^b$, and let $B = \{p_n^b, ..., p_{j+1}^b\}$. Since $1 \leq i \leq k < j \leq n$, the set of paths $\mathcal{P} \setminus \{B\}$ contains at least two paths. Let $v \in p_L$ be the first node from the leaf to the root that has at least two children in $Supp(\mathcal{P} \setminus \{B\})$. There are two possibilities:

I. $v \notin l_0$ *i.e.* there is at least one path different of p_L in $\mathcal{P} \setminus \{B\}$ that has v as node or

II. $v \in l_0$.

In both cases, $w'_j(u) = 0$ for all u in the path p_L from v to the root.

Consider case *I*. Let $p_{L'} \in \mathcal{P} \setminus \{B\}$ be a path different from p_L that contains v in it. Let p_v be the path from the root to v. Since $p_L = p_j^b$

$$\sum_{u \in p_L \setminus p_v} w'_j(u) = \sum_{u \in p_L} w'_j(u) \leq \sum_{u \in p_{L'}} w'_j(u) \tag{14}$$

$$= \sum_{u \in p_{L'} \setminus p_v} w'_j(u). \tag{15}$$

(16)

On the other hand, since $p_L = p_i^f$

$$\sum_{u \in p_L} w_i(u) \geq \sum_{u \in p_{L'}} w_i(u). \tag{17}$$

Next, we need to show that following holds:

$$\sum_{u \in p_{L'} \setminus p_v} w'_j(u) \leq \sum_{u \in p_{L'} \setminus p_v} w_i(u). \tag{18}$$

To do this, suppose that:

$$\sum_{u \in p_{L'} \setminus p_v} w'_j(u) > \sum_{u \in p_{L'} \setminus p_v} w_i(u).$$

It implies that there is at least one node v' that has $w'_j(v') > 0$ and $w_i(v') = 0$. Since $w_i(v') = 0$, a path that contains v' and is different of $p_{L'}$ was yielded by the forward algorithm before $p_{L'}$.

However, this implies that there are at least two paths that has v' as node at step j in the backward algorithm, then $w'_j(v') = 0$. This gives a contradiction.

It is straightforward to see

$$\sum_{u \in p_L \setminus p_v} w_i(u) \leq \sum_{u \in p_L \setminus p_v} w'_j(u). \tag{19}$$

Let us suppose that the inequality in (14) is strict, *i.e.*

$$\sum_{u \in p_L \setminus p_v} w'_j(u) < \sum_{u \in p_{L'} \setminus p_v} w'_j(u). \tag{20}$$

We have

$$\begin{aligned} \sum_{u \in p_L} w_i(u) &= \sum_{u \in p_v} w_i(u) + \sum_{u \in p_L \setminus p_v} w_i(u) \\ &\leq_{(19)} \sum_{u \in p_v} w_i(u) + \sum_{u \in p_L \setminus p_v} w'_j(u) \\ &<_{(20)} \sum_{u \in p_v} w_i(u) + \sum_{u \in p_{L'} \setminus p_v} w'_j(u) \\ &\leq_{(18)} \sum_{u \in p_v} w_i(u) + \sum_{u \in p_{L'} \setminus p_v} w_i(u) \\ &= \sum_{u \in p_{L'}} w_i(u) \end{aligned}$$

which is a contradiction to equation (17). Therefore, equation (14) has to be an equality, *i.e.*

$$\sum_{u \in p_L \setminus p_v} w'_j(u) = \sum_{u \in p_{L'} \setminus p_v} w'_j(u). \tag{21}$$

If one or both inequalities

$$\sum_{u \in p_{L'} \setminus p_v} w'_j(u) < \sum_{u \in p_{L'} \setminus p_v} w_i(u)$$

and

$$\sum_{u \in p_L \setminus p_v} w_i(u) < \sum_{u \in p_L \setminus p_v} w'_j(u),$$

holds, then the result follows in the same way as above.

Finally, let us suppose

$$\sum_{u \in p_{L'} \setminus p_v} w'_j(u) = \sum_{u \in p_{L'} \setminus p_v} w_i(u)$$

and

$$\sum_{u \in p_L \setminus p_v} w_i(u) = \sum_{u \in p_L \setminus p_v} w'_j(u),$$

which implies that

$$\sum_{u \in p_{L'}} w'_j(u) = \sum_{u \in p_L} w'_j(u) \text{ and } \sum_{u \in p_{L'}} w_i(u) = \sum_{u \in p_L} w_i(u).$$

Now, since $p_i^f = p_L$, we have $p_L > p_{L'}$. And, since $p_j^b = p_L$, we have $p_L < p_{L'}$. Which is a contradiction.

In the case II, where $v \in l_0$, let v' be the last node from the root to the leaf in p_L that belongs to l_0. Take $p_{L'} \in \mathcal{P} \setminus \{B\}$ as a different path of p_L, and v'' as the last node from the root to the leaf in $p_{L'}$ that belongs to l_0. Let $p_{v'}$ be the unique path from the root to the node v' and $p_{v''}$ the unique path from the root to the node v''. Since $p_{v'}$ and $p_{v''}$ are contained in l_0, we have

$$\begin{aligned}\sum_{u \in p_{v'}} w_i(u) &= \sum_{u \in p_{v''}} w_i(u) = \sum_{u \in p_{v'}} w'_j(u) \\ &= \sum_{u \in p_{v''}} w'_j(u) = 0.\end{aligned}$$

Since $p_L = p_j^b$

$$\sum_{u \in p_L} w'_j(u) \leq \sum_{u \in p_{L'}} w'_j(u) \tag{22}$$

On the other hand, since $p_L = p_i^f$

$$\sum_{u \in p_L} w_i(u) \geq \sum_{u \in p_{L'}} w_i(u). \tag{23}$$

Similar to case I, we can see that 22 is an equality. This gives a contradiction. □

This proposition motivates the following theorem:

THEOREM 5.18. *For each $1 \leq k \leq n$ the k^{th} PC tree-line obtained by the forward algorithm is equal to the k^{th} BPC tree-line obtained by the backward algorithm.*

PROOF. By the proposition 5.17, we have that at step $n-1$ of the forward algorithm there is no tree-line yielded by the forward algorithm equal to L_n^b, then $L_n^b = L_n^f$. At the step $n-2$, there is no tree-line yielded by the forward algorithm equal to L_n^b or L_{n-1}^b. Since $L_n^b = L_n^f$, we have the $L_{n-1}^b = L_{n-1}^f$. We continue iteratively until step 1. At the end, we will have $L_k^b = L_k^f$ for all $1 \leq k \leq n$. □

This result guarantees the comparability of principal components obtained by either method, enabling the analyst to use them interchangeably depending on which type of analysis is appropriate at the time.

Just like in PCA of Euclidean space, when two or more candidate directions exist for the next principal component, a tie-breaking scheme should exist to resolve ambiguity. For the equivalence of forward and backward PCA, the tie breaking rule for each method should mirror each other. The tie situation is not very likely in Euclidean space, but because the tree space is discrete, it may arise more often. In our numerical analyses, we see that support trees carry nodes that only exist in one data tree. These tend to correspond to the principal components that the BPCA removes first, or forward PCA selects last. The tie breaking rule is utilized to decide the order of pruning these. We did not encounter tie situations for the more important components.

6. Numerical Analysis

In this section we will analyze two different data sets with tree structure. The first data set consists of branching structures of brain arteries belonging to 98 healthy subjects. An earlier version of this data set was used in [**4**] to illustrate the forward tree-line PCA ideas. In that study they have shown that a significant correlation exists between the branching structure of brain arteries and the age of subjects. Later on, 30 more subjects are added to that data set, and the set went through a data cleaning process described in [**5**]. In our study we will use this updated data set.

The second data set describes the organizational structure of a large company. The details of this data set are propriety information, therefore revealing details will be held back. We will investigate the organizational structural differences between business units, and differences between types of departments.

The Backward Algorithm used to conduct this analysis and produce visualizations is written in MATLAB. The software is available by request.

6.1. Visualization. To visually present all the principal components on a support tree of a data set, we present a visualization technique called *Radial PC Drawing*. The radial PC drawing of one of the subpopulations of the artery data is given in Figure 19, and the radial drawings of company data are given in Figure 24. (See [**7**] for details on relevant techniques for graph visualization.)

In radial drawing of rooted trees, the root node is at the origin. The root is surrounded by concentric circles centered at the origin. We plot our nodes on these circles, each circle is reserved for the nodes in one level of the tree. The coordinate of each node on a circle is determined by the number of descendants count. For example, for the nodes on the second level, the 360 degrees available on the circle is distributed to the nodes with respect to the number of descendants they have. Nodes with more descendants get more space. The nodes are put at the middle of the arc on the circle corresponding to the degrees set for that node. The children of that node in the next circle share these degrees according to their own number of descendants. This scheme allows the allocation of most space on the graph to the largest sub-trees and the distribution of nodes on the graph space as evenly as possible.

The principal components are expressed through the coloring scheme. A color scala starting from dark red, going through shades of light red, orange and yellow is used. The components that have higher sum of weights ($\sum w'(k)$) are colored with the shades on the red side, and lower sum of weights get the cooler shades. Since the backward principal components are ordered from low sum of weights $\sum w'(k)$ to higher, this means the earlier BPC's (lower impact components) are shown in yellow, while the stronger components are in red tones of the scala. The color bar on the right of the support tree shows which $\sum w'(k)$ corresponds to which shade.

6.2. Brain Artery Data Set.

6.2.1. **Data Description.** The data is extracted from Magnetic Resonance Angiography (MRA) images of 98 heathy subjects of both sexes, ranging from 18 to 72. This data can be found at [**14**]. [**6**] applied a tube tracking algorithm to construct $3D$ images of brain arteries from MRA images. See also [**11**] for further results on this set.

The artery system of the brain consists of 4 main systems, each feeding a different region of the brain. In Figure 18 they are indicated by different colors: gold for the Back, cyan for the Left, blue for the Right and red for the Front regions. The system feeding each of the regions are

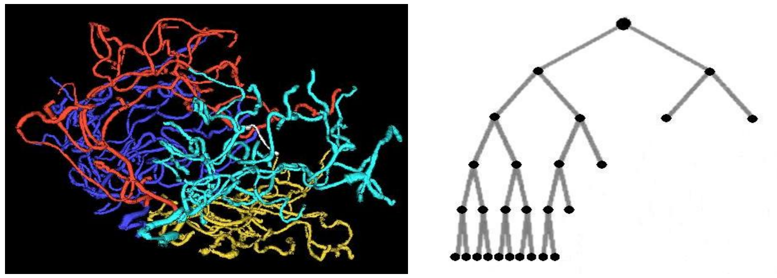


FIGURE 18. Left panel: Reconstructed set of trees of brain arteries. The colors indicate regions of the brain: Back (gold), Right (blue), Front (red), Left (cyan). Right panel: An example binary tree obtained from one of the regions. Only branching information is retained.

represented as trees, reduced from the $3D$ visuals seen in Figure 18. The reason for this is to focus on the branching structure only. Each node in a tree represents a vessel tube between two split points in the $3D$ representation. The two tubes formed by this split become the children nodes of the previous tube. The initial main artery that enters the brain, and feeds the region through its splits, constitutes the root node in the tree. The tree provided in Figure 18 (right panel) is an example tree extracted from a $3D$ image through this process.

The Back tree sets obtained after this process consist of 13634 nodes where the maximum depth observed is 37 levels. Other subpopulations have similar sizes.

The correspondence issue for this data set is solved as follows. At each split, the child with the higher number of nodes that descend from it is determined to be the left child, and the other node becomes the right child. This scheme is called descendant correspondence.

The study of brain artery structure is important in understanding how various factors affect this structure, and how they are related to certain diseases. The correlation between aging and branching structure was shown in previous studies ([**4**], [**11**]). The brain vessel structure is known to be affected by hypertension, atherosclerosis, retinal disease of prematurity, and with a variety of hereditary diseases. Furthermore, results of studying this structure may lead to establishing ways to help predict risk of vessel thrombosis and stroke. Another very important implication regards malignant brain tumors. These tumors are known to change and distort the artery structure around them, even at stages where they are too small to be detected by popular imaging techniques. Statistical methods that might differentiate these changes from normal structure may help earlier diagnoses. See [**10**] and the references therein for detailed medical studies focusing on these subjects.

6.2.2. **Analysis of Artery Data.** The forward tree-line PCA ideas were previously applied to an earlier version of this data set. Our first theoretical contribution of this thesis, extension of tree-line PCA to rooted labeled trees, does not affect this particular data set since all trees in it are binary. Therefore we focus on the dimension reduction approach. In [**4**], only first 10 principal components were computed, and the age effect was presented through the first 4 components. In general, the main philosophy of our dimension reduction or backward technique is to determine how many dimensions need to be removed for enough noise to get cleared from the data set before the statistical correlations become visible or significant. We ask this question for the brain artery data set and the effect of aging on it, on the updated brain artery data set. Also, [**4**] had used the intersection trees as the starting point in calculating the principal components. In this numerical study, we will use the root node as the starting point of the tree-lines.

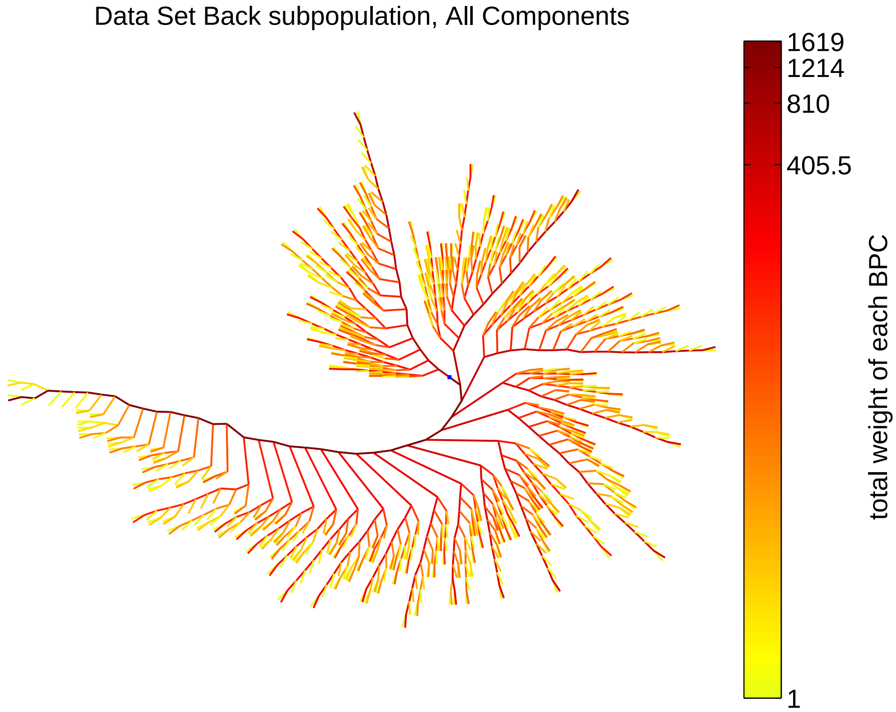


FIGURE 19. Radial PC drawings of the support tree of Back subpopulation. The root node is at the center. The principal components are represented through colors: Earlier BPC's start from the cold yellow end of the color scala while the latter BPC's go towards the red end. Nodes that are in multiple components are colored with respect to the highest total weighted component they are in. The color bar on the right of each panel shows the coloring scheme according to the total weight of each BPC.

The radial PC drawing of the Back subpopulation is given in Figure 19. Other subpopulations present similar radial PC drawings, so they are omitted here.

An observation on this data set, or any data set consisting of large trees is the abundance of leaves. Many of the leaves of the trees exist in one or few number of data trees. This leads to support trees that are much larger than any of the original data trees. The underlying structures are expected to be seen in upper levels, and most of the leaves can in fact be considered as noise. In our setting, the leaves that only exist in one or few data trees make up the first backward components. A question to ask is, what percentage of tree bodies is created by the low-weight leaves, and what percentage is due to the high-weight nodes, or underlying shape? Figure 20 provides two plots that illustrate an answer.

In Figure 20, the number of backward components removed from the tree space data is in, versus the total coverage explained by the remaining subspace is shown (left panel). The Y values at the $X = 0$ point correspond to the total coverage before any components are removed. This value is different for each subpopulation, as the sizes of their support trees are different. As backward components are removed from each of the sub-spaces, the coverage decreases. We can

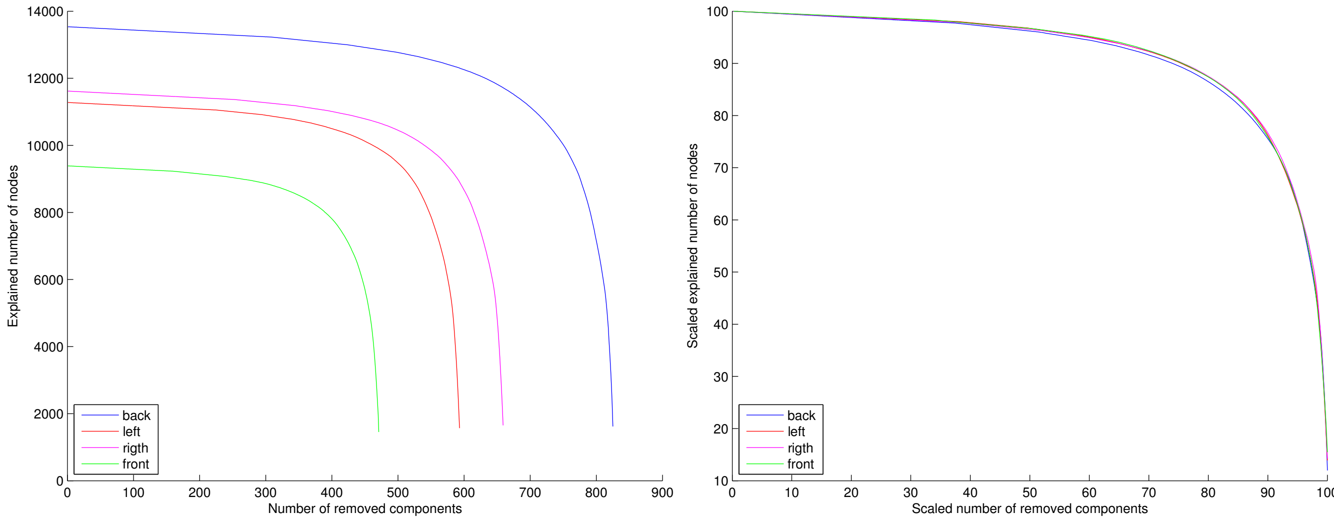


FIGURE 20. Left panel: X axis represents the total number of backward principal components removed from data. Y axis represents the number of nodes (coverage) explained by the remaining subspace after removal. Four subpopulations are shown: Back (blue), Left (red), Right (magenta), Front (green). Right panel: Same information as the left panel is used. For each subpopulation, the total coverage and the number of total backward principal components are scaled so that the maximum is 100.

observe that the initial backward components carry very little information, and therefore result in a very small drop in the coverage of the remaining sub-space. This is caused by the very large amount of leaves that aren't part of any underlying structure. The $Y = 0$ points for each of the curves mark the total number of principal components that cover the whole data. This number is in fact equal to the number of leaves on the support trees of each of the subpopulations.

On the right panel, we see the same information, only the X and Y axes for each of the curves are scaled so that the maximum corresponds to 100. The first observation we see in this graph is that the curves are almost plotted on top of each other: even if the sizes of their support trees are much different, the same percentage of coverage is explained by the same percentage of principal components in each of these data sets. We can conclude from this that the underlying structures are distributed similarly for each of these subpopulations. The second observation is that the majority of the principal components explain very little number of nodes. In the right panel of Figure 20, we see that for all the subpopulations, the first 80% of the principal components only cover 10% of the nodes, and the last 10% of these components explain about 80%. This data set is known to be very high-dimensional: over 800 components are needed to cover the Back sub-population. However, Figure 20 shows that a very small ratio of them are actually necessary to preserve the underlying structures.

Our next focus is to see, during the backward elimination process, at which points the age-structure correlation is visible.

It was established previously that the branching of brain arteries are reduced with age. [**10**] noted an observed trend on this phenomenon, while [**4**] showed this effect on Left subpopulation

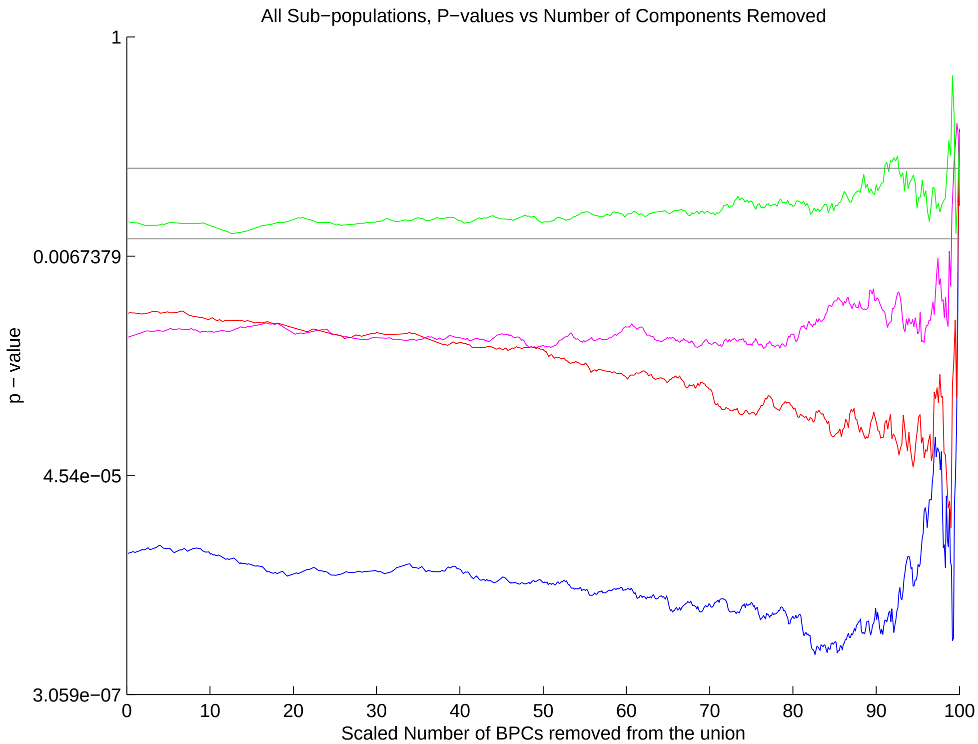


FIGURE 21. X axis represents the scaled number of backward principal components removed from the subspace of each of the subpopulations. At each X value, the data points are projected onto the remaining subspace. The sizes of these projections, plotted against age, show a downward trend (not shown here). Statistical significance of this downward trend is tested by calculating the standard linear regression p-value (Y axis) for the null hypothesis of 0 slope. Y axis is scaled using natural logarithm, while the Y axis ticks are given in original values. The grey horizontal lines indicate 0.05 and 0.01 p-value levels. The subpopulations are colored as: Back (blue), Left (red), Right (magenta), Front (green). A statistically significant age effect is observed for subpopulations Back, Left and Right.

using principal components. In this thesis, for each subpopulation, we start from the whole subspace and reduce it gradually by removing backward principal components. At each step the data trees are projected onto the remaining subspace. The relationship between the age of each data point and the norm of the data tree projection is explored by fitting a linear regression line to these two series. One example is provided in Figure 22. Others are omitted to save space, but they exibit very similar patterns.

This line tends to show a downward slope, suggesting that the projection sizes are reduced by age. To measure the statistical significance of the observation, a Student's t-test was applied to the data with the null hypothesis of 0 slope. Validity of t-test depends on several assumptions: Normality of the residuals, homoscedasticity and linearity. Normality and homoscedasticity was confirmed using appropriate tests (Kolmogorov-Smirnov and Breusch–Pagan tests). Linearly is assumed based on visual inspections. Note that the t-test results for the slope are known to be robust against deviations from linearity.

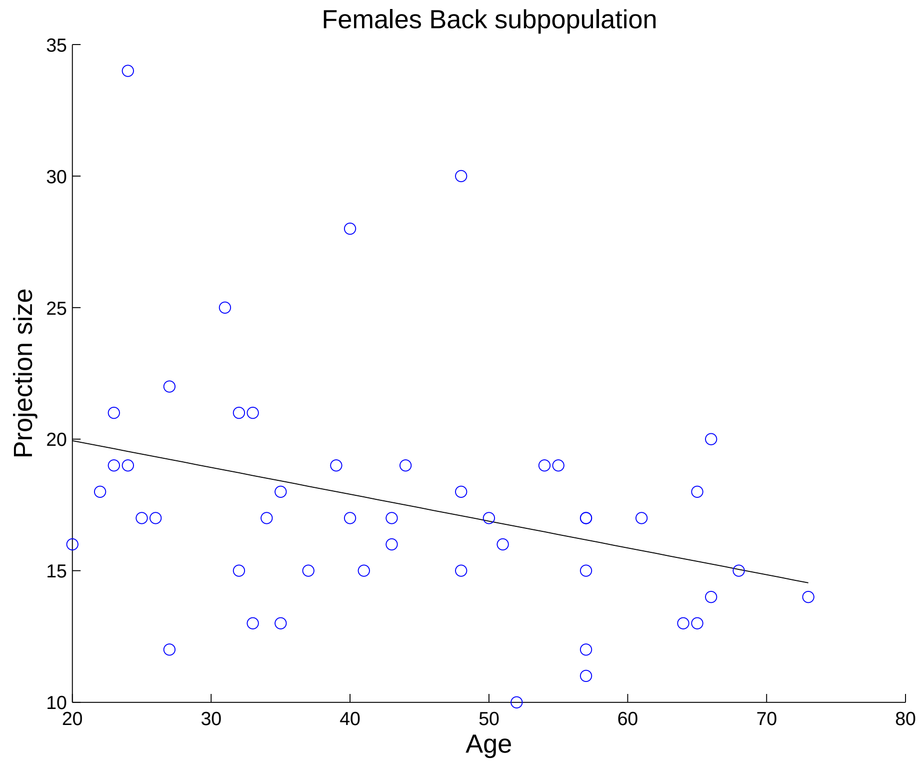


FIGURE 22. For the female sub-population, the size of the data projections (Y axis) onto the first PC. The X axis gives the age of each subject. The black solid line represents the linear regression line fitted to these points. The slope of this line is negative, implying the projection sizes may be reduced by age.

Figure 21 shows the plots of p-values at each step of removing BPC's, for each subpopulation. The p-values are scaled using natural logarithm while the Y axis ticks are left at their original values. A rule-of-thumb for the p-value is that 0.01 or less is considered significant. We will use this value to interpret our results. For a somewhat looser test, 0.05 can also be used. Figure 21 provides grey lines for both of these levels for reference.

In Figure 21 we see that, the Front subpopulation does not reach the p-value levels that are considered significant at any sub-space. The Front region of the brain, unlike the other regions, does not get fed by a direct artery entering the brain from below, but gets fed by vessels extending from other regions. (See Figure 18). Therefore it is not surprising that the Front vessel subpopulation does not carry a structural property presented by the other three subpopulations.

For other subpopulations, we identify two different kinds of age-structure dependence. First, for Left and Back subpopulations, the age versus projection size relationship is very sharp until the last 5% of the components are left. Most of the early BPC's correspond to the small artery splits that are abundant in younger population, which people tend to lose as age increases ([**10**]). Therefore the overall branchyness of the artery trees are reduced. Figure 21 is consistent with this previous observation. The p-value significance gets volatile at the last 5% of the components, where the BPC's corresponding to the small artery splits are removed, and only the largest components remain in the subspace. These largest components correspond to the main arteries that branch the most. The location-specific relationship between structure and age, noted in [**4**] can be observed

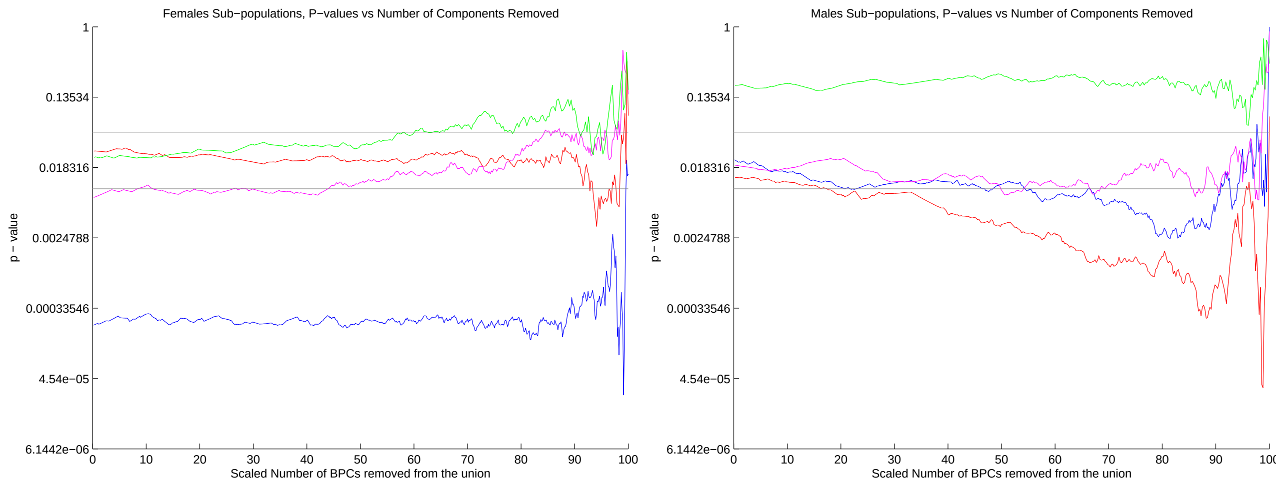


FIGURE 23. The left and right panels are the p-value versus subspace plots for female and male populations. The axes are as explained in Figure 21. The subpopulations are colored as: Back (blue), Left (red), Right (magenta), Front (green). For males, a statistically significant age effect is observed for subpopulations Back, Left and Right. No such effect is observed for females.

for Left, Right and Back subpopulations towards the end of the X axis. This is the second kind of dependence we observe in the data sets.

Our second focus is to repeat the question of age-structure relationship for the male and female subpopulations. Our data set consists of 49 male, 47 female and 2 trans-gender subjects. We run our analysis for the largest two groups to see how aging effects males and females separately. In Figure 23, the p-value versus subspace graphs are given for the male and female subpopulations. As before, the Front subpopulation does not show any statistical significance.

For the female group, the first kind of structural affect of age (overall branchyness) can be observed for Back and Right. For the location-specific relationship (branchyness of the main arteries), Back and Left subpopulations present significant p-values.

For the male group, the age versus overall branchyness is not significant at 0.01 p-value level, although the Back, Right and Left subpopulations are very close. The location-specific relationship can again be observed for these three subpopulations at significant levels.

The study on the full data set implies that two kinds of age-structure relationships can be observed in the whole population using this method. Subsequent analysis of male and female groups shows that the overall branchyness effect is observed more strongly in the male group. These results suggest that the brain vessel anatomy of male and females may respond slightly differently to aging.

6.3. Company Organization Data Set.

6.3.1. **Data Description.** In this analysis, we use a company organization data set of a large US company. This data set is a snapshot of the employee list taken sometime during the last ten years. It also includes the information on hierarchical structure and the organizations that employees belong to. The set includes more than two hundred thousand employees active at the time when the snapshot was taken. In this section we will explain the general aspects of the data set that are relevant to our analysis, but we will hold back any specifics due to privacy reasons.

The original company structure can be considered as one giant tree. Each employee is represented as a node. The CEO of the company is the root node. The child-parent relationships are established through the reporting structure: the children of a node are the employees that directly report to that person in the company. Since every employee directly reports to exactly one person (except the CEO, the root node), this system naturally lends itself to a tree representation. Moreover, this organization tree is not binary, but a general rooted tree. It has a maximum depth of 13 levels.

In this study we will focus on populations of different departments across the company that are assigned to a similar type of job. When the whole organization tree is considered, the directors of these departments are at the fifth level of that tree. The upper levels of the main tree represent organization divisions of the company based on main business activity, geographical locations around the world, etc.

To form our data set, we gathered the list of all the directors in the company who are at the fifth level. Then, based on the organization codes, we determined the main job focus of the departments that the directors are leading. We selected four main groups of jobs to compare for our study: finance, marketing, R&D, and sales. Other departments that focus on different jobs, like legal affairs or IT support, are left out. For each category, each department assigned to that category forms one data point. The director of that department is taken as the root node of the data tree representing the department, and the people who work at that department are nodes of this tree. The structure of the tree is determined by the reporting structure within the department.

The correspondence issue within the data sets requires some attention. A job-based correspondence scheme between two data trees would involve determining which individuals in one department perform a similar function to which individuals at the same reporting level in another department. With the exception of the directors (who form the root nodes and naturally correspond to each other), this kind of matching is virtually impossible for this data set. For this reason, we employ the descendant correspondence for the data points.

The data set of the finance departments constructed in this fashion consists of 37 data trees, with a maximum depth of 6 levels. The marketing set has 60 trees, maximum depth of 5, sales has 41 trees, maximum depth 5, and R&D data set has 20 trees, maximum depth 6. The support trees of these sets can be seen in Figure 24.

To visualize the company organization data, we will again employ radial PC drawing method. The depth of these trees is not very large: 6 levels for the deepest data point. However, the node population at each level is very dense.

6.3.2. **Analysis of Company Organization Data.** The comparative structural analysis of these four organization data sets is conducted via the principal component tree-lines. We have run the dimension reduction method for general rooted trees as described in Section 4, although the forward method of Section 3 would have given the same set of components, as shown in Section 4. The principal components obtained with this analysis are shown in Figure 24. The first conclusions on the differences across types of departments come from the comparison of their support tree structure. It can be clearly seen that the sales departments are larger than others in population. Another clear distinction is in the flatness of each organization type. Typically, a flat organization does not have many levels of hierarchy, and most of the workers do not have subordinates. This is common in organizations of a technical focus. In Figure 24, we can see that the R&D departments are visibly flatter than other three types: most of the nodes are at the leaves and not at the interim levels. This is due to the fact that most of the employees in these departments do engineering-research type of work, for which a strongly hierarchical organizational model is less efficient. The other three data sets, finance, marketing and sales have most of their employees on interim levels, pointing to a strong hierarchy. This seems especially strong in sales departments.

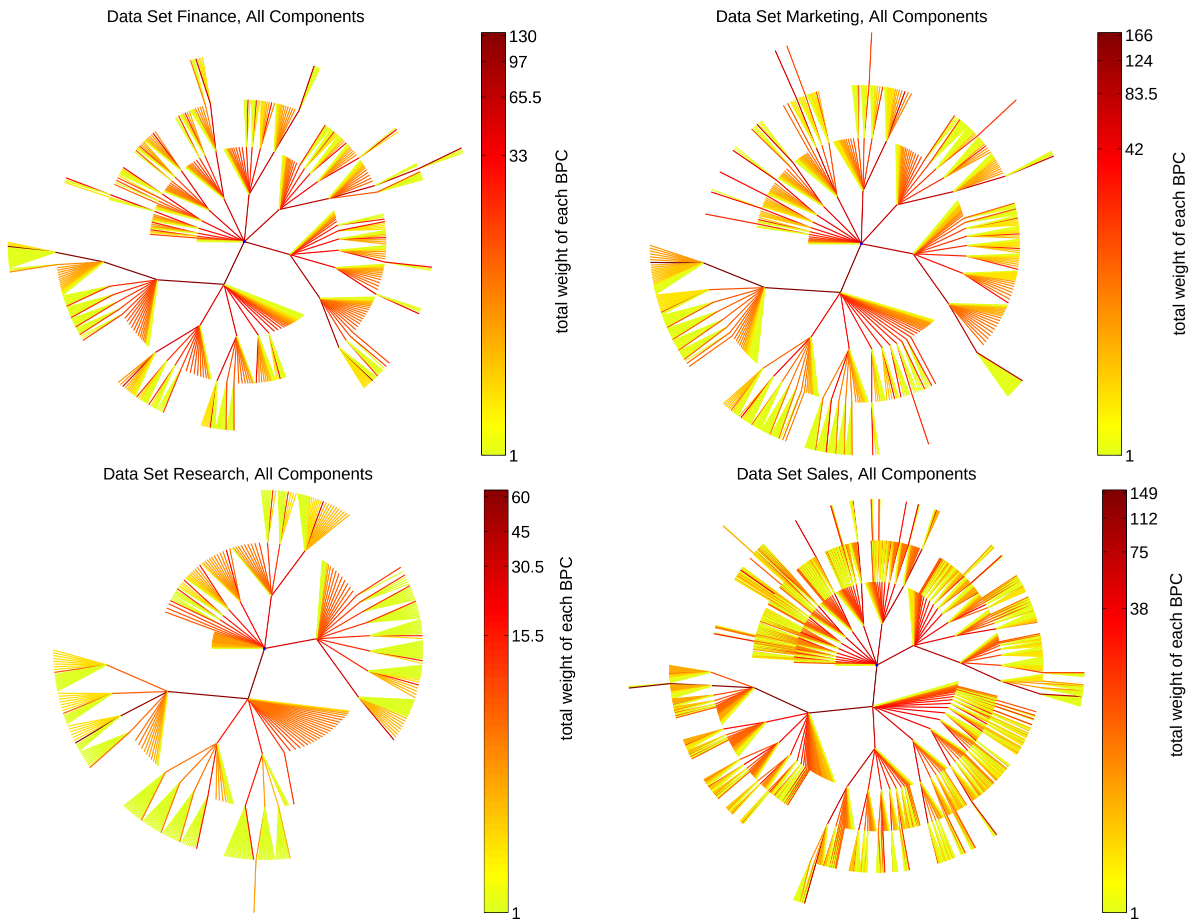


FIGURE 24. Radial PC drawings of the support trees of four organization subsets: Finance, Marketing, R&D and Sales. The root nodes are at the center. The principal components are represented through colors: Earlier BPC's start from the blue end of the color scala while the latter BPC's go towards the red end. Nodes that are in multiple components are colored with respect to the highest total weighted component they are in. The color bar on the right of each panel shows the coloring scheme according to the total weight of each BPC.

In the next figure (Figure 25), the effect of reducing the principal components gradually on the amount of nodes explained is shown. This figure is constructed in the same way as Figure 20, right panel. Figure 25 shows that none of the organization data sets have a very concave coverage-versus-components curve like the brain artery set did. Therefore for the organizational structure setting, the earlier BPC's have more potential to carry information compared to the artery setting. Between the organization data sets, we see that the curves belonging to R&D and sales are very close to each other (the less concave pair), while the curves of finance and marketing are shape-wise close (the more concave pair). The concavity of these curves depend on what percentage of the

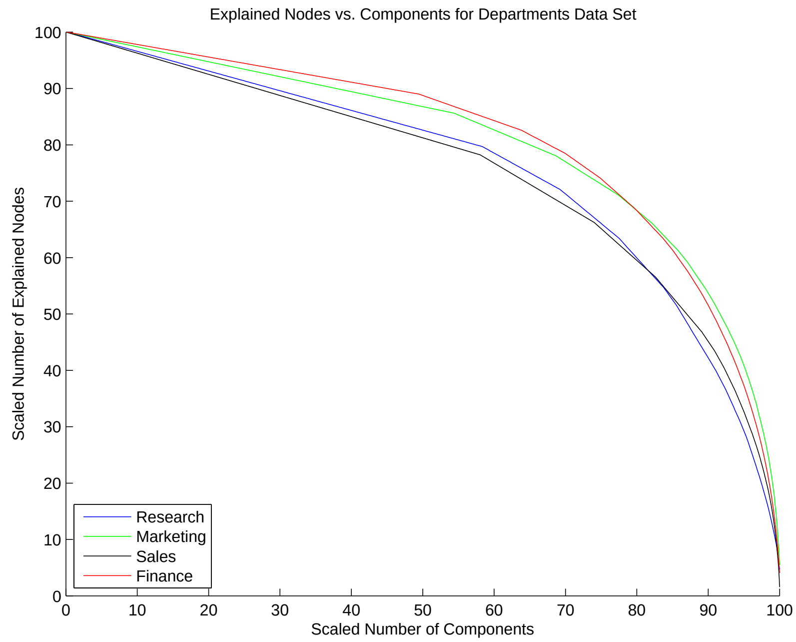


FIGURE 25. The X axis is the number of backward principal components subtracted from the subspace. The Y axis is the amount of nodes that can be explained by the remaining subspace at each X level. Both of the axes scaled within themselves so that the highest X and Y coordinates for all of the organization curves are 100. The blue curve is for R&D, green is for marketing, black is for sales and red is for finance.

structure is explained by the early BPC's, and what percentage by the later, stronger components. A very concave curve means that most of the nodes of the data set can in fact be expressed through a small number of principal components. This means that the structures within the data points are not very diverse: the data trees of the set structurally look like each other, allowing a smaller number of PC's to explain more of the nodes. Vice versa, a less concave curve points to a data set where a small portion of the principal components are not enough to explain many nodes due to the diversity in the structures of the data points. Figure 25 shows that finance and marketing departments are more uniformly structured than R&D and sales departments, i.e., two random finance data trees are more likely to have a shorter distance to each other than two random R&D data trees.

A coverage-versus-components curve is helpful in establishing the trend in the distribution of variability within the data set: the earlier BPC's express nodes that are not common across the data points, and the later BPC's cover the nodes that are common to most data points. The next, and more in-depth question is: How are these more common and less common nodes distributed among the data points themselves? To answer this question, we divide the set of all BPC's into two subsets. The first 90% of the BPC's on the X axis of Figure 25 form the one set (SET 2). These BPC's collectively represent the subspace where the less-common-nodes are in. The remaining 10% of the BPC's form the other set (SET 1). These BPC's express the subspace where the more common structures are in. For any data tree t, the projection of it onto SET 1 ($P_{SET1}(t)$) represents the portion of the tree that is more common with other data trees in the data set. The projection of t onto SET 2 ($P_{SET2}(t)$) carries the nodes of it that are less common with others.

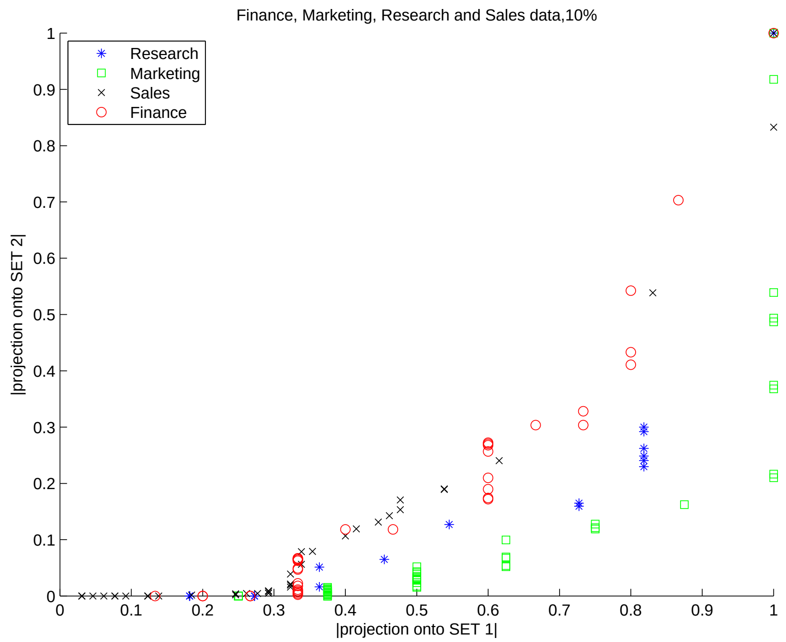


FIGURE 26. The data points of each of the data sets: R&D (blue stars), marketing (green squares), sales (black crosses) and finance (red circles). For each of the data points, the length of its projection onto SET 2 is on the Y axis, and the length of its projection onto SET 1 is given on the X axis. Each of these axes are scaled such that the highest coordinate for each data set is 1 on each of the axes.

Since these two sets are complementary, the two projections of t would give t itself when combined: $P_{SET1}(t) \bigcup P_{SET2}(t) = t$.

Figure 26 shows how the nodes in SET 1 and 2 are distributed among the data trees for each of the organization data sets. For each data point, the length of its projection onto SET 2 is on the Y axis, and the length of its projection onto SET 1 is given on the X axis. Each of these axes are scaled such that the highest coordinate for each data set is 1 on each of the axes. Blue stars denote the R&D data points, green squares are marketing data points, black crosses are sales data points and red circles are finance data points. In Figure 26, it can be seen that none of the data points are above the 45 degree line. This is an artifact of the descendant correspondence.

A very interesting aspect of Figure 26 is that, the data points of each data set visually separate from each other. This is especially true for the marketing departments which follow a distinctly more convex pattern compared to other kinds of departments.

For finance departments, we observe an almost linear trend, starting from around $X = 0.3$. The bottom left data points are trees that are small in general: they contain little of the common nodes set and almost none of the non-common set. As we go top-right, the trees grow in SET 1 and SET 2 spaces proportionally. A similar pattern is there for sales departments, with the exception of a group of data points lying on the X axis, pointing to a group of very small departments that only consist of the main structure nodes. The R&D departments follow a lower angle pattern. However, this might be due to the one outlier department at the coordinate $(1, 1)$, pushing all others to the left/bottom of the graph.

The most significant pattern on this graph belongs to the marketing group. Unlike other departments, there is no linear alignment trend. The set seemingly consists of two kinds of departments: First is the group with very little projection on SET 2, and varying sizes of projection on SET 1. These are relatively small departments. The second is a group of departments that contain all the nodes represented by SET 1 (therefore the "common structure" part of the trees are common to all of these trees), and varying, but large amounts of nodes represented in SET 2. These trees are much larger than the first trees of the group. These two different modes of structure within this group may be due to the particular kind of marketing activity, product family, etc they focus on. The details of activities of each department is not part of our data set, therefore we are not able to offer a reason for this separation. Note that two data sets that are shown to be structurally similar in Figure 25, finance and marketing, are the furthest apart sets in Figure 26. This is because Figure 25 focuses on the overall dispersion of coverage, while Figure 26 focuses on the relative differences between the individual data trees.

Bibliography

[1] C.A. Alfaro, B. Aydın, C.E. Valencia, E. Bullitt, A. Ladha, Dimension Reduction in Principal Component Analysis for Trees, Computational Statistics and data analysis 74:157–179, 2014.

[2] B. Aydin, K. Guler, C.A. Alfaro, C.E. Valencia, Data tree reduction, Patent application. Patent number: US 13/560,435.

[3] Allen B.L. and Steel M., Subtree transfer operations and their induced metrics on evolutionary trees. Ann. Comb., 5:1–15, 2001.

[4] B. Aydın, G. Pataki, H. Wang, E. Bullitt. and J.S. Marron, A Principal Component Analysis For Trees, Annals of Applied Statistics **4** vol. 3 1597-1615 (2009)

[5] B. Aydın, G. Pataki, H. Wang, A. Ladha, E. Bullitt and J.S. Marron, Visualizing the Structure of Large Trees. Electronic Journal of Statistics, 5, 405-420 (2011)

[6] S. Aylward and E. Bullitt, Volume Rendering of Segmented Image Objects. IEEE Trans. Medical Imaging, 21:998-1002 (2002)

[7] G.D. Battista, P. Eades, and R. Tamassia, I.G. Tollis, Graph Drawing Algorithms for the Visualization of Graphs. Prentice Hall, Upper Saddle River, NJ. (1999)

[8] P. Bille, A survey on tree edit distance and related problems. Theoretical Computer Science, 337, 217-239 (2005)

[9] L.J. Billera, S.P. Holmes and K. Vogtmann, Geometry of the Space of Phylogenetic Trees. Adv. in Appl. Math., 27:733-767 (2001)

[10] E. Bullitt, G. Gerig, S.M. Pizer and S.R. Aylward, Measuring Tortuosity of the Intracerebral Vasculature from MRA Images. IEEE-TMI 22:1163-1171 (2003)

[11] E. Bullitt, D. Zeng, A. Ghosh, S.R. Aylward, W. Lin, B.L. Marks, K. Smith, The Effects of Healthy Aging on Intracerebral Blood Vessels Visualized by Magnetic Resonance Angiography, Neurobiology of Aging, 31(2):290300 (2010)

[12] G. Chartrand and L. Lesniak, Graphs and digraphs, Third Edition, Chapman & Hall/CRC (2000)

[13] A. Feragen, P. Lo, M. de Bruijne, M. Nielsen and F. Lauze, Towards a theory of statistical tree-shape analysis, IEEE Trans Pattern Anal Mach Intell, to appear.

[14] Handle, http://hdl.handle.net/1926/594 (2008)

[15] J. Hein, Reconstructing evolution of sequences subject to recombination using parsimony. Math. Biosci., 98:185–200, 1990.

[16] H. Heumann and G. Wittum, The tree-edit-distance, a measure for quantifying neuronal morphology. Neuroinform, 7:179-190 (2009)

[17] H. Hotelling, Analysis of a Complex of Statistical Variables Into Principal Components. Journal of Educational Psychology, 24:417-441,498–520 (1933)

[18] I. Jolliffe, Principal Component Analysis, Second Edition, Springer (2002)

[19] S. Jung, X. Liu, J.S. Marron and S.M. Pizer, Generalized PCA via the Backward Stepwise Approach in Image Analysis, Brain, Body and Machine, Proceedings on an International Symposium on the Occasion of the 25th Anniversary of the McGill Centre for Intelligent Machines, Montreal, (J. Angeles, et al., eds.), Springer, New York, 111-123 (2010)

[20] K.V. Mardia, J.T. Kent and J.M. Bibby, Multivariate Analysis, Academic Press (1973)

[21] J.S. Marron, S. Jung and I.L. Dryden, Speculation on the Generality of the Backward Stepwise View of PCA, Proceedings of MIR 2010: 11th ACM SIGMM International Conference on Multimedia Information Retrieval, Association for Computing Machinery, Inc., Danvers, MA, 227-230 (2010)

[22] E. Miller, M. Owen and J.S. Provan, Averaging metric phylogenetic trees, preprint.

[23] K. Pearson, On Lines and Planes of Closest Fit to Systems of Points in Space, Philosophical Magazine 2 (6): 559–572 (1901)

[24] M. Owen and J. Provan, A fast algorithm for computing geodesic distances in tree space, Transactions on Computational Biology and Bioinformatics, vol. 8, pp. 2-13, 2011.

[25] D.F. Robinson, Comparison of labeled trees with valency three. J. Combinatorial Theory, 11:105–119, 1971.

[26] D.F. Robinson and L.R. Foulds, Comparison of phylogenetic trees. Math. Biosci., 53:131–147, 1981.

[27] J. Tschirren, K. Palágyi, J.M. Reinhardt, E.A. Hoffman and M. Sonka, Segmentation, Skeletonization and Branchpoint Matching a Fully Automated Quantitative Evaluation of Human Intrathoracic Airway Trees. Proc. Fifth International Conterence on Medical Image Computing and Computer-Assisted Intervention, Part II. Lecture Notes in Comput. Sci., 2489:12-19 (2002)

[28] H. Wang and J.S. Marron, Object Oriented Data Analysis: Sets of Trees, Annals of Statistics 35 1847–1873 (2007)

Index