UNIVERSITAT DE BARCELONA
Facultat de Matemàtiques
i Informàtica

GRAU DE MATEMÀTIQUES

Treball final de grau

The role of poorly connected layers in multiplex networks: suitability and impact for the study of human past

Autor: Clara Galceran Puig

Director: Dr. Xavier Jarque Ribera
Realitzat a: Departament de Matemàtiques i Informàtica

Barcelona, 15 de gener de 2025

Contents

Abstract

This work explores the theoretical and practical implications of poorly connected layers in multiplex networks, advancing their mathematical formalization and application in historical research. Graph theory concepts are extended to the multilayer network framework, focusing on structural transitions in two-layer multiplex networks and the role of poorly connected layers in enhancing overall connectivity. This framework is applied to analyze the evolution of ancient road and river networks in Southern Etruria and Latium Vetus, showcasing the value of the multiplex formalism in interpreting complex historical phenomena.

Key contributions include new proofs related to the algebraic connectivity of two-layer multiplex networks and a new proposition of the definition of the Laplacian matrix for supra walks, providing a basis for further research into these type of networks.

Resum

Aquest treball explora les implicacions teòriques i pràctiques de la presència de capes pobrament connectades en xarxes multiplex, avançant en la seva formalització matemàtica i aplicació en la recerca històrica. Conceptes de teoria de grafs són estesos al marc de les xarxes multicapa, centrant-se en transicions estructurals en xarxes multiplex de dues capes i el rol de capes mal connectades en millorar la connectivitat total de la xarxa. Aquest marc de treball és aplicat per analitzar l'evolució d'antigues xarxes de camins i rius del Sud d'Etrúria i el Laci, mostrant així el valor del formalisme de les xarxes multiplex en interpretar fenòmens històrics complexos.

S'aporten noves proves relacionades amb la connectivitat algebraica de xarxes multiplex de dues capes i una nova proposta per la definició de la matriu Laplaciana per supra-camins, obrint la porta a futura recerca relacionada amb aquesta tipologia de xarxes.

Acknowledgments

I would like to thank Emanuele Cozzo for his guidance on working with multiplex networks, Luce Prignano for her advice on applying the multiplex formalism to ancient transportation networks and Xavier Jarque for his support on the overall redaction of this work.

I would also like to thank my family, partner and friends for their unconditional support during the years of my mathematics degree, which have culminated in this work.

2020 Mathematics Subject Classification. 05C40, 05C50, 05C82

Introduction

The natural and man-made world exhibit numerous of dynamics and phenomena rooted on the relationships among individuals that conform them. In order to analyze these relationships, a network framework serves as a powerful tool, offering a comprehensive understanding of the complexity of the relationships and providing insights into phenomena that may be challenging to understand if other methods are used. This approach offers us an unified methodology for examining vastly different types of relationships [1], ranging from communication networks [2] to biological processes [3].

The mathematical study of networks began in 1736 with Leonhard Euler's work on the Seven Bridges of Köningsberg problem, which marked the bases of graph theory. However, the application of these mathematical models to analyze the dynamics and relationships of the world around us originated in the early 20th century within sociology, through the works of Émile Durkheim and Georg Simmel [4]. While these initial studies adopted a largely theoretical approach on the matter, further developments made throughout the 20th century provided enough analytical methods to enable a systematic use of social network analysis. This field experienced significant growth in the 1990s with the involvement of physicists, computer scientists, economists and political scientists, the contributions of whom introduced new models and data analysis techniques, expanding the applicability of network studies beyond theoretical mathematics to a broader range of disciplines.

One of the advancements that contributed to the growth of the field is the development of the multilayer network framework [5]. This approach extends the traditional framework, which typically considers only a single type of relationship among the components of a system, by incorporating multiple types of relationships through distinct layers. In this way, each layer captures a specific type of interaction among the system's components, allowing for the study of both intra-layer connections (within a single layer) and inter-layer connections (across different layers). This framework enables a more comprehensive analysis of systems influenced by diverse aspects and relationships, such as the network of different modes of communications (considering relationships via e-mail or in-person) within different social groups (considering familial, friendship or professional relationships).

The specific constraints placed on intra-layer and inter-layer connections lead to various types of multilayer networks. Among these, a particularly valuable type is the *multiplex network*, which focuses on a single aspect while restricting inter-layer connections to objects that represent the same entity across different layers. These constraints make multiplex networks especially suitable for studying systems such as transport networks [6] or the dynamics of disease spread [7].

Within multiplex networks, the inclusion of a poorly connected layer alongside a layer with stronger connection properties presents an interesting area of study. While the addition of a weakly connected layer might initially appear to reduce the robustness of the system, it can, in fact, enhance the overall network functionality in cases where the weaker layer captures connections that provide significant benefits to the overall network. The theoretical analysis of this phenomenon is specially valuable when studying dynamics related to the maintenance of infrastructures that facilitate communication and movement. This work aims to explore the suitability and impact of maintaining poorly connected layers within multiplex networks, with a final specific focus in their application to the viability and likelihood of maintaining both road and river transportation networks in ancient regional transport infrastructures.

To achieve this, Chapter 1 introduces the basic concepts of graph theory necessary for analyzing single-layer networks, with a particular focus on their matrix representation, including the *adjacency matrix* and *Laplacian matrix*, and their role in studying connectivity. Connectivity is presented as a metric to quantify the network's robustness, with a focus on *algebraic connectivity* as a lower bound of the network's connectivity and as an indicator of the behavior of diffusion models within the network, derived from its Laplacian matrix. Moreover, the chapter introduces the concept of *network quotient* as a method for simplifying a network by coarsening it according to a node partition. The resulting quotient network preserves spectral properties of the original network, a feature that will later prove essential for deriving results related to the maintenance of poorly connected layers.

Chapter 2 extends the framework to multilayer networks, generalizing concepts and results from Chapter 1 to systems that operate across different types of relationships. The *set of layers* is introduced to encapsulate the various type of aspects the network operates on, along with potential constraints that can be applied to multilayer networks. The quotient results from the previous chapter are then applied to this framework by considering natural node partitions that arise from considering either individual layers or the entities that represent nodes across layers. These different approaches result in the derivation of the *network of layers* and the *aggregate network*, respectively.

From the constraints discussed in the previous chapter, Chapter 3 introduces multiplex networks, a specific type of multilayer network where inter-layer connections are restricted to objects that represent the same entity across different layers. This chapter examines structural transitions in two-layer multiplex networks, using algebraic connectivity as an indicator. By varying the *inter-layer* *weight p*, two main behavior regions are identified: one where the layers act as disconnected networks and another where the system behaves as an unified network, with algebraic connectivity bounded above by $2p$ and by that of the average network. The analysis highlights how the poorer connected layer determines the multiplex's overall behavior and connectivity. To account for the potential value of the poorly connected layer, a *weighting factor* α is introduced to scale the importance of its connections, studying how algebraic connectivity changes with α and characterizing the regions in the parameter space (α, p) where the multiplex outperforms the well-connected layer. Lastly, the chapter introduces *supra-walks*, which impose restrictions on inter-layer steps, thus bringing the need for a redefinition of the Laplacian matrix that defines the system.

The theoretical results obtained in Chapter 3 are applied in Chapter 4, where the multiplex framework is used to study the transportation infrastructures of archaeological sites with different connectivity characteristics in Southern Etruria and Latium Vetus. This analysis examines how the viability of maintaining both road and river networks evolves over time and across regions, providing insights into how the development of a robust road infrastructure gradually detrimentally affects the importance placed on the river infrastructures in the articulation of the territory.

This work extends the understanding of multiplex network dynamics in the presence of poorly connected layers by formalizing and presenting new proofs for the algebraic connectivity of a two-layer multiplex network as a function of its inter-layer weight p and weighting factor α associated with the poorly connected layer. Additionally, it proposes how to calculate the bound for α to characterize the region in the (α, p) space where the multiplex outperforms the well-connected layer. Furthermore, the work proposes a new definition of the Laplacian matrix that accounts for the constraints introduced by supra-walks.

On the practical side, this study shows that the multiplex network formalism can serve as an effective tool for analyzing ancient transport infrastructures, highlighting the broader applicability of the multiplex framework to fields such as archaeology, history and anthropology, where understanding the interplay between different types of connections is essential.

Natural directions for future work include a deeper analysis of the suitability of the proposed Laplacian matrix for supra-walks in multiplex networks, and the replication the algebraic connectivity results derived in this study using the new Laplacian definition. Moreover, it would be valuable to extend the multiplex formalism to additional network metrics, such as global efficiency, to further characterize these networks.

Chapter 1

Networks

To stablish a solid mathematical foundation for studying phenomena and systems characterized by the relationships among its constituents, this first chapter provides a brief introduction to the main definitions and results of complex network theory. These definitions and results will serve as an essential tool in the following chapters for obtaining significant insights when examining specific types of real-world networks.

Accordingly, it is important to begin by defining the primary mathematical object from which we will work on: the graph, as this object allows us to easily study the relationships among the different elements of a system. A graph is often physically represented by a network, where its components correspond to tangible entities. In this work, the terms "graph" and "network" are used interchangeably to refer to the same object, as a matter of language convenience.

Definition 1.1. A *network* (or *graph*) $\mathcal{G} = (V, E)$ is a mathematical object consisting of a set of nodes (or vertices) $V = \{v_i\}_i$ and a set of edges $E \subseteq \{(v_i, v_j) | v_i, v_j \in V\}$, where (v_i, v_j) represents a connection between nodes v_i and v_j.

If there exists an edge in $\mathcal{G}$ between two nodes u and v, i.e., $(u, v) \in E$, the two nodes are said to be *adjacent*. Likewise, an edge is said to be *incident* to a node if it starts or ends at that node.

Various type of networks arise from this basic definition, depending on the characteristics of their edges and connections.

- A network is called *directed* if the edges are ordered pairs: (v_i, v_j) denotes an edge from v_i to v_j.

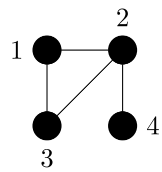
- A network is called *undirected* if the edges are unordered pairs: $(v_i, v_j) = (v_j, v_i)$.

- A *weighted network* is a network in which each edge is assigned a weight, that is, each edge is mapped to a real number by a function $w : E \to \mathbb{R}$.

Note that two nodes can be connected by more than one edge; in this case, the collection of these edges is referred to as a *multiedge*. Furthermore, an edge may connect a node to itself, in which case it is called *self-edge* or *self-loop*. When a network has neither self-edges nor mutiedges, it is called a *simple network*.

Unless otherwise stated, this work will only consider simple networks.

Example 1.2. Consider the following network.



It is a simple undirected unweighted network with a node set $V = \{1, 2, 3, 4\}$ and an edge set $E = \{(1, 2), (1, 3), (2, 3), (2, 4)\}$.

To characterize certain structural properties of a network related to its robustness, it is helpful to introduce two main concepts regarding the structure's ability to link its different components.

Definition 1.3. Let $\mathcal{G} = (V, E)$ be a network. A *path* is a sequence of non-repeated nodes $P = \{v_1, v_2, \dots, v_k\}$ where $v_i \in V$ and $(v_i, v_{i+1}) \in E$ for $i = 1 \dots k - 1$. If the sequence allows repeated nodes, it is called a *walk*. The elementary component of a walk is an edge, also referred to as a *step*, which connects two adjacent nodes.

This definition allows us to classify networks based on wether there exists a path between every pair of nodes. This distinction motivates the following definition.

Definition 1.4. Let $\mathcal{G} = (V, E)$ be a network. $\mathcal{G}$ is *connected* if, for all $u, v \in V$, there exists a path from u to v. If $\mathcal{G}$ is not connected, it is called a *disconnected* network.

Note that the network in Example 1.2 is a connected network, and a possible path from $u = 1$ to $v = 4$ is $P = \{1, 3, 2, 4\}$.

1.1 Matrix representation

When studying the structural properties of a network, working solely with edge and node sets can be somewhat cumbersome. A more effective way to analyze networks is to represent the set of nodes and edges in a matrix format using the adjacency matrix as a way to encapsulate the entire structure of the network.

Definition 1.5. The *adjacency matrix* of a network $\mathcal{G}$ on n nodes is an $n \times n$ matrix, $A_{\mathcal{G}} = (a_{ij})$, where

- $a_{ij} \neq 0$ if there is an edge between v_i and v_j of weight a_{ij}.
- $a_{ij} = 0$ if there is no edge between v_i and v_j.

For an unweighted network, the entries of the matrix are binary: $a_{ij} = 1$ indicating the presence of an edge between v_i and v_j and $a_{ij} = 0$ the absence of an edge.

Note that for undirected networks the adjacency matrix is symmetrical, and for networks without self-loops all the diagonal elements on the adjacency matrix are zero.

Example 1.6. The adjacency matrix of the network used in Example 1.2 is

$$A_{\mathcal{G}} = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

While the adjacency matrix contains all the structural information about a network, a new matrix called the *Laplacian matrix* is also convenient for describing undirected networks.

Definition 1.7. For an undirected network, $\mathcal{G}$ with n nodes, that satisfies $a_{ij} \geq 0$, the *Laplacian matrix* is defined as $L_{\mathcal{G}} = D - A_{\mathcal{G}}$. Where $D = \text{diag}(d_1, \ldots, d_n)$ is the diagonal matrix of the node degrees, with $d_i = \sum_j a_{ij}$.

The Laplacian matrix arises naturally in the study of diffusion processes (i.e., simple models of spread across a network) on undirected networks [8, Section 6.13].

Consider a condition of some kind affecting the nodes of a network $\mathcal{G} = (V, E)$, where each node v_i is assigned an amount ϕ_i. Suppose this condition spreads along the edges, flowing from node j to an adjacent node i at a rate proportional to $C(\phi_j - \phi_i)$, where C is a constant known as the *diffusion constant*. The rate of change of ϕ_i is given by

$$\frac{d\phi_i}{dt} = C \sum_j a_{ij}(\phi_j - \phi_i). \tag{1.1}$$

Here a_{ij} are the elements of the adjacency matrix $A_{\mathcal{G}}$, ensuring that only terms corresponding to node pairs that are actually connected by an edge appear in the sum.

By splitting the terms in (1.1), we obtain

$$\frac{d\phi_i}{dt} = C\sum_j a_{ij}\phi_j - C\phi_i \sum_j a_{ij} = C\sum_j a_{ij}\phi_j - C\phi_i d_i = C\sum_j (a_{ij} - \delta_{ij} d_i)\phi_j \tag{1.2}$$

where d_i is the node degree of v_i, and δ_{ij} is the Kronecker delta (which is 1 if $i = j$ and 0 otherwise).

Equation (1.2) can be written in matrix form as:

$$\frac{d\phi}{dt} = C(A_{\mathcal{G}} - D)\phi \tag{1.3}$$

with $D = \text{diag}(d_1, ..., d_n)$ being the diagonal matrix of the node degrees.

Using Definition 1.7, Equation (1.3) takes the form:

$$\frac{d\phi}{dt} + CL_{\mathcal{G}}\phi = 0 \tag{1.4}$$

which has the same form as the ordinary diffusion equation for a gas*, where the Laplacian operator ∇^2 is replaced by the matrix $L_{\mathcal{G}}$, which can be viewed as a discretized version of the Laplacian operator.

Example 1.8. The Laplacian of the network used in Example 1.2 is

$$L_{\mathcal{G}} = \begin{bmatrix} 2 & -1 & -1 & 0 \\ -1 & 3 & -1 & -1 \\ -1 & -1 & 2 & 0 \\ 0 & -1 & 0 & 1 \end{bmatrix}$$

The set of eigenvalues of $L_{\mathcal{G}}$ provides crucial information on the stationary behavior of the system. Taking into account that all eigenvalues of $L_{\mathcal{G}}$ are positive, except the smallest, which is 0 [9]:

The diffusion equation 1.4 can be solved by expressing the vector ϕ as a linear combination of the Laplacian matrix's eigenvectors $\mathbf{v}_i$

$$\phi(t) = \sum_i a_i(t)\mathbf{v}_i \tag{1.5}$$

with the coefficients a_i evolve over time. Substituting (1.5) in (1.4) and using $L\mathbf{v}_i = \lambda_i \mathbf{v}_i$ (with λ_i as the eigenvalue corresponding to the eigenvector $\mathbf{v}_i$) we get:

$$\sum_i \left(\frac{da_i}{dt} + C\lambda_i a_i\right)\mathbf{v}_i = 0. \tag{1.6}$$

*The ordinary diffusion equation of a gas with a density $\phi(r,t)$ is $\frac{d\phi(r,t)}{dt} - C\nabla^2\phi(r,t)$=0, note that in (1.4) there is a plus sign instead of a minus (due to standardized notation).

Since the eigenvectors of a symmetric matrix, such as the Laplacian, are orthogonal, this equation holds for each i independently:

$$\frac{da_i}{dt} + C\lambda_i a_i = 0. \tag{1.7}$$

Which has as the solution:

$$a_i(t) = a_i(0)e^{-C\lambda_i t}. \tag{1.8}$$

Thus,

$$\phi(t) = \sum_i a_i(0)e^{-C\lambda_i t}. \tag{1.9}$$

Since $L_\mathcal{G}$ always has a null eigenvalue[†], $\lambda_1 = 0$ associated with the eigenvector $\mathbf{v}_1 = (1, \ldots, 1)$, and all other eigenvalues are positive, it follows that as $t \to \infty$, $\phi(t) \to a_1(0)$. This represents the stationary solution, reached over a characteristic timescale $\tau = \frac{1}{\lambda_2}$, where λ_2 is the smallest nonzero eigenvalue.

1.2 Connectivity

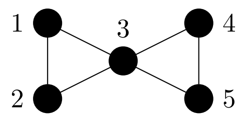
One of the main aspects of interest when analyzing a connected network is assessing how well connected the nodes of the network are. A way to quantify this is by measuring how many edge or node eliminations a connected network can withstand before becoming disconnected.

Definition 1.9. The *node connectivity* of a connected network $\mathcal{G}$, denoted $k(\mathcal{G})$ is the minimum number of nodes whose removal results in a disconnected graph or a trivial graph (i.e. a graph with a single node).

In general, a network $\mathcal{G}$ with a larger node connectivity $k(\mathcal{G})$ is more robust, as it indicates greater resistance to node failures and, therefore, a stronger network.

Definition 1.10. The *edge connectivity* of a connected network $\mathcal{G}$, denoted $\lambda(\mathcal{G})$ is the minimum number of edges whose removal results in a disconnected graph or a trivial graph.

Example 1.11. The node and edge connectivity of



are $k(\mathcal{G}) = 1$, $\lambda(\mathcal{G}) = 2$.

[†]Following convention, the n different eigenvalues of the Laplacian are numbered in ascending order: $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$.

Remark 1.12. For a network $\mathcal{G} = (V, E)$, if there exists a set of edges of size $\lambda(\mathcal{G})$ whose removal disconnects the network, then by selecting one endpoint of each of these edges, we can form a set of nodes whose removal also disconnects the graph. Since this node set has a size less than or equal to $\lambda(\mathcal{G})$, the following inequality holds: $k(\mathcal{G}) \leq \lambda(\mathcal{G})$.

The value of both edge and node connectivity provide important insights into the characteristics of the studied network. Generally, we aim to achieve high values for these parameters, as they indicate greater resilience in the network. However, calculating the exact values of edge and node connectivity can be computationally challenging. Therefore, it is beneficial to find alternative values that can serve as a lower bound for the edge or node connectivity.

1.2.1 Algebraic connectivity

A valuable lower bound for the edge connectivity of a network $\mathcal{G}$ can be obtained by analyzing the Laplacian spectrum of the network. Additionally, important connectivity insights are obtained by studying this spectrum [8, section 6.13.3].

- The network $\mathcal{G}$ is connected if and only if the second smallest eigenvalue of $L_\mathcal{G}$, $\lambda_2(\mathcal{G})$, is positive.
- The number of zero eigenvalues of $L_\mathcal{G}$ corresponds to the number of connected components in the network $\mathcal{G}$.

The second smallest eigenvalue of the Laplacian matrix, $\lambda_2(\mathcal{G})$, is known as the *algebraic connectivity* or *Fiedler value* of $\mathcal{G}$. This value serves as a good indicator of the network's connectedness, as it provides a lower bound for the edge connectivity.

For any eigenvector v of $L_\mathcal{G}$ corresponding to a non-zero eigenvalue λ we have $v^T L_\mathcal{G} = \lambda v^T$. Multiplying both sides by the all-ones n-dimensional vector $\mathbf{1}_n$ we get

$$v^T L_\mathcal{G} \mathbf{1}_n = \lambda v^T \mathbf{1}_n.$$

Using the fact that $L_\mathcal{G} \mathbf{1}_n = 0$, we get $\lambda v^T \mathbf{1}_n = 0$. Since $\lambda \neq 0$, it follows that $v^T \mathbf{1}_n = 0$.

This condition allows us to express the algebraic connectivity of a network $\mathcal{G}$ on n nodes as the following optimization problem

$$\lambda_2(\mathcal{G}) = \min_{\substack{v^T \mathbf{1}_n = 0 \\ v \neq 0}} \frac{v^T L_\mathcal{G} v}{\|v\|^2}. \tag{1.10}$$

Lemma 1.13. For a network $\mathbf{G} = (V, E)$, if $\mathcal{G}_1$ is obtained by removing n nodes (and their incident edges), then $\lambda_2(\mathcal{G}_1) \geq \lambda_2(\mathcal{G}) - n$

Proof. See [9, Theorem 3.2].

Proposition 1.14. The node connectivity $k(\mathcal{G})$, edge connectivity $\lambda(\mathcal{G})$, and algebraic connectivity $\lambda_2(\mathcal{G})$ of a non-complete network [‡] $\mathcal{G} = (V, E)$ satisfy the following inequality

$$\lambda_2(\mathcal{G}) \leq \lambda(\mathcal{G}) \leq k(\mathcal{G}). \tag{1.11}$$

Proof. Consider V_1 as a node cut of size $\lambda(\mathcal{G})$, such that removing V_1 and its incident edges from $\mathcal{G}$ results in a disconnected network $\mathcal{G}_1$. Since $\mathcal{G}$ is not complete, $\mathcal{G}_1$ is non-empty and disconnected, so $\lambda_2(\mathcal{G}_1) = 0$. Applying Lemma 1.13 we get

$$0 \geq \lambda_2(\mathcal{G}) - \lambda(\mathcal{G}).$$

The second inequality is obtained from Remark 1.12.

□

1.3 Network quotients

In network analysis, there are various situations in which it is useful to work with a subnetwork formed by coarsening the original network according to a node partition. The resulting network, known as the *quotient network*, preserves important spectral relationships of the initial network, which will be crucial for the developments that follow in this work. The main concepts we will explore in this section are interlacing and the quotient of a network, based on the approach outlined in [10].

Definition 1.15. Given integers $m, n \in \mathbb{N}$ with $m < n$, consider two ordered sets of numbers $\mu_1 \leq \ldots \leq \mu_m$ and $\lambda_1 \leq \ldots \leq \lambda_n$. The first set is said to *interlace* the second if:

$$\lambda_i \leq \mu_i \leq \lambda_{i+(n-m)} \text{ for } i = 1, \ldots, m.$$

Example 1.16. The set $A_1 = \{-4, -2, -1, 0, 1, 2, 4\}$ interlaces $A_2 = \{-3, 0, 3\}$ since $-4 \leq 3 \leq 1$, $-2 \leq 0 \leq 2$ and $-1 \leq 3 \leq 4$.

To define the quotient of a network, we first introduce the concepts of *partition* of the node set and its associated *characteristic matrix*.

[‡]A network in which at least two distinct nodes are not connected by an edge.

Definition 1.17. A *partition* of a set X is a set of non-empty subsets of $\{X_1, \ldots, X_m\}$ such that every element $x \in X$ belongs to exactly one subset X_i. For a network $\mathcal{G} = (V, E)$, and a partition $\{V_1, \ldots, V_m\}$ of the node set V, the *characteristic matrix* of the partition, $S = (s_{ij})$ is a $|V| \times m$ matrix with $s_{ij} = 1$ if $i \in V_j$ and 0 otherwise.

Definition 1.18. Let $\mathcal{G} = (V, E)$ be a network with adjacency matrix $A = (a_{ij})$, and let $\{V_1, \ldots, V_m\}$ be a partition of the node set V, where $|V_i| = n_i$. The *quotient network* $\mathcal{Q}$ of $\mathcal{G}$ is defined as the network coarsened according to this partition. In $\mathcal{Q}$, each partiton V_i is represented as a single node, and an edge from V_i to V_j is weighted by the average connectivity between nodes in V_i and V_j:

$$b_{ij} = \frac{1}{\sigma} \sum_{\substack{k \in V_i \\ l \in V_j}} a_{kl}, \tag{1.12}$$

where various choices can be made for the size parameter σ: $\sigma_i = n_i$, $\sigma_j = n_j$ or $\sigma_{ij} = \sqrt{n_i n_j}$. The resulting network is referred to as the *left quotient*, *right quotient* or *symmetric quotient*, respectively.

Note that the remaining quotient network can have self-edges and, while the symmetric quotient is undirected, the left and right quotients are generally directed unless all clusters have the same size ($n_i = n_j$ for all i and j).

Let $S = (s_{ij})$ be the characteristic matrix of the partition, $\Lambda = diag(n_1, \ldots, n_m)$ and A the adjacency matrix of the original network. The adjacency matrix of the quotient network, denoted $\mathcal{Q}(A) = (b_{ij})$[§], can be expressed as:

$$\mathcal{Q}_l(A) = \Lambda^{-1} S^T A S,\ \mathcal{Q}_r(A) = S^T A S \Lambda^{-1} \text{ and } \mathcal{Q}_s(A) = \Lambda^{-1/2} S^T A S \Lambda^{-1/2}.$$

Proposition 1.19. Let $\mathcal{G} = (V, E)$ be a network with adjacency matrix A, and let $\{V_1, \ldots, V_m\}$ be a partition of the node set V. The adjacency matrices of the left, right and symmetric quotients with respect to the partition, $\mathcal{Q}_l(A)$, $\mathcal{Q}_r(A)$, $\mathcal{Q}_s(A)$, share the same eigenvalues.

Proof. Let $B = S^T A S$. Then we have $\mathcal{Q}_l(A) = \Lambda^{-1} B$, $\mathcal{Q}_r(A) = B \Lambda^{-1}$, and $\mathcal{Q}_s(A) = \Lambda^{-1/2} B \Lambda^{-1/2}$. Using the Weinstein-Aronszajn identity [13, Appendix B.1]

$$\det(\lambda Id - \Lambda^{-1} B) = \det(\lambda Id - B \Lambda^{-1}) = \det(\lambda Id - \Lambda^{-1/2} B \Lambda^{-1/2}).$$

Thus, all three matrices $\mathcal{Q}_l(A)$, $\mathcal{Q}_r(A)$ and $\mathcal{Q}_s(A)$ share the same characteristic polynomial, implying they have the same eigenvalues.

□

[§]When specifying the type of quotient, we write $\mathcal{Q}_l$, $\mathcal{Q}_r$ and $\mathcal{Q}_s$ for the left, right, or symmetric quotient, respectively.

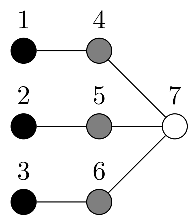
Moreover, the main spectral result is that the adjacency eigenvalues of a quotient interlace the adjacency eigenvalues of the initial network. This result is a direct consequence of the following theorem.

Theorem 1.20 ([11, Theorem 2.1])**.** Let A be a symmetric matrix of order n, and let U be an $n \times m$ matrix such that $U^T U = Id$. Then the eigenvalues of $U^T A U$ interlace those of A.

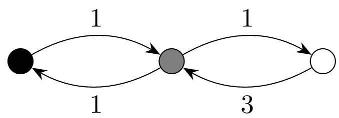
Using this theorem, we obtain the following result.

Corollary 1.21 ([11, Corollary 2.3])**.** Let B be the quotient matrix of A with respect to a given partition. Then, the eigenvalues of B interlace the eigenvalues of A.

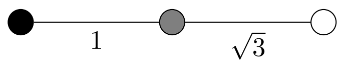
Example 1.22. A possible quotient of the following network $\mathcal{G}$:



can be obtained by considering the partition set $V_1 = \{1,2,3\}$, $V_2 = \{4,5,6\}$, $V_3 = \{7\}$. The left and symmetric quotients are:



left quotient



symmetric quotient

The spectrum of the adjacency matrix for the initial network is $\Lambda(A_{\mathcal{G}}) = \{-2,-1,-1,0,1,1,2\}$ while the spectrum for the quotient network is $\Lambda(\mathcal{Q}) = \{-2,0,2\}$. Thus, the eigenvalues of the adjacency matrix of the quotient network interlace those of the original network.

As the quotient network is generally not represented by an undirected matrix, a proper definition of the Laplacian for quotient networks is necessary to derive analogous interlacing results. The following definition of the quotient Laplacian allows us to obtain the same interlacing results that we have previously mentioned ¶.

¶See [10, Appendix 4 a] for a broad discussion of the choice of this Laplacian.

Definition 1.23. For a network $\mathcal{G} = (V, E)$ with adjacency matrix $A_{\mathcal{G}} = (a_{ij})$, consider the left quotient $\mathcal{Q}_l(A)$ with respect to a partition $\{V_1, \ldots, V_m\}$ of the node set. Define the row sums of $\mathcal{Q}_l(A)$ as $\overline{d_i} = \frac{1}{n_i} \sum_{k \in V_i} d_k$, which represent the average degree of nodes in V_i. Let $\overline{D}$ denote the diagonal matrix of these average degrees. The *quotient Laplacian* $L_{\mathcal{Q}}$ is defined by:

$$L_{\mathcal{Q}} = \overline{D} - \mathcal{Q}_l(A). \tag{1.13}$$

An equivalent definition is valid considering the left or symmetric quotients, independently of the type of quotient considered, as the spectrum of the quotient Laplacian remains consistent across these types.

Using this definition, the eigenvalues of the quotient Laplacian interlace with those of the Laplacian of the original network [10, Appendix 4 a. Theorem 1].

Note that, as this definition of the Laplacian matrix ignores self-loops, since the diagonal values of the adjacency matrix are also counted in the diagonal values of $\overline{D}$), we have that for $\tilde{\mathcal{Q}}$, the loopless quotient of $\mathcal{G}$ (i.e., the quotient network of $\mathcal{G}$ with all the self-loops removed), the quotient laplacian remains unchanged: $L_{\tilde{\mathcal{Q}}} = L_{\mathcal{Q}}$.

This observation is helpful when analyzing quotient networks where self-loops may or may not be present, as it ensures that removing self-loops form $\mathcal{Q}$ does not affect the interlacing properties of the quotient Laplacian's spectrum.

1.3.1 Regular quotients

If additional conditions on the node partition are imposed, a stronger relationship between the quotient network and the original network can be obtained.

Definition 1.24. For a network $\mathcal{G} = (V, E)$, a partition of the node set $\{V_1, \ldots, V_m\}$ is called *equitable* if the number of edges (accounting for edge weights) from a node V_i to any node in V_j is independent of the choice of node in V_i. Formally, this means:

$$\sum_{l \in V_j} a_{kl} = \sum_{l \in V_j} a_{k'l} \text{ for all } k, k' \in V_i,$$

for all i and j. When a partition is equitable, the resulting quotient is called a *regular quotient*.

For networks with equitable partitions, the spectral connection between the quotient and the original network is particularly strong.

Specifically, the eigenvalues of the adjacency matrix of the regular quotient are a subset of the eigenvalues of the adjacency matrix of $\mathcal{G}$. Furthermore, an eigenbasis of $\mathcal{G}$ can be constructed with m eigenvectors derived from *lifting* the eigenvectors of the quotient to $\mathcal{G}$ (by repeating the coordinates on each cluster). The remaining $n-m$ eigenvectors are orthogonal to the partition, meaning that the sum of their coordinates within each layer is zero [10, Appendix 3]. This spectral phenomenon is known as *lifting*.

Example 1.25. The network used in example 1.22 is an equitable partition. Note that the spectrum of the quotient is a subset of the spectrum of the original network.

A similar spectral relationship can be obtained for the Laplacian matrix when slightly relaxed conditions on the partition are used:

Definition 1.26. For a network $\mathcal{G} = (V, E)$, a partition of the node set $\{V_1, \ldots, V_m\}$ is called *almost equitable* if the condition specified in Definition 1.24 is satisfied for all $i \neq j$ but not necessarily for $i = j$. In other words, the regularity condition is satisfied when ignoring the intra-cluster edges. When a partition is almost equitable, the resulting quotient is called an *almost regular quotient*.

Note that the quotient $\mathcal{Q}$ being almost regular is equivalent to the loopless quotient $\tilde{\mathcal{Q}}$ being regular.

For networks with almost equitable partitions, the Laplacian eigenvalues of $\mathcal{Q}$ are a subset of the Laplacian eigenvalues of $\mathcal{G}$. Moreover, we can construct a Laplacian eigenbasis of $\mathcal{G}$ consisting of m Laplacian eigenvectors of the quotient (either $\mathcal{Q}$ or $\tilde{\mathcal{Q}}$) lifted to $\mathcal{G}$, with the remaining $n-m$ eigenvectors orthogonal to the partition [10, Appendix 4 a].

Chapter 2

Multilayer networks

When working with networks $\mathcal{G} = (V, E)$ whose components operate across different levels or when different types of nodes and edges are present (i.e., when considering different *aspects* in the network connections), it is often useful to introduce an additional structural element: a set of *layers*. Following the approach presented in [5], a first general view of this *multilayer network* framework is obtained by allowing each node to belong to any subset of layers and permitting edges to form pairwise connections across all possible combination of nodes and layers. For instance, consider a network representing communications among individuals that considers two aspects: the type of relationship (e.g., family, friends or colleagues) and the mode of communication (e.g., letters, online or in-person). This layered perspective allows us to clearly differentiate the diverse interactions more comprehensively than using a traditional single-layer network.

In a multilayer network associated with d aspects, each aspect a has associated a set of *elementary layers* L_a. This new multilayer structure is defined by the sequence of the set of elementary layers $\mathbf{L} = \{L_a\}_{a=1}^{d}$, with the *set of layers* represented as the Cartesian product $L_1 \times \ldots \times L_d$, describing all possible combinations of aspects.

To indicate whether a node in V is present in a given layer, the Cartesian product $V \times L_1 \times \ldots \times L_d$ is used, defining a subset $V_M \subseteq V \times L_1 \ldots L_d$ to represent all valid *node-layer tuple* combinations among the different layers, requiring that each node must appear in at least one layer. Thus, the node-layer tuple $(v, \alpha_1, \ldots, \alpha_d)$* represents the node v on layer $(\alpha_1, \ldots, \alpha_d)$.

When defining connections between pairs of node-layer tuples in this multilayer framework, the starting and ending layers of each edge must be specified, apart from the starting and ending node. Therefore, the *edge set* in a multilayer net-

*Usually the array of elementary layers $(v, \alpha_1, \ldots, \alpha_d)$ is simply denoted by $\boldsymbol{\alpha}$.

work is defined as the set of all possible pairs node-layer tuples: $E_M \subseteq V_M \times V_M$. For instance, the edge $((v, \boldsymbol{\alpha}), (u, \boldsymbol{\beta}))$ connects node v in layer $\boldsymbol{\alpha}$ with node u in layer $\boldsymbol{\beta}$.

Definition 2.1. The quadruplet $\mathcal{M} = (V_M, E_M, V, \mathbf{L})$ defines a *multilayer network*.

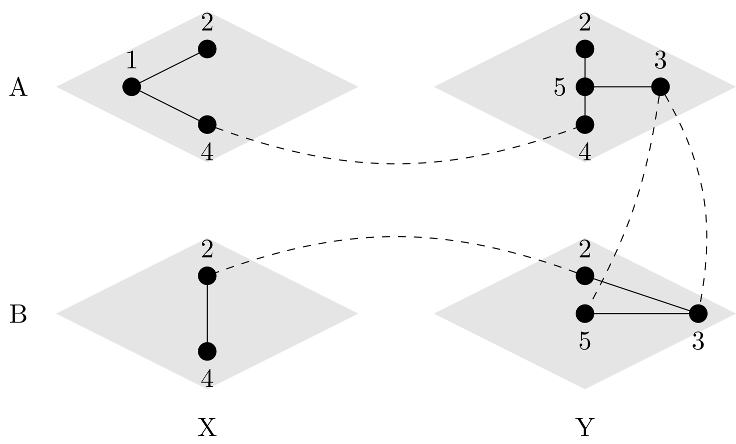
Note that we have two different entities representing the components of a multilayer network: nodes and node-layer tuples. The first one corresponds to a "physical object" (e.g., an individual), while the node-layer tuples are different instances of the same object (e.g., the same individual under a specific type of relationship and communication mode).

If no aspects are considered, the number of aspects d is formally zero, as there are no elementary layers. In this case, the multilayer network reduces to a *monoplex* (i.e., a single-layer network), where $V_M = V$, making the set of node-layer tuples V_M redundant.

A multilayer network is said to be *node-aligned* (or "*fully interconnected*") if all of the layers contain all nodes, i.e., $V_M = V \times L_1 \times \ldots \times L_d$.

As with monoplex networks, the term *adjacency* is used to describe a direct connection via an edge between a pair of node-layers and the term *incidence* is used to describe the connection between a node-layer and an edge.

Example 2.2. An example of this general type of multilayer network $\mathcal{M}$ is illustrated below



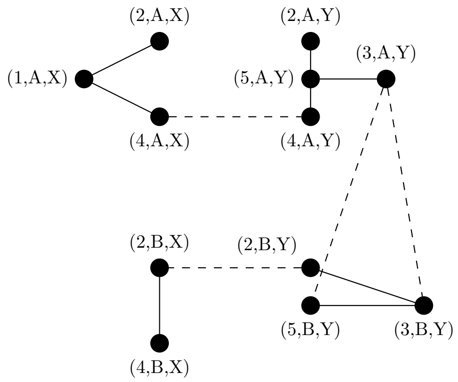
This multilayer network has a total of five nodes $V = \{1, 2, 3, 4, 5\}$ and two aspects, corresponding to the elementary layer sets $L_1 = \{A, B\}$ and $L_2 = \{X, Y\}$, resulting in four distinct layers: (A, X), (A, Y), (B, X), (B, Y). Examples of elements in the set of node-layer tuples V_M include $(1, A, X)$ and $(5, B, Y)$. Similarly, examples of elements in the edge set E_M include $((2, B, X), (4, B, X))$ and $((5, B, Y), (3, A, Y))$.

This network could model a system that considers $d = 2$ aspects: a social media platform (e.g. A = Twitter, B = Facebook) and interaction types (e.g. X = Like and Y = Share). For instance, user 3 has shared posts in Twitter from their own Facebook account and user 5's Facebook account.

The first two components of the multilayer network $\mathcal{M}$ form an underlying graph $\mathcal{G}_M = (V_M, E_M)$, allowing us to interpret a multilayer network as a network with nodes that are labeled in a specific way. This interpretation allows us to generalize some basic concepts from monoplex networks to multilayer networks.

- A multilayer network is *directed* (or *undirected*) if the underlying graph $\mathcal{G}_M = (V_M, E_M)$ is directed (or undirected).
- A *weighted multilayer network* is one in which weights are assigned to the edges of the underlying graph $\mathcal{G}_M$.
- The multilayer network includes *self-edges* or *self-loops* if they are present in $\mathcal{G}_M$.

Example 2.3. The underlying graph $\mathcal{G}_M$ of the multilayer network from Example 2.2 is



Moreover, it is convenient to distinguish edges that connect nodes within the same layer from those that connect nodes across different layers. To do so, we can partition the edge set into two categories, *intra-layer edges* $E_A = \{((u, \boldsymbol{\alpha}), (v, \boldsymbol{\beta})) \in E_M | \boldsymbol{\alpha} = \boldsymbol{\beta}\}$ and *inter-layer edges* $E_L = E_M \setminus E_A$. A subset of the intra-layer edge set can also be defined for edges connecting the same entity in different layers. These are known as *coupling edges* $E_C = \{((u, \boldsymbol{\alpha}), (v, \boldsymbol{\beta})) \in E_L | u = v\}$.

This categorization allows us to define three distinct graphs associated with the multilayer network, the *intra-layer graph* $\mathcal{G}_A = (V_M, E_A)$, the *inter-layer graph* $\mathcal{G}_L = (V_M, E_L)$ and the *coupling graph* $\mathcal{G}_C = (V_M, E_C)$.

Further inherent graphs can be obtained from the original multilayer by considering subsets of the intra-layer edge set for edges within the same layer $\boldsymbol{\alpha}$. These subsets are denoted as $E_{\boldsymbol{\alpha}} = \{((u, \boldsymbol{\alpha}), (v, \boldsymbol{\alpha})) \in E_A\}$. Using the node-layer tuple set from the layer $\boldsymbol{\alpha}$, we can obtain the corresponding *layer-graph*, denoted as $\mathcal{G}_{\boldsymbol{\alpha}} = (V_{\boldsymbol{\alpha}}, E_{\boldsymbol{\alpha}})$.

2.1 Tensor representation

When working with multilayer networks and different channels of interaction within a system, tensor notation provides a robust framework for handling these multidimensional problems. This formalism generalizes concepts such as scalars, vectors and matrices, which naturally arose when modeling interactions within simple networks (e.g, the use of adjacency matrices) into a structure that can effectively represent the additional complexity of interactions within multilayer networks. Thus, for multilayer networks, the equivalent of the adjacency matrix in simple networks is the *adjacency tensor*.

Definition 2.4. For a node-aligned multilayer network $\mathcal{M} = (V_M, E_M, V, \mathbf{L})$ on d aspects, the *adjacency tensor* is a rank-$2(d+1)$ tensor,

$$\mathcal{A} \in \{0, 1\}^{|V| \times |V| \times |\mathbf{L_1}| \times |\mathbf{L_1}| \times \ldots \times |\mathbf{L}_d| \times |\mathbf{L}|_d}$$

where the tensor element $\mathcal{A}_{uv\boldsymbol{\alpha}\boldsymbol{\beta}} := \mathcal{A}_{uv\alpha_1\beta_1\ldots\alpha_d\beta_d}$ has a value of 1 if and only if $((u, \boldsymbol{\alpha}), (v, \boldsymbol{\beta})) \in E_M$; otherwise, $\mathcal{A}_{uv\boldsymbol{\alpha}\boldsymbol{\beta}}$ is 0.

Similarly, a *weighted adjacency tensor*, $\mathcal{W}$ can be defined, where each element $\mathcal{W}_{uv\boldsymbol{\alpha}\boldsymbol{\beta}}$ corresponds to the weight of the edge $((u, \boldsymbol{\alpha}), (v, \boldsymbol{\beta})) \in E_M$, and 0 if no edge exists.

Since this tensor representation relies on the existence of each node in every layer, it is technically only appropriate for node-aligned networks. However, many of the tensor-based methods for multilayer network analysis can be adapted to work with non-node-aligned networks by introducing extraneous nodes in each layer (that are called *empty nodes*)[†]. These nodes do not participate in any connections of the layer but enable a consistent tensor structure across layers. This approach allows the tensor formalism to be extended to networks where certain nodes may not be present in every layer by effectively "projecting" each node into every layer.

This projection creates a structure that is node-aligned from a mathematical perspective. However, the resulting tensors must be interpreted carefully, as they

[†]A discussion of the employment of tensors to represent non-node-aligned networks can be found in [5] and [12].

can lead to misleading characterizations of network properties. For example, computing the mean node degree or connectivity of the network may yield inaccurate results unless empty nodes are properly accounted for.

2.1.1 Constraints

In node-aligned multilayer networks, certain constraints can simplify the tensor representation by setting to zero, allowing the network to be represented with a tensor of lower rank than would otherwise be required.

An important case of a constrained network is found when the multilayer network has *diagonal couplings*, meaning that inter-layer edges are permitted only between two representations of the same node across different layers. In this case the adjacency-tensor element $\mathcal{A}_{uv\boldsymbol{\alpha}\boldsymbol{\beta}}$ is forced to be zero whenever $u \neq v$ and $\boldsymbol{\alpha} \neq \boldsymbol{\beta}$ (i.e., when the nodes and layers differ at the same time).

Under this restriction, the multilayer network can be expressed as a combination of an *intra-layer adjacency tensor* with elements $\mathcal{A}_{uv\boldsymbol{\alpha}} := \mathcal{A}_{uv\boldsymbol{\alpha}\boldsymbol{\beta}}$ (encapsulating the connections in the same layer) and a *coupling tensor* with elements $\mathcal{C}_{u\boldsymbol{\alpha}\boldsymbol{\beta}} := \mathcal{A}_{uv\boldsymbol{\alpha}\boldsymbol{\beta}}$ (encapsulating the connections across layers).

Further constraints can be obtained by disallowing inter-aspect couplings, as the adjacency tensor has null elements when the layer indices differ in more than one aspect. We will not delve further into these possible constraints as they will not be used later on this work.

2.1.2 Tensor flattening

As outlined in [5, Section 2.2.2], the number of aspects in an adjacency tensor can be reduced by combining two aspects i and j into a single new aspect h.

This transformation, known as *flattening*, allows a node-aligned multilayer network $\mathcal{M} = (V_M, E_M, V, \mathbf{L})$ with d aspects to be mapped into a new node-aligned multilayer network $\mathcal{M}' = (V'_M, E'_M, V, \mathbf{L}')$ with $d-1$ aspects. In this new configuration, the new aspect h is defined by $L'_h = L_i \times L_j$, ensuring the total number of elements remains consistent, as $|L'_h| = |L_i||L_j|$. This process establishes a bijection between the elements of the original and flattened tensors, preserving the original network structure.

Consider a node-aligned multilayer network $\mathcal{M} = (V_M, E_M, V, \mathbf{L})$ with d aspects, where the layers are denoted using integers starting from one. Without loss of generality, suppose we flatten aspects $d-1$ and d of the adjacency tensor $\mathcal{A}$ of $\mathcal{M}$ to obtain a new flattened tensor $\mathcal{A}'$. The elements of the corresponding mapping are given by

$$\mathcal{A}_{uv\alpha_1\beta_1\ldots\alpha_{d-1}\beta_{d-a}\alpha_d\beta_d} = \mathcal{A}'_{uv\alpha_1\beta_1\ldots((\alpha_{d-1}-1)|L_d|+\alpha_d)((\beta_{d-1}-1)|L_d|+\beta_d)}.$$

This flattening can be repeatedly iteratively to further reduce the number of aspects, as it can be useful both conceptually and when implementing software algorithms to analyze multilayer networks.

2.2 Matrix representation

An extreme case of tensor flattening occurs when the multilayer network is reduced to $d = 0$ aspects by combining all aspect layers and node indices into a single dimension. This process produces a rank-2 adjacency tensor, which corresponds to a matrix known as the *supra-adjacency matrix* of the multilayer network.

Definition 2.5. For a multilayer network $\mathcal{M} = (V_M, E_M, V, \mathbf{L})$, the *supra-adjacency matrix* corresponds to the adjacency matrix of its underlying graph $\mathcal{G}_M = (V_M, E_M)$. Other "supra-matrices" are defined in an analogous way.

An advantage of using supra-adjacency matrices over adjacency tensors is that they provide a natural way to represent multilayer networks that are not necessarily node-aligned without having to add empty nodes. However, by flattening the multilayer network to obtain the supra-adjacency matrix, some information of the aspects is lost.

Some of this information can be retained by considering the partition of the edge set of the multilayer network into intra-layer edges, inter-layer edges and coupling edges and working with the *intra-layer supra-adjacency matrix*, *inter-layer supra-adjacency matrix* and *couping supra-adjacency matrix* (obtained respectively from the graphs $\mathcal{G}_A$, $\mathcal{G}_L$ and $\mathcal{G}_C$), along with the *supra-adjacency matrix*.

A *supra-Laplacian matrix* can be derived from the supra-adjacency matrix in a manner analogous to how the Laplacian matrix is constructed for monoplex networks, providing insights into the network's connectivity, dynamics, diffusion processes and other structural and functional properites.

Different definitions of the Laplacian matrix can be considered depending on special constraints and characteristics of the considered multilayer network. Still, an example of a general definition of a supra-Laplacian matrix suitable to any type of multilayer network is $L_M = D_M - A_M$, where D_M is the diagonal supra-matrix containing the weighted degrees of the graph $\mathcal{G}_M$ and A_M the adjacency matrix of $\mathcal{G}_M$.

Since this definition is analogous to that used for monoplex networks, the eigenvectors and eigenvalues of the supra-Laplacian matrix are important indicators of several structural features of the network, providing important insights into dynamical processes that evolve on top of it. Analogously, the second smallest eigenvalue of the supra-Laplacian matrix is known as the *algebraic connectivity* (or *Fiedler value*) of $\mathcal{M}$.

Depending on the nature of the inter-layer couplings of the network, one can differentiate distinct behaviors based on the relative strengths of the inter-layer and intra-layer edges. For instance, in a node-aligned multilayer network with diagonal couplings, two regimes emerge, separated by a discontinuous phase transition known as a "structural transition". In one regime, the algebraic connectivity is independent of the intra-layer structure and is instead determined by the inter-layer edges. In the other regime, the algebraic connectivity of the multilayer network is bounded above by a constant multiple of the algebraic connectivity of the unweighted superposition of the layers (see Subsection 3.2 for broader details on structural transition in the case of multiplex networks, a special case of multilayer networks).

2.3 Dimensionality reduction

The multilayer framework formalism induces two naturally occurring quotients: the *network of layers* and the *aggregate network*, which simplify the analysis of a multilayer network with d aspects by reducing it to a simple network structure.

Using the results discussed in Section 1.3, strong relationships between the multilayer network and the structure of its layers can be established, allowing us to work with simple networks to analyze the original multilayer structure. This dimensionality reduction facilitates the application of the tools introduced in Chapter 1 to obtain meaningful results of the original network.

2.3.1 Network of layers

Consider a multilayer network $\mathcal{M} = (V_M, E_M, V, \mathbf{L})$ with d aspects and let its underlying graph be $\mathcal{G}_M = (V_M, E_M)$. Each layer of the multilayer network can be viewed as a subgraph $\mathcal{G}_{\boldsymbol{\alpha}} = (V_{\boldsymbol{\alpha}}, E_{\boldsymbol{\alpha}})$, where $V_{\boldsymbol{\alpha}} = \{(u, \boldsymbol{\alpha}) \in V_M\}$ and $E_{\boldsymbol{\alpha}} = \{((u, \boldsymbol{\alpha}), (v, \boldsymbol{\alpha})) \in E_M\}$. Note how the layers of the multilayer network partition the node set, making it natural to consider the quotient induced by this partition. Let $\{V_{\boldsymbol{\alpha}}\}_{\boldsymbol{\alpha} \in L_1 \times \ldots \times L_d}$ be the partition of the multilayer node-layer set by the layers with $n_{\boldsymbol{\alpha}} = |V_{\boldsymbol{\alpha}}|$.

Definition 2.6. The *average inter-layer degree* from $\boldsymbol{\alpha}$ to $\boldsymbol{\beta}$ is defined as

$$d^{\boldsymbol{\alpha}\boldsymbol{\beta}} = \frac{1}{n_{\boldsymbol{\alpha}}} \sum_{\substack{i \in V_{\boldsymbol{\alpha}} \\ j \in V_{\boldsymbol{\beta}}}} a_{ij}, \tag{2.1}$$

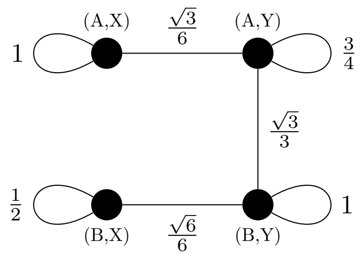
where $A_{\mathcal{G}_M} = (a_{ij})$ is the adjacency matrix of $\mathcal{G}_M$.

This metric represents the average connectivity from a node in $\mathcal{G}_{\boldsymbol{\alpha}}$ to any node in $\mathcal{G}_{\boldsymbol{\beta}}$. When $\boldsymbol{\alpha} = \boldsymbol{\beta}$ we write $d^{\boldsymbol{\alpha}}$ instead of $d^{\boldsymbol{\alpha}\boldsymbol{\alpha}}$ and refer to it as the *average intra-layer degree*.

The quotient with respect to the partition $\{V_{\boldsymbol{\alpha}}\}_{\boldsymbol{\alpha} \in L_1 \times \ldots \times L_d}$ forms a weighted directed network, with adjacency matrix $(d^{\boldsymbol{\alpha}\boldsymbol{\beta}})$. This resulting network is called the *network of layers*, a directed network where each node corresponds to a layer of $\mathcal{M}$, with a self-loop on layer $\boldsymbol{\alpha}$ weighted by the average intra-layer degree $d^{\boldsymbol{\alpha}}$, and a directed edge from layer $\boldsymbol{\alpha}$ to layer $\boldsymbol{\beta}$ is weighted by the average inter-layer degree $d^{\boldsymbol{\alpha}\boldsymbol{\beta}}$. Note that the network of layers is undirected if all layers contain the same number of nodes and that if weights and self-loops are ignored, this network simply represents the layer connection configuration.

As described in Section 1.3, one can construct a spectrally equivalent symmetric quotient by replacing $\frac{1}{n_{\boldsymbol{\alpha}}}$ by $\frac{1}{\sqrt{n_{\boldsymbol{\alpha}} n_{\boldsymbol{\beta}}}}$ in Equation (2.1).

Example 2.7. The undirected network of layers of the multilayer network from Example 2.2 is



Applying the spectral results of Section 1.3, the eigenvalues of both the adjacency and Laplacian matrices of the network of layers interlace those of the supra-adjacency and supra-Laplacian matrices of the multilayer network. That is, for a multilayer network with n node-layer tuples and m layers, if $\mu_1 \leq \ldots \leq \mu_m$ are the adjacency (resp. Laplacian) matrix eigenvalues of the network of layers, then

$$\lambda_i \leq \mu_i \leq \lambda_{i+(n-m)} \quad \text{for} \quad i = 1, \ldots, m, \tag{2.2}$$

where $\lambda_1 \leq \ldots \leq \lambda_n$ are the supra-adjacency (resp. supra-Laplacian) eigenvalues of the multilayer network.

It is particularly convenient to explore whether the layer partition is equitable, as this condition provides stronger relationships between the eigenvalues of the network of layers and those of the multilayer network. For this condition to hold, each layer must be a $d^{\boldsymbol{\alpha}}$-regular graph, a very strong requirement that is unlikely satisfied in real-world multilayer networks and is the reason why this condition is relaxed.

Definition 2.8. A multilayer network $\mathcal{M} = (V_M, E_M, V, \mathbf{L})$ with d aspects is *regular* if the layer partition $\{V_{\boldsymbol{\alpha}}\}_{\boldsymbol{\alpha} \in L_1 \times \ldots \times L_d}$ is almost equitable. That is, the inter-layer connections are independent of the chosen vertices.

This is a more natural condition that can be found in multilayer networks with *all-to-all* (all nodes are connected), *empty* or *one-to-one* (each node connects to only one node) connections with homogeneous weights.

If the multilayer network is regular, then in addition to the interlacing property of eigenvalues, the Laplacian eigenvalues of the network of layers form a subset of the Laplacian eigenvalues of the multilayer network. Moreover, a Laplacian eigenbasis of the network of layers can be lifted to form a supra-Laplacian eigenbasis of the multilayer network, as shown in Subsection 1.3.1.

2.3.2 Aggregate network

Consider a multilayer network $\mathcal{M} = (V_M, E_M, V, \mathbf{L})$ with d aspects, and let $\mathcal{G}_M = (V_M, E_M)$ be its underlying graph. For a node $u \in V$, the corresponding node-layer tuples represent the same entity in several layers, i.e., $(u, \boldsymbol{\alpha})$, $(u, \boldsymbol{\beta}) \in V_M$ with $\boldsymbol{\alpha} \neq \boldsymbol{\beta}$. Note how one can consider a partition of V_M into subsets of node-layer tuples that represent the same object across layers. This identification is particularly meaningful for multilayer networks where the behavior of a node in one layer critically depends on its behavior in another layer and vice versa.

Definition 2.9. For a multilayer network $\mathcal{M} = (V_M, E_M, V, \mathbf{L})$ with d aspects a *supra-node* of a node $u \in V$ is defined as $\tilde{u} = \{(u, \boldsymbol{\alpha}) \in V_M | \boldsymbol{\alpha} \in L_1 \times \ldots \times L_d\}$.

Note that the collection of supra-nodes $\{\tilde{u}\}_{u \in V}$ forms a partition of the multilayer node-layer set. We define $k_{\tilde{u}} = |\tilde{u}|$ as the *multiplexity degree* of the supra-node $\tilde{u}$, representing the number of layers in which the same object u is present.

Definition 2.10. The *average connectivity* between supra-nodes $\tilde{u}$ and $\tilde{v}$ is defined as

$$d_{\tilde{u}\tilde{v}} = \frac{1}{k_{\tilde{u}}} \sum_{\substack{i \in \tilde{u} \\ j \in \tilde{v}}} a_{ij}, \tag{2.3}$$

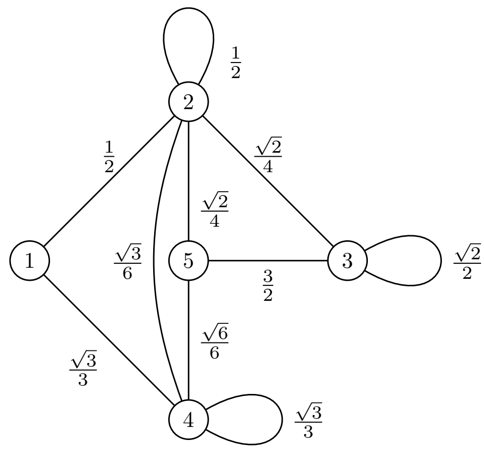
where $A_{\mathcal{G}_M} = (a_{ij})$ is the adjacency matrix of $\mathcal{G}_M$.

This metric represents the average connectivity between the node-layer tuples aggregated into their corresponding supra-nodes. When $\tilde{u} = \tilde{v}$ we write $d_{\tilde{u}}$ instead of $d_{\tilde{u}\tilde{u}}$.

The quotient with respect to the partition $\{\tilde{u}\}_{u \in V}$ forms a weighted directed network, with adjacency matrix $(d_{\tilde{u}\tilde{v}})$. This resulting network is called the *aggregate network*, a directed network where each node corresponds to a supra-node, with a self-loop on supra-node $\tilde{u}$ weighted by $d_{\tilde{u}}$ and a directed edge from $\tilde{u}$ to $\tilde{v}$ weighted by $d_{\tilde{u}\tilde{v}}$. Note that the aggregate network will be undirected if every supra-node has the same multiplexity degree.

As described in Section 1.3, one can construct a spectrally equivalent symmetric quotient by replacing $\frac{1}{k_{\tilde{u}}}$ by $\frac{1}{\sqrt{k_{\tilde{u}}k_{\tilde{v}}}}$ in Equation (2.3).

Example 2.11. The undirected aggregate network of the multilayer network from Example 2.2 is



Applying the spectral results of Section 1.3, the eigenvalues of both the adjacency and Laplacian matrices of the aggregate network interlace those of the supra-adjacency and supra-Laplacian matrices of the multilayer network. That is, for a multilayer network with n node-layer tuples and $\tilde{n}$ supra-nodes, if $\mu_1 \leq \ldots \leq \mu_{\tilde{n}}$ are the adjacency (resp. Laplacian) matrix eigenvalues of the aggregate network, then

$$\lambda_i \leq \mu_i \leq \lambda_{i+(n-\tilde{n})} \quad \text{for} \quad i = 1, \ldots, \tilde{n}, \tag{2.4}$$

where $\lambda_1 \leq \ldots \leq \lambda_n$ are the supra-adjacency (resp. supra-Laplacian) eigenvalues of the multilayer network.

Note that in this type of quotient, it is not usual to obtain stronger relationships between the eigenvalues of the aggregate network and those of the multilayer network. This is because having a regular or almost regular partition in this case is very restrictive, as every pair of nodes would need to connect in the same uniform way on every layer, which is unlikely to happen in real-world multilayer networks.

Chapter 3

Multiplex networks

In real-world scenarios, networks often exhibit inherent constraints that simplify their representation, as discussed in Section 2.1.1. These constraints not only lead to a more intuitive matrix representation of the network, but also reveal interesting properties unique to such constrained networks.

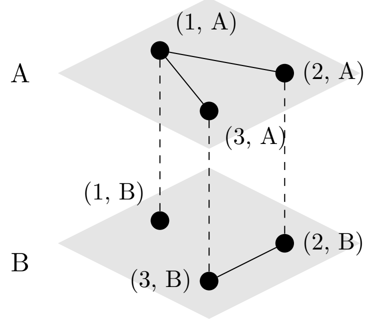
A usual constraint that arises naturally when considering communication networks, such as public transportation systems, is having only diagonal couplings. For instance, consider a multilayer network with a single aspect ($d = 1$) corresponding to the type of transportation mode (e.g., tram, bus or metro). Then, each layer in the multilayer network represents the connections between stops for a specific transportation mode, while the inter-layer connections indicate stops that serve as transfer points, where passengers can switch from one mode of transport to another. This specific type of multilayer network is an example of a *multiplex network*. The properties and concepts of multiplex networks have been comprehensively introduced in [14], whose notation and terminology are followed in this chapter.

Definition 3.1. A *multiplex network* is a multilayer network $\mathcal{M} = (V_M, E_M, V, \mathbf{L})$, where inter-layer edges between two node-layer tuples (u, α) and (v, β) exist if and only if $u = v$.

Even though there is not a restriction on the number of aspects d in the multiplex formalism, due to the nature of its connections, usually only one aspect $d = 1$ is considered. This work will only consider multiplex networks on one aspect, therefore the sequence of elementary layers consists of only one element, $\mathbf{L} = L$. If $|L| = m$, we say that the multiplex network has m layers.

Moreover, as only diagonal couplings are allowed, the inter-layer edges E_C coincide with the coupling edges $E_{\tilde{C}}$. Consequently, the underlying graph $\mathcal{G}_M$ can be expressed as $\mathcal{G}_M = \mathcal{G}_A \cup \mathcal{G}_{\tilde{C}}$, which is known as the *supra-graph* of the multiplex network.

Example 3.2. An example of a node-aligned multiplex network $\mathcal{M} = (V_M, E_M, V, L)$ with $V = \{1, 2, 3\}$ and $L = \{A, B\}$ is illustrated below



3.1 Matrix representation

Given a multiplex network $\mathcal{M} = (V_M, E_M, V, L)$ with m layers, consider the adjacency matrix of the layer-graph $\mathcal{G}_\alpha = (V_\alpha, E_\alpha)$ of layer α, where $|V_\alpha| = n_\alpha$. This matrix, denoted as $A^{(\alpha)} := A_{\mathcal{G}_\alpha}$ is an $n_\alpha \times n_\alpha$ matrix.

The matrix obtained from the direct sum of the different adjacency matrices of the layer-graphs is defined as

$$\mathcal{A} = \begin{pmatrix} A^{(1)} & 0 & \dots & 0 \\ 0 & A^{(2)} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & A^{(m)} \end{pmatrix} = \bigoplus_{\alpha \in L} A^{(\alpha)} \tag{3.1}$$

and is called the *intra-layer adjacency matrix*. Note that the $\overline{\mathcal{A}}$ is also the adjacency matrix of the intra-layer graph $\mathcal{G}_A$, i.e., $\mathcal{A} = A_{\mathcal{G}_A}$.

If we consider the adjacency matrix of the coupling graph $\mathcal{G}_C$, denoted by $\mathcal{C} := A_{\mathcal{G}_C}$, the *supra-adjacency matrix* of the multiplex $\mathcal{M}$ is defined as

$$\overline{\mathcal{A}} = \bigoplus_{\alpha \in L} A^{(\alpha)} + \mathcal{C} = \mathcal{A} + \mathcal{C}. \tag{3.2}$$

For an unweighted node-aligned multiplex network with m layers and n nodes, the supra-adjacency matrix takes a simpler form

$$\overline{\mathcal{A}} = \mathcal{A} + \mathbf{K}_m \otimes \mathbf{I}_n \tag{3.3}$$

where $\otimes$ denotes the Kronecker product, $\mathbf{K}_m$ is the adjacency matrix of a complete graph* on m nodes and $\mathbf{I}_n$ the $n \times n$ identity matrix.

*A graph where every two distinct nodes are connected by an edge.

As discussed in Subsection 2.1, we can always work with a node-aligned multiplex network by introducing empty nodes. Still, it is important to account for the presence of such empty nodes when analyzing the properties of their matrix representations.

Example 3.3. The supra-adjacency matrix of the multiplex used in Example 3.2 is

$$\overline{\mathcal{A}} = \left(\begin{array}{ccc|ccc} 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 \\ \hline 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 \end{array}\right)$$

To study the diffusion and movement across the multiplex, the Laplacian matrix also arises naturally, in an analogous than for simple networks (see Subsection 1.1).

For a multiplex network $\mathcal{M} = (V_M, E_M, V, L)$ with m layers, a supra-adjacency matrix $\overline{\mathcal{A}} = (\overline{a}_{ij})$ with diagonal blocks $A^{(\alpha)}$ and a total number of node-layer tuples $|V_M| = N$, a first definition of the *supra-Laplacian matrix* is defined as

$$\overline{\mathcal{L}} = \overline{\mathcal{D}} - \overline{\mathcal{A}} \tag{3.4}$$

where $\overline{\mathcal{D}} = \text{diag}(D_1, \dots, D_N)$, and D_i is the *degree of a node-layer tuple* $i^\dagger$, given by

$$D_i = \sum_j \overline{a}_{ij}, \tag{3.5}$$

which counts the number of node-layer tuples connected to i in $\mathcal{G}_M$.

Due to the clear distinction between intra-layer edges and inter-layer edges, the degree of a node-layer tuple can be expressed as

$$D_{i(\alpha)} = \sum_j \overline{a}_{ij} = d_{i(\alpha)} + c_{i(\alpha)} \tag{3.6}$$

where $d_{i(\alpha)}$ is the *layer-degree* and $c_{i(\alpha)}$ is the *coupling-degree* of the node-layer tuple $i(\alpha)$, defined as

$$d_{i(\alpha)} = \sum_j a_{ij}^{\alpha} = \sum_{j=1+e_\alpha}^{n_\alpha+e_\alpha} \overline{a}_{ij}, \tag{3.7}$$

$$c_{i(\alpha)} = \sum_j c_{ij} = \sum_{\substack{j<e_\alpha \\ j>n_\alpha+e_\alpha}} \overline{a}_{ij}, \tag{3.8}$$

†Sometimes $i(\alpha)$ is used instead of i to indicate that the node-layer tuple i is in layer α.

where $A^{(\alpha)} = (a^{\alpha}_{ij})$ is the adjacency matrix of $\mathcal{G}_\alpha$, $\mathcal{C} = (c_{ij})$ is the adjacency matrix of $\mathcal{G}_C$ and $e_\alpha = \sum_{\beta<\alpha} n_\beta$ is the *excess index* of layer α.

Since $D_{i(\alpha)} = d_{i(\alpha)} + c_{i(\alpha)}$, we can rewrite $\overline{\mathcal{D}}$ as

$$\overline{\mathcal{D}} = \operatorname{diag}(d_1 + c_1, \ldots, d_N + c_N) = \bigoplus_{\alpha\in L} D^{(\alpha)} + \Delta, \tag{3.9}$$

where $D^{(\alpha)}$ is the diagonal matrix of node degrees in the layer-graph $\mathcal{G}_\alpha$ and $\Delta = \operatorname{diag}(c_1, \ldots, c_N)$. From this, it follows that

$$\overline{\mathcal{L}} = \bigoplus_{\alpha\in L} D^{(\alpha)} + \Delta - \bigoplus_{\alpha\in L} A^{(\alpha)} - \mathcal{C} = \bigoplus_{\alpha\in L} L^{(\alpha)} + \Delta - \mathcal{C} \tag{3.10}$$

where $L^{(\alpha)} = D^{(\alpha)} - A^{(\alpha)}$ is the Laplacian matrix of the layer-graph $\mathcal{G}_\alpha$. This naturally leads to the decomposition of the supra-Laplacian matrix

$$\overline{\mathcal{L}} = \mathcal{L}_L + \mathcal{L}_C, \tag{3.11}$$

where $\mathcal{L}_L = \bigoplus_{\alpha\in L} L^{(\alpha)}$ is the *supra-Laplacian of the independent layers* and $\mathcal{L}_C = \Delta - \mathcal{C}$ is the *inter-layer supra-Laplacian*.

For an unweighted node-aligned multiplex network with m layers and n nodes, the supra-Laplacian matrix takes a simpler form

$$\overline{\mathcal{L}} = \bigoplus_{\alpha\in L} (L^{(\alpha)} + (m-1)\mathbf{I}_N) - \mathbf{K}_m \otimes \mathbf{I}_n. \tag{3.12}$$

This decomposition will be particularly useful in the following section, as it allows us to analyze different structural properties of the multiplex network in terms of the properties of both the layer-graphs $\mathcal{G}_\alpha$ and coupling graph $\mathcal{G}_C$.

3.2 Structural transitions

One of the main interests in the research of complex networks is focused in comprehending the relationships between the multiplex's topology and the behavior of the processes occurring within it. These behaviors can often be quantified through the network's connectivity, which is strongly related to its Laplacian through the algebraic connectivity, i.e., the smallest non-zero eigenvalue of the supra-Laplacian matrix. Abrupt changes in the algebraic connectivity leading to the distinction of different operational phases for multiplex networks that have no counterpart for traditional single-layer networks have been studied in previous works such as [15] and [16].

In this section, we examine the structural behavior of multiplex networks as a function of different variables that define them, deriving and examining various upper bounds for the algebraic connectivity of a two-layer multiplex network. This enables us to identify and characterize distinct behavioral regions: in some, the multiplex system behaves as one interconnected system, while in others, the layers effectively decouple and behave as independent networks.

3.2.1 Coupling edges

Consider an unweighted multiplex network $\mathcal{M} = (V_M, E_M, V, L)$ with m layers. We study its behavior as a function of the importance of inter-layer edges. To do so, following the approach made in [17], we introduce a weight parameter p that adjusts the relative strength of the inter-layer coupling compared to the intra-layer connectivity. By incorporating this weight parameter p into the coupling edges, Equation 3.11 takes the following form

$$\overline{\mathcal{L}} = \bigoplus_{\alpha \in L} L^{(\alpha)} + p\mathcal{L}_C \tag{3.13}$$

where L^{α} represents the Laplacian of the intra-layer graph $\mathcal{G}_{\alpha}$ and $\mathcal{L}_C$ is the inter-layer supra-Laplacian.

To study the different behavioral regions in a two-layer, node-aligned multiplex network with n nodes on each layer as a function of the relative strength of the inter-layer coupling, we explicitly write the matrix form of Equation 3.13

$$\overline{\mathcal{L}}(p) = \begin{pmatrix} L^{(1)} + p\mathbf{I}_n & -p\mathbf{I}_n \\ -p\mathbf{I}_n & L^{(2)} + p\mathbf{I}_n \end{pmatrix}. \tag{3.14}$$

Lemma 3.4. For a connected two-layer, node-aligned multiplex network $\mathcal{M} = (V_M, E_M, V, L)$ with n nodes on each layer and inter-layer weight p, the algebraic connectivity λ_2 of the multiplex satisfies $\lambda_2 \leq 2p$.

Proof. The Laplacian matrix of the network of layers of the multiplex $\mathcal{M}$ is

$$L^{(L)} = \begin{pmatrix} p & -p \\ -p & p \end{pmatrix}. \tag{3.15}$$

The eigenvalues of this Laplacian are $\mu_1 = 0, \mu_2 = 2p$. From (2.2) we get

$$\lambda_2 \leq 2p.$$

□

Lemma 3.5. For a connected two-layer, node-aligned multiplex network $\mathcal{M} = (V_M, E_M, V, L)$ with n nodes on each layer and inter-layer weight p, if $\lambda_2^{(i)}$ is the algebraic connectivity of the intra-layer graph $\mathcal{G}_i$, then the algebraic connectivity λ_2 of the multiplex satisfies $\lambda_2 \leq \lambda_2^{(i)} + p$ for $i = 1, 2$.

Proof. Using Equation 1.10 and the fact that for a connected network its algebraic connectivity is positive, we know that for any non-zero eigenvector v of the supra-Laplacian matrix of the multiplex, the condition

$$v^T \overline{\mathcal{L}}(p) v - \lambda_2 \|v\|^2 \geq 0 \tag{3.16}$$

holds.

Let us write $v^T = (v_1^T, v_2^T)$, where v_1 and v_2 are n-dimensional vectors. Consider the normalized eigenvector u_1 of $L^{(1)}$ corresponding to $\lambda_2^{(1)}$, that is, $L^{(1)} u_1 = \lambda_2^{(1)} u_1$ with $\|u_1\| = 1$.

Expanding the expression for $v^T \overline{\mathcal{L}}(p) v$ we have

$$v^T \overline{\mathcal{L}}(p) v = v_1^T (L^{(1)} + p\mathbf{I}_n) v_1 + v_2^T (L^{(2)} + p\mathbf{I}_n) - 2p v_1^T v_2. \tag{3.17}$$

Consider the following decomposition of v_1

$$v_1 = \beta u_1 + y_1 \tag{3.18}$$

where $\beta \in \mathbb{R}$ and $u_1^T y_1 = 0$.

Using this decomposition, we compute

$$v_1^T (L^{(1)} + p\mathbf{I}_n) v_1 = \beta^2 (\lambda_2^{(1)} + p) + y_1^T (L^{(1)} + p\mathbf{I}_n) y_1. \tag{3.19}$$

Additionally, from the norm of v and Equation 3.18, we have

$$\|v\|^2 = \|v_1\|^2 + \|v_2\|^2 = \beta^2 + y_1^T y_1 + \|v_2\|^2. \tag{3.20}$$

Substituting Equations 3.19 and 3.20 into Inequality 3.16, we obtain

$$\beta^2 (\lambda_2^{(1)} + p) + \ldots - \lambda_2 (\beta^2 + y_1^T y_1 + \|v_2\|^2) \geq 0. \tag{3.21}$$

Since this inequality must hold for all $\beta \in \mathbb{R}$, the coefficient of β^2 must be non-negative. Therefore we obtain:

$$\lambda_2 \leq \lambda_2^{(1)} + p.$$

The inequality for $\lambda_2^{(2)}$ is obtained analogously.

□

Lemma 3.6. For a connected two-layer, node aligned multiplex network $\mathcal{M} = (V_M, E_M, V, L)$ with n nodes on each layer and inter-layer weight p, let $L^{(1)}$ and $L^{(2)}$ be the Laplacian matrices of the intra-layer graphs $\mathcal{G}_1$ and $\mathcal{G}_2$, respectively. If λ_A is the algebraic connectivity of the average graph $\frac{1}{2}(\mathcal{G}_1 + \mathcal{G}_2)$, then the connectivity of the multiplex λ_2 satisfies $\lambda_2 \leq \lambda_A$

Proof. The Laplacian matrix of the aggregate network of the multiplex $\mathcal{M}$ is

$$L^{(A)} = \frac{1}{2}(L^{(1)} + L^{(2)}), \tag{3.22}$$

which corresponds to the Laplacian matrix of the average graph $\frac{1}{2}(\mathcal{G}_1 + \mathcal{G}_2)$.

If $\lambda_2^{(A)}$ is the smallest non-zero eigenvalue of $L^{(A)}$, from (2.4) we get

$$\lambda_2 \leq \lambda_2^{(A)} = \lambda_A.$$

□

These upper bounds for the algebraic connectivity of the multiplex network reveal different behaviors of the multiplex depending on the value of the inter-layer coupling p, the algebraic connectivity of each layer ($\lambda_2^{(1)}$ and $\lambda_2^{(2)}$), and the algebraic connectivity of the average graph (λ_A):

- Weak coupling ($p \leq \min(\lambda_2^{(1)}, \lambda_2^{(1)}, \lambda_A)$): the layers behave in a disconnected way, and the algebraic connectivity of the multiplex increases linearly as $2p$. The easiest way to disconnect the multiplex is by separating the multiplex into its two individual layers (deleting the inter-layer couplings).
- Intermediate coupling ($\min(\lambda_2^{(1)}, \lambda_2^{(2)}) < p < \lambda_A$): if $\min(\lambda_2^{(1)}, \lambda_2^{(2)}) < \lambda_A/2$, the system enters a transitional phase where the algebraic connectivity of the multiplex will be roughly given by $p + \min(\lambda_2^{(1)}, \lambda_2^{(2)})$. The easiest way to disconnect the multiplex is no longer by separating the layers, but by splitting the least connected layer (i.e., the layer with the smallest algebraic connectivity) and cutting off the inter-layer connections between one of those parts and the other layer.
- Strong coupling ($p > \lambda_A$): the algebraic connectivity of the multiplex is approximately that of the aggregate network of both layers. The system behaves as a single unified network, with no distinction between the individual layers.

Finally, note that if $\min(\lambda_2^{(1)}, \lambda_2^{(2)})$ is sufficiently small, the first regime (weak coupling) does not effectively exist.

Example 3.7. The following graphs illustrate three different scenarios on the behavior regions of a two-layer, node-aligned multiplex $\mathcal{M} = (V_M, E_M, V, L)$ with $n = 50$ nodes and different connectivity properties of the intra-layer graph structures, formed by Erdös-Renyi networks [18] with 50 nodes and different edge probabilities.

- The three behavioral phases are present. The second layer of the multiplex has a significantly lower connectivity than the first, satisfying $\lambda_2^{(2)} \leq \lambda_A/2$. However, $\lambda_2^{(2)}$ is sufficiently large to distinguish the region where $\lambda_2 \approx 2p$.

  To visualize this behavior on a randomly generated multiplex, the first layer is constructed as an Erdös-Renyi network with 50 nodes and an edge probability of 0.5, while the second layer is an Erdös-Renyi network with 50 nodes and an edge probability of 0.2.

  The resulting multiplex has $\lambda_2^{(1)} \approx 12.288$, $\lambda_2^{(2)} \approx 3.519$ and $\lambda_A \approx 10.195$.

  Figure 3.1 shows one transition point at $p = \lambda_2^{(2)} \approx 3.519$. Up from this point λ_2 gradually converges to λ_A.

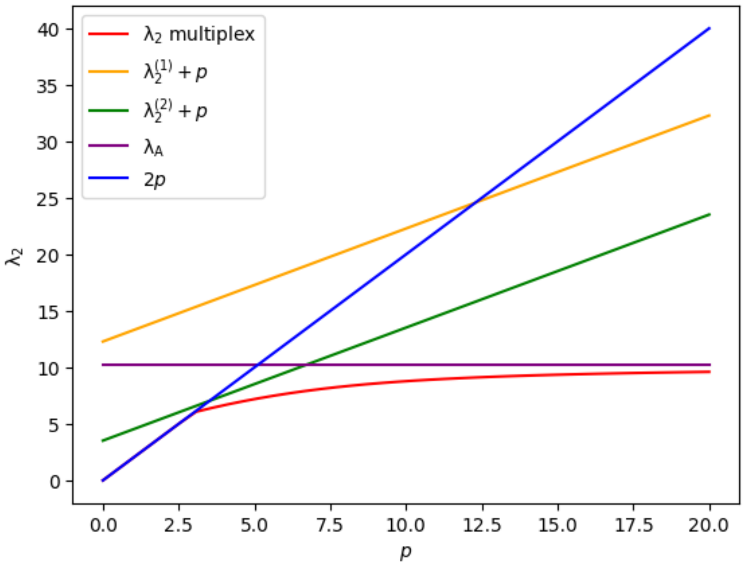


Figure 3.1: Plot of the second smallest eigenvalue for a two-layer multiplex as a function of the inter-layer weight p.

- There is no weak coupling. The second layer of the multiplex has a significantly lower connectivity than the first, also satisfying $\lambda_2^{(2)} \leq \lambda_A/2$. However, $\lambda_2^{(2)}$ is so small that the region where $\lambda_2 \approx 2p$ is not distinguishable.

To visualize this behavior on a randomly generated multiplex, the first layer is constructed as an Erdös-Renyi network with 50 nodes and an edge probability of 0.5, while the second layer is an Erdös-Renyi network with 50 nodes and an edge probability of 0.005.

The resulting multiplex has $\lambda_2^{(1)} \approx 16.043$, $\lambda_2^{(2)} = 0.212$ and $\lambda_A = 8.359$.

Figure 3.2 shows no appreciable transition point. Instead, λ_2 smoothly converges to λ_A.

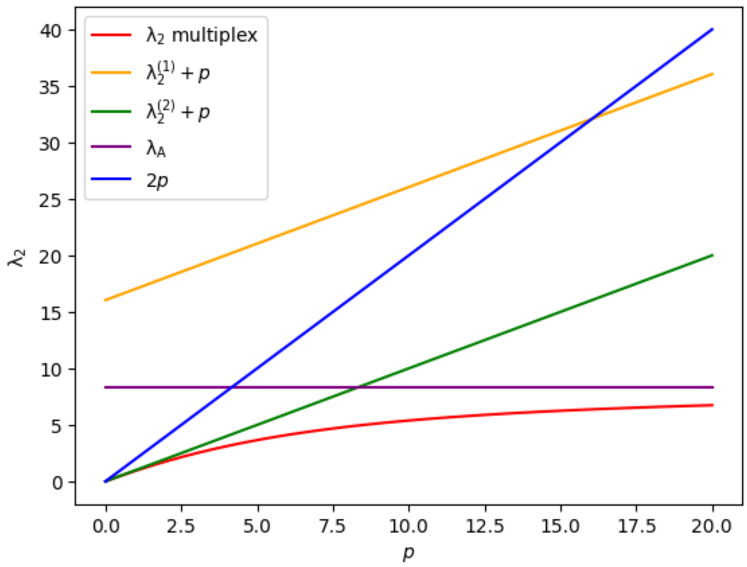


Figure 3.2: Plot of the second smallest eigenvalue for a two-layer multiplex as a function of the inter-layer weight p.

- There is no intermediate coupling. The second layer of the multiplex has a connectivity comparable to that of the first, satisfying $\lambda_2^{(2)} > \lambda_A/2$. Therefore, there is an abrupt change in the behavior of λ_2 with no transitional phase.

  To visualize this on a randomly generated multiplex, the first layer is constructed as an Erdös-Renyi network with 50 nodes and an edge probability of 0.5, while the second layer is an Erdös-Renyi network with 50 nodes and an edge probability of 0.4. The resulting multiplex has $\lambda_2^{(1)} \approx 13.807$, $\lambda_2^{(2)} = 10.137$ and $\lambda_A = 13.498$.

  Figure 3.3 shows a sharp transition point at $p = \lambda_A/2 \approx 6.749$.

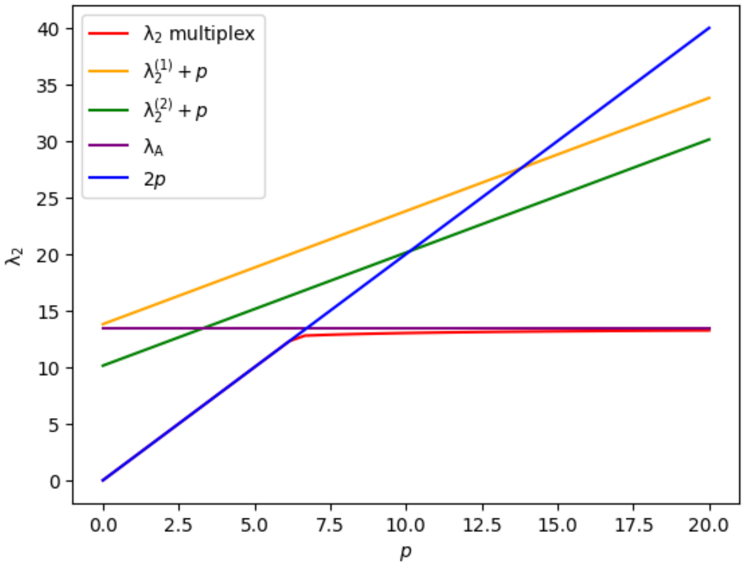


Figure 3.3: Plot of the second smallest eigenvalue for a two-layer multiplex as a function of the inter-layer weight p.

3.2.2 The value of poorly connected layers

The results from Example 3.7 show that when adding a significantly less connected network ($L^{(2)}$) to a well-connected one ($L^{(1)}$), the behavior of the multiplex depends on the less connected layer (for this reason, we will refer to it as the dominant layer) and the algebraic connectivity of the multiplex is, at best, as resilient as the aggregated network. However, the algebraic connectivity of the multiplex will not surpass the algebraic connectivity of the better-connected layer.

Nonetheless, there are situations where including a less connected network brings significant benefits to the overall multiplex. This is because the less connected network may possess traits that compensate for its low algebraic connectivity, hence enhancing the functionality of the multiplex. For example, consider a transportation multiplex network, where one layer consists of walking paths and the other of boat routes. While the boat routes layer may be less connected, it provides value to the general transportation multiplex by enabling faster movement, making it a crucial addition to the system.

To account for this added value provided by the less connected layer, we introduce a weighting factor α, which is applied to the edges of the poorer-connected layer, allowing us to adjust the relative importance of this layer, ranging form $\alpha = 1$, where both layers are treated as equal, to $\alpha > 1$, where the dominant layer is considered more advantageous or influential than the better-connected layer.

To study how this weighting factor influences the behavior of a two-layer multiplex, consider a multiplex network $\mathcal{M} = (V_M, E_M, V, L)$ with $m = 2$ layers, $L^{(1)}$ and $L^{(2)}$ with $\lambda_2^{(1)} > \lambda_2^{(2)}$. We study its behavior as a function of the importance of the dominant layer by introducing the weighting factor α to $L^{(2)}$. By doing so, Equation 3.11 takes the following form

$$\overline{\mathcal{L}} = \begin{pmatrix} L^{(1)} & 0 \\ 0 & \alpha L^{(2)} \end{pmatrix} + \mathcal{L}_C. \tag{3.23}$$

If we consider that the multiplex is node-aligned and that the coupling-edges have a weight p, Equation 3.23 is explicitly written as

$$\overline{\mathcal{L}}(\alpha) = \begin{pmatrix} L^{(1)} + p\mathbf{I}_n & -p\mathbf{I}_n \\ -p\mathbf{I}_n & \alpha L^{(2)} + p\mathbf{I}_n \end{pmatrix}. \tag{3.24}$$

Lemma 3.8. For a connected two-layer, node-aligned multiplex network $\mathcal{M} = (V_M, E_M, V, L)$ with n nodes on each layer, if the intra-layer graph $\mathcal{G}_2$ has a weighting factor α, each inter-layer edge has a weight p and $\lambda_2^{(2)}$ is the algebraic connectivity of the intra-layer graph $\mathcal{G}_2$, then the algebraic connectivity λ_2 of the multiplex satisfies $\lambda_2 \leq \alpha\lambda_2^{(2)} + p$.

Proof. Using the same reasoning as in the proof of Lemma 3.5, Inequality 3.16 holds.

Let us write $v^T = (v_1^T, v_2^T)$, where v_1 and v_2 are n-dimensional vectors. Consider the normalized eigenvector u_2 of $L^{(2)}$ corresponding to $\lambda_2^{(2)}$, that is, $L^{(2)}u_2 = \lambda_2^{(2)}$, with $\|u_2\| = 1$.

Expanding the expression for $v^T\overline{\mathcal{L}}(\alpha)v$ we have

$$v^T\overline{\mathcal{L}}(\alpha)v = v_1^T L^{(1)} v_1 - 2pv_1^T v_2 + p\|v_1\|^2 + \alpha v_2^T L^{(2)} v_2 + p\|v_2\|^2. \tag{3.25}$$

Consider the following decomposition of v_2

$$v_2 = \beta u_2 + y_2 \tag{3.26}$$

where $\beta \in \mathbb{R}$ and $u_2^T y_2 = 0$.

Using this decomposition, we compute

$$\alpha v_2^T L^{(2)} v_2 = \alpha\beta^2\lambda_2^{(2)} + \alpha y_2^T L^{(2)} y_2. \tag{3.27}$$

Additionally, from the norm of v and Equation 3.26, we have

$$\|v\|^2 = \|v_1\|^2 + \|v_2\|^2 = \|v_1\|^2 + \beta^2 + y_2^T y_2. \tag{3.28}$$

Substituting Equations 3.27 and 3.28 into Inequality 3.16, we obtain

$$\beta^2(\alpha\lambda_2^{(2)} + p) + \ldots - \lambda_2(\beta^2 + y_2^T y_2 + \|v_1\|^2) \geq 0. \tag{3.29}$$

Since this inequality must hold for all $\beta \in \mathbb{R}$, the coefficient of β^2 must be non-negative. Therefore we obtain:

$$\lambda_2 \leq \alpha\lambda_2^{(2)} + p.$$

□

To study the different behavioral phases of the multiplex as a function of the weighting factor α, note that Lemma 3.4 also serves as an upper bound and, so the algebraic connectivity of the multiplex is upper bounded by $\lambda_2 \leq 2p$. Therefore, we can distinguish two different behaviors with a transition point

$$\alpha = \frac{p}{\lambda_2^{(2)}}.$$

If $p/\lambda_2^2 > 1$, for $1 \leq \alpha < p/\lambda_2^{(2)}$, the algebraic connectivity of the multiplex will be upper-bounded by $\alpha\lambda_2^{(2)} + p$, and the best way to disconnect the multiplex is by splitting the dominant layer and cutting off the inter-layer connections between one of those parts and the other layer.

For $\lambda_2^{(2)} \geq p/\lambda_2^{(2)}$, the algebraic connectivity of the multiplex will be the same no matter the relative importance of the dominant layer, and the best way to disconnect the multiplex is by separating it into its two individual layers.

Note that, as $\alpha \geq 1$, if $p/\lambda_2^{(2)} < 1$ the multiplex will only show the second behavior.

Example 3.9. For a two-layer node-aligned multiplex $\mathcal{M} = (V_M, E_M, V, L)$ with $n = 50$ nodes and no inter-layer weights (i.e., $p = 1$) formed by an Erdös-Renyi network with an edge probability of 0.5 in the first layer and an Erdös-Renyi network with an edge probability of 0.1 in the second one. The resulting multiplex has $\lambda_2^{(1)} \approx 15.109$ and $\lambda_2^{(2)} \approx 0.498$. The behavior of its algebraic connectivity as a function of the weight factor α is plotted below. Since $\lambda_2^{(2)} < 1$, Figure 3.4 shows a sharp transition point at $\alpha = 1/\lambda_2^{(2)} \approx 2.008$, as expected.

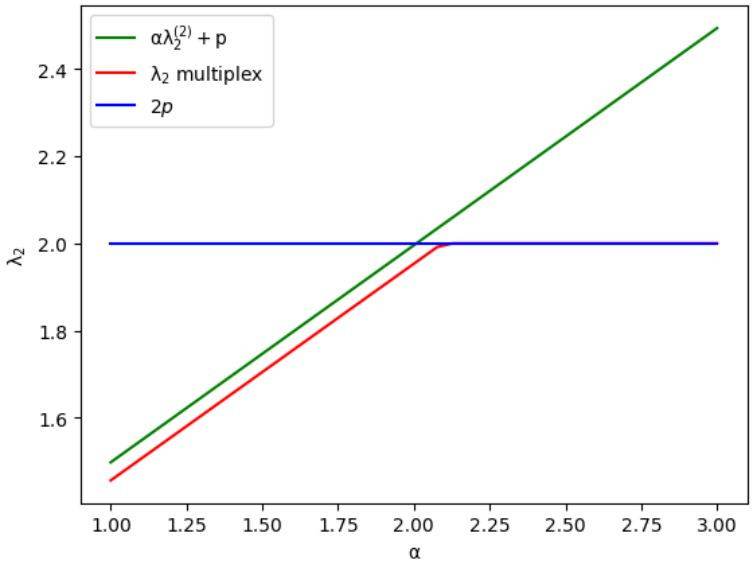


Figure 3.4: Plot of the second smallest eigenvalue for a two-layer multiplex as a function of the weighting factor α of the intra-layer graph $\mathcal{G}_2$.

3.2.3 Convenience of adding a poorly connected layer

As shown, whether the dominant layer enhances the algebraic connectivity of the multiplex depends on two key factors: the inter-layer weight, p, and the weighting factor α. To evaluate the conditions under which the algebraic connectivity of the multiplex, λ_2, exceeds that of the well-connected layer, $\lambda_2^{(1)}$, it is useful to plot the region in the parameter space of p and α where $\lambda_2 > \lambda_2^{(1)}$.

This visualization facilitates the comparison of different scenarios, each defined by distinct characteristics of the layers that conform the multiplex network, and evaluating when maintaining the entire multiplex network is more advantageous than preserving only the well-connected layer.

To understand the frontier of this region for a two-layer multiplex network, it is convenient to use the bounds previously derived in Subsections 3.2.1 and 3.2.2:

- Limit for p. Since $\lambda_2 \leq 2p$ is always an upper bound for the algebraic connectivity of the multiplex, on the boundary of the region (i.e. the values of p and α where $\lambda_2 = \lambda_2^{(1)}$) we get

$$p \geq \frac{\lambda_2^{(1)}}{2}. \tag{3.30}$$

- Limit for α. Using Lemma 3.6, $\lambda_2 \leq \lambda_2^{(A)}$, where $\lambda_2^{(A)}$ is the algebraic connectivity of the aggregate network. Considering that the supra-Laplacian matrix has the form given in Equation 3.24, the Laplacian matrix of the aggregate network is

$$L^{(A)} = \frac{1}{2}(L^{(1)} + \alpha L^{(2)}), \tag{3.31}$$

where $L^{(1)}$ and $L^{(2)}$ are the Laplacian matrices of the layers that conform the multiplex. Since $\alpha \geq 1$, we can rewrite

$$L^{(A)} = \frac{1}{2}(L^{(1)} + L^{(2)}) + \frac{\alpha - 1}{2} L^{(2)}. \tag{3.32}$$

The first component of the sum corresponds to the Laplacian matrix of the average graph $\frac{1}{2}(\mathcal{G}_1 + \mathcal{G}_2)$, where $\mathcal{G}_1$ and $\mathcal{G}_2$ are the intra-layer graphs associated with $L^{(1)}$ and $L^{(2)}$, respectively.

Therefore, the eigenvalues of $L^{(A)}$ are those of the Laplacian matrix of the average graph, with an added perturbation that depends only on α. Since obtaining the explicit expression of the algebraic connectivity of $L^{(A)}$, $\lambda_2^{(A)}$, can be cumbersome, approximations for values of α close to one and its asymptotic behavior can be derived using perturbation theory or approaches similar to those in [19] and [15].

Since the boundary region $\lambda_2 = \lambda_2^{(1)}$ is often achieved for intermediate values of α (i.e., neither $\alpha \approx 1$ nor $\alpha \gg 1$), these analytical approximations are not usually useful enough and numerical approximations can provide a more accurate characterization of the bound for α.

Example 3.10. To visualize the different types of region shapes obtained for a two-layer node-aligned multiplex network $\mathcal{M} = (V_M, E_M, V, L)$ on $n = 20$ nodes, the following cases, randomly generated using Erdös-Renyi networks on each layer, are analyzed and plotted in Figure 3.5.

- 1. Similar values for $\lambda_2^{(1)}$, $\lambda_2^{(2)}$ and λ_A: both layers have similar connectivity properties (edge probability: 0.5). The obtained layers have the following algebraic connectivities: $\lambda_2^{(1)} \approx 2.4909$, $\lambda_2^{(2)} \approx 3.4894$, and the algebraic connectivity of the average graph is $\lambda_A \approx 3.3942$.

- 2. $\lambda_2^{(2)} < \lambda_A \lessapprox \lambda_2^{(1)}$: the second layer is less connected than the first one, but the average graph has a similar connectivity to the first one (edge probability of layer 1: 0.5, edge probability of layer 2: 0.3). The obtained layers have the following algebraic connectivities: $\lambda_2^{(1)} \approx 5.1765$, $\lambda_2^{(2)} \approx 1.5933$, and $\lambda_A \approx 4.7554$.

- 3. $\underline{\lambda_2^{(2)} < \lambda_A < \lambda_2^{(1)}}$: the second layer is significantly less connected than the first one, so the average graph has a smaller connectivity than the first layer (edge probability of layer 1: 0.5, edge probability of layer 2: 0.15). The layers of the represented regions have the following algebraic connectivities: $\lambda_2^{(1)} \approx 5.8873$, $\lambda_2^{(2)} \approx 0.5067$, and $\lambda_A \approx 3.9830$.

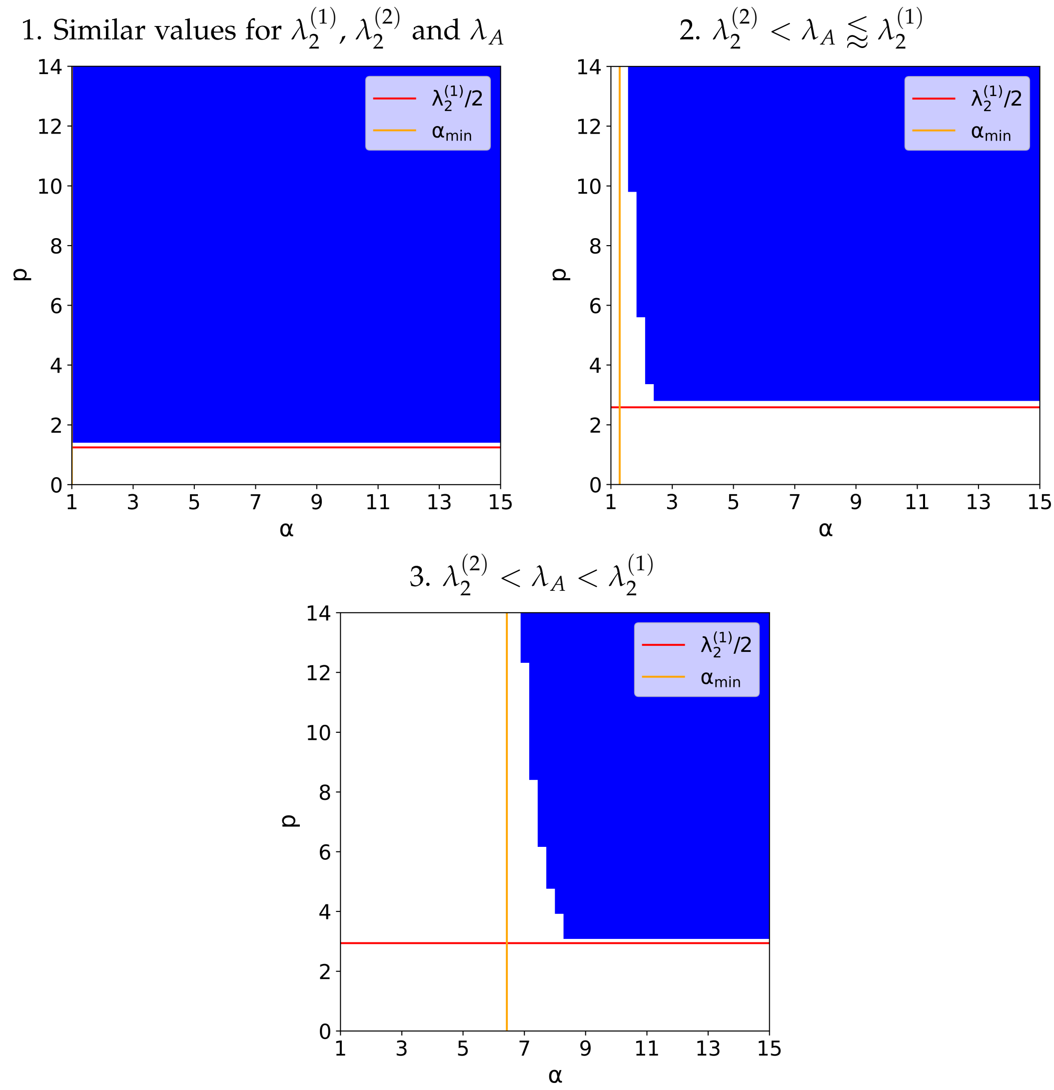


Figure 3.5: Algebraic connectivity figures of two-layer multiplexes with different connectivity properties. The colored region means an improvement in algebraic connectivity when adding a poorly connected layer as a function of the weighting factor α of the poorly connected layer and the inter-layer weight p.

The regions where it is beneficial to maintain the entire multiplex (i.e., where $\lambda_2 > \lambda_2^{(1)}$) are shown in blue. The plots also include the bounds of the regions $p = \lambda_2^{(1)}/2$ and $\alpha = \alpha_{\text{min}}$, where α_{min} is numerically computed from the eigenvalues of Equation 3.32.
As one would expect, the shaded region is increasingly smaller as the second layer becomes comparatively poorly connected.

3.3 Supra-walks

In a similar manner to the definition for simple networks (see Definition 1.3), we can define a walk in a node-aligned multiplex network $\mathcal{M} = (V_M, E_M, V, L)$ with n nodes on each layer, and call it a *supra-walk*. A supra-walk is an ordered list of nodes in which either before of after each intra-layer step, a walk can either continue on the same layer or switch to an adjacent layer. The choice of wether to stay on the same layer or change layers is represented by the matrix

$$\hat{\mathcal{C}}(\beta, \gamma) = \beta \mathbf{I}_n + \gamma \mathcal{C} \tag{3.33}$$

where $\mathbf{I}_n$ is the $n \times n$ identity matrix, $\mathcal{C}$ is the adjacency matrix of the coupling graph, $\beta > 0$ is a weight that accounts for the walk staying within the current layer, and $\gamma > 0$ is a weight that accounts for the walk stepping to another layer.

In a supra-walk, a *supra-step* consists either of a single intra-layer step or of a step that includes both an intra-layer step and changing from one layer to another (either before or after having an intra-layer step), disallowing two consecutive inter-layer steps. In other words, supra-walks are walks on the supra-graph $\mathcal{G}_M$ with the restriction that there cannot be two consecutive inter-layer steps.

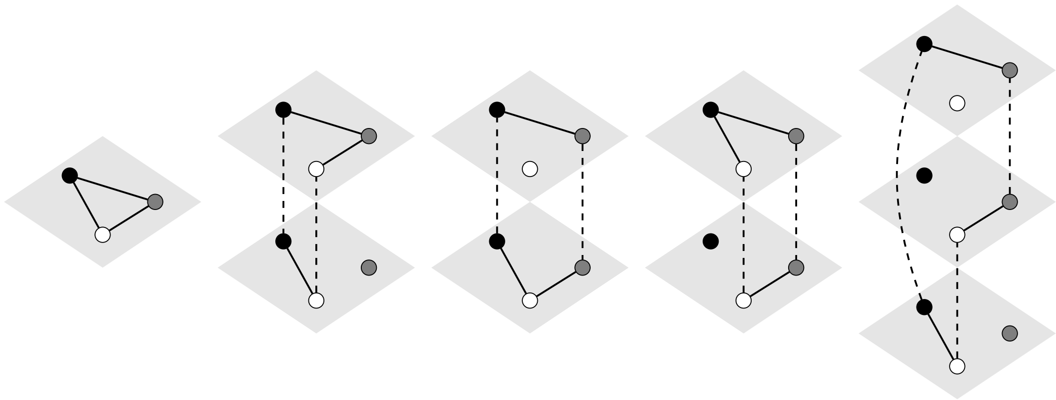
To encode the permissible steps in a multiplex network, a *multiplex walk matrix* is used. Depending on the order of the inter-layer and intra-layer steps, different multiplex walk matrices are defined:

- The matrix $\mathcal{A}\hat{\mathcal{C}}$ represents steps where either only an intra-layer step is performed (represented by the matrix $\mathcal{A}$) or an inter-layer step is made first, followed by an intra-layer step (represented by the matrices $\mathcal{C}\mathcal{A}$).

- The matrix $\hat{\mathcal{C}}\mathcal{A}$ represents steps where either only an intra-layer step is performed (represented by the matrix $\mathcal{A}$) or an intra-layer step is made first, followed by an inter-layer step (represented by the matrices $\mathcal{A}\mathcal{C}$).

Note that, by definition, the supra-adjacency matrix $\overline{\mathcal{A}}$ is also a walk matrix, with no restriction on the type of steps that are performed in the multiplex. Still, it is often of interest to treat intra and inter-layer edges differently, as changing a layer is an action of different nature with respect to moving between nodes in the same layer.

Example 3.11. A sketch of the possible allowed elementary cycles (i.e., the minimum possible paths where only the first and last nodes are equal) for the multiplex walk matrix $\mathcal{A}\hat{\mathcal{C}}$, going from black - grey - white - black nodes, is the following one:



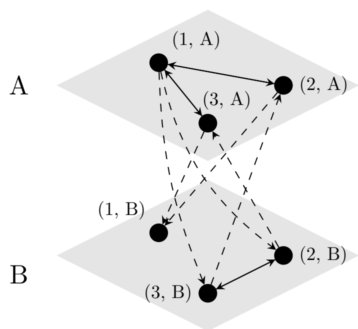
$\mathcal{AAA}$ $\mathcal{AACAC}$ $\mathcal{ACAAC}$ $\mathcal{ACACA}$ $\mathcal{ACACAC}$

The intra-layer edges are represented by solid lines and the inter-layer edges are represented by dotted lines.

Note that both matrices $\mathcal{A}\hat{\mathcal{C}}$ and $\hat{\mathcal{C}}\mathcal{A}$ can be interpreted as adjacency matrices of directed (and possibly weighted, if $\beta \neq 1$ or $\gamma \neq 1$) graphs $\mathcal{G}_M$ that correspond to the underlying graph of directed multilayer networks. The asymmetry introduced by the directed edges and potentially different weights ($\beta \neq \gamma$) may result in complex eigenvalues for these matrices.

Example 3.12. Consider the multiplex network used in Example 3.2. If the multiplex walk matrix $\mathcal{A}\hat{\mathcal{C}}$ is used with $\beta = \gamma = 1$, we obtain the following multiplex walk matrix (left) and a representation of its corresponding directed multilayer network (right)

$$\mathcal{A}\hat{\mathcal{C}} = \left(\begin{array}{ccc|ccc} 0 & 1 & 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 \end{array}\right)$$



If instead $\beta = 1$ and $\gamma = 2$ are used, the resulting multiplex walk matrix becomes

$$\mathcal{A}\hat{\mathcal{C}} = \left(\begin{array}{ccc|ccc} 0 & 1 & 1 & 0 & 2 & 2 \\ 1 & 0 & 0 & 2 & 0 & 0 \\ 1 & 0 & 0 & 2 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 1 \\ 0 & 2 & 0 & 0 & 1 & 0 \end{array}\right).$$

The eigenvalues of this matrix are: $\lambda_1 = -1$, $\lambda_2 \approx -0.814 - 1.272\,i$, $\lambda_3 \approx -0.814 + 1.272\,i$, $\lambda_4 = \lambda_5 = 0$, $\lambda_6 \approx 2.629$.

For directed networks, the previous definition of node degree used for an undirected network $\mathcal{G} = (V, E)$, which emerged in Equation 1.2 ($d_i = \sum_j a_{ij}$ for a node $v_i \in V$), is insufficient as it does not distinguish between edges entering and leaving a node. To address this issue, two measures are defined: the *indegree* of a node, which counts the sum of the weights of the edges pointing to it, and the *outdegree*, which counts the sum of the weights of the edges originating from it.

Definition 3.13. Let $\mathcal{G} = (V, E)$ be a directed network with adjacency matrix $A = (a_{ij})$. For a node $v_i \in V$, its indegree and outdegree are given by $d_i(i) = \sum_j a_{ji}$ and $d_o(i) = \sum_j a_{ij}$, respectively. Moreover, the *degree* of a node $v_i \in V$, $d(i)$, is the sum of its indegree and outdegree: $d(i) = d_i(i) + d_o(i)$.

If the Laplacian matrix is defined as $L_M = D_M - A_M$ (as in Section 2.2), where A_M is the adjacency matrix of $\mathcal{G}_M$ and $D_M = \mathrm{diag}(d(1), \ldots, d(n))$ is the diagonal supra-matrix containing the weighted directed degrees of $\mathcal{G}_M$, the resulting Laplacian matrix is generally asymmetric. This asymmetry can lead to complex eigenvalues, making this definition of Laplacian matrix unsuitable for defining algebraic connectivity in the context of multiplex-walk matrices.

Example 3.14. Consider the multiplex network used in Example 3.2. Using the Laplacian definition $L_M = D_M - A_M$ for the multilayer network derived from the multiplex walk matrix $\mathcal{A}\hat{\mathcal{C}}$ with $\beta = 1$ and $\gamma = 2$, we obtain

$$L_M = \left(\begin{array}{ccc|ccc} 8 & -1 & -1 & 0 & -2 & -2 \\ -1 & 6 & 0 & -2 & 0 & 0 \\ -1 & 0 & 6 & -2 & 0 & 0 \\ \hline 0 & 0 & 0 & 4 & 0 & 0 \\ 0 & 0 & -2 & 0 & 6 & -1 \\ 0 & -2 & 0 & 0 & -1 & 6 \end{array}\right).$$

The eigenvalues of this matrix are: $\lambda_1 \approx 3.847$, $\lambda_2 = 4$, $\lambda_3 = 6$, $\lambda_4 = 7$, $\lambda_5 \approx 7.576 - 0.549\,i$, $\lambda_6 \approx 7.576 + 0.549\,i$.

This example illustrates how the current Laplacian definition is inadequate for multiplex networks with constraints on inter-layer transitions in supra-walks, motivating the development of a new Laplacian matrix definition that accounts for these constraints.

3.3.1 Possible definitions for the Laplacian matrix

A natural first approach to define the Laplacian matrix of the multiplex network derived from the multiplex walk matrices is to consider the outdegrees of its underlying graph $\mathcal{G}_M$, obtaining with it the *outdegree Laplacian matrix*.

Definition 3.15. For a directed multilayer network $\mathcal{M} = (V_M, E_M, V, L)$ with underlying graph $\mathcal{G}_M$, its *outdegree Laplacian matrix*, L_o, is given by

$$L_o = D_o - A_M,$$

where A_M is the adjacency matrix of $\mathcal{G}_M$ and $D_o \text{diag}(d_o(1), \dots, d_o(n))$ is the diagonal supra-matrix containing the weighted outdegrees of $\mathcal{G}_M$.

Using this definition, the *algebraic connectivity of a directed graph* is defined as the following real number

$$\lambda_2(\mathcal{G}_{\mathcal{M}}) = \min_{\substack{v^T \mathbf{1}_n = 0 \\ v \in \mathbb{R}^n, v \neq 0}} \frac{v^T L_o v}{\|v\|^2}. \tag{3.34}$$

As extensively discussed in [20], this definition generalizes the algebraic connectivity to directed graphs, preserving several of the properties previously mentioned in Subsection 1.2.1.

Example 3.16. Consider the multiplex network used in Example 3.2. Using the Laplacian definition $L_o = D_o - A_M$, with the multilayer network derived from the multiplex walk matrix $\mathcal{A}\hat{\mathcal{C}}$ with $\beta = 1$ and $\gamma = 2$, we obtain

$$L_o = \left(\begin{array}{ccc|ccc} 6 & -1 & -1 & 0 & -2 & -2 \\ -1 & 3 & 0 & -2 & 0 & 0 \\ -1 & 0 & 3 & -2 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -2 & 0 & 3 & -1 \\ 0 & -2 & 0 & 0 & -1 & 3 \end{array}\right).$$

The eigenvalues of this matrix are: $\lambda_1 = 0$, $\lambda_2 = 1$, $\lambda_3 = 3$, $\lambda_4 = \lambda_5 = 4$, $\lambda_6 = 6$. Thus, the algebraic connectivity of the multiplex network derived from the multiplex walk matrix $\mathcal{A}\hat{\mathcal{C}}$ is $\lambda_2 = 1$.

Still, this definition may fail to completely capture the dynamics occurring in the multiplex walk matrix, as only the outdegrees are considered.

To overcome the possible limitations of L_o, we propose a Laplacian matrix derived from the *symmetrization* of $\mathcal{G}_M$, denoted as $\mathcal{G}_S$. The graph $\mathcal{G}_S$ is obtained by replacing each directed edge in $\mathcal{G}_M$ with an undirected edge of half the weight.

If A_M is the adjacency matrix of $\mathcal{G}_M = (V_M, E_M)$, then the adjacency matrix of $\mathcal{G}_S = (V_S, E_S)$ (where $V_S = V_M$) is defined as $A_S = 1/2(A_M + A_M^T)$. Since $\mathcal{G}_S$ is undirected, we can use the usual Laplacian matrix definition to define its Laplacian.

Definition 3.17. For a directed multilayer network $\mathcal{M} = (V_M, E_M, V, L)$ with underlying graph $\mathcal{G}_M$ and a symmetrized graph $\mathcal{G}_S$, its *symmetrized Laplacian matrix*, L_S is given by

$$L_S = D_S - A_S,$$

where $D_S = \text{diag}(d_1, \ldots, d_n)$, and d_i is the node degree of $v_i \in V_S$.

This definition ensures that all eigenvalues of L_S are real, so the original definition of algebraic connectivity (Equation 1.10) still applies.

Example 3.18. Consider the multiplex network used in Example 3.2. Using the Laplacian definition $L_S = D_S - A_S$ for the multilayer network derived from the multiplex walk matrix $\mathcal{A}\hat{\mathcal{C}}$ with $\beta = 1$ and $\gamma = 2$, we obtain

$$L_S = \left(\begin{array}{ccc|ccc} 4 & -1 & -1 & 0 & -1 & -1 \\ -1 & 3 & 0 & -1 & 0 & -1 \\ -1 & 0 & 3 & -1 & -1 & 0 \\ \hline 0 & -1 & -1 & 2 & 0 & 0 \\ -1 & 0 & -1 & 0 & 3 & -1 \\ -1 & -1 & 0 & 0 & -1 & 3 \end{array}\right).$$

The eigenvalues of this matrix are $\lambda_1 = 0$, $\lambda_2 = \frac{1}{2}(7 - \sqrt{13})$, $\lambda_3 = \frac{1}{2}(7 - \sqrt{5})$, $\lambda_4 = 4$, $\lambda_5 = \frac{1}{2}(7 + \sqrt{5})$, $\lambda_6 = \frac{1}{2}(7 + \sqrt{13})$. Thus, the algebraic connectivity of the multiplex network derived from the multiplex walk matrix $\mathcal{A}\hat{\mathcal{C}}$ is $\lambda_2 = \frac{1}{2}(7 - \sqrt{13}) \approx 1.697$.

This definition of the Laplacian matrix using the symmetric part of $\mathcal{G}_M$ has the advantage of accounting for both indegree and outdegree contributions, offering a more balanced view of the dynamics of the supra-walks in the multiplex network. However, some problems may arise as it also reduces the directed multilayer network to an undirected one, loosing part of the supra-walk constraints, potentially influencing the computation of the algebraic connectivity. For this reason, further investigation is needed to correctly evaluate its advantages and drawbacks and to assess its suitability for analyzing the dynamics of supra-walks in multiplex networks.

Chapter 4

Multiplex approach in ancient regional transport infrastructures

The goal of this chapter is to apply the algebraic connectivity results for multiplex networks obtained in Chapter 3 to study the structural properties of transportation networks in the Iron Age archaeological settlements of Latium Vetus and Southern Etruria located in the Italian peninsula. Using the multiplex formalism introduced in this work, we aim to analyze the viability and maintenance of road and river networks in these regions.

The use of network science in archaeology can be a powerful tool when employed correctly, but one must be aware of the inherent incompleteness of the data and the simplifications required when analyzing large time spans and tackling the complexity of human interaction, as discussed in [21]. Still, by following a correct approach to data processing, we can apply the theoretical results obtained in Section 3.2. The work started in [22] and [23], along with the multiplex approach made in [24] and [25] provide the foundations for this analysis.

4.1 Road and river networks

As with any aspect related to the complexity of human interactions, the distribution of settlements and the various relationships between them are dynamic and evolve over time. For example, new settlements are established, new roads are built, old ones are neglected, and the use and navigation of rivers change. Since obtaining precise information on these matters is difficult and a discretization of the time periods is necessary for a meaningful analysis of regional dynamics and change, [22] identifies five major periods during which settlements and connections remain stable without major changes:

- Early Iron Age 1 Early (EIA1E): 950/925-900 BC.
- Early Iron Age 1 Late (EIA1L): 900-850/825 BC.
- Early Iron Age 2 (EIA2): 850/825-730/720 BC.
- Orientalizing Age (OA): 730/720-580 BC.
- Archaic Period (AA): 580-500 BC.

This simplification allows us to focus on five time stamps during which the set of settlements and the different connections among them can be considered constant, reducing the analysis from a continuously evolving network of settlements to five static networks [22].

For the road networks in Latium Vetus (LV) we used the data presented in [23], and for the the road networks in Southen Etruria (SE) we used the data presented in [22]. Regarding the river networks, we used the data presented in [26].

In both the road and river networks, each node represents a settlement, and its location corresponds to the settlement's geographic coordinates. For the road network, an edge exists between two nodes if there is archaeological evidence of a road directly connecting their corresponding settlements without passing through another settlement. For the river network, a settlement is included if it was located near a river, and an edge exists between two nodes if geological evidence allows inferring the navigability of the river between those two settlements without passing through another settlement, even if there is no direct evidence of its use for transportation purposes, making this model an upper bound of the real river network [24]. Short navigable distances between coastal settlements have also been included as edges in the river network. Figure 4.1 illustrates the resulting road and river networks using this criteria for the EIA1L period.

4.1.1 Network quantification

Both the road and river networks are treated as weighted networks, where the weight is proportional to the reliance of each road or river connecting a settlement. The unit of quantifying the cost of the maintenance of the infrastructure is calculated based on the cost of maintaining one kilometer of road, thus the cost of a road is proportional to its length (in kilometers) and the reliance* of a road between two settlements is inversely proportional to the distance between them. In estimating the distances of road or river routes, the geodesic distance between the nodes they connect has been considered.

*We understand as reliance of a road or river the "probability" of the connection not failing: the shorter the connection the more secure it is.

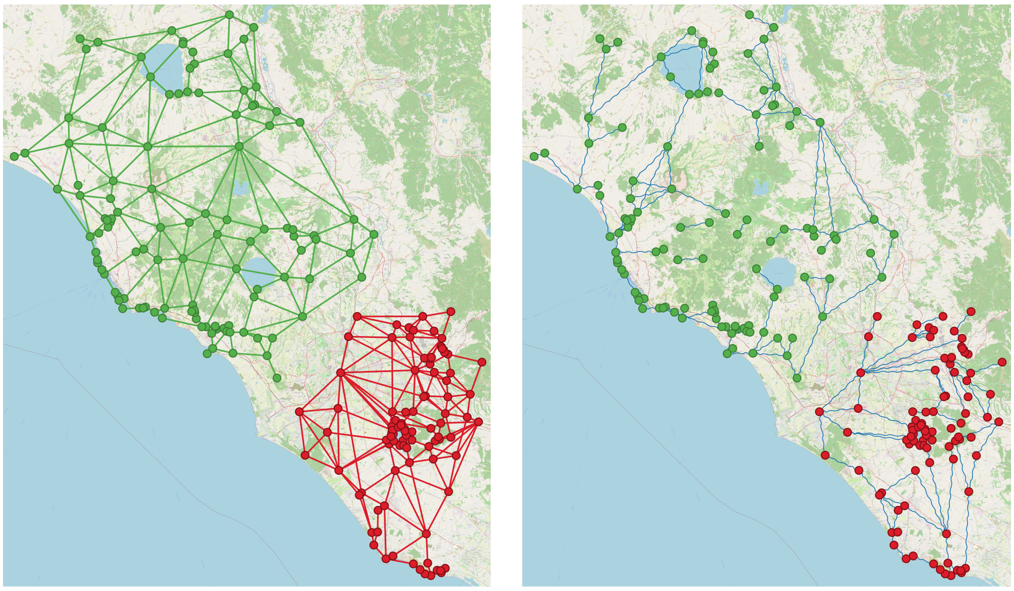


Figure 4.1: Road networks (left) and river networks (right) of Southern Etruria (upper region, green nodes) and Latium Vetus (lower region, red nodes) for EIA1L period.

While this approach may lack precision, it is suitable for our analysis, as it avoids introducing biases toward less-studied areas, while aligning with the focus of this study, which is to understand the general structural properties of the system, rather than focusing on the specific details of individual paths [22].

As introduced in Section 1.2, both the node and edge connectivity of a network serve as indicators of its robustness: the greater the connectivity, the more resistant the network is to node or edge failures. Moreover, the algebraic connectivity of the network provides a lower bound for both of these values, as well as setting the time scale for diffusion processes in the network, representing the connection between the structural and dynamical robustness of a network [16].

To quantify the characteristics of each network, we calculated the algebraic connectivity of the greatest connected component[†] for each transportation mode and period. Table 4.1 shows how the river networks, as one would expect given that new rivers generally cannot be created, have poorer connectivity properties compared to the road networks, as the algebraic connectivity of the river networks are generally one order of magnitude smaller than those of the road networks.

[†]We work with the greatest connected component to satisfy the premises of the results discussed in Section 1.2.

Road	**EIA1E**	**EIA1L**	**EIA2**	**OA**	**AA**
SE λ_2^*	$8.06 \cdot 10^{-3}$	$8.25 \cdot 10^{-3}$	$7.17 \cdot 10^{-3}$	$5.31 \cdot 10^{-3}$	$5.23 \cdot 10^{-3}$
LV λ_2	$1.16 \cdot 10^{-2}$	$1.03 \cdot 10^{-2}$	$1.44 \cdot 10^{-2}$	$2.29 \cdot 10^{-2}$	$2.99 \cdot 10^{-2}$
River	**EIA1E**	**EIA1L**	**EIA2**	**OA**	**AA**
SE λ_2^*	$3.1 \cdot 10^{-4}$	$4.8 \cdot 10^{-4}$	$1.15 \cdot 10^{-3}$	$2.6 \cdot 10^{-4}$	$2.2 \cdot 10^{-4}$
LV λ_2^*	$2.3 \cdot 10^{-3}$	$1.9 \cdot 10^{-3}$	$1.0 \cdot 10^{-3}$	$2.1 \cdot 10^{-3}$	$1.6 \cdot 10^{-3}$

Table 4.1: Algebraic connectivity for empirical road and river networks of Latium Vetus (LV) and Southern Etruria (SE). $*$ accounts for the algebraic connectivity of the giant connected component.

4.1.2 Convenience of adding the river network

These poor connectivity characteristics of the river networks compared to the road networks motivate the study of the viability and likelihood of maintaining the river network, following the same approach detailed in Subsections 3.2.2 and 3.2.3.

We model these networks using a two-layer multiplex network where, using the same notation as in Chapter 3, $\mathcal{G}_1$ corresponds to the intra-layer graph of the road network and $\mathcal{G}_2$ corresponds to the intra-layer graph of the river network. Each inter-layer edge of the multiplex $\mathcal{M}$, which physically can be conceived as a *port*, i.e. a connection linking the road and the river, is assigned a weight p. As shown in Table 4.1, the river layer is the dominant layer ($\lambda_2^{\text{river}} < \lambda_2^{\text{road}}$). However, since the archaeological record provides evidence of the existence of the river network, we introduce the weighting factor α for the river intra-layer network $\mathcal{G}_2$, which we will call the *river convenience factor*, which scales the maintenance of waterways relative to the unitary cost of road maintenance: the higher the value of α, the lower the maintenance cost of rivers compared to roads [25]. A high river convenience factor increases the weights of river connections, effectively increasing the algebraic connectivity of $\mathcal{G}_2$ as $\alpha\lambda_2^{\text{river}}$. Thus $\lambda_2^{\text{min}}(\alpha) = \min(\alpha\lambda_2^{\text{river}}, \lambda_2^{\text{road}})$, giving us a transition marked by

$$\overline{\alpha} = \frac{\lambda_2^{\text{road}}}{\lambda_2^{\text{river}}} \tag{4.1}$$

at which the river network can be considered more resilient than the road network. Beyond this value of α, the addition of the river layer no longer contributes to the system's connectivity as a weaker layer. Lower values of $\overline{\alpha}$ indicate more similarity in the quality of both road and river networks, thus a high value of $\overline{\alpha}$ would suggest the need of a highly efficient maintenance of ports and rivers, as the connectivity of the river layer is significantly smaller than that of the road. As Table 4.1.2 shows, the values of $\overline{\alpha}$ are generally higher for the Etruscan region, indicating that the fluvial infrastructure could hinder terrestrial transportation [25].

Region	**EIA1E**	**EIA1L**	**EIA2**	**OA**	**AA**
SE $\overline{\alpha}$	26.0	17.2	6.2	21.2	23.8
LV $\overline{\alpha}$	5.0	5.4	14.4	10.9	18.7

Table 4.2: $\overline{\alpha}$ values derived form the empirical road and river networks of Latium Vetus (LV) and Southern Etruria (SE).

Furthermore, from Subsection 3.2.3, the following bound for p of the frontier of the region holds

$$p \geq \frac{\lambda_2^{\text{road}}}{2} \tag{4.2}$$

Note that the inter-layer edge weight p accounts for the relative strength of the inter-layer coupling, and, following the reasoning made in Subsection 4.1.1 we consider that $1/p$ is a parameter which quantifies the expense of maintaining port infrastructures relative to the cost of maintaining one kilometer of road [25].

Considering that the algebraic connectivity of the multiplex network formed by the road and river layers (λ_2) depends on λ_2^{road}, λ_2^{river}, α and p, and that the maintenance of the joint network is advantageous only if the algebraic connectivity of the multiplex exceeds that of the road network alone, since λ_2^{road} and λ_2^{river} are constant for each period, the relevance of maintaining the river network depends entirely on the factors α and p.

Although the exact values of α and p are unknown and likely varied across different regions and periods, our analysis aims to evaluate under what specific circumstances the addition of the river routes, as an auxiliary mode, enhances the overall transportation system effectiveness beyond what terrestrial routes could achieve on their own, using α and $1/p$ as indicative factors.

Figure 4.2 illustrates the regions in the parameter space of $1/p$ and α where $\lambda_2 > \lambda_2^{\text{road}}$ for Southern Etruria and Latium Vetus during the five different periods considered. A larger colored area means good interplay between layers.

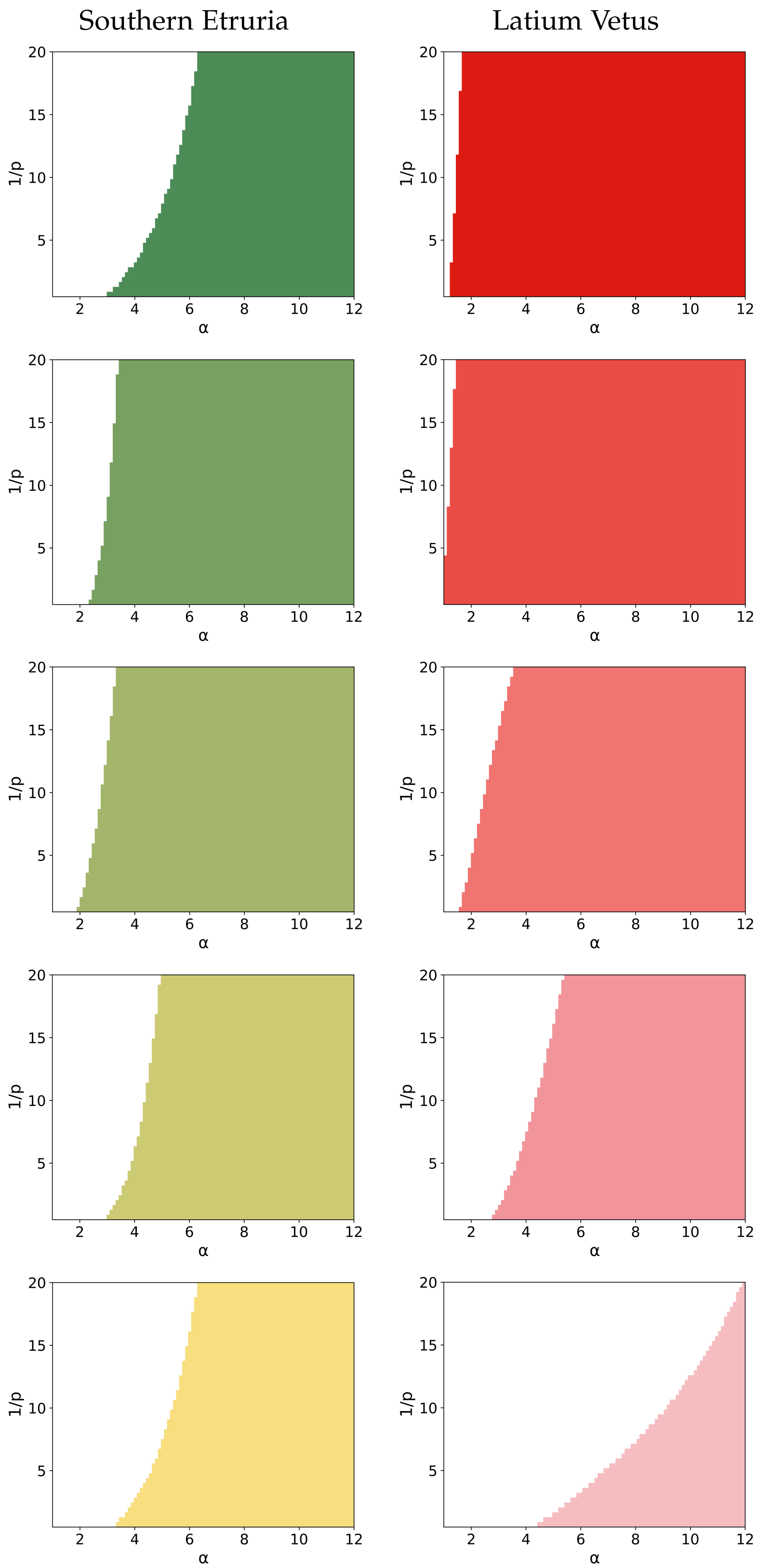


Figure 4.2: Algebraic connectivity figures of road and river multiplexes for periods EIA1E, EIA1L, EIA2, OA and AA from top to bottom. The colored region means an improvement in algebraic connectivity when adding the river layer as a function of the river convenience factor, α, and the port cost, $1/p$ (with p the inter-layer weight).

The y-axis represents the port maintenance costs, $1/p$, with values plotted in the range $[0.5, 20]$. This interval ensures visibility of the most relevant dynamics, as the smallest theoretical upper bound $1/p \leq 2/\lambda_2^{\text{road}}$, derived from Table 4.1 is much higher ($1/p \approx 60$ for LV AA).

In Latium Vetus, Figure 4.2 reveals that good interplay between rivers and roads was most beneficial during the initial stages of Iron Age (EIA1E and EIA1L). Over time, as terrestrial infrastructure improved, reliance on fluvial connections decreased and, by the Archaic Period (AA), the terrestrial infrastructure had become sufficiently resilient by itself, greatly reducing the critical role that the fluvial network had had in the past. In contrast for the Soutern Etruria region, while the river networks consistently show poor connection qualities, the terrestrial connections were also never strong enough on their own, thus the river networks consistently contributed to the overall network connectivity (in a variable way) across the years.

Since river infrastructure cannot be constructed, its impact greatly depends on natural geography. In Latium Vetus, where rivers naturally provided better connections, the fluvial network played a key role in early stages when road infrastructure was not yet properly developed. However, as the Latin articulation of the territory became stronger, partly by developing a robust road network, the importance of rivers diminished significantly. In contrast, the Southern Etruria region had weaker river connections, so it never depended as heavily on fluvial connections. Nonetheless, the river network remained relevant because the road infrastructure did not develop as rapidly or effectively as in Latium Vetus, due to geographical and political factors beyond the scope of this work.

These results are consistent with previous studies were the multiplex formalism was not used [22] [23], confirming archaeological theories regarding the importance of road and river infrastructure in vertebrating these Italian regions and open the door to applying the multiplex framework to other historical or archaeological contexts with more limited material evidence.

Bibliography

[1] Borgatti, S. P., Mehra, A., Brass, D. J., & Labianca, G. (2009). Network Analysis in the Social Sciences. *Science, 323*(5916), 892-895.

[2] Prignano, L., Font-Pomarol, L., Morer, I., & Lozano, S. (2023). Infrastructures connecting people: A mechanistic model for terrestrial transportation networks. *Environment and Planning B: Urban Analytics and City Science, 50*(8), 2254-2272.

[3] Ideker, T., & Sharan, R. (2008). Protein networks in disease. *Genome Research, 18*(4), 644-652.

[4] Segre, S. (2004). A Durkheimian Network Theory. *Journal of Classical Sociology, 4*(2), 215-235.

[5] Kivelä, M., Arenas, A., Barthelemy, M., Gleeson, J. P., Moreno, Y., & Porter, M. A. (2014). Multilayer networks. *Journal of Complex Networks*, *2*(3), 203-271.

[6] Li, Z., Yuan, C., Tang, J., Zhu, K., & Pan, X. (2023). Understanding the mobility of public transport systems based on weighted multiplex networks. *Physica A: Statistical Mechanics and its Applications, 624*, 128977.

[7] Wang, W., Liu, Q.H., Cai, S.M., Tang, M., Braunstein, L. A., & Stanley, H. E. (2016). Suppressing disease spreading by using information diffusion on multiplex networks. *Nature Scientific Reports, 6*(1), 29259.

[8] Newman, M. E. J. (2010). *Networks*. Oxford University Press.

[9] Fiedler, M. (1989). Laplacian of graphs and algebraic connectivity. *Banach Center Publications*, 25(1), 57-70.

[10] Sánchez-García, R. J., Cozzo, E., & Moreno, Y. (2014). Dimensionality reduction and spectral properties of multilayer networks. *Physical Review E*, 89(5), 052815.

[11] Haemers, W. H. (1995). Interlacing eigenvalues and graphs. *Linear Algebra and its Applications*, 226-228, 593-616.

[12] De Domenico, M., Solé-Ribalta, A., Cozzo, E., Kivelä, M., Moreno, Y., Porter, M. A., Gómez, S., & Arenas, A. (2013). Mathematical Formulation of Multilayer Networks. *Physical Review E, 3*(4), 041022.

[13] Pozrikidis, C. (2014). *An Introduction to Grids, Graphs, and Networks*. Oxford University Press.

[14] Cozzo, E., Ferraz de Arruda, G., Aparecido Rodrigues, F., & Moreno, Y. (2018). *Multiplex Networks. Basic Formalism and Structural Properties.* Springer.

[15] Solé-Ribalta, A., Domenico, M. D., Kouvaris, N. E., Diaz-Guilera, A., Gomez, S., & Arenas, A. (2013). Spectral properties of the Laplacian of multiplex networks. *Physical Review E*.

[16] Cozzo, E., Ferraz de Arruda, G., Aparecido Rodrigues, F., & Moreno, Y. (2019). Layer degradation triggers an abrupt structural transition in multiplex networks. *Physical Review E, 100*(1), 012313.

[17] Cozzo, E., & Moreno, Y. (2016). Characterization of multiple topological scales in multiplex networks through supra-Laplacian eigengaps. *Physical Review E, 94*, 052318.

[18] Erdös, P. & Rényi, A. (1959). On Random Graphs I. *Publicationes Mathematicae Debrecen*, 6, 290-297.

[19] Ferraz de Arruda, G., Cozzo, E., Rodrigues, F. A., & Moreno, Y. (2018). A polynomial eigenvalue approach for multiplex networks. *New Journal of Physics, 20*(9), 095004.

[20] Wu, C. W. (2005). Algebraic connectivity of directed graphs. *Linear and Multilinear Algebra, 53*(3), 203-223.

[21] Prignano, L., Morer, I., & Diaz-Guilera, A. (2017). Wiring the Past: A Network Science Perspective on the Challenge of Archeological Similarity Networks. *Frontiers in Digital Humanities*, 4.

[22] Prignano, L., Morer, I., Fulminante, F., & Lozano, S. (2019). Modelling terrestrial route networks to understand inter-polity interactions (southern Etruria, 950-500 BC). *Journal of Archaeological Science, 105*, 46-58.

[23] Fulminante, F., Prignano, L., Morer, I., & Lozano, S. (2017). Coordinated Decisions and Unbalanced Power. How Latin Cities Shaped Their Terrestrial Transportation Network. *Frontiers in Digital Humanities, 4*(4).

[24] Candelas Peñalba, P. (2022). *Ancient regional transportation infrastructures as co-evolving complex systems: a multiplex network approach* [Master's Dissertation, Universitat de Barcelona]. Dipòsit Digital de la Universitat de Barcelona.

[25] Prignano, P., Candelas Peñalba, P., Fulminante, F., Cozzo, E. (under revision). Roads and rivers. The importance of regional transportation networks for early urbanization in central Italy (1000-500 BC). *Journal of Archaeological Science*.

[26] Fulminante, F. (2023). *The Rise of Early Rome: Transportation Networks and Domination in Central Italy, 1050-500 BC*. Cambridge University Press.