SCUOLA DI DOTTORATO
UNIVERSITÀ DEGLI STUDI DI MILANO-BICOCCA

Department of Statistics and Quantitative Methods (DISMEQ)

PhD in Statistics and Mathematical Finance - Cycle XXXIII

Curriculum in Mathematical Finance

Community detection, risk exposure and multilayer structure in economic and financial complex networks

Candidate: Paolo Bartesaghi

Registration number: 827252

Tutor: prof.ssa Rosanna Grassi

Supervisor: prof.ssa Rosanna Grassi

Coordinator: prof.ssa Emanuela Rosazza

ACADEMIC YEAR 2019-20

Abstract

Complex networks theory, although a relatively new research field, has already proved to be a powerful tool for the description of real systems of different nature. Its transversal nature makes it flexible enough to be applied to extremely diversified contexts, not least economic and financial systems. One of the ideas that have proved particularly effective is the concept of communicability between nodes, introduced by E. Estrada in the field of graph theory. This thesis collects some original contributions that exploit this mathematical tool on many different levels. The first one proposes an extension of the concept of communicability that makes it possible to introduce a new class of centrality measures on networks. These measures turn out to be depending not only on the topology of the network, but also on an external stress factor which affects the level of risk-exposure of the single nodes and of the whole network. Communicability also induces a non-standard network metric. As a second result, we used this metric for the first time from a community detection perspective, that is, in order to identify strongly interacting clusters of nodes, specifically in the World Trade Network. This new methodology will be compared with others already known in the literature and with a new multi-attribute approach we propose here to cluster communities of countries that play a comparable role within the international trade network. Finally, we introduce a preliminary tensor analysis of the multi-level structure of a network, with particular attention to clustering problems, in order to go deeper into the sectoral structure of the world trade network and the interconnections between different trade sectors.

to my parents

Acknowledgements

I would like to acknowledge prof.ssa Rosanna Grassi of Università degli Studi di Milano - Bicocca and prof. Gian Paolo Clemente of Università del Sacro Cuore di Milano for their constant support during the past three years.

Contents

Introduction

Modern economic and financial systems are characterized by a vast collection of interacting agents. Their increasing complexity, the interdependence among highly interconnected entities, the mutual interaction between institutions of different nature and at different levels make the theory of complex networks a suitable tool for describing their local and global behaviour, from both a theoretical and an empirical perspective.

In network theory, vertices (equivalently, nodes) represent these entities - be they individuals, banks, firms, countries, etc. - and edges (equivalently, links) account for the relationships between them [1].

A great effort has been expended in the literature to study the topological properties of networks. Sometimes this is called *static analysis* since it does not assume any mechanism of transmission of effects between nodes. Among such studies, it is frequent to find analyses of clusters formed by groups of institutions, as well as investigations on the centrality of individual nodes. Centrality measures, indeed, represent one of the most useful topological characterizations of the nodes and their role in a network.

In the analysis of financial and economic networks, classical centrality measures have some limitations, as they provide a static description of the interactions between nodes and even measures based on dynamic processes, such as random walks based centralities [2], do not capture the changing conditions to which networks could be submitted in time.

A network can host multiple dynamical processes of different types and their study plays a role of primary importance. Typical examples are represented by diffusion, percolation or contagion processes, like the spread of a virus, a disease, an opinion or a financial default. Although a contagion process is typically a discrete dynamic process in which each node can be in a binary state - susceptible or infected, for instance - the most widely used contagion models describe the evolution, mostly in time, of a continuous variable like the probability of each node to get infected. This probability in turn depends on external factors such as the rate of infection and this rate may

change in time, or, more in general, according to different environmental conditions.

As an illustrative example, let us consider an interbank network. Typically we are interested in analysing the risk-dependent exposure of the different entities involved. Any centrality measure will point out a specific and static ranking of the nodes. However, a bank, which is very central - and so more exposed to a contagion risk - at a low infection rate, may not be as central when the rate is higher. The study of the propagation of financial shocks through these networks and their dependence on external risk conditions is therefore crucial and it is usually known as *dynamic analysis* [3, 4, 5, 6].

This is the reason why, in the first chapter of this work, we develop a mathematical model to account for the risk exposure of entities, in economic or financial networked systems, based on the relation between the Susceptible-Infected (SI) epidemiological model and the so-called communicability functions of a network. Communicability functions belong to a class of matrix functions which are widely used in the description of network properties and which proved to be a powerful tool in many different fields. Through the link between SI models and communicability functions, we propose new centrality indices that quantify the level of risk an entity is exposed to, as a function of the global external stress on the network. This global stress can be identified for instance with the value of the infection rate, which plays the same role of a temperature if we imagine the network embedded in a thermal bath. Our approach takes advantage of the benefits of both static and dynamic analyses. Indeed, unlike the standard approaches, these risk-dependent centralities may vary according to the change of the external global stress and, very peculiarly, it turns out that the ranking of the nodes in terms of risk exposure depends on this external stress. A node, which is at low (high) level of risk under given external conditions, can be at high (low) level under different conditions.

We test our model by using two different systems, a network of assets based on the daily returns of the components of the S&P 100 for the period ranging from January 2001 to December 2017 and a network representing the interconnection between companies in the US top corporates according to Forbes in 1999. For the first one, we extract the essential information about asset correlations through the minimum spanning tree. We measure how the centrality of the assets reacts to different values of the external stress. What emerges is a high volatility in the rankings during the financial crisis of 2007-2008, when the node centrality proves to be more sensitive to the external risk. For the second one, we analyze a sample of significant companies, looking for a correlation between the shareholders value creation (SVC) and their behavior during and after the crisis period at which data were collected. We find that a remarkable

increase in their risk-centrality ranking during a crisis corresponds to a less resilient reaction to the external market turmoil [7].

In the second chapter, we exploit a further consequence of the idea of communicability - the fact that it induces a proper metric on the network - to propose a new approach to a long-standing problem in network theory, the detection of communities. The community structure of a network reveals how it is internally organized, highlighting the presence of special relationships between nodes, that might not be revealed by direct empirical analyses. In particular, we aim at describing the inner structure of the World Trade Network. International trade is based on a set of complex relationships between different countries. Both connections between countries and bilateral trade flows can be modelled as a dense network of interrelated and interconnected agents. Our goal is to highlight subsets of nodes among which the interactions are stronger than average.

In this framework, a critical role is assumed by the communicability distance between nodes. The neighbours of a given node are immediately connected to such a node and they can affect its status in the most direct way. Nonetheless, more distant nodes can influence this node while passing through intermediary ones. In the economic field, a network perspective is actually based on the idea that indirect trade relationships may be important [8]. For instance, it well known the impact of shocks on a given country coming from indirect trade links. A measure of the distance between nodes that takes into account also indirect connections is therefore crucial to grasp the deep interdependencies between trading countries. In this work, we will focus on two measures of distance or metrics on the network: the Estrada communicability distance [9] and the vibrational communicability distance [10]. They both go beyond the limits of the immediate interaction between neighbours and they look simultaneously, albeit differently, at all the possible channels of interactions between nodes. The nearest two nodes are in each metric, the stronger is their interaction or, in other words, the higher is the level of communicability between them. By using communicability and vibrational communicability metrics, we group nodes whose mutual distances are below a given threshold, i.e. whose interactions are stronger than a given value. Then we identify the optimal partition according to a maximum quality function criterion. It is well-known that modularity, for instance, is a way to measure if a specific mesoscopic description of the network in terms of communities is more or less accurate, or, at least, more or less useful. But, unlike the most widely applied Girvan-Newman approach [11], we refer to the quality function proposed in [12] for general metric spaces. We call it *partition quality index* in order to immediately clarify the different nature, specifically metric,

of the quality function we adopt here with respect to classical Newman modularity or its immediate modifications. In this way, we can exploit the additional information contained in the metric structure of the network. Among all the different partitions we get at different thresholds, we select the one providing the maximum quality index, according to the criterion described in [12].

Our proposal provides several advantages and represents a viable alternative to classical methodologies for community detection. Firstly, the method is very efficient from a computational viewpoint. Indeed, given the specific distance matrix, the optimal solution can be easily evaluated varying the threshold. Classical Girvan-Newman methodology is instead a NP-hard problem due to the fact that the space of possible partitions grows faster than any power of the system size. For this reason, several heuristic search strategies have been provided in the literature to restrict the search space while preserving the optimization goal, see e.g. [13] or [14]. Secondly, we cluster nodes going beyond the interactions between neighbours and considering all possible channels of interaction between them. Thirdly, we allow for a degree of flexibility by introducing a threshold. Varying the threshold, it is possible to depart from the optimal solution so that only the strongest (or the weakest) channels of communications emerge. Finally, the procedure offers a set of indicators that allow to exploit main characteristics of the communities detected as well as the relevance of countries inside the community and in the whole network.

In the third chapter, we change our paradigm as we provide a new methodology for clustering countries in the World Trade Network based on a multi-criteria assessment of several topological indicators. That is we look for a specific way to detect nodes having a peculiar common feature and playing a similar role inside the network. The method consists of two steps. In the first step, we rank countries according to a set of centrality measures. In the second one, we compute a similarity index, based on those rankings, between countries and then we apply the clustering algorithm based on the Clique Partition model to detect communities.

More specifically, in the first step, and unlike classical methodologies, we consider all the most prominent centrality definitions proposed in the literature and relevant to international trade. Rather than advocate the superiority of one of them, we aggregate this rich multi-criteria assessment by defining a proper measure of similarity/dissimilarity between nations using their ranking positions. Next, we group together countries that have common structural features in terms of those rankings. The main advantage of our proposal is that we do not focus on a single and specific indicator of centrality, nor we come out with a detailed countries ranking. Rather, we are able

to identify groups of countries that have similar structural properties. A specific tool developed for our project is a new heuristic algorithm to find clusters, based on the Clique Partition model [15]. The Clique Partition model consists of partitioning the vertices of a graph into the smallest number of cliques. First, a measure of similarity or dissimilarity between units must be established. This measure can take both positive and negative values, respectively if two units are similar or dissimilar, and which play the role of gain or cost in grouping together such elements. Then units must be partitioned in subsets, in such a way to maximize the similarity between them. This model has some advantages over, for instance, the classical k-means or hierarchical models. First of all, the clique partition model does not require either that the number of clusters were fixed in advance, e.g. the parameter k, or that the user should arbitrarily analyse the chart of the hierarchical clusters. Rather, the number of clusters results by the optimization of an objective function. Moreover, outliers are not forced to be in a cluster, but they can form peculiar groups of a single element. Finally, the principle of the method is that cluster are composed of mutually homogeneous data, while the k-means models first try to establish cluster's centres and then groups are composed by units that are similar to the centres.

In the last chapter, we propose an extension of some of the methodologies discussed in the previous chapters to multilevel networks. More specifically, we illustrate how it is possible to extend to multilayer networks three different approaches to community detection and we suggest a new way to look at clustering coefficients on multilayer networks. Multiplex and multilayer networks are an extremely challenging research topic. The nature of the connections between the same set of nodes may be of different kind. According to the nature of the connections, different networks are generated on the same set of nodes. Each one can be interpreted as a level in a more complex object, called multilevel network. Different levels highlights a different nature in the interaction between nodes and each one is called monoplex. We can move from one node to another on the same level following the links on that level. But we can also move from one level to another one. The easiest way to do that is to imagine a jump from a node on a level to the same node on a different level. When this is the only possibility to switch levels, we call the multilevel network a multiplex. Nothing prevents us from conjecturing a jump from a node on one level to a completely different node on a second level. When this is allowed the multilevel network is called multilayer network. This is the most general case and it is the one we are interested in. For instance, cities and roads connecting them, on one side, and the same cities and railways connecting them, on the another side, represent two possible monoplexes. When a traveller gets

in a city by train, he can rent a car to go to another city. That is, he is switching from the first monoplex (roads) to the second one (railways). Together they constitute a multiplex. Of course, he cannot move from a city to a different one without using some form of transportation; so this is not a multilayer network. On the contrary, the industrial chemical sector in different countries is a monoplex, as it is the sector of the pharmaceutical industry in different countries, but the former can supply the latter with materials and chemicals within the same country or to foreign nations in the same way. This is an example of multilayer network.

In our work, we start from the tensorial approach in [16], to extend to multilayer networks the community detection methodologies seen in the previous chapters, specifically based on modularity, communicability and communicability distance. These extended methodologies have been tested on some simple toy-models and, in a future work, they will be applied to real multilayer networks too. Finally, since the problem of community detection is strictly linked to how much the network is clustered or not, at the end of the chapter we propose an extension of all the best known clustering coefficients in the literature to multilayer networks. In particular, it will be shown how, in the tensorial setting, it is possible to give them a substantially unified writing and how the choice of a single reference tensor allows to switch from one to the other.

Note for the readers. Each chapter consists of an independent paper and can be read separately from the others. For this reason, we preferred to keep in the individual chapters the notations of the original corresponding paper, which of course are completely consistent, even if this entails a not full uniformity in the notations on the overall body of the thesis.

Chapter 1

Risk-Dependent Centrality in Economic and Financial Networks.

1.1 Introduction

Modern economic and financial systems are characterized by a vast collection of interacting agents [17, 18, 19, 20, 21, 22]. In economic systems, for instance, the interdependence among entities characterizes the trade and exchange of goods in non-anonymous markets as well as in risk sharing agreements in developing countries [17]. In this framework, the agents' interaction is responsible for the nature of the relations between the individual behaviour and the aggregate behaviour [22].

The human factor which underlies these economic and financial systems is also characterized by the interconnectivity. The existence of networks of interpersonal relations has been empirically observed to constitute a fundamental factor in shaping the inter-institution networks, or in accounting for the networks of risk-sharing agreements [23], the formation of buyer-seller networks [24, 25, 26], product adoption decisions [27, 28], diffusive processes [29, 30, 31], industrial organization [21], trade agreements [20] and even for the existence of interbank networks [17]. This is not surprising as humans are responsible for the execution of deals between the institutions to which they belong to [32, 33, 34].

From a mathematical perspective all these interdependencies between economic and financial entities can be captured by the formal concept of network, in which nodes represent the entities (individuals, firms, countries, etc.) and edges account for the

relations between such entities, ranging from social relations to trade agreements [1]. Hence, it is possible to use the tools of network theory to analyze the structure, the evolution and the dynamic processes that take place on these systems. On one side, researchers have studied the topological properties of these networks (sometimes called *static analysis*), which do not assume mechanisms of transmission of effects through the economic and financial entities [35, 36, 37]. Among such studies, it is frequent to find analyses of clusters formed by groups of institutions, as well as the centrality of individual nodes in the networks [38, 39, 40]. Specifically, centrality measures (see Chapter 5 in [1] for a detailed analysis) are topological characterizations of the nodes and their neighborhood in a network. In the analysis of financial and economic networks, the use of centrality measures is not so effective, as the classical ones provide a static view of the network and even other measures based on dynamic processes, such as random walks based centralities [2], do not capture the changing conditions to which these networks could be submitted in relatively short periods of time. As an illustrative example, let us consider a hypothetical interbank network for which we are interested in analyzing the risk-dependent exposure of the various entities of the system. Any centrality measure will point out a specific and static ranking of the nodes. However, a bank which is very central at a low-level of external risk is not necessarily central when such level of external risk increases, and vice versa. On the other hand, the propagation of shocks through these networks is considered and it is usually known as *dynamic analysis* [3, 4, 5, 6, 41, 42, 43]. In these studies, a specific way of transmission of these shocks through the network is assumed – as in the case of "Susceptible-Infected" and "Susceptible-Infected-Recovered" epidemiological models [44, 45, 46] – and then a systemic risk analysis is based on the contagion effects observed through such models.

In this work we develop a mathematical model to account for the risk exposure of an entity in a networked (economic or financial) system. This model is based on the relation between the Susceptible-Infected epidemiological model and the so-called communicability functions of a network [47]. Using this connection we derive new centrality indices that quantify the level of risk at which an entity is exposed to as a function of the global external level of risk. Our approach takes advantage of the benefits of both static and dynamic analyses. Indeed, unlike the standard approaches followed in the literature, these risk-dependent centralities are not static indices, as most of centrality indices are, but they vary with the change of the external global risk level at which the system is submitted to. More importantly, the ranking of the nodes in these networks also depends on this global external level of risk. This means that an entity – a node in the network – which is at low (high) level of risk under external

conditions can be at high (low) level under different conditions.

We test our model by using two different systems, a network of assets based on the daily returns of the components of the S&P 100 for the period ranging from January 2001 to December of 2017 and a network representing the interconnection between companies in the US top corporates according to Forbes in 1999. In the first case we extract the essential information about asset correlations through the minimum spanning tree. We measure how the centrality of the assets changes at different values of the external risk. What emerges is a high volatility in the rankings during the financial crisis of 2007-2008, when the node centrality proves to be more sensitive to the external risk. In the case of the corporate network we analyze a sample of significant companies, looking for a correlation between the shareholders value creation (SVC) and their behaviour during and after the crisis period at which data were collected. We find that a remarkable increase in their risk-centrality ranking during a crisis corresponds to a less resilient reaction to the external market turmoil.

The chapter is structured as follows. In Subsection 1.1.1 we recall the main literature about the use of epidemiological models for modeling financial contagion and we motivate the choice of a Susceptible-Infected model. The necessary mathematical preliminaries are given in Section 1.2. Therefore, we describe a Susceptible-Infected (SI) model on a financial network (Section 1.3) and we define the risk-dependent centrality proving some mathematical properties (Section 1.4). We perform numerical analyses of the proposed centrality for random networks (Section 1.5), then we apply the proposed measure to real-world financial networks (Section 1.6) and we analyze the ranking interlacement problem (Section 1.7). Section 1.8 remarks how the proposed model could provide additional insights in the analysis of the economic and financial impacts of the crisis related to the diffusion of the new coronavirus named SARS-COV-2. Conclusions follow in Section 1.9.

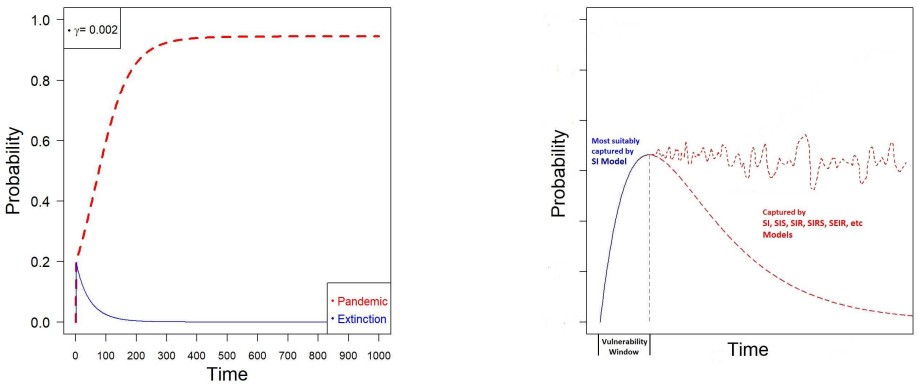
1.1.1 Related literature and motivations

The process in which one financial institution spreads negative effects to another institution resembles very much the propagation of epidemics on networks [42, 48]. The fact that such processes are known as "financial contagion" already captures part of these similarities. Then, it is not strange that epidemiological models are frequently used to capture the subtleties of financial contagion processes. There are many of such compartmental models in epidemiology, but the most widely used for modeling financial contagion are the Susceptible-Infected-Recovered (SIR) [49, 50, 51, 52, 53, 54] and the Susceptible-Infected-Susceptible (SIS) [55, 56] ones. They are not only used to model

financial contagion per se, but also for the propagation of rumors and innovations of interest for financial institutions [57, 58]. These models are well-suited in depicting financial contagion because they do not require arbitrary assumption on loss rates and balance sheets. As remarked by Toivanen [59], they capture the psychological aspects of contagion process "by relating a bank's relative financial strength with the perceived counterparty risk and expectations".

The previously mentioned SIS/SIR models and their variants are mainly used in studying the dynamics of contagion in a system in a post-mortem way. As it is well-known, both SIS and SIR models are characterized by the presence of a threshold τ, which is defined as the reciprocal of the principal eigenvalue λ_1 of the adjacency matrix. The below-the-threshold or above-the-threshold behaviour of the spreading process depends on whether the effective infection rate is less than or greater than such a threshold. Below the threshold, we have the extinction of the contagion and above the threshold a non-zero fraction of infected nodes persists in the network even over a wide range of timescales. The effective infection rate depends on both the infection rate per link γ and on the curing or recovering rate δ. For instance in Figure 1.1(a) we illustrate the evolution of a contagion dynamics for an Erdős-Rényi graph with 100 nodes and connection probability 0.1 by using the SIS model. The principal eigenvalue of the adjacency matrix is $\lambda_1 \approx 10.71$ so that the epidemic threshold is $\tau \approx 0.093$. The infectivity rate per link is 0.002 for both curves and the initial infection probability is 0.2 (20 nodes over 100 initially infected). The curing rate is 0.001 for the dashed red line (epidemic) and 0.04 for the solid blue line (extinction). Then, the effective infection rate is $2 > 0.093$ for the the dashed red line (epidemic) and $0.05 < 0.093$ for the solid blue line (extinction).

In this work we are interested in the very early signals that the system can provide for alerting about a propagation of a financial contagion. In this case it is very important to consider the window of vulnerability between the time the contagion phenomenon is firstly recognized and the time an action is taken to face the infection. This window could be arbitrarily wide. In any real condition, there is a non negligible time interval in which a recovery tool is not available yet and the recovering rate is equal to zero. It has been recently shown by Lee et al. [60] that, within this window, the spreading phenomenon is better described by a SI model than by any other model with a non-null recovering rate, e.g., SIR and SIS (see Figure 1.1(b)). In this framework, a key point is to predict the "most at risk" nodes in the network. Therefore, we are interested in the early times of the epidemic where it is possible to limit or avoid the distress propagation by introducing specific measures on the risky nodes in the network.



(a) **(b)**

Figure 1.1: (a) Illustrative example of an epidemic SIS process over an Erdős-Rényi graph; probability on the vertical axis represents the fraction of infected nodes; (b) Evolution of contagion before and after the window of vulnerability. Adapted from Lee et al.[60]

Moreover, in order to be effective in reducing the spreading phenomenon, the curing rate has to be large enough. More precisely, since $\lambda_1 > \max\left(\bar{k}, \sqrt{k_{\max}}\right)$ (being $\bar{k}$ the mean degree and $k_{\max}$ the maximum degree), the curing rate has to be at least $\delta > \gamma \cdot \max\left(\bar{k}, \sqrt{k_{\max}}\right)$ to get a below-the-threshold behaviour [61]. But for a big real network $\sqrt{k_{\max}}$ can be very large, even if the mean degree is small. This implies that δ has to be significantly bigger than γ, or in other words, the infection significantly weaker than the self-recovering process. This fact could be totally unlikely in a real contagion process on a real network and it makes the use of the SIS or SIR model extremely unrealistic as remarked by Lee et al. [60]. Even when δ is small and the node infection process is dominant, the corresponding epidemic dynamic is better captured by the SI model. From an application point of view, this is possibly true over a wide range of timescales under constrained environments where applying massive action to limit contagion is practically infeasible.

As mentioned before, the detection of risky nodes in a network could be relevant for limiting the risk propagation effects (see, e.g., [62, 63]). Hence, centrality of a given institution as best spreader node in a contagion process has been widely explored (see, for instance, [44]) in order to identify the most dangerous crisis epicenter. The idea of best spreader node has also been studied in [64] in terms of topological centralities, which was previously investigated under the name of *vibrational centrality* (see, e.g.,

[65]). Centralities have been also used as measures to assess contagion in the interbank market. In this framework, Dimitros and Vasileios [66] recommended the use of well-established centrality measures as a way to identify the most important variables in a network. Battiston et al. [38] introduce DebtRank, a centrality measure that accounts for distress in one or more banks, based on the possibility of losses occurring prior to default. The concept that some banks might be too central to fail originates from this work (see, e.g., [38]).

1.2 Preliminaries

Here we use indistinctly the terms graphs and networks. Most of the network theoretic concepts defined hereafter can be found in [1]. A *graph* $\Gamma = (V, E)$ is defined by a set of n nodes (vertices) V and a set of m edges $E = \{(u, v)|u, v \in V\}$ between the nodes. $(u, u) \in E$ is a loop starting and ending in u. The *degree* of a node, denoted by k_u, is the number of edges incident to u in Γ. The *adjacency matrix* of the graph $A = (A_{uv})_{n\times n}$ with entries $A_{uv} = 1$ if $(u, v) \in E$ or zero otherwise. We consider here simple graphs, i.e. without loops and multiedges. The theoretical model will be developed for unweighted networks, we also recall here the definition of weighted graphs, as we consider in the chapter two empirical real examples for which the network is weighted. A weighted *graph* $\Gamma' = (V, E, W)$ is a graph in which $w_{uv} \in W$ is a positive number assigned to the corresponding edge $(u, v) \in E$. In this case the sum of the weights for all edges incident to a node is known as the *weighted degree* or *strength*. We consider here only undirected networks, such that $(u, v) \in E$ implies that $(v, u) \in E$. In this case the matrix A can be expressed as $A = U\Lambda U^T$ where $U = \left[\vec{\psi}_1 \cdots \vec{\psi}_n\right]$ is an orthogonal matrix of the eigenvectors of A and Λ is the diagonal matrix of eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. The entries of $\vec{\psi}_j$ are denoted by $\psi_{j,1}, \ldots, \psi_{j,n}$.

An important quantity for studying communication processes in networks is the communicability function [47], defined for a pair of nodes u and v as

$$G_{uv} = \sum_{k=0}^{\infty} \frac{\left(A^k\right)_{uv}}{k!} = \left(\exp\left(A\right)\right)_{uv} = \sum_{j=1}^{n} e^{\lambda_j} \psi_{j,u} \psi_{j,v}. \tag{1.1}$$

It counts the total number of walks starting at node u and ending at node v, weighted in decreasing order of their length by a factor of $\frac{1}{k!}$. A *walk* of length k in Γ is a set of nodes $i_1, i_2, \ldots, i_k, i_{k+1}$ such that for all $1 \leq l \leq k$, $(i_l, i_{l+1}) \in E$. A *closed walk* is a walk for which $i_1 = i_{k+1}$. Therefore, G_{uv} is considering shorter walks as more influential than longer ones. The matrix exponential is an example of a general class

of matrix functions which are expressible as

$$\big(f(A)\big)_{uv} = \sum_{k=0}^{\infty} c_k \big(A^k\big)_{uv}, \tag{1.2}$$

where c_k are coefficients giving more weight to the shorter than to the longer walks, and making the series converge. The term G_{uu}, which counts the number of closed walks starting at the node u giving more weight to the shorter than to the longer ones, is known as the subgraph centrality of the node u.

We also consider here a Susceptible-Infected (SI) model over an undirected network. Each susceptible node becomes infected at the infection rate γ *per link* times the number of infected neighboring nodes. Let t^* be the instant in which a node i is infected. Node i remains in this state $\forall t \geq t^*$ and does not come back susceptible. Let us introduce a random variable $X_i(t)$ denoting the state of a node i at time t

$$X_i\left(t\right) = \begin{cases} 1 & \text{if } t \geq t^* \\ 0 & \text{otherwise} \end{cases} \tag{1.3}$$

Then we define

$$x_i(t) = P[X_i(t) = 1] = \mathbb{E}[X_i(t)] \in [0, 1], \tag{1.4}$$

which is the probability that node i is infected at time t. In other words, node i is healthy at time t with probability $1 - x_i(t)$. For the whole network, we define the vector of probabilities:

$$\vec{x}(t) = [x_1(t), \ldots, x_n(t)]^T. \tag{1.5}$$

1.3 Model

Let us consider a SI model on a financial network. The nodes of a graph $\Gamma = (V, E)$ represent financial institutions and the edges connecting them represent an interaction that can transmit a "disease" from one institution to another. A node can be susceptible and then get infected from a nearest neighbor or it is infected and can transmit the infection to other susceptible nodes. Let γ be the infection rate and let $x_i\left(t\right)$ be the probability that node i get infected at time t from any infected nearest neighbor. Then,

$$\frac{dx_i\left(t\right)}{dt} = \dot{x}_i\left(t\right) = \gamma\left[1 - x_i\left(t\right)\right] \sum_{j=1}^{n} A_{ij} x_j\left(t\right) \tag{1.6}$$

which in matrix-vector form becomes:

$$\dot{\vec{x}}(t) = \gamma \left[1 - \mathrm{diag}\left(\vec{x}(t)\right)\right] A\vec{x}(t), \tag{1.7}$$

with initial condition $\vec{x}(0) = \vec{x}_0$.

It is well-known that on a general strongly connected network[1][45]:

1. if $\vec{x}_0 \in [0,1]^n$ then $\vec{x}(t) \in [0,1]^n$ for all $t > 0$;
2. $\vec{x}(t)$ is monotonically non-decreasing in t;
3. there are two equilibrium points: $\vec{x} = \vec{0}$, i.e. no epidemic, and $\vec{x} = \vec{1}$ (the vector of all ones), i.e. full contagion;
4. the linearization of the model around the point $\vec{0}$ is given by

$$\dot{\vec{x}}(t) = \gamma A\, \vec{x}(t) \tag{1.8}$$

   and it is exponentially unstable; in fact, since, in a non-empty undirected graph, A has at least one positive eigenvalue, any solution component in the direction of the corresponding eigenvector grows unboundedly as t increases;
5. each trajectory with $\vec{x}_0 \neq \vec{0}$ converges asymptotically to $\vec{x} = \vec{1}$, i.e. the epidemic spreads monotonically to the entire network.

In particular, the linearized problem comes from the following observation. It can be checked that

$$\dot{x}_i(t) = \gamma[1 - x_i(t)] \sum_{j=1}^{n} A_{ij} x_j(t) \leq \gamma \sum_{j=1}^{n} A_{ij} x_j(t) \tag{1.9}$$

or

$$\dot{\vec{x}}(t) \leq \gamma A\, \vec{x}(t), \tag{1.10}$$

$\forall i$ and $\forall t$. Then, we can use the linear dynamical system

$$\dot{\vec{x}}^{\star}(t) = \gamma A \vec{x}^{\star}(t), \tag{1.11}$$

[1] Although in what follows we will refer only to undirected networks, we recall the following definition: A graph $\Gamma = (V, E)$ is strongly connected if and only if for each pair of nodes $i, j \in V$ there is a directed walk starting at i and ending at j, and a directed walk starting at j and ending at i

as an upper-bound for the original non-linear dynamical system, that has been used in the literature (see [45]) as an approximation of the exact problem. One of its main advantages is that it can be solved analytically and its solution $\vec{x}^{\star}(t)$ can be written as:

$$\vec{x}^{\star}(t) = e^{\gamma t A}\vec{x}_0^{\star}, \tag{1.12}$$

which using the spectral decomposition of A can be written as

$$\vec{x}^{\star}(t) = \sum_{j=1}^{n} e^{\gamma t \lambda_j}\vec{\psi}_j\vec{\psi}_j^{T}\vec{x}_0^{\star}. \tag{1.13}$$

This solution to the linearized model is affected by the following main problems:

1. $\vec{x}^{\star}(t)$ grows quickly without bound, in spite of the fact that $\vec{x}^{\star}(t)$ is a vector of probabilities which should not exceed the unit;

2. $\vec{x}^{\star}(t)$ is an accurate solution to the nonlinear SI problem only if $t \to 0$ and $\vec{x}_0^{\star} \to 0$.

The mathematical properties of the linear dynamical system 1.8 as well as of the solution 1.12 have been extensively studied by Mugnolo in [67]. We direct the reader to this reference for the details.

Hereafter we will follow the recent work of Lee et al. [60], who proposed the following change of variable to avoid the aforementioned problems with the solution of the linearized SI model:

$$y_i(t) \coloneqq -\log\left(1 - x_i(t)\right), \tag{1.14}$$

which is an increasing convex function. Then, as $1 - x_i(t)$ is the probability that node i is not infected at a given time t, the new variable $y_i(t)$ can be interpreted as the information content of the node i or surprise of not being infected (see, e.g., [68]). According to [60], the SI model 1.6 can be now written as

$$\frac{dy_i(t)}{dt} = \dot{y}_i(t) = \gamma\sum_{j=1}^{n} A_{ij}x_i(t) \tag{1.15}$$

or

$$\vec{\dot{y}}(t) = \gamma A\,\vec{x}(t). \tag{1.16}$$

The approximate solution to the SI model provided by [60] is then given by

$$\vec{x}(t) = \vec{1} - e^{-\vec{y}(t)}, \tag{1.17}$$

where $e^{-\vec{y}(t)}$ is the vector in which the ith entry is $e^{-y_i(t)}$ and

$$\begin{aligned}\vec{y}\,(t) = \quad & e^{\gamma t A \mathrm{diag}\left(\vec{1}-\vec{x}_0\right)}\left[-\log\left(1-\vec{x}_0\right)\right] \\ & + \sum_{j=0}^{\infty} \frac{(\gamma t)^{j+1}}{(j+1)!}\left[A\mathrm{diag}\left(\vec{1}-\vec{x}_0\right)\right]^j A\left(\vec{x}_0 + \left(\vec{1}-\vec{x}_0\right)\log\left(\vec{1}-\vec{x}_0\right)\right).\end{aligned} \tag{1.18}$$

As stressed by [60], the interesting case of the dynamics is when $\vec{x}_0 < \vec{1}$, in which case the solution simplifies to

$$\vec{y}(t) = \vec{y}_0 + \left[e^{\gamma t A \mathrm{diag}\left(\vec{1}-\vec{x}_0\right)} - I\right] \cdot \mathrm{diag}\left(\vec{1}-\vec{x}_0\right)^{-1} \vec{x}_0. \tag{1.19}$$

Now, we can make the further assumption that the initial probabilities of being infected are equal for every node, i.e. that at the beginning every node has the same probability β to be infected and to be the one from which the epidemic starts. This means that we are asking for

$$x_{0i} = \beta = \frac{c}{n}, \ \forall i = 1, \ldots, n \tag{1.20}$$

for some scalar constant c. In this case $\mathrm{diag}\left(\vec{1}-\vec{x}_0\right) = \left(1-\frac{c}{n}\right) I = (1-\beta)I$. If we set $\alpha = 1-\beta$, the approximate solution of the SI on the network becomes:

$$\vec{y}(t) = \vec{y}_0 + \frac{1-\alpha}{\alpha}\left[e^{\alpha\gamma t A} - I\right]\vec{1}. \tag{1.21}$$

and since $\vec{y}_0 = (-\log\alpha)\,\vec{1}$,

$$\vec{y}(t) = \left(\frac{1}{\alpha} - 1\right) e^{\alpha\gamma t A}\,\vec{1} - \left(\log\alpha + \frac{1-\alpha}{\alpha}\right)\vec{1}. \tag{1.22}$$

The component $(e^{\alpha\gamma t A}\vec{1})_i$ is called total communicability of node i and it will be denoted by $\mathscr{R}_i$. Hence, component-wise we have:

$$y_i\,(t) = \left(\frac{1}{\alpha} - 1\right)\mathscr{R}_i - \left(\log\alpha + \frac{1-\alpha}{\alpha}\right). \tag{1.23}$$

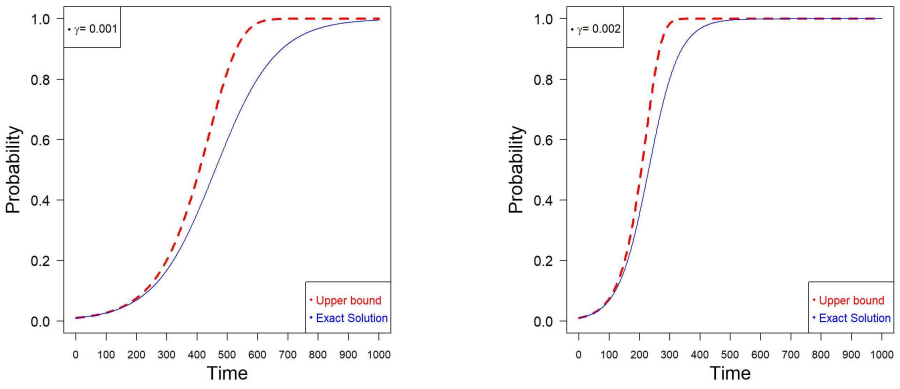
Keeping in mind that $-\log\alpha = y_i(0)$ and $\alpha = 1-\beta$, we can write the previous equation also as

$$\Delta y_i\,(t) = y_i\,(t) - y_i\,(0) = \frac{\beta}{\alpha}(\mathscr{R}_i - 1), \tag{1.24}$$

which means that $\mathscr{R}_i - 1$ at time t is proportional to the variation in the information content of node i from time 0 to time t. Finally, the probability of node i of being infected at time t can be expressed in terms of $\mathscr{R}_i$ as

$$x_i(t) = 1 - (1-\beta)e^{-\frac{\beta}{1-\beta}(\mathscr{R}_i - 1)}. \tag{1.25}$$

When the parameter β is fixed, the number of infected nodes depends only on the term $e^{\alpha\gamma t A}\vec{1}$ and then on the total communicabilities $\mathscr{R}_i$. It is worth noticing that the probability given by 1.25 for a node i represents an upper bound for the exact solution of the SI model. Hence, in this way we do not underestimate the contagion probabilities. Let us consider, for instance, the time evolution of an infection propagation on an Erdős-Rényi network with 100 nodes and edge density $\delta = 0.1$. Results are illustrated in figure 1.2 for two different values of the infectivity rate, $\gamma = 0.001$ (left) and $\gamma = 0.002$ (right). The dashed red lines represent the mean probability that a node is infected at time t as given by equation 1.25. The solid blue lines represent the same probability as given by the exact solution of the Kermack-McKendrick SI model with the same mean degree. In both plots, the initial probability is $\beta = 0.01$.



(a) **(b)**

Figure 1.2: Simulation of the progression of a SI epidemics on an Erdős-Rényi network with 100 nodes and edge density $\delta = 0.1$. The parameters used in the model are: $\beta = 0.01$ and $\gamma = 0.001$ (left) and $\gamma = 0.002$ (right). Dashed (red) lines represent the upper bound given by 1.25; solid (blue) lines represent the value of the same probability in a Kermack-McKendrick SI model with the same mean degree $\bar{k} = (n-1)\delta$.

1.4 Risk-dependent centrality

Let us designate $\zeta = \alpha\gamma t$, which determines the level of risk to which the whole network is submitted at time t. For instance, for $\gamma = 0$, i.e. $\zeta = 0$, there is no risk of infection on the network as a node cannot transmit the disease to a nearest neighbor. This situation corresponds to the case of isolated nodes (no edges). When $\zeta \to \infty$ the risk of infection is very high due to the fact that for a fixed value of c the infectivity is infinite. Therefore, we call $\mathscr{R}_i = \left(e^{\zeta A}\vec{1}\right)_i$ the risk-dependent centrality of the node i. That is, the values of $\mathscr{R}_i$ reflects how central a node is in "developing" the epidemics on the network. As the networks considered are undirected, this centrality accounts for both the facility with which the node gets infected as well as the propensity of this node to infect other nodes. The index $\mathscr{R}_i$ can be expressed as

$$\mathscr{R}_i = \left[\left(I + \zeta A + \zeta^2\frac{A^2}{2!} + \zeta^3\frac{A^3}{3!} + \cdots\right)\vec{1}\right]_i, \tag{1.26}$$

which indicates that it counts the number of walks of different lengths, that have started at the corresponding node, weighted by a factor $\dfrac{\zeta^k}{k!}$. It is straightforward to realize from the definition of the risk-dependent centrality that it can be split into two contributions. That is, $\mathscr{R}_i$ is composed by a weighted sum of all closed walks that start and end at i, $\left(e^{\zeta A}\right)_{ii}$ and by the weighted sum of walks that start at the node i and end elsewhere, $\sum_{j\neq i}\left(e^{\zeta A}\right)_{ij}$

$$\mathscr{R}_i = \left(e^{\zeta A}\right)_{ii} + \sum_{j\neq i}\left(e^{\zeta A}\right)_{ij} := \mathscr{C}_i + \mathscr{T}_i, \tag{1.27}$$

where the first term in the right-hand side represents the circulability of the disease around a given node and the second one represents the transmissibility of the disease from the given node to any other in the network. The circulability is very important because it accounts for the ways the disease has to become endemic. For instance, a large circulability for a node i implies that the disease can infect its nearest neighbors and will keep coming back to i over and over again in a circular way. We start now by proving some results about these risk-dependent centralities as functions of ζ.[1] The following theorem is a special case of results found, for instance, in [69].

Theorem 1. *The node ranking given by the risk dependent centralities $\mathscr{R}_i(\zeta)$, with $i = 1, \ldots, n$, reduces to the ranking given by the degree k_i in the limit as the risk $\zeta \to 0$, and to the ranking given by eigenvector centrality as $\zeta \to \infty$.*

[1] When needed, we will explicitate the dependence of $\mathscr{R}_i$ on ζ as $\mathscr{R}_i(\zeta)$.

Proof. We begin by observing that the ranking of nodes, in terms of their risk-dependent centrality, is unaffected if all the centralities $\mathscr{R}_i$ are shifted and rescaled by the same amount. That is, the same ranking is obtained using either $\mathscr{R}_i$ or the equivalent measure

$$\hat{\mathscr{R}}_i = \frac{\mathscr{R}_i - 1}{\zeta},$$

where $\zeta > 0$. Now, we have

$$\hat{\mathscr{R}}_i = \left[\left(A + \frac{\zeta}{2!}A^2 + \cdots\right)\vec{1}\right]_i = k_i + \frac{\zeta}{2!}(A^2\vec{1})_i + O(\zeta^2). \tag{1.28}$$

Hence, in the limit of $\zeta \to 0$, the ranking given by $\mathscr{R}_i$ is identical to degree ranking.

To study the limit for ζ large we write

$$\mathscr{R}_i = \left[e^{\zeta A}\vec{1}\right]_i = \sum_{k=1}^{n} e^{\zeta\lambda_k}(\psi_k^T\vec{1})\psi_{k,i} = e^{\zeta\lambda_1}(\psi_1^T\vec{1})\psi_{1,i} + \sum_{k=2}^{n} e^{\zeta\lambda_k}(\psi_k^T\vec{1})\psi_{k,i}. \tag{1.29}$$

We note again that for ranking purposes we can use the equivalent measure obtained by dividing all risk-dependent centralities by the same quantity, $e^{\zeta\lambda_1}(\psi_1^T\vec{1})$, which is strictly positive. That is, we can use

$$\tilde{\mathscr{R}}_i = \psi_{1,i} + \frac{1}{\psi_1^T\vec{1}}\sum_{k=2}^{n} e^{\zeta(\lambda_k-\lambda_1)}(\psi_k^T\vec{1})\psi_{k,i}. \tag{1.30}$$

Since the network is connected, the Perron–Frobenius Theorem insures that $\lambda_1 > \lambda_2 \geq \cdots \geq \lambda_n$. Hence, each term $e^{\zeta(\lambda_k-\lambda_1)}$ for $k = 2, \ldots, n$ vanishes in the limit as $\zeta \to \infty$, and we see from 1.30 that the risk-dependent centrality measure gives the same ranking as eigenvector centrality for ζ large. □

It is interesting to observe that the risk-dependent centrality of every node also depends on the (strictly positive) quantity

$$\psi_1^T\vec{1} = \sum_{j=1}^{n}\psi_{1,j},$$

see equation 1.29. The larger this quantity is, the higher is the risk-dependent centrality of each node. Assuming that the dominant eigenvector is normalized so as to have Euclidean norm equal to 1, it is well known that this quantity is always between 1 and $\sqrt{n}$. The value 1 is never attained for a connected graph. It can only be approached in the limit as all the eigenvector centrality is concentrated on one node, say node i, where it takes values arbitrarily close to 1, with the values $\psi_{1,j}$ for all $j \neq i$ taking

arbitrarily small values. An example of this would be the star graph[1] S_n for $n \to \infty$. The maximum value is attained in the case where all nodes have the same eigenvector centrality: $\psi_{1,1} = \psi_{1,2} = \cdots = \psi_{1,n}$ (i.e., in the case of regular graphs).

Let us return to the decomposition $\mathscr{R}_i = \mathscr{C}_i + \mathscr{T}_i$ of the risk-dependent centrality of a node into its two components, circulability and transmissibility. Similar considerations apply to these quantities. We summarize them in the following result.

Theorem 2. *The node rankings given by the degree k_i and the eigenvector centrality represent the limiting cases of the ranking based on the risk-dependent circulability $\mathscr{C}_i(\zeta)$ as the external level of risk $\zeta \to 0$ and $\zeta \to +\infty$, respectively. The same is true for the risk dependent transmissibility $\mathscr{T}_i(\zeta)$.*

Proof. The proof for the circulability is a straightforward adaptation of that for the total communicability; see also [69].

We give the details for the transmissibility, which has not been analyzed before. We have for $i \neq j$ that

$$\left(e^{\zeta A}\right)_{ij} = \zeta A_{ij} + \frac{\zeta^2}{2!} w_{i,j}^{(2)} + O(\zeta^3),$$

where $w_{i,j}^{(2)}$ denotes the number of walks of length two between node i and node j. Dividing by $\zeta > 0$, summing over all $j \neq i$ and taking the limit as $\zeta \to 0$, we find

$$\zeta^{-1} \mathscr{T}_i = \zeta^{-1} \sum_{j \neq i} \left(e^{\zeta A}\right)_{ij} \to \sum_{j \neq i} A_{ij} = k_i,$$

where we have used the fact that $A_{ii} = 0$, for all i. Hence, transmissibility is equivalent to node degree in the small ζ limit. For the large ζ limit we write

$$\mathscr{T}_i = \sum_{j \neq i} \sum_{k=1}^{n} e^{\zeta \lambda_k} \psi_{k,i} \psi_{k,j} = e^{\zeta \lambda_1} \psi_{1,i} \sum_{j \neq i} \psi_{1,j} + \sum_{k=2}^{n} e^{\zeta \lambda_k} \left[\sum_{j \neq i} \psi_{k,i} \psi_{k,j} \right].$$

Dividing by the positive constant $e^{\zeta \lambda_1} \sum_{j \neq i} \psi_{1,j}$ and taking the limit as $\zeta \to \infty$, the second part of the right-hand side vanishes and we obtain again the eigenvector centrality $\psi_{1,i}$ of node i. □

Remark 1. *A natural question is how rapidly the degree (for $\zeta \to 0$) and eigenvector (for $\zeta \to \infty$) centrality limits are approached if the number of nodes n in the network goes to infinity. From the Taylor expansions (see for example equation 1.28) we see that the degree limit is reached more slowly if the row sums of A^2 grow as $n \to \infty$.*

[1] We recall that the star graph S_n consists of $n-1$ nodes $v_1, \ldots, v_{n-1}$, each attached to a central node v_n by an edge.

In this case, as n increases ζ must be taken smaller and smaller before the ranking reduces to the one given by the degree. On the other hand, if the network grows in such a way that the maximum degree of any node remains uniformly bounded, then the rate of convergence is independent of the number n of nodes, at least asymptotically.

The rate of convergence to the eigenvector centrality ranking is largely determined by the spectral gap, $\lambda_1 - \lambda_2$. If the gap remains bounded below by a positive constant as $n \to \infty$, the value of ζ necessary to reach the eigenvector centrality limit is easily seen to grow at most like $O(\ln n)$, and in practice the rate of convergence is scarcely affected by the size of the network. If, on the other hand, the gap closes as $n \to \infty$, then the rate of convergence to the eigenvector centrality will become arbitrarily slow. The faster the gap closes for $n \to \infty$, the more rapidly the rate of convergence deteriorates.

We conclude this section with some comments on the measures $\mathscr{R}_i$, $\mathscr{C}_i$ and $\mathscr{T}_i$. While they all display the same limiting behaviour and provide identical rankings in the small and large ζ limits, they provide different insights on the network structure (and therefore on node risk). For instance, it is well known that subgraph centrality (which is the same as circulability, see [1, 70]) can discriminate between the nodes of certain regular graphs, that is, graphs in which all the nodes have the same degree. The same holds for transmissibility. Total communicability, on the other hand, is unable to discriminate between the nodes of regular graphs (and neither are degree and eigenvector centrality, of course). These measures are also different from a computational viewpoint. One advantage of the risk centrality based on total communicability is that it only requires the computation of the action of the matrix exponential $e^{\zeta A}$ on the vector $\vec{1}$. The entries of the resulting vector can be computed efficiently without having to compute any entry of $e^{\zeta A}$, see [71]. Modern Krylov-type iterative methods (like those based on the Lanczos or Arnoldi process) can handle huge networks (with many millions of nodes) without any difficulty. In contrast, the computation of the circulability requires the explicit computation of the diagonal entries of $e^{\zeta A}$ (the node transmissibility is then easily obtained by subtracting the circulability from the total communicability). Although there are techniques that can handle fairly large graphs (see [72]), these calculations are much more expensive than those for the total communicability. This limits the size of the networks that they can be applied to. However, for most financial networks the computation of the circulability is still feasible.

A final consideration regards the values assumed by the external risk parameter ζ. Although, in principle, it can vary between 0 and infinity, for the purposes of most of the applications that follow, it may be sufficient to vary ζ between 0 and 1. The rationale for using the interval $[0, 1]$ relies on the fact that, at $\zeta = 1$, the rankings

given by $\mathscr{R}_i$ are already stabilizing around those provided by eigenvector centrality and therefore no more interlacings between rankings are possible. As we will show, we typically observe a single point of interlacement and it usually occurs before reaching the value $\zeta = 1$. Furthermore, this choice is equivalent to fix $t = 1$ in the epidemic model solution 1.22, and, already as ζ approaches 1, all the probabilities involved in that model become completely negligible or equal to 1.

1.5 Risk-dependent centrality on a random network

For the analysis of real-world (financial and economic) networks it is necessary to investigate how informative the results obtained are with respect to the real system under analysis. This significance is typically addressed by comparing to those properties obtained from network null models. As such null models we consider here Erdős-Rényi (ER) random networks $\Gamma_{ER}(n, p)$ with n nodes and wiring probability p (see [73, 74]), for which, in this section, we provide a series of analytical results. We start by generating a family of simulated ER graphs and discarding simulations for which the obtained graph is not connected.

In particular, we aim at testing how the external risk ζ and the probability p, and hence the expected graph density δ, affect the results. For this purpose, we generate 1000 graphs $\Gamma_{ER}(n; p)$ with $n = 100$ at different values of p. For each graph, we compute the main measures for alternative values of ζ. Firstly, we report in figure 1.3 the behaviour of risk-dependent centrality $\mathscr{R}_i$, circulability $\mathscr{C}_i$ and transmissibility $\mathscr{T}_i$ as functions of the density, assuming a fixed high level of external risk, $\zeta = 1$. Since the values of $\mathscr{R}_i$ are significantly increasing when the density of the graph increases, we display, in figure 1.3(a), the distributions of the ratio between the risk-dependent centrality of each node $\mathscr{R}_i$ and its average value $\mathbb{E}(\mathscr{R}_i)$.

As might be expected, the centralities of nodes tend to be similar when $\delta \to 1$ and we move towards the complete graph, i.e. we observe a lower variability of the distribution of the ratios. Similar behaviours are also observed for $\mathscr{C}_i$ and $\mathscr{T}_i$, with an higher volatility for the circulability (see figures 1.3(b) and 1.3(c)). In 1.3(d), we show the distributions of the incidence of the circulability $\mathscr{C}_i$ on the risk-dependent centrality $\mathscr{R}_i$, that is the distribution of the ratio $\frac{\mathscr{C}_i}{\mathscr{R}_i}$ again as a function of the density δ. When $\zeta = 1$, for all the graphs analyzed, the average value is around $\frac{1}{n}$, implying that the transmissibility has an average incidence of $\frac{n-1}{n}$ on $\mathscr{R}_i$. It is noteworthy to look at the variability of the distributions. When the density is extremely low, i.e. we refer to a very sparse graph, the heterogeneity of the nodes degree affects the ratio $\frac{\mathscr{C}_i}{\mathscr{R}_i}$. For

instance, when $\delta = 0.1$, the circulability of a node ranges approximately from 0.15% to 2.5% of the risk-dependent centrality for the same node. A lower variability is observed for higher densities. For instance, for $\delta = 0.5$, the ratio $\frac{\mathscr{C}_i}{\mathscr{R}_i}$ varies between 0.6% and 1.3%. For $\delta = 0.95$, we observe a ratio between 0.9% and 1.15%.

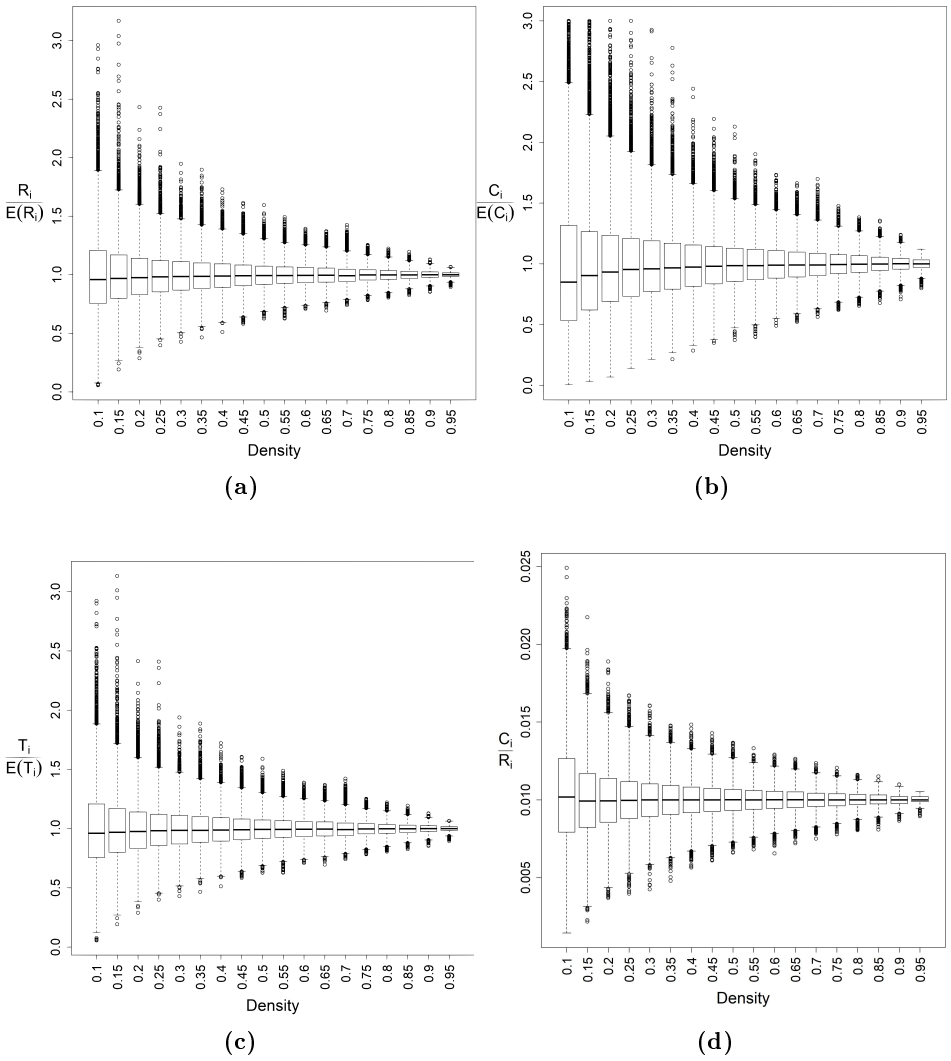


Figure 1.3: Figure a) displays the distributions of the ratios between the risk-dependent centrality of each node $\mathscr{R}_i$ and the average risk-dependent centrality $\mathbb{E}(\mathscr{R}_i)$, computed assuming $\zeta = 1$. Figure b) and c) display the analogous distributions for circulability and transmissibility. Figure d) shows the distributions of the ratios between the circulability $\mathscr{C}_i$ and the risk-dependent centrality of each node $\mathscr{R}_i$, computed assuming $\zeta = 1$. All Figures are based on 1000 randomly generated ER networks $\Gamma_{ER}(n; p)$ with a density varying between 0.10 and 0.95.

In figure 1.4, we show the corresponding behaviours of risk-dependent centrality $\mathscr{R}_i$, circulability $\mathscr{C}_i$ and transmissibility $\mathscr{T}_i$ as functions of the density, but assuming a fixed low level of external risk, $\zeta = 0.1$. Again all Figures are based on 1000 randomly generated ER networks $\Gamma_{ER}(n;p)$ with δ varying between 0.10 and 0.95.

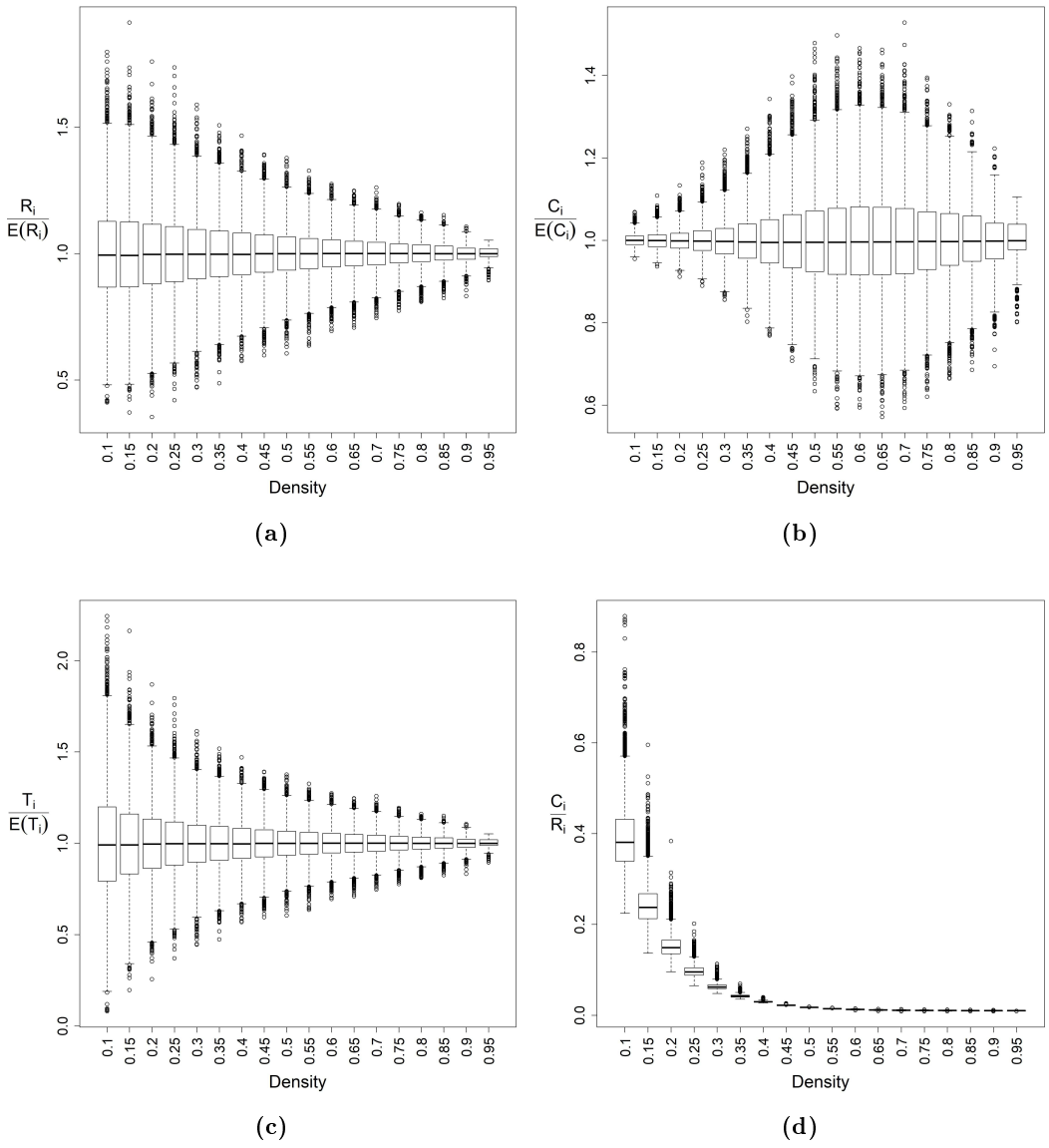


Figure 1.4: Figures a), b), c) and d) display the distributions of ratios $\frac{\mathscr{R}_i}{\mathbb{E}(\mathscr{R}_i)}$, $\frac{\mathscr{C}_i}{\mathbb{E}(\mathscr{C}_i)}$, $\frac{\mathscr{T}_i}{\mathbb{E}(\mathscr{T}_i)}$ and $\frac{\mathscr{C}_i}{\mathscr{R}_i}$ respectively, computed in case of a low external risk ($\zeta = 0.1$). All Figures are based on 1000 randomly generated ER networks $\Gamma_{ER}(n;p)$ with a density varying between 0.10 and 0.95.

Focusing on the risk-dependent centrality ratio $\frac{\mathscr{R}_i}{\mathbb{E}(\mathscr{R}_i)}$, we observe that the standard deviation between nodes is lower in the low-risk framework ($\zeta = 0.1$) than in the high-risk one ($\zeta = 1$). For instance, when the density is equal to 0.1, the standard deviation of the ratio moves from 0.20 for $\zeta = 0.1$ to 0.37 for $\zeta = 1$. At a phenomenological level, this behaviour can be justified by the fact that differences between nodes tend to be enhanced when the network is highly risk-exposed. Furthermore, the pattern of $\frac{\mathscr{C}_i}{\mathbb{E}(\mathscr{C}_i)}$ for $\zeta = 0.1$ is very peculiar. In this case, when the network is very sparse, nodes show a similar circulability, while higher differences are observed when the density is around 0.5. Lastly, in figure 1.5, we focus on the ratio $\frac{\mathscr{C}_i}{\mathscr{R}_i}$ and we report the incidence of the circulability on the risk-dependent centrality as a function of the external risk ζ. In case of sparse networks (1.5(a)), when the external risk is low, we have that the infection remains in larger part circulating in a loopy way around the nodes, while only a lower proportion of risk tends to be transmitted to other nodes. This is due to the fact that, for A sparse and ζ small, the matrix $e^{\zeta A} = I + \zeta A + \frac{\zeta^2}{2}A^2 + O(\zeta^3)$ is strongly diagonally dominant. When the external risk is high, as already observed, we have an average incidence of the circulability $\mathscr{C}_i$ on the risk-dependent centrality around $\frac{1}{n}$. On the contrary, when a very dense network is considered, the ratio $\frac{\mathscr{C}_i}{\mathscr{R}_i}$ is very little affected by the external risk. In this case, both $\mathscr{C}_i$ and $\mathscr{R}_i$ increase on average at the same rate when ζ increases. However, the decreasing behaviour of $\frac{\mathscr{C}_i}{\mathscr{R}_i}$ is noticeable for very low values of ζ.

In what follows we provide an exhaustive proof of the behaviours observed so far. Let us start with the pattern of the ratio $\frac{\mathscr{C}_i}{\mathscr{R}_i}$ at high density (see figures 1.3(d), 1.4(d) and 1.5(b)).

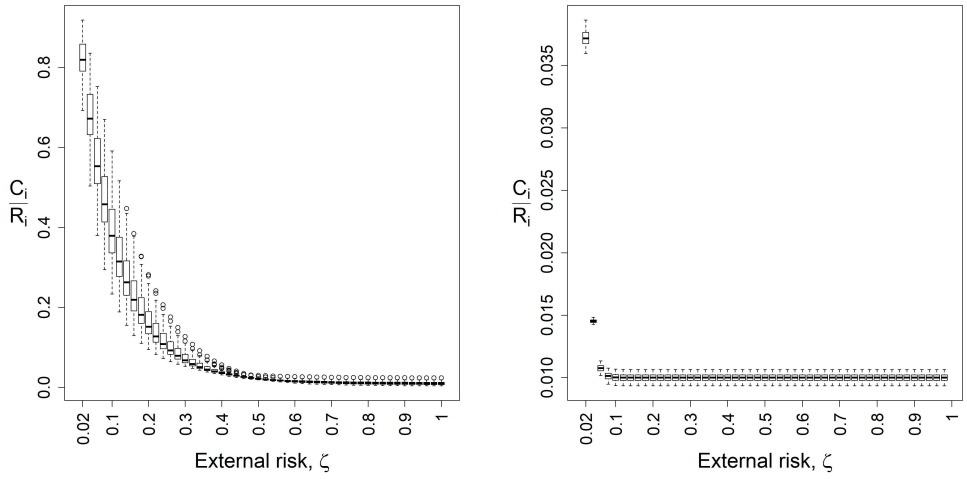
The asymptotic behaviour of this ratio can be explained as a consequence of Theorem 5 in Appendix A, where we derive the close expressions of the three risk-dependent centrality measures for a complete graph. In fact, as $\delta \to 1$, the ER network approaches a complete network and, for ζ increasing, the ratio $\frac{\mathscr{C}_i}{\mathscr{R}_i}$ approaches $1/n$, as shown in A.1.

Nonetheless, this result can be generalized. In fact, for an ER network which is dense enough, the following property holds for any ζ.

Theorem 3. *Let $\Gamma_{ER}(n;p)$ be an Erdős-Rényi random graph with n nodes and probability p. If the edge density of the graph is $\delta > (\log n)^6/n$ and $p(1-p) > (\log n)^4/n$, then for any node i*

$$\lim_{n\to\infty} \frac{n\mathscr{C}_i}{\mathscr{R}_i} = 1, \tag{1.31}$$

independently of ζ.



(a) **(b)**

Figure 1.5: Distribution of the ratios between $\mathscr{C}_i$ and $\mathscr{R}_i$, computed for different ζ and by using generated ER graphs with a density equal to 0.1 (Figure a) and 0.9 (Figure b), respectively. Both Figures are based on 1000 randomly generated ER networks $\Gamma_{ER}(n;p)$.

Proof. Let us consider as usual that $\lambda_1 > \lambda_2 \geq \cdots \geq \lambda_n$ in a connected graph. It is known that in an ER graph the spectral gap $(\lambda_1 - \lambda_2) \gg 0$. Indeed, as proved in [75], $\lim_{n\to\infty} \frac{\lambda_1}{np} = 1$, while λ_2 and λ_n grow more slowly as $\lim_{n\to\infty} \frac{\lambda_2}{n^{\varepsilon}} = 0$ and $\lim_{n\to\infty} \frac{\lambda_n}{n^{\varepsilon}} = 0$ for every $\varepsilon > 0.5$, respectively.

Then, since we have $np\,(1-p) > (\log n)^4$ for n large enough, all but the largest eigenvalue lie with high probability in the interval $\sqrt{np\,(1-p)}\,[-2+o\,(1)\,,+2+o\,(1)]$ (see [76] and [77]). Therefore,

$$\lim_{n\to\infty} \frac{\mathscr{C}_i}{\mathscr{R}_i} = \lim_{n\to\infty} \frac{\psi_{1,i}^2 e^{\zeta\lambda_1} + \sum_{k=2}^{n} \psi_{k,i}^2 e^{\zeta\lambda_k}}{\psi_{1,i}\left(\vec{\psi}_1^T \vec{1}\right) e^{\zeta\lambda_1} + \sum_{k=2}^{n} \psi_{k,i}\left(\vec{\psi}_k^T \vec{1}\right) e^{\zeta\lambda_k}} = \frac{\psi_{1,i}}{\sum_{j=1}^{n} \psi_{1,j}}. \tag{1.32}$$

The edge density of an ER graph is $\delta = p$. In [78], it was proved that for $np > (\log n)^6$, there exists a positive constants C such that the following inequality holds

$$\left\| \vec{\psi}_1 - \frac{1}{\sqrt{n}} \vec{1} \right\|_{\infty} < C \frac{1}{\sqrt{n}} \frac{\log n}{\log (np)} \sqrt{\frac{\log n}{np}}, \tag{1.33}$$

which in plain words means that an ER graph of density $\delta > (\log n)^6 / n$ is "almost" regular when $n \to \infty$. That is $\lim_{n\to\infty} \sqrt{n}\psi_{1,i} = 1$ for every node i. Thus, the result immediately follows. □

It is worth pointing out that, when the density of an ER network is very low, the standard deviation of the ratio $\frac{\mathscr{C}_i}{\mathscr{R}_i}$ is very large with respect to that of ER networks with large densities (as shown in figure 1.3(d)). As we have proved before, the convergence of this ratio to the value n^{-1} takes place only when the density of the graph is relatively large. Let us now analyze what happens when the edge density is very small for large graphs. In this case, we observe a slower decay of the ratio $\frac{\mathscr{C}_i}{\mathscr{R}_i}$ as a function of the external risk in the range $[0,1]$ (see figure 1.5(a)). This fact can be easily proven as follows. In general, both the numerator and denominator of this ratio can be expressed as infinite series of the type:

$$\mathscr{C}(\zeta)_i = Q(\zeta) = 1 + a_2\zeta^2 + \cdots + a_k\zeta^k + \cdots,$$

$$\mathscr{R}(\zeta)_i = H(\zeta) = 1 + b_1\zeta + (a_2 + b_2)\zeta^2 + \cdots + (a_k + b_k)\zeta^k + \cdots = Q(\zeta) + L(\zeta),$$

where a_k counts the number of closed walks of length k starting and ending at node i and b_k counts all the open walks of length k starting at i and ending at any node $j \neq i$. Let us consider

$$\begin{aligned}\frac{d}{d\zeta}\left(\frac{Q(\zeta)}{Q(\zeta)+L(\zeta)}\right) &= \frac{L(\zeta)Q'(\zeta) - L'(\zeta)Q(\zeta)}{[Q(\zeta)+L(\zeta)]^2} \\ &= \frac{\left(2a_2b_1\zeta^2 + \cdots + 2a_2b_k\zeta^{k+1} + \cdots\right) - \left(b_1 + 2b_2\zeta + a_2b_1\zeta^2 + \cdots + b_1a_k\zeta^k + \cdots\right)}{[Q(\zeta)+L(\zeta)]^2}\end{aligned}$$

Then, for certain $\zeta < 1$ the numerator of the previous expression is negative, which means that the ratio $\frac{\mathscr{C}_i(\zeta)}{\mathscr{R}_i(\zeta)}$ is monotonically decreasing with ζ. For instance, let us make a second order approximation to the polynomials $Q(\zeta)$ and $H(\zeta)$. Then, we have

$$\frac{Q(\zeta)}{H(\zeta)} = \frac{1 + \frac{1}{2}\zeta^2 k_i}{1 + \zeta k_i + \frac{1}{2}\zeta^2(k_i + P_{2,i})},$$

where $P_{2,i}$ is the number of paths of length 2 (wedges) starting at node i. In an ER graph $\mathbb{E}(k_i) = (n-1)p$ and $\mathbb{E}(P_{2,i}) = (n-1)^2p^2 - (n-1)p$. Thus,

$$\frac{Q(\zeta)}{H(\zeta)} \approx \frac{1 + \frac{1}{2}\zeta^2(n-1)p}{1 + \zeta(n-1)p + \frac{1}{2}\zeta^2(n-1)^2p^2} = \frac{1 + \frac{\bar{k}}{2}\zeta^2}{1 + \bar{k}\zeta + \frac{\bar{k}^2}{2}\zeta^2},$$

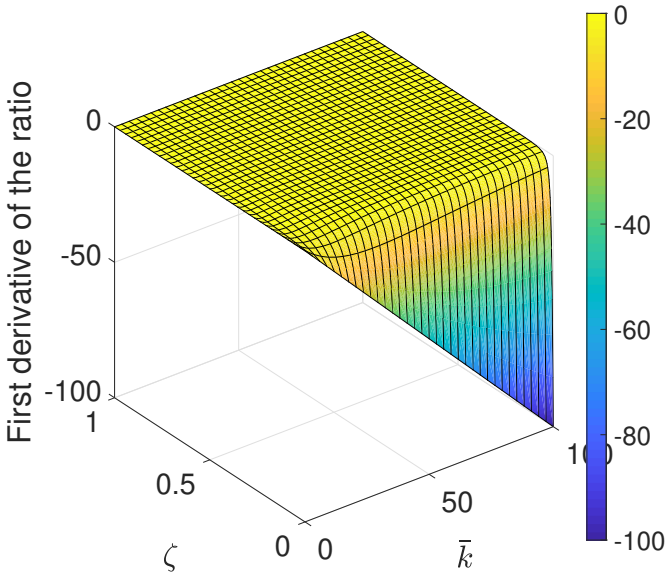


Figure 1.6: Illustration of the behaviour of the derivative of the ratio $\frac{\mathscr{C}_i(\zeta)}{\mathscr{R}_i(\zeta)}$ for values of $0 \leq \zeta \leq 1$ and for the parameter $\bar{k} \geq 1$.

where $\bar{k} = (n-1)\,p$ is the mean degree. The first derivative of this rational function is

$$\frac{d}{d\zeta}\left(\frac{Q(\zeta)}{H(\zeta)}\right) = \frac{2\bar{k}^2\zeta^2 - \left(4\bar{k}\left(\bar{k}-1\right)\zeta + 4\bar{k}\right)}{\left(2 + 2\bar{k}\zeta + \bar{k}^2\zeta^2\right)^2},$$

which is always negative for any $\bar{k} \geq 1$ and $0 \leq \zeta \leq 1$ as can be seen in figure 1.6. Moreover, the absolute value of this derivative increases as $\bar{k}$ decreases, implying a slower decay in the function $\frac{\mathscr{C}_i(\zeta)}{\mathscr{R}_i(\zeta)}$ for lower densities.

To conclude this section, we want to focus on the rankings produced by the two main centrality measures $\mathscr{R}_i$ and $\mathscr{C}_i$ and on the similarities between them. In particular, we are interested in determining if, or for what type of networks, the different centrality measures provide similar rankings. To this end, we display in table 1.1 the Spearman correlation coefficient between the risk dependent centrality $\mathscr{R}_i$ and the circulability $\mathscr{C}_i$ for different graph densities and for various values of ζ. On average, we observe a strong positive monotonic dependence between the two centrality measures. As expected, the two measures tend towards the perfect monotonicity as the density arises. It is noteworthy the behaviour with respect to ζ. The higher dependence is observed in a low-risk framework ($\zeta = 0.1$), while a slight reduction is noticeable when higher risk contexts are analyzed, providing again an empirical evidence of the fact that differences between nodes are increased in stressed conditions. Furthermore, this result is in line

with the higher incidence of $\mathscr{C}_i$ on $\mathscr{R}_i$ as ζ vanishes, discussed in the previous lines. For the sake of brevity, we do not report the Spearman correlation between $\mathscr{R}_i$ and $\mathscr{T}_i$. However, in all cases, the coefficient is larger than 0.9999.

Table 1.1: Spearman correlation coefficients between $\mathscr{C}_i$ and $\mathscr{R}_i$ in ER graphs with 100 vertices at different densities and different values of ζ.

		Density				
		0.1	0.3	0.5	0.7	0.9
ζ	0.1	0.9947	0.9967	0.9971	0.9994	0.9998
	0.5	0.9844	0.9950	0.9966	0.9994	0.9998
	1.0	0.9813	0.9950	0.9966	0.9994	0.9998

1.6 Analysis of real-world financial networks

In this section, we perform some empirical studies in order to assess the effectiveness of the proposed approaches. We consider two different families of networks. In the first one, we collected daily returns of a dataset referred to the time-period ranging from January 2001 to December 2017, that includes 102 leading U.S. stocks constituents of the $S\&P$ 100 index at the end of 2017. Data have been downloaded from Bloomberg. Returns have been split by using monthly stepped six-months windows. It means that the data of the first in-sample window of width six-month are used to build the first network. The process is repeated rolling the window one month forward until the end of the dataset is reached, obtaining a total of 199 networks. The first network, denoted as "1-2001"covers the period 1^{st} of January 2001 to 30^{th} of June 2001. The latter one ("7-2017") covers the period 1^{st} of July 2017 to 31^{th} of December 2017.

Hence, for each window, we have a network $\Gamma_t = (V_t, E_t)$ (with $t = 1, ..., 199$), where assets are nodes and links are weighted by computing the correlation coefficient ${}_t\rho_{i,j}$ between the empirical returns of each couple of assets. Notice that the number of assets can vary over time. Indeed, as mentioned, we have considered the 102 assets constituents of the $S\&P$ 100 index at the end of 2017. Some of these assets have no information available for some specific time periods. Therefore, in each window, we have considered only assets, whose observations are sufficiently large to assure a significant estimation of the correlation coefficient. However, it is not the aim of this work to deal with the effects of alternative estimation methods. As a consequence, the number of nodes in the 199 networks varies from 83 to 102 during the time-period.

Then, we follow the methodology proposed in [79, 80] and we use the non-linear

transformation, based on distances ${}_td_{i,j}$: ${}_td_{i,j} = \sqrt{2(1 - {}_t\rho_{i,j})}$. The distance matrix $D_t = [{}_td_{i,j}]_{i,j\in V_t}$, with elements $0 \leq {}_td_{i,j} \leq 2$, becomes the weighted adjacency matrix of the graph Γ_t. As proposed in [80], we extract the minimum spanning tree T_t. This is a simple connected graph that connects all n_t nodes of the graph with $n_t - 1$ edges such that the sum of all edge weights $\sum_{{}_td_{i,j}\in T_t} {}_td_{i,j}$ is minimum. As shown in [80], this minimum spanning tree, as a strongly reduced representative of the whole correlation matrix, bears the essential information about asset correlations. Furthermore, the study of the centrality of nodes and the analysis of the evolution of the tree over time are two critical issues in portfolio selection problem (see [80, 81, 82]).

The second dataset consists of a network of the top corporates in US in 1999 according to Forbes magazine. The network is constructed as follows. First we consider a bipartite network in which one set of nodes consists of companies and the other of directors of such companies. As one director can be in more than one company, we make a projection of this bipartite graph into the company-company space. In this way, the nodes represent corporations and two corporations are joined by an edge if they share at least one director. We consider two versions of this network, in the first we use the number of directors shared by two companies as an edge weight, and in the second we use the binary version of the first. We will refer to these as to the weighted and binary network, respectively. The network has 824 nodes, made up of one giant component of 814 nodes. We selected the giant component, with its binary and weighted adjacency matrices. For a comprehensive description of this network see, for instance, [83]. Networks, derived by both datasets, have been studied by computing the total communicability, circulability and transmissibility for each node with ζ varying in $(0, 1]$ with step 0.01.

1.6.1 Network of assets

Starting from the asset trees T_t, we measure the relevance of each node by using the risk-dependent centrality $\mathscr{R}_i$ and by testing different values of ζ. We consider in figure 1.7 the rankings' distribution of each asset. Different outcomes of each distribution have been obtained by computing the rankings based on $\mathscr{R}_i$ for alternative values of ζ in the interval $(0, 1]$ with step 0.01. These results regard the first network "1-2001", namely, the network based on data that cover the period 1^{st} of January 2001 to 30^{th} of June 2001. We observe that some nodes show a significant variability according to different values of ζ. Indeed, some assets have climbed more than 20 positions in the ranking when ζ increases. For instance, Amazon (node 7 in figure 1.7) moved from position 66 to 41 in case of low and high risk, respectively. Vice versa, Exelon Group

(node 32 in 1.7) lowered its ranking from 15 to 46. On the other hand, the most central nodes in the network remain very central also when external risk is very high. We have indeed that the top 6 is quite stable for different values of ζ. Top assets only exchange a bit their position, preserving their central role. For instance, United Technologies Corporation (node number 79 in 1.7) is at the top of the ranking, independent of ζ.

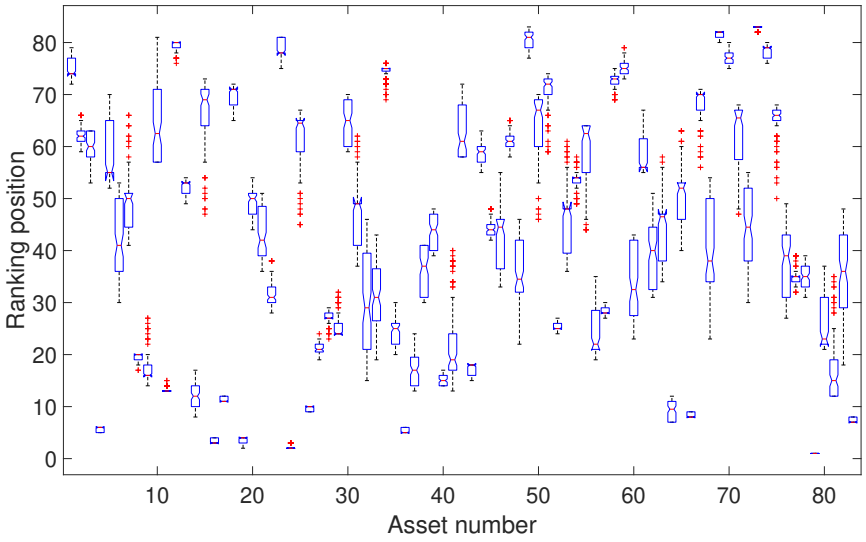


Figure 1.7: Figure reports the distribution of nodes' rankings based on $\mathscr{R}_i$ with respect to ζ. For each distribution, the set of outcomes is given by the rankings of $\mathscr{R}_i$ computed for alternative values of ζ. Results regard the network T_1, i.e. the asset-tree in the first window $1-2001$.

If we consider the period of the global financial crisis of 2007-2008 (see figures 1.9 and 1.10), we observe an increase in the rankings' volatility. In shock periods, centrality of nodes is more affected by the value of ζ. In particular, to catch rankings' volatility, we report in figure 1.8 the standard deviations of rankings of each asset computed varying ζ. In shocks periods, results confirm higher average volatility as well as positive skewed distributions because of a greater number of assets whose ranking is highly affected by the value of ζ. We also tested that differences in average volatility are significant by means of a paired t-test, useful for comparing the same sample of assets at different time periods. When the network 1-2001 is compared with the two networks covering period of crisis (End 2007 or End 2008), we obtain p-values around 10^{-5} and 10^{-8} that confirm strong evidence against the null hypothesis that the average difference between the two samples is zero. As expected, the test is not statistically significant (p-value is

0.31) when networks covering period of the global financial crisis are compared.

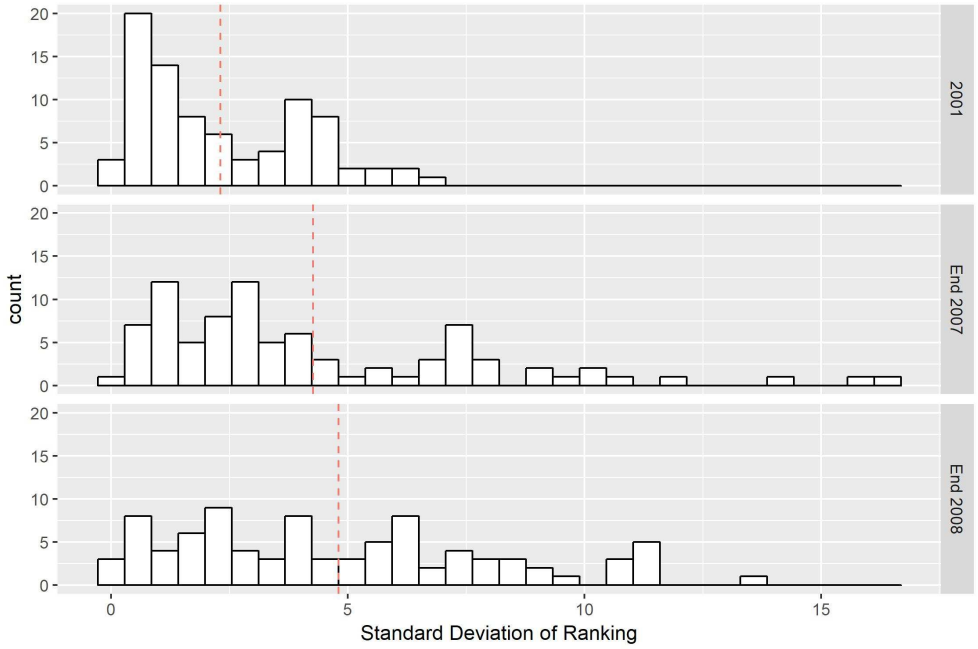


Figure 1.8: Figure reports the distribution of standard deviations of nodes' rankings based on $\mathscr{R}_i$ with respect to ζ. For each distribution, the set of outcomes is given by the standard deviation of rankings of $\mathscr{R}_i$ computed for alternative values of ζ. Results regard respectively the network in the first window $1-2001$, at the end of 2007 and at the end of 2008. The dotted red lines indicate the average standard deviation: values are equal to 2.31, 4.27 and 4.80, respectively.

Concerning the behaviour of specific assets, we observe, for instance, that some assets move down by approximately 60 positions from a low risk to an high risk framework. Two examples are represented by Danaher Corporation and Honeywell International (assets 28 and 43, respectively, in figure 1.9). Instead, Accenture PLC (node 3 in 1.9) increased its ranking from position 61 to 11.
Even top central nodes are affected by ζ as the volatilities of their rankings show. It is instead confirmed the relevance of United Technologies Corporation (node number 82 in 1.9 and 83 in 1.10) that is again at the top of the ranking at the end of 2017, independent of ζ. At the end of 2008, the centrality of this asset is also confirmed, although, a bit of variability in the ranking is observed for this firm.

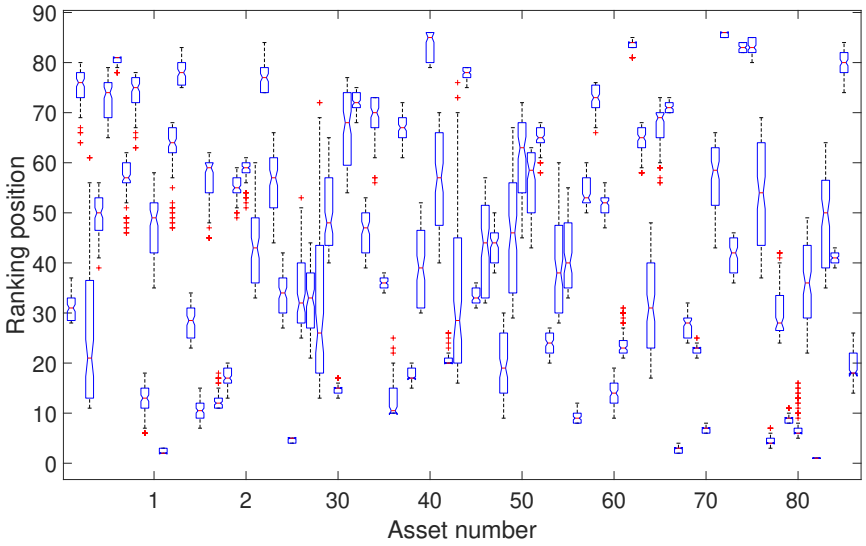


Figure 1.9: Figure reports the distribution of nodes' rankings based on $\mathscr{R}_i$ with respect to ζ. For each distribution, the set of outcomes is given by the rankings of $\mathscr{R}_i$ computed for alternative values of ζ. Results regard the asset-tree at the end of 2007.

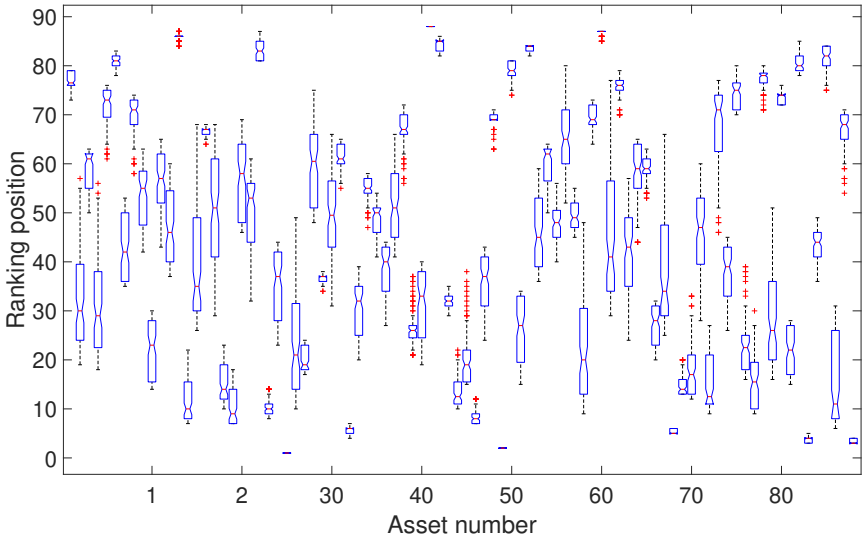


Figure 1.10: Figure reports the distribution of nodes' rankings based on $\mathscr{R}_i$ with respect to ζ. For each distribution, the set of outcomes is given by the rankings of $\mathscr{R}_i$ computed for alternative values of ζ. Results regard the asset-tree at the end of 2008.

1.6.2 US corporate network

We now analyze the network of US top corporates in 1999 according to Forbes magazine. Before starting our analysis let us explain the importance of studying a spreading dynamics on this network. According to this network, the board of directors of a given corporation is formed by a few members, some of which are also present in the board of other corporations. Then, such directors serving on more than one board can act as spreaders of information between the corresponding corporations. Such information can be about future (favorable or unfavorable) economic situations, alarms, market opportunities, or anything that could be of interest to the companies in which the director is. Due to the global connectivity of the system, such "information" can be spread across the whole network "infecting" all the corporations in a relatively short time. As we have mentioned before, epidemiological models have also being used for modeling such propagation dynamics (see Section 1.1.).
Hence, we devote this section to the investigation about whether a significant increase of the risk-dependent centrality is a proxy of the vulnerability of the corporate to financial infections propagating on the network. At first, we should remark the fact that the network we are considering here was built based on data corresponding to year 1999. At this year the level of stress of the international economic system was relatively high due to the fact that the East Asian financial crisis occurred in the years 1997–1998, which was also followed by the Russian default of 1998. The two aforementioned financial crises had a ripple effect on the US market. In the literature, for instance, it is well-documented the so-called "fire-sale" FDI (Foreign Direct Investment) phenomenon, that is, the surge of massive foreign acquisitions of domestic firms during a financial crisis [84, 85].Thus, the level of stress and infectability of the system for the next few years after 1999 (we will eventually see that these correspond to the period 2000-2002) is expected to be significantly larger than in the subsequent years when the effects of these crises gradually relaxed. Therefore, we proceed our analysis by considering that the level of infectability in 1999 is high and we investigate the effects of relaxing such a condition to lower levels of stress. That is, we start by assuming that in 1999 the external market turmoil could be represented by a value of $\zeta = 1$ and we want to find out how the companies change their ranking positions in term of risk-dependent centrality[1] $\mathscr{R}_i$ as ζ vanishes. To this purpose, we set up different initial conditions in the contagion model described by 1.15, assigning to each year a different

[1]The analysis has been also developed for circulability and transmissibility, but, since the significantly high rank correlation between $\mathscr{R}_i$ and $\mathscr{T}_i$ (with Spearman correlation coefficients larger than 0.99), we focus here only on $\mathscr{R}_i$.

value of the infectability parameter γ, according to the environmental conditions of the market. Therefore, we let ζ factors reduce year by year in order to reflect a reduction in the overall stress on the network. In particular, we decrease ζ linearly from 1 to 0 in the period 1999-2003. Therefore, rankings based on the risk-dependent centrality computed for $\zeta = 1$ allow to assess the relevance of each corporate in 1999. Lowering ζ, we test how the positions of firms vary over time when the external risk reduces. It is noteworthy that the connection between this parameter and the risk could be quite loose but as provided by the following analysis the model seems to work quite well in describing firms that reduce their SVC in the period.
The variation of rankings is then compared with the pattern of the shareholder value creation (SVC) over time. According to the *OECD Principles of Corporate Governance*, corporations should be run, first and foremost, in the interests of shareholders (OECD 1999). Therefore, companies should work to increase their shareholder values. Increasing shareholders value cannot be done without risk. It is known [86] that in the shareholder value model, companies usually take more risk than needed in order to maximize SVC. As a consequence of this additional risk, companies acquire debts which could make them unstable and more exposed to the risk of bankruptcy. Acquiring large debts is seen as conductive to increasing shareholder value, due to the potential of the company to increase value when it has started from a low baseline. Thus, there is a relation between SVC and risk, because in searching for large SVC the companies increase their risks to attract more investors and increasing potential value gain, but, at the same time, the risk also puts the company in a more vulnerable position to bankruptcy and collapse.
To support our interpretation, we make use of SVCs of the companies[1] in the S&P500 for the period 1999-2003, that have been collected by FernÃ¡ndez and Reinoso (see [87]). Hence, we use SVC as a proxy for risk. Indeed, the global average of SVC reflects very well what happens for the period 1999-2003. After the financial crisis of 1998 the world was at a higher level of risk which is reflected by a dramatic drop of the SVC in year 2000 from a positive value in 1999 to a negative one in 2000. This situations remained until 2002, but eventually recovered to positive in 2003 (see figure 1.11). It is noteworthy that the data for SVC was reported by Fernandez and Reinoso for the years 1993-2003. From this long period we select the segment 1999-2003 which contains exactly the valley produced from the financial crisis of 1998 and also because the data used for building the corporate network is of 1999. That is, it corresponds to a segment

[1] In particular, we use a sample of 337 companies in our network whose SVCs are made available in the dataset available in [87].

in time in which the world economy drop due to a crisis and then eventually recovered from it.

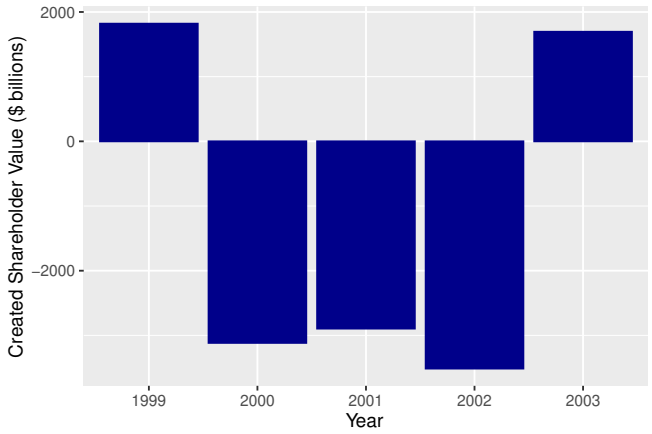


Figure 1.11: Total created shareholder value ($ billion) of firms constituent the index S&P500 for the period 1999-2003 (data taken from [87]).

We focus our statistical analysis on the predictability of the risk-dependent centrality on the evolution of the SVC. We consider the evolution of the SVC of a company for the period 1999-2003, which is the period immediately after the network of corporate elite in US was built. As a proxy for the evolution of the SVC of a company we consider the Pearson correlation coefficient ρ of the ranking position of the company based on SVC versus the reciprocal of the year. In this case, a negative (positive) value of ρ indicates that the corresponding company decreases (increases) its SVC from 1999 to 2003 when the global external infectability decreases. Therefore, we apply a Linear Discriminant Analysis (LDA) to classify the companies into two groups: (i) those with negative trend in the SVC for this period, and (ii) those with a positive one. The only predictor used for this classification is the parameter $\Delta\text{Rank}(\mathscr{R}_i)$. This parameter is the difference between the ranking position of the company i when $\zeta = 1$ and the ranking position of the same company when $\zeta = 0.01$. In other words, a negative (positive) value of $\Delta\text{Rank}(\mathscr{R}_i)$ means that the company dropped (increased) its exposure to risk when the infectability of the system is lower.

Before proceeding with the application of the LDA on the whole sample at disposal, we eliminate a few companies whose correlation coefficient between SVC and the reciprocal of the year is marginal (i.e., close to zero). We test empirically the effect produced by the removal of companies for which $|\rho| < a$ for different values of the threshold a, e.g., $a = 0.01, 0.025, 0.05, 0.075, 0.1$. The best classification of the compa-

nies into the two groups analyzed is obtained by eliminating those companies for which $|\rho| < 0.05$. In this case the total accuracy of the LDA model is 60.5%. That is, 200 out of 332 of the companies are classified correctly in their respective groups representing their trends in shrinking SVC or expanding it. In particular, the fitted LDA model is $\hat{Y}_i = -0.3177 + 0.0102\Delta\text{Rank}(\mathscr{R}_i)$, where $\hat{Y}_i$ is the predicted response variable of our analysis that allows to classify companies in their respective group. The positive coefficient of the variable $\Delta\text{Rank}(\mathscr{R}_i)$ indicates that: (i) increasing the exposure to risk ($\Delta\text{Rank}(\mathscr{R}_i) > 0$) tends to expand the SVC of the company, and (ii) decreasing the exposure to risk ($\Delta\text{Rank}(\mathscr{R}_i) < 0$) tends to shrink the SVC of the company. For both groups, we report in figure 1.12(a) a comparison between the predicted value with the LDA and the observed value for each firm. Red squares below the line and blue circles over the line are well classified, while blue circles below the line and red squares over the line are wrongly classified. Furthermore, in figure 1.12(b) we report the related confusion plot,where the number of true negative and true positive are on the anti-diagonal (bottom and upper parts, respectively) and the number of false negative and false positive are on the main diagonal (bottom and upper parts, respectively). It is noticeable the low classification performance of the model, when only companies that expand their SVC are considered (in this regard, see in figure 1.12(b) companies that belong to the observed class denoted with the sign +). For instance, from 147 companies in the network which increase their SVC in the period 1999-2003 only 36 are correctly predicted by $\Delta\text{Rank}(\mathscr{R}_i)$ in their class. On the contrary, from the 175 companies that shrink their SVC in the period 1999-2003, the variable $\Delta\text{Rank}(\mathscr{R}_i)$ correctly predicts 157 companies in this class. That is, the risk-dependent centrality of the companies clearly identifies about 90% of the companies which will shrink their SVC in the period 1999-2003, using only data referring to the year 1999. In plain words, our results indicate that diminishing the exposure to risk when the external conditions of infectability are low, with high probability, reduces the SVC of a company.

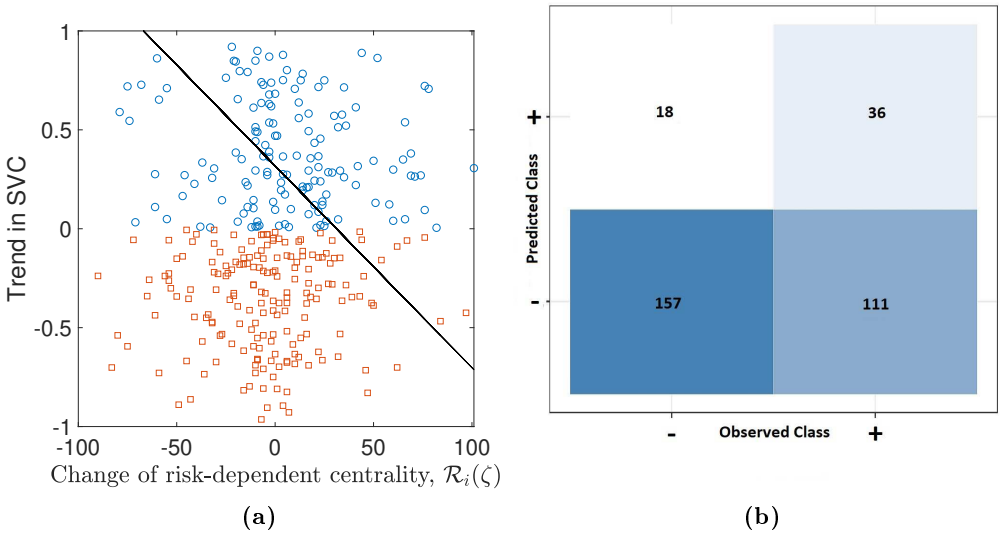
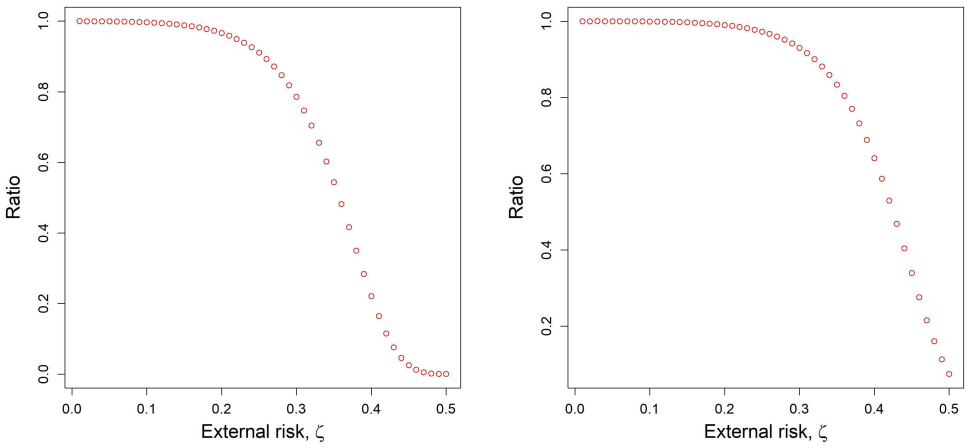


Figure 1.12: (a) Illustration of the linear discriminant analysis (LDA) model classifying the trend of corporations into those shrinking their SVC (red squares) and those expanding it (blue circles). The black line represents the LDA model based on the change of $\mathscr{R}_i(\zeta)$ for the values of $\zeta = 0.01$ and $\zeta = 1.0$ to predict the trend in the SVC. Red squares below the line and blue circles over the line are well classified, while blue circles below the line and red squares over the line are wrongly classified. The LDA classifies correctly about 90% of all companies who shrank their SVC (red squares). (b) Plot of the confusion matrix. On the x-axis we report the true class (Observed Class), on the y-axis the predicted class (Output Class). The number of true negative and true positive cases are on the anti-diagonal (bottom and upper parts, respectively) and the number of false negative and false positive cases are on the main diagonal of the matrix (bottom and upper parts, respectively). In the class minus (plus) we consider companies with negative (positive) trend in the SVC for the period 1999-2003

Let us conclude with the following remark. Even if "good" companies increase their risk-centrality ranking as ζ vanishes, it is worth noting that this occurs when the global stress in the market is very low. When the infectability rate is very low, the absolute probability of getting infected also remains very low for both "good" and "bad" companies. To show this fact, let us consider that, according to our model, the probability that a given corporate is not affected by a crisis propagating inside the network is given by $1 - x_i(t) = \alpha e^{-\frac{\beta}{\alpha}(\mathscr{R}_i - 1)}$, where again β and $\alpha = 1 - \beta$ are the initial probabilities to have infected and not-infected nodes, respectively. Hence, the ratio between the probabilities of two nodes i and j to pass successfully through a crisis is given by $e^{\frac{\beta}{\alpha}(\mathscr{R}_j - \mathscr{R}_i)}$. We compute these ratios for different couples of corporates

operating in a similar sector, a "good" one and a "bad" one (see figure 1.13).



(a) (b)

Figure 1.13: Figures display the ratios between the probabilities of not being infected by a crisis for two different couples of Corporates: a) Lucent Technologies Inc. over General Electric Co. b) Morgan Stanley Co. over Bank One Corp. It is noteworthy that Lucent Technologies Inc. and Morgan Stanley Co. reduced their rankings over time, while General Electric Co. and Bank One Corp. increased their rankings.

As expected, at low ζ the probability of not being infected by a crisis is the same for both high and low risk-centrality companies. But this ratio decreases very quickly as ζ increases and this means that for companies that reduced their risk (e.g., Lucent Technologies, Morgan Stanley, Union Carbide and American Express) the probabilities to stay safe during a crisis are very small if compared with the analogous probabilities for companies that increased their risk (e.g., General Electric, Bank One, Ashland and Bank of America).

1.7 Ranking interlacement

During the analysis of the two real-world networks studied above, we have noticed that with the change of ζ some nodes vary their ranking significantly, to the point of changing their positions relative to each other. For instance, in figure 1.14 we illustrate six pairs of corporates that interlace their positions with the change of the global infectability in the network. In the first pair, 1.14(a), we see that at low levels of infectability, i.e., $\zeta \to 0$, J.P. Morgan&Co Inc. (red) occupies a position in the ranking of $\mathscr{C}_i$ more at the bottom

than Bank of America Corp. (blue). That is, at low global infectability J.P. Morgan&Co is exposed to less risk than Bank of America. However, when the global infectability in the network increases ($\zeta \to 1$), Bank of America is exposed to less risk than J.P. Morgan&Co. A similar interlacement is observed between the other couples in figure 1.14. For instance, in 1.14(f), the interlacement between rankings for General Motors Corp. (red) and Boeing Co. (blue) occurs at a smaller value of ζ than for the previous cases. Before proceeding with the analysis of this phenomenon, we would like to remark that the existence of ranking interlacement means that the ranking of the nodes in a network based on the risk-dependent centralities is not unique and fixed as in the case of other classical centrality measures, e.g., degree, eigenvector, closeness, betweenness. Here instead the ranking of nodes depends on the global external conditions to which the network is submitted.

In order to shed light on the issue of ranking interlacement we will make use of different representations of the risk-dependent total communicability $\mathscr{R}_i(\zeta)$ and circulability $\mathscr{C}_i(\zeta)$ measures (the transmissibility is obtained as the difference of these two and can be treated accordingly). First, expanding the matrix exponential in a power series gives the representation

$$\mathscr{R}_i(\zeta) = \left(e^{\zeta A}\vec{1}\right)_i = \sum_{k=0}^{\infty} \frac{\zeta^k}{k!} w_i^{(k)}, \tag{1.34}$$

where $w_i^{(k)} = \left(A^k\vec{1}\right)_i$ denotes the number of walks of length k starting from the node i, with $w_i^{(0)} = 1$. In particular, $w_i^{(1)} = k_i$, the degree of node i. Similarly,

$$\mathscr{C}_i(\zeta) = \left(e^{\zeta A}\right)_{ii} = \sum_{k=0}^{\infty} \frac{\zeta^k}{k!} w_{i,i}^{(k)}, \tag{1.35}$$

where now $w_{i,i}^{(k)} = \left(A^k\right)_{ii}$ is the number of closed walks of length k through node i; in particular, $w_{i,i}^{(0)} = 1$, $w_{i,i}^{(1)} = 0$, $w_{i,i}^{(2)} = k_i$, and $w_{i,i}^{(3)} = 2t_i$, where t_i is the number of triangles node i participates in.

Second, we recall that the spectral theorem yields the formulas

$$\mathscr{R}_i(\zeta) = \sum_{k=1}^{n} e^{\zeta\lambda_k}\left(\psi_k^T\vec{1}\right)\psi_{k,i}, \quad \mathcal{C}_i(\zeta) = \sum_{k=1}^{n} e^{\zeta\lambda_k}\left(\psi_{k,i}\right)^2. \tag{1.36}$$

Using 1.34-1.35, we readily see that both functions of ζ are *absolutely monotonic* for $\zeta > 0$, i.e. they are positive and infinitely differentiable on $(0,\infty)$, with all the derivatives being nonnegative. In particular, both functions are strictly increasing and strictly convex.

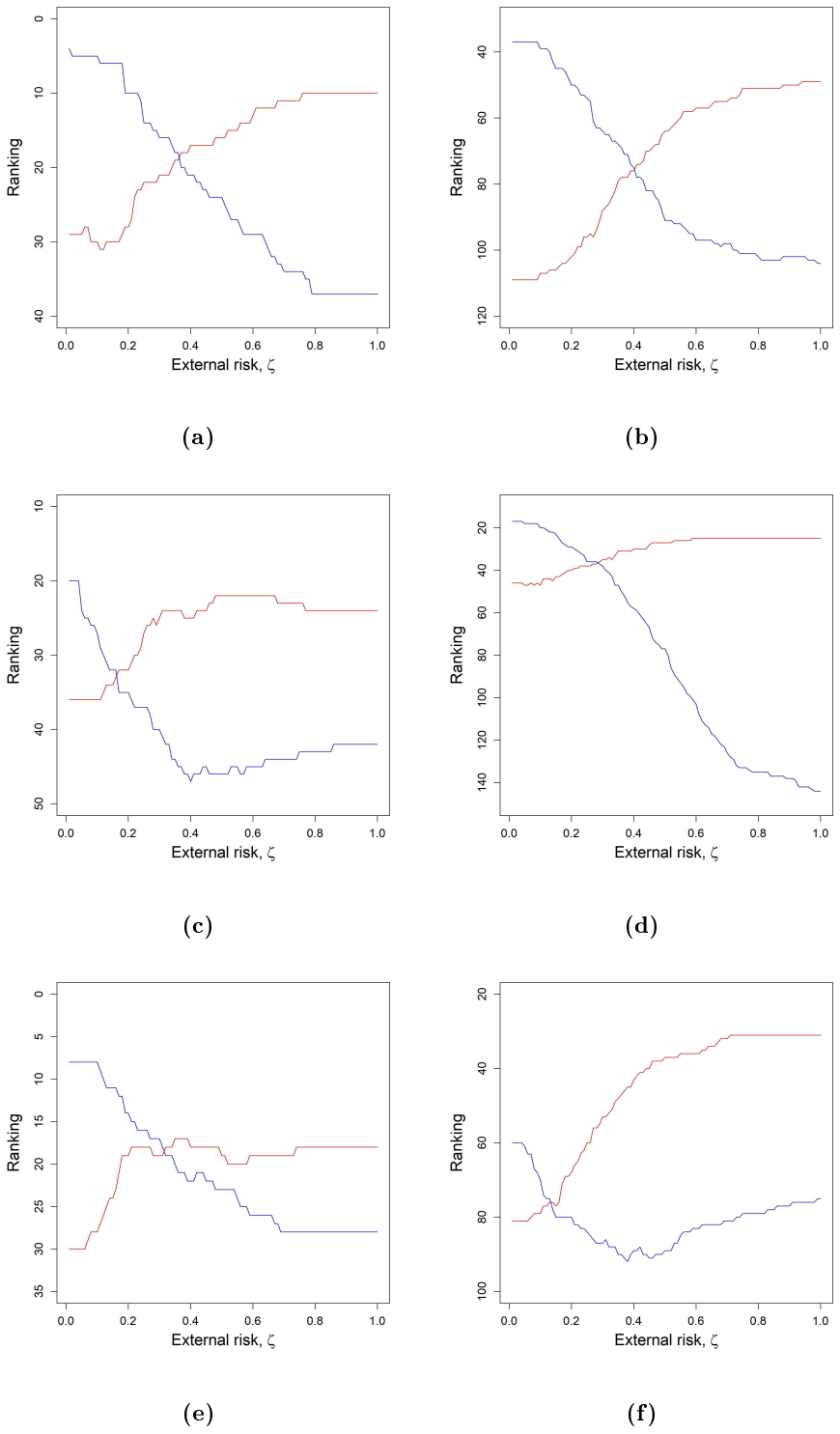


Figure 1.14: Illustration of the Circulability Ranking Interlacement for a) J.P. Morgan&Co Inc. (red) and Bank of America Corp. (blue) b) Pfizer Inc. (red) and Ashland Inc (blue) c) Morgan Stanley & Co. (red) and Bank One Corp. (blue) d) AT&T Corp. (red) and Airtouch Communications Inc. (blue) e) Union Carbide Corp. New (red) and AON Corp. (blue) f) General Motors Corp. (red) and Boeing Co. (blue)

Definition 1. *We say that the rankings of node i and node j based on the circulability interlace at $\zeta^* > 0$ if $\mathscr{C}_i(\zeta^*) = \mathscr{C}_j(\zeta^*)$ and there exists an $\varepsilon > 0$ such that $\mathscr{C}_i(\zeta) - \mathscr{C}_j(\zeta)$ changes sign exactly once in $(\zeta^* - \varepsilon, \zeta^* + \varepsilon)$.*

In other words, nodes i and j interlace at $\zeta^* > 0$ if the plots of $\mathscr{C}_i(\zeta)$ and $\mathscr{C}_j(\zeta)$ cross for $\zeta = \zeta^*$. We note that, in principle, it is possible to have $\mathscr{C}_i(\zeta^*) = \mathscr{C}_j(\zeta^*)$ for some value of ζ^* without interlacing taking place. Two cases are possible: in the first one, the two curves touch at the isolated point ζ^* (without crossing), and in the second one the two functions are identical on an open neighborhood of ζ^* and, therefore, for all ζ since they are analytic functions. In practice, either scenario is very unlikely to occur, at least for real world networks. Note that points of tangency must satisfy the additional condition $\mathscr{C}_i'(\zeta^*) = \mathscr{C}_j'(\zeta^*)$.

An analogous definition can be given for the ranking based on other ζ-dependent measures, like the total communicability $\mathscr{R}_i(\zeta)$. In the following we limit our discussion to the interlacing of rankings according to the circulability, but analogous observations hold for the total communicability and transmissibility functions.

Identifying the interlacing points (if they exist) requires to find the roots of the transcendental equation $\mathscr{C}_i(\zeta) - \mathscr{C}_j(\zeta) = 0$, or

$$\Psi(\zeta) := \sum_{k=1}^{n} e^{\zeta\lambda_k} \left[\psi_{k,i}^2 - \psi_{k,j}^2\right] = 0.$$

Even if we knew the eigenvalues and eigenvectors of A explicitly, there is no general closed form expression for the roots of the transcendental function Ψ. Of course one could resort to numerical root-finding techniques, but this would be impractical for large networks. Here and below we give a qualitative discussion followed by a heuristic approach that yields approximations that seem to work well in practice.
We begin with the following result. It applies to both circulability and total communicability based rankings, and in fact for a much larger class of parameter-dependent centrality ranking functions, including Katz centrality [88]. We remind the reader that we restrict the risk rate ζ to positive values.

Theorem 4. *Let i and j be two nodes with different eigenvector centrality: $\psi_{1,i} \neq \psi_{1,j}$. Then the number of interlacing points for i and j is necessarily finite (possibly zero).*

Proof. Let us assume that there is at least one pair of nodes, i and j, whose rankings interlace, so that $\Psi(\zeta) = 0$ has at least one positive root. Observe that the ranking of node i provided by $\mathscr{C}_i(\zeta)$ is identical to that obtained using

$$\hat{\mathscr{C}}_i(\zeta) = e^{-\zeta\lambda_1}\mathscr{C}_i(\zeta) = \psi_{1,i}^2 + \sum_{k=2}^{n} e^{\zeta(\lambda_k - \lambda_1)}\psi_{k,i}^2.$$

As this quantity tends monotonically to $\psi_{1,i}^2$ for $\zeta \to \infty$, there exists a $\bar{\zeta}$ such that no rank interlacing with node j can occur for $\zeta > \bar{\zeta}$, since all the node rankings must stabilize on the eigenvector rankings in the large ζ limit. Hence, all interlacing points must fall within the compact interval $[0, \bar{\zeta}]$. Suppose that the number of interlacing points is infinite. By the Bolzano-Weierstrass Theorem, this set has a point of accumulation. But since $\hat{\Psi}(\zeta) := e^{-\zeta\lambda_1}\Psi(\zeta)$ is analytic, and zero on this set, it must be identically zero everywhere, which contradicts the assumption that there is at least one interlacing point in $(0, \infty)$. □

As a consequence:

Corollary 4.1. *If all nodes in the network have different eigenvector centralities, the total number of interlacing points is finite (possibly zero).*

A sufficient condition for the existence of at least one interlacing point for the pair of nodes i and j is that $k_i \geq k_j$ (or $k_j \geq k_i$) while $\psi_{1,i} < \psi_{1,j}$ (resp., $\psi_{1,i} > \psi_{1,j}$). This follows from Theorem 2: since $\mathscr{C}_i(\zeta)$ interpolates smoothly between degree centrality and eigenvector centrality, the only way that a node with higher degree can have lower eigenvector centrality than another node is that the corresponding circulabilities interlace at some value $\zeta^* > 0$. If more than one interlacing point exists, this number must be odd, for otherwise the node with higher degree would also have higher eigenvector centrality than the other node. That the above condition is not necessary is made clear considering the possibility of an even number of interlacing points. A necessary condition for the existence of at least one interlacing point is that there exist at least two values of k, say k_1 and k_2, for which $(A^{k_1})_{ii} - (A^{k_1})_{jj}$ and $(A^{k_2})_{ii} - (A^{k_2})_{jj}$ have different sign. Indeed, it is obvious from equations 1.34 and 1.35 that if (say) $(A^k)_{ii} \geq (A^k)_{jj}$ for all k, then no rank interlacing point exists. That this condition may not be sufficient is suggested by the fact that the series expansions contain an infinity of terms.

We mention that the same problem has been studied, for a different centrality function (the Katz resolvent), by [89] independently of us.

1.7.1 A back of envelop approach

We now consider heuristics based on truncated series expansions. Let $k_0 \geq 3$ be the smallest value of k such that the sequence of values $\{(A^k)_{ii} - (A^k)_{jj}\}_{k \geq 2}$ undergoes a sign change (here zero is considered positive). If no such k_0 exists, then no interlacing can take place, as we already observed. We consider approximating $\mathscr{C}_i(\zeta)$ with its

truncation to an order $k \geq k_0$:

$$\mathscr{C}_i(\zeta) \approx 1 + \frac{1}{2!}\zeta^2 w_{i,i}^{(2)} + \frac{1}{3!}\zeta^3 w_{i,i}^{(3)} + \cdots + \frac{1}{k!}\zeta^k w_{i,i}^{(k)} = \tilde{\mathscr{C}}_i(\zeta), \tag{1.37}$$

where we recall that $w_{i,i}^{(k)} = (A^k)_{ii}$. We emphasize that this polynomial approximation assumes that ζ is small, since the error in it is $O(\zeta^{k+1})$. In alternative we can also use as a surrogate for $\mathscr{C}_i$ the same polynomial shifted by 1 and divided by ζ^2:

$$\frac{\tilde{\mathscr{C}}_i(\zeta) - 1}{\zeta^2} = \frac{1}{2!}w_{i,i}^{(2)} + \frac{1}{3!}\zeta w_{i,i}^{(3)} + \cdots + \frac{1}{k!}\zeta^{k-2} w_{i,i}^{(k)},$$

where now the error is $O(\zeta^{k-1})$. We can now use these polynomial approximations to try to locate, approximately, any interlacing points sufficiently small in magnitude. This requires finding the (positive) roots, if any, of the polynomial equation of degree $k-2$:

$$\begin{aligned} q(\zeta) = &\frac{(w_{i,i}^{(k)} - w_{j,j}^{(k)})}{k!}\zeta^{k-2} + \frac{(w_{i,i}^{(k-1)} - w_{j,j}^{(k-1)})}{(k-1)!}\zeta^{k-3} + \cdots \\ &+ \frac{(w_{i,i}^{(3)} - w_{j,j}^{(3)})}{3!}\zeta + \frac{(w_{i,i}^{(2)} - w_{j,j}^{(2)})}{2!} = 0. \end{aligned} \tag{1.38}$$

It is well known that for degree greater than or equal to 5 there is no closed form expression of the solutions of an algebraic equation involving only arithmetic operations and root extractions, so in general if $k \geq 7$ we will have to resort to numerical methods for solving 1.38. Evaluation of the coefficients requires computing the diagonal entries of powers of the adjacency matrix A, which can be expensive for very large graphs and large values of k.

As the simplest possible example, we consider the case where $w_{i,i}^{(2)} > w_{j,j}^{(2)}$ and $w_{i,i}^{(3)} < w_{j,j}^{(3)}$ (or vice-versa), i.e., $k_0 = 3$. Taking $k = k_0$, equation 1.38 becomes the linear equation

$$\frac{(w_{i,i}^{(3)} - w_{j,j}^{(3)})}{3!}\zeta + \frac{(w_{i,i}^{(2)} - w_{j,j}^{(2)})}{2!} = 0,$$

which admits the unique solution $\zeta^* = \frac{3(w_{i,i}^{(2)} - w_{j,j}^{(2)})}{w_{i,i}^{(3)} - w_{j,j}^{(3)}}$, which is of course positive. In terms of the degree of the nodes and the number of triangles in which they take place, this can be written in the form:

$$\zeta^* = \frac{3}{2}\left|\frac{k_i - k_j}{t_i - t_j}\right|. \tag{1.39}$$

In the case of weighted networks, the degree is replaced by the weighted degree or strength, and the number of triangles is replaced by the weighted number of cycles of

length 3, i.e., the weight of a cycle of length 3 is the product of the weights at its three edges. A priori, there is no reason to expect that this value is close to an actual interlacing point (assuming it even exists), since the behaviour of higher order terms may more than offset the influence of the negative term involving $t_i - t_j$. Better approximations might be obtained by considering higher order approximations; for example using $k = 4$ leads to an easily solved quadratic equation in ζ, $k = 5$ leads to a cubic, and so forth. In any case, these are heuristics whose usefulness can only be assessed experimentally on concrete examples. We emphasize that the use of power series truncation requires knowledge of k_0, since truncating the series at orders lower than k_0 would lead to an equation devoid of positive solutions and therefore to concluding that no interlacing points exist for a given pair of nodes, even if such points do exist.

It is also worth recalling Descartes's Rule of Signs, according to which the number of positive real roots of a polynomial (counted with their multiplicities) is equal to the number of sign changes in the (nonzero) coefficients or less than that by an even whole number, when the powers are ordered in descending order. If, moreover, the polynomial is known to have only real roots (as in the case of a symmetric adjacency matrix, i.e., of undirected networks) then the number of sign changes is exactly equal to the number of positive roots. It is then obvious that if the power series is truncated at order k_0, i.e., as soon as we observe the first sign change in the coefficients, then there will be exactly one positive root and therefore only one (approximate) interlacing point can be found by this method. A polynomial truncation of higher degree $k > k_0$ may have more than one positive root, depending on the number of changes in the coefficients (assuming the network is undirected). We will come back to this case shortly.

To exemplify the previous finding let us consider a pair of nodes with a small difference in their degree, e.g., $k_i - k_j = 2$, then $-(k_i - 2)^2 \leq (t_i - t_j) \leq k_i^2$, such that if, for instance, $k_i \leq 10$ and we let ζ vary from 0 to 0.1 we obtain the plot given in figure 1.15(a). As can be seen there are certain values of $\Delta = t_i - t_j < 0$ for which we can obtain positive and negative values of $\mathscr{C}_i - \mathscr{C}_j$. This is illustrated in 1.15(b) where we can see that when $-100 \leq \Delta \leq -40$ there are both positive and negative values of $\mathscr{C}_i - \mathscr{C}_j$. In other words, it is possible to find pairs of nodes for which $\mathscr{C}_i(\zeta_1) > \mathscr{C}_j(\zeta_1)$ and then $\mathscr{C}_i(\zeta_2) < \mathscr{C}_j(\zeta_2)$, which means that these nodes will change their ranking position in terms of the risk-dependent centrality when the values of ζ change even for a relatively narrow window. Notice that if $k_i - k_j = 2$, and $\Delta \geq -30$ such change is not observed for the corresponding range of ζ analyzed.

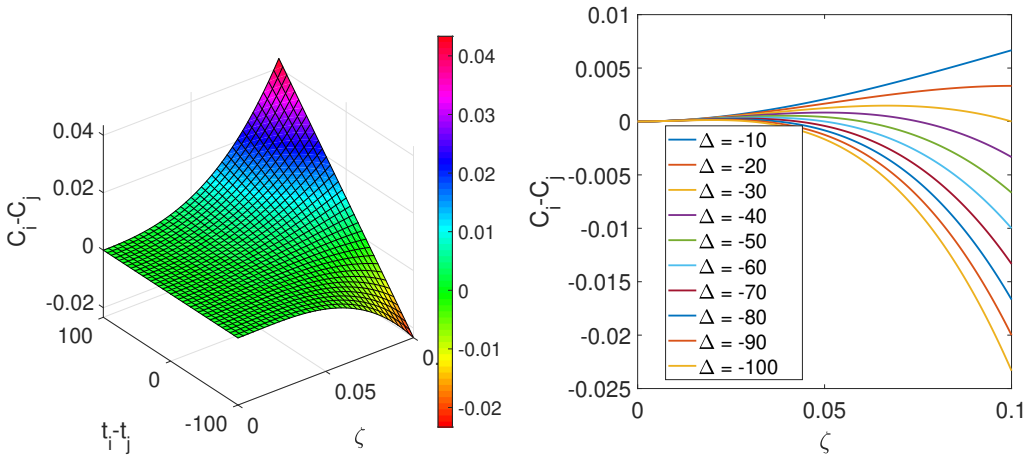


Figure 1.15: (a) Illustration of the change in the difference in the risk-dependent centrality of nodes having a small difference in degrees, $k_i - k_j = 2$, as a function of the difference in the number of triangles, $t_i - t_j$, and of the network infectivity risk ζ. (b) Some of the curves obtained for $k_i - k_j = 2$ and a given value of $\Delta = t_i - t_j$ as a function of ζ.

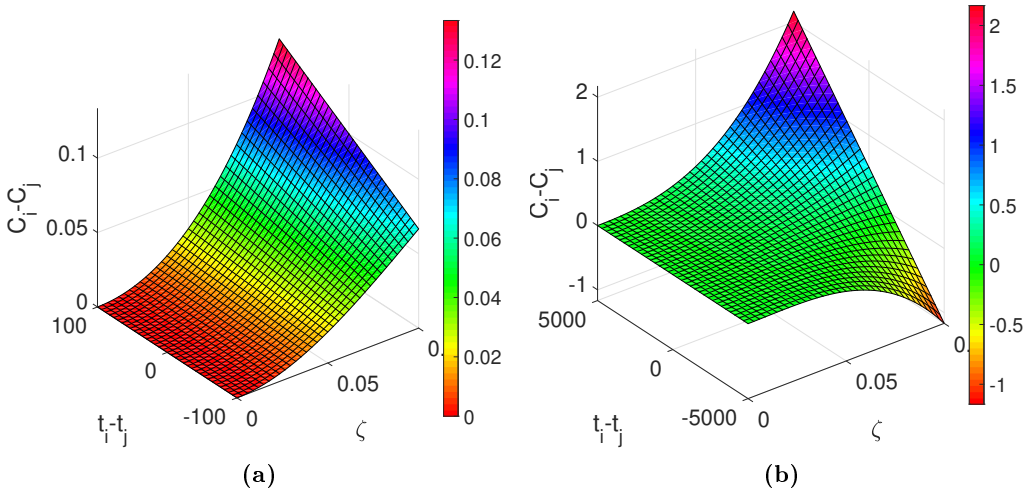


Figure 1.16: Illustration of the change in the difference in the risk-dependent centrality of nodes having a small difference in degrees, $k_i - k_j = 100$, as a function of the difference in the number of triangles, $-100 \leq \Delta \leq 100$ (a) and $-5000 \leq \Delta \leq 5000$ (b), and of the network infectivity (risk) ζ.

If we now consider a large difference in the node degrees, e.g., $k_i - k_j = 100$, and the same range of change for the difference in the number of triangles, e.g., $-100 \leq \Delta \leq 100$

we do not observe any variation in the ranking of pairs of nodes as can be seen in figure 1.16(a). In this case the range of Δ must be increased dramatically to obtain inversions in the ranking of pairs of nodes (see 1.16(b)).

To illustrate how well the estimate 1.39 performs, we use it for approximating the interlacement point for several pairs of corporates and compare them with the observed values in 1.2 for the weighted version of the US corporate network.

Table 1.2: Calculation of the crossing point ζ^* **calc** of ranking interlacement for several pairs of corporates in the US corporates network of 1999 as well as the observed values ζ^* **obs** at which such interlacements occur.

Plot	**Corporate 1**	**Corporate 2**	ζ^* **calc**	ζ^* **obs**
(a)	J.P. Morgan&Co Inc.	Bank of America Corp.	0.375	0.37
(b)	Pfizer Inc.	Ashland Inc.	0.441	0.41
(c)	Morgan Stanley & Co.	Bank One Corp.	0.176	0.17
(d)	AT&T Corp.	Airtouch Communications	0.273	0.27
(e)	Union Carbide Corp. New	AON Corp.	0.353	0.32
(f)	General Motors Corp.	Boeing Co.	0.214	0.14

A few more general considerations on the validity of the power series truncation heuristic can be made. The size of the interval containing any interlacing points is dictated to a large extent by how quickly the rankings based on the measures $\mathscr{C}_i(\zeta)$ (or $\tilde{\mathscr{C}}_i(\zeta)$) stabilize near the rankings obtained using eigenvector centrality. This, in turn, depends on the spectral gap $\lambda_1 - \lambda_2$: the larger the gap, the faster the eigenvector centrality rankings are approached for increasing values of ζ. Hence, in the case of relatively large gaps, we expect any interlacing values to occur for fairly small values of ζ. In this case, the heuristics based on polynomial approximations may be justified, since interlacing is likely to occur already for small values of ζ. As is well known, however, it is not easy to determine when the spectral gap is "sufficiently large". On the other hand, when the spectral gap is tiny, then the interval $[0, \bar{\zeta}]$ is going to be larger and therefore there is "more room" for the occurrence of interlacing. Unfortunately, in this case it is not clear that polynomial truncation will be effective in approximately locating the interlacing points. In this case, a possible solution is to expand the functions $\mathscr{C}_i(\zeta)$ not around the value $\zeta = 0$, but also around a few values $\zeta_0 > 0$. This strategy can also be used to find a possible second point of interlacing after having found a first such

point ζ^*. Expanding around ζ^* leads to

$$\Psi(\zeta^*+\eta) = \mathscr{C}_i(\zeta^*+\eta) - \mathscr{C}_j(\zeta^*+\eta) = \\ \frac{1}{2!}(w_{i,i}^{(2)} - w_{j,j}^{(2)})\eta^2 + \frac{1}{3!}(w_{i,i}^{(3)} - w_{j,j}^{(3)})\eta^3 + \cdots + \frac{1}{k!}(w_{i,i}^{(k)} - w_{j,j}^{(k)})\eta^k + O(\eta^{k+1}).$$

Dividing by η^2 and setting the result equal to zero leads to an algebraic equation of degree $k-2$ for η; the smallest positive root η^* of this equation, if there are any, leads to the approximation $\zeta^* + \eta^*$ for the next interlace point, and so forth.

Completely analogous considerations apply to the approximation of interlacing points when the ranking of nodes is done according to the risk-based total communicability measure $\mathscr{R}_i(\zeta)$. In this case the transcendental equation to be solved is given by

$$\chi(\zeta) = \mathscr{R}_i(\zeta) - \mathscr{R}_j(\zeta) = \sum_{k=1}^{n} e^{\zeta\lambda_k}\left(\psi_k^T\vec{1}\right)[\psi_{k,i} - \psi_{k,j}] = 0.$$

Let $w_i^{(k)} = (A^k\vec{1})_i$. Then, truncating the series expansion 1.34 and dividing by $\zeta > 0$ leads to the approximation

$$\frac{(w_i^{(k)} - w_j^{(k)})}{k!}\zeta^{k-1} + \cdots + \frac{(w_i^{(2)} - w_j^{(2)})}{2!}\zeta + (w_{i,i}^{(2)} - w_{j,j}^{(2)}) = 0 \tag{1.40}$$

for the equation whose smallest positive solution approximates the first interlacement value for the rankings of nodes i and j, assuming it exists; here again $k \geq k_0$ where now $k_0 \geq 2$ is the smallest integer value for which the sequence $\{w_i^{(k)} - w_i^{(k)}\}_k$ changes sign. The simplest possible case is when $k = k_0 = 2$, which occurs when $w_{i,i}^{(2)} - w_{j,j}^{(2)}$ and $w_i^{(2)} - w_j^{(2)} = (A^2\vec{1})_i - (A^2\vec{1})_j$ have different sign. In this case 1.40 reduces to the linear equation

$$\frac{(w_i^{(2)} - w_j^{(2)})}{2}\zeta + (w_{i,i}^{(2)} - w_{j,j}^{(2)}) = 0,$$

with the unique root

$$\zeta^* = 2\frac{w_{i,i}^{(2)} - w_{j,j}^{(2)}}{w_j^{(2)} - w_i^{(2)}} > 0.$$

1.8 Risk prediction and COVID-19

Starting on December 2019, a pandemic has been expanding worldwide from the city of Wuhan, Hubei province of China [90, 91]. This disease is produced by a new coronavirus named SARS-CoV-2 [92] and has affected in about three months more than 200 countries around the world. The major problem, at the time this thesis is written, is of

a health and medical nature, but as stated by Balwing and Weder di Mauro this coronavirus is "as contagious economically as it is medically" [93]. One of the most important characteristics of this pandemic in comparison with recent ones is that it is hitting very strongly the most important economies in the world: China, USA, Germany, Italy, Spain. There are some preliminaries studies about the macroeconomic impacts of this pandemic (see for instance [93]). However, it is important to apply mathematical and computational techniques to forecast, at regional, national and international level, the impact of this crisis on financial institutions, corporations and small companies. All of them are highly interconnected in a globally dependent economy, forming series of complex networks. In this new scenario the current work represents an opportunity for modelers to advance predictions on the potential risks to which different institutions are submitted to in the current situation. This modeling scenario consists of the networks of interactions between the institutions under analysis assuming a high infectability in the network. Using the transmissibility and circulability measures defined here, the modeler can understand how at risk of transmitting the crisis to others or, respectively, of staying in a cycle of repeated economic difficulties, a company is. At the same time, the current work allows to model how different palliative measures taken by regional or global financial institutions in the European Union, USA or China can impact these companies. In this case, the modeler should drop the infectivity of the system and analyze how the ranking of risk for the different companies changes to gain insights about their potential recovery or bankruptcy.

1.9 Conclusions

In general, node centrality in networks are of either of two types: (i) node centrality in networks of time-invariant topology [1], or (ii) node centrality in networks of time-dependent topology (aka Holme2012) [94]. In this work we have developed a new concept of node centrality, depending on both the topology of the network and the external conditions to which the network as a whole is submitted. In particular, we have focused on global risk as the external factor by which an economic and financial network is affected. We started by considering the "Susceptible-Infected" model and its connection to the communicability functions of nodes and edges in a network. Then, we developed a few centrality measures which depend not only on the local and global topological environment of a node but also on the level of infectivity stressing the system as a whole. In this way we have been able to make predictions in financial and economic systems about the changes in the risk-dependent centralities of nodes

as a function of the global level of infectivity in the system. We observe that without altering the topology of the network, i.e., without varying any connection between the nodes, the ranking of the nodes, according to these new centrality measures, changes significantly as the infectivity rate changes. In the real-world networks studied here we have been able to associate those changes in the risk-dependent centrality of nodes with events of the real financial and economic worlds in which these networks are embedded. In closing, we provide here both theoretical, computational and empirical evidences that the node centrality is not a static function even when the topology of the system is not varying at all. This new paradigm is expected to play a fundamental role in assessing the robustness of financial and economic systems to the variation of the external conditions which they are submitted to.

Chapter 2

Community structure in the World Trade Network based on communicability distances

2.1 Introduction

International trade is based on a set of complex relationships between different countries. Both connections between countries and bilateral trade flows can be modelled as a dense network of interrelated and interconnected agents. A long-standing problem in this field is the detection of communities, namely subset of nodes among which the interactions are stronger than average. Indeed, the community structure of a network reveals how it is internally organized, highlighting the presence of special relationships between nodes, that might not be revealed by direct empirical analyses.
In this framework, a specific role is assumed by the distance between nodes. Indeed, the neighbours of a given node are immediately connected to such a node and they can affect its status most directly. Nonetheless, more distant nodes can influence this node while passing through intermediary ones. In the economic field, a network perspective is actually based on the idea that indirect trade relationships may be important (see, e.g., [8]). For instance, the authors in [95] explain the impact of shocks on a given country by indirect trade links. Based on a global VaR approach, [96] shows that countries that do not trade (very much) with the U.S. are largely influenced by its dominance over other trade partners linked with the U.S. via indirect spillovers. In [97], the bilateral trade is assumed not independent of the production, consumption, and trading decisions made by firms and consumers in third countries. A measure of

the distance between nodes that also considers indirect connections is therefore crucial to catch deep interconnections between nodes. In this work, we will focus on two measures of distance or metrics on the network: the Estrada communicability distance [9] and the vibrational communicability distance [10]. They both go beyond the limits of the immediate interaction between neighbours and they look simultaneously, albeit differently, at all the possible channels of interactions between nodes. The nearest two nodes are in each metric, the stronger is their interaction or, in other words, the higher is the level of communicability between them.

With this chapter we contribute to the literature by proposing a specific methodology that exploits such metrics to inspect the mesoscale structure of the network, in search for strongly interacting clusters of nodes. Indeed, our purpose is twofold. We reveal hidden relationships between nodes due to non-immediate connections and long-range interactions and we show how this approach turns out to be particularly suitable when applied to a dense network like the World Trade Network (WTN). More specifically, we exploit communicability and vibrational communicability metrics to group nodes whose mutual distances are below a given threshold, i.e. whose interactions are stronger than a given value. Then we identify the optimal partition according to a maximum quality function criterion. It is well-known that classical modularity is a way to measure if a specific mesoscopic description of the network in terms of communities is more or less accurate. But, unlike the Girvan-Newman approach [11], we will refer to the partition quality index proposed in [12] for general metric spaces. In this way, we can exploit the additional information contained in the metric structure of the network. Among all the different partitions we get at different thresholds, we select the one providing the maximum quality index, according to the criterion described in [12]. Our proposal is very efficient from a computational viewpoint. Indeed, given the specific distance matrix, the optimal solution can be easily evaluated varying the threshold. We cluster nodes going beyond the interactions between neighbours and considering all possible channels of interaction between them. We allow for a degree of flexibility by introducing a threshold. Varying the threshold, it is possible to depart from the optimal solution so that only the strongest (or the weakest) channels of communications emerge.

The chapter is organised as follows. After a short review of the literature in Section 2.2, main preliminaries and the definitions of the communicability functions are revised in Section 2.3. These functions lead to two important metrics on networks, which are described in Section 2.4. Section 2.5 contains the description of the proposed

methodology, which is also tested on a suitable toy-model. In Section 2.6, we apply our methodology to the World Trade Network. In particular, main characteristics of the network are described in Section 2.6.1. The steps of the methodology are summarized in Section 2.6.2. We report in Section 2.6.3 main results based on communicability and resistance distance, respectively. We show how the proposed methodology is able in capturing key economic clusters as well as in providing additional insights into intracluster and intercluster characteristics and of countries' relevance both in the community and in the whole network. Conclusions follow. Technical details are left in Appendices A and B.

2.2 Literature Review

Community detection is an important topic in the analysis of the topological structure of complex systems. Its importance has grown over time in light of the remarkable progress in the description of large networks, together with the development of new powerful data analysis tools [98]. These advances have made it possible to extend the field of applicability of the theory not only to networks of enormous dimensions but also to weighted networks and direct networks [99, 100, 101, 102]. Various methods and algorithms to detect communities on networks have been studied. Some methods are algorithm-based, such as methods based on hierarchical clustering or edge removal [14]. Other methods are based on the optimization of specific criteria over all possible network partitions. In this context, it is well known the optimization of a modularity function according to Newman's definition [11]. An exhaustive review about methods and algorithms can be found in [103] and [104]. Some authors proposed to detect communities by means of a quality measure called surprise [105, 106]. Inspired by this literature, recently the authors in [107] deal with detection of general mesoscale structures, such as core-periphery structures.

More recently the role of non-local interactions between nodes has been highlighted, that is interactions that do not exclusively involve the immediate neighbours of a given node. In particular, results connected to the idea of communicability introduced by Estrada in 2004 have proved to be extremely effective [9, 47, 70, 108]. All the more so by allowing a metric different from the shortest path metric to be introduced on the network. The purpose of this new metric is precisely to take into consideration long-range interactions between institutions. Some important similarities can be found between this new metric and the resistance distance, a well-known metric in network theory derived from the study of electric circuits [10, 60, 109], and its interpretation in

terms of vibrational communicability [64, 65, 110, 111].

An area in which these concepts allow us to gain a deep insight into the hidden structures of the network is properly the WTN. The topology of the world trade web has been extensively analysed over time [112, 113, 114, 115, 116, 117]. The behaviour of international trade flows, the impact of globalization on the international exchanges, the presence of a core-periphery structure or the evolution of the community centres of trade, are just some of the issues addressed by the recent developments [118, 119, 120, 121, 122]. Many works have dealt with the network from a multi layers perspective [123, 124] or aim to emphasize financial implications of the world trade or contagion processes on the network [125, 126, 127, 128, 129, 130, 131, 132, 133, 134].

The impact of topology and metric properties on the stability and resilience of an economic or financial system has been widely studied in order to describe the large-scale pattern of dynamical processes inside the network [135, 136, 137]. These processes determine the subsequent diversification of the export of a country, which can be compared with descriptive empirical indices of its potential growth, such as the one introduced in a very fruitful way in [138].

2.3 Communicability in complex networks

The idea of communicability on a network is based on the ways in which a pair of nodes can communicate, namely through walks connecting them. In the literature, two different definitions of communicability have been introduced: the Estrada Communicability and the Vibrational Communicability [47, 65]. We recall them in this section.

2.3.1 Preliminary definitions

First of all, we briefly remind some preliminary definitions. A network is formally represented by a graph $\mathscr{G} = (V, E)$ where V and E are the sets of n nodes and m edges, respectively. Two nodes i and j are adjacent if there is an edge $(i, j) \in E$ connecting them. The network is undirected if (j, i) is an element of E whenever (i, j) is such. A $i - j$-path is a sequence of distinct vertices and edges between i and j. The shortest path, or geodesic, between i and j is a path with the minimum number of edges. The length of a geodesic is called geodesic distance or shortest path distance $d(i, j) = d_{ij}$. A graph $\mathscr{G}$ is connected if, $\forall i, j \in V$, a $i - j$-path connecting them exists.

Adjacency relationships are represented by a binary symmetric matrix $\mathbf{A}$ (adjacency matrix). Graphs considered here will be always connected and without self-loops; in

this case $a_{ii} = 0\ \forall i = 1, ..., n$. We denote with $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ the eigenvalues of $\mathbf{A}$, and $\varphi_i, i = 1, ..., n$ the corresponding eigenvectors.
The degree k_i of a node i is the number of edges incident on it. The diagonal matrix whose diagonal entries are k_i is $\mathbf{K}$. The Laplacian matrix is $\mathbf{L} = \mathbf{K} - \mathbf{A}$. $\mathbf{L}$ is a positive semidefinite symmetric matrix. We denote the eigenvalues of $\mathbf{L}$ by $\mu_1 \geq \mu_2 \geq \cdots > \mu_n = 0$ and $\psi_i, i = 1, ..., n$ the corresponding eigenvectors.

A graph $\mathscr{G}$ is weighted when a positive real number $w_{ij} > 0$ is associated with the edge (i, j). We define the strength s_i as the sum of the weights of the edges adjacent to i. The definition of geodesic path still holds, and it is a weighted path with the minimum sum of edge weights. In this case, the adjacency matrix is a non-negative symmetric matrix $\mathbf{W}$. When $w_{ij} = 1$ if $(i, j) \in E$, then the graph is unweighted. Thus, the unweighted case can be viewed as a particular weighted one.

2.3.2 Estrada Communicability

The *Estrada communicability* [47] between two nodes i and j is defined as:

$$G_{ij} = \sum_{k=0}^{+\infty} \frac{1}{k!} [\mathbf{A}^k]_{ij} = \left[e^{\mathbf{A}} \right]_{ij}. \tag{2.1}$$

As the ij-entry of the k-power of the adjacency matrix $\mathbf{A}$ counts the number of walks of length k starting at i and ending at j, G_{ij} accounts for all channels of communication between two nodes, giving more weight to the shortest routes connecting them. It can also be interpreted as a measure of the likelihood that a particle starting at i ends up at j after wandering randomly on the complex network. The communicability matrix is denoted by $\mathbf{G}$.

By definition, it follows that $G_{ij} > 0$. Moreover, G_{ij} can be conveniently expressed using the spectral decomposition of $\mathbf{A}$ as follows [47]:

$$G_{ij} = \sum_{k=1}^{n} \varphi_k(i) \varphi_k(j) e^{\lambda_k},$$

where $\varphi_k(i)$ is the i-component of the k-th eigenvector associated with λ_k.

It is worth noting that since G_{ii} characterizes the importance of a node according to its participation in all closed walks starting and ending at it, we recover the so-called *subgraph centrality* (see [70]).

In the case of a weighted network the communicability function is defined as

$$G_{ij} = \sum_{k=0}^{+\infty} \frac{1}{k!} [(\mathbf{S}^{-\frac{1}{2}} \mathbf{W} \mathbf{S}^{-\frac{1}{2}})^k]_{ij} = \left[e^{(\mathbf{S}^{-\frac{1}{2}} \mathbf{W} \mathbf{S}^{-\frac{1}{2}})} \right]_{ij} \tag{2.2}$$

where $\mathbf{S}$ is the diagonal matrix whose diagonal entries are the strengths of the nodes. We will call this quantity weighted communicability.

2.3.3 Vibrational Communicability

Vibrational communicability represents an alternative definition of communicability, different from Estrada communicability, and which can be introduced through the following model. Let us suppose that nodes of the network are objects of negligible identical mass connected by springs in a plane grid. Nodes can oscillate in the direction perpendicular to the plane and the displacement of the node i from its rest position is z_i. The elastic force applied to node i is given by $F_i = \mathcal{K} \sum_j A_{ij}(z_i - z_j)$, where $\mathcal{K}$ is the common elastic constant of each spring. An elastic potential energy can be assigned to each perturbed spring and the potential energy of all the springs connected with node i is given by $U_i = \frac{1}{2} \mathcal{K} \sum_j A_{ij}(z_i - z_j)^2$.
The overall potential energy of the network is therefore

$$U = \frac{1}{4} \mathcal{K} \sum_{i,j} A_{ij}(z_i - z_j)^2 = \frac{1}{2} \mathcal{K} \sum_{i,j} z_i L_{ij} z_j \tag{2.3}$$

where L_{ij} is the ij-entry of $\mathbf{L}$.

The reciprocal influence of two nodes i and j in their positions z_i and z_j is computed by means of the Green's function, according to the classical Boltzmann's distribution [10, 65]. This mutual influence can be interpreted as the correlation function between the displacements z of two nodes in the network:

$$G^v_{ij}(\beta) = \langle z_i z_j \rangle = \frac{1}{\mathcal{Z}} \int z_i z_j e^{-\beta U} d\mathbf{z}$$

where β is a constant and $\mathcal{Z} = \int e^{-\beta U} d\mathbf{z}$ is the partition function. Using the non-zero eigenvalues of $\mathbf{L}$, $\mathcal{Z}$ can be expressed as

$$\mathcal{Z} = \int e^{-\frac{1}{2}\beta \mathcal{K} \sum_{ij} z_i L_{ij} z_j} \prod_k dz_k = \prod_{k=1}^{n-1} \sqrt{\frac{2\pi}{\beta \mathcal{K} \mu_k}} \tag{2.4}$$

so that the correlation function can be rewritten in the final form

$$G_{ij}^{v}(\beta) = \sum_{k=1}^{n-1} \frac{\psi_k(i)\psi_k(j)}{\beta \mathcal{K} \mu_k} \tag{2.5}$$

where ψ_k is the eigenvector associated with μ_k. Introducing the Moore-Penrose pseudo-inverse of the Laplacian $\mathbf{L}^+$ [110, 139], the vibrational communicability between nodes i and j is defined as

$$G_{ij}^{v}(\beta) = \frac{1}{\beta \mathcal{K}} L_{ij}^{+} \tag{2.6}$$

The vibrational communicability matrix is denoted by $\mathbf{G}^v$. In the remainder of the chapter we will assume $\beta = 1$ and $\mathcal{K} = 1$, so that $G_{ij}^{v} = L_{ij}^{+}$.
The detailed computations for previous formulas are reported in Appendix B.

2.4 Metrics on networks

Metric properties play an important role in the study of the structure and dynamics of networks. The best known metric is the so-called shortest path distance. In the literature other metrics have been defined, each one stressing different features of the network. We remind the definitions of communicability distance and resistance distance, in view of their following application to the WTN.

2.4.1 Communicability Distance

The communicability distance ξ_{ij} is defined as (see [108]):

$$\xi_{ij} = G_{ii} - 2G_{ij} + G_{jj}. \tag{2.7}$$

As already observed, G_{ii} is the subgraph centrality of i and it measures the amount of information that starts from and returns to node i after having wandered through the network. On the other hand, G_{ij} measures the amount of information transmitted from i to j. Notice that the word *information* is meant in its broadest sense. Therefore, information flow can be any kind of flow along edges: money, current, traffic and so on. Thus, the quantity ξ_{ij} accounts for the difference in the amount of information that returns to the nodes i and j and the amount of information exchanged between them.

The greater is G_{ij}, the larger the information exchanged and the nearer are the nodes; the greater are G_{ii} or G_{jj}, the larger the information that comes back to the nodes and the farther are the nodes. In a matrix form, ξ_{ij} can be expressed as follows:

$$\mathbf{\Xi} = \mathbf{g}\mathbf{u}^T - 2\mathbf{G} + \mathbf{u}\mathbf{g}^T$$

where $\mathbf{g} = [G_{11}, \ldots, G_{nn}]^T$ is the vector of subgraph centralities and $\mathbf{u}$ the all 1's $n-$vector. Since ξ_{ij} is a metric, then $G_{ii} + G_{jj} \geq 2G_{ij}$, i.e., no matter what the structure of the network is, the amount of information absorbed by a pair of nodes is always larger than the amount of information transmitted between them.

2.4.2 Resistance Distance

The vibrational communicability distance between i and j is defined as (see [10, 64]):

$$\omega_{ij} = G^v_{ii} - 2G^v_{ij} + G^v_{jj}. \tag{2.8}$$

Formula 2.8 can be written in a more suitable way. Indeed, recalling that $G^v_{ij} = L^+_{ij}$, we have:

$$\begin{aligned}
\omega_{ij} &= L^+_{ii} - 2L^+_{ij} + L^+_{jj} \\
&= (\mathbf{e}_i - \mathbf{e}_j)^T \mathbf{L}^+ (\mathbf{e}_i - \mathbf{e}_j) \\
&= (\mathbf{e}_i - \mathbf{e}_j)^T \left[\left(\mathbf{L} + \frac{1}{n}\mathbf{J} \right)^{-1} - \frac{1}{n}\mathbf{J} \right] (\mathbf{e}_i - \mathbf{e}_j) \\
&= (\mathbf{e}_i - \mathbf{e}_j)^T \left(\mathbf{L} + \frac{1}{n}\mathbf{J} \right)^{-1} (\mathbf{e}_i - \mathbf{e}_j)
\end{aligned} \tag{2.9}$$

where $\mathbf{e}_k$, $k = 1, \ldots, n$, is the standard basis in $\mathbb{R}^n$ and $\mathbf{J} = \mathbf{u}\mathbf{u}^T$ is the matrix whose entries are all 1. Note that in the previous chain of equalities we made use of the following expression of the pseudo-inverse $\mathbf{L}^+ = \left(\mathbf{L} + \frac{1}{n}\mathbf{J}\right)^{-1} - \frac{1}{n}\mathbf{J}$, proved in [139].

Equation 2.9 offers an interesting interpretation of the resistance distance. We synthesize here the main idea, referring to Appendix C for a more detailed discussion. Let $\mathbf{v} = [v_1, v_2, \ldots, v_n]^T$ be a vector representing attributes of the nodes – for instance, the Gross Domestic Product (GDP) of a country or the assets of a financial institution – and suppose that there are currents or flows (of money, for instance) along the edges of the network. The operator $\left(\mathbf{L} + \frac{1}{n}\mathbf{J}\right)^{-1}$ allows to obtain the state vector that gives rise to a given set of flows. In formula 2.9, the vector $(\mathbf{e}_i - \mathbf{e}_j)$ refers to a global flow equal to $+1$ from node i, a flow equal to -1 into node j and a flow equal to 0 for the other ones. When we apply $\left(\mathbf{L} + \frac{1}{n}\mathbf{J}\right)^{-1}$ to $(\mathbf{e}_i - \mathbf{e}_j)$, we get the state vector $\mathbf{v} = [v_1, v_2, \ldots, v_n]^T$ of attributes on nodes that gives rise to these flows. Finally, the

left inner product with $(\mathbf{e}_i - \mathbf{e}_j)$ in formula 2.9 gives $v_i - v_j$, namely, the difference between attributes of nodes i and j. This gradient produces exactly the flow $+1$ from node i and -1 to node j. If $v_i - v_j$ is big, we need a big difference in order to produce such a unit flow and so we have a big resistance between nodes i and j. If $v_i - v_j$ is small, it is enough a low difference in order to produce such a unit flow and so we have a low resistance between nodes i and j. If ω_{ij} is big we have a high resistance distance between i and j. Therefore, these two nodes do not communicate easily. Vice versa a low value of ω_{ij} means a high level of communication between the nodes. ω_{ij} is called *effective resistance* between nodes i and j and $\mathbf{\Omega} = [\omega_{ij}]$ is the resistance matrix.

In literature, it is known an important close form for $\mathbf{L}^+$ in terms of $\mathbf{\Omega}$:

$$\mathbf{L}^+ = \frac{1}{2}\left[\frac{1}{n}(\mathbf{\Omega J} + \mathbf{J\Omega}) - \frac{1}{n^2}\mathbf{J\Omega J} - \mathbf{\Omega}\right]$$

which allows us to rewrite the diagonal elements of the matrix $\mathbf{L}^+$ in a useful form[1]

$$L^+_{ii} = \frac{1}{n}\sum_j \omega_{ij} - \frac{R}{n^2}$$

where

$$R = \frac{1}{2}\sum_{i,j}\omega_{ij} = \sum_{i=1}^{n}\sum_{j=i+1}^{n}\omega_{ij} = \frac{1}{2}\mathbf{u}^T\mathbf{\Omega u} = n\,\mathrm{tr}\,\mathbf{L}^+ = n\sum_{k=1}^{n-1}\frac{1}{\mu_k}$$

is the *effective graph resistance* (or *Kirchhoff index*) of the network, i.e. the sum of the resistances between all possible pairs of nodes in the graph (see, e.g., [109]). R reflects the overall transport capability of the network: the lower R, the better the network conducts flows. In particular, it has been shown that this index is able to catch the average vulnerability of a connection between a pair of nodes and, therefore, it is a suitable tool for assessing the ability of a network to well react when it is subject to failure and/or attack (see [140, 141, 142]).

Effective resistances allow to give a specific definition of the centrality of a node in the network. Indeed, the *best spreader* (or best connected) node in the network is the node $i^\star$ that minimizes the quantity $\sum_{j=1}^{n}\omega_{i^\star j} = (\mathbf{\Omega u})_{i^\star}$, i.e. the sum of all its resistance distances from any other node in the network. Since L^+_{ii} equals the difference between the average resistance between node i and all the other nodes in the network and the overall network mean resistance, then the best spreader node $i^\star$ is the one such

[1] $L^+_{ii} = \frac{1}{2n}(\mathbf{\Omega J})_{ii} + \frac{1}{2n}(\mathbf{J\Omega})_{ii} - \frac{1}{2n^2}(\mathbf{J\Omega J})_{ii} - \frac{1}{2}(\mathbf{\Omega})_{ii} = \frac{1}{2n}\sum_j \omega_{ij}J_{ji} + \frac{1}{2n}\sum_j J_{ij}\omega_{ji} - \frac{1}{2n^2}\sum_{jk} J_{ij}\omega_{jk}J_{ki} - 0 = \frac{1}{2n}\sum_j \omega_{ij} + \frac{1}{2n}\sum_j \omega_{ij} - \frac{1}{2n^2}\sum_{jk}\omega_{jk} = \frac{1}{n}\sum_j \omega_{ij} - \frac{R}{n^2}$

that $L^+_{i^\star i^\star} \leq L^+_{jj}$ for any $j \neq i^\star$. Node $i^\star$ can be regarded as the best diffuser of a flow to the rest of the network, and, to some extent, it is the most influential with respect to a diffusion process inside the network, since it guarantees the highest flow toward other nodes (see [64]). Best diffuser means that most of the information coming out from this node is absorbed by other nodes. If L^+_{ii} is big, then most of this information comes back to node i and doesn't reach other nodes. The reciprocal of L^+_{ii} can then be regarded as a centrality measure of a node and it is called *vibrational centrality*.

2.5 Community detection based on communicability metrics

2.5.1 The model

As discussed in the previous section, $\xi_{ij} = G_{ii} - 2G_{ij} + G_{jj}$ and $\omega_{ij} = G^v_{ii} - 2G^v_{ij} + G^v_{jj}$ represent the two metrics induced on the network by the Estrada communicability and the vibrational communicability, respectively.

In an economical context, referring to the international trade network, they measure how well two countries, or companies, communicate in terms of commercial and trade exchanges. For instance, the attributes on nodes may be identified with the GDP and the currents along nodes with the total trade or money flow between two countries. Information on the network may be replaced by money flow. Therefore the quantity ξ_{ij} of equation 2.7 accounts for the difference in the amount of money flow that returns to the nodes i and j and the amount of money flow exchanged between them. The bigger is G_{ij}, i.e. the money flow exchanged, the nearer are the nodes; the bigger are G_{ii} or G_{jj}, i.e. the amount of money flow that comes back to the each node, the farther they are. A similar interpretation holds for ω_{ij}. In a trade network ω_{ij} accounts for the difference between the mean resistance to export a given money flow from each country and the correlation between them. The bigger is G^v_{ij}, the more interconnected they are and the nearer they are in the resistance metric; the bigger are G^v_{ii} and G^v_{jj}, the more isolated they are in the network and between them and the farther they are.

In light of these observations, we formulate our proposal[1], considering as members of the same cluster nodes whose mutual distance is below a given threshold ξ_0. Specifically, we construct a new community graph where the elements of the adjacency matrix $\mathbf{M} = [m_{ij}]$ are given by:

[1] In what follows, we will refer to the communicability distance ξ, but similar arguments may be repeated identically for the resistance distance ω

$$m_{ij} = \begin{cases} 1 & \text{if } \xi_{ij} \leq \xi_0 \\ 0 & \text{otherwise} \end{cases}$$

with ξ_0 threshold distance such that $\xi_0 \in [\xi_{\min}, \xi_{\max}]$, being $\xi_{\min}$ and $\xi_{\max}$ the minimum and the maximum distances between couples of nodes, respectively. In this way, clustered groups of nodes that strongly communicate emerge, in dependence of the threshold. If ξ_0 is high enough, all nodes in the network are at a mutual distance lower than the threshold and the whole network behaves like a unique community. As ξ_0 decreases, there will be nodes too far, such that to be considered disconnected and then members of different clusters, entailing the emergence of islands of connected nodes. Hence, the number of communities depends on the threshold, precisely it increases as ξ_0 decreases.

It is important to observe that, with the proposed methodology, we do not choose any *a priori* optimal number of communities. Our approach is more in line with the classic Girvan-Newman approach [11].

The optimal partition is determined according to an optimization problem whose objective function is based on the idea of cohesion between nodes. Specifically, since we deal with distances, following the approach for clustering in metric spaces proposed by [12], we provide a cohesion measure γ_{ij} between two nodes i and j, as follows:

$$\gamma_{ij} = \left(\bar{\xi}_j - \bar{\xi}\right) - \left(\xi_{ij} - \bar{\xi}_i\right)$$

where $\bar{\xi}_i = \frac{1}{n-1} \sum_{k \neq i} \xi_{ik}$ is the average distance between i and nodes other than i and $\bar{\xi}$ is the average distance over the whole network. Thus, $\xi_{ij} - \bar{\xi}_i$ represents the *relative distance* between nodes i and j and $\bar{\xi}_j - \bar{\xi}$ represents the *relative distance* from a random node to the node j.

Two nodes i and j are said to be cohesive (or incohesive) if $\gamma_{ij} \geq 0$ ($\gamma_{ij} \leq 0$). Notice that $\gamma_{ij} \geq 0$ yields $\xi_{ij} + \bar{\xi} \leq \bar{\xi}_i + \bar{\xi}_j$, i.e., intuitively, two nodes are cohesive if they are close to each other and, on average, they are both far away from the other nodes. In other words, γ_{ij} can be interpreted as the gain (when positive) or the cost (when negative) related to the grouping of nodes i and j in the same cluster of a given partition.

We assume to maximize an objective function that represents the global cohesion function based on the mutual relative distances between every pairs of nodes. Therefore, we refer to a specific partition quality index defined as

$$Q = \sum_{i,j} \gamma_{ij} x_{ij} \tag{2.10}$$

where x_{ij} is a binary variable equal to 1 if two nodes are in the same cluster and 0 otherwise and γ_{ij} is the cohesion measure between nodes i and j. It is worth to notice that when the partition is made up of a unique community, equal to the entire network, $x_{ij} = 1\ \forall i, j$. In this case[1]

$$\begin{aligned} Q = \sum_{i,j} \gamma_{ij} &= \sum_{i,j} \bar{\xi}_j + \sum_{i,j} \bar{\xi}_i - \sum_{i,j} \bar{\xi} - \sum_{i,j} \xi_{ij} \\ &= n \sum_{j} \bar{\xi}_j + n \sum_{i} \bar{\xi}_i - n^2 \bar{\xi} - \sum_{i,j} \xi_{ij} \\ &= 2n^2 \bar{\xi} - n^2 \bar{\xi} - n(n-1)\bar{\xi} \\ &= n\bar{\xi}. \end{aligned}$$

On the other hand, when the partition consists of n isolated nodes, $x_{ij} = 0\ \forall i \neq j$ then

$$Q = \sum_{i} \gamma_{ii} = \sum_{i} (\bar{\xi}_i - \bar{\xi}) - (\xi_{ii} - \bar{\xi}_i) = 2 \sum_{i} \bar{\xi}_i - n\bar{\xi} = n\bar{\xi}.$$

Therefore, the partition quality index Q assigns to the two extreme cases the same value $n\bar{\xi}$. This property will be confirmed in the next illustrative example and applications to the World Trade Network.

2.5.2 An illustrative example

We start by testing our methodology on a simple example. Let us consider the weighted undirected network displayed in Figure 2.1. The network has 10 nodes and 32 edges. The thickness of links is proportional to weights. The network allows to easily identify two natural communities, which are highlighted by the two closed lines containing nodes 1 to 5 (on the left) and nodes 6 to 10 (on the right).

We compute the Estrada communicability matrix $\mathbf{G}$, then we get the communicability distance matrix $\mathbf{\Xi}$. The nearest nodes are 1 and 3 with a communicability distance equal to $\xi_{\min} = \xi_{13} = 1.18$ and farthest nodes are 3 and 6 with a communicability distance equal to $\xi_{\max} = \xi_{36} = 1.49$. Figure 2.2 summarizes the number of communities identified at different thresholds. The blue line represents the number of communities while the red line represents the quality index Q of the corresponding partition. When the threshold is greater than or equal to $\xi_0 = 1.38$ all nodes are connected and the network is partitioned in a single community, with quality index $Q = n\bar{\xi}$. As the threshold decreases below 1.38, the network begins to split into disconnected components. When

[1] Notice that $\bar{\xi} = \frac{1}{n}\sum_i \bar{\xi}_i = \frac{1}{n(n-1)}\sum_{i,j} \xi_{ij}$

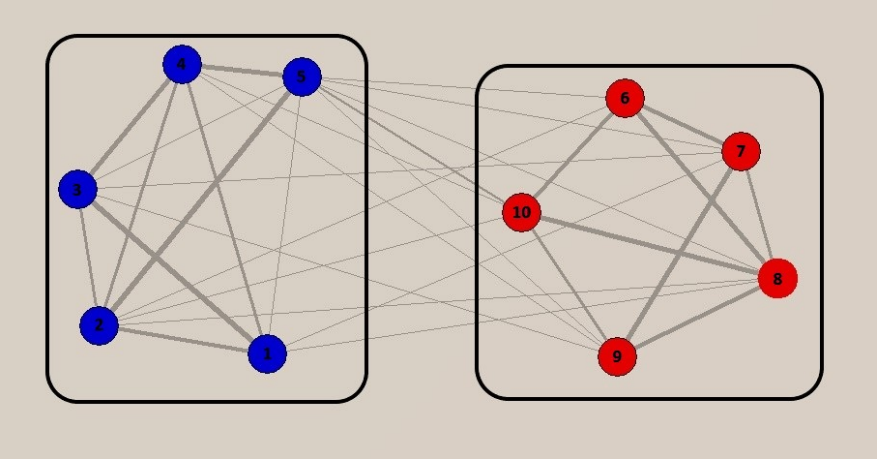


Figure 2.1: A weighted undirected network with 10 nodes and 32 edges. Edges weights have been randomly sampled with replacement from integers between 1 and 6. The thickness of edges is proportional to the weights. Nodes of two relevant communities are highlighted in blue and red.

the threshold becomes lower than the minimum distance, the network is partitioned into ten communities and each node belongs to a different community. The best partition according to the maximum quality index criterion splits the network into two clusters, which are easily identified with the two expected natural communities. The composition of the communities for alternative thresholds is reported in Figure 2.3. It is noticeable that, lowering the threshold, the procedure allows to disentangle tightest relationships. For instance, when $\xi_0 = 1.23$ only nodes connected by edges with highest weights are kept in the same community.

Similar results are derived by applying the procedure based on the vibrational communicability. The nearest nodes are 1 and 3 with a resistance distance equal to $\omega_{\min} = \omega_{13} = 1.22$ and farthest nodes are 3 and 8 with a resistance distance equal to $\omega_{\max} = \omega_{38} = 1.69$. Again if we move the threshold from the maximum distance to the minimum distance, we get an increasing number of communities from 1, the whole network, to 10, isolated nodes. The best partition according to the maximum quality index criterion splits the network into the two expected communities, as shown in Figure 2.4.

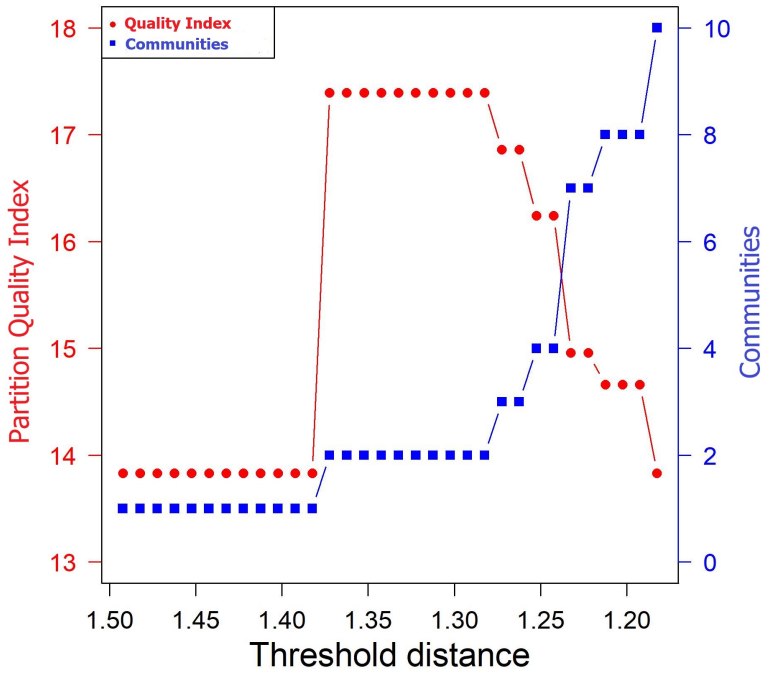
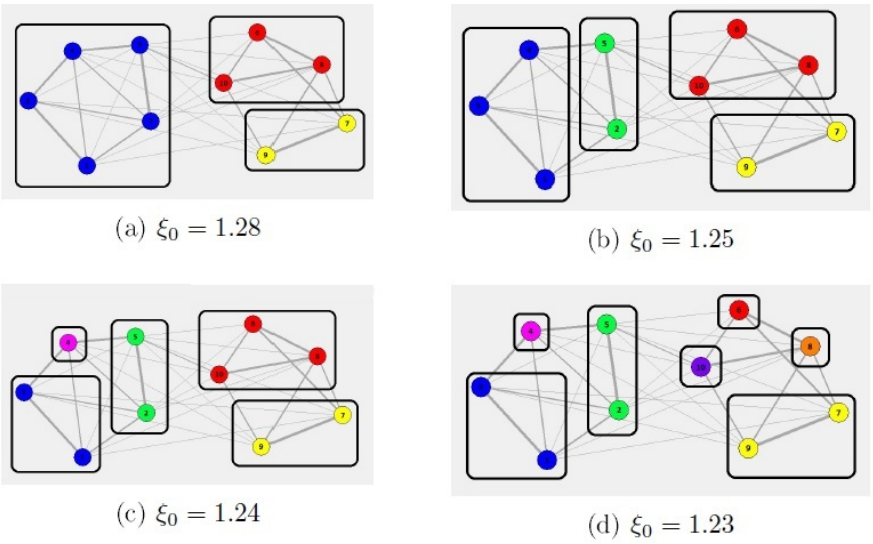


Figure 2.2: Quality index Q of the partition computed according to formula 2.10 and number of components (on the secondary scale) for different threshold values. The communicability distance has been used for the identification of the communities.

2.6 Application to the World Trade Network

In this Section, we apply the proposed model in order to detect relevant communities of countries in the WTN. As described before, the method aims at grouping strongly interacting countries by means of their mutual distances. Two alternative distance functions will be tested. On the one hand, we find clusters exploiting communicability distance. Therefore we detect how much two countries are close in the network considering all possible weighted walks connecting them. On the other hand, we select clusters by means of resistance distance. In this case countries are grouped together if they have a similar relevance in the network in terms of vibrational centralities as well as if they are correlated in terms of their expositions towards common countries.
We start with a general description of the dataset and the main characteristics of the



(a) $\xi_0 = 1.28$ (b) $\xi_0 = 1.25$

(c) $\xi_0 = 1.24$ (d) $\xi_0 = 1.23$

Figure 2.3: Community structure at different thresholds.

WTN. Then, we briefly summarize the primary steps of the methodology, providing a pseudo-code of the algorithm. Finally, we report the results in terms of community structure with the related discussion.

2.6.1 Dataset and main characteristics of the WTN

We refer to the World Trade Data, available on the Observatory of Economic Complexity database[1]. The database has been developed by the Research and Expertise Center on the World Economy at a high level of product disaggregation and it is based on original data provided by the United Nations Statistical Division (UN Comtrade). In particular, a harmonization procedure, that reconciles the declarations of exporters and importers, enables to extend considerably the number of countries for which trade data are available, as compared to the original dataset. In this analysis, we refer to the last version published in 2017, based on the Harmonized Commodity Description and Coding System, and that provides aggregated bilateral values of exports for each couple of origin and destination countries, expressed in billion dollars. We focus on the aggregated data of last available year, namely, 2016.
Hence, we construct a weighted network where each node is a country and weighted links represent the amount of product traded between couple of countries (see Figure 2.5). The mutually exchanged products between two countries are different in terms

[1]The Observatory of Economic Complexity (OEC) is the world's leading data visualization tool for international trade data. Data can be found at: https://atlas.media.mit.edu/en/

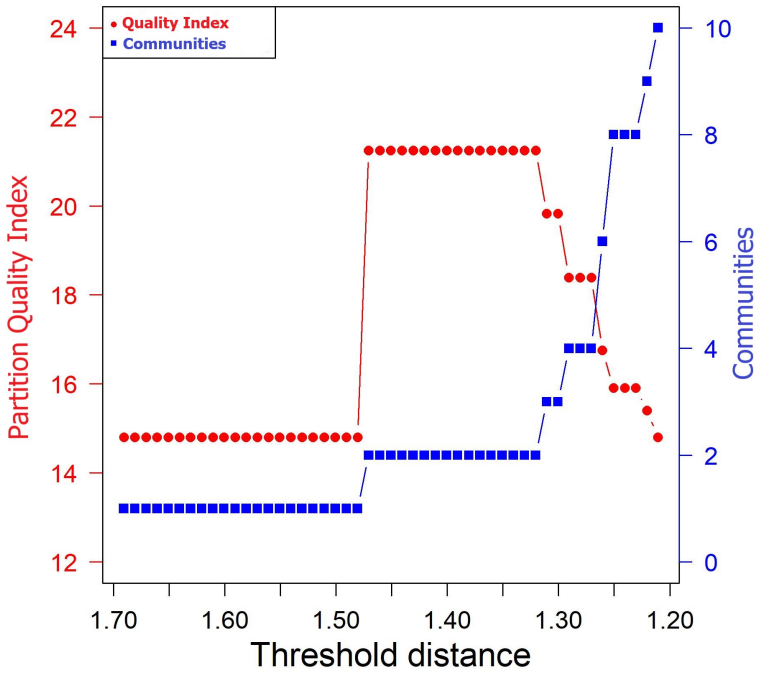


Figure 2.4: Quality index Q of the partition and number of components (on the secondary scale) for different thresholds. The resistance distance has been used for the identification of the communities.

of entity, so that they can be better represented by oriented links from a country to another one. However, we observed a strict relation between in and out strength distribution with a Spearman correlation coefficient equal to 0.956. Hence, countries are ranked in a very similar way in terms of in and out strength. Thus, we perform all the analysis assuming the network as undirected.

The undirected network is characterized by 221 nodes and 14933 links. The network is connected and its density is approximately 0.614: on average, each country has trades with more than a half of the entire network. However, the network is not regular and is far from being complete or, in other words, most countries do not trade with all the others, but they rather select their partners. Furthermore, main trade flows tend to be concentrated in a specific sub-group of countries and a small percentage of the

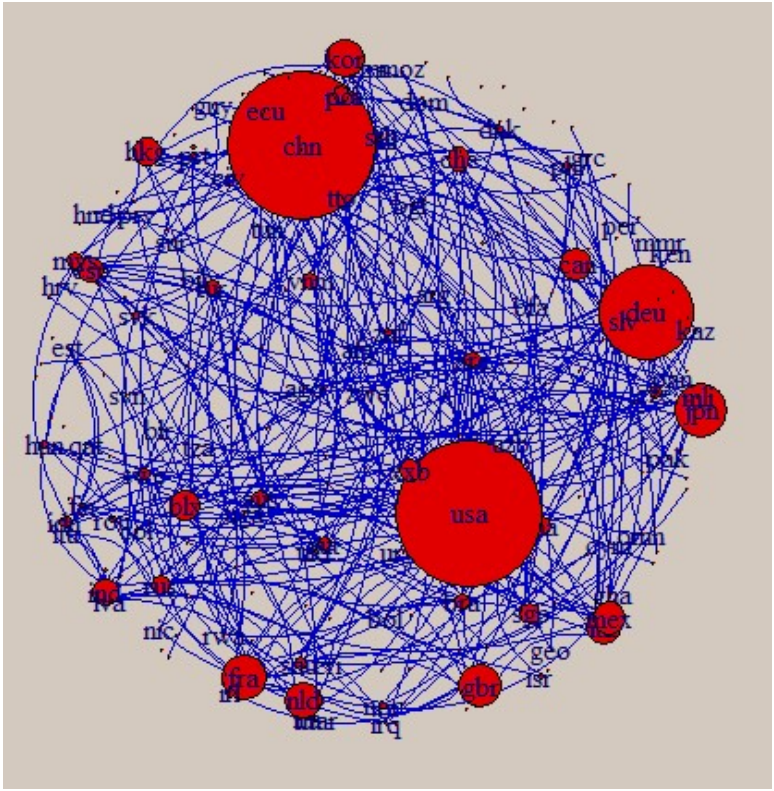


Figure 2.5: WTN based on 2016 data. Nodes are countries and links are product trades between pair of countries. The size of the node is proportional to its strength.

total number of flows accounts for a disproportionately large share of world trade. For instance, the top 10 countries export more than 50% of the total flow. The maximum weight corresponds to the channel between China and USA and its value amounts to 277 billion dollars. Minimum, non null, weights are involved in the trade between a number of very small countries, far from each others, and they are approximately around 1 thousand dollars.

Finally, we expect that several countries trade with their geographical neighbours so that we investigate the correlation between flows and geographical distance of countries. We computed the Spearman rank correlation between link weights (i.e. monetary flows between countries in the network) and the great circle distance between capital cities in kilometers. We obtained a rank correlation of -0.27, that confirms a little preference for trading with physical neighbours. However, as stressed before, our aim is to go beyond immediate neighbours by means of both communicability and resistance distances.

2.6.2 Summary of the methodology

In this section we summarize by means of a pseudo-code the main steps of the methodology we are proposing. The code has been written taking into account the communicability distance matrix $\mathbf{\Xi}$, but the same procedure can be easily applied by considering the resistance matrix $\mathbf{\Omega}$.

1. let $\mathscr{G}$ be the original directed weighted network with n nodes and weighted adjacency matrix $\mathbf{W}$;
2. build the undirected weighted network $\mathscr{G}_1$ with a symmetric adjacency matrix defined as $\mathbf{W}_1 = \frac{1}{2}(\mathbf{W} + \mathbf{W}^T)$;
3. build the undirected weighted network $\mathscr{G}_2$ with normalised weighted adjacency matrix $\mathbf{W}_2 = \mathbf{S}^{-1/2}\mathbf{W}_1\mathbf{S}^{-1/2}$, where $\mathbf{S}$ is the diagonal matrix of the strengths of the network $\mathscr{G}_1$;
4. construct the distance matrix $\mathbf{\Xi} = \mathbf{g}\mathbf{u}^T - 2\mathbf{G} + \mathbf{u}\mathbf{g}^T$ based on the communicability matrix $\mathbf{G}$;
5. define the threshold interval $[\xi_{\min}, \xi_{\max}]$, where $\xi_{\min}$ and $\xi_{\max}$ represent the minimum and the maximum communicability distances between couples of nodes, respectively and set $\xi_h = \xi_{\min}$, with $h = 0$;
6. define a $n \times n$ matrix $\mathbf{M}_h = [m_{ij}]$ such that
$$m_{ij} = \begin{cases} 1 & \text{if } \xi_{ij} \leq \xi_h \text{ and } i \neq j \\ 0 & \text{otherwise} \end{cases};$$
7. build the undirected unweighted network $\mathscr{G}_{3,h}$ from the binary adjacency matrix $\mathbf{M}_h$;
8. select the partition P_h given by the components of the network $\mathscr{G}_{3,h}$;
9. compute the quality index $Q = \sum_{i,j} \gamma_{ij} x_{ij}$ of the network $\mathscr{G}_2$ with respect to the partition P_h;
10. set the number of iterations r, compute $k = \frac{\xi_{\max} - \xi_{\min}}{r}$, set $\xi_h = \xi_{h-1} + k$ and $h = h + 1$ and repeat steps 6-9 while $\xi_h \leq \xi_{\max}$;
11. select the optimal partition $P_h^\star$ as the partition P_h that provides the maximum quality index Q.

We stress some key points of the presented methodology. We aim at clustering countries on the basis of a specific distance. The two distances we have chosen highlight relationships of a different nature between countries and the different community structure emerging will support this fact. Varying the threshold we can disentangle the role of very tight relationships between couples of countries. Of course, reducing the threshold distance a great number of isolated nodes may appear. They are typically very small countries whose trade volume is very low and whose commercial partners are few. They play a marginal role in the WTN and they do not affect in a significant way the structure of the network in terms of relevant communities. This is the reason why we will focus our attention on the main communities that are produced by our methodology.

2.6.3 Results

2.6.3.1 Results in terms of communicability metric

We initially applied the methodology described in Section 2.6.2 by using the communicability distance. The rationale for using the communicability metric on the WTN is the following. Two countries share a total volume of trade because they exchange a given set of products, of any kind. But they can be linked even if they don't exchange each other a given product, that is there is no direct flow of such product between them. A higher order exchange may occur between them. For instance, a country A exports some raw materials - let's say, iron - to a country B; country B produces mechanical parts from iron and exports them to country C. A and C communicate via a higher order walk and they depend on each other even if the two countries are not neighbours in the network. Indeed, communicability takes into account precisely all possible weighted walks between two nodes.

Therefore, we calculate the communicability matrix $\mathbf{G}$ on the normalised network $\mathscr{G}_2$ and the corresponding communicability distance matrix $\mathbf{\Xi}$. Using this metric, we find that the nearest countries are USA and Canada with a distance $\xi_{\min} = 1.242$ and the farthest countries are USA and Seychelles Islands with a distance $\xi_{\max} = 1.470$. Lowering the threshold distance value from maximum to minimum with a 0.001 step, we look at the corresponding partition in communities. In Figure 2.6, we plot the value of partition quality index Q (in red) and the number of communities (in blue), counting each isolated node as an independent one. Both values are expressed as functions of the threshold ξ_h. The maximum of Q is reached at a threshold distance $\xi_h = 1.392$. It corresponds to 106 communities, among which we have 87 isolated nodes. Hence, we

observe 19 significant communities other than isolated nodes.

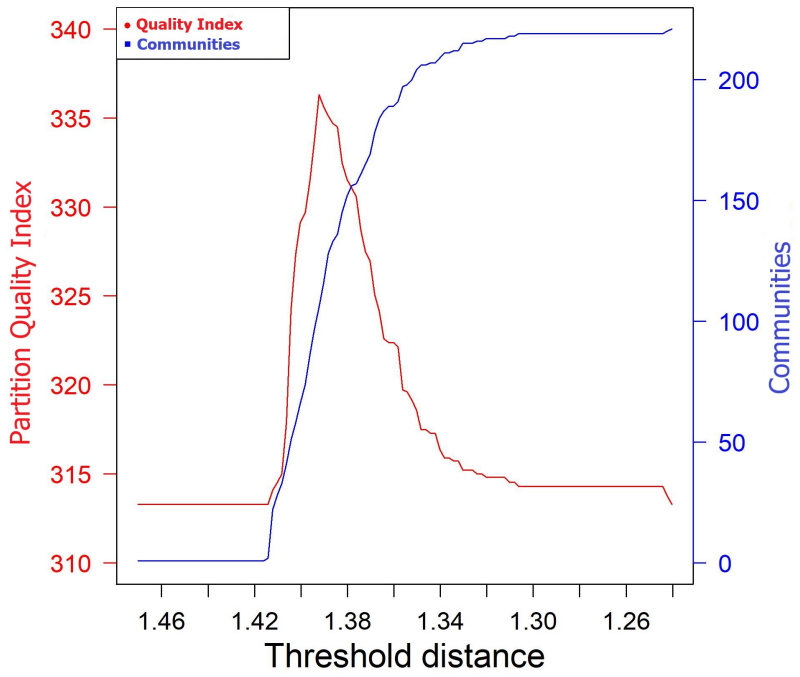


Figure 2.6: Partition quality index Q (red line) and number of communities (blue line) as functions of the threshold communicability distance ξ_h. Maximum Q is observed for $\xi_h = 1.392$.

We display in Figure 2.7 communities in the optimal partition and we list in Table 2.1 the countries belonging to the ten biggest communities in terms of numerousness.

Going deeper into the composition of the communities, the biggest one (see community 1 in blue) includes almost all continental European countries, with Great Britain and Ireland. This community acts on the screen of the global network as single player. It is worth pointing out also the presence of Morocco, confirming positive effects of bilateral trade agreements (see, e.g., [143]). We also notice the presence of South Asian countries that are economically linked together by the South Asian Association for Regional Cooperation. Presence of these countries in the community is also an effect of the bilateral foreign relations between the European Union (EU) and the Association

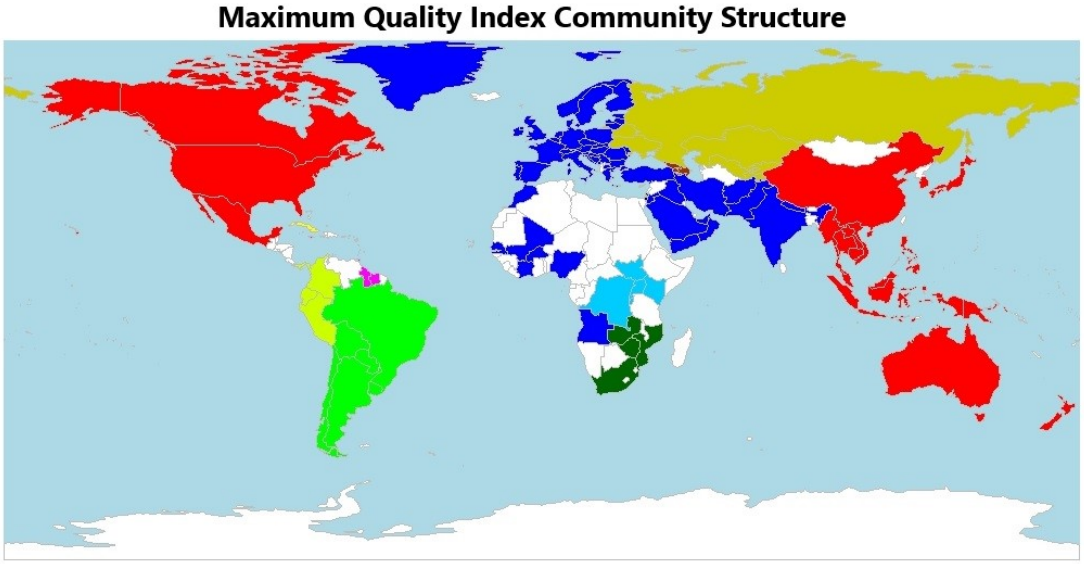


Figure 2.7: Optimal community structure based on communicability distance. Results have been derived by means of a maximum quality index criterion, with threshold $\xi_h = 1.39$ (Max Q Partition). Isolated nodes appear in white.

of Southeast Asian Nations (ASEAN). The partnership between the EU and ASEAN dates back to 1972 when the EU countries became ASEAN's first formal dialogue partner. Finally, to the same community belong African countries that are characterized by close economic and cultural ties to European countries, in particular to France (see, for instance, Ivory Coast, Burkina Faso, Angola, Senegal).
Opposed to this community, we see the second largest community (see, community 2 in red) which sees United States and China as main actors. This means that in Europe there are preferential channels of internal exchanges, whereas, outside Europe, most communication channels seem to be polarized around the exchange channel between China and the US and all their satellites countries. Moreover, we can recognize other well-identified and coherent communities.
Furthermore, it is interesting the decomposition of post-Soviet States. While Baltic and Eastern Europe States (except for Ukraine) have main partners in European countries, Central Asian countries have Russia as their leading trade and economic partner (see community 3). Although a positive trade balance and a priority of Russian government of an increasing participation in the economic relations of Asia-pacific region (see [144]), at moment, results show preferential channels with border countries. Transcaucasia is instead detected as a separate community (see community 10).

Except for Mexico, characterized by strong ties with United States, the Latin American and the Caribbean Economic System is decomposed into four relevant communities (see communities 4, 6, 7 and 8). In particular, it is noticeable community 4 developed on the basis of the South Common Market, namely the so-called MERCOSUR. Mercosur's purpose is to promote free trade and the fluid movement of goods, people, and currency in south America. Since its foundation, Mercosur's functions have been updated and amended many times; it currently confines itself to a customs union, in which there is free intra-zone trade and a common trade policy between member countries. In 2019, the Mercosur had generated a nominal gross domestic product (GDP) of around 4.6 trillion US dollars, reaching the fifth economy of the world.
Finally, significant blocks are also observed in central and south Africa (communities 5 and 9, respectively), polarized around Democratic Republic of the Congo and Republic of South Africa.

	Size	Members
Community 1	54	AFG AGO ARE AUT BFA BGR BHR BIH BLX CHE CIV CYP CZE DEU DNK ESP EST FIN FRA GBR GRC GRL HRV HUN IND IRL IRN IRQ ITA JOR LTU LVA MAR MDA MKD MLI MNE NGA NLD NOR NPL OMN PAK POL PRT ROU SAU SEN SRB SVK SVN SWE TUR YEM
Community 2	21	AUS CAN CHN HKG IDN JPN KHM KOR LAO MEX MHL MMR MYS NZL PHL PNG SGP THA USA VNM XXB
Community 3	7	BLR KAZ KGZ RUS TJK UKR UZB
Community 4	6	ARG BOL BRA CHL PRY URY
Community 5	6	BDI COD KEN RWA SSD UGA
Community 6	5	BES COL ECU PAN PER
Community 7	5	CRI GTM HND NIC SLV
Community 8	4	GUY JAM SUR TTO
Community 9	4	MOZ ZAF ZMB ZWE
Community 10	3	ARM AZE GEO

Table 2.1: Members of the top ten communities in terms of number of countries (for the names of the countries we refer to the current officially assigned ISO 3166-1 alpha-3 codes.)

If we reduce the threshold, we let very strong channels of communication between countries emerge. For instance, Figures 2.8 and 2.9 show the community structure lowering the threshold distance (equal to $\xi_h = 1.37$ and $\xi_h = 1.35$, respectively). Moving from 1.39 to 1.37 some loose connections are lost (see Figure 2.8). Scandinavia and the Nordic Region split up from community 1 creating a separate cluster together. The South East Asian and former Yugoslavia appear as separate communities characterized only by most relevant partnerships, Australia goes out from community 2, and the strong community in the South of Africa loses some country. Furthermore, in South America, only the relation between Brazil and Argentina survives. This result is in line with the fact that the strategic relationship between Argentina and Brazil is considered to be at the highest point in history: Brazil accounts indeed for Argentina's largest export and import market.
Reducing further the threshold to 1.35, only the most closely interrelated communities survive. The strongest community counts now, among its members, all North America, Mexico, China and Japan (in red in Figure 2.9). In Europe two communities are saved. On the one hand, the relation between Spain and Portugal is preserved. On the other hand, a community emerged in central Europe around the channel between France and Germany. Finally community 3 in Table 2.1, including Russia and Central Asian countries, resists also when the threshold is lowered.

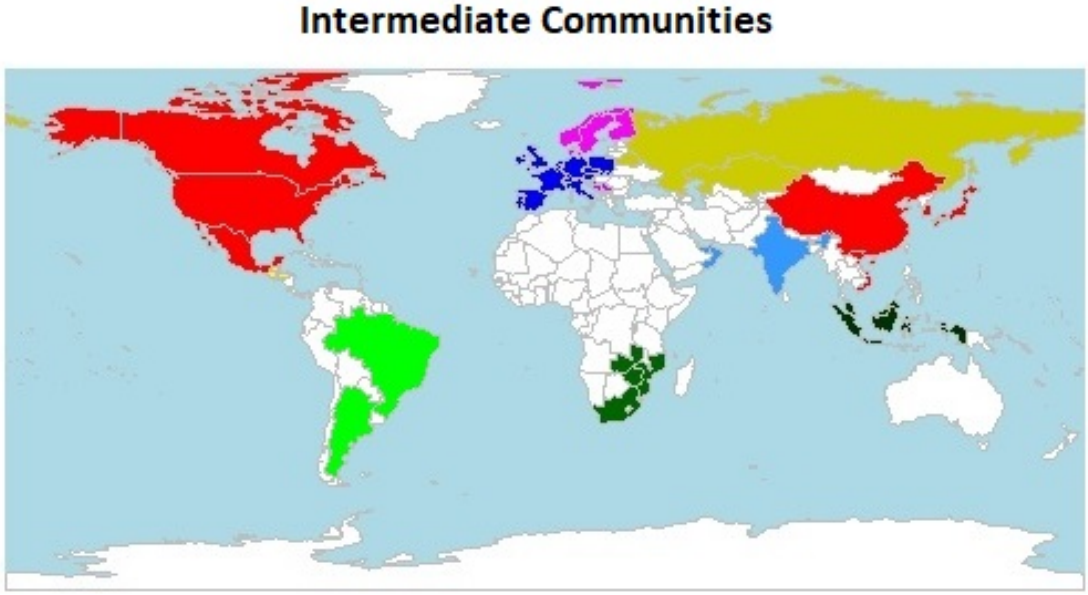


Figure 2.8: Intermediate Connected Community Structure - $\xi_h = 1.37$

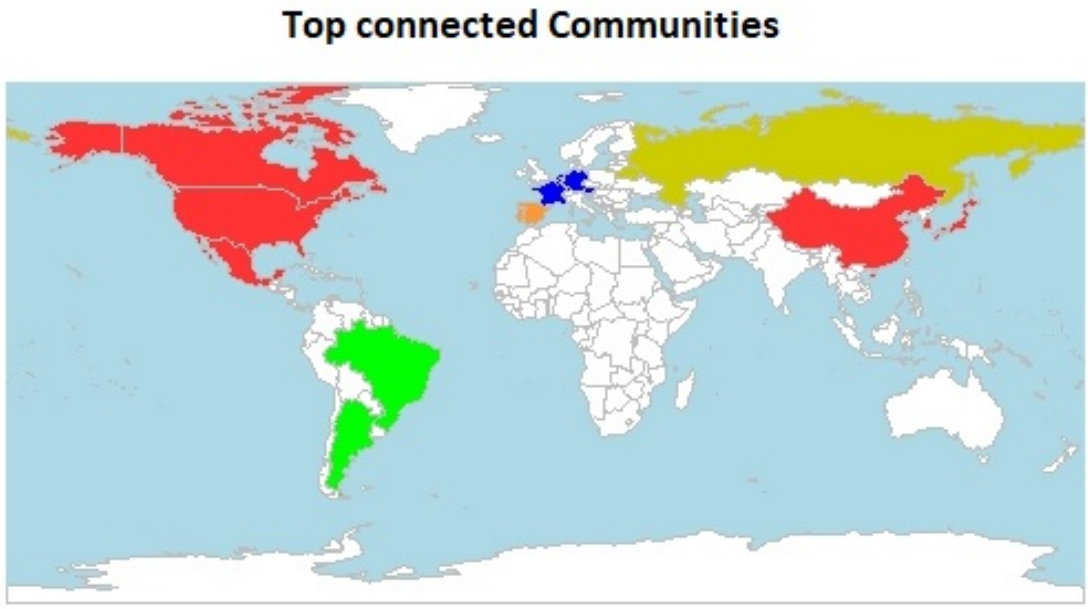


Figure 2.9: Top Connected Community Structure - $\xi_h = 1.35$

A significant feature of our approach is the fact that it allows to get deeper insight into the internal structure of each community and to give a measure of the mutual relationships between communities. Let us refer now to the clusters depicted in Figure 2.7 and detected with the maximum quality index criterion. In this regard, we display in Figure 2.10 the distributions of the communicability distances between pair of countries that belong to the same community. In particular, we compare the distributions for the first two relevant communities listed in Table 2.1

In fact, if we focus, for instance, on communities 1 and 2, we can inspect and compare their internal structure by providing some synthetic indicators in Table 2.2. From the analysis of Figure 2.10 and of the values shown in Table 2.2, we can say that the community 2 (let's say, USA-China) shows slightly more intense interactions than community 1 (let's say, Europe) since in the former the average intracluster distance is slightly lower than in the latter. However, although the largest number of countries that belong to community 1, a more compact distribution is observed with a lower volatility. Trading interactions between countries in community 1 appear indeed somehow more homogeneous than between countries in community 2. This is partially related to the geographical distribution of the countries inside the two communities. We have indeed that community 2 can be interpreted as the aggregation of different blocks mainly developed around USA, China and Japan.

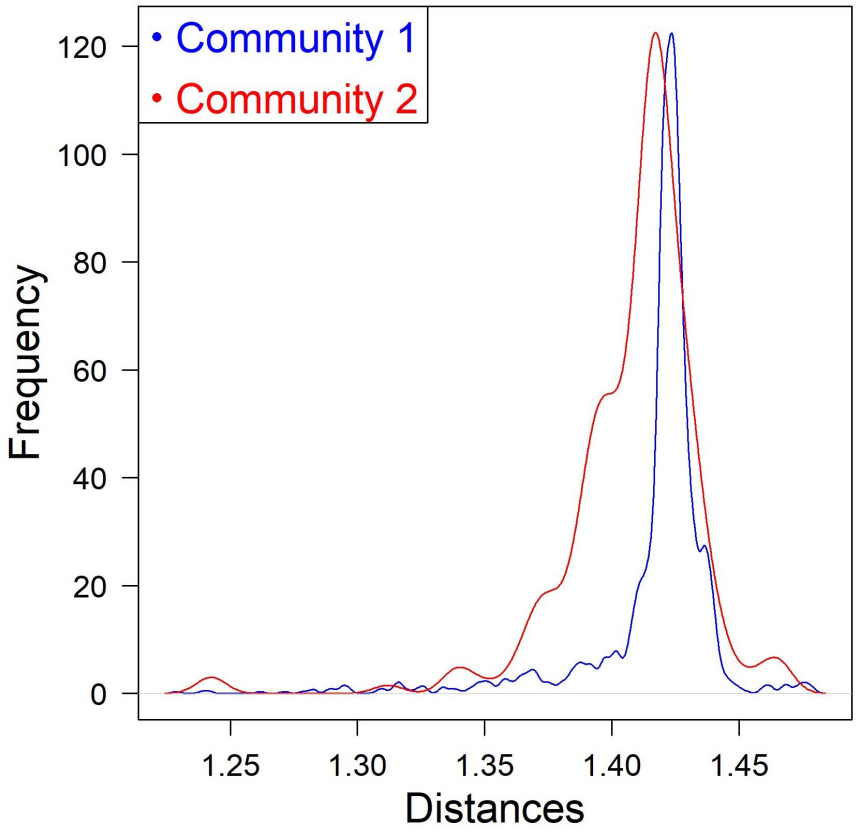


Figure 2.10: Distributions of Communicability Distances between countries of the same community. We display only the distributions related to the two main communities summarized in Table 2.1

Last column of Table 2.2 provides the same indicators computed on intercluster basis. This analysis allows to provide additional information in terms of heterogeneity in the group and between groups. It is worth pointing out the lower intercluster standard deviation. It means that couple of countries that belong to a different community has a similar distance between them.

It is noteworthy that additional insights can be provided by assessing the relevance of each country in the community. Indeed, communicability distance matrix provides a metric on the network and on each subnetwork, like a community. Therefore, we adapt the idea of closeness to our context, by providing the following communicability closeness to assess how effectively a node is supposed to spread trade flows through the network. Similarly to the definition of closeness, we define the *communicability*

	Intracluster		Intercluster
	Community 1	Community 2	Community 1 vs 2
Number of Nodes	54	21	—
Mean Distance	1.414	1.409	1.423
Min Distance	1.325	1.242	1.393
Closest Countries	NLD-BLX	USA-CAN	SAU-KOR
Max Distance	1.444	1.467	1.469
Furthest Countries	DEU-AFG	USA-LAO	USA-MNE
Standard Deviation	0.012	0.028	0.011

Table 2.2: Intercluster and Intracluster characteristics of the distributions of communicability distances. Columns Community 1 and Community 2 refer to the intracluster properties of the two main detected communities, in terms of number of nodes. Last column reports the corresponding intercluster properties computed between the same two communities.

closeness as:

$$C_i = \frac{1}{\sum_{j \in \mathscr{C}} \xi_{ij}} \tag{2.11}$$

where the sum is over all the internal nodes of the cluster $\mathscr{C}$ to which the node i belongs.

To exemplify, we rank in Figure 2.11 (left-hand side) the top 20 countries of community 2 on the basis of values of C_i. It is worth to stress that the centre of this community is located in China, Japan and South Korea and not in the North American sub-community. The three Asian nations are nowadays major traders and their high-level economic cooperation has been strengthened also because of the speed-up of the negotiations on the trilateral Free Trade Agreement. The three parties unanimously agreed to further increase the level of trade and investment liberalization based on the consensus reached in the Regional Comprehensive Economic Partnership Agreement[1].

Moreover, it is interesting to see that most central country in a community has not necessarily the same relevance on the whole network. We have indeed that, in terms of

[1]See "Fifteenth Round of Negotiations on a Free Trade Agreement among Japan, China and the Republic of Korea", April 12, 2019, , Ministry of Foreign Affairs of Japan and Free Trade Agreement (FTA) and Economic Partnership Agreement (EPA), 4 November 2019, Ministry of Foreign Affairs of Japan

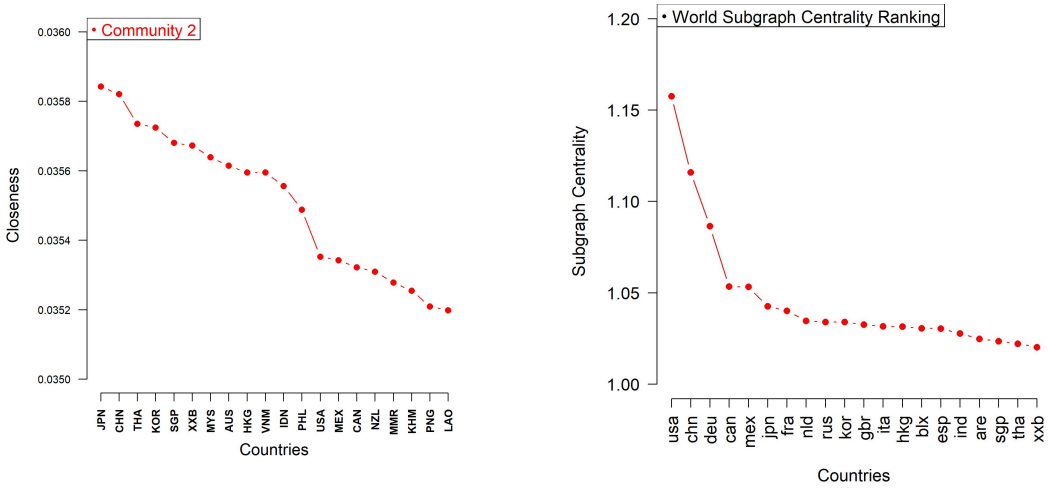


Figure 2.11: On the left-hand side, values of communicability closeness C_i for the top 20 countries inside community 2; on the right-hand side, world top 20 countries according to subgraph centrality rankings.

subgraph centrality, when we deal with the whole network (see Figure 2.11, right-hand side), USA appears as the key player followed by China and Germany. This ranking is inline with the top three countries provided by the World Trade Organizations, in terms of World's leading traders of goods and services [145].

Additionally, it is interesting to highlight that the relevance of countries reported in Figure 2.11 (right-hand side) is consistent with the Economic Complexity Index (ECI), introduced by [138]. The ECI allows to rank countries in the WTN according to the diversification of their export flows, which reflects the amount of knowledge that drives their growth. The higher is the ECI, the more advanced and diversified is an economy. In particular, countries whose economic complexity is greater than expected (on the basis of their global income), tend to grow faster than rich countries with a low ECI. In this perspective, ECI represents a suitable tool for comparing countries in the WTN independently of their total output and it has been extensively validated as a relevant economic measure by showing its capability to predict future economic changes and to explain international differences in countries incomes.
Although the network we analysed in the present work is based on the total normalised output and this fact prevents us from comparing directly their values with the ECI for a given country, there is a positive correlation between them. All the top 20 countries in Figure 2.11 (right-hand side) show a positive and high value of ECI. More specifically,

they kept a high value of ECI during the years preceding the year to which the network refers (2016) and this can justify the high value in the aforementioned centrality measures.

Finally, from the point of view a single country, it is worth to look for the closest trade partners, that is the nearest nodes in terms of communicability distance. Figures 2.12 show the distance profiles for China and Germany, respectively. For instance, looking at Figure 2.12 (right-hand side), we can notice countries, as Austria, Poland, Czech Republic that are characterized by a condition of strong dependence on Germany, that is a major player in the network. Similarly Figure 2.12 (left-hand side) shows how strong is the commercial relationship between China and Hong Kong, also as a result of the trade agreements between the two countries, like CEPA (Closer Economic Partnership Arrangement) aimed at eliminating duties on large categories of products. Indeed, it is well-known that, for the Chinese trade market, Hong Kong plays a crucial role since foreign companies use Hong Kong as a springboard to invest in China thanks to its infrastructure network that has no equal in the world, investor protection, transparent and efficient judicial system, legal certainty.

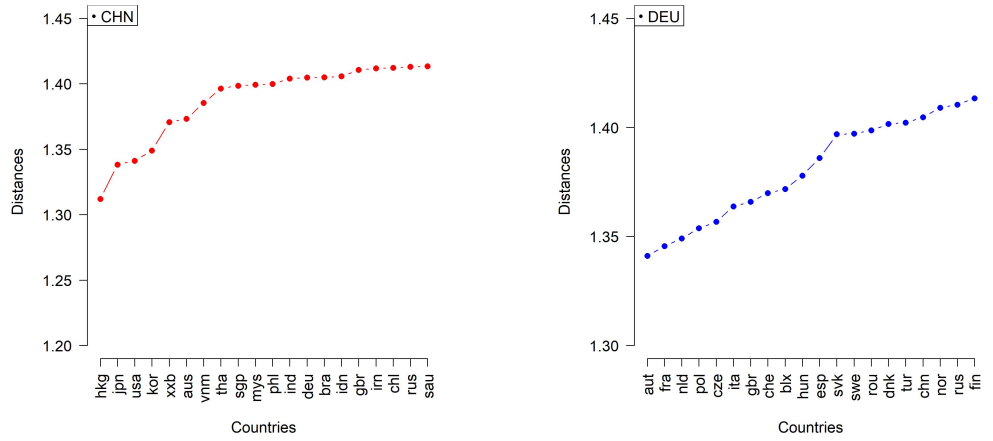


Figure 2.12: Top 20 nearest countries for China (left) and Germany (right)

2.6.3.2 Results in terms of resistance metric

The methodology described in Section 2.6.2 has also been applied using the resistance distance ω. In this case, we consider the total trade of a given country as flow of the

global wealth that has been produced during a year. Therefore, the Gross Domestic Product (GDP) is the attribute of interest on each node. In this regard, the effective resistance of an edge expresses how easily (or not) a unit flow moves from a country to another one, i.e. how easily two countries trade a unit of wealth, independently of its nature. It is noteworthy that, according to formula 2.8, the resistance distance between a pair of countries depends on the values of the vibrational centralities of both countries (the more central these countries are in the network, the less is the resistance distance between them) and on the value of their mutual correlation (the more correlated they are and again the less is their distance).

Therefore, we construct the vibrational communicability matrix $\mathbf{G}^v$ on the normalised network $\mathscr{G}_2$, and the corresponding resistance distance matrix $\mathbf{\Omega}$. Using this metric, we find that the nearest countries are, again, USA and Canada with a distance $\omega_{\min} = 1.238$ and the farthest countries are USA and Germany with a distance $\omega_{\max} = 1.497$. For each value of the threshold distance between minimum and maximum, we obtain the corresponding partition in communities. The maximum partition quality index Q corresponds to 15 communities plus isolated nodes. In Figure 2.13, we plot the value of Q in red and the number of communities, counting each isolated node as an independent one, in blue as functions of the threshold ω_h. The maximum quality index Q is reached at a threshold distance $\omega_h = 1.365$. The main characteristic of this partition is the presence of a giant component of 127 nodes e 14 other components with few nodes.
Main results in terms of geographical distribution are displayed in Figure 2.14 and, as in the previous Section, we summarize in Table 2.3 main composition of top communities in terms of number of constituents.

With respect to results based on communicability, we have that the first community has a larger number of countries (equal to 127). Additionally, the larger community includes again main Asian and Oceanian countries as well as several African countries. It is noteworthy that North America behaves as a separate cluster. This result is in line with the literature that emphasizes the interesting economic relation between Asia and Oceania. Several works showed that the Asia-Oceania community collapsed after China entered the WTO in 2001 and built strong trade relationships with other communities, especially with the external cores, (i.e. the United States and Germany). China then became regionally attractive and restored the Asia-Oceania community as the community leader after it gained a significant portion of trade globally (see, e.g., [146]).
Significant differences are also observed for the European community (see community

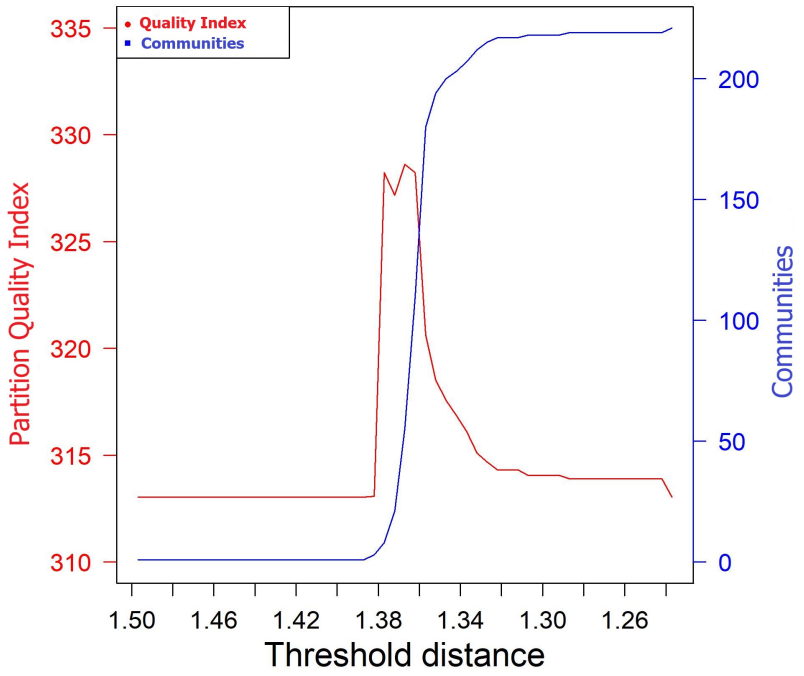


Figure 2.13: Partition quality index Q (red line) and number of communities (blue line) as functions of the threshold resistance distance. Maximum Q is observed for $\omega_h = 1.365$.

2 in Table 2.3). Norway and Sweden and Great Britain and Ireland provide indeed two separate groups with respect to main European economic groups.

It is worth pointing out that communities detected above represent groups of countries showing a positive correlation in their trade strength, whereas members of different clusters show a negative correlation. Being strongly anti-correlated means that when the total trade deficit of a country grows, the total trade surplus of a second country grows too. For instance, Japan and USA have been classified by the methodology in different communities. Indeed, in the literature, empirical analyses show a negative correlation coefficient between normalised trade strengths of these countries (see, e.g., [147] and [148]). Similar arguments can be extended also to other pairs of countries. For instance, Germany is negatively correlated with USA (see [147]) and show a high positive correlation with Belgium and France (see [148]), that belong to the same com-

munity.

If we disentangle communities characterized by very tight relationships between countries, the results seem strictly related to the ECI index. We may expect that, if two countries communicate well, then their ECI's could be similar. That is, if their mutual distance is small, both in terms of communicability metric and resistance metric, then they display similar values of ECI. In fact, the existence of multiple channels of trade exchange between them would result in a similar diversification of their output. This means that countries inside each community (could) share homogeneous values of ECI. Concerning Table 2.3, we notice small clusters whose components show homogeneous values of the ECI index. For instance, community 6 is formed by Russia (with an ECI of 0.855 in 2016) and Belarus (with an ECI of 0.744 in the same year). Similarly Canada (1.084), Mexico (1.160) and USA (1.781); Norway (1.199) and Sweden (1.862); UK (1.549) and Ireland (1.409); Brazil (0.648) and Argentina (0.380) that constitute communities 3, 4, 5 and 7, respectively.

Figure 2.14: Communities detected by using the procedure based on the resistance matrix and considering the threshold ω_h that maximize the partition quality index Q.

As in the previous Section, we explore main characteristics of two most relevant communities (see Table 2.4), It is noticeable that, although the two groups show a very similar mean distance, European countries are characterized by a higher heterogeneity. Focusing on intercluster indicators, we notice also a lower similarity between the two

	Size	Members
Community 1	127	AUS CHN HKG IDN IND IRN IRQ JPN KOR LAO PHL THA and others
Community 2	11	AUT BLX CZE DEU ESP FRA HUN ITA NLD POL PRT SVK
Community 3	3	CAN MEX USA
Community 4	2	NOR SWE
Community 5	2	GBR IRL
Community 6	2	BLR RUS
Community 7	2	ARG BRA

Table 2.3: Members for the seven main communities in terms of number of countries, obtained by applying the procedure based on the resistance distance. (For the names of the countries we refer to the current officially assigned ISO 3166-1 alpha-3 codes.)

communities with respect to Table 2.2 based on communicability.

	Intracluster		Intercluster
	Community 1	Community 2	Community 1 vs 2
Number of Nodes	127	12	—
Mean Distance	1.414	1.391	1.418
Min Distance	1.290	1.324	1.395
Closest Countries	CHN-HKG	AUT-DEU	SMR-DEU
Max Distance	1.429	1.424	1.466
Furthest Countries	ARE-HKG	AUT-PRT	JPN-DEU
Standard Deviation	0.008	0.026	0.010

Table 2.4: Intercluster and intracluster characteristics of the distributions of resistance distances. Columns Community 1 and Community 2 refer to the intracluster properties of the two main detected communities, in terms of number of nodes. Last column reports the corresponding intercluster properties computed between the same two communities.

The relevance of a country can be now assessed in terms of vibrational centrality. To this end, we display in Figure 2.15, the top 20 countries, calculated over the whole network. China, USA and Germany are again in the top 3, with China playing as the best spreader node. Also in this case, almost all the top 20 has a positive ECI. A comparison between Figures 2.11 and 2.15 confirms the different role played by USA and China in the global network. As confirmed by [145], USA is the leading commercial

service provider and in such a way it is widespread well-integrated in the global market; on the other side, China plays the role of hub for goods and represents the leading merchandise trader and this gives to the country a very robust position which makes it less vulnerable to market turmoil.

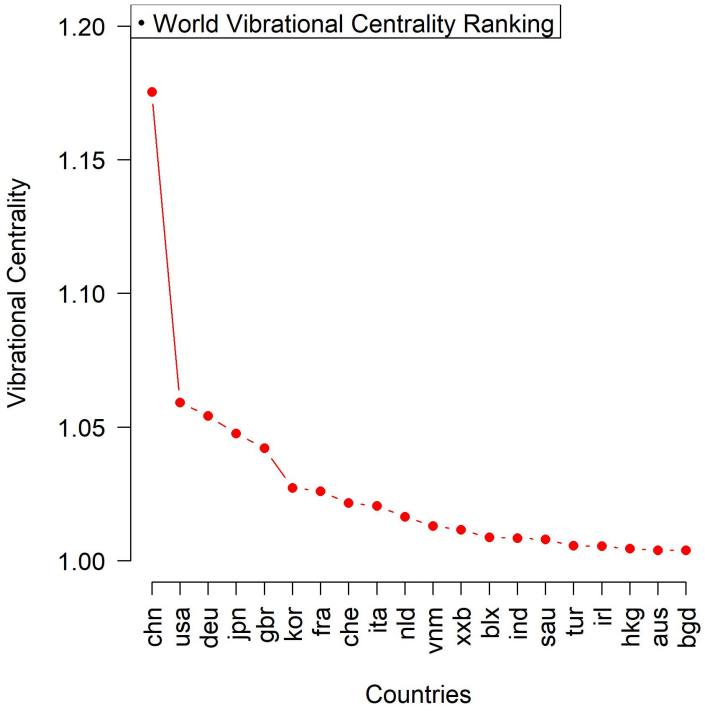


Figure 2.15: World top 20 countries according to vibrational centrality rankings

Finally, from the point of view a single country, it is worth to look for the closest trade partners, that is the nearest nodes in terms of resistance distance. Figure 2.16 shows the distance profiles for the most central country of community 1 and 2, respectively. These plots can be interpreted as the list, in decreasing order, of countries that are most positively correlated with the selected centre, China or Germany. For instance, while in terms of communicability distance China is well-communicating with USA (third position in Figure 2.12), USA does not belong to the top 20 most correlated countries with China. Rather, the left-hand side in figure 2.16 clearly shows a driving and synchronizing effect of the Chinese giant in the entire South-East Asia area. Simi-

larly, figure 2.16 (right-hand side) confirms the role of Germany in the European Union and the strong correlation with Austria, Czech Republic and Poland.

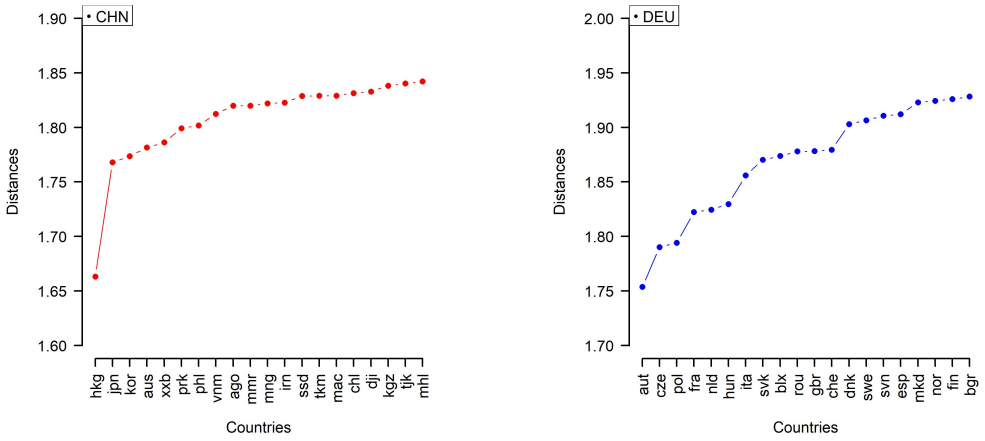


Figure 2.16: Top 20 nearest countries for China (left) and Germany (right)

2.6.4 Comparison with different approaches applied to the same network

It is worth briefly comparing our results with those obtained by other methodologies on the same network (see [124] and [149]). In particular, in [149], several approaches are proposed to analyse the community structure of the WTN at different times. The authors showed that the recognition of mesoscale structures is increasingly difficult also because the world is becoming increasingly global over time. This makes even more compelling the search for a method that forces even slight deviations from a random structure to emerge. Both directed and undirected networks have been tested, although no significant differences have been found. As in our case, results reported in [149] show that geographical proximity still matters for international trade, jointly with trade agreements, common language or religion, and traditional partnerships. In particular, focusing on the application of a classical maximum modularity criterion, the authors find in 2008 (the most recent year of their analyses) three big communities containing 68, 66, and 47 countries, with the largest cluster associated with Asia and Oceania. This is partially in line with our result in which a large relevant community including China, Oceania and North America is observed. On the other hand, by using either communicability or resistance distance, we found a higher level of granularity.

Additionally, our approach provides a higher flexibility allowing to emphasize stronger connections when the threshold decreases.

The authors in [149] also adopt a notion of distance among nodes based on random walks by row-normalizing the weighted matrix. Modelling the WTN by stochastic matrix corresponds to moving from absolute to relative trade values. That distance between nodes is defined by complementing a similarity measure. A dendrogram is computed initially by defining groups containing single nodes and then by iteratively linking pairs of groups with minimal distance. This approach looks similar to ours being based on a varying threshold. They choose to maximize the so-called cophenetic correlation coefficient, which is defined as the linear correlation between the distances and the cophenetic distances, which are the heights of the link joining (directly or indirectly) nodes in the dendrogram.
Some common evidences are noticeable also in this case. The United States and Canada form one of the strongest partnerships: their distance in the dendrogram stays constantly very small over time. France is strongly connected to some of its former colonies, as we also pointed out above, whereas Germany is close to other European countries. Main differences are related to the behaviour of very small countries. While, in our case, small countries are often classified as isolated nodes, in [149], very small countries are connected to much larger ones as an effect of the disassortativity observed in the WTN. These links tend to be small in absolute terms, given the small economic size of the countries, but they appear as relevant in relative terms, because the strong preference for a given partner.

The authors in [149] also used stability and persistence to confirm their results. A random walker starting in a community is likely to remain for quite a long time within that community, before leaving it to enter another one. The analysis of the persistence probabilities induced in a network by a given partition has recently been proven to be an effective tool for testing the existence and significance of communities. Also in this case, we observe that communities with high persistence probability have common features with our results. Indeed, the top communities identified in [149] considers the entire set of European countries, plus a number of minor non-European partners, that is in line with the top community selected by the communicability approach. Similarly, the second large community with a high persistence probability includes the entire North America and most of Central and South America, plus China, Australia, and many others. Although less granular, this community is fully comparable with community 2 detected by the communicability approach.

A quantitative correlation between the world partition in communities obtained by

a modularity criterion and geographical distances has been investigated in [124]. The authors, both at an aggregate level and at a number of commodity-specific levels, compare the two maximum modularity partitions of the input-output network and of the weighted network of the geographical closenesses. They find a high similarity between aggregate trade and geography-based communities, greater than, for instance, communities determined by regional trade agreements. They conclude that geographically-related factors explain the patterns of global trade more than political determinants. Although a positive correlation is present between monetary flows and geographical closenesses, we noticed that the geographical distances are less relevant when indirect relationships are also considered via either communicability or resistance distances[1]. As a consequence, the community structure we find appears more granular than the groups found in [124] and the composition cannot be explained only by geographical patterns. Other factors are involved as historical relationships, trade agreements and strategic economic alliances.

To conclude, although some common results with [124] and [149] are observed, our methodology has the advantage of clearly highlighting even small differences and forcing the emergence of very strong ties between different countries through the use of a distance threshold. Furthermore the partition quality index Q we applied turns out to be a simple and flexible tool, more homogeneous to the context of a network interpreted as a metric space.

2.7 Conclusions and further research

Community detection is a key topic in the analysis of complex systems, where discovering the inner structure plays a relevant role. In particular, the centrality of countries and the relationships between them assume specific relevance in the World Trade Network, where economical and geopolitical phenomena affect over time the structure of the global network. In this framework, this work aimed at detecting different levels of clustered communities in the network on the basis of both communicability and resistance distances. The proposed methodology allows to discover the hidden hierarchical structure of the network, as it presents a degree of flexibility highlighting very tight relationships by varying the threshold parameter, and revealing in this way the clusters of nodes that more easily communicate. Moreover, it performs well also for weighted and extremely dense network, as the case of the WTN.

[1] The rank correlation between these distances and the geographical distance between capital cities is lower than 0.15.

Features and properties of each community can be exploited in order to compare the characteristics of different clusters and to detect the most central countries inside the single community as well in the whole network.
Numerical results depict the structure of the economic trade detecting main relevant communities. In particular, main community sees United States and China as main actors. Most flows are polarized around the exchange channel between China and USA and all their satellite countries. However, focusing on the correlation between trades, the procedure emphasizes the different role of these two countries. In particular, it is worth mentioning the emerging of China-Oceania community when deep links emerge. Furthermore, it is confirmed that Germany plays a key role in Europe and preferential channels of internal exchanges are observed in the European market. In line with [146], emphasizing tight links, we obtain that although the strong trade relationships with USA and Germany, China became regionally attractive and restored the leadership of Asia-Oceania community. European community is highly centralized around founding members of the European Economic Community with the central role of Germany. High income countries in Northern Europe are instead in a separate community with a less relevant role in the network.

Chapter 3

Multi-attribute community detection in International Trade Network

3.1 Introduction

In network theory, a specific way to detect vertices having a peculiar common feature is termed clustering or community detection. Formally, a cluster, or a community, is a subgraph whose similarity or internal connections are stronger than the ones with the rest of the graph [103]. In recent years there was a surge of interest on the community structure of international trade [112, 113, 114, 115, 118, 124, 150]. The classical approach consists in finding sets of countries which are densely connected, through preferential economic relationships. A typical representation of this phenomenon is through a directed and weighted network, where nodes are countries and weighted links represent the aggregate trade flows. This representation is named in the literature as the International Trade Network (ITN).

Under this perspective, it becomes important to map the input-output interrelations among the countries through an inspection of the communities, where two countries share the same community if they have a comparable intensity in the trade flows or if they have preferential trade flows.

International trade has been widely studied in the literature showing that main characteristics have changed over time, with an acceleration of modifications occurring in the last decades. In particular, over the years, the composition of trade flows changed making countries even more deeply interconnected. The geographical distribution of

trade also varied, with an increasing role of the emerging countries, especially in Asia.[1]

To detect the network structure, a key function is played by the vertex centrality. The idea of centrality is quite simple to grasp: a numerical score is assigned to each node of the network so that the higher the score, the more central the node in the network. The literature has highlighted the importance to be central in an economic network (see [122, 130]). In particular, centrality may be associated with countries that are the most important hub of the ITN, even though they are not leading import or export countries [121, 122]. There are different metrics describing centrality, but it has been shown that different measures (degree, coreness, etc.) identify different influential nodes [111]. For instance, a node could be central if it is directly connected with many other nodes, if it has an intermediary role in communication, and so on. Indeed, there is no consensus on an univocal definition of network centrality, because each measure considers only one specific concept (see, e.g., [151]). But, resorting to only one of them is discarding a large amount of the whole information available. Related to centrality, the clustering coefficient is also an important index to measure the interconnections within a community. This coefficient has been developed in all the cases of weighted, unweighted, directed and undirected networks (see [99, 101, 152, 153, 154, 155]). In particular, [156] discusses the clustering coefficient in presence of already established communities for directed networks and [102] presents a concept of clustering coefficient which also includes the presence of missing indirect links in the construction of triangles. The association between communities and clustering coefficients is quite natural. Triangles are the easiest geometric visualization of communities, providing a picture of non-exclusive interactions among different agents. The relevance of this coefficient has been investigated also in the context of ITN (see, e.g., [120, 121, 132, 133]). As stressed in [124], detecting the community structure of the ITN and how it correlates with country-specific variables and geography (e.g., distances between countries) is crucial from an international-trade perspective. Indeed, finding communities in the ITN means identifying clusters of countries that carry tightly interrelated trade linkages among them, while being relatively less interconnected with countries outside the cluster. In this work, we provide a new methodology for clustering countries based on a multi-criteria assessment of several topological indicators of centrality. The method consists of two steps. In the first step, we rank countries in ITN, according to various centrality measures. In the second one, based on those rankings, we compute the similarities between countries and then we apply the clustering algorithm based on the Clique Partition model.

[1] https : //www.wto.org/english/res_e/publications_e/anrep10_e.htm

More specifically, in the first step, and unlike classical methodologies, we consider all the most prominent centrality definitions proposed in the literature that are relevant to international trade. Rather than advocate the superiority of one of them, we aggregate this rich multi-criteria assessment by defining a proper measure of similarity/dissimilarity between nations using their ranking positions. Next, we group together countries that have common structural features in terms of those rankings. The main advantage of our proposal is that we do not focus on a single and specific indicator of centrality, nor we come out with a detailed countries ranking. Rather, we are able to identify groups of countries that have similar structural properties in the ITN. A specific tool developed for our project is a new heuristic algorithm to find clusters, based on the Clique Partition model [15, 157, 158]. The Clique Partition model consists of partitioning the vertices of a graph into the smallest number of cliques. First, a measure of similarity/dissimilarity between units must be established. This measure can take both positive and negative values, respectively if two units are similar or dissimilar. Then units must be partitioned in subsets, in such a way to maximize the similarity between them. This model has some advantages over the classical k-means or hierarchical models. First of all, the clique partition model does not require either that the number of clusters were fixed in advance, e.g. the parameter k, or that the user should arbitrarily analyse the chart of the hierarchical clusters. Rather, the number of cluster results by the optimization of an objective function. Moreover, outliers are not forced to be in a clusters, but they can form peculiar groups of a single element. Finally, the principle of the method is that cluster are composed of mutually homogeneous data, while the k-means models first try to establish cluster's centres and then groups are composed by units that are similar to centres. Conversely, the clique partitioning forms groups of similar units. Experimental comparison between the clique partition and other clustering methods can be found in [159].

The chapter is organized as follows. In Section 3.2, we recall main literature related to network theory, analysis of ITN and main solution methods for clique partitioning problems. In Section 3.3, we describe the methodological framework and the integer linear programming problem. In Section 3.3.2, we define the maximum clique partition problem as well as the algorithm applied for identifying the optimal solution. In Section 3.4, a numerical application is developed by using the paradigmatic case of the ITN. Conclusions follow in Section 3.5.

3.2 Related literature

In this section we briefly remind the main literature related to network theory and International Trade, as well as clique partitioning problems and the main solution methods.

Network theory has been traditionally used in sociology and political science in order to investigate international trade relations, being an effective tool in revealing the core-periphery structure of the countries or in studying the impact of the globalization on the international trade structure [123, 135, 160]. The topological and statistical properties of the international trades, also in a time perspective, have been deeply studied in several works (see for instance, [112, 116, 117]). More recently, complex networks have also been used to investigate economic and financial implications of the world trade. For instance, Kali and Reyes [136, 161] study the country's role in the ITN deducing important implications in terms of economic growth and explaining the phenomenon of financial contagion. Both international trade and financial integration patterns are investigated by Fagiolo et al. [127]. Another important issue is the identification of communities in the trade network. Barigozzi et al. [124] deeply study the topology of the international trade multi-network, aiming at discovering its community structure. In [119], the authors analyse the evolution of communities ("islands"): from two large trading communities, centred on UK and US, to a fairly heterogeneous "archipelago" of trade, that seems to reflect a phenomenon of globalization. Finally, dissimilarities between different layers of an international trade multiplex network have been studied in [162]. The authors characterize each layer as a commodity network in a specific time period. The definition of communities can be naturally associated with a partition in clusters, and one of the most important model of community detection is the clique partition. The presence of communities inside the network is revealed by the modularity index (see [11, 163]), that corresponds to the objective function of a clique partition model. By maximizing the partition modularity, one can determine the community structure of the network [14, 164, 165, 166, 167]. The clique partition model, as a combinatorial approach to cluster qualitative data, had a methodological development independent of the problem of community detection, as it has been introduced in [15, 157, 158, 168] and its applications range in many different fields (see, for instance, [169]). It has been recognized that it is a NP-hard problem, implying that the exact solution cannot be computed in polynomial time, unless P=NP. In practice, exact methods can solve instances that do not exceed one hundred nodes [167, 170], so that the use of heuristic procedure is necessary in our applications [163, 171].

3.3 The model

In this section, we describe our methodology for clustering countries on the basis of the similarity attributes.
A network is described by a graph $G = (V, E)$ where V and E are respectively the set of n vertices and m links (or edges). Two nodes are adjacent if there is a link (i, j) connecting them. The degree d_i of a node i is the number of links incident to it. If a weight $w_{ij} > 0$ is associated with each link (i, j), a weighted graph $G = (V, E, W)$ is obtained, being W the set of weights. In general, both adjacency relationships between vertices of G and weights on the links are described by a nonnegative, real n-square matrix $\mathbf{W}$. In the unweighted case, matrix $\mathbf{W}$ is simply the classical binary adjacency matrix $\mathbf{A}$, of entries a_{ij}, where $a_{ij} = 1$ if $(i, j) \in E$, 0 otherwise. Since we consider network without loops, $a_{ii} = 0$ (or $w_{ii} = 0$). The $(i, j)-$element of the $k-$power of $\mathbf{A}$ is the number of walks of length k from i to j. The Laplacian matrix is defined as $\mathbf{L} = \mathbf{D} - \mathbf{A}$, where $\mathbf{D}$ is the diagonal matrix having the vertex degrees on the diagonal entries.
A network is directed if each link is directed, that is an arc $(i, j) \in E$ means that there is a link starting from i and ending in j. The in-degree d_i^{in} (out-degree d_i^{out}) of a node i is the number of arcs pointing towards (starting from) i. The degree d_i^{tot} of a vertex is then the sum of the in and out-degree. In the directed case, matrices $\mathbf{A}$, for a binary network, and $\mathbf{W}$, for a weighted network, are not symmetric.

3.3.1 Network attributes and rankings

We are interested in specific characteristics of the nodes, such as their centrality or their level of interconnection within the network. Since the network is weighted and directed, we need appropriate measures that take into accounts both weights and directions. Thus, according to the four dimensions classification of centrality indices in [172], we focus on four class of network indicators, each one computed using both incoming and outgoing links. These are in and out-strength, in and out-clustering, hub and authority and Laplacian centrality.

The strength (in and out) is the natural extension to the weighted and directed case of the degree centrality. It counts both the number of ties and their intensity. Formally, for a node i, we have:

$$s_i^{in} = (\mathbf{A}^T\mathbf{W})_{ii} = \mathbf{W}_i^T\mathbf{1} \tag{3.1}$$

$$s_i^{out} = (\mathbf{A}\mathbf{W}^T)_{ii} = \mathbf{W}_i\mathbf{1} \tag{3.2}$$

where $\mathbf{W}_i$ corresponds to the $i-th$ row of the matrix $\mathbf{W}$.

In particular, in our application, the in-strength s_i^{in} measures the total trade flows incoming to the country i, that is the import. The out-strength s_i^{out} measures the total trade flows outgoing from the country i, that is the export.

Clustering coefficient measures the tendency of a node to be well interconnected with its neighbours. Local clustering coefficient of a node i counts the number of observed weighted directed triangles connected to i, divided by all its potential unweighted directed triangles:

$$c_i(\tilde{\mathbf{W}}) = \frac{\frac{1}{2}[(\tilde{\mathbf{W}}^{[\frac{1}{3}]} + (\tilde{\mathbf{W}}^T)^{[\frac{1}{3}]}]^3_{ii}}{d_i^{tot}(d_i^{tot}-1) - 2d_i^{\leftrightarrow}}, \tag{3.3}$$

where $\tilde{\mathbf{W}} = [\tilde{w}_{ij}]_{i,j\in V}$ is the normalized weighted matrix whose elements are defined as $\tilde{w}_{ij} = \frac{w_{ij}}{max(w_{ij})}$ and $d_i^{\leftrightarrow} = \sum_{j\neq i} a_{ij}a_{ji}$ is the degree of bilateral arcs between the node i and its adjacent nodes.

As pointed out in [101] and [99], we have four types of directed triangles to which i could belong. They generate four types of clustering coefficients, that can be separately computed.

Formula (3.3) includes all the four coefficients described in [99]. Nevertheless, the country i is part of the in-type and out-type triangles, highlighting the presence/role of the node i in import/export between its neighbouring countries. Thus, in our analysis, in-clustering and out-clustering coefficients seem more appropriate in capturing the role of the node i in the exchanges between the closest countries, distinguishing between import and export:

$$c_i^{in}(\tilde{\mathbf{W}}) = \frac{\frac{1}{2}(\tilde{\mathbf{W}}^T\tilde{\mathbf{W}}^2)_{ii}}{d_i^{in}(d_i^{in}-1)}, \tag{3.4}$$

$$c_i^{out}(\tilde{\mathbf{W}}) = \frac{\frac{1}{2}(\tilde{\mathbf{W}}^2\tilde{\mathbf{W}}^T)_{ii}}{d_i^{out}(d_i^{out}-1)}. \tag{3.5}$$

In order to model the influence, or the prominence, of a country in a global scenario of trade flows, the eigenvector centrality is the most suitable measure. The generalization of this measure to directed networks allows to associate with a node two status: authority and hubness. The idea arises in the context of web page search to rank the importance of a page [173]. A web page is an authority if it is pointed by many other

pages. Hubs are pages that link to many authoritative pages. Formally, let a_i and h_i be the authority and hub scores respectively. Then, the following relations hold:

$$a_i = (\mathbf{W}^T \mathbf{h})_i \tag{3.6}$$

and

$$h_i = (\mathbf{W}\mathbf{a})_i \tag{3.7}$$

where the vectors **a** and **h** collect respectively authorities and hubs scores of all nodes.

By formulas (3.6) and (3.7), definitions of hubs and authorities are characterized by a mutually reinforcing relationship: essentially, a good hub is a page that points to many good authorities; a good authority is a page that is pointed to by many good hubs. The use of these measures is motivated by their interpretation: on one hand, authorities are central countries as they import in turn from central countries. On the other hand, hubs are central as they export towards central countries.

To compute the scores (3.6) and (3.7), an iterative algorithm (HITS - Hyperlink Induced Topic Search) is proposed in [173]. Starting with initial score vectors $\mathbf{a}^0$ and $\mathbf{h}^0$, through the power iteration method on $\mathbf{A}\mathbf{A}^T$ and $\mathbf{A}^T\mathbf{A}$, the process converges to the principal eigenvectors **a*** and **h*** of the matrices $\mathbf{A}\mathbf{A}^T$ and $\mathbf{A}^T\mathbf{A}$.

The idea behind the Laplacian centrality is that the importance of a vertex i is related to the network ability to adapt itself to the deletion of the vertex, i.e. its resilience. The Laplacian centrality of a vertex i is reflected by the drop of the Laplacian energy of the network deriving by the deletion of i from the network. According to [174], the definition[1] of the Laplacian energy is:

$$E_L(G) = \sum_k \lambda_k^2 \tag{3.8}$$

where λ_k are the eigenvalues of the Laplacian $\mathbf{L}$. Therefore, the Laplacian centrality is (see [176]):

$$l_i = \frac{E_L(G) - E_L(G_i)}{E_L(G)} = \frac{(\Delta E)_i}{E_L(G)}. \tag{3.9}$$

Since the denominator $E_L(G)$ has the same value for all vertices, we focus on the numerator $(\Delta E)_i$, that is always nonnegative for the interlacing property of the eigenvalues of the Laplacian matrix (see [177]). The Laplacian energy can be re-expressed

[1] It is noteworthy that an alternative definition of Laplacian energy has been provided in the literature (see [175]). Although this alternative definition has been widely explored in the literature, we focus on the original version defined in [174] because it is related to the Laplacian centrality measure.

in terms of strength[1] (see [176], Th. 1):

$$E_L(G) = \sum_k s_k^2 + 2\sum_{k<j} w_{kj}^2. \tag{3.10}$$

Hence, the difference $(\Delta E)_i$ is:

$$(\Delta E)_i = s_i^2 + \sum_{k\in N(i)} (w_{ki}^2 + 2s_k w_{ki}) \tag{3.11}$$

where $N(i)$ is the set of neighbours of the node i. This expression allows the following interpretation of the Laplacian centrality of i. This centrality depends (in a quadratic way) on the strength and on the weights of the neighbours of i.

As stressed in [176] and [178], compared with other standard centrality measures proposed for weighted networks (e.g. strength or betweenness centrality), the Laplacian centrality is an intermediate measure between global and local characterization of the importance of a vertex. The generalization to directed and weighted case follows[2], giving an expression for weighted and directed Laplacian centrality (in and out) l_i^{in} and l_i^{out} derived by formula (3.11).

In our analysis, we intend to aggregate different indicators. Indeed, as already stressed, each measure has peculiarities and characteristics that highlight various aspects of the exchange relations between countries,

This heterogeneity requires an approach that cannot be simply based on the direct comparison among extremely different measures.

Given that each index has specific unit measures and range of variations, we will focus on the various country centrality rankings rather than their absolute values. More specifically, first we calculate the country rankings according to any index, then we cluster countries according to their positions on those rankings. Indeed, each indicator induces a ranking which represents the structural importance of a single node in the network. Rankings analysis allows us to compare more than one centrality simultaneously. The comparison will be developed by computing a distance function between rankings. In particular in this work we refer to the Minkowski distance, also known as L_p-norm distance.

Let us order the scores of each node obtained for each centrality measure k and let

[1] In case of unweighted graphs, formula (3.10) gives the result provided in [174]: $E_L(G) = \sum_k d_k\,(d_k+1) = \sum_k d_k^2 + 2m$. The use of entries of the Laplacian matrix, instead of eigenvalues, is meaningful especially for large networks.

[2] See [179] and [180] for two definitions of Laplacian energy for directed graphs.

r_i^k be the position of the node i with respect to k. The Minkowski distance $d(\mathbf{r}_i, \mathbf{r}_j)$ is

$$d(\mathbf{r}_i, \mathbf{r}_j) = ||\mathbf{r}_i - \mathbf{r}_j||_p = \left(\sum_{k=1}^{K} \left| r_i^k - r_j^k \right|^p \right)^{1/p} \tag{3.12}$$

being $\mathbf{r}_i$ the rankings vector of node i, K the number of considered centrality measures and p any real value such that[1] $p \geq 1$.
This distance measure is commonly used in the literature for computing the dissimilarity of objects described by numeric attributes. It is a generalized distance metric that includes others as special cases. In fact, although theoretically infinite measures exist by varying the value of p, just three have gained importance (Manhattan distance for $p = 1$, Euclidean distance for $p = 2$ and Chebyshev distance for $p \to \infty$).

A remarkable feature of this distance consists in grouping more than one objects, namely it allows to consider all the network indicators simultaneously, producing a global fictitious distance between couple of nodes ranking. Furthermore, this distance allows to exploit several values of p in order to better highlight the general features of the analysed data (see [181, 182]). For instance, [182] highlights how different configurations of data concentration can be caught varying p, so that Minkowski distance can be used for effectively tackling data analysis problems.

In our context, we use this distance to construct a complete network K_n having the same node set and weighted adjacency matrix $\mathbf{\Omega}$, whose entries are defined as:

$$\omega_{ij} = \begin{cases} \frac{1}{1+d(\mathbf{r}_i, \mathbf{r}_j)} & \text{for } i \neq j \\ 0 & \text{for } i = j \end{cases} . \tag{3.13}$$

These weights range in $[0; 1]$ and turn out to be effective in describing the similarities between countries. Indeed, the more two countries have a similar behaviour, the smaller is the distance and the higher is the weight.

3.3.2 The Maximum Clique Partition Problem

The Clique Partition (CP) problem, as applied to our model, is defined as follows. The complete undirected graph $G = (V, E)$ is given, with $V = \{1, \ldots, n\}$. For each $(i, j) \in E$, gains/costs g_{ij} are defined, which can take both positive and negative values. In our application, positive values of g_{ij} are similarities, negative values are dissimilarities. Let $P = \{V_1, V_2, \ldots, V_q\}$ be a partition of V and let $\pi(V_k) = \sum_{i,j \in V_k} g_{ij}$ be the gains/costs

[1] Although p can be any real number, when $p < 1$ the formula does not define a metric, being the triangle inequality not satisfied.

sum of subset V_k, for $1 \leq k \leq q$. The CP problem consists of finding the node partition P that maximizes the objective function $f(P) = \sum_{k=1}^{q} \pi(V_k)$.

It is important to note that values g_{ij} must be both positive and negative, otherwise there is no incentive to discard negative values and the best partition would be the total set $P = \{V\}$. Therefore, we calculate g_{ij} as the difference between ω_{ij} (that are positive and bounded between 0 and 1) and benchmark values ω^*_{ij}, representing a neutral threshold. Neutral thresholds are calculated as follows. Let $\omega = \sum_{ij} \omega_{ij}$ be the total network similarities and let $\omega_i = \sum_j \omega_{ij}$ the sum of similarities appointed to unit i. The probability that a unit x of network similarity would be allocated to node i is $\mathrm{P}[x \text{ incident to } i] = \omega_i/\omega$. If similarity has no structure, that is, it is independent of pairs (i, j) because data do not have clusters, then:

$$\begin{aligned} &\mathrm{P}[x \text{ incident to } i \cap x \text{ incident to } j] = \\ &\mathrm{P}[x \text{ incident to } i] \times \mathrm{P}[x \text{ incident to } j] = \\ &\omega_i \omega_j / \omega^2. \end{aligned} \tag{3.14}$$

Then, if similarities are independent, the expected similarity between i and j should be: $\omega^*_{i,j} = 2\frac{\omega_i \omega_j}{\omega}$. So, we can calculate gain/cost g_{ij} as the difference between the actual and the hypothetical similarity: $g_{ij} = \omega_{ij} - \omega^*_{ij}$. In this way we obtain values g_{ij} that are both positive and negative. The integer linear programming formulation of the Clique Partition is then:

$$\max \sum_{i \neq j} g_{ij} x_{ij} \tag{3.15}$$

subject to

$$\begin{cases} -x_{ij} + x_{ik} + x_{jk} \leq 1, & \forall i < j < k, \quad i, j, k \in V \\ -x_{ik} + x_{jk} + x_{ij} \leq 1, & \forall i < j < k, \quad i, j, k \in V \\ -x_{jk} + x_{ij} + x_{ik} \leq 1, & \forall i < j < k, \quad i, j, k \in V \\ x_{ij} \in \{0, 1\}, & i < j, \quad i, j \in V \end{cases}$$

where x_{ij} is equal to 1 if two nodes are in the same cluster and 0 otherwise.

We experimented very long computational times when we tried to solve it through Integer Linear Programming. Therefore, we implemented a heuristic procedure based on shrinking the vertices of the graph. Shrink is the subroutine by which we take two vertices, representing single units or clusters, and we merge them together to obtain a single cluster. Shrink is described in Algorithm 1. Input is a data structure

$G^h =< V^h, g^h, \pi^h >$, in which V^h is the active node set, each node representing a set of the partition, g^h are the shrunken costs, defined for every pair $i, j \in V^h$, π^h are the clique costs, defined for every active node $i \in V^h$. Output is a data structure $G^q =< V^q, g^q, \pi^q >$ in which $|V^q| = |V^h| - 1$. When we shrink $i, j \in V^q$, we delete j from the active nodes, see Line 1, and the clique profit π_i^h of i increases by the arc profit g_{ij}^h, while all others remain the same, see Lines 2 and 3. In the next steps, the profit of i inherits the profits of j's connections, see Lines 5-7.

Algorithm 1: SHRINK

Input: The data structure: $G^h =< V^h, g^h, \pi^h >$, the pair $i, j \in V^h$
Output: The data structure: $G^q =< V^q, g^q, \pi^q >$
$V^q \leftarrow V^h - j$
$\pi^q \leftarrow \pi^h$
$\pi_i^q \leftarrow \pi_i^h + g_{ij}^h$
$g^q \leftarrow g^h$
for $k \in V^h$ **do**
  $g_{jk}^q \leftarrow 0$
  $g_{ik}^q \leftarrow g_{ik}^q + g_{jk}^h$
return G^q

Subroutine Shrink is used to join nodes or clusters every time we find an improvement of the objective function, that is, when we find a pair (i, j) such that $g_{ij}^h > 0$. The procedure is described in Algorithm 2. At the beginning, Lines 1 and 2, the partition V^q is composed of subsets of one element and the profits π associated to them are null. Then, in the loop 3-9, the greatest profit g_{ij} is selected and, if positive, vertices (i, j) are shrunken. Otherwise, the algorithm stops. The objective function is calculated in Line 10.

We found that Algorithm 2 calculates quickly good quality solution. However, it can be the case that the selected partition is suboptimal. Therefore, we implemented a version of the Neighborhood Search procedure proposed in [183]. The procedure starts with a feasible partition P, in our case the one calculated through Algorithm 2. Then we select at random k vertices of V and try to relocate them to different clusters, searching for an improvement of the objective function. The procedure is repeated several time and for different values of k, until no improvement are found for many consecutive attempts. But in our data, we found that most of the times the results of Algorithm 2 were not improved.

Algorithm 2: Clique Partition

Input: The CP Problem, defined with input V, g.

Output: The Clique Partition V^q, clique costs π^q, objective function f^q

$V^q \leftarrow \{1, \ldots, n\}$

$\pi^q \leftarrow 0$

while $stop = False$ **do**

  $g^q_{ij} \leftarrow \max\{g^q_{kl} | k, l \in V^q\}$

  if $g^q_{ij} > 0$ **then**

    $G^h \leftarrow \text{Shrink}(V^q, i, j)$

    $G^q \leftarrow G^h$

  else

    $stop \leftarrow \text{True}$

$f^q \leftarrow \sum_{i \in V^q} \pi^q_i$

return G^q, f^q

3.3.3 A summary of the Ranking Aggregation/Clique Partitioning procedure

The next pseudo-code (see Algorithm 3) summarizes the methodology that we are proposing:

Algorithm 3: Aggregation and Partition

Calculate rankings $\mathbf{r}^k$, for every centrality measure $k = 1, \ldots, K$

Calculate similarity/dissimilarity ω_{ij} between every countries pairs i, j.

Calculate the gain/cost g_{ij} for all i, j pairs.

Solve the Clique Partition model whose input are g_{ij}'s.

In Step 1, we have K centrality measures, as defined in Subsection 3.3.1. For every measure k, $(k = 1, ..., K)$, we obtain the ranking $\mathbf{r}^k$, whose element r^k_i is the position of country i in the ranking according to the measure k. In Step 2, we calculate values ω_{ij} according to Formula (3.13). In Step 3, we calculate the gains/costs needed to define the Clique Partition model explained in Subsection 3.3.2. Lastly, in Step 4, we apply the Algorithm 2.

3.4 Numerical application

3.4.1 International Trade Network

In this section, we apply the model previously described in order to study the structure of the ITN. We focus on a World Trade dataset, made available by the Observatory of Economic Complexity[1]. In particular, data regard the world trade database developed by the research and expertise centre on the world economy (CEPII) at a high level of product disaggregation. Original data are provided by the United Nations Statistical Division (UN Comtrade) and then the dataset is constructed by CEPII using an original procedure that reconciles the declarations of the exporter and the importer. This harmonization procedure enables to extend considerably the number of countries for which trade data are available, as compared to the original dataset (see [184]).
In particular, we consider the last version published in 2017, based on the Harmonized Commodity Description and Coding System, and that provides aggregated bilateral values of exports for each couple of origin and destination countries. We focus on the aggregated data of the last available year, namely, 2014.
Hence, we construct a directed and weighted network (see Figure 3.1), where each node is a country and weighted links represent the amount of product trades between couple of countries expressed in US dollars. This network is characterized by 220 countries and 26034 links. Its arc density is approximatively 0.54, because on average each country has a large number of trade partners and the entire system is intensely connected. However, the network is far from being complete or, in other words, most countries do not trade with all other countries, but they rather select their partners. Furthermore, world trade tends to be concentrated among a sub-group of countries and a small percentage of the total number of flows accounts for a disproportionally large share of world trade. We have indeed that, on average, each country has trades with more than an half of the other countries in the world, but the top 10 countries export more than 50% of the total flow. To this end, key importers and exporters, classified in terms of strength, are displayed in Figure 3.2. Differences between import and export ranking are remarkable. United States, China, Japan, South Korea and some European countries (namely, France, Germany, Italy, Netherlands and United Kingdom) are world largest importers and exporters. Russia and Canada display instead a top ranking in terms of volume of exports. In particular, Russia is characterized by a significant positive trade balance, equal to approximatively 30% of its total exportations.
Furthermore, as expected, greater countries have more partners and they account for a

[1] See https://atlas.media.mit.edu/en/

generally larger share of world trade. However, the relationship between the economic size and the number of partners is far from perfect, as indicated by the correlation, around 0.5, between the total value of (in or out) flows and the number of partners for each country.

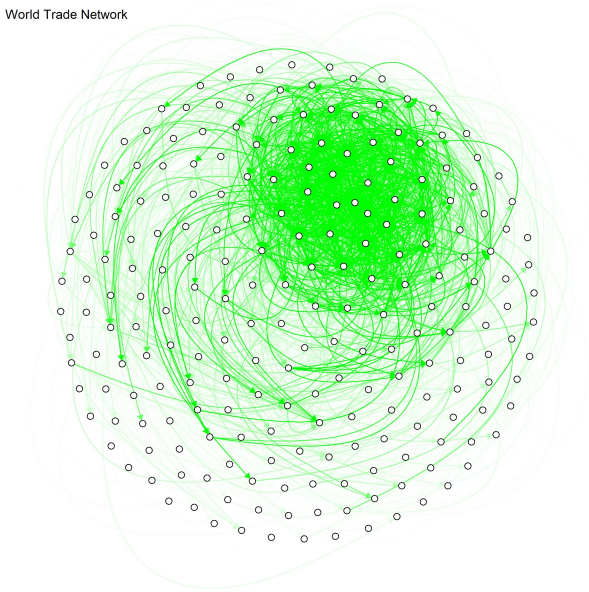


Figure 3.1: World Trade Network of imports and exports at the end of 2014.

3.4.2 Numerical results and discussion

As described in Section 3.3, we aggregate the centrality indexes through a community detection method. As a result, communities are determined by the Clique Partition model, whose input is a weighted network constructed by the original one, in which weights are determined taking into account all the topological indicators in a multi-criteria approach. Four class of network indicators are initially computed by using the network depicted in Figure 3.1. We report in Figure 3.3 the scatter plots of each couple of centrality measures and the Spearman's rank-order correlation, in order to assess the strength and the direction of association between different ranked indicators. All the correlation are positive, because a country with a high volume of exports is also highly interconnected in the network. However, there are not fully correlated couples and, in many cases, the correlation is far from one. It is also noteworthy the strong dependence between in and out versions of the same indicator. This is mainly explained by the similar patterns of imports and exports for several countries. Only hubs and

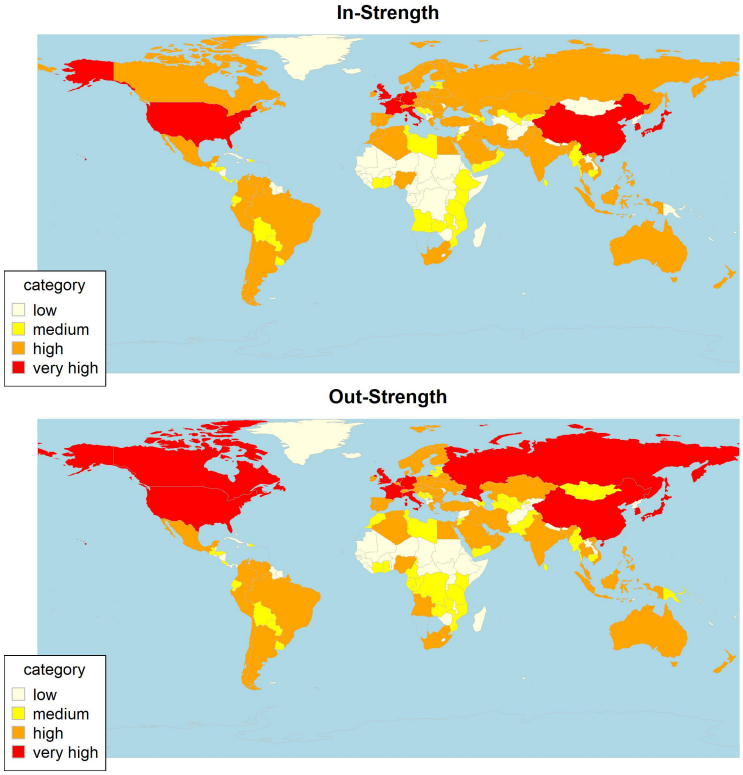


Figure 3.2: In and out-strength of countries in world trade network. Categories are based on the following classes $[0-q_{50}]$, $(q_{50}-q_{75}]$,$(q_{75}-q_{95}]$,$(q_{95}-q_{100}]$ where q_p is the p-quantile of the in-strength and out-strength distribution, respectively.

authorities seem to emphasize the presence of specific exceptions. Table 3.1 reports the top ten countries according to the rankings of the four used indicators. The rankings reflect the results about the correlations and they exemplify the differences in the role of each country as importer or exporter.

Laplacian In	Laplacian Out	In-Strength	Out-Strength	In-Clustering	Out-Clustering	Hubs	Authority
FRA	THA	USA	CHN	USA	CHN	CHN	USA
SGP	BLX	CHN	USA	CHN	DEU	CAN	HKG
CZE	NLD	DEU	DEU	DEU	USA	MEX	JPN
USA	FRA	JPN	JPN	ARE	JPN	DEU	CHN
GBR	GBR	GBR	KOR	GBR	SAU	JPN	DEU
POL	DEU	FRA	FRA	JPN	RUS	USA	GBR
BLX	USA	NLD	NLD	SAU	FRA	KOR	KOR
NLD	SGP	HKG	ITA	NLD	ITA	FRA	FRA
THA	ITA	KOR	GBR	ITA	KOR	GBR	CAN
CAN	CAN	ITA	RUS	FRA	GBR	ITA	MEX

Table 3.1: The top ten countries for each network indicator.

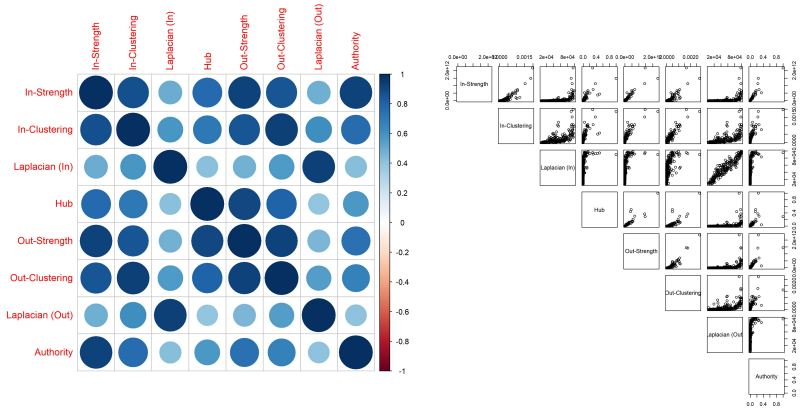


Figure 3.3: On the left-hand side, spearman correlation between each couple of measures. On the right-hand side, matrix of scatter plots between different indicators.

By applying the methodology[1] described in Section 3.3, we obtain at the first step three communities, characterized by 69, 87 and 64 countries, respectively. We display in Figures 3.4 the communities initially identified by the algorithm. These three clusters are also well separated in terms of countries' centrality. We have indeed that countries belonging to community 1 have an average ranking of 38, the second community has an average ranking of 113, while countries that belong to the lowest community have an average ranking around 185. In other words, the most central countries are all included in the top community. We also notice that the three clusters are characterized by a very different intra-group density. We have indeed that the density of the subgraphs (of the original ITN) induced by the countries belonging to the three clusters is 0.97, 0.53, 0.05, respectively. This behaviour can be partially explained by the fact that central countries tend to concentrate a high number of transactions between them.

Since in several contexts this initial division could be too raw, we can refine the procedure in order to reduce the heterogeneity in each group. To this end, at the subsequent step, we separately consider the ranking of centralities of countries, applying the proposed method for community detection to the single group. Specifically, at step 2 we apply the proposed algorithm within each community detected at the previous step. In other words, at this step the algorithm takes into account how a specific country is ranked with respect to other countries of the same subgroup on the basis of

[1] In the application we set $p = 2$ for the computation of the Minkoski distance. Similar results have been obtained by using other values of p.

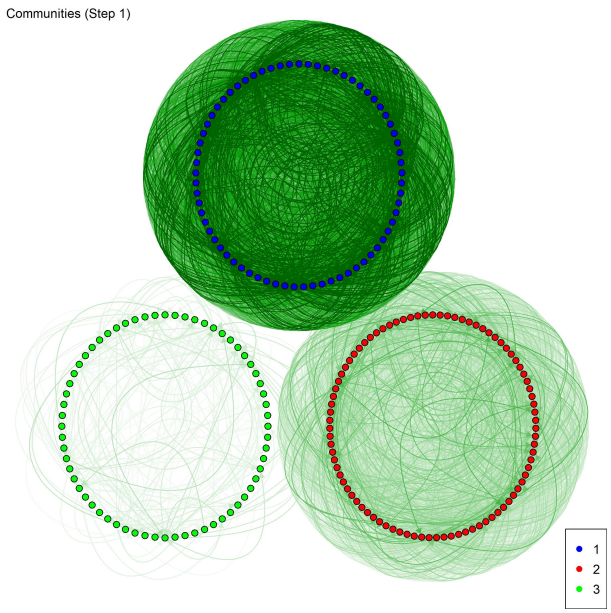


Figure 3.4: Clusters of countries identified at the first step by the community detection algorithm. The communities are ordered in terms of average ranking.

the centrality indicators computed on the whole network. The ranking position of each country may change, but the global ranking remains the original one. For instance, the community 1, characterized by 69 countries, splits into two groups of 32 and 37 countries, respectively. The two groups obtained have an average ranking of 19 and 55. The procedure is repeated in a similar way also for the other two communities identified at the step 1, resulting in 8 communities at step 2 (see dendrogram in Figure 3.5 and top left-hand side in Figure 3.6).

Further reductions of the heterogeneity in each cluster are possible of course, repeating again this process at the next steps and, in general, a stopping criterion is needed. A possible one consists in looking at the volatility of the ranking inside each cluster. If we focus on community with larger standard deviation, we tend to produce a more refined breakdown between low-ranking countries. Vice versa, looking at a measure of relative volatility (as the coefficient of variation (CV)), we deal with a higher decomposition of top-ranking clusters. Here we follow this second approach and, at each step, we further divide a community only if the CV of countries' average rankings is lower than 7.5%.

The complete structure representing the various division steps is represented by the dendrogram in Figure 3.5. We notice that the number of communities increases at each step, leading to 22 communities at step 4. As expected, the criterion based on CV leads

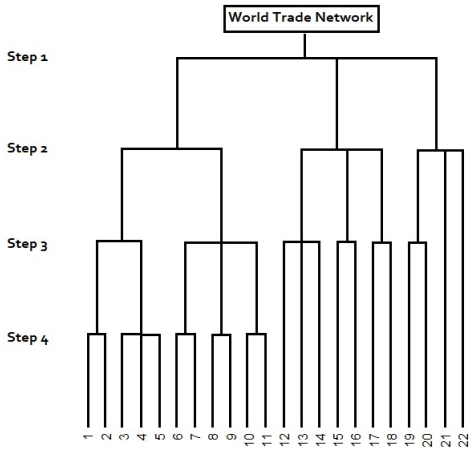


Figure 3.5: Dendrogram that illustrates the arrangement of clusters by applying the algorithm at four different levels. Communities are ordered in terms of average ranking.

to a more granular breakdown for clusters characterized by a higher average ranking. In this way, we are able to classify key countries in different clusters. In Figure 3.6 we report the subnetworks induced by the clusters. The analysis confirms a tendency of top communities in showing a higher intra-group density. For instance, the top community at step 3 and the three higher ranking communities at step 4 are complete, that is all central countries trade each other. However, there is not a monotonic behaviour between ranking and intra-density. For instance, at step 2 community 4 has a higher average ranking than community 5 (124 against 128), but a significant lower intra density (0.05 against 0.58). This peculiar behaviour can be justified by the composition of the groups[1]. Indeed, we are grouping countries on the basis of similarity in terms of their central role in the network instead of using preferential economic relationships.

[1]Community 5 at step 2 is indeed characterized by various groups of countries that trade each other. For instance, in this group, we have several countries, originated after the breakup of Jugoslavia and Russia.

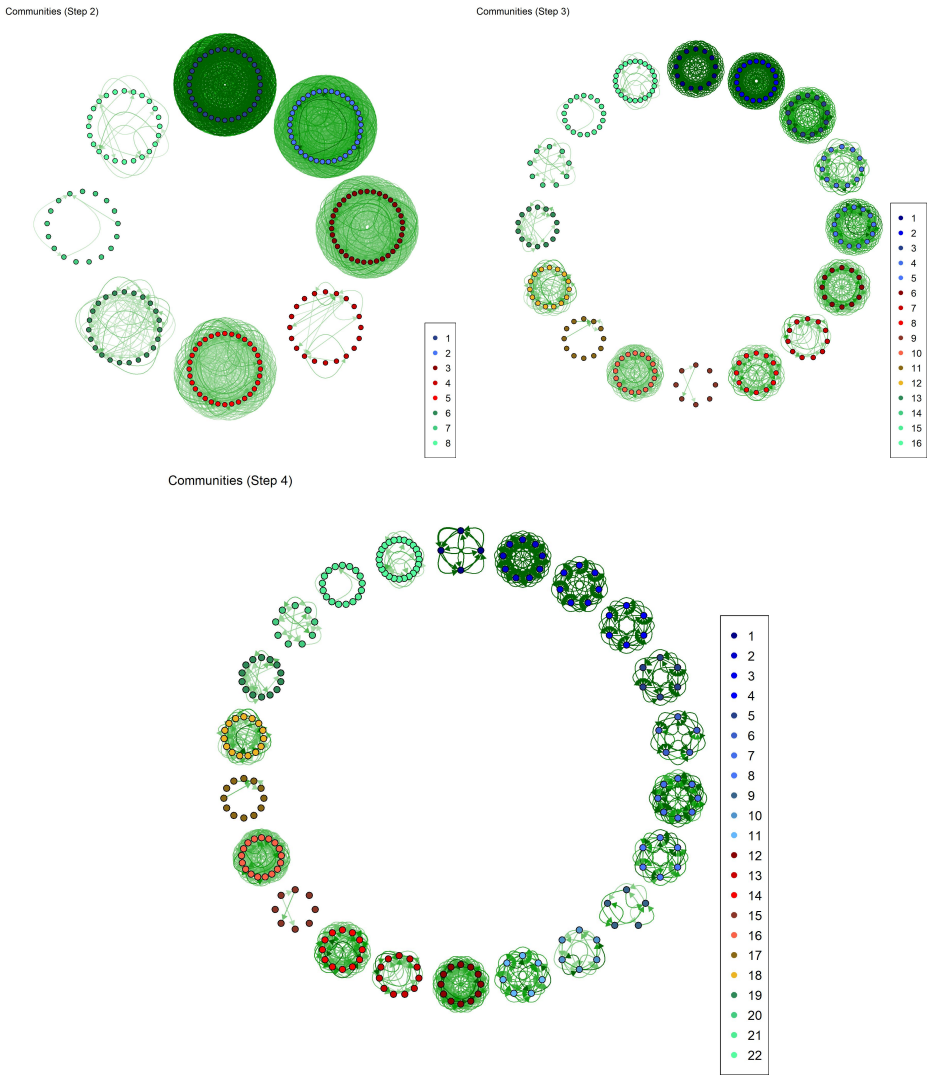


Figure 3.6: Clusters of countries identified at the second, third and fourth step, respectively, by the community detection algorithm. The communities are ordered in terms of average ranking.

It is worth to compare our results with a well-known country-classification method based on the Economic Complexity Index (ECI). This index, introduced by Hidalgo and Hausmann [138], allows to rank countries in the ITN according to the diversification of their export flows, which reflects the amount of knowledge that drives their growth. The higher is the ECI, the more advanced and diversified is an economy. In particular, countries whose economic complexity is greater than expected (on the basis of their global income), tend to grow faster than rich countries with a low ECI. In this perspective, ECI represents a suitable tool for comparing countries in the ITN independently of their total output and it provides an independent measure of similarity. For instance, in Table 3.2, we list the values of the ECI for the countries in the top four clusters detected. As shown in Table 3.3, the mean value of such an index for each cluster is positively correlated with their ranking in the final partition we found at step 4. However, some exceptions are noticeable. For instance, China, in cluster 1, is characterised by a lower ECI than some countries in cluster 2 (e.g. UK and Italy) because of a lower diversification of exported commodities. Indeed, its wealth comes from a more homogeneous set of assets than UK and Italy, which can express a wider diversification in their total output. This could explain why the Standard Deviation inside each one of our communities is significantly high.

Now, we focus on the countries' role within the network. As shown in Figure 3.7, the initial breakdown in communities gives a general feeling of the relevance of different macro-regions in the whole trade network. We have indeed that the top cluster, characterized by 69 countries at step 1, includes all the most developed European countries[1], largest economies in Asia and Middle East, several countries in South America, Canada, Mexico, USA, Australia and New Zealand. Furthermore, Algeria, Angola, Egypt, Morocco, Nigeria and South Africa are included for the African continent. Except for some small countries, this community includes all the advanced economies identified in the World Economic Outlook (WEO) by the International Monetary Fund (IMF)[2] and the emerging economies identified by IMF and by other analysts[3].
At the end of the procedure, we obtain that the most central group is composed by

[1] 28 European Countries are included in community 1. Gibraltar, San Marino and Andorra and some countries originated after the breakup of Jugoslavia and Russia are not included.

[2] List of advanced countries according to WEO are available at:
https://www.imf.org/external/pubs/ft/weo/2019/01/weodata/groups.htm#ea

[3] Various sources list countries as "emerging economies" exist. A few countries appear in every list (BRICS, Mexico, Turkey). While there are no commonly agreed upon parameters on which the countries can be classified as "Emerging Economies", several firms have developed detailed methodologies to identify the top performing emerging economies every year.

China, Germany, Japan and United States. Higher volumes of trades are indeed moved by these countries (e.g., see ranking of in and out-strength in Table 3.1) and, at the same time, they also show the highest levels of interconnections.
In the second group, we have countries which either are positioned at a slightly lower level (as GBR, FRA, ITA and NLD) or are outstanding for one specific indicator, but, on average, they show a less relevant role in the network. For instance, Canada has the second position in terms of hubs centrality (see Table 3.1), but shows an average ranking around 14, because of a lower clustering. This is in line with its low value of the ECI.

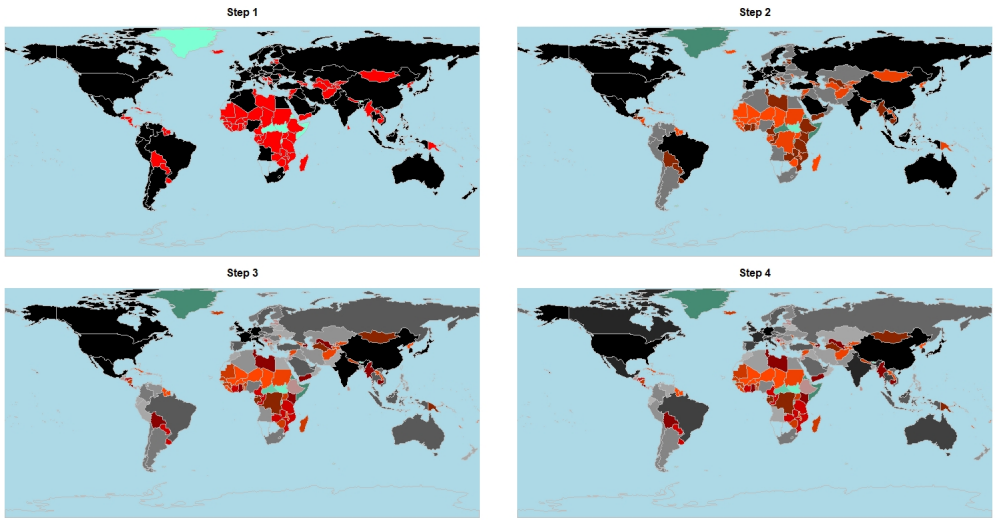


Figure 3.7: Structure of communities at different steps. Darker colours are associated to communities with an higher average ranking. The number of communities is respectively equal to 3, 8, 16, 22.

Country	Step 1	Step 2	Step 3	Step 4	ECI
CHN	1	1	1	1	1.16379
DEU	1	1	1	1	1.81367
JPN	1	1	1	1	2.31842
USA	1	1	1	1	1.30167
BLX	1	1	1	2	0.90581
CAN	1	1	1	2	0.411362
FRA	1	1	1	2	1.15748
GBR	1	1	1	2	1.40296
IND	1	1	1	2	-0.014696
ITA	1	1	1	2	1.24155
KOR	1	1	1	2	1.90646
MEX	1	1	1	2	0.953003
NLD	1	1	1	2	0.756212
AUS	1	1	2	3	-0.846322
BRA	1	1	2	3	-0.151225
CHE	1	1	2	3	1.99456
ESP	1	1	2	3	0.701443
MYS	1	1	2	3	0.828817
SGP	1	1	2	3	1.71171
THA	1	1	2	3	0.955651
AUT	1	1	2	4	1.64981
CZE	1	1	2	4	1.52129
IDN	1	1	2	4	-0.014696
POL	1	1	2	4	0.839266
SWE	1	1	2	4	1.6459
TUR	1	1	2	4	0.378481
ARE	1	1	2	4	-0.502072
HKG	1	1	2	4	1.35236
RUS	1	1	2	4	0.008439
SAU	1	1	2	4	-0.369927
VNM	1	1	2	4	-0.129961
XXB	1	1	2	4	NA

Table 3.2: Composition of top four clusters (in terms of average ranking) derived at step 4. Last column displays the ECI for each country.

Community	Mean ECI	SD ECI
1	1.6493875	0.526404666
2	0.968904556	0.559587598
3	0.742090571	0.990344256
4	0.579899091	0.844314087

Table 3.3: Mean and standard deviation of ECI inside each of the four top clusters

3.5 Conclusions

Community detection is a widely discussed topic in network theory. The analysis of the mesoscale structure of a real network throws light on its inner structure. This plays an even more significant role when applied to ITN, in view of its multiple implications. This work aimed at clustering countries according to similarities in their role in the global market, rather than using only the preferential channels of exchange between them. Centrality measures have represented, by now, a classical tool to rank such a role in the network. In particular, each centrality measure expresses a different information about the nodes position. We proposed a way to collect all the informa-

tion content, represented by suitable centrality measures, through a distance measure between countries.

Among all possible similarity-dissimilarity distances, the Minkowski distance allows to grasp different data distributions, depending on a specific parameter p. In this way, we constructed a weighted complete network where nodes are countries and weighted links are related to similarities between them. By means of this similarity-network, we set up a classical Clique Partitioning problem to identify the community structure that maximizes the modularity. We proposed here a new algorithm which, loosely speaking, merges different nodes or clusters and shrinks the network in such a way to get polynomial times for its solution.

When applied to the ITN in the year 2014, the optimal solution shows three big clusters, more or less equivalent in size but very different in terms of intra-cluster density. This has been easily interpreted since the rate of exchanges between top countries is far more intense than for poor ones. We iterated the same methodology to each cluster, in order to reduce the internal heterogeneity. This allows to build a dendrogram tree stemming at each step.

The top leader economies in the world result to be those of China, Japan, USA and Germany. This is not unexpected but our proposal shows that these countries also play a very similar role in the world economy on the basis of the set of selected indicators, making our approach suitable for other network applications.

Chapter 4

Community Detection in Multilayer Networks

4.1 Introduction

In this last chapter we want to address the fact that networks are often more complex objects than what may have emerged, for instance, from the analyses contained in the previous chapters. A network is a set of interconnected nodes but the nature of the connections between nodes in the same set may be of different kind. According to the nature of the connections, different networks are generated on the same set of nodes. Each one can be interpreted as a level in a more complex object, called *multilevel network*. Different levels highlight a different nature in the interaction between nodes and each one is called *monoplex*. When we focus on a single level, it is as if we are observing a detail through a magnifying glass, enlightening a particular aspect of the possible relationships between nodes. However, it is clear that a deeper insight into the complexity of these relationships plays an important role for a global description of the network. Multilevel networks are precisely the effective tool to catch such a complexity.

Generally speaking, in a multilevel network, we can move from one node to another on the same level following the links on that level. But we can also move from one level to another one. The easiest way to do that is to imagine a jump from a node on a level to the same node on a different level. When this is the only possibility to switch levels, we call the multilevel network a *multiplex*. But nothing prevents us from conjecturing a jump from a node on one level to a completely different node on a second level. When this is allowed the multilevel network is called *multilayer network*. This is the most general case and it is the one we are interested in. [185, 186, 187, 188].

For instance, cities and roads connecting them, on one side, and the same cities and railways connecting them, on the another side, represent two possible monoplexes. When a traveller gets in a city by train, he can rent a car to go to another city. That is, he is switching from the first monoplex (roads) to the second one (railways). Together they constitute a multiplex. Of course, he cannot move from a city to a different one without using some form of transportation; so this is not a multilayer network. On the contrary, the industrial chemical sector in different countries is a monoplex, as it is the sector of the pharmaceutical industry in different countries, but the former can supply the latter with materials and chemicals within the same country or to foreign nations in the same way. This is an example of multilayer network.

This chapter will be organized as follows. We start by recalling some mathematical and notational issues. In particular, we adopt and extend the tensorial approach in [16] to give a description as compact as possible to all the quantities involved. The central purpose is then the extension to multilayer networks of the community detection methods described in the previous chapters, with particular reference to three different approaches that rely respectively on modularity, on communicability and on a specific metric partition quality index. Despite requiring further developments, the proposed methodologies prove to work well when applied to some toy-model networks.

Finally, since the problem of community detection is strictly linked to how much the network is clustered or not, at the end of the chapter we propose an extension of all the best known clustering coefficients in literature to multilayer networks. In particular, it will be shown how, in the tensorial language, it is possible to give them a substantially unified writing and how the choice of a single reference tensor allows to switch from one to the other. Also in this case, these coefficients have been tested only on some toy-model networks, reserving, as we will say at the end, to extend their application to real networks in future works.

4.2 Preliminaries

4.2.1 Notations

In order to clarify the tensorial notations that will be used in the following, we list here the symbols we will adopt for the indices of tensors, matrices and vectors. We will consider a multilayer network with N nodes on each level and L levels. Links can connect different nodes within the same level, the same node on different levels, or different nodes on different levels. In general, we will use Latin letters for objects, nodes or levels, and Greek letters for components of tensors, matrices and vectors.

Specifically:

- $i, j, k, l, \ldots$ for the names of the nodes
- $a, b, c, d, \ldots$ for the names of the levels
- $\mu, \nu, \rho, \sigma, \ldots$ for components of quantities associated with nodes
- $\alpha, \beta, \gamma, \delta, \ldots$ for components of quantities associated with levels

For instance, $v^\mu(i)$ and $v_\nu(i)$ represent respectively the contravariant and covariant μ- and ν-component of a vector v related to node i (see, e.g, [189]); or $W^\mu_\nu(a, b)$ represents the $\mu\nu$ - entry of the matrix W related to levels a and b. Throughout this chapter we will adopt Einstein summation convection over repeated indices: when an index variable appears twice in a single term, it implies summation of that term over all the values of the index. For example: $c_\mu x^\mu$ means $\sum_{\mu=1}^{n} c_\mu x^\mu$.

Let now $e^\mu(i)$ be the canonical basis in $\mathbb{R}^N$, that is

$$e^\mu(i) = \begin{cases} 1 & \text{if } \mu = i \\ 0 & \text{if } \mu \neq i \end{cases} \tag{4.1}$$

We will use the following notations for the canonical tensors:

- $E^\mu_\nu(i, j) = e^\mu(i)e_\nu(j)$ the canonical second order tensor basis in $\mathbb{R}^{N\times N}$. This tensor is represented by a N-square matrix whose (i, j)-entry is 1, and the other entries are 0.
- $E^\alpha_\beta(a, b) = e^\alpha(a)e_\beta(b)$ the canonical second order tensor basis in $\mathbb{R}^{L\times L}$. This tensor is represented by a L-square matrix whose (a, b)-entry is 1, and the other entries are 0.
- $E^{\mu\alpha}(i, a) = e^\mu(i)e^\alpha(a)$ the canonical second-order tensor basis in $\mathbb{R}^{N\times L}$. This tensor is represented by a $N \times L$ matrix whose (i, a)-entry is 1, and the other entries are 0.
- $E^{\mu\alpha}_{\nu\beta}(i, j; a, b) = e^\mu(i)e_\nu(j)e^\alpha(a)e_\beta(b)$ the canonical fourth-order tensor basis in $\mathbb{R}^{N\times N\times L\times L}$, whose (i, j, a, b)-entry is 1, and the other entries are 0.

4.2.2 Adjacency tensor

To describe adjacency relations between nodes in a multilayer network, we need a forth-order adjacency tensor. Let us focus on two nodes on two levels: node μ on level α and node ν on level β. The adjacency tensor is then defined as

$$M_{\nu\beta}^{\mu\alpha} = \sum_{a,b=1}^{L} W_{\nu}^{\mu}(a,b) E_{\beta}^{\alpha}(a,b) \tag{4.2}$$

where $W_{\nu}^{\mu}(a,b)$ is the adjacency matrix between levels a and b, that is a matrix whose entries are the weights of the links between nodes on level a and nodes on level b. This matrix can be expressed as

$$\begin{aligned} W_{\nu}^{\mu}(a,b) &= \sum_{i,j=1}^{N} w_{ij}(a,b) e^{\mu}(i) e_{\nu}(j) \\ &= \sum_{i,j=1}^{N} w_{ij}(a,b) E_{\nu}^{\mu}(i,j) \quad \text{for} \quad a,b = 1,\dots,L \end{aligned} \tag{4.3}$$

so that the adjacency tensor in 4.2 can be written equivalently as

$$M_{\nu\beta}^{\mu\alpha} = \sum_{a,b=1}^{L} \sum_{i,j=1}^{N} w_{ij}(a,b) E_{\nu\beta}^{\mu\alpha}(i,j;a,b) \tag{4.4}$$

In particular, the element $w_{ij}(a,b)$ represents the intensity of the relationship between node i in level a and node j in level b, i.e. it is the scalar weight of the link between the node (i,a) and the node (j,b) and, as before, $E_{\nu\beta}^{\mu\alpha}(i,j;a,b) = e^{\mu}(i) e_{\nu}(j) e^{\alpha}(a) e_{\beta}(b)$ is the fourth-order tensor canonical basis in $\mathbb{R}^{N\times N\times L\times L}$.

We will set $W_{\nu}^{\mu}(a) := W_{\nu}^{\mu}(a,a)$ and $w_{ij}(a) := w_{ij}(a,a)$ within a single level, that is for the usual adjacency matrix of a single layer. Moreover, when $W_{\nu}^{\mu}(a,b)$ is diagonal, i.e. when $w_{ij}(a,b) \neq 0$ if and only if $i = j$, each node is connected only with its counterparts in different levels and the multilayer network is a multiplex. We exclude the possibility of self-loops, that is $w_{ii}(a) = 0,\ \forall i = 1,\dots,N$ and $\forall a = 1,\dots,L$.

4.2.2.1 Contractions and Projected Networks

Contraction operation allows to reduce the dimensions of a given tensor by adding over repeated indices and it is useful to compute, for example, some elementary quantities like the number of nodes or the number of edges in the network.

Let us define, throughout this chapter, $u^\mu = (1, \ldots, 1)^T$ and $u_\nu = (1, \ldots, 1)$ the contravariant and covariant 1st order 1-tensor (all 1's vectors) and $U^\mu_\nu = u_\nu u^\mu$ the 2nd order 1-tensor (all 1's matrix).

The number of nodes per level is given by $\delta^\mu_\mu = N$, where δ^μ_ν is the Kronecker delta tensor. The number of edges between level a and level b is given by $W^\mu_\nu(a,b)U^\nu_\mu$ and the number of edges on level a by $W^\mu_\nu(a)U^\nu_\mu$.

In order to extract a single level adjacency matrix, we need to project the tensor $M^{\mu\alpha}_{\nu\beta}$ onto the canonical tensor $E^\alpha_\beta(c,c)$, where c is the level of interest:

$$M^{\mu\alpha}_{\nu\beta}E^\beta_\alpha(c,c) = \sum_{a,b=1}^{L} W^\mu_\nu(a,b)E^\alpha_\beta(a,b)E^\beta_\alpha(c,c) = W^\mu_\nu(c,c) = W^\mu_\nu(c)$$

The levels in a multilayer network can be collapsed into a single layer network in two different ways.

The *monoplex projected network* is the monoplex network whose adjacency matrix is given by

$$P^\mu_\nu = M^{\mu\alpha}_{\nu\beta}U^\beta_\alpha = \sum_{a,b=1}^{L} W^\mu_\nu(a,b)E^\alpha_\beta(a,b)U^\beta_\alpha = \sum_{a,b=1}^{L} W^\mu_\nu(a,b) \tag{4.5}$$

Note that $E^\alpha_\beta(a,b)U^\beta_\alpha = 1 \in \mathbb{R}$. The sum in equation 4.5 is extended over all the possible couples of levels, i.e. including both couples made up of the same level and couples made up of different levels. This means that both intralayer and interlayer links are included. Specifically, links between homologous nodes in different levels collapse into loops in the monoplex projected network.

The *overlay network* is the monoplex network whose adjacency matrix is given by

$$O^\mu_\nu = M^{\mu\alpha}_{\nu\alpha} = \sum_{a,b=1}^{L} W^\mu_\nu(a,b)E^\alpha_\alpha(a,b) = \sum_{a=1}^{L} W^\mu_\nu(a,a) \tag{4.6}$$

Note that in $E^\alpha_\alpha(a,b) = \sum_{\alpha=1}^{L} E^\alpha_\alpha(a,b)$ we sum along the diagonal of $E^\alpha_\beta(a,b)$, and this sum is 1 if and only if $a = b$, i.e. $\sum_{\alpha=1}^{L} E^\alpha_\alpha(a,b) = \delta^a_b$. In the overlay network interlayer links are excluded so that they do not collapse into loops in the final projection.

Finally, we can deal with the *network of levels*, whose weighted adjacency tensor is given by

$$\Psi^\beta_\alpha = M^{\mu\alpha}_{\nu\beta}U^\nu_\mu \tag{4.7}$$

In this network an entire level collapses into a single node and links only survive if the levels are immediate neighbours.

4.2.2.2 Centrality measures: degree and strength

The multidegree centrality vector for a multilevel undirected binary network is the $N-$vector whose components are the total degree of each node in all levels:

$$k^{\mu} = M^{\mu\alpha}_{\nu\beta} U^{\beta}_{\alpha} u^{\nu} \tag{4.8}$$

Indeed,

$$\begin{aligned}
M^{\mu\alpha}_{\nu\beta} U^{\beta}_{\alpha} u^{\nu} &= \sum_{a,b=1}^{L} W^{\mu}_{\nu}(a,b) E^{\alpha}_{\beta}(a,b) U^{\beta}_{\alpha} u^{\nu} \\
&= \sum_{a,b=1}^{L} \sum_{i,j=1}^{N} w_{ij}(a,b) E^{\mu}_{\nu}(i,j) E^{\alpha}_{\beta}(a,b) U^{\beta}_{\alpha} u^{\nu} \\
&= \sum_{a,b=1}^{L} \sum_{i,j=1}^{N} w_{ij}(a,b) E^{\mu}_{\nu}(i,j) u^{\nu} \\
&= \sum_{a,b=1}^{L} \sum_{i,j=1}^{N} w_{ij}(a,b) e^{\mu}(i) \\
&= \sum_{a,b=1}^{L} \sum_{i=1}^{N} \left(\sum_{j=1}^{N} w_{ij}(a,b) \right) e^{\mu}(i) \\
&= \sum_{a,b=1}^{L} k^{\mu}(a,b) = k^{\mu}
\end{aligned}$$

where $\sum_{j=1}^{N} w_{ij}(a,b)$ represents the degree of a node $i \in a$ obtained by counting only links from i whose second end is node j that lies on level b. By multiplying by $e^{\mu}(i)$ we assign this degree to a component μ of a given vector. The degree of a specific node i is then given by $k(i) = k^{\nu} e_{\nu}(i)$.

We can also define a degree centrality matrix for multilevel undirected binary network the $N \times L$ matrix with the degree of each node in each level:

$$K^{\mu\alpha} = M^{\mu\alpha}_{\nu\beta} u^{\beta} u^{\nu} \tag{4.9}$$

Similarly, we can define a multistrength vector and strength centrality matrix for multilevel undirected weighted network as

$$s^{\mu} = M^{\mu\alpha}_{\nu\beta} U^{\beta}_{\alpha} u^{\nu} \tag{4.10}$$

and

$$S^{\mu\alpha} = M^{\mu\alpha}_{\nu\beta} u^{\beta} u^{\nu} \tag{4.11}$$

It is straightforward to extend the previous definitions to a directed network. If we refer to a weighted network we may have an out-strength, an in-strength and a total strength. All of them can be given for the whole multilayer network or for a single level. Finally, a strength related to the bilateral links only can be provided too. All of them may be computed by means of the formulae synthetically listed below, where we denoted by $A^{\mu\alpha}_{\nu\beta}$ the binary version of the weighted adjacency tensor:
Multi-level out-strength s_{out} contravariant vector:

$$s^{\mu}_{\text{out}} = M^{\mu\alpha}_{\nu\beta} U^{\beta}_{\alpha} u^{\nu} = P^{\mu}_{\nu} u^{\nu} \tag{4.12}$$

Multi-level in-strength s^{in} covariant vector:

$$s^{\text{in}}_{\nu} = u_{\mu} M^{\mu\alpha}_{\nu\beta} U^{\beta}_{\alpha} = u_{\mu} P^{\mu}_{\nu} \tag{4.13}$$

Multi-level total strength:

$$s = s_{\text{out}} + s^{\text{in}} \tag{4.14}$$

Multi-level strength of bilateral arc on node i:

$$s_{\text{bil}}(i) = M^{\mu\alpha}_{\rho\beta} A^{\rho\beta}_{\nu\alpha} e_{\mu}(i) e^{\nu}(i) \tag{4.15}$$

Single-level out-strength matrix:

$$S^{\mu\alpha}_{\text{out}} = M^{\mu\alpha}_{\nu\beta} u^{\beta} u^{\nu} \tag{4.16}$$

Single-level in-strength matrix:

$$S^{\text{in}}_{\nu\beta} = u_{\alpha} u_{\mu} M^{\mu\alpha}_{\nu\beta} \tag{4.17}$$

Single-level total strength matrix:

$$S = S_{\text{out}} + S^{\text{in}} \tag{4.18}$$

Single-level strength of bilateral arc on node (i, a) matrix:

$$S_{\text{bil}}(i, a) = M^{\mu\alpha}_{\rho\gamma} A^{\rho\gamma}_{\nu\beta} E_{\mu\alpha}(i, a) E^{\nu\beta}(i, a) \tag{4.19}$$

4.3 Community detection on multilayer networks

A large part of chapters 2 and 3 of this thesis is devoted to the problem of community detection on monoplex networks. In particular in chapter 2 we proposed a new methodology based on the idea of communicability metric to identify clusters of strongly interacting nodes. The purpose of the present section is to lay the foundations for a community detection on multilayer networks providing some extensions of the discussed methodologies.

In particular, we aim to test and compare here three different approaches to community detection on multilayer networks, based respectively on an extension to these networks of a) the classical Newman modularity and the widely used Girvan-Newman approach; b) the methodology based on the Estrada communicability graph; c) the methodology we proposed in chapter 2 based on communicability metrics.

Following [16], each one of the proposed methodologies will be formulated in terms of adjacency tensor. For the purpose of an easier visualization, hereafter we refer to the multilayer binary and weighted networks in figures 4.1 and 4.2, respectively, with $N = 4$ and $L = 3$ and we will use these sample networks to test our methodologies.

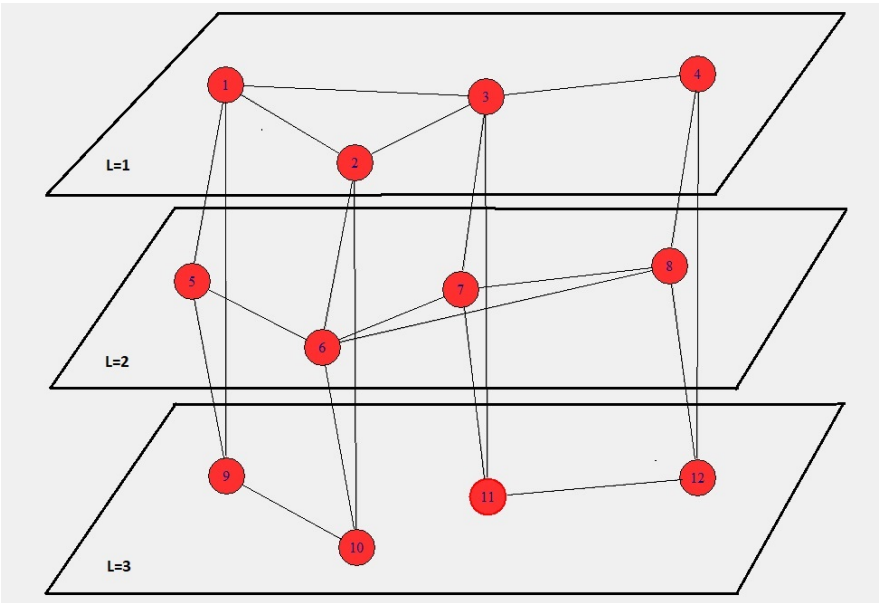


Figure 4.1: Binary Multilayer Network

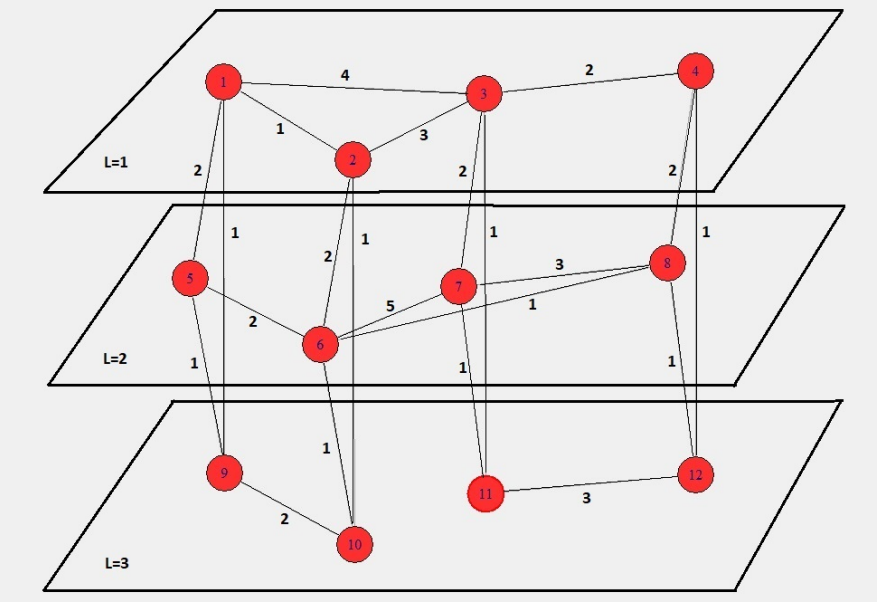


Figure 4.2: Weighted Multilayer Network. The number on each link represents its weight.

4.3.1 Community detection on multilayer networks based on Modularity

Here we refer to the classic community detection based on modularity (see, for instance, [190]). In particular we define the multilevel version of modularity and propose a method to find the optimal partition of a multilayer network, the one for which modularity is maximum.

4.3.1.1 Undirected Networks

Let $W^{\mu\alpha}_{\nu\beta}$ be the actual adjacency tensor and $M^{\mu\alpha}_{\nu\beta}$ a null-model tensor that encodes the random connections against which we compare the actual connections of the multilayer network. Let $\mathscr{P} = \{\mathscr{P}_c\}$, $c = 1, \ldots, C$ be a given partition into C clusters of the multi-layer network. Let us define $Q^{\mu\alpha}_c \in \mathbb{R}^{N\times L\times C}$ such that

$$Q^{\mu\alpha}_c = \begin{cases} 1 & \text{if } (\mu,\alpha) \in \mathscr{P}_c \\ 0 & \text{if } (\mu,\alpha) \notin \mathscr{P}_c \end{cases} \tag{4.20}$$

equal to 1 if node μ on level α belongs to cluster $\mathscr{P}_c$, 0 otherwise. The modularity of partition $\mathscr{P}$ is then given by the scalar

$$\mathcal{Q} = \frac{1}{\mathcal{S}} Q^c_{\mu\alpha} B^{\mu\alpha}_{\nu\beta} Q^{\nu\beta}_c \tag{4.21}$$

where $B^{\mu\alpha}_{\nu\beta} = W^{\mu\alpha}_{\nu\beta} - M^{\mu\alpha}_{\nu\beta}$ and $\mathcal{S}$ is the total strength of the network, that is $\mathcal{S} = \sum_{\mu=1}^{N} s^{\mu} = W^{\mu\alpha}_{\nu\beta} U^{\nu\beta}_{\mu\alpha}$ where s^{μ} is the multistrength centrality vector defined in formula 4.10. Alternative choices of the null model can be made, that is many different null model tensors $M^{\mu\alpha}_{\nu\beta}$ can be used. Let us choose

$$B^{\mu\alpha}_{\nu\beta} = W^{\mu\alpha}_{\nu\beta} - \frac{S^{\mu\alpha} S_{\nu\beta}}{\mathcal{S}} \tag{4.22}$$

where $S^{\mu\alpha} = W^{\mu\alpha}_{\nu\beta} u^{\beta} u^{\nu}$ as in 4.11.

4.3.1.2 Directed Networks

Let us outline how we can write modularity in the directed case. Let's define

$$B^{\mu\alpha}_{\nu\beta} = W^{\mu\alpha}_{\nu\beta} - \frac{2S^{\mu\alpha} S_{\nu\beta}}{\mathcal{S}} \tag{4.23}$$

where $\mathcal{S} = s_{\text{out}} + s_{\text{in}} = \sum_{\mu=1}^{N} s^{\mu} + \sum_{\nu=1}^{N} s_{\nu}$; now $s^{\mu} = M^{\mu\alpha}_{\nu\beta} U^{\beta}_{\alpha} u^{\nu}$ is the multilevel out-strength (contravariant vector, s_{out}) defined in 4.12 and $s_{\nu} = u_{\mu} M^{\mu\alpha}_{\nu\beta} U^{\beta}_{\alpha}$ is the multilevel in-strength (covariant vector, s_{in}) defined in 4.13.

Let us notice that $S^{\mu\alpha} S_{\nu\beta}$ represents automatically the tensor product between in-strength and out-strength matrices. Now, let B^T be the transpose of tensor B according to rule $\left[B^{\mu\alpha}_{\nu\beta}\right]^T = B^{\nu\beta}_{\mu\alpha}$ which transposes both nodes and levels and let us set $\hat{B} = B + B^T$. Then the modularity of a given partition $\mathscr{P} = \{\mathscr{P}_c\},\ c = 1, \ldots, C$ is given by

$$\mathcal{Q} = \frac{1}{\mathcal{S}} Q^{c}_{\mu\alpha} \hat{B}^{\mu\alpha}_{\nu\beta} Q^{\nu\beta}_{c} \tag{4.24}$$

4.3.1.3 Application

When we apply the maximum modularity approach to detect optimal communities to the two toy networks in figures 4.1 and 4.2, we obtain the partitions represented in figures 4.3 and 4.4 for the binary and weighted version respectively. For a straightforward visualization, we have simply grouped nodes belonging to the same community by means of a closed blue line. As can be seen from figure 4.3 in the binary case, the network is partitioned into two transversal communities due to the presence of strong complete triangles between homologous nodes on different levels. When we add weights, as in figure 4.4, their role becomes dominant and, for example, the whole first level acts as a community in itself. We will add further comments later when comparing the different methodologies.

Binary case:

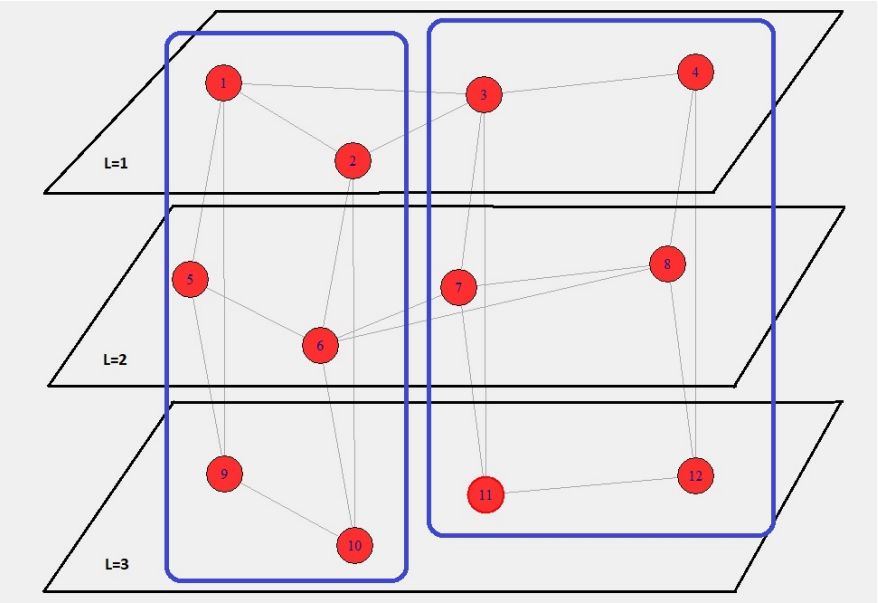


Figure 4.3: Binary Multilayer Network Communities obtained by modularity approach.

Weighted case:

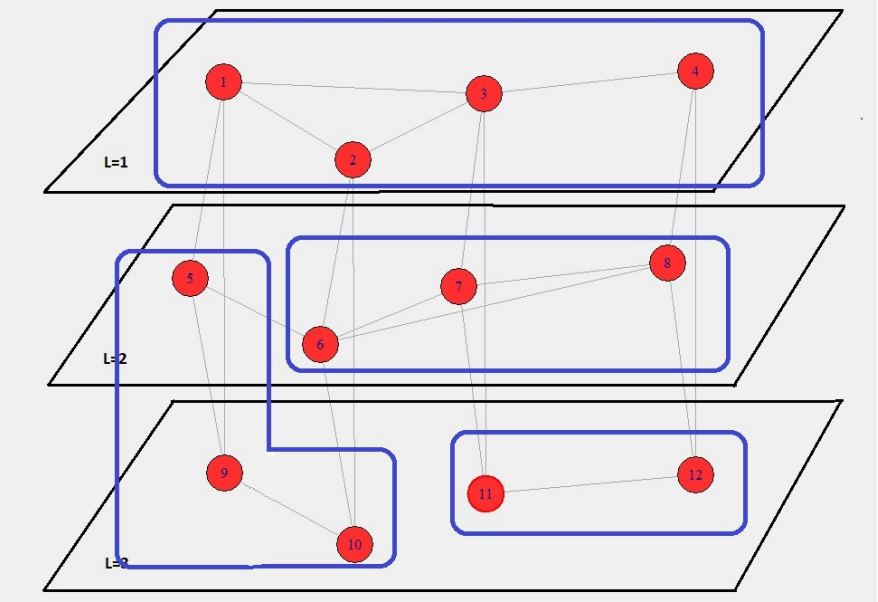


Figure 4.4: Weighted Multilayer Network Communities obtained by modularity approach.

4.3.2 Community detection on multilayer networks based on Communicability Graph

Communicability can be used to identify network communities. The intuition behind this methodology is that nodes inside a community communicate better than nodes belonging to different communities. In this context, a community is then defined as a group of nodes in which each pair has larger intracluster communicability than intercluster one. Let us briefly remind some details about how to exploit this idea on monoplex networks; next, we will extend it to multilayer networks.

4.3.2.1 Monoplex network

Communicability between a pair of nodes μ and ν has been defined in chapter 1, formula 1.1, and in chapter 2, formula 2.1. For the present purposes, it is convenient to write communicability by using the spectral decomposition of the adjacency matrix A as [47]

$$G_{\mu\nu} = \sum_{i=1}^{n} \varphi_\mu(i)\varphi_\nu(i)e^{\lambda_i} \tag{4.25}$$

where $\varphi(i)$ are the eigenvectors of A and λ_i are the corresponding eigenvalues, with $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. Note that the expression $\varphi_\mu(i)$ is used here to represent component μ of the eigenvector i. Let us rewrite this function in the following way:

$$\begin{aligned} G_{\mu\nu} = {} & \varphi_\mu(1)\varphi_\nu(1)e^{\lambda_1} + \\ & + \sum_i \varphi_\mu^+(i)\varphi_\nu^+(i)e^{\lambda_i} + \sum_i \varphi_\mu^+(i)\varphi_\nu^-(i)e^{\lambda_i} \\ & + \sum_i \varphi_\mu^-(i)\varphi_\nu^+(i)e^{\lambda_i} + \sum_i \varphi_\mu^-(i)\varphi_\nu^-(i)e^{\lambda_i} \end{aligned} \tag{4.26}$$

where $\varphi_\mu^+(i)$ and $\varphi_\mu^-(i)$ represent components of the $i-$th eigenvector having positive and negative sign, respectively and each sum is made over components with the sign pattern specified in its general term. Now, we can interpret the sign of the eigenvector components as a *state* of the corresponding node. For instance, if $\varphi_\mu(i) > 0$, we say that node μ is in a positive state for the eigenstate corresponding to λ_i. This state can be visually represented as a small arrow on the node pointing in a given verse, let's say the positive one. Then, $\varphi_\mu(i) < 0$, means that the arrow on node is pointing in the opposite verse, let's say the negative one. We can also interpret these states by considering that an individual in a social or economic network has a positive or negative position with

respect to some criterion depending on whether the sign of $\varphi_\mu(i)$ is positive or negative, respectively (or neutral in case $\varphi_\mu(i) = 0$).

Then, the first term $\varphi_\mu(1)\varphi_\nu(1)e^{\lambda_1}$ in 4.26 represents the *consensus configuration* in which all the nodes share the same state or point in the same verse, since components $\varphi_\mu(1)$ have all the same sign.

In the first and in the forth sum in 4.26, components μ and ν of the eigenvectors (those actually considered in the sums) have the same sign, both positive or both negative. In other words, nodes μ and ν are in a state where the arrows all point in the same verse. Consequently, we can consider μ and ν in the same cluster.

The second and third sums, on the other hand, represent a lack of consensus in the states of the nodes μ and ν, i.e. their eigenvector components have different signs, and their states point in different verses. Then, we can consider that they belong to different clusters.

As a consequence of the previous sign pattern analysis, we call *intracluster communicability* and *intercluster communicability* between a pair of nodes, respectively

$$G_{\mu\nu}^{\text{intracluster}} = \sum_i \varphi_\mu^+(i)\varphi_\nu^+(i)e^{\lambda_i} + \sum_i \varphi_\mu^-(i)\varphi_\nu^-(i)e^{\lambda_i} \tag{4.27}$$

and

$$G_{\mu\nu}^{\text{intercluster}} = \sum_i \varphi_\mu^+(i)\varphi_\nu^-(i)e^{\lambda_i} + \sum_i \varphi_\mu^-(i)\varphi_\nu^+(i)e^{\lambda_i} \tag{4.28}$$

We stress again that the sums above are extended only to the eigenvectors whose components μ and ν comply with the indicated criterion for signs.

The consensus configuration does not give us any information about the community structure of a network. In that state the whole network behave as a single community. Consequently, we can write this term off communicability, consider $G_{\mu\nu} - \varphi_\mu(1)\varphi_\nu(1)e^{\lambda_1}$ and focus on the difference between the intra- and intercluster communicability:

$$\Delta G_{\mu\nu} = G_{\mu\nu}^{\text{intracluster}} + G_{\mu\nu}^{\text{intercluster}} = |G_{\mu\nu}^{\text{intracluster}}| - |G_{\mu\nu}^{\text{intercluster}}| \tag{4.29}$$

or, equivalently, in matrix form

$$\Delta G = e^A - e^{\lambda_1}\varphi(1)\varphi(1)^T \tag{4.30}$$

If $\Delta G_{\mu\nu} > 0$, two nodes display larger intracluster than intercluster communicability and they are members of the same communicability cluster; if $\Delta G_{\mu\nu} < 0$, two nodes display larger intercluster than intracluster communicability and they belong to

different communicability clusters. A community is therefore a subset of nodes $C \subseteq V$ such that the intracluster communicability is greater than the intercluster one for all the couples $(\mu, \nu) \in C$.

Let us notice that the product $e^{\lambda_1}\varphi(1)\varphi(1)^T$ plays the role of null model like the last term in equation 4.22 for classical modularity. Indeed, we are shifting each communicability value according to its value, in such a way to get a list of positive/negative gains/costs in coupling two nodes into the same community. Moreover, the parallelism between the two terms is further confirmed by the fact that while, the first contains the product of the eigenvector centralities, the second, typical of modularity, contains the product of the degree or strength centralities.

In view of this observation we can extend to this new context the approach described in the previous section. That is, we can use the entries of (the upper triangle of) the matrix ΔG as gains/costs to set a linear programming problem to look for the partition that maximises a new partition quality function.This quality function can be given the same structure as modularity in equation 4.21

$$\mathcal{Q}^{\text{comm}} = Q^c_\mu \Delta G^\mu_\nu Q^\nu_c \tag{4.31}$$

where we arranged indices position so to apply the usual sum convention and Q^μ_c is equal to 1 if node μ belong to the community c in the partition $\mathscr{P}_c$, 0 otherwise.

Note that the original proposal by Estrada was different. He defines the binary matrix $\mathcal{C}$ as $\mathcal{C}_{\mu\nu} = \Theta(\Delta G_{\mu\nu})$ where Θ is the Heavyside step function

$$\Theta(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x \leq 0 \end{cases} \tag{4.32}$$

$\mathcal{C}$ is the adjacency matrix of the so called *communicability graph*. A communicability community is given by a clique in the communicability graph. This means that, in the original, methodology, we find communicability communities looking for cliques in the communicability graph. Our choice makes this approach more homogeneous to the first one, described in terms of classical modularity in the previous section - and to the next one in terms of network metrics - and lends itself to being extended, at least in terms of writing equivalence, to the case of multilevel networks. Finally, although communicability can also be defined on directed networks, we prefer to limit the discussion here to the case of undirected networks, both binary or weighted.

4.3.2.2 Multilayer networks

Let $W_{\nu\beta}^{\mu\alpha}$ be the adjacency tensor of a multilayer network. The communicability tensor can be defined as

$$G_{\nu\beta}^{\mu\alpha} = e^{W_{\nu\beta}^{\mu\alpha}} \tag{4.33}$$

where the exponential of a tensor is defined as

$$e^{W_{\nu\beta}^{\mu\alpha}} = \sum_{k=0}^{+\infty} \frac{1}{k!} (W_{\nu\beta}^{\mu\alpha})^k \tag{4.34}$$

and $(W_{\nu\beta}^{\mu\alpha})^k = W_{\rho_1\gamma_1}^{\mu\alpha} W_{\rho_2\gamma_2}^{\rho_1\gamma_1} \dots W_{\nu\beta}^{\rho_{k-1}\gamma_{k-1}}$. In the case of a weighted network, the adjacency tensor can be preliminarily normalised as $T_{\rho\gamma}^{\mu\alpha} W_{\sigma\delta}^{\rho\gamma} T_{\nu\beta}^{\sigma\delta}$ where $T_{\nu\beta}^{\mu\alpha} = [(\operatorname{diag} S)_{\nu\beta}^{\mu\alpha}]^{-1/2}$ and $S = S^{\mu\alpha}$ is defined as in 4.11. Note that $\operatorname{diag} S$ transforms a 2^{nd} order matrix into a 4^{th} order diagonal tensor. Now, by the singular value decomposition of the tensor $W_{\nu\beta}^{\mu\alpha}$, we get the eigenvalues λ and the eigen-matrices $\Phi_{\mu\alpha}$:

$$W_{\nu\beta}^{\mu\alpha} \Phi_{\mu\alpha} = \lambda \Phi_{\nu\beta} \tag{4.35}$$

In particular we identify $\Phi_{\mu\alpha}^{(1)}$ corresponding to the maximum eigenvalue λ_1 and we can consider the null model $e^{\lambda_1} \Phi_{(1)}^{\mu\alpha} \Phi_{\nu\beta}^{(1)}$ to build

$$\Delta G_{\nu\beta}^{\mu\alpha} = G_{\nu\beta}^{\mu\alpha} - e^{\lambda_1} \Phi_{(1)}^{\mu\alpha} \Phi_{\nu\beta}^{(1)} \tag{4.36}$$

We can now use the entries of $\Delta G_{\nu\beta}^{\mu\alpha}$ to set a linear programming problem to look for the partition that maximises the communicability quality function

$$\mathcal{Q}^{\text{comm}} = Q_{\mu\alpha}^{c} \Delta G_{\nu\beta}^{\mu\alpha} Q_{c}^{\nu\beta} \tag{4.37}$$

When we apply the maximum communicability quality function approach to detect optimal communities to the two toy networks in figures 4.1 and 4.2, we obtain the partitions represented in figures 4.5 and 4.6 for the binary and weighted version, respectively. Again, for a straightforward visualization, we have simply grouped nodes belonging to the same community by means of a closed blue line. Although in the binary case an optimal partition equal to that by modularity is gained, in the weighted case a different partition is produced as a consequence of the strong intracluster communicability inside triangles $1-2-3$ on level 1 and $6-7-8$ on level 2.

Binary case:

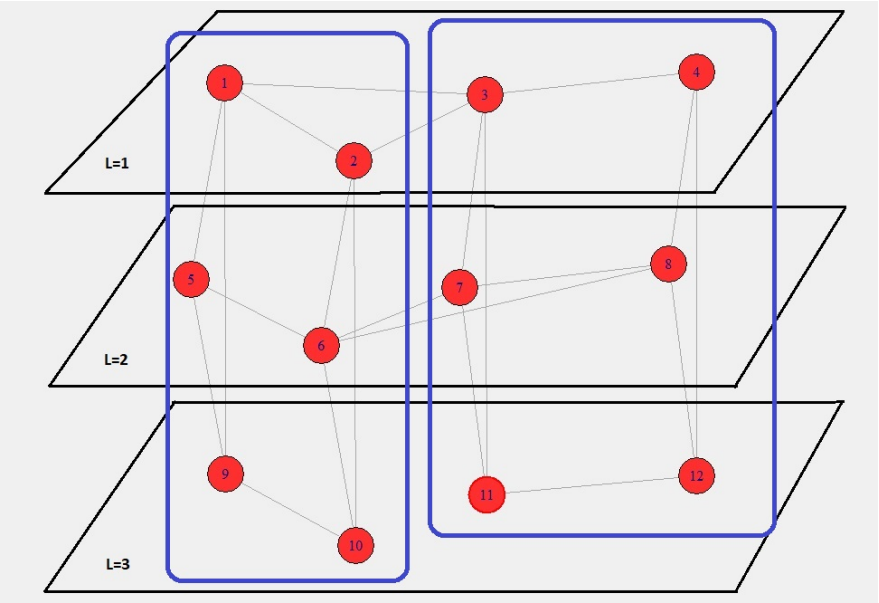


Figure 4.5: Binary Multilayer Network Communities by communicability approach.

Weighted case:

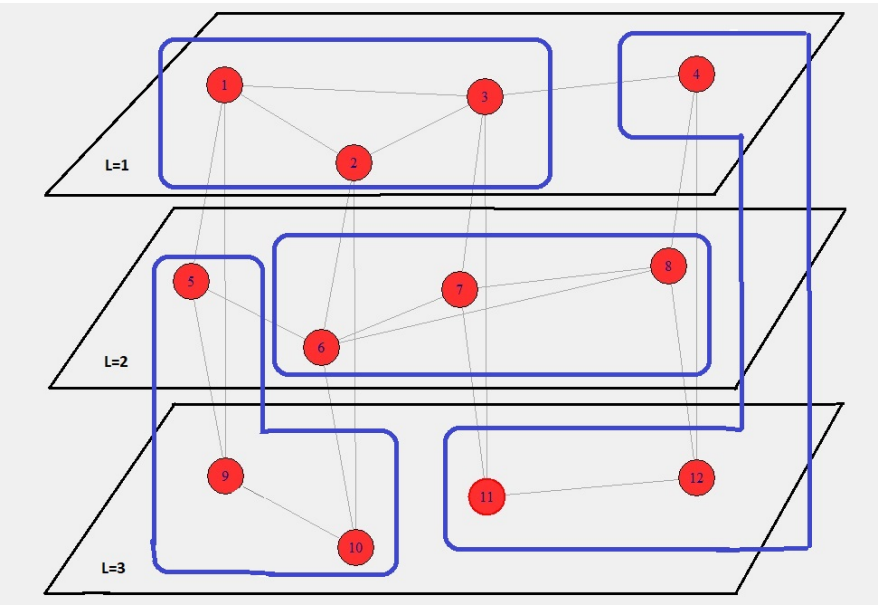


Figure 4.6: Weighted Multilayer Network Communities by communicability approach.

4.3.3 Community detection on multilayer network based on Communicability Distance

The communicability distance $\xi_{\mu\nu}$ has been defined in chapter 2, by means of equation 2.7 (see also [108]). For two nodes μ and ν in a monoplex network, it can be written as

$$\xi_{\mu\nu} = G_{\mu\mu} - 2G_{\mu\nu} + G_{\nu\nu}. \tag{4.38}$$

The diagonal element $G_{\mu\mu}$ represents the subgraph centrality of node μ and it measures the amount of information that starts from and returns to node μ after having wandered through the network. On the other hand, $G_{\mu\nu}$ measures the amount of information transmitted from μ to ν. Notice that the word *information* is meant in its broadest sense. Therefore, information flow can be understood as any kind of flow along edges: money, current, traffic and so on. Thus, the quantity $\xi_{\mu\nu}$ accounts for the difference in the amount of information that returns to the nodes μ and ν and the amount of information actually exchanged between them.

In a matrix form, $\xi_{\mu\nu}$ can be expressed as in equation 2.4.1:

$$\mathbf{\Xi} = \mathbf{g}\mathbf{u}^T - 2\mathbf{G} + \mathbf{u}\mathbf{g}^T \tag{4.39}$$

If we set a distance measure on the network, we can exploit the partition quality index Q defined in [12] for general metric spaces. It is built starting from the cohesion coefficient defined as:

$$\gamma_{\mu\nu} = \bar{\xi}_\mu + \bar{\xi}_\nu - \xi_{\mu\nu} - \bar{\xi} \tag{4.40}$$

where $\bar{\xi}_\mu$ is the mean communicability distance of node μ from all the other nodes in the network, and $\bar{\xi}$ is the mean distance over the whole network. Coefficient $\gamma_{\mu\nu}$ can be interpreted as a cohesion measure between nodes μ and ν. Two nodes μ and ν are said to be cohesive (or incohesive) if $\gamma_{\mu\nu} \geq 0$ ($\gamma_{\mu\nu} \leq 0$). In other words, $\gamma_{\mu\nu}$ represents the gain (when positive) or the cost (when negative) related to the grouping of nodes μ and ν in the same cluster of a given partition $\mathscr{P} = \{\mathscr{P}_c\}$.

In equation 2.10, we introduced an objective function that represents the global cohesion function in the form

$$\mathrm{Q}^{\mathrm{metric}} = \sum_{\mu,\nu} \gamma_{\mu\nu}\, x_{\mu\nu} \tag{4.41}$$

where $x_{\mu\nu}$ are binary variables equal to 1 if two nodes are in the same cluster and 0 otherwise. We can give this function an alternative expression, more suitable for the

present discussion, equivalent to equation 4.31:

$$\mathrm{Q}^{\mathrm{metric}} = Q^c_\mu \, \gamma^\mu_\nu \, Q^\nu_c \tag{4.42}$$

Let us notice that, according to our definition, in both the extreme cases, i.e. the partition $\mathscr{P}_c,\ c = 1$, made up of a unique community equal to the entire network and the partition $\mathscr{P}_c,\ c = 1, \ldots, n$, made up af all isolated nodes, Q reduces to $n\bar{\xi}$, as proved in Chapter 3, section 2.5.1.

In order to extend this methodology to multilayer networks, we refer to the so called unfolding procedure (named also flattening procedure or matricization) discussed in Appendix D. A multilayer network can be equivalently described in terms of adjacency tensor, as we did in the previous paragraphs, and in terms of a block matrix, with L^2 square blocks each one of order N, called *supradjacency matrix*. This matrix is the unfolding of the adjacency tensor $W^{\mu\alpha}_{\nu\beta}$: it contains in the diagonal block the adjacency matrices of each level and in the out-of-diagonal block the adjacencies between different levels.

Supradjacency matrix can be used, like an ordinary adjacency matrix, to compute distances $\xi^{\nu\beta}_{\mu\alpha}$ between node μ on level α and node ν on level β.[1]

At this point we are able to calculate the cohesion tensor as

$$\gamma^{\nu\beta}_{\mu\alpha} = \bar{\xi}_{\mu\alpha} + \bar{\xi}^{\nu\beta} - \xi^{\nu\beta}_{\mu\alpha} - \bar{\xi} \tag{4.43}$$

and we can use the γ's coefficients as gain/costs for the partition quality index $\mathrm{Q}^{\mathrm{metric}}$ in a usual linear programming problem, being now

$$\mathrm{Q}^{\mathrm{metric}} = Q^c_{\mu\alpha} \, \gamma^{\mu\alpha}_{\nu\beta} \, Q^{\nu\beta}_c \tag{4.44}$$

We look then for the maximum value for 4.44 in order to select the optimal partition.

When we apply the maximum metric quality factor approach to detect optimal communities to the two toy networks in figures 4.1 and 4.2, we obtain the partitions represented in figures 4.7 and 4.8 for the binary and weighted version respectively. For a straightforward visualization, we have again simply grouped nodes belonging to the same community by means of a closed blue line. The same partition obtained by the previous methods is confirmed in the binary case. In the weighed case, results are

[1] Let us notice that in this way, for the sake of simplicity, we are implicitly assigning the same meaning to distances between nodes in the same level, distances between versions of the same node in different levels, and distances between different nodes in different levels. This could be a significant limitation to be overcome in the future by refining the discussion.

consistent with those by the method based on the communicability graph, although two previous distinct communities are now unified in single one. This difference could be explained in this way. The cohesion tensor depends not only on the distance but also on the greater or lesser centrality of the nodes in the network through the average value of their distances from all the other nodes as shown in the equation 4.43. Therefore, for more peripheral nodes, with large average distance, it is not necessary to have a very small distance for the gamma coefficient to be positive and the nodes to be placed in the same cluster.

Binary case:

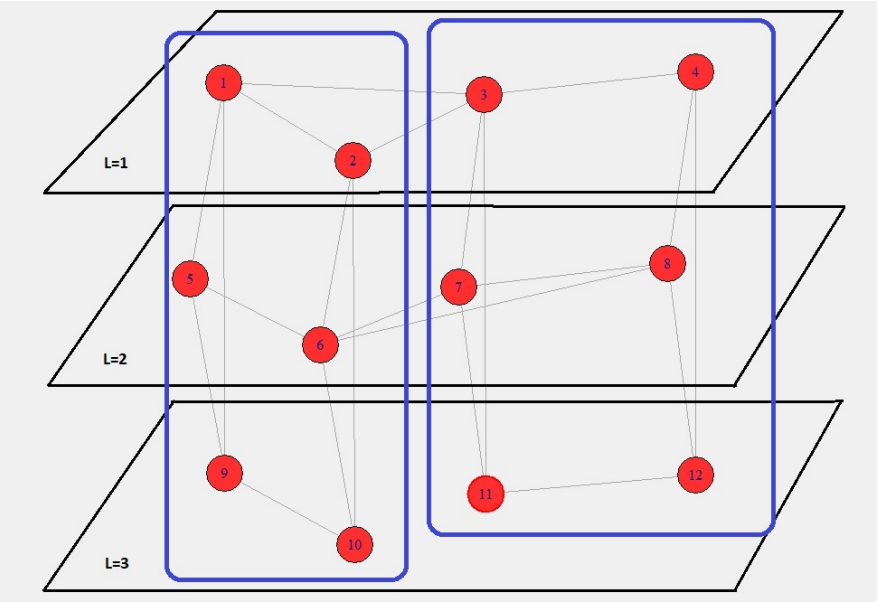


Figure 4.7: Binary Multilayer Network Communities obtained by communicability distance approach.

Weighted case:

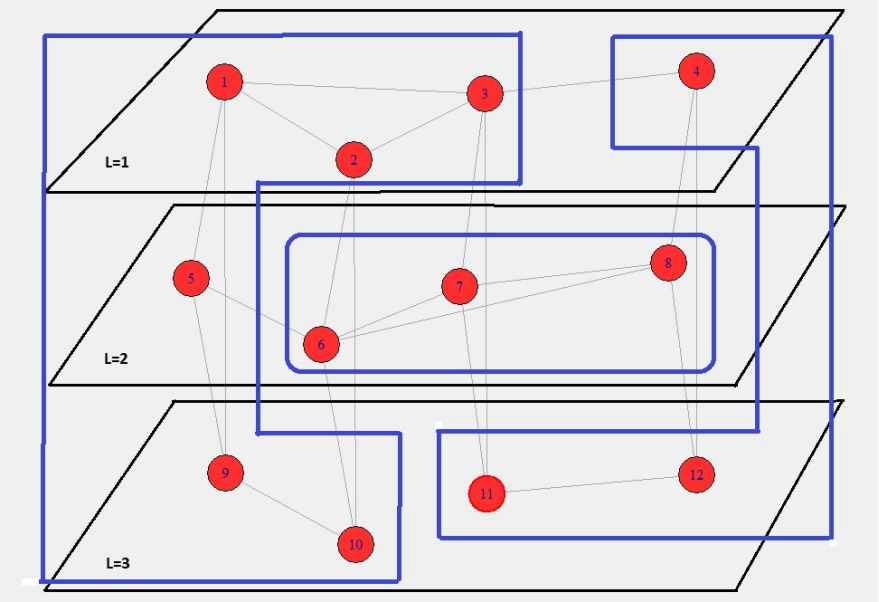


Figure 4.8: Weighted Multilayer Network Communities obtained by communicability distance approach.

4.3.4 Some remarks

We conclude this paragraph about community detection on multilayer networks with some observations about our findings and some future perspectives. Firstly, when applied to the same binary network in 4.1 the three different methodologies gave the same optimal partition: the links between levels strongly tie nodes 1 and 2 on level 1 and their counterparts on levels 2 and 3; similarly for nodes 3 and 4 and their counterparts. Actually the choice in the arrangements of links made a node and its counterparts linked in a cycle, a triangle; moreover nodes $9-10$ and nodes $11-12$ are disconnected each other. The structural bonds, that is the unweighted triangles, are so strong that they emerge immediately in any methodology and the bi-partition as optimal partition is not unexpected. Secondly, the presence of weights on links in network 4.2 changes things quite radically, producing three different optimal partitions. This fact highlights how the three methods allow to underline different aspects of the interaction between nodes in the network. In all cases we detect the presence of a community made up of the triangle $6-7-8$ on level 2 and, as a whole, the partitions obtained are consistent with each other.

Although having been described in tensorial terms, all the quantitative results were obtained both by using the tensorial approach and by using the unfolding procedure, and of course in both procedures results are always equal. The second and third methodology still haven't been checked on directed networks but since communicability matrix is defined on directed networks as well, we expect a quite easy extension to them too. In all the methodologies applied here, we solved in exact form the underlying linear programming problem by using R software. For larger networks, there would be non trivial computational problems and heuristic algorithms, like the one we adopted in chapter 3, would be necessary. This fact opens up the possibility of applying and testing these methods to very dense networks, in particular to the World Trade Network, in order to compare any new results with those obtained in the previous chapters.

To conclude we would like to emphasize that, in light of the complexity of the multilayer networks, the proposed methodologies represent valid options to highlight, according to the different nature of the optimized function, different aspects of the mesoscale structure of the network, especially when the role of weights is crucial.

4.4 Clustering coefficients

The last section of this chapter is devoted to the study of the clustering coefficients in multilayer networks. Our purpose is twofold. In the first instance, we aim at extending the well-known definitions of clustering coefficients in the literature to multilayer networks, in each case providing a version of such a coefficient that takes into account the role of a node in the entire network or in the level in which it is located, or by providing global coefficients for the level and for the entire network. In the second instance, we want to show that all these coefficients can be re-expressed by means of a unified formula, provided that a certain null reference model is appropriately chosen for the denominator. Indeed, the language of tensors allows to give them a unique expression that differs only in the choice and nature of the tensors involved. But since, in general, any clustering coefficient is related to the number of actual and potential triangles in the network, the premise is the very definition of triangle we will adopt for multilayer networks.

4.4.1 Triangles in multilayer networks

Before introducing the definition of clustering coefficients, we need to define what a triangle in a multilayer network is. A triangle in a multilayer network is a closed triplet (a three-cycle) i, j, k such that the three nodes can belong to up to three different levels

and they are connected by inter or intra-layers links. By this definition, we mean to include all possible closed triplet, moving in all directions, along inter or intra-layer links. This definition extends the one adopted in [16] for a monoplex unweighted network. In particular, in a monoplex unweighted network, the number of actual triangles to which node i belongs is given by

$$t(i) = W_{\nu}^{\mu} W_{\rho}^{\nu} W_{\sigma}^{\rho} e_{\mu}(i) e^{\sigma}(i)$$

where W_{ν}^{μ} is the weighted adjacency matrix. Similarly, in a multilayer unweighted network the number of actual triangles to which node i on level a belongs, taking into account all possible three-cycles as claimed above, with vertices on the same level or not, is given by

$$t(i,a) = M_{\nu\beta}^{\mu\alpha} M_{\rho\gamma}^{\nu\beta} M_{\sigma\delta}^{\rho\gamma} E_{\mu\alpha}(i,a) E^{\sigma\delta}(i,a) \tag{4.45}$$

where $M_{\nu\beta}^{\mu\alpha}$ is the weighted adjacency tensor defined in 4.2.

By contracting over all the levels on which node i lies, we get the total number of three-cycles to which that node belongs

$$t(i) = M_{\nu\beta}^{\mu\alpha} M_{\rho\gamma}^{\nu\beta} M_{\sigma\alpha}^{\rho\gamma} E_{\mu}^{\sigma}(i) \tag{4.46}$$

Finally, summing up over all nodes, we obtain the total number of triangles in the multilayer network:

$$t = M_{\nu\beta}^{\mu\alpha} M_{\rho\gamma}^{\nu\beta} M_{\mu\alpha}^{\rho\gamma} \tag{4.47}$$

In a weighted network, the previous definitions represent the weighted number of triangles, where the weight of a single triangle is given by the product of the weights of the edges that make it up.

4.4.2 General definitions

In this section we introduce the general definitions for local and global clustering coefficients on multilayer networks. The clustering coefficient is typically the ratio between the number or weight of actual triangles to which a node belongs, and the number or weight of potential triangles to which it could belong. But in a multilayer network we have at first to decide where to take actual or potential triangles around a given node. That's why, in this framework, we can define three different versions of local clustering coefficient depending on which nodes and levels are taken into account and one global clustering coefficient for the whole network. Specifically, we will denote with

1. $C(i,a)$ the clustering coefficient of node i on level a;
2. $C(i)$ the clustering coefficient of node i over all its levels;
3. $C(a)$ the clustering coefficient of level a over all its nodes;
4. C the global clustering coefficient of the network, over all nodes and levels.

We provide here a general expression for each of the four types of clustering coefficients. In the next subsections we will adapt this general definitions to the cases of an undirected weighted network and a directed weighted network respectively.

Let everywhere $F^{\nu\beta}_{\rho\gamma} = U^{\nu\beta}_{\rho\gamma} - \delta^{\nu\beta}_{\rho\gamma}$ be the adjacency tensor of the complete multilayer network. In this network, a node in one level is connected with all other nodes in all levels except itself. Let us notice that the weights of a complete network will always be understood to be 1.

Let H be any adjacency-like tensor with components $H^{\mu\alpha}_{\nu\beta}$. By adjacency-like tensor we mean a possible and useful modification of the binary or weighted adjacency tensor. The actual choice of $H^{\mu\alpha}_{\nu\beta}$ will be depending on the specific definition of an already existing clustering coefficient in literature or on the new definitions we will provide in the present work. We aimed firstly at showing how all existing definitions can be traced back to a unified writing. Let $[H_k]^{\mu\alpha}_{\nu\beta},\ k = 1,\dots,5$ be five not necessary equal adjacency-like tensors all referring to the same network. Again their choice will be depending on the particular coefficient we are going to describe and they could be binary or weighted, normalised or not, symmetrised or not, accordingly. Now, we define

1. Clustering coefficient of node i on level a:

$$C(i,a) = \frac{[H_1]^{\mu\alpha}_{\nu\beta}[H_2]^{\nu\beta}_{\rho\gamma}[H_3]^{\rho\gamma}_{\sigma\delta}E_{\mu\alpha}(i,a)E^{\sigma\delta}(i,a)}{[H_4]^{\mu\alpha}_{\nu\beta}[F]^{\nu\beta}_{\rho\gamma}[H_5]^{\rho\gamma}_{\sigma\delta}E_{\mu\alpha}(i,a)E^{\sigma\delta}(i,a)} \tag{4.48}$$

2. Clustering coefficient of node i (over the whole network):

$$C(i) = \frac{[H_1]^{\mu\alpha}_{\nu\beta}[H_2]^{\nu\beta}_{\rho\gamma}[H_3]^{\rho\gamma}_{\sigma\alpha}E^{\sigma}_{\mu}(i)}{[H_4]^{\mu\alpha}_{\nu\beta}[F]^{\nu\beta}_{\rho\gamma}[H_5]^{\rho\gamma}_{\sigma\alpha}E^{\sigma}_{\mu}(i)} \tag{4.49}$$

3. Clustering coefficient of the level a (over all the nodes on the level):

$$C(a) = \frac{[H_1]^{\mu\alpha}_{\nu\beta}[H_2]^{\nu\beta}_{\rho\gamma}[H_3]^{\rho\gamma}_{\mu\delta}E^{\delta}_{\alpha}(a)}{[H_4]^{\mu\alpha}_{\nu\beta}[F]^{\nu\beta}_{\rho\gamma}[H_5]^{\rho\gamma}_{\mu\delta}E^{\delta}_{\alpha}(a)} \tag{4.50}$$

4. Global clustering coefficient of the whole network:

$$C = \frac{[H_1]^{\mu\alpha}_{\nu\beta}[H_2]^{\nu\beta}_{\rho\gamma}[H_3]^{\rho\gamma}_{\mu\alpha}}{[H_4]^{\mu\alpha}_{\nu\beta}[F]^{\nu\beta}_{\rho\gamma}[H_5]^{\rho\gamma}_{\mu\alpha}} \tag{4.51}$$

These definitions encode in a unique general formula all the coefficients already existing in the literature for monoplex network, and extend them to unweighted or weighted, undirected or directed, multilayer networks, by properly setting the adjacency tensors $[H_k]^{\mu\alpha}_{\nu\beta}$. In the next two sections, we will show how to recover them in each case.

4.4.3 Weighted undirected networks

In a weighted multilayer network, both intra-layer links and inter-layer links are weighted and all tensors are symmetric.[1]

Let us denote by

- M the weighted adjacency tensor;
- A the corresponding binary adjacency tensor;
- $\tilde{M} = \frac{1}{\mathcal{W}}M$ the normalised adjacency tensor, where $\mathcal{W} = \max_{\mu\nu\alpha\beta} M^{\mu\alpha}_{\nu\beta}$;
- $\hat{M} = \tilde{M}^{1/3}$ the classical entry-wise cubic root of $\tilde{M}$.

We refer here to formula 4.48 but formulae 4.49, 4.50 and 4.51 can be adapted in a similar manner. In monoplex weighted undirected networks the most important clustering coefficients in the literature are provided by De Domenico, Barrat and Onnela coefficients (see [16, 154, 155] respectively). These coefficients can be recovered in our context as follows:

1. De Domenico Clustering Coefficient:

   we set $H_k = \tilde{M}$, for each $k = 1, \ldots, 5$, and

$$C(i,a) = \frac{\tilde{M}^{\mu\alpha}_{\nu\beta}\tilde{M}^{\nu\beta}_{\rho\gamma}\tilde{M}^{\rho\gamma}_{\sigma\delta}E_{\mu\alpha}(i,a)E^{\sigma\delta}(i,a)}{\tilde{M}^{\mu\alpha}_{\nu\beta}F^{\nu\beta}_{\rho\gamma}\tilde{M}^{\rho\gamma}_{\sigma\delta}E_{\mu\alpha}(i,a)E^{\sigma\delta}(i,a)} \tag{4.52}$$

[1] Symmetries of 4^{th} order tensors present a richer set of possibilities than the symmetry of 2^{nd} order tensors, since a number of symmetries can be defined by applying different 'symmetry rules' on the four coefficient indices. Indeed, we may have major symmetry, minor symmetry and total symmetry. We refer here to the major symmetry whose rule is $H^{\mu\alpha}_{\nu\beta} = H^{\nu\beta}_{\mu\alpha}$

Formula 4.52 generalizes the clustering coefficient introduced by De Domenico et al. in [16].

2. Barrat Clustering Coefficient:

   we set $H_1 = H_4 = M$ and $H_2 = H_3 = H_5 = A$, and

$$C(i,a) = \frac{M^{\mu\alpha}_{\nu\beta} A^{\nu\beta}_{\rho\gamma} A^{\rho\gamma}_{\sigma\delta} E_{\mu\alpha}(i,a) E^{\sigma\delta}(i,a)}{M^{\mu\alpha}_{\nu\beta} F^{\nu\beta}_{\rho\gamma} A^{\rho\gamma}_{\sigma\delta} E_{\mu\alpha}(i,a) E^{\sigma\delta}(i,a)} \tag{4.53}$$

   Formula 4.53 generalizes the clustering coefficient introduced by Barrat et al. in [154].

3. Onnela Clustering Coefficient:

   we set $H_1 = H_2 = H_3 = \hat{M}$ and $H_4 = H_5 = A$, and

$$C(i,a) = \frac{\hat{M}^{\mu\alpha}_{\nu\beta} \hat{M}^{\nu\beta}_{\rho\gamma} \hat{M}^{\rho\gamma}_{\sigma\delta} E_{\mu\alpha}(i,a) E^{\sigma\delta}(i,a)}{A^{\mu\alpha}_{\nu\beta} F^{\nu\beta}_{\rho\gamma} A^{\rho\gamma}_{\sigma\delta} E_{\mu\alpha}(i,a) E^{\sigma\delta}(i,a)} \tag{4.54}$$

   Formula 4.54 generalizes the clustering coefficient introduced by Onnela et al. in [155].

4.4.4 Weighted directed networks

In a weighted directed multilayer network, both intra-layer links and inter-layer links are weighted and tensors can be asymmetric. Let us denote by

- M the weighted adjacency tensor;
- A the corresponding binary adjacency tensor;
- $\tilde{M}_{\text{out}} = \frac{1}{\mathcal{W}} M$ and $\tilde{M}_{\text{in}} = \frac{1}{\mathcal{W}} M^T$ where $\mathcal{W} = \max_{\mu\nu\alpha\beta} M^{\mu\alpha}_{\nu\beta}$ and, by definition, H^T is given by $\left[H^{\mu\alpha}_{\nu\beta}\right]^T = H^{\nu\beta}_{\mu\alpha}$;
- Let $\tilde{M} = \frac{1}{2}\left(\tilde{M}_{\text{out}} + \tilde{M}_{\text{in}}\right)$
- Let $\hat{M}_{\text{out}} = \left(\tilde{M}_{\text{out}}\right)^{1/3}$ and $\hat{M}_{\text{in}} = \left(\tilde{M}_{\text{in}}\right)^{1/3}$;
- Let $\hat{M} = \frac{1}{2}\left(\hat{M}_{\text{out}} + \hat{M}_{\text{in}}\right)$;
- Let similarly $\tilde{A} = \hat{A} = \frac{1}{2}\left(A_{\text{out}} + A_{\text{in}}\right)$.

We refer again to formula 4.48 but formulae 4.49, 4.50, and 4.51 can be adapted in a similar manner.

1. De Domenico Clustering coefficient:

   we set $H_k = \tilde{M}$, for each $k = 1, \ldots, 5$, and

   $$C(i,a) = \frac{\tilde{M}^{\mu\alpha}_{\nu\beta}\tilde{M}^{\nu\beta}_{\rho\gamma}\tilde{M}^{\rho\gamma}_{\sigma\delta}E_{\mu\alpha}(i,a)E^{\sigma\delta}(i,a)}{\tilde{M}^{\mu\alpha}_{\nu\beta}F^{\nu\beta}_{\rho\gamma}\tilde{M}^{\rho\gamma}_{\sigma\delta}E_{\mu\alpha}(i,a)E^{\sigma\delta}(i,a)} \tag{4.55}$$

   Formula 4.55 is the immediate generalisation of formula 4.52 to the directed case and it is introduced here for the first time.

2. Clemente-Grassi Clustering Coefficient:

   we set $H_1 = H_4 = \tilde{M}$ and $H_2 = H_3 = H_5 = \tilde{A}$, and

   $$C(i,a) = \frac{\tilde{M}^{\mu\alpha}_{\nu\beta}\tilde{A}^{\nu\beta}_{\rho\gamma}\tilde{A}^{\rho\gamma}_{\sigma\delta}E_{\mu\alpha}(i,a)E^{\sigma\delta}(i,a)}{\tilde{M}^{\mu\alpha}_{\nu\beta}F^{\nu\beta}_{\rho\gamma}\tilde{A}^{\rho\gamma}_{\sigma\delta}E_{\mu\alpha}(i,a)E^{\sigma\delta}(i,a)} \tag{4.56}$$

   Formula 4.56 generalizes the clustering coefficient introduced by Clemente and Grassi in [101].[1]

3. Fagiolo Clustering Coefficient:

   we set $H_1 = H_2 = H_3 = \hat{M}$ and $H_4 = H_5 = \hat{A}$, and

   $$C(i,a) = \frac{\hat{M}^{\mu\alpha}_{\nu\beta}\hat{M}^{\nu\beta}_{\rho\gamma}\hat{M}^{\rho\gamma}_{\sigma\delta}E_{\mu\alpha}(i,a)E^{\sigma\delta}(i,a)}{\hat{A}^{\mu\alpha}_{\nu\beta}F^{\nu\beta}_{\rho\gamma}\hat{A}^{\rho\gamma}_{\sigma\delta}E_{\mu\alpha}(i,a)E^{\sigma\delta}(i,a)} \tag{4.57}$$

   Formula 4.57 generalizes the clustering coefficient introduced by Fagiolo in [99].[2]

[1] Remind that, denoting by $K(i,a)$ the total degree of node i on level a, $S(i,a)$ the total strength of node i on level a, and $K_{\text{bil}}(i,a)$ the total bilateral degree of node i on level a, the local clustering coefficient defined by Clemente and Grassi is given by:

$$c(i,a) = \frac{1}{2}\frac{\hat{M}^{\mu\alpha}_{\nu\beta}\hat{A}^{\nu\beta}_{\rho\gamma}\hat{A}^{\rho\gamma}_{\sigma\delta}E_{\mu\alpha}(i,a)E^{\sigma\delta}(i,a)}{S(i,a)\big(K(i,a)-1\big) - 2S_{\text{bil}}(i,a)}$$

.

[2] Remind that, denoting by $K(i,a)$ the total degree of node i on level a and $K_{\text{bil}}(i,a)$ the total bilateral degree of node i on level a, the local clustering coefficient defined by Fagiolo is given by:

$$c(i,a) = \frac{\hat{M}^{\mu\alpha}_{\nu\beta}\hat{M}^{\nu\beta}_{\rho\gamma}\hat{M}^{\rho\gamma}_{\sigma\delta}E_{\mu\alpha}(i,a)E^{\sigma\delta}(i,a)}{K(i,a)\big(K(i,a)-1\big) - 2K_{\text{bil}}(i,a)}$$

4.4.5 Summary

We can summarize, in the following table, all the definitions of clustering coefficient given in the previous paragraphs. We refer to the notations introduced in those paragraphs and we stress that they are written here in a synthetic and symbolic form in order to highlight the extremely compact writing made possible by the introduction of suitably modified adjacency-like tensors.

Definition		
Name		**Formula**
Undirected network	**Directed network**	
De Domenico	De Domenico	$\frac{\tilde{M}\tilde{M}\tilde{M}}{\tilde{M}F\tilde{M}}$
Barrat	Clemente-Grassi	$\frac{MAA}{MFA}$
Onnela	Fagiolo	$\frac{\hat{M}\hat{M}\hat{M}}{AFA}$

Table 4.1: Symbolic recap table of all the clustering coefficients described in text.

Clustering coefficients described in the previous sections via tensorial representation can also be expressed by using the so called unfolding procedure (named also flattening procedure or matricization). We leave the details of this parallel and equivalent description in the appendix D.

4.4.6 Interpretation

The clustering coefficients defined in the previous paragraphs point out different aspects of the way a node is embedded in the network.

De Domenico clustering coefficients take into account the weights of all the links in the actual triangles and multiply them to assign a total weight to each triangle. The weights of the links are normalized dividing by the maximum one, so that the total weight of each triangle falls into $[0, 1]$. Due to the product of the three weights in each triangle, the weights distribution could be very skewed and its mean could be close to zero. Moreover, the weight of the link that completes the potential triangles in the denominator is always 1, i.e. equal to the maximum one. Therefore, since this clustering coefficient is the ratio between the sum of the total weights of actual triangles and the sum of the total weights of potential triangles, it tends to be very close to zero. Of course, the closer is the weight of the actual triangles to that of potential triangles,

the higher is the clustering coefficient. The most significant aspect is the fact that it emphasizes more the role of the total weight of potential triangles than their number.

Onnela and Fagiolo coefficients take into account the geometric mean of the weights of the links in the actual triangles, each weight being normalised like in the De Domenico coefficient. In the numerator we find the sum of this geometric means used as total weights of the actual triangles. In the denominator, we find the number of potential triangles so that Onnela and Fagiolo coefficients actually return the arithmetic mean of the total weight of the actual triangles, assigning a value equal to zero to non existing triangles. This fact results again in a compression towards zero of the values of the resulting clustering coefficients.

Barrat and Clemente-Grassi coefficients take into account only the weights of the two links between the node and its adjacent nodes in an actual triangle. In the numerator, they take the sum of the arithmetic means of these weights and, in the denominator, the mean strength multiplied by the number of potential triangles. Hence, in this way, weights do not need to be normalized and the role of the strength of the node is emphasized. These coefficients typically assume higher values than the previous ones and their proposal seems more consistent when the distribution of weights is very skewed. On the other hand, when the network is almost complete, this coefficient is always very close to 1 independently of the weights on the triangles.

Let us consider the simple undirected monoplex graph in figure 4.9.

The difference between De Domenico, Onnela and Barrat coefficients depends on how we allocate the weights and their distribution among the links. If we give all the links weight 1, the three coefficients give the same result, that is in a binary network all of them give the same description of the clustering of nodes.

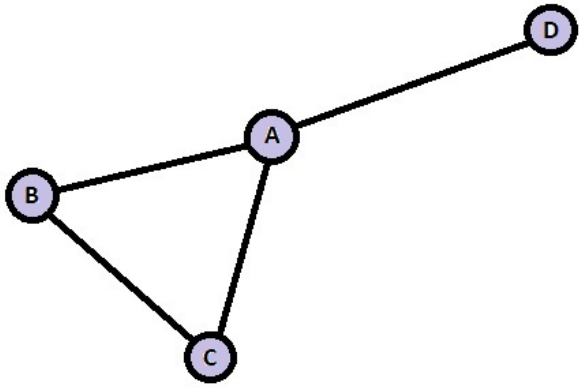


Figure 4.9: Illustrative monoplex network

Consider now node A. Node A belongs to one actual triangle and to three potential triangles. The weight of the link $A - D$ does not influence the Onnela coefficient (except in the case in which it affects normalization), while it influences De Domenico and Barrat coefficients: e.g. assume that the weights of $A - B$, $A - C$ and $B - C$ are equal to 0.5 and the weight $A - D$ equal to 1. For the Onnela coefficient, we have that the geometric mean is 0.5 and it is divided by 3, the number of potential triangles. Whereas, in Barrat coefficient the link $A - D$ contributes to the total strength. In this case the weight of the actual triangle is 0.5 but it is divided by the average strength $\frac{2}{3}$ multiplied by the number of potential triangles, i.e. it is divided by 2.

Vice versa, the weight of the $B - C$ link does not affect Barrat coefficient, while it affects Onnela coefficients. In fact, the former is determined only by the weights of the links $A - B$ and $A - C$. Therefore, if the distribution of the weights on links is very skewed (e.g. we have one very large weight and all the others are very small), Onnela and Fagiolo are more affected by that skewness. For example, let's put $A - D$ equal to 1 and the other links equal to 0.01. The Onnela coefficient of A will be 0.01/3. Conversely Barrat is less affected by that skewness because the weights are also present in the denominator and in this case we have: $0.01/(0.01 + 0.505 + 0.505) = 0.01/1.02$. If a node is involved in a number of triangles equal to the potential ones, its Barrat coefficient will always be 1, whatever the weights, while Onnela coefficient will depend on the weights. For instance, consider node C. Whatever the choice of weights, its Barrat coefficient will always be 1, whereas its Onnela coefficient will depend on the geometric mean of the weights of the triangle $A - B$, $B - C$, $A - C$ divided by 1.

To sum up, the advantage-disadvantage of Onnela coefficient is that in a very dense graph we get very low clustering coefficients if the weights are concentrated in a specific way. So this is more affected by the weights than by the number of triangles. The advantage-disadvantage of Barrat coefficient is that it depends more on the number of triangles than on their weights. The choice of the coefficients strongly depends on the empirical data we are dealing with and what we are interested in when we describe a given network.

In the next section, we will test these observations on a simple multilayer network to which we apply formulae in sections 4.4.3 and 4.4.4.

4.4.7 An illustrative example

Let us consider the undirected multilayer network in figure 4.10, where a set of three nodes are connected in two different layers ($N = 3$ and $L = 2$). The weights of the links are: $M_{21}^{11} = 2$, $M_{31}^{11} = 4$, $M_{31}^{21} = 3$, $M_{22}^{12} = 1$, $M_{32}^{12} = 2$, $M_{32}^{22} = 5$, $M_{12}^{11} = 1$ and

$M_{32}^{11} = 4$. Tables 4.2 and 4.3 collect the values of the clustering coefficients discussed in section 4.4.3.

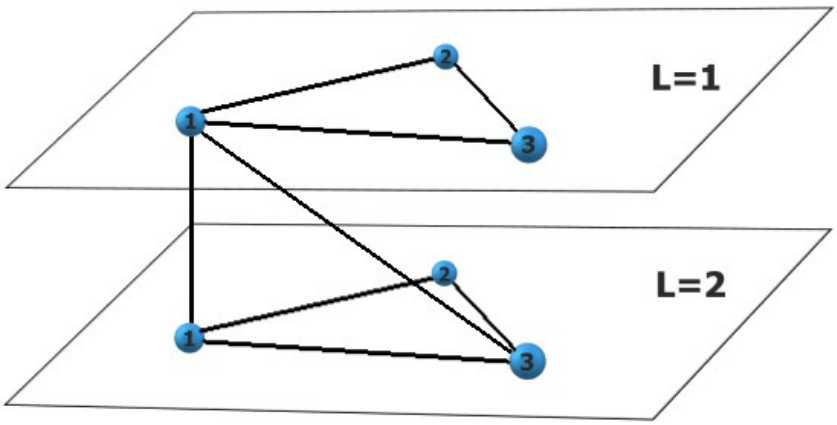


Figure 4.10: Illustrative example for undirected multilayer network

As shown in the tables, for the binary version of the network all coefficients return the same value, as expected. For the weighted version, Barrat coefficients are always greater than others; node 3 in level 2, for instance, has a very small clustering coefficient with De Domenico definition because the weight of actual triangles is very low with respect to that of potential triangles; Onnela coefficient for node 1 in level 2 and node 3 in level 2 are equal because actual triangles are the same and this coefficient is not affected by the weight of potential triangles.

Local Clustering Coefficients $C(i,a)$						
Version	**Level 1**			**Level 2**		
	Node 1	Node 2	Node 3	Node 1	Node 2	Node 3
Binary	0.333	1.000	1.000	0.667	1.000	0.667
De Domenico	0.152	0.800	0.400	0.720	0.400	0.095
Onnela	0.163	0.577	0.577	0.277	0.431	0.277
Barrat	0.333	1.000	1.000	0.750	1.000	0.591

Table 4.2: Values of different clustering coefficients for the network shown in figure 4.10. Binary refers to the common value of local clustering coefficients for the binary version of the network.

Clustering Coefficients						
Version	$C(i)$			$C(a)$		C
	Node 1	Node 2	Node 3	Level 1	Level 2	Network
Binary	0.444	1.000	0.750	0.500	0.714	0.600
De Domenico	0.213	0.618	0.168	0.267	0.192	0.233
Onnela	0.201	0.504	0.352	0.266	0.299	0.282
Barrat	0.415	1.000	0.690	0.511	0.694	0.593

Table 4.3: Values of different clustering coefficients for the network shown in figure 4.10. Binary refers to the common value of local clustering coefficients for the binary version of the network.

Let us consider now in figure 4.11 a directed version of the multilayer network in figure 4.10. The weights of the oriented links are: $M^{11}_{21} = 2$, $M^{11}_{31} = M^{31}_{11} = 4$, $M^{21}_{31} = M^{31}_{21} = 3$, $M^{12}_{22} = M^{22}_{12} = 1$, $M^{32}_{12} = 2$, $M^{22}_{32} = 5$, $M^{11}_{12} = M^{12}_{11} = 1$ and $M^{11}_{32} = 4$. Tables 4.4 and 4.5 collect the values of the clustering coefficients discussed in section 4.4.4.

Also in the direct case, it is confirmed the fact that Clemente-Grassi coefficient is typically greater than the others. Finally, when all the weights are set equal to 1 but the orientation is maintained, again all coefficients return the same value.

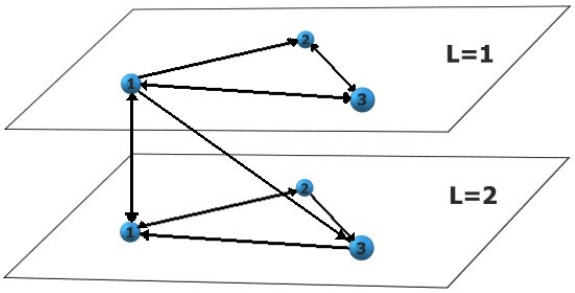


Figure 4.11: Illustrative example for directed multilayer network

Local Clustering Coefficients $C(i,a)$						
Version	**Level 1**			**Level 2**		
	Node 1	Node 2	Node 3	Node 1	Node 2	Node 3
Binary	0.231	1.000	0.500	0.250	0.500	0.667
De Domenico	0.133	0.800	0.200	0.300	0.200	0.095
Fagiolo	0.120	0.577	0.288	0.104	0.215	0.277
Clemente-Grassi	0.243	1.000	0.500	0.300	0.500	0.591

Table 4.4: Values of different clustering coefficients for the network shown in figure 4.11 Binary here refers to the common value of local clustering coefficients for a version of the directed network in which all the weights are set equal to 1 but the orientation is maintained.

Clustering Coefficients						
Version	$C(i)$			$C(a)$		C
	Node 1	Node 2	Node 3	Level 1	Level 2	Network
Binary	0.238	0.750	0.571	0.368	0.384	0.375
De Domenico	0.154	0.527	0.153	0.211	0.153	0.194
Fagiolo	0.114	0.396	0.284	0.203	0.161	0.186
Clemente-Grassi	0.256	0.727	0.540	0.380	0.463	0.407

Table 4.5: Values of different clustering coefficients for the network shown in figure 4.11. Binary here refers to the common value of local clustering coefficients for a version of the directed network in which all the weights are set equal to 1 but the orientation is maintained.

4.5 Conclusions

The previous discussion shows how it is possible to trace the clustering coefficients expressed up to now by means of *ad hoc* formulas to a single writing provided that the necessary choice of the reference tensors is made. When applied to some simple networks taken as a model to test their effectiveness they work very well and always reduce to the known coefficients when the network consists of only one level. Future perspectives will see above all their application to real multilayer networks and the identification of criteria to prefer the application of one or the other of the described coefficients according to the nature and properties of the network.

Appendices

Appendix A

The following theorem follows theorem 3 in Chapter 1 and it provides a close expression for $\mathscr{R}_i$, $\mathscr{C}_i$ and $\mathscr{T}_i$ for a complete network.

Theorem 5. *The risk-dependency $\mathscr{R}_i$ for each node in a complete graph is given by*

$$\mathscr{R}_i = e^{(n-1)\zeta}$$

and the circulability and transmissibility are given by

$$\mathscr{C}_i(\zeta) = \frac{n-1}{n}\left[\frac{e^{(n-1)\zeta}}{n-1} + \frac{1}{e^{\zeta}}\right] \qquad \mathscr{T}_i(\zeta) = \frac{n-1}{n}\left[e^{(n-1)\zeta} - \frac{1}{e^{\zeta}}\right]$$

Proof. For a complete graph, $\psi_j^T \cdot \vec{1} = 0,\ j \neq 1$, because of the mutual orthogonality between $\psi_j,\ j \neq 1$ and the principal eigenvector ψ_1 of constant components. That is, $\mathscr{R}_i$ is completely determined by the eigenvector centralities $\psi_{1,i}$ which of course are equal for every node and equal to $\psi_{1,i} = \frac{1}{\sqrt{n}}$. Since $\lambda_1 = n-1$, we obtain:

$$\mathscr{R}_i = e^{\zeta\lambda_1}\left(\psi_1^T \cdot \vec{1}\right)\psi_{1,i} + 0 = e^{(n-1)\zeta}\left(\frac{1}{\sqrt{n}} \cdot n\right)\frac{1}{\sqrt{n}} = e^{(n-1)\zeta}$$

Subgraph centrality close expression for a complete graph is provided in [70]:

$$\mathscr{C}_i(1) = SC(i) = \frac{1}{n}\left[e^{n-1} + \frac{n-1}{e}\right]$$

Multiplying each entry in A by ζ and summing up the power series, we get

$$\mathscr{C}_i(\zeta) = \frac{n-1}{n}\left[\frac{e^{(n-1)\zeta}}{n-1} + \frac{1}{e^{\zeta}}\right]$$

By difference, we get $\mathscr{T}_i(\zeta)$. □

An important remark concerns the ratio $\frac{\mathscr{C}_i}{\mathscr{R}_i}$. Indeed:

$$\lim_{\zeta\to+\infty}\frac{\mathscr{C}_i}{\mathscr{R}_i}=\lim_{\zeta\to+\infty}\frac{\frac{n-1}{n}\left[\frac{e^{(n-1)\zeta}}{n-1}+\frac{1}{e^{\zeta}}\right]}{e^{(n-1)\zeta}}=\lim_{\zeta\to+\infty}\left[\frac{1}{n}+\frac{n-1}{n}\frac{1}{e^{n\zeta}}\right]=\frac{1}{n}. \tag{A.1}$$

Similarly,

$$\lim_{\zeta\to+\infty}\frac{\mathscr{C}_i}{\mathscr{T}_i}=\frac{\frac{e^{(n-1)\zeta}}{n-1}+\frac{1}{e^{\zeta}}}{e^{(n-1)\zeta}-\frac{1}{e^{\zeta}}}=\frac{1}{n-1}.$$

Appendix B

We report here thorough computations and proofs of formulae 2.3, 2.4 and 2.5 in section 2.3.3. The expression 2.3 of the total potential energy U can be handled in the following way:

$$\begin{aligned}
U &= \frac{1}{4}\mathcal{K}\sum_{i,j} A_{ij}[z_i^2 - 2z_i z_j + z_j^2] \\
&= \frac{1}{4}\mathcal{K}\left[\sum_i z_i^2 \sum_j A_{ij} - 2\sum_{i,j} A_{ij} z_i z_j + \sum_j z_j^2 \sum_i A_{ij}\right] \\
&= \frac{1}{4}\mathcal{K}\left[\sum_i z_i^2 k_i - 2\sum_{i,j} A_{ij} z_i z_j + \sum_j z_j^2 k_j\right] = \frac{1}{2}\mathcal{K}\left[\sum_i z_i^2 k_i - \sum_{i,j} A_{ij} z_i z_j\right] \\
&= \frac{1}{2}\mathcal{K}\left[\sum_{ij} z_i (\mathbf{K\text{-}A})_{ij} z_j\right] = \frac{1}{2}\mathcal{K}\left[\sum_{ij} z_i L_{ij} z_j\right].
\end{aligned}$$

We compute now the expression of the partition function $\mathcal{Z}$ in formula 2.4. Using the spectral decomposition of the Laplacian matrix $\mathbf{L} = \mathbf{M}\mathbf{\Lambda}\mathbf{M}^{\mathbf{T}}$, where $\mathbf{\Lambda}$ is the diagonal matrix of the eigenvalues and $\mathbf{M}$ is the corresponding matrix of the eigenvectors, we have the following chain of equalities:

$$\begin{aligned}
\mathcal{Z} &= \int e^{-\frac{1}{2}\beta\mathcal{K}\sum_{ij} z_i (\mathbf{M}\mathbf{\Lambda}\mathbf{M}^T)_{ij} z_j} \prod_k dz_k = \int e^{-\frac{1}{2}\beta\mathcal{K}\mathbf{z}^T(\mathbf{M}\mathbf{\Lambda}\mathbf{M}^T)\mathbf{z}} \prod_k dz_k \\
&= \int e^{-\frac{1}{2}\beta\mathcal{K}(\mathbf{M}^T\mathbf{z})^T\mathbf{\Lambda}(\mathbf{M}^T\mathbf{z})} \prod_k dz_k = \int e^{-\frac{1}{2}\beta\mathcal{K}\mathbf{x}^T\mathbf{\Lambda}\mathbf{x}} \prod_k dx_k \\
&= \int e^{-\frac{1}{2}\beta\mathcal{K}\sum_k \mu_k x_k^2} \prod_k dx_k
\end{aligned}$$

where we set $\mathbf{x} = \mathbf{M}^T\mathbf{z}$ and $d\mathbf{z} = |\det \mathbf{M}| d\mathbf{x} = d\mathbf{x}$.

As usual in literature, we remove the contribution from $\mu_n = 0$, providing the modified partition function we still call $\mathcal{Z}$, (see [1] pag. 117); this yields:

$$\mathcal{Z} = \prod_{k=1}^{n-1} \int e^{-\frac{1}{2}\beta\mathcal{K}\mu_k x_k^2} dx_k = \prod_{k=1}^{n-1} \sqrt{\frac{2\pi}{\beta\mathcal{K}\mu_k}}$$

the last equality being valid because all integrals are Gaussian with $\mu_k > 0$, $k = 1, ..., n-1$.

Finally we compute $G_{ij}^v(\beta)$ (formula 2.5):

$$\begin{aligned} G_{ij}^v(\beta) &= \frac{1}{\mathcal{Z}} \int z_i z_j e^{-\beta U} d\mathbf{z} \\ &= \frac{1}{\mathcal{Z}} \int z_i z_j e^{-\frac{1}{2}\beta\mathcal{K}\sum_{ij} z_i L_{ij} z_j} \prod_k dz_k \\ &= \frac{1}{\mathcal{Z}} \int (\mathbf{Mx})_i (\mathbf{Mx})_j e^{-\frac{1}{2}\beta\mathcal{K}\sum_k \mu_k x_k^2} \prod_k dx_k \\ &= \frac{1}{\mathcal{Z}} \int \left(\sum_{k=1}^n \psi_k(i) x_k\right) \left(\sum_{k=1}^n \psi_k(j) x_k\right) \prod_{k=1}^n e^{-\frac{1}{2}\beta\mathcal{K}\mu_k x_k^2} dx_k \end{aligned}$$

Notice that, computing the product of the two sums inside the integral above, all the integrals involving mixed terms are null, as the integrand is an odd function and the integral is extended to $\mathbb{R}$ for each x_k. Then, only the squared terms remain inside the integral, so that:

$$\begin{aligned} G_{ij}^v(\beta) &= \frac{1}{\mathcal{Z}} \int \left(\psi_1(i)\psi_1(j)x_1^2 + \cdots + \psi_n(i)\psi_n(j)x_n^2\right) \prod_{k=1}^n e^{-\frac{1}{2}\beta\mathcal{K}\mu_k x_k^2} dx_k \\ &= \frac{1}{\mathcal{Z}} \psi_1(i)\psi_1(j) \int x_1^2 e^{-\frac{1}{2}\beta\mathcal{K}\mu_1 x_1^2} dx_1 \cdot \int e^{-\frac{1}{2}\beta\mathcal{K}\mu_2 x_2^2} dx_2 \cdot \ \ldots \ \cdot \int e^{-\frac{1}{2}\beta\mathcal{K}\mu_n x_n^2} dx_n + \cdots + \\ &\quad \frac{1}{\mathcal{Z}} \psi_n(i)\psi_n(j) \int e^{-\frac{1}{2}\beta\mathcal{K}\mu_1 x_1^2} dx_1 \cdot \int e^{-\frac{1}{2}\beta\mathcal{K}\mu_2 x_2^2} dx_2 \cdot \ \ldots \ \cdot \int x_n^2 e^{-\frac{1}{2}\beta\mathcal{K}\mu_n x_n^2} dx_n \end{aligned}$$

We remove once again the contribution from $\mu_n = 0$, then computing the integrals we have:

$$
\begin{aligned}
G_{ij}^{v}(\beta) = & \frac{1}{\mathcal{Z}}\psi_1(i)\psi_1(j)\frac{\sqrt{2\pi}}{\sqrt{(\beta\mathcal{K}\mu_1)^3}}\cdot\frac{\sqrt{2\pi}}{\sqrt{\beta\mathcal{K}\mu_2}}\cdot\ \ldots\ \cdot\frac{\sqrt{2\pi}}{\sqrt{\beta\mathcal{K}\mu_{n-1}}}+\cdots+ \\
& \frac{1}{\mathcal{Z}}\psi_{n-1}(i)\psi_{n-1}(j)\cdot\frac{\sqrt{2\pi}}{\sqrt{\beta\mathcal{K}\mu_1}}\cdot\frac{\sqrt{2\pi}}{\sqrt{\beta\mathcal{K}\mu_2}}\cdot\ \ldots\ \cdot\frac{\sqrt{2\pi}}{\sqrt{(\beta\mathcal{K}\mu_{n-1})^3}} \\
= & \frac{1}{\mathcal{Z}}\prod_{k=1}^{n-1}\sqrt{\frac{2\pi}{\beta\mathcal{K}\mu_k}}\left[\frac{\psi_1(i)\psi_1(j)}{\beta\mathcal{K}\mu_1}+\cdots+\frac{\psi_{n-1}(i)\psi_{n-1}(j)}{\beta\mathcal{K}\mu_{n-1}}\right]=\sum_{k=1}^{n-1}\frac{\psi_k(i)\psi_k(j)}{\beta\mathcal{K}\mu_k}.
\end{aligned}
$$

Appendix C

The expression of the pseudo-inverse of the Laplacian $\mathbf{L}^+ = \left(\mathbf{L} + \frac{1}{n}\mathbf{J}\right)^{-1} - \frac{1}{n}\mathbf{J}$ allows an interesting interpretation of the resistance distance ω_{ij} in an economic, or financial, networked system.

Suppose that, to each node, a value of a given attribute is assigned through a state vector $\mathbf{v} = [v_1, v_2, \ldots, v_n]^T$ (such an attribute could be, for instance, the GDP of a Country or the assets of a financial institution), and let $I_{ij} = v_i - v_j$ be the flow of such an attribute from node i to node j. We denote by I_i the total outgoing flow from the node i to its adjacent nodes, i.e. $I_i = \sum_{j=1}^{n} a_{ij}(v_i - v_j)$.
In matrix form, the total outgoing flow of the nodes attribute is then

$$\mathbf{I} = (\mathbf{K} - \mathbf{A})\mathbf{v} = \mathbf{L}\mathbf{v}.$$

The Laplacian matrix transforms nodes attributes v_i, $i = 1, ..., n$ into outgoing flows from nodes I_i, under the assumption that a flow I_{ij} along a given edge is equal to the gradient $\Delta v_{ij} = v_i - v_j$. This assumption is equivalent to choose an effective resistance equal to 1 along all edges. Of course, we may have both outgoing and ingoing currents according to the sign of Δv_{ij}: positive for outgoing flows from i and negative for ingoing flows into i.
A similar meaning can be given to $\left(\mathbf{L} + \frac{1}{n}\mathbf{J}\right)\mathbf{v}$. Indeed,

$$\left(\mathbf{L} + \frac{1}{n}\mathbf{J}\right)\mathbf{v} = \mathbf{L}\mathbf{v} + \frac{1}{n}\mathbf{J}\mathbf{v} = \mathbf{I} + \overline{v}\mathbf{u},$$

where $\overline{v} = \frac{1}{n}\sum_{k=1}^{n} v_k$, that is, the operator $\mathbf{L} + \frac{1}{n}\mathbf{J}$ adds to the flows a constant term given by the mean value of all the attributes of the nodes. Then, the matrix $\left(\mathbf{L} + \frac{1}{n}\mathbf{J}\right)$ transforms nodes attributes $\mathbf{v}$ into total outgoing flows $\mathbf{I}$ in the network, up to an additive constant.
In a similar way, the inverse $\left(\mathbf{L} + \frac{1}{n}\mathbf{J}\right)^{-1}$ acts on a current vector $\mathbf{I}$ and produces a state vector $\mathbf{v}$, which can be interpreted as the cause of such currents in the network. Specifically

$$\mathbf{v} = \mathbf{L}^{+}\mathbf{I} = \left[\left(\mathbf{L} + \frac{1}{n}\mathbf{J}\right)^{-1} - \frac{1}{n}\mathbf{J}\right]\mathbf{I} = \left(\mathbf{L} + \frac{1}{n}\mathbf{J}\right)^{-1}\mathbf{I} - \bar{\mathrm{I}}\mathbf{u}$$

where, once again, the term $\frac{1}{n}\mathbf{J}\mathbf{I} = \bar{\mathrm{I}}\mathbf{u}$ is the average value of the outgoing currents coming from every node.

Suppose now that in the system there are an outgoing flow equal to 1 from a node (node 1, for instance), an ingoing flow equal to -1 into another node (for instance, node 2), whereas for all the other nodes the flow is zero. This is equivalent to a current vector equal to $\mathbf{I} = [1, -1, 0, \ldots, 0]^T = \mathbf{e}_1 - \mathbf{e}_2$. Loosely speaking, a unit information is coming out from node 1 and goes entirely into node 2. To produce these flows, we have to start from an initial attributes vector on nodes given by

$$\mathbf{v} = \mathbf{L}^{+}(\mathbf{e}_1 - \mathbf{e}_2) = \left[\left(\mathbf{L} + \frac{1}{n}\mathbf{J}\right)^{-1} - \frac{1}{n}\mathbf{J}\right]\begin{pmatrix} 1 \\ -1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \left(\mathbf{L} + \frac{1}{n}\mathbf{J}\right)^{-1}\begin{pmatrix} 1 \\ -1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

where the last equality holds because $\mathbf{J}(\mathbf{e}_1 - \mathbf{e}_2) = \mathbf{0}$, that is $\bar{\mathrm{I}} = 0$. Thus, the resistance distance between nodes 1 and 2 is given by

$$\omega_{12} = (\mathbf{e}_1 - \mathbf{e}_2)^T \left(\mathbf{L} + \frac{1}{n}\mathbf{J}\right)^{-1}(\mathbf{e}_1 - \mathbf{e}_2) = v_1 - v_2 = \Delta v_{12}.$$

If Δv_{12} is small, a small gradient is enough to transmit such a unit flow from node 1 to node 2; whereas, if $v_1 - v_2$ is big, a high gradient is needed in order to produce the same unit flow. More in general, let's imagine that in the node i the value v_i is positive. Then the fact that another attribute v_j with $j \neq i$ is positive means that node i and node j are strongly correlated since it is enough a low attribute difference to subtract from node i a unit flow. This means that these two nodes communicate a lot. Whereas, if for another node k with $k \neq i$, the corresponding component v_k is negative this implies that node i and node k are strongly anti-correlated since, in order to produce a unit flow from node i, node k has to be at a negative attribute, i.e. the attribute difference between i and k must be high. This means that the two nodes don't communicate well. The signs of the components of the vector $\mathbf{v}$ indicate nodes that are positively or negatively correlated with node i according to the fact these components have the same sign as v_i or not. Let us observe that, in general,

$\mathbf{v} = \mathbf{L}^{+}\mathbf{I} = \mathbf{L}^{+}(\mathbf{e}_i - \mathbf{e}_j) = L_i^{+} - L_j^{+}$ with L_i^{+} i-th column of the matrix L^{+}. That is, if we want to decrease by 1 the attribute of node i and increase by 1 the attribute of node j, we have to take an initial distribution of attributes on nodes equal to the difference between i-th column of $\mathbf{L}^{+}$ and j-th column of $\mathbf{L}^{+}$, and these columns are also the values of vibrational communicability $\mathbf{G}^{v}$ between nodes, as defined in the text.

Appendix D

Clustering coefficients introduced in sections 4.4.3 and 4.4.4 via tensorial representation can also be expressed by using the so called unfolding procedure (named also flattening procedure or matricization). It consists in representing the adjacency tensor by a block matrix, with L^2 square blocks each one of order N, called *supradjacency matrix*

$$\mathbf{W} = \begin{bmatrix} \mathbf{W}_1 & \mathbf{W}_{12} & \cdots & \mathbf{W}_{1L} \\ \mathbf{W}_{21} & \mathbf{W}_2 & \cdots & \mathbf{W}_{2L} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{W}_{L1} & \mathbf{W}_{L2} & \cdots & \mathbf{W}_L \end{bmatrix}$$

where the diagonal blocks $\mathbf{W}_a$, $a = 1, ..., L$ represent the weighted adjacency matrix of each layer and the out off diagonal blocks $\mathbf{W}_{ab}$, $a, b = 1, ..., L$, represent the weighted adjacency relations between nodes on layers a and nodes on layer b.

Let us observe that, for instance, the fourth order tensor generated by a tensorial product like $M^{\mu\alpha}_{\nu\beta} M^{\nu\beta}_{\rho\gamma} M^{\rho\gamma}_{\sigma\delta}$ can be represented by the square block matrix $\mathbf{W}^3$ of order NL. It is also straightforward to observe that the number of triangles $t(i, a)$ provided in formula 4.45 is given by the i-diagonal entry of the block a in $\mathbf{A}^3$, that is by $[(\mathbf{A}^3)_a]_{ii}$. In particular, the supradjacency matrix $\mathbf{F}$, corresponding to the adjacency tensor $F^{\mu\alpha}_{\nu\beta}$ of the complete multilayer network, is the square matrix of order NL having 1 in all positions but the diagonal entries, where it has 0.

Weighted undirected networks

In order to show how to represent coefficients in terms of supradjacency matrices, let us refer to node-level coefficient in formula 4.48 for the Barrat version of the clustering coefficient (formula 4.53). $C(i, a)$ can be expressed in terms of the supradjacency matrices $\mathbf{W}$ and $\mathbf{A}$ as

$$C(i, a) = \frac{[(\mathbf{W}\mathbf{A}^2)_a]_{ii}}{[(\mathbf{W}\mathbf{F}\mathbf{A})_a]_{ii}} \tag{D.1}$$

where $[(\mathbf{WA}^2)_a]_{ii}$ is the ii-entry of the a-diagonal block of the matrix $\mathbf{WA}^2$ and $[(\mathbf{WFA})_a]_{ii}$ is the ii-entry of the a-diagonal block of the matrix $\mathbf{WFA}$. Observe that the numerator of D.1 counts the number of actual triangles node i in the layer a belongs to (triangles laying completely in layer a or with vertices in other layers) weighted with the average weight of the two links on i. Notice that $[(\mathbf{WFA})_a]_{ii} = s_{i,a}(k_{i,a} - 1)$, where $k_{i,a}$ and $s_{i,a}$ are degree and strength of i in the layer a.

Formula D.1 is the natural extension of the classical representation, in matrix terms, of the local clustering coefficient for monoplex networks.

Following this idea, in the same way we can represent the clustering coefficient of the node i in the whole network (formula 4.49) for the Barrat coefficient:

$$C(i) = \frac{\sum_{a=1}^{L}[(\mathbf{WA}^2)_a]_{ii}}{\sum_{a=1}^{L}[(\mathbf{WFA})_a]_{ii}} \tag{D.2}$$

In this case it is easy to prove that $\sum_{a=1}^{L}[(\mathbf{WFA})_a]_{ii}$ is equal to $s_i(k_i - 1)$, where k_i and s_i are degree and strength of i in the whole multilayer network.

The clustering coefficient of level a over all its nodes (formula 4.50) for the Barrat coefficient is:

$$C(a) = \frac{\sum_{i=1}^{N}[(\mathbf{WA}^2)_a]_{ii}}{\sum_{i=1}^{N}[(\mathbf{WFA})_a]_{ii}} \tag{D.3}$$

In this case, $\sum_{i=1}^{N}[(\mathbf{WFA})_a]_{ii} = s_a(k_a - 1)$, where k_a and s_a are the total degree and strength on the level a.

Finally, the global clustering coefficient in formula 4.51 again for the Barrat version of the coefficient is given by:

$$C = \frac{\sum_{a=1}^{L}\sum_{i=1}^{N}[(\mathbf{WA}^2)_a]_{ii}}{\sum_{a=1}^{L}\sum_{i=1}^{N}[(\mathbf{WFA})_a]_{ii}} = \frac{\mathrm{tr}(\mathbf{WA}^2)}{\mathrm{tr}(\mathbf{WFA})} \tag{D.4}$$

Similarly, we express formula 4.52 as:

$$C(i,a) = \frac{[(\tilde{\mathbf{W}}^3)_a]_{ii}}{[(\tilde{\mathbf{W}}\mathbf{F}\tilde{\mathbf{W}})_a]_{ii}} \tag{D.5}$$

where $\tilde{\mathbf{W}} = \frac{1}{\mathcal{W}}\mathbf{W}$ and formula 4.54 as

$$C(i,a) = \frac{[(\hat{\mathbf{W}}^3)_a]_{ii}}{[(\mathbf{AFA})_a]_{ii}} \tag{D.6}$$

where $\hat{\mathbf{W}} = \tilde{\mathbf{W}}^{1/3}$.

Weighted directed networks

Consistently with notations in section 4.4.4, let us define: $\tilde{\mathbf{W}}_{\text{out}} = \frac{1}{W}\mathbf{W}$, $\tilde{\mathbf{W}}_{\text{in}} = \frac{1}{W}\mathbf{W}^T$ and $\tilde{\mathbf{W}} = \frac{1}{2}\big(\mathbf{W}_{\text{out}} + \mathbf{W}_{\text{in}}\big)$. Similarly, $\tilde{\mathbf{A}} = \frac{1}{2}\big(\mathbf{A}_{\text{out}} + \mathbf{A}_{\text{in}}\big)$.

The coefficient of formula 4.56 may be expressed as:

$$C(i,a) = \frac{[(\tilde{\mathbf{W}}\tilde{\mathbf{A}}^2)_a]_{ii}}{[(\tilde{\mathbf{W}}\mathbf{F}\tilde{\mathbf{A}})_a]_{ii}} \tag{D.7}$$

The numerator counts all the actual oriented triangles node i belongs to, on the layer a or with neighbours in other layers. Triangles are weighted with the average weight of the links connecting node i to its adjacent nodes. Notice that

$$[(\tilde{\mathbf{W}}\mathbf{F}\tilde{\mathbf{A}})_a]_{ii} = s^{\text{tot}}_{i,a}(k^{\text{tot}}_{i,a} - 1) - 2s^{\leftrightarrow}_{i,a}$$

where $k^{\text{tot}}_{i,a}$ and $s^{\text{tot}}_{i,a}$ are total degree and total strength of i in the layer a, whereas $s^{\leftrightarrow}_{i,a}$ is the strength of bilateral links on node i in layer a. The denominator represents all possible (appropriately weighted) directed triangles that node i could form.

Similarly, the coefficient of formulas (4.55) and 4.57 can be rewritten as:

$$C(i,a) = \frac{[(\tilde{\mathbf{W}}^3)_a]_{ii}}{[(\tilde{\mathbf{W}}\mathbf{F}\tilde{\mathbf{W}})_a]_{ii}} \tag{D.8}$$

$$C(i,a) = \frac{[(\hat{\mathbf{W}}^3)_a]_{ii}}{[(\mathbf{A}\mathbf{F}\mathbf{A})_a]_{ii}} \tag{D.9}$$

References

[1] E. Estrada. *The structure of complex networks: theory and applications*. Oxford University Press, 2012. 1, 8, 12, 21, 49, 150

[2] M.E. J. Newman. **A measure of betweenness centrality based on random walks**. *Social Networks*, **27**(1):39 – 54, 2005. 1, 8

[3] Hamed Amini, Rama Cont, and Andreea Minca. **RESILIENCE TO CONTAGION IN FINANCIAL NETWORKS**. *Mathematical Finance*, **26**(2):329–365, 2016. 2, 8

[4] R. Cont and A. Minca. **Credit default swaps and systemic risk**. *Annals of Operations Research*, **247**(2):523–547, 2016. 2, 8

[5] A. Lehar H. Elsinger and M. Summer. **Using market information for banking systems**. *Int. J. Central Bank.*, **27**:137–165, 2006. 2, 8

[6] Prasanna Gai and Sujit Kapadia. **Contagion in Financial Networks**. *Proc. Royal Soc. A*, **466**:2401–23, 2010. 2, 8

[7] Paolo Bartesaghi, Michele Benzi, Gian Paolo Clemente, Rosanna Grassi, and Ernesto Estrada. **Risk-Dependent Centrality in Economic and Financial Networks**. *SIAM Journal on Financial Mathematics*, **11**(2):526–565, 2020. 3

[8] G Fagiolo, J Nicoll Victor, M Lubell, and AH Montgomery. **The international trade network: Empirics and modeling**. *The Oxford Handbook of Political Networks*, pages 173–193, 2015. 3, 51

[9] Ernesto Estrada and Naomichi Hatano. **Communicability Graph and Community Structures in Complex Networks**. *Appl. Math. Comput.*, **214**(2):500–511, August 2009. 3, 52, 53

[10] E. Estrada and N. Hatano. *Resistance Distance, Information Centrality, Node Vulnerability and Vibrations in Complex Networks*. Springer, London, London, 2010. 3, 52, 53, 56, 58

[11] Mark EJ Newman and Michelle Girvan. **Finding and evaluating community structure in networks**. *Physical review E*, **69**(2):026113, 2004. 3, 52, 53, 61, 92

[12] C. Chang, W. Liao, Y. Chen, and L. Liou. **A Mathematical Theory for Clustering in Metric Spaces**. *IEEE Transactions on Network Science and Engineering*, **3**(1):2–16, 2016. 3, 4, 52, 61, 129

[13] Alex Arenas, Jordi Duch, Alberto Fernández, and Sergio Gómez. **Size reduction of complex networks preserving modularity**. *New Journal of Physics*, **9**(6):176, 2007. 4

[14] Mark EJ Newman. **Fast algorithm for detecting community structure in networks**. *Physical review E*, **69**(6):066133, 2004. 4, 53, 92

[15] M. Grötschel and Y. Wakabayashi. **A cutting plane algorithm for a clustering problem**. *Mathematical Programming*, **45**(1-3):59–96, 1989. 5, 91, 92

[16] Manlio De Domenico, Albert Solé-Ribalta, Emanuele Cozzo, Mikko Kivelä, Yamir Moreno, Mason A Porter, Sergio Gómez, and Alex Arenas. **Mathematical formulation of multilayer networks**. *Physical Review X*, **3**(4):041022, 2013. 6, 114, 120, 134, 136, 137

[17] Franklin Allen and Ana Babus. **Networks in Finance**. *History of Finance eJournal*, 2008. 7

[18] Stefano Battiston, Domenico Delli Gatti, Mauro Gallegati, Bruce Greenwald, and Joseph Stiglitz. **Liaisons dangereuses: Increasing connectivity, risk sharing, and systemic risk**. *Journal of Economic Dynamics and Control*, **36**(8):1121–1141, 2012. 7

[19] Paola Bongini, Gian Paolo Clemente, and Rosanna Grassi. **Interconnectedness, G-SIBs and network dynamics of global banking**. *Finance Research Letters*, **27**:185–192, 2018. 7

[20] Taiji Furusawa and Hideo Konishi. **Free trade networks**. *Journal of International Economics*, **72**(2):310 – 335, 2007. 7

[21] S. Goyal and J.L. Moraga-Gonzalez. **R & D Networks**. *The RAND Journal of Economics*, **32**(4):686–707, 2001. 7

[22] A. Kirman. **The economy as an evolving network**. *J. Evol. Econ.*, **7**(4):339–53, 1997. 7

[23] M. Fafchamps and F. Guberti. **The formation of risk sharing networks**. *J. Dev. Econ.*, (83(2)):326–350, 2007. 7

[24] Margarida Corominas-Bosch. **Bargaining in a network of buyers and sellers**. *Journal of Economic Theory*, **115**(1):35 – 77, 2004. 7

[25] Douglas M. Gale and Shachar Kariv. **Financial Networks**. *American Economic Review*, **97**(2):99–103, May 2007. 7

[26] **A theory of buyer-seller networks**. *Am. Econ. Rev.*, **91**(3, pages =). 7

[27] Nicholas Economides. **The economics of networks**. *International Journal of Industrial Organization*, **14**(6):673 – 699, 1996. 7

[28] Michael L. Katz and Carl Shapiro. **Systems Competition and Network Effects**. *Journal of Economic Perspectives*, **8**(2):93–115, June 1994. 7

[29] A. Galeotti and S. Goyal. **A Theory of Strategic Diffusion**. *FEEM Working Paper*, **70**, 2007. 7

REFERENCES

[30] Benjamin Golub and Matthew O. Jackson. **NaÃ¨ve Learning in Social Networks and the Wisdom of Crowds**. *American Economic Journal: Microeconomics*, **2**(1):112–49, February 2010. 7

[31] Dunia Lopez-Pintado. **Diffusion in complex social networks**. *Games and Economic Behavior*, **62**(2):573–590, 2008. 7

[32] Lauren Cohen, Andrea Frazzini, and Christopher Malloy. **The Small World of Investing: Board Connections and Mutual Fund Returns**. *Journal of Political Economy*, **116**(5):951–979, 2008. 7

[33] **Social networks in the boardroom**. *J. Eur. Econ. Assoc*, **11**(4, pages =). 7

[34] B. Nguyen-Dang. **Does the rolodex matter. Corporate Elite's Small World and the Effectiveness of Boards of Directors**. *SSRN Working Paper*, 2007. 7

[35] Prasanna Gai, Andrew Haldane, and Sujit Kapadia. **Complexity, concentration and contagion**. *Journal of Monetary Economics*, **58**(5):453 – 470, 2011. Carnegie-Rochester Conference on public policy: Normalizing Central Bank Practice in Light of the credit Turmoi, 12-13 November 2010. 8

[36] Masayasu Kanno. **Assessing systemic risk using interbank exposures in the global banking system**. *Journal of Financial Stability*, **20**:105 – 130, 2015. 8

[37] MarianoTirado. **Complex network for a crisis contagion on an interbank system**. *International Journal of Modern Physics C*, **23**, 09 2012. 8

[38] Stefano Battiston, Michelangelo Puliga, Rahul Kaushik, Paolo Tasca, and Guido Caldarelli. **DebtRank: Too central to fail? financial networks, the fed and systemic risk**. *Scientific reports*, **2**:srep00541, 2012. 8, 12

[39] Michael Boss, Helmut Elsinger, Martin Summer, and Stefan Thurner. **An Empirical Analysis of the Network Structure of the Austrian Interbank Market**. *Financial Stability Report*, (7):77–87, 2004. 8

[40] R. Seliger C. Puhr and M. Sigmund. **Contagiousness and vulnerability in the Austrian interbank market**. *Financial Stability Report - Oesterreichische National-bank*, **24**, 2012. 8

[41] João F. Cocco, Francisco Gomes, and Nuno C. Martins. **Lending relationships in the interbank market**. *Journal of Financial Intermediation*, **18**(1):24–48, 2009. 8

[42] A.G. Haldane and R.M. May. **Systemic Risk in Banking Ecosystems**. *Nature*, (469):351–355, 2011. 8, 9

[43] Paul Glasserman and H. Peyton Young. **How likely is contagion in financial networks?** *Journal of Banking & Finance*, **50**(C):383–399, 2015. 8

[44] Kyu-Min Lee, Jae-Suk Yang, Gunn Kim, Jaesung Lee, Kwang-Il Goh, and In-mook Kim. **Impact of the Topology of Global Macroeconomic Network on the Spreading of Economic Crises**. *PLOS ONE*, **6**(3):1–11, 03 2011. 8, 11

[45] Wenjun Mei, Shadi Mohagheghi, Sandro Zampieri, and Francesco Bullo. **On the dynamics of deterministic epidemic propagation over networks**. *Annual Reviews in Control*, **44**:116 – 128, 2017. 8, 14, 15

[46] Romualdo Pastor-Satorras, Claudio Castellano, Piet Van Mieghem, and Alessandro Vespignani. **Epidemic processes in complex networks**. *Rev. Mod. Phys.*, **87**:925–979, Aug 2015. 8

[47] Ernesto Estrada and Naomichi Hatano. **Communicability in complex networks**. *Phys. Rev. E*, **77**:036111, Mar 2008. 8, 12, 53, 54, 55, 124

[48] S.A. Levin R. M. May and G. Sugihara. **Complex systems: ecology for bankers**. *Nature*, **451**:893–895, 2008. 9

[49] Nikolaos Demiris, Theodore Kypraios, and L. Vanessa Smith. **On the epidemic of financial crises**. *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, **177**(3):697–723, 2014. 9

[50] Eloy Fisher. **A biological approach for financial network contagion based on the Susceptible - Infected - Recovered (SIR) model**. Number 28(69), pages 109–128, 2013. 9

[51] Robert Peckham. **Contagion: epidemiological models and financial crises**. *Journal of Public Health*, **36**(1):13–17, 08 2013. 9

[52] **The spread of a financial virus through Europe and beyond**. *AIMS Mathematics*, **4**(Math-04-01-086, pages =). 9

[53] A. Leontitsis D. Philippas, Y. Koutelidakis. **Insights into European interbank network contagion.** *Managerial Finance*, **10**, 2015. 9

[54] Antonios Garas, Panos Argyrakis, Céline Rozenblat, Marco Tomassini, and Shlomo Havlin. **Worldwide spreading of economic crisis**. *New Journal of Physics*, **12**(11):113043, nov 2010. 9

[55] A. Arenas P. Fleurquin J. Nin JJ. Ramasco E. TomÃ¡s A. Barja, A. MartÃnez. **Assessing the risk of default propagation in interconnected sectoral financial networks**. *EPJ Data Science, Dec 1, 8(1), 32*, 2019. 9

[56] D. Liuzzi S. Marsiglio A. Bucci, D. La Torre. **Financial contagion and economic development: An epidemiological approach**. *J. Econ. Behavior & Organization*, **162**:211–228, 2019. 9

[57] M. Nekovee, Y. Moreno, G. Bianconi, and M. Marsili. **Theory of rumour spreading in complex social networks**. *Physica A: Statistical Mechanics and its Applications*, **374**(1):457 – 470, 2007. 10

[58] I. Hull. **The development and spread of financial innovations**. *Quantitative Economics*, (7(2)):613–36, 2016. 10

[59] M. Toivanen. **Contagion in the interbank network: An epidemiological approach**. *Bank of Finland ResearchDiscussion Paper*, **23**, 2013. 10

[60] Chul-Ho Lee, Srinivas Tenneti, and Do Young Eun. **Transient Dynamics of Epidemic Spreading and its Mitigation on Large Networks**. *CoRR*, **abs/1903.00167**, 2019. 10, 11, 15, 16, 53

[61] M. E. J. Newman. **Spread of epidemic disease on networks**. *Phys. Rev. E*, **66**:016128, Jul 2002. 11

[62] Reuven Glick and Andrew Rose. **Contagion and trade: Why are currency crises regional?** *Journal of International Money and Finance*, **18**(4):603–617, 1999. 11

[63] Stefano Battiston, Domenico Delli Gatti, Mauro Gallegati, Bruce Greenwald, and Joseph Stiglitz. **Default cascades: When does risk diversification increase stability?** *Journal of Financial Stability*, **8**(3):138–149, 2012. 11

[64] P. Van Mieghem, K. Devriendt, and H. Cetinay. **Pseudoinverse of the Laplacian and best spreader node in a network**. *Physical Review E*, **96**, 09 2017. 11, 54, 58, 60

[65] Ernesto Estrada and Naomichi Hatano. **A vibrational approach to node centrality and vulnerability in complex networks**. *Physica A: Statistical Mechanics and its Applications*, **389**(17):3648 – 3660, 2010. 12, 54, 56

[66] Kydros Dimitrios and Oumbailis Vasileios. **A Network Analysis of the Greek Stock Market**. *Procedia Economics and Finance*, **33**:340 – 349, 2015. The Economies of Balkan and Eastern Europe Countries in the Changed World (EBEEC 2015). 12

[67] Delio Mugnolo. **Dynamical systems associated with adjacency matrices**. *Discrete & Continuous Dynamical Systems - B*, **23**:1945, 2018. 15

[68] T.M. Cover and J.A. Thomas. *Elements of Information Theory*. Wiley & Sons, 2006. 15

[69] Michele Benzi and Christine Klymko. **On the limiting behavior of parameter-dependent network centrality measures**. *SIAM J.Matrix Anal.Appl.*, **36**:686–706, 2015. 18, 20

[70] Ernesto Estrada and Juan Alberto Rodriguez-Velazquez. **Subgraph Centrality in Complex Networks**. *Physical review. E, Statistical, nonlinear, and soft matter physics*, **71**:056103, 06 2005. 21, 53, 55, 147

[71] Michele Benzi and Christine Klymko. **Total communicability as a centrality measure**. *Journal of Complex Networks*, **1**(2):124–149, 05 2013. 21

[72] Michele Benzi and Paola Boito. **Quadrature rule-based bounds for functions of adjacency matrices**. *Linear Algebra and its Applications*, **433**(3):637 – 652, 2010. 21

[73] P. Erdős and A. RÃšnyi. **On the evolution of random graphs**. *Publ. Math. Inst. Hungar. Acad. Sci.*, (5):17–61, 1960. 22

[74] P. Erdős and A. RÃšnyi. **On the strength of connectedness of a random graph**. *Acta Math. Hung.*, (12):261–267, 1961. 22

[75] T. Luczak S. Janson and A. Rucinski. *Random Graphs*. Wiley & Sons, 2000. 26

[76] V. H. Vu. **Spectral Norm of Random Matrices**. *Proceedings of the Thirty-seventh Annual ACM Symposium on Theory of Computing*, **STOC '05**, 2005. 26

[77] A. Knowles and R. Rosenthal. **Eigenvalue confinement and spectral gap for random simplicial complexes**. *Random Struct. Alg.*, **51**:506–537, 2017. 26

[78] Laszlo Erdős, Antti Knowles, Horng-Tzer Yau, and Jun Yin. **Spectral statistics of Erdős-RÃšnyi graphs I: Local semicircle law**. *Annals of Probability*, **41**(3B):2279–2375, 05 2013. 26

[79] R.N. Mantegna. **Hierarchical structure in financial markets**. *Eur. Phys. J. B.*, **11**(1):193–197, 1999. 29

[80] J.P. Onnela, A. Chakraborti, K. Kaski, J. Kertesz, and A. Kanto. **Dynamics of market correlations: Taxonomy and portfolio analysis**. *Physical Review E*, **68**(5):056110, 2003. 29, 30

[81] Gustavo Peralta and Abalfazl Zareei. **A network approach to portfolio selection**. *Journal of Empirical Finance*, **38**:157 – 180, 2016. 30

[82] T. Di Matteo F. Pozzi and T. Aste. **Spread of risk across financial markets: better to invest in the peripheries**. *Scientific reports*, 2013. 30

[83] Gerald F. Davis, Mina Yoo, and Wayne E. Baker. **The Small World of the American Corporate Elite, 1982-2001**. *Strategic Organization*, **1**(3):301–326, 2003. 30

[84] Ron Alquist, Rahul Mukherjee, and Linda Tesar. **Fire-sale FDI or Business as Usual?** Working Paper 18837, National Bureau of Economic Research, February 2013. 34

[85] Sebastian Edwards. *Capital Flows and the Emerging Economies: Theory, Evidence, and Controversies*. National Bureau of Economic Research, Inc, 2000. 34

[86] William Lazonick and Mary O'Sullivan. **Maximizing shareholder value: a new ideology for corporate governance**. *Economy and Society*, **29**(1):13–35, 2000. 35

[87] P. Fernandez and L. Reinoso. **Shareholder Value Creators in the S &P 500: Year 2003**. *SSRN Electronic Journal.*, (10.2139/ssrn.506102), 2004. 35, 36

[88] L. Katz. **A new status index derived from sociometric data analysis**. *Psychometrika*, **18**:39–43, 1953. 42

[89] K. Kloster. **Talk delivered at the International Conference on Industrial and Applied Mathematics (ICIAM 2019) and personal communication**. 2019. 43

[90] B. Yu Y. M. Chen W. Wang Z. G. Song Y. Hu Z. W. Tao J. H. Tian Y. Y. Pei F. Wu, S. Zhao and M. L. Yuan. **A new coronavirus associated with human respiratory disease in China**. *Nature*, **579**:265–269, 2020. 48

REFERENCES

[91] X. G. Wang B. Hu L. Zhang W. Zhang H. R. Si Y. Zhu B. Li C. L. Huang P. Zhou, X. L. Yang and H. D. Chen. **A pneumonia outbreak associated with a new coronavirus of probable bat origin.** *Nature*, **579**:270–273, 2020. 48

[92] S. Baker A. Gorbalenya and R. Baric. **Coronaviridae Study Group of the International Committee on Taxonomy of Viruses: The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2.** *Nature Microbiol.*, **3**(04), 2020. 48

[93] R. Balwin and B. Weder di Mauro. *Economics in the Time of COVID-19*. CEPR Press, London, 2020. 49

[94] Petter Holme and Jari Saramaki. **Temporal networks**. *Physics Reports*, **519**(3):97–125, 2012. 49

[95] Tilak Abeysinghe and Kristin Forbes. **Trade linkages and output-multiplier effects: A structural VAR approach with a focus on Asia**. *Review of International Economics*, **13**(2):356–375, 2005. 51

[96] Stéphane Dées and Arthur Saint-Guilhem. **The role of the United States in the global economy and its evolution over time**. *Empirical Economics*, **41**(3):573–591, 2011. 51

[97] Michael D Ward, John S Ahlquist, and Arturas Rozenas. **Gravity's rainbow: A dynamic latent space model for the world trade network**. *Network Science*, **1**(1):95–118, 2013. 51

[98] Andrea Lancichinetti and Santo Fortunato. **Community detection algorithms: a comparative analysis**. *Physical review E*, **80**(5):056117, 2009. 53

[99] G. Fagiolo. **Clustering in complex directed networks**. *Physical Review E*, **76**(2), 2007. 53, 90, 94, 138

[100] Matthew J Rattigan, Marc Maier, and David Jensen. **Graph clustering with network structure indices**. In *Proceedings of the 24th international conference on Machine learning*, pages 783–790. ACM, 2007. 53

[101] Gian Paolo Clemente and Rosanna Grassi. **Directed clustering in weighted networks: A new perspective**. *Chaos, Solitons & Fractals*, **107**:26–38, 2018. 53, 90, 94, 138

[102] R. Cerqueti, G. Ferraro, and A. Iovanella. **A new measure for community structure through indirect social connections**. *Expert Systems with Applications*, **114**:196–209, 2018. 53, 90

[103] Santo Fortunato. **Community detection in graphs**. *Physics reports*, **486**(3-5):75–174, 2010. 53, 89

[104] Santo Fortunato and Darko Hric. **Community detection in networks: A user guide**. *Physics reports*, **659**:1–44, 2016. 53

[105] V. A. Traag, R. Aldecoa, and J.C. Delvenne. **Detecting communities using asymptotical surprise**. *Physical Review E*, **92**(2):022816, 2015. 53

[106] C. Nicolini, C Bordier, and A. Bifone. **Community detection in weighted brain connectivity networks beyond the resolution limit**. *Neuroimage*, **146**(28-39), 2017. 53

[107] J. Van Lidth de Jeude, R. Di Clemente, G. Caldarelli, F. Saracco, and T. Squartini. **Reconstructing Mesoscale Network Structures**. *Complexity*, 2019. 53

[108] Ernesto Estrada. **Complex networks in the Euclidean space of communicability distances**. *Phys. Rev. E*, **85**:066122, Jun 2012. 53, 57, 129

[109] Douglas Klein and Milan Randic. **Resistance Distance**. *Journal of Mathematical Chemistry*, **12**:81–95, 12 1993. 53, 59

[110] Enrico Bozzo. **The Moore-Penrose inverse of the normalized graph Laplacian**. *Linear Algebra and its Applications*, **439**(10):3038 – 3043, 2013. 54, 57

[111] G. Ferraz de Arruda, A. Luiz Barbieri, P. M. RodrÃguez, F. A. Rodrigues, Y. Moreno, and L. da Fontoura Costa. **The role of centrality for the identification of influential spreaders in complex networks**. *Physical Review E*, **90**, 2014. 54, 90

[112] M Angeles Serrano and Marián Boguñá. **Topology of the world trade web**. *Physical Review E*, **68**(1):015101, 2003. 54, 89, 92

[113] Xiang Li, Yu Ying Jin, and Guanrong Chen. **Complexity and synchronization of the world trade web**. *Physica A: Statistical Mechanics and its Applications*, **328**(1-2):287–296, 2003. 54, 89

[114] Diego Garlaschelli and Maria I Loffredo. **Fitness-dependent topological properties of the world trade web**. *Physical review letters*, **93**(18):188701, 2004. 54, 89

[115] Diego Garlaschelli and Maria I Loffredo. **Structure and evolution of the world trade network**. *Physica A: Statistical Mechanics and its Applications*, **355**(1):138–144, 2005. 54, 89

[116] Diego Garlaschelli, Tiziana Di Matteo, Tomaso Aste, Guido Caldarelli, and Maria I Loffredo. **Interplay between topology and dynamics in the World Trade Web**. *The European Physical Journal B*, **57**(2):159–164, 2007. 54, 92

[117] Giorgio Fagiolo, Javier Reyes, and Stefano Schiavo. **On the topological properties of the world trade web: A weighted network analysis**. *Physica A: Statistical Mechanics and its Applications*, **387**(15):3868–3873, 2008. 54, 92

[118] M Angeles Serrano, Marián Boguñá, and Alessandro Vespignani. **Patterns of dominant flows in the world trade web**. *Journal of Economic Interaction and Coordination*, **2**(2):111–124, 2007. 54, 89

[119] Irena Tzekina, Karan Danthi, and Daniel N Rockmore. **Evolution of community structure in the world trade web**. *The European Physical Journal B*, **63**(4):541–545, 2008. 54, 92

[120] Giorgio Fagiolo, Javier Reyes, and Stefano Schiavo. **The evolution of the world trade web: a weighted-network analysis**. *Journal of Evolutionary Economics*, **20**(4):479–514, 2010. 54, 90

[121] Luca De Benedictis and Lucia Tajoli. **The world trade network**. *The World Economy*, **34**(8):1417–1454, 2011. 54, 90

[122] Florian Blõchl, Fabian J Theis, Fernando Vega-Redondo, and Eric O'N Fisher. **Vertex centralities in input-output networks reveal the structure of modern economies**. *Physical Review E*, **83**(4):046127, 2011. 54, 90

[123] David Snyder and Edward L Kick. **Structural position in the world system and economic growth, 1955-1970: A multiple-network analysis of transnational interactions**. *American journal of Sociology*, **84**(5):1096–1126, 1979. 54, 92

[124] Matteo Barigozzi, Giorgio Fagiolo, and Giuseppe Mangioni. **Identifying the community structure of the international-trade multi-network**. *Physica A: statistical mechanics and its applications*, **390**(11):2051–2066, 2011. 54, 84, 86, 89, 90, 92

[125] Allen Wilhite. **Bilateral trade and 'small-world' networks**. *Computational Economics*, **18**(1):49–64, 2001. 54

[126] Javier Reyes, Stefano Schiavo, and Giorgio Fagiolo. **Assessing the evolution of international economic integration using random walk betweenness centrality: The cases of east asia and latin america**. *Advances in Complex Systems*, **11**(05):685–702, 2008. 54

[127] Stefano Schiavo, Javier Reyes, and Giorgio Fagiolo. **International trade and financial integration: a weighted network analysis**. *Quantitative Finance*, **10**(4):389–399, 2010. 54, 92

[128] Giorgio Fagiolo, Tiziano Squartini, and Diego Garlaschelli. **Null models of economic networks: the case of the world trade web**. *Journal of Economic Interaction and Coordination*, **8**(1):75–107, 2013. 54

[129] Ying Fan, Suting Ren, Hongbo Cai, and Xuefeng Cui. **The state's role and position in international trade: A complex network perspective**. *Economic Modelling*, 2014. 54

[130] Luis M Varela, Giulia Rotundo, Marcel Ausloos, and Jesús Carrete. **Complex network analysis in socioeconomic models**. In *Complexity and Geographical Economics*, pages 209–245. Springer, 2015. 54, 90

[131] Paolo Giudici and Alessandro Spelta. **Graphical network models for international financial flows**. *Journal of Business & Economic Statistics*, **34**(1):128–138, 2016. 54

[132] L. De Benedictis and L. Tajoli. **Comparative Advantage and Centrality in the World Network of Trade and Value Added: An Analysis of the Italian Position**. *Rivista di Politica Economica*, **66**((3)), 2016. 54, 90

[133] Freddy Cepeda-López, Fredy Gamboa-Estrada, Carlos León, and Hernán Rincón-Castro. **The evolution of world trade from 1995 to 2014: A network approach**. *The Journal of International Trade & Economic Development*, **28**(4):452–485, 2019. 54, 90

[134] Roy Cerqueti, Gian Paolo Clemente, and Rosanna Grassi. **A Network-Based Measure of the Socio-Economic Roots of the Migration Flows**. *Social Indicators Research*, **146**, 11 2019. 54

[135] David A Smith and Douglas R White. **Structure and dynamics of the global economy: network analysis of international trade 1965–1980**. *Social forces*, **70**(4):857–893, 1992. 54, 92

[136] Raja Kali and Javier Reyes. **The architecture of globalization: a network approach to international economic integration**. *Journal of International Business Studies*, **38**(4):595–620, 2007. 54, 92

[137] Carlo Piccardi and Lucia Tajoli. **Complexity, centralization, and fragility in economic networks**. *PLOS ONE*, **13**:1–13, 11 2018. 54

[138] Ricardo Hausmann, César A Hidalgo, Sebastián Bustos, Michele Coscia, Alexander Simoes, and Muhammed A Yildirim. *The atlas of economic complexity: Mapping paths to prosperity*. Mit Press, 2014. 54, 77, 108

[139] Ivan Gutman and W Xiao. **Generalized inverse of the Laplacian matrix and some applications**. *Bulletin (Académie serbe des sciences et des arts. Classe des sciences mathématiques et naturelles. Sciences mathématiques)*, pages 15–23, 2004. 57, 58

[140] W. Ellens, F.M. Spieksma, P. Van Mieghem, A. Jamakovic, and R.E. Kooij. **Effective graph resistance**. *Linear algebra and its applications*, pages 2491–2506, 2011. 59

[141] X. Wang, E. Pournaras, R.E. Kooij, and P. Van Mieghem. **Improving Robustness of Complex Networks via the Effective Graph Resistance**. *The European Physical Journal B*, **87**(9):221, 2014. 59

[142] Gian Paolo Clemente and Alessandra Cornaro. **A novel measure of edge and vertex centrality for assessing robustness in complex networks**. *Soft Computing*, 2019. 59

[143] S Van Berkum. *Trade effects of the EU-Morocco Association Agreement*. LEI, onderdeel van Wageningen UR, 2013. 70

[144] Natalia Victorovna Kuznetsova, Ekaterina Victorovna Kocheva, and Nikolay Anatolievich Matev. **The analysis of foreign trade activities of Russia and Asia-Pacific region**. *International Journal of Economics and Financial Issues*, **6**(2):736–744, 2016. 71

[145] WTO. **World Trade Statistical Review**. Technical report, World Trade Organizations, 2017. 77, 82

[146] Zhen Zhu, Federica Cerina, Alessandro Chessa, Guido Caldarelli, and Massimo Riccaboni. **The rise of China in the international trade network: a community core detection approach**. *PloS one*, **9**(8):e105496, 2014. 79, 87

REFERENCES

[147] G. Kozmetsky and P. Yue. *Global Economic Competition: Today's Warfare in Global Electronics Industries and Companies*. Springer US, 2012. 80

[148] Rita Maria del Rio-Chanona, Jelena Grujic, and Henrik Jeldtoft Jensen. **Trends of the World Input and Output Network of Global Trade**. *PLOS ONE*, **12**(1):1–14, 01 2017. 80

[149] Carlo Piccardi and Lucia Tajoli. **Existence and significance of communities in the World Trade Web**. *Phys. Rev. E*, **85**:066119, Jun 2012. 84, 85, 86

[150] Carlo Piccardi. **Finding and Testing Network Communities by Lumped Markov Chains**. *PLOS ONE*, **6**:1–13, 11 2011. 89

[151] M. E. J. Newman. *Networks: an introduction*. Oxford university press, 2010. 90

[152] S. Wasserman and K. Faust. *Social Network Analysis: Methods and Applications*. Cambridge University Press, New York, NY., July 1994. 90

[153] D. J. Watts and S. H. Strogatz. **Collective dynamics of 'small-world 'networks**. *Nature*, **393**(6684):440–442, 1998. 90

[154] A. Barrat, M. Barthélemy, R. Pastor-Satorras, and A. Vespignani. **The architecture of complex weighted networks**. *Proceedings of the National Academy of Sciences*, **101**(11):3747–3752, 2004. 90, 136, 137

[155] J.P. Onnela, J. Saramäki, J. Kertész, and K. Kaski. **Intensity and coherence of motifs in weighted complex networks**. *Physical Review E*, **71**(6), 2005. 90, 136, 137

[156] Giulia Rotundo and Marcel Ausloos. **Organization of networks with tagged nodes and biased links: A priori distinct communities: The case of intelligent design proponents and Darwinian evolution defenders**. *Physica A: Statistical Mechanics and its Applications*, **389**(23):5479–5494, 2010. 90

[157] M. Grötschel and Y. Wakabayashi. **Facets of the clique partitioning polytope**. *Mathematical Programming*, **47**(1-3):367–387, 1990. 91, 92

[158] S.G. de Amorim, J.-P. Barthélemy, and C.C. Ribeiro. **Clustering and clique partitioning: Simulated annealing and tabu search approaches**. *Journal of Classification*, **9**(1):17–41, 1992. 91, 92

[159] H. Wang, T. Obremski, B. Alidaee, and G. Kochenberger. **Clique partitioning for clustering: A comparison with K-means and latent class analysis**. *Communications in Statistics: Simulation and Computation*, **37**(1):1–13, 2008. 91

[160] Sangmoon Kim and Eui-Hang Shin. **A longitudinal analysis of globalization and regionalization in international trade: A social network approach**. *Social forces*, **81**(2):445–468, 2002. 92

[161] Raja Kali and Javier Reyes. **Financial contagion on the international trade network**. *Economic Inquiry*, **48**(4):1072–1101, 2010. 92

[162] Xiaohang Zhang, Huiyuan Cui, Ji Zhu, Yu Du, Qi Wang, and Wenhua Shi. **Measuring the dissimilarity of multiplex networks: An empirical study of international trade networks**. *Physica A: Statistical Mechanics and its Applications*, **467**:380–394, 2017. 92

[163] Rafael Santiago and LuÃs C. Lamb. **Efficient modularity density heuristics for large graphs**. *European Journal of Operational Research*, **258**(3):844 – 865, 2017. 92

[164] Aaron Clauset, Mark EJ Newman, and Cristopher Moore. **Finding community structure in very large networks**. *Physical review E*, **70**(6):066111, 2004. 92

[165] Vincent D Blondel, Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre. **Fast unfolding of communities in large networks**. *Journal of statistical mechanics: theory and experiment*, **2008**(10):P10008, 2008. 92

[166] Leon Danon, Albert Díaz-Guilera, and Alex Arenas. **The effect of size heterogeneity on community identification in complex networks**. *Journal of Statistical Mechanics: Theory and Experiment*, **2006**(11):P11010, 2006. 92

[167] Daniel Aloise, Sonia Cafieri, Gilles Caporossi, Pierre Hansen, Sylvain Perron, and Leo Liberti. **Column generation algorithms for exact modularity maximization in networks**. *Physical Review E*, **82**(4):046112, 2010. 92

[168] Jeffrey Pattillo, Nataly Youssef, and Sergiy Butenko. **On clique relaxation models in network analysis**. *European Journal of Operational Research*, **226**(1):9 – 18, 2013. 92

[169] S. Butenko and W.E. Wilhelm. **Clique-detection models in computational biochemistry and genomics**. *European Journal of Operational Research*, **173**:1 – 17, 2006. 92

[170] A. Mehrotra and M.A. Trick. **Cliques and clustering: A combinatorial approach**. *Operations Research Letters*, **22**(1):1–12, 1998. 92

[171] Rachid Chelouah and Patrick Siarry. **Tabu Search applied to global optimization**. *European Journal of Operational Research*, **123**:256 – 270, 2000. 92

[172] U. Brandes and T. Erlebach. *Network Analysis. Methodological Foundations*. Springer, 2005. 93

[173] Jon M Kleinberg. **Authoritative sources in a hyperlinked environment**. *Journal of the ACM (JACM)*, **46**(5):604–632, 1999. 94, 95

[174] Mirjana Lazić. **On the Laplacian energy of a graph**. *Czechoslovak Mathematical Journal*, **56**(4):1207–1213, Dec 2006. 95, 96

[175] Ivan Gutman and Bo Zhou. **Laplacian energy of a graph**. *Linear Algebra and its applications*, **414**(1):29–37, 2006. 95

[176] Xingqin Qi, Eddie Fuller, Qin Wu, Yezhou Wu, and Cun-Quan Zhang. **Laplacian centrality: A new centrality measure for weighted networks**. *Information Sciences*, **194**:240–253, 2012. 95, 96

[177] W.H. Haemers. **Interlacing eigenvalues and graphs**. *Linear Algebra and its Applications,*, **226-228**:593–616, 1995. 95

[178] D. Baruah and A. Bharali. **A Comparative Study of Vertex Deleted Centrality Measures**. *Annals of Pure and Applied Mathematics*, **14**(1):199–205, 2017. 96

[179] C. Adiga and M. Smitha. **On the skew Laplacian energy of a digraph**. *International Mathematical Forum*, **4**(3):1907–1914, 2009. 96

[180] P. Kissani and Y. Mizoguchi. **Laplacian energy of directed graphs and minimizing maximum out-degree algorithms**. Technical report, Kyushu University Institutional Repository, 2010. 96

[181] Renato Cordeiro de Amorim and Boris Mirkin. **Minkowski metric, feature weighting and anomalous cluster initializing in K-Means clustering**. *Pattern Recognition*, **45**:1061–1075, 2012. 97

[182] Cynthia Rudin. **The p-norm push: A simple convex ranking algorithm that concentrates at the top of the list**. *Journal of Machine Learning Research*, **10**(Oct):2233–2271, 2009. 97

[183] Michael J Brusco and Hans-Friedrich Köhn. **Clustering qualitative data based on binary equivalence relations: neighborhood search heuristics for the clique partitioning problem**. *Psychometrika*, **74**(4):685, 2009. 99

[184] S. Gaulier, G.; Zignago. **BACI: International Trade Database at the Product-Level. The 1994-2007 Version**. Technical Report 2010-23, CEPII„ 2010. 101

[185] Ginestra Bianconi. **Statistical mechanics of multiplex networks: Entropy and overlap**. *Phys. Rev. E*, **87**:062806, Jun 2013. 113

[186] S. Gómez, A. Díaz-Guilera, J. Gómez-Gardeñes, C. J. Pérez-Vicente, Y. Moreno, and A. Arenas. **Diffusion Dynamics on Multiplex Networks**. *Phys. Rev. Lett.*, **110**:028701, Jan 2013. 113

[187] Charles D. Brummitt, Kyu-Min Lee, and K.-I. Goh. **Multiplexity-facilitated cascades in networks**. *Phys. Rev. E*, **85**:045102, Apr 2012. 113

[188] Kyu-Min Lee, Jung Yeol Kim, Won kuk Cho, K-I Goh, and I-M Kim. **Correlated multiplexity and connectivity of multiplex random networks**. *New Journal of Physics*, **14**(3):033027, mar 2012. 113

[189] T. Levi-Civita, E. Persico, and M. Long. *The Absolute Differential Calculus (calculus of Tensors)*. Absolute (DC Comics). Dover Publications, 1977. 115

[190] M. E. J. Newman. **The structure and function of complex networks.** *SIAM Review*, **6**:28384, June 2003. 121

Declaration

I herewith declare that I have produced these papers without the prohibited assistance of third parties and without making use of aids other than those specified; notions taken over directly or indirectly from other sources have been identified as such. These papers have not previously been presented in identical or similar form to any other examination board.

The thesis work was conducted from 2018 to 2020 under the supervision of Rosanna Grassi at Università degli Studi di Milano Bicocca.

Milan, October 2020