NAVIGABILITY AND SYNCHRONIZATION IN COMPLEX NETWORKS: A COMPUTATIONAL APPROACH

Pau Erola Cañellas

Pau Erola Cañellas

NAVIGABILITY AND SYNCHRONIZATION IN COMPLEX NETWORKS

A COMPUTATIONAL APPROACH

PH.D. DISSERTATION

directed by Dr. Sergio Gómez Jiménez
Dr. Alejandro Arenas Moreno

Department of
Computer Engineering and Mathematics

UNIVERSITAT ROVIRA i VIRGILI

Tarragona
2016

UNIVERSITAT ROVIRA I VIRGILI
NAVIGABILITY AND SYNCHRONIZATION IN COMPLEX NETWORKS: A COMPUTATIONAL APPROACH
Pau Erola Cañellas

Departament d'Enginyeria
[DΣIM]
Informàtica i Matemàtiques

Departament d'Enginyeria Informàtica i Matemàtiques
Campus Sescelades
Av. dels Països Catalans, 26
43007 Tarragona
Tel. 977 559 703
Fax 977 559 710

We STATE that the present study, entitled "Navigability and synchronization in complex networks: a computational approach", presented by Pau Erola Cañellas for the award of the degree of Doctor, has been carried out under our supervision at the Department of Computer Engineering and Mathematics of this University, and that it fulfills all the requirements to receive the International Doctor distinction.

Tarragona, the 7th of January 2016

Doctoral Thesis Supervisors

Dr. Sergio Gómez Jiménez

Dr. Alejandro Arenas Moreno

UNIVERSITAT ROVIRA I VIRGILI
NAVIGABILITY AND SYNCHRONIZATION IN COMPLEX NETWORKS: A COMPUTATIONAL APPROACH
Pau Erola Cañellas

Acknowledgments

"Italians have a little joke, that the world is so hard a man
must have two fathers to look after him, and that's why they have godfathers."
– Mario Puzo, The Godfather

First of all, I have to thank the *godfathers* of this thesis, Sergio and Alex, for their guidance and immense patience. You both have discovered to me a new world of science I have never imagined.

Secondly, I am very grateful to my hosts in the inspiring University of Oxford. Sang Hoon and Mason, thank you for your teachings, thoughts and *moments*. I learned a lot with both of you.

I also would like to give my thanks to my co-authors Albert, Javi and Raúl, and to Sandra, Gemma, Roser and Juanjo for their kindliness, kindness and encouragements during this long-tail journey.

And last but not least, to my beloved Anna and Cua, because even the deeper dejection disappears with a hug and a few tail wags.

UNIVERSITAT ROVIRA I VIRGILI
NAVIGABILITY AND SYNCHRONIZATION IN COMPLEX NETWORKS: A COMPUTATIONAL APPROACH
Pau Erola Cañellas

Contents

UNIVERSITAT ROVIRA I VIRGILI
NAVIGABILITY AND SYNCHRONIZATION IN COMPLEX NETWORKS: A COMPUTATIONAL APPROACH
Pau Erola Cañellas

List of Tables

UNIVERSITAT ROVIRA I VIRGILI
NAVIGABILITY AND SYNCHRONIZATION IN COMPLEX NETWORKS: A COMPUTATIONAL APPROACH
Pau Erola Cañellas

List of Figures

UNIVERSITAT ROVIRA I VIRGILI
NAVIGABILITY AND SYNCHRONIZATION IN COMPLEX NETWORKS: A COMPUTATIONAL APPROACH
Pau Erola Cañellas

Abstract

Complex networks are a powerful tool to study many real systems, partly thanks to the increasing capacity of computational resources. In this dissertation we address computationally a broad scope of problems that are framed in two parts.

The first part is motivated by the issues posed by the rapid evolution of the Internet. On one side, the exponential growth of the network is compromising its scalability due to dependencies on global routing tables. In Chapter 4 we propose a decentralized routing scheme that exploits the TSVD projection of the mesoscopic structure of the network as a map. The results show that, using only local information, we can achieve good success rates in the routing process. Additionally, Chapter 3 evaluates the reliability of this projection when network topology changes. The results indicate that this map is very robust and does not need continual updates.

On the other side, the increasing bandwidth demand is a potential trigger for congestion episodes. In Chapter 5 we extend a dynamic traffic-aware routing scheme to the context of multiplex networks, and we conduct the analysis on synthetic networks with different coupling assortativity. The results show that considering the traffic load in the transmission process delays the congestion onset. However, the uniform distribution of traffic produces an abrupt phase transition from free-flow to congested state. Withal, assortative coupling is depicted as the best consideration for optimal network designs.

The second part is motivated by the current global financial crises. Chapter 6 presents a study on the spreading of economic crises using a simple model of networked integrate-and-fire oscillators and we characterize synchronization process on the evolving trade network. The results show the emergence of a globalization process that dilutes the topological borders and accelerates the spreading of financial crashes.

UNIVERSITAT ROVIRA I VIRGILI
NAVIGABILITY AND SYNCHRONIZATION IN COMPLEX NETWORKS: A COMPUTATIONAL APPROACH
Pau Erola Cañellas

Resum

Les xarxes complexes han demostrat ser una eina molt valuosa per estudiar sistemes reals, en part, gràcies a la creixent capacitat de computació. En aquesta tesi abordem computacionalment diversos problemes dividits en dos blocs.

El primer bloc està motivat pels problemes que planteja la ràpida evolució de la Internet. D'una banda, el creixement exponencial de la xarxa està comprometent la seva escalabilitat per les dependències a les taules d'enrutament globals. Al Capítol 4 proposem un esquema d'enrutament descentralitzat que fa servir la projecció TSVD de l'estructura mesoscòpica de la xarxa com a mapa. Els resultats mostren que fent servir informació local podem guiar amb èxit en l'enrutament. Al Capítol 3 també avaluem la fiabilitat d'aquesta projecció davant el creixement de la xarxa. Els resultats indiquen que aquest mapa és robust i no necessita actualitzacions contínues.

D'altra banda, la creixent demanda d'ample de banda és un factor potencial per produir congestió. Al Capítol 5 estenem un esquema d'enrutament dinàmic en el context de les xarxes *multiplex*, i l'analitzem amb xarxes sintètiques amb diferents assortativitats d'acoblament. Els resultats mostren que tenir en compte el volum de trànsit en l'enrutament retarda l'inici de la congestió. Tot i això, la distribució uniforme del trànsit produeix una transició de fase abrupta. Amb tot, l'acoblament assortatiu es presenta com la millor opció per a dissenys de xarxes òptimes.

El segon bloc ve motivat per l'actual crisi financera mundial. Al Capítol 6 proposem estudiar la propagació de les crisis econòmiques utilitzant un model simple de xarxa formada per oscil·ladors *integrate-and-fire*, i caracteritzar la seva sincronització durant l'evolució de la xarxa de comerç. Els resultats mostren l'aparició d'un procés de globalització que dilueix les fronteres topològiques i accelera la propagació de les crisis financeres.

UNIVERSITAT ROVIRA I VIRGILI
NAVIGABILITY AND SYNCHRONIZATION IN COMPLEX NETWORKS: A COMPUTATIONAL APPROACH
Pau Erola Cañellas

A la mama i el papa.

Als avis.

UNIVERSITAT ROVIRA I VIRGILI
NAVIGABILITY AND SYNCHRONIZATION IN COMPLEX NETWORKS: A COMPUTATIONAL APPROACH
Pau Erola Cañellas

1 Introduction

"I think the next century will be the century of complexity."
– Stephen Hawking, January 23, 2000

After Barabási and Albert presented their model [20] in 1999 the scientific community began to see networks everywhere. Its strong predictive capabilities in all fields of science has made networks earn the label of ubiquitous. But beyond the mathematical model, networks represent a paradigm shift in science, from *individual* to *collective*, where "the whole becomes not only more than the sum of but very different from the sum of the parts" [7]. No matter if you study power grids, financial markets, telephone systems, the Internet, human genetics, ant colonies or social clubs, now you will be in front of a system of entities and their interconnections, a network. Nowadays, network science constitutes one of the most successful and multidisciplinary research areas.

The interest of computer engineering in networks has been strongly motivated by the technological development itself, with the emergence of the Internet, the World Wide Web (WWW) and social networks. But also because network science is a heavily computational subject [166] with big amounts of data gathered but without analytical solutions for all the arising questions. In this document will use computational tools to answer, hopefully, a few of them.

This first chapter wants to provide a gentle introduction to the topics that encompass this thesis, as well as its motivation and contributions. In Section 1.1 we briefly introduce network science. Then we are going to focus on the navigability (Section 1.2) and

synchronization (Section 1.3) in complex networks. Finally, Section 1.4 will describe shortly each chapter's contributions.

1.1 Introduction to network science

In 1735, the Swiss mathematician and physicist Leonhard Euler proposed a solution to the Königsberg bridges problem (see Fig. 1.1) establishing the seed of graph theory [5], the primary analytical method used to study networks. He reformulated the problem in abstract terms, eliminating all features except the list of landmasses and bridges. This simplified abstraction proved to be useful for many problems, and in the 19th century scientists like Thomas Kirkman, William Hamilton, Gustav Kirchhoff and Arthur Cayley made further advances in this field, primary on topological measures of graph properties [90].

Social scientists started using these tools at the beginning of the 20th century to study the relationships among social entities, members of a group, trade and economic transactions between corporations [33]. Particularly interesting, it was in 1967 when the American social psychologist Stanley Milgram published the results of his famous small-world experiment [138, 208] that showed the existence of short paths between apparently distant individuals [150]. Some years later, Watts and Strogatz saw that this phenomenon was not unique to social networks stating that

> "Ordinarily, the connection topology is assumed to be either completely regular or completely random. But many biological, technological and social networks lie somewhere between these two extremes ... We find that these systems can be highly clustered, like regular lattices, yet have small characteristic path lengths, like random graphs. We call them 'small-world' networks, by analogy with the small-world phenomenon (popularly known as six degrees of separation)" [212],

and they proposed the so called Watts-Strogatz (WS) network model. Those networks with non-trivial topological features were called *complex network*.

Nevertheless, network theory gained much more popularity one year later, after the introduction of the Barabási-Albert (BA) model of preferential attachment [20]. Barabási and Albert observed also that the distribution in the number of connections of each element of the network is highly heterogeneous, with a large majority of nodes with few connections and a reduced number of nodes highly connected, and presented a mechanism

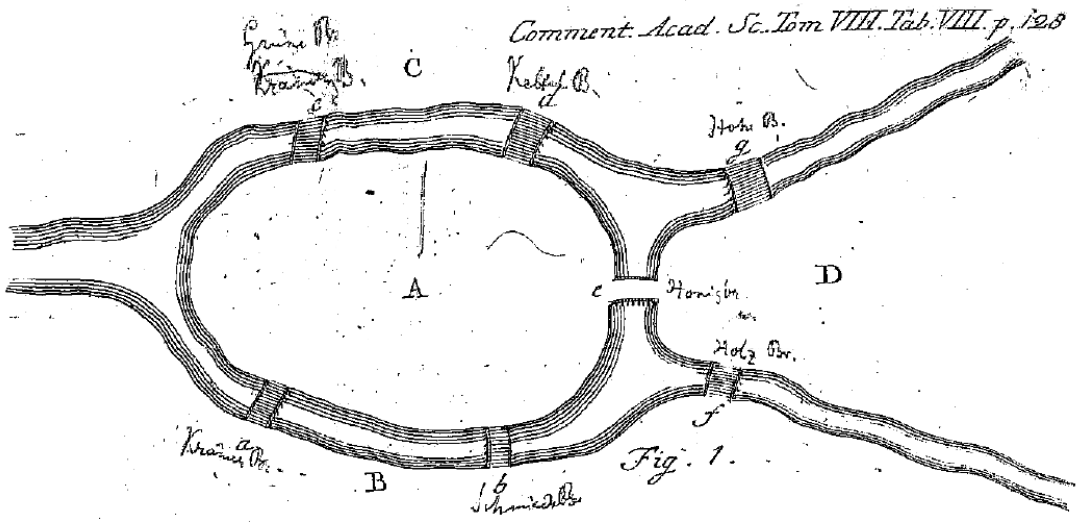


Figure 1.1: The seven bridges of Königsberg problem consists in to find a walk through the city, set on both sides of the Pregel River, that crosses each bridge once and only once. Euler proposed an abstraction where each landmass is a vertex, and each bridge is represented by an edge. Map representation from [71].

explaining the emergence of this scale-free property. Surprisingly, BA model fitted very well in real-world systems as diverse as metabolic and protein-protein interaction networks [29], the human brain [194], transportation networks [18], the Internet [162] or the collaborations between scientists [91], earning the label of pervasive.

Since then, a lot of active research has been done to discover the basic principles of networked systems, and to elucidate the relation between structure and dynamics and their function. But networks are inherently difficult to understand because its structure is complex and can evolve over time, connections can have diverse weights, directions and signs, and there may be different kinds of nodes or they could be nonlinear dynamical systems, or any meta-complication such, for instance, how does the structure affect dynamics on top of networks [197]. Of course, there is no unique answer to this question, since complex networks respond differently to disparate dynamical processes [146]. Here we will have to cope with several of these complications.

We have divided this document in two parts. First part of this thesis will revolve around the design of schemas for packet delivery in communications networks, considering its structure and evolution, and different dynamics. The second part will study the dynamics of a synchronization process on an economical network, the structure of which varies over time due to the globalization process. The following sections detail both issues.

1.2 Navigability in complex networks

We understand navigability as the quality of being suitable in transportation, measured by system and actors effectiveness and efficiency in the delivery of information or transportation of persons and goods. Obviously, it is a crucial quality of networked infrastructures such as the Internet, power grids and road systems.

The optimization of these communication and transportation networks has attracted much interest. In transportation networks, many efforts have focused on the design of efficient systems from the perspective of their structural characteristics [18], the pedestrian and vehicle traffic [44, 109], and their sustainability [137]. On the other hand, in communication networks, Tadić and Rodgers [199] studied the kinetic properties of packet transport on scale-free networks, and Guimerà, Díaz-Guilera, Vega-Redondo, Cabrales and Arenas [103] characterized the optimal network topologies for local search. Congestion phenomena has also been studied extensively [54, 95, 142], and several adaptive routing strategies have been proposed to avoid, or delay, it [61, 62, 89]. In the first part of this document we will study communication networks, with special focus on the Internet, without loss of generality.

The Internet has risen as a self-organized worldwide communication network, having a high impact in current society. Its rapid evolution and success have spurred fundamental research on network topology, and its easy accessibility has also attracted the research on traffic flows. It is noteworthy, that this large network lacks a centralized administration and design, working in a semi-cooperative fashion through protocols that provide a common substrate for communication. However, those protocols were not designed for a network with an exponential growth [162], and they are currently compromising the scalability of the Internet [136]. This issue motivates part of our work.

In 2000, Kleinberg pointed out that in many complex networks, in addition to have small-world structure, “individuals operating with purely local information are very adept at finding these chains” [117], the paths to distant nodes. Thus, they are easily navigable. Based on this premise, we will propose and evaluate a novel routing scheme for the Internet based on network theory that, in the same way described above, forwards packets using only local information. Further, we will characterize our solution under the network growth to ensure its scalability.

Another important aspect for the efficient performance in communication networks is the ability to avoid congestion. Congestion arises when the number of packets to be processed by a node exceeds its limited capacity [59, 103] and it affects the travel time of

the paths that pass through [187]. An example of this occurred in 1998, when connections to ".net" domain failed for several hours due to a misconfigured critical database server that incorrectly referred all queries to a wrong secondary server [124].

As anticipated, several works have studied congestion in the context of monoplex complex networks. However, this traditional abstraction can not take into account the complicated patterns that show many natural and engineered systems, that encompass multiple types of relationships and changes in time [116], and it can lead to misleading results. For this reason, last few years network concepts have been generalized to a new multilayer framework, but a lot of work remains to be done.

In this multilayer context, only a few works have addressed the congestion phenomena. Zhou, Yan and Lai [220] have studied the load balancing in communication systems made of two heterogeneous layers, a wired layer and a wireless layer. Tan, Wu, Xia and Tse [203] have explored the effect of coupling assortativity and probability in network performance. A more theoretical work has been carried out by Solé-Ribalta, Granell, Gómez and Arenas [193], where they have studied the diffusion of information in multiplex networks with community structure. In this point, we will contribute to the understanding of "networks of networks" with the assessment of adaptive routing strategies designed to avoid congestion.

1.3 Synchronization in complex networks

Synchronization is the adjustment of rhythms of objects whose internal state evolves periodically in time, due to their weak interaction [164]. The achievement of a final synchronized state, if any, strongly depends on the interaction pattern of the system [12].

Synchronization processes are encountered in circadian rhythms, ecology, neuronal networks, wireless communications, power grids, financial systems, or even networks of opinion formation, to name a few. So it is not surprising that many efforts have devoted to understand this phenomena, and recent complex network theory has played an important role in this. For instance, several studies have characterized synchronization patterns of networks with homogeneous and heterogeneous degree distributions of pulse-coupled oscillators [40, 176, 207] and Kuramoto oscillators [99]. Complex networks display enhanced synchronizability suggesting that perhaps there are some fundamental relationships between traffic and diffusion dynamics and synchronization [144, 211]. Others have studied synchronization under networks whose structure evolve in time [158, 168, 169, 210].

Current interest in the economic literature has focused also in the study of synchronization of economic cycles, the rises and falls in stock options, economic sectors, Gross Domestic Product (GDP) or other economic parameters [105, 114, 115]. In the last part of this dissertation we will aim our attention to these synchronization phenomena using real data about trade markets as a substrate of potential interaction channels between countries. The evolution of these interactions over time, due to globalization effects, will allow us to present a case study where we hypothesize about financial crises spreading. Despite the growing interest in understanding financial crises we still know very little about them, and our theoretical work will contribute to their characterization.

1.4 Thesis outline

This dissertation is organized as follows. First, Chapter 2 is devoted to provide an overview of concepts and mathematical techniques on complex networks. This is not intended to be an exhaustive review, but to describe those points required for the rest of the document.

Next two chapters explore jointly an alternative routing approach intended to address the scalability issues of the Internet. In Chapter 3 we present a technique based on projection theory to create maps with structural information of the network. We also evaluate the reliability of these maps when are used in growing scale-free networks and we compare them with other recent proposals.

Chapter 4 starts with a review of previous efforts in the design of routing schemes for the Internet. Then we propose a novel routing algorithm that uses local information provided by the previously studied maps. At last, we asses its performance in actual snapshots of the Internet network.

The material presented in Chapter 5 is inspired by the emergence of multiplex networks and the increasing demands for bandwidth on the Internet. We present an extension of an adaptive routing strategy to control congestion in traditional single-layer networks, and we evaluate its performance in synthetic multiplex networks composed of two scale-free layers. Our proposal uses topological information and data about the traffic status in a local area.

In Chapter 6 we turn our attention to the synchronization processes of economic cycles. There we study the possible effects of the globalization of trade markets in the spreading of financial crises. Using a model where each country is represented by a pulse-coupled

oscillator, we characterize the synchronization behavior of the system depending on how network topology has evolved in time.

Finally, Chapter 7 concludes this dissertation with a brief summary of our results.

UNIVERSITAT ROVIRA I VIRGILI
NAVIGABILITY AND SYNCHRONIZATION IN COMPLEX NETWORKS: A COMPUTATIONAL APPROACH
Pau Erola Cañellas

2 Concepts of network theory

This chapter is devoted to provide a brief introduction to the basic concepts, measures and models in network theory that will be used along this work. For a more exhaustive introduction to the structure and dynamics of complex networks, we would refer the reader to these excellent works [3, 12, 21, 19, 30, 58, 116, 150, 147, 166].

2.1 Definitions and metrics

The simplest network consists of a set of *vertices*, or *nodes*, connected by *edges*, or *links*, that indicate whether exist or don't exist a mutual relationship between two elements. This network is called *unweighted* and *undirected*. If two nodes i and j have a link connecting them, it is said that they are *adjacent*, or *neighbors*.

One of the most used representations of a network is the *adjacency matrix*. Suppose we have an unweighted undirected network with N nodes. Then, the adjacency matrix A is a $N \times N$ symmetric matrix, whose element a_{ij} is 1 if there is a link between nodes i and j, and 0 otherwise. For example, the adjacency matrix for the network in Fig. 2.1a is

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 \end{pmatrix} . \tag{2.1}$$

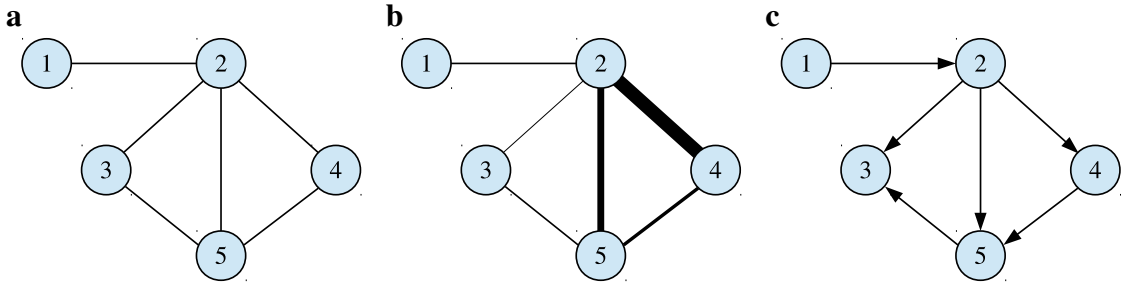


Figure 2.1: Illustration of **(a)** an unweighted undirected network, **(b)** a weighted undirected network and **(c)** an unweighted directed network.

If links have a weight or label associated the network is called *weighted* network. Note that weighted networks have a *topology* given by an unweighted adjacency matrix A and an aggregate *geometry* given by a weighted adjacency matrix W [166]. The elements of W are denoted by w_{ij}. For example, the network depicted in Fig. 2.1b has the weighted adjacency matrix

$$W = \begin{pmatrix} 0 & 1.0 & 0 & 0 & 0 \\ 1.0 & 0 & 0.5 & 6.0 & 2.5 \\ 0 & 0.5 & 0 & 0 & 1.0 \\ 0 & 6.0 & 0 & 0 & 1.5 \\ 0 & 2.5 & 1.0 & 1.5 & 0 \end{pmatrix} . \tag{2.2}$$

One might think also of links that indicate a one-directional tie, for instance, the hyperlinks that have a web page i to a web page j in the WWW network. If links account for the directionality of the connections the network is *directed*, and we refer to the connections as *arcs*. And, in this case, the adjacency matrix is not symmetric. For example, the adjacency matrix for Fig. 2.1b is

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix} . \tag{2.3}$$

Similarly, we can find weighted directed network. This is the case, for instance, of the world trade network, where a link accounts for the amount of exports from one country to another.

Degree, degree distribution and assortativity

In undirected networks, the sum of each row in A is the *degree* of the node, i.e. the number of connections a node has to other nodes. So

$$k_i = \sum_{j=1}^{N} a_{ij} \,, \tag{2.4}$$

is the degree of node i, and the *mean degree* is then

$$\langle k \rangle = \frac{1}{N} \sum_{i=1}^{N} k_i \,. \tag{2.5}$$

If the network is weighted, the weighted degree is called *strength* and is given by

$$s_i = \sum_{j=1}^{N} w_{ij} \,. \tag{2.6}$$

In the case of directed networks, however, one have to distinguish between incoming and outgoing links. Thus, for each node i we have to calculate the *in-degree*

$$k_i^{\text{in}} = \sum_{j=1}^{N} a_{ji} \,, \tag{2.7}$$

and the *out-degree*

$$k_i^{\text{out}} = \sum_{j=1}^{N} a_{ij} \,. \tag{2.8}$$

Analogously, in the case of weighted directed networks one can calculate the *in-strength* and *out-strength*.

The *degree distribution* is defined as the fraction of nodes in the network that have degree k, or, equivalently, as the probability that a vertex chosen uniformly at random has degree k [150]. That is

$$p_k = \frac{1}{N} \sum_{i=1}^{N} \delta(k_i, k) \,, \tag{2.9}$$

where $\delta(a, b)$ is the Kronecker function which is 1 if $a = b$, and 0 otherwise.

The degree distribution is a very important property that has gained a lot of interest since

Network	type	N	$\langle k \rangle$	L	γ	C	r
film actors	und.	449 913	113.43	3.48	2.3	0.78	0.208
sexual contacts	und.	2 810			3.2		
WWW Altavista	dir.	203×10^6	10.46	16.18	2.1/2.7		
Internet	und.	10 697	5.98	3.31	2.5	0.39	−0.189
email messages	dir.	59 912	1.44	4.95	1.5/2.0	0.16	
citation network	dir.	783 339	8.57		3.0		
word co-occurrence	und.	460 902	70.13		2.7	0.44	
train routes	und.	587	66.79	2.16	–	0.69	−0.033
power grid	und.	4 941	2.67	18.99	–	0.08	−0.003
protein interactions	und.	2 115	2.12	6.80	2.4	0.07	−0.156
marine food web	dir.	135	4.43	2.05	–	0.23	−0.263
neural network	dir.	307	7.68	3.97	–	0.28	−0.226

Table 2.1: Properties of some published scale-free networks: type of graph, directed or undirected; N, number of vertices; $\langle k \rangle$, mean degree; L, average shortest path; γ, exponent of degree distribution if it follows a power law (or "–" if not); C, clustering coefficient; r, assortativity. Blank entries indicate unavailable data. Network examples extracted from [150].

the discovery that many real networks exhibit heavy-tailed degree distributions. However, measuring the tail of this right-skewed distributions is somewhat tricky. For convenience, specially when data is sparse such in the tail, it is normal to use the complementary cumulative distribution function (CCDF) of the degree, which is the cumulative histogram of the degrees of nodes

$$P_k = \sum_{k'=k}^{\infty} p_{k'} \,, \tag{2.10}$$

that gives the probability that degree is greater than or equal to k.

The *degree correlation*, or *assorativity*, is a special case of assortative mixing that measures which nodes pair up with which others according to the degree. One of the most used definitions of degree correlation was introduced by Newman [149]. He proposed to measure this property by calculating the Pearson correlation coefficient of the degrees at each side of all the links, which is positive for assortatively mixed networks and negative for disassortative ones. Surprisingly, it has been observed that all social networks appear to be assortative, but other networks appear to be disassortative. Table 2.1 shows some examples.

Path, geodesic path and diameter

We define a *path* in a network as a sequence of nodes $v_1, v_2, \ldots, v_m$ such that from each of these nodes there is a link to the next node in the sequence. This path has a *length* equal to the sum of the weights of the links in the path, that in unweighted networks is $m-1$. Interestingly, using the adjacency matrix we can calculate the number of paths of length r between i and j as

$$N_{ij}^{(r)} = [A^r]_{ij} \,. \tag{2.11}$$

The *geodesic path*, or *shortest path*, between a pair of nodes i and j, if any, is the path from i to j with the shortest length. And the shortest path length d_{ij} is the corresponding addition of the weight of links in the shortest path. One can also calculate the average distance between all pairs of nodes of a network as the average shortest path

$$L = \frac{1}{N(N-1)} \sum_{i \neq j} d_{ij} \,. \tag{2.12}$$

It is certainly worthy of comment that most real networks tend to have a small average shortest path, as it is shown in Table 2.1.

Another important distance measure is the maximal shortest path over the whole network, it is referred as the *diameter*

$$D = \max\{d_{ij}\} \,. \tag{2.13}$$

Note that if there is no possible path between some nodes, the average shortest path and the diameter are set as ∞.

Clustering coefficient

Evidence suggests that in many real networks, specially in social networks, nodes tend to create clusters with a relatively high density of ties. Watts and Strogatz [212] introduced a measure to quantify the extent to which neighbors of node i are also adjacent to each other, the *clustering coefficient*. It was defined as

$$C_i = \frac{2t_i}{k_i(k_i-1)} \,, \tag{2.14}$$

where $k_i(k_i-1)/2$ is the number of links that can exist at most between the k_i neighbors of node i, and t_i is the actual number of existing links between them; in other words, t_i

is the number of triangles around node i.

One can also define the global clustering coefficient C, that measures the cliquishness of the network, such that

$$C = \langle C_i \rangle = \frac{1}{N} \sum_{i=1}^{N} C_i \,. \tag{2.15}$$

Alternatively, Newman [148] proposed to define the global clustering coefficient as the fraction of connected triples of nodes (nodes connected to two other nodes) which also form triangles. This measure can be directly calculated as

$$C = \frac{3 \times \text{number of triangles in the network}}{\text{number of possible triangles in the network}} \,, \tag{2.16}$$

where the factor of 3 is a compensation for the fact that each triangle contributes to three triplets. Examples of the value of this global clustering coefficient for real networks are shown in Table 2.1.

Centrality measures

Measures of *centrality* are used to quantify the importance of a node in the network. Depending on what we mean by more important, or more central, we should use one type or another of centrality measure. The most simplest one is, obviously, the *degree centrality*.

A centrality measure based on shortest path is the *closeness centrality* [178], that prioritizes nodes with shorter mean distance to every other reachable node, such as

$$C_c(i) = \sum_{j=1}^{N} 2^{-L_{ij}} \,. \tag{2.17}$$

Another important measure, also based on shortest path, is the *betweenness centrality* [84], that is often used to estimate the traffic load in communication networks [162]. It is defined as

$$C_b(i) = \sum_{jz} \frac{\Psi_{jz}(i)}{\Psi_{jz}} \,, \tag{2.18}$$

where Ψ_{jz} is total number of shortest paths from node j to node z and Ψ_{jz}(i) is the number of those paths that pass through i. Hence, betweenness centrality quantifies the importance of a node for traversing between different nodes of a network.

2.2 Network models

2.2.1 Random networks

Random networks provide simple models that allow us to use probability theory as a tool to formally study and predict the behavior of real systems that are not amenable to analysis with a deterministic approach [83]. One of the early models, and perhaps the most influential, was proposed in 1959 by Paul Erdös and Alfréd Rényi [63]. They defined a random graph as a set of N labeled nodes connected (or not) by links with some fixed probability p (see Fig. 2.2a). We will call this model the ER model.

Interestingly, many properties of ER networks showed to be almost always predictable [64]. Mean degree is $(N-1)p$, and the degree distribution follows a binomial distribution, since each link is placed with independent fixed probability, defined as

$$p_k = \binom{N-1}{k} p^k (1-p)^{N-1-k} . \tag{2.19}$$

In the limit of large graph size, while keeping the mean degree, Eq. (2.19) approaches a Poisson distribution with mean value $\lambda = \langle k \rangle$ as

$$p_k = \frac{\lambda^k}{k!} e^{-\lambda} . \tag{2.20}$$

Unfortunately, despite some links in real networks may be randomly set up, randomness is not a main feature in ordinary systems. Thus, many properties of ER networks, such as clustering, do not correspond well to that found in real world networks.

In 1998, inspired by Milgram's small-world experiment, Duncan Watts and Steven Strogatz [212] showed that real networks are neither regular nor random (see Fig. 2.2b), and they proposed a model to construct a tractable family of toy networks currently named the Watts and Strogatz (WS) model. The model works as follows. Assume we have N nodes in a regular low dimensional space, e.g. a ring lattice. Then, with a certain probability p, we rewire each of the links at random, in such a way that for $p = 0$ we recover the regular low dimensional lattice, and for $p = 1$ we obtain a completely random network.

For a range of small values of p, the clustering coefficient is large because most of the links are still established with neighbors in the lattice, but the average path length

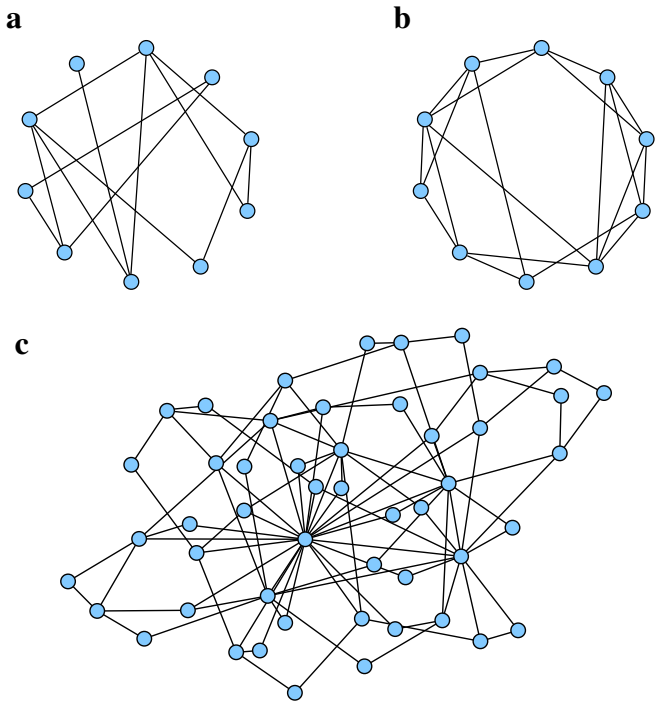


Figure 2.2: Illustration of **(a)** a ER random network, **(b)** a WS small-world network and **(c)** a BA scale-free network (after [106]).

between nodes decreases dramatically (*small-world* property), as random long range links are acting as shortcuts. In fact, for $N \to \infty$ the average shortest path scales as $\log N$.

Despite the fact that the work by Watts and Strogatz attracted a lot of interest, the model is still too simple to capture all the phenomena observed in real networks, the most relevant being the scale-free structure and evolution.

2.2.2 Scale-free networks

In 1999, Albert-László Barabási and Réka Albert were studying the degree distribution of the WWW network when they discovered that on a log-log scale the data points were forming an approximate straight line, suggesting that the degree distribution of the WWW is well approximated with a power law [19]. That is, that the probability that a given node in the network has k connections is

$$P(k) \sim k^{-\gamma} \,. \tag{2.21}$$

After finding that many other networks also exhibit heavy-tailed degree distributions, they named such class of networks *scale-free networks* (see Fig. 2.2c). Table 2.1 shows some examples of real scale-free networks.

Then an important question arose: what mechanism is responsible for the emergence of power-law distributions? Barabási and Albert proposed a generative model to explain the appearance of such kind of complex structures, in which nodes are connected according a mechanism called *preferential attachment*. Preferential attachment states that the probability Π_i that a new node will be connected to node i depends on the degree k_i as

$$\Pi_i = \frac{k_i}{\sum_j k_j} \,. \tag{2.22}$$

This model put emphasis on capturing the network dynamics, describing an open system that grows by the continuous addition of new nodes. In the limit $t \to \infty$, a large graph size, the BA model produces a degree distribution that follows a power law with exponent $\gamma = 3$. In addition, analytical results predicted that the average distance in BA scale-free networks is smaller than in random graphs, being $L \sim \frac{\log N}{\log(\log N)}$, and that the clustering coefficient vanishes with the system size as $C \sim N^{-0.75}$.

2.3 Networks with modular structure

Common experience tells us that some networks seem to have groups of nodes more connected between them than with the rest of the network. For example, we can picture groups of interest in a social network, web pages on related topics in the WWW, the papers on a single topic in a citation network or a functional group in a metabolic network [91].

The ability to detect this group structure, called *modular structure* or *community structure*, could enrich our understanding of the structure-function relationships [80, 167]. Several techniques have been proposed to this purpose: mechanisms based on link removal that try to split the network into subsets of nodes likely to be in the same group [82, 91, 155, 173]; conversely, agglomerative methods that join up nodes in meaningful communities [17, 159, 180]; techniques that use the eigenvalue spectra of the Laplacian matrix of the network [57, 153], which tells us about the node centrality; approaches that classify nodes using random walks [165, 219]; and methods based in the optimization of *modularity* [49].

Modularity is widespread global descriptor that quantifies a given partition into communities and it rests on the intuitive idea that random networks should not exhibit community structure by definition [152]. It is calculated as the probability of having edges falling within groups in the network minus the expected probability in an equivalent network (null case) with the same number of nodes, and edges placed at random preserving the strength of each node:

$$Q = \sum_{\alpha} \left(e_{\alpha\alpha} - a_{\alpha}^2 \right) , \tag{2.23}$$

where $e_{\alpha\beta}$ are the elements of a matrix E that represent the total fraction of links starting at a node in community α and ending at a node in community β, and therefore, $e_{\alpha\alpha}$ is the fraction of links starting and ending in the same community. And $a_\alpha = \sum_\beta e_{\alpha\beta}$, i.e. the fraction of links of E connected to community α, being a_α^2 the expected number of intra-community links.

The larger the modularity the best the partitioning is, because more deviates from the null case. However, modularity optimization is a NP-hard problem [35] and cannot be performed by exhaustive search since the number of different partitions grow at least exponentially in the number of nodes [25]. As a consequence heuristics for the optimization of modularity have become the only feasible, in computational time, reliable and accurate methods to detect modular structure up to now. Some of the most successful heuristics are outlined in [15, 45, 60, 102, 152, 154, 170], and a comparison of different methods was performed in [49]. Note also that modularity relies on a global null model that imposes a certain resolution limit [81, 127, 101], which has led to extend its definition [15].

2.4 Multiplex networks

So far, we have seen networks as simplified mathematical representations that assume that all links coexist in time or are of the same type. However, many real networks are time-dependent or multiplex. Consider, for example, the different kind of ties between users of online social platforms, such as Facebook or Twitter. Or more generally, a social network where individuals constantly drop ties and replace them by new ones in a highly unpredictable fashion [94]. Neglecting the existence of multiple online roles or the time-dependence on relationships can alter the topology of networks and the dynamics on top of it [28]. Mucha, Richardson, Macon, Porter and Onnela [145] introduced a formulation called *multiplex networks* to study these situations.

A multiplex network is a network where nodes are tied to each other by means of

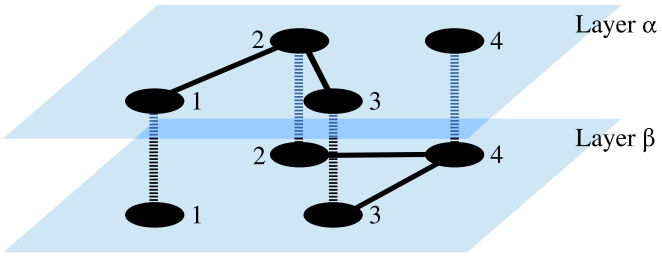


Figure 2.3: Illustration of a multiplex network with 4 nodes and 2 layers.

M different types of links, or at M different snapshots in time. Each type of link is represented as a layer, and we assume that each node of the network has M replicas, one in each layer. Note that "multilayer network" is the general term for a network with multiple layers, and a multiplex network is a multilayer network with the same nodes replicated in all the layers [116]. Thus, we can define a multiplex network with a set of intralayer adjacency matrices $\mathcal{A} = [A^\alpha, A^\beta, \dots, A^M]$, where A^α is the ordinary adjacency matrix of the network in layer α, and a set of interlayer adjacency matrices $\mathcal{C} = [C_1, C_2, \dots, C_N]$, that encodes for the connections among layers. Hence, the element a_{ij}^α of $\mathcal{A}$ is 1 if there is a link of type α between nodes i and j, and 0 otherwise. And the element $c_i^{\alpha\beta}$ of $\mathcal{C}$ is 1 if node i has a tie between layers α and β, and 0 otherwise. Note that, for the sake of clarity, we use latin letters $(i, j, \dots)$ to indicate nodes, and greek letter $(\alpha, \beta, \dots)$ to indicate layers. See [50] for a mathematical formulation of multilayer networks based on tensors. Figure 2.3 shows a schematic of a multiplex network.

In multiplex networks the degree of a node i in the layer α is defined as the *multilayer degree*

$$\kappa_i^\alpha = k_i^\alpha + r_i^\alpha \,, \tag{2.24}$$

where $k_i^\alpha = \sum_{j=1}^N a_{ij}^\alpha$ is the *intralayer degree*, i.e. the number of ties of node i to other nodes in layer α, and $r_i^\alpha = \sum_{\beta=1}^M c_i^{\alpha\beta}$ is the *interlayer degree*, i.e. the number of links from node i in layer α to node i in other layers.

Equivalently to the case of single-layer networks, one can also define a path as an ordered sequence of nodes which starts from node i in layer α and finishes in node j in layer β.

Several other network concepts and methods have been recently generalized to the multiplex framework. So long, the clustering coefficient for multilayer networks is presented in [47], several correlation measures are detailed in [156], and the extension of centrality measures has been done in [23, 52, 53, 104, 191]. A few techniques of community detection have also been proposed [24, 26, 145].

UNIVERSITAT ROVIRA I VIRGILI
NAVIGABILITY AND SYNCHRONIZATION IN COMPLEX NETWORKS: A COMPUTATIONAL APPROACH
Pau Erola Cañellas

3 Mapping the structure of complex networks

A geographic map is a simplified representation of the real world that aims to clearly convey a specific message. To build a map it is necessary to choose about what to represent, at what level of detail and over what time period [134]. Recent studies on complex networks have used mapping as an information extraction technique to understand and decipher the relationship between structure and function [8, 32, 41, 183]. However, similarly to geographic maps, these diagrams depict static information of a time period of natural and artificial dynamic networks, meaning that under network evolution these maps may become uncertain.

In this chapter we will study how network growing affects the reliability of maps. First, we will review two methods to map the network topology into geometric spaces (Sections 3.1 and 3.2), allowing the use of properties that do not exist in the topological domain. Afterwards, in Section 3.3 will study the stability of projections based on Singular Value Decompostion (SVD) when applied to growing scale-free networks. Lastly, Section 3.4 will present the conclusions.

3.1 Mapping the structure of complex networks in hyperbolic space

Transitivity is a very important property in many real-world complex networks and specially in social networks [111, 146] . It implies that, if node i is friends with node j, and node j is also friends with node k, then there is a high probability that i and k are also friends. If this edge indeed exists, it is said that i, j and k form a triangle.

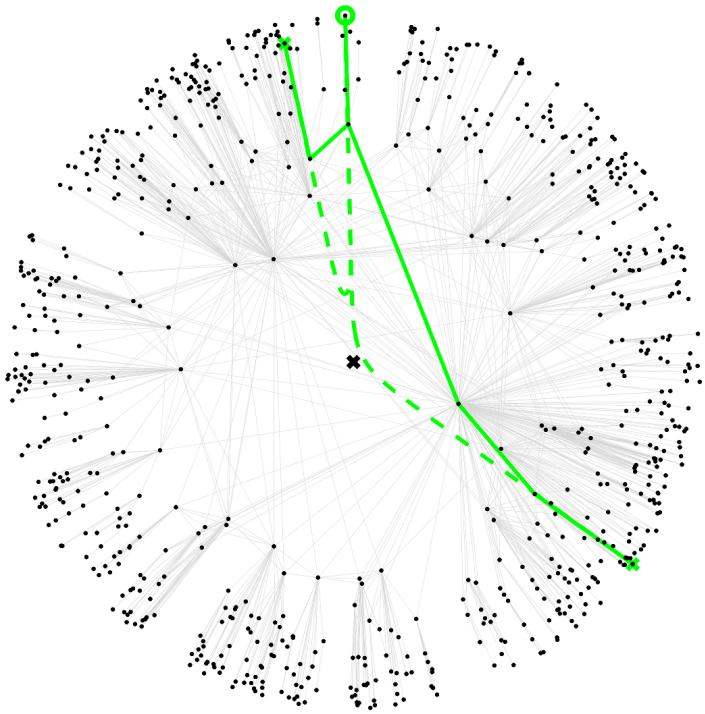


Figure 3.1: Hyperbolic map of a synthetic network, reprinted from [160]. The dashed green lines correspond to two straight lines in the hyperbolic space, and the solid lines represent two greedy paths that try to mimic the geodesics.

Serrano, Krioukov and Boguñá [184] studied this interesting phenomena finding that transitivity may be a topological reflection of the triangle inequality of an underlying hidden metric space where nodes reside. That means that two nodes that are close in this underlying space, are also very likely to be connected in the network. Krioukov, Papadopoulos, Vahdat and Boguñá [123] later proved that hyperbolic geometry matches strong heterogeneity, in terms of the power-law degree distribution exponent, and transitivity properties of complex network topologies. Thus, further studies [122, 32, 183] embedded the Internet network and metabolic networks in hyperbolic planes using a combination of geometric features, a distance measure between nodes inversely proportional to their probability of being connected, and topological characteristics, in this case the degree of the nodes. Figure 3.1 depicts an example of a synthetic network mapped in the hyperbolic space.

The mapping method used in these studies is as follows. In a first step nodes of the network were placed in a hyperbolic disk of radius R, uniformly distributed through the angular component $\theta \in [0, 2\pi]$, and with radial coordinates r inversely proportional to the degree of each node. Nodes with higher degree have smaller r values and are closer to the center of the disk, and low degree nodes are more external. Secondly, the

angular component θ was computed to satisfy the requirement that nearby nodes in the hyperbolic space are connected, i.e. the probability $p(x_{ij})$ that two nodes i and j are connected decreases with the distance x_{ij} between them. This distance can be calculated using the hyperbolic law of cosines

$$\cosh x_{ij} = \cosh r_i \cosh r_j - \sinh r_i \sinh r_j \cos\left(\theta_i - \theta_j\right) . \tag{3.1}$$

The estimation of the coordinates (θ_i, r_i) was performed by maximization of the likelihood that the network topology is produced by this model, that is given by

$$L = \prod_{i<j} p\left(x_{ij}\right)^{a_{ij}} \left[1 - p\left(x_{ij}\right)\right]^{a_{ij}} , \tag{3.2}$$

where a_{ij} are the elements of the adjacency matrix of the network, and $p(x_{ij})$ is the relationship between probabilities and distances

$$p\left(x_{ij}\right) = \left(1 + e^{\frac{x_{ij}-R}{2T}}\right)^{-1} , \tag{3.3}$$

where T is a parameter related with the clustering of the network.

It is interesting to note that in these embedded hyperbolic spaces there exists a congruence between the geodesic distances and the shortest paths of the network [122]. Next chapter reviews a practical application of this characteristic to design an embedded greedy routing on the Internet network.

3.2 Mapping the modular structure of complex networks using TSVD

Singular Value Decomposition (SVD) is a valuable tool for information extraction from matrix-shaped data which has been extraordinarily successful in many applications [97]. Given a rectangular $N \times M$ (real or complex) matrix A, SVD stands for the factorization into the product of three other matrices,

$$A = U\Sigma V^{\dagger} , \tag{3.4}$$

where U is an unitary N-by-N matrix (left singular vectors), and describes the original row entities as vectors of derived orthogonal factor values; Σ, the singular values, is a diagonal N-by-M matrix containing scaling values; and $V^{\dagger}$ denotes the conjugate

transpose of V, an M-by-M unitary matrix, which describes the original column entities in the same way as U.

A practical use of SVD is dimensional reduction approximation via truncation, TSVD. It consists in keeping only some of the largest singular values to produce a least squares optimal, lower rank order approximation. For example, severe dimensional reduction is a condition for success in machine learning SVD applications [97], in particular for the analysis of relationships between a set of documents and the words they contain [56, 27, 129]. In those cases, the decomposition yields information about word-word, word-document and document-document semantic associations; the technique is known as latent semantic indexing [27] (LSI) or latent semantic analysis [128] (LSA).

In the field of complex networks, Arenas, Borge-Holthoefer, Gómez and Zamora-López [8] recently introduced SVD as a useful method to scrutinize the modular structure in complex networks. The idea behind this work is to compute the projection of the connectivity of nodes into the space spanned by the first two left singular vectors, i.e. a rank $r = 2$ approximation. This dimensional reduction is particularly interesting to depict results in a two-dimensional plot for visualization purposes, and complies that the first two left and right singular vectors are unique (up to a sign) if the ordered singular values σ_i of the matrix Σ satisfy $\sigma_1 > \sigma_2 > \sigma_3$ [97].

The object of analysis used to characterize the mesoscopic structure of a network is the *contribution matrix* C of N nodes to M modules; where a module (or community) is a set of nodes with more connections between them than with the rest of the network. The rows of C correspond to nodes, the columns to modules and the elements $C_{i\alpha}$ are the number of links that node i dedicates to module α. C is obtained by multiplying the network's adjacency matrix A and the *partition matrix* S

$$C_{i\alpha} = \sum_{j=1}^{N} A_{ij} S_{j\alpha} \,, \tag{3.5}$$

where $S_{j\alpha} = 1$ if node j belongs to module α, and $S_{j\alpha} = 0$ otherwise. Note that certain changes in the topology might not be reflected in the values of C, for example the rewiring of the connections of a node towards other nodes in the same community.

Once we have calculated the contribution matrix C, we have to project the connectivity of this matrix in the projection space called $\mathcal{U}_2$, where each node i has a coordinate pair or contribution projection vector $\tilde{v}_i$, using TSVD. Figure 3.2b shows an example of a network projected in $\mathcal{U}_2$. More precisely, if U_2 is the matrix formed by the first two columns of U, then $\tilde{n}_i$ is the i-th row of U_2. The map obtained gives an accurate

representation of the main characteristics of the original data, visualizable and easier to scrutinize. Arenas et al. highlighted that this approach has essential differences with classical pattern recognition techniques based on TSVD such as Principal Components Analysis (PCA) because columns of C can not be independently shifted to mean zero without loosing its original meaning.

In order to unveil the signification of the projection's outcome, given that the comparison between input and output is difficult, Arenas et al. proposed to calculate for each coordinate pair the polar coordinates (R_i, θ_i) where R_i is the length of the contribution projection vector $\tilde{v}_i$, and θ_i is the angle between $\tilde{v}_i$ and the hortizontal axis. Moreover, each community has a unique intramodular projection $\tilde{e}_\alpha$, a distinguished direction line of the projection of its internal nodes, those that have links exclusively inside the community, see Fig. 3.2c. Using these values, they came up with a new pair (R_i, ϕ_i), where ϕ is an absolute distance in angle calculated as

$$\phi_n = |\theta_n - \theta_{\tilde{e}_\alpha}|, \tag{3.6}$$

and the descriptors

$$R_{\text{int}} = R \sin \phi, \tag{3.7}$$

and

$$R_{\text{ext}} = R \cos \phi. \tag{3.8}$$

These values, derived from the geometrical properties of the projection, are essential to interpret the mapping of matrix C. Arenas et al. demonstrated that R_{int} informs about the internal contribution of nodes to their corresponding communities, and R_{ext} reflects the boundary structure of communities besides large values of R correspond to highly connected nodes. In simpler words, this two-dimensional plane characterizes the structure of communities and their boundaries.

Further in Chapter 4, we will use this breakthrough interpretation as some kind of map to design a decentralized routing algorithm for the Internet network.

3.3 Quantifying the reliability of TSVD on growing networks

Up to this point, we have discussed some of the applications of TSVD assuming somehow that we are characterizing an invariant object. However, in real applications a common

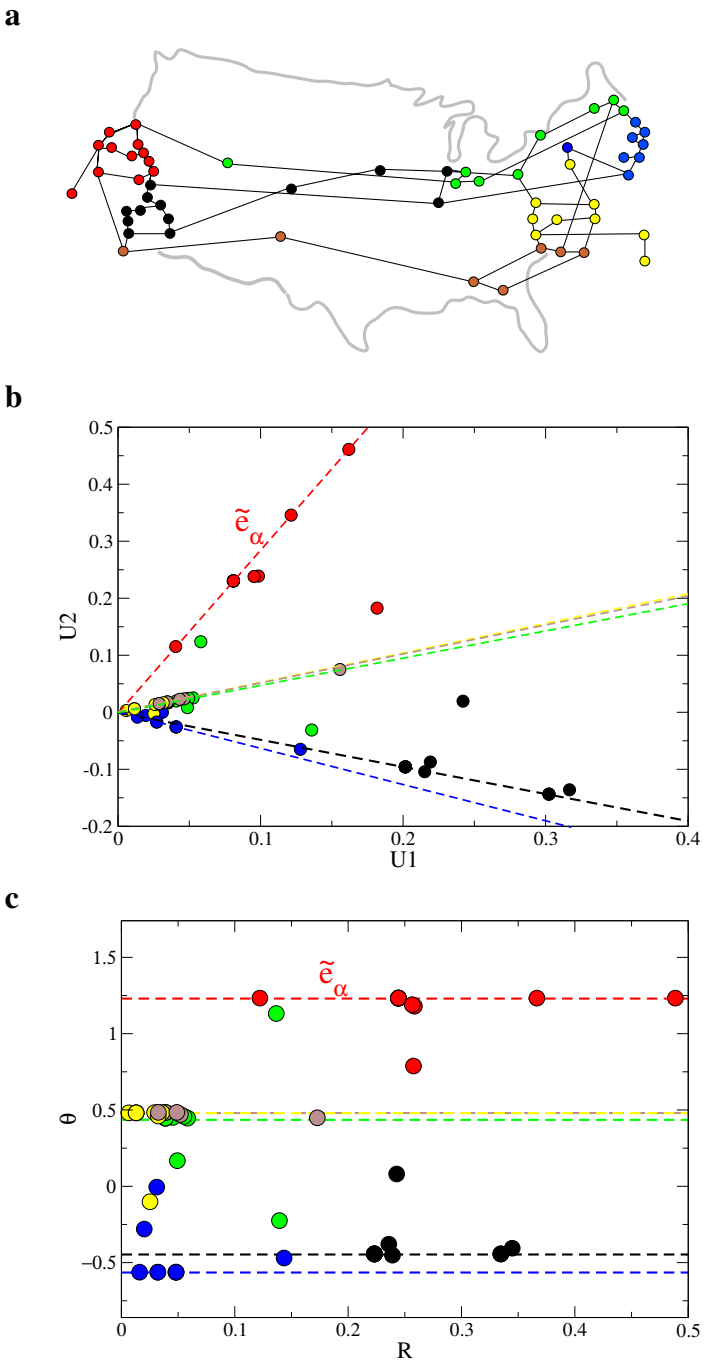


Figure 3.2: Scheme of the Arpanet network at July 1976 **(a)** and their projection in space $\mathcal{U}_2$ **(b)** and in the plane R-θ **(c)**. The color of nodes correspond to the communities detected using the algorithm Extremal Optimization [60] implemented in [171]. The dotted lines are the singular directions $\tilde{e}_\alpha$ of each community.

characteristic is their dynamic nature. For instance, databases rarely stay the same and in order to attain successful information retrieval LSI or LSA must rely on the fact that SVD of textual resources is always up to date. Addition and/or removal of information is constant, meaning that catalogs and indexes quickly become obsolete or incomplete. Turning to networks, the question is equally pertinent: both natural and artificial networks are dynamic, in the sense that they change through time (and so do their modular structures). Paradigmatic examples of this fact are the Internet, the World Wide Web or knowledge databases like Wikipedia: all of them have been object of study from a graph-theoretical point of view [162, 42, 222]. Given this realistic scenario, a major question arises, namely, for how long TSVD stands as a reliable projection of evolving data.

Here we will answer to this question for the TSVD-based network projection method reviewed above using a classical model of network growth, the Barabási-Albert's (BA) scale-free network [20]. This implies that TSVD will be computed on an initial network of size N_0, and then re-computed for successive node additions up to a final size $N_f = 2N_0$. In Fig. 3.3 we show the R–θ planes of an evolving network to get a visual intuition of the map's stability: as the network grows the mapping gets distorted.

To this end, in following sections we propose two measures to quantify the effect of the growth on the TSVD projection at two different levels of study: global and local [65]. The reliability of the projection will be measured considering the "worst case scenario", where each node belongs to its own community, mathematically when $C = A$. Establishing that TSVD is robust to change in these circumstances will settle the fact that TSVD is robust to change on a coarse-grained structure.

3.3.1 Measuring global reliability

In the sequence of computed TSVD projections of an evolving network the nodes' coordinates in $\mathcal{U}_2$ space change. These variations can be assessed by the difference of vectors $\tilde{n}_i$ between the initial and evolved network projections as shown in Fig. 3.4, that maps out the projection of a growing network on the space $\mathcal{U}_2$. We have fixed our attention in two time-shots of the evolution corresponding to growths of 30% and 80%, and we have computed the differences between positions of the same nodes at different stages (z) as $\tilde{v}_i = \tilde{n}_i^0 - \tilde{n}_i^z$, producing a field map that accounts for the changes. This field map is shown in the insets of Fig. 3.4. With an 80% increase of the initial size, the vectors $\tilde{v}_i$ are longer than with a 30% increment, which evidences a larger variability, i.e. a progressive degradation in the TSVD reliability.

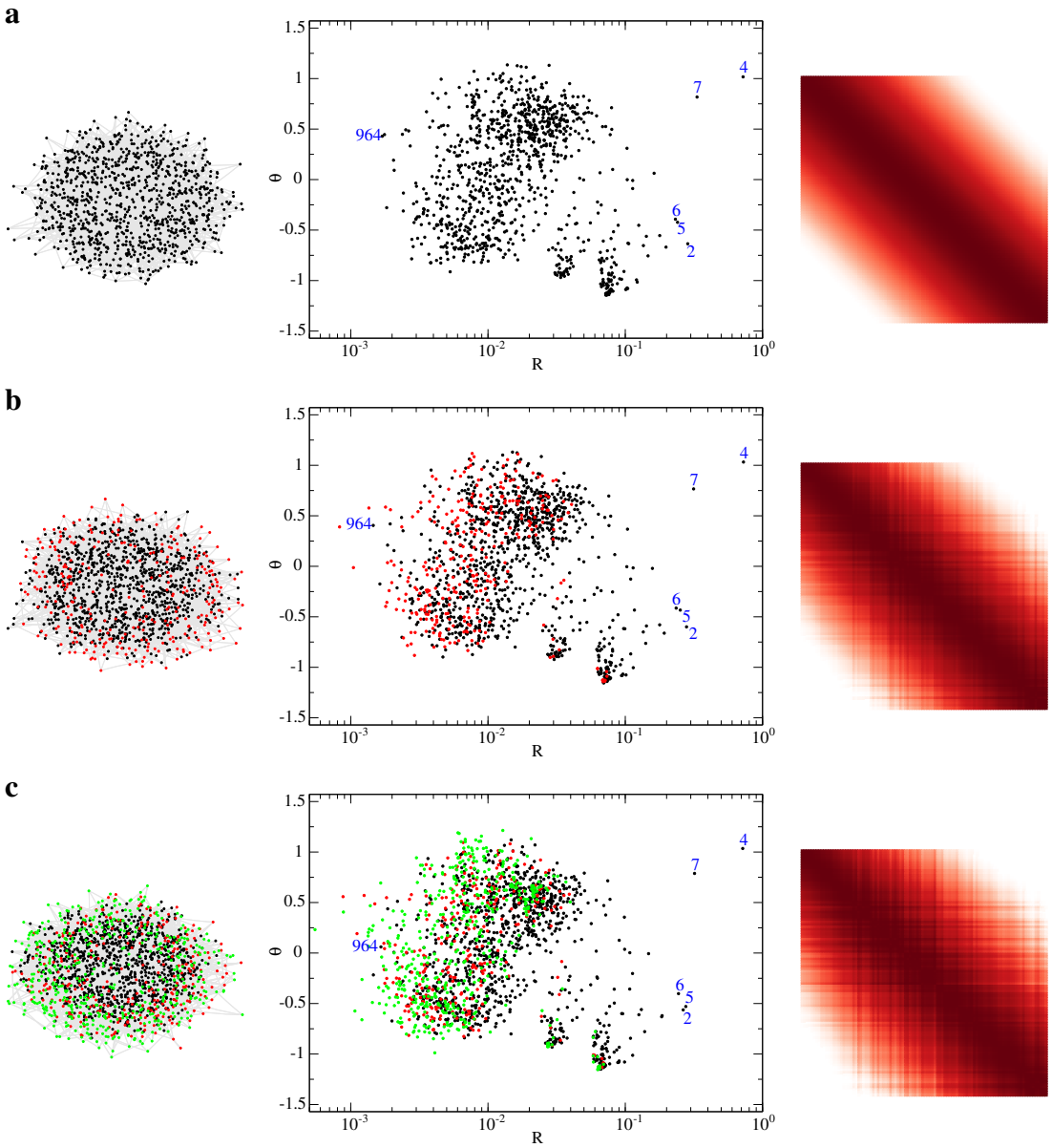


Figure 3.3: Three snapshots of a growing network (left side), their corresponding projection on R–θ plane (center) and the θ-overlapping matrices (right side; see [8] for details). For the sake of clarity, the initial set of nodes **(a)** $N = 1000$ are drawn in black; the second snapshot **(b)** represents a growth of a 30% of nodes, $N = 1300$, new nodes are drawn in red. Finally **(c)** represents a network with $N = 1800$, last arrived nodes are depicted in green. Some nodes from the initial set have been highlighted (2, 4, 5, 6, 7, 964) in the R–θ plane, to get a visual intuition of the map's stability. Note that nodes with a high value of R (2, 4, 5, 6, 7) remain almost unchanged throughout the topology's growth; whereas node 964 undergoes much change from an absolute point of view. The rightmost matrices illustrate the amount of change of nodes with respect to their θ angles: as nodes are added in the structure the cosine overlaps between them increasingly distorts the original figure.

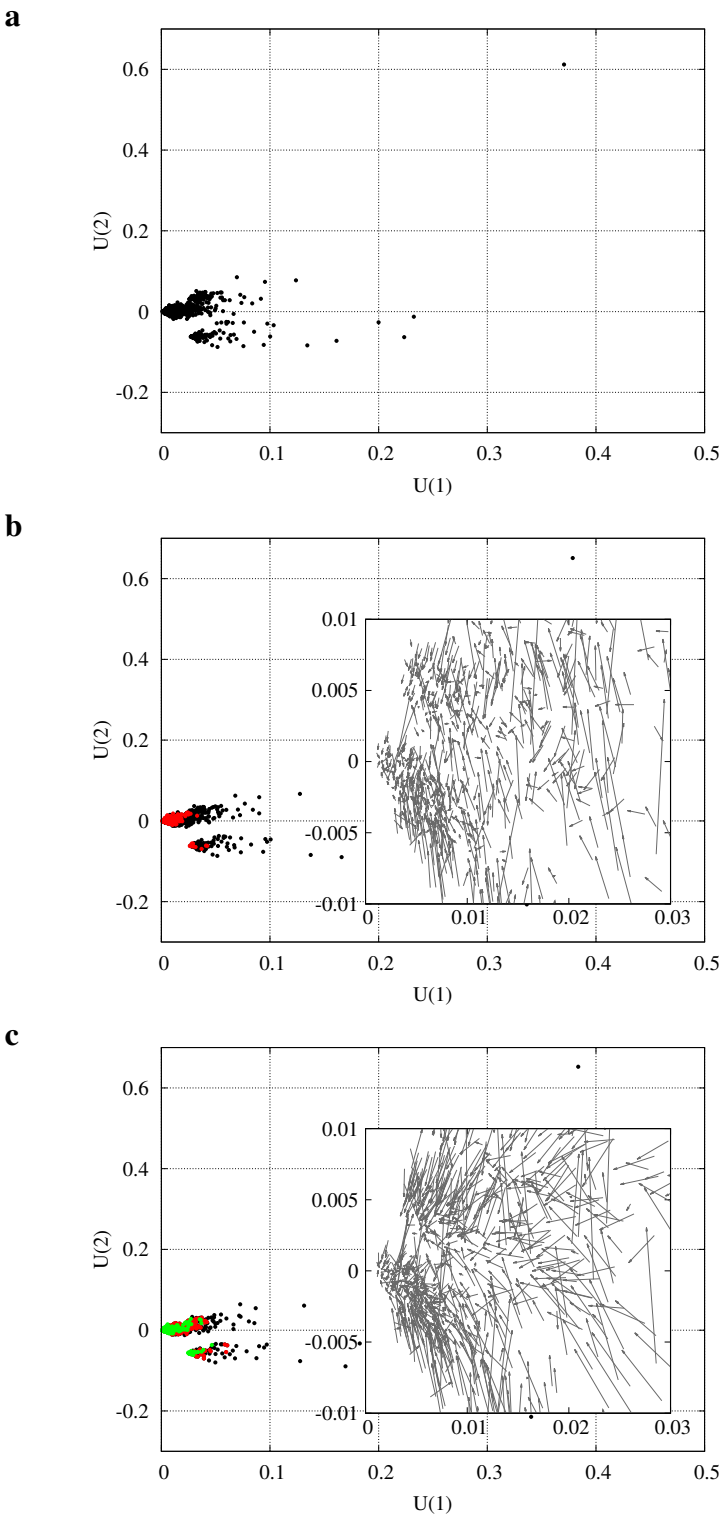


Figure 3.4: Projection in space $\mathcal{U}_2$ of the evolving network presented in Fig. 3.3 (nodes are colored alike). U(1) and U(2) are respectively the first and the second components of the vectors $\tilde{n}_i$. The insets for $N = 1300$ **(b)** and $N = 1800$ **(c)** trace the vectors $\tilde{v}_i$ between the projected coordinates of each node on the grown network and the original coordinates on the initial network.

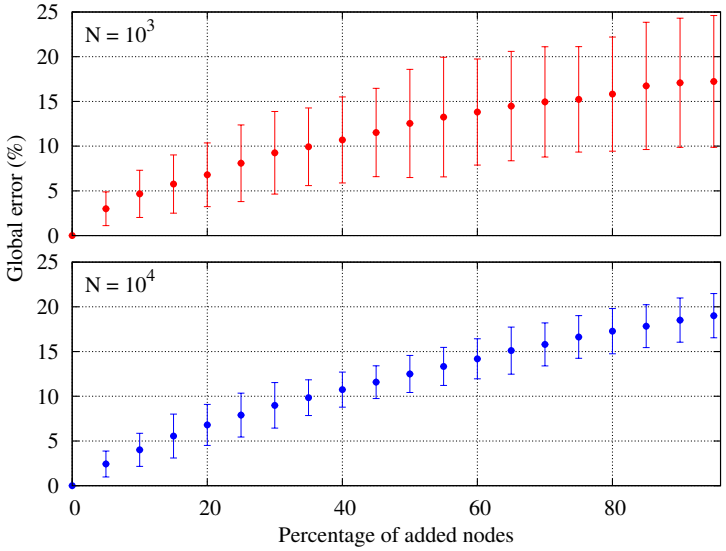


Figure 3.5: Global error on two growing networks with initial sizes $N_0 = 1000$ (above) and $N_0 = 10000$ (below). For each network we have computed the error by increments of 5% in growth up to double the size. In both cases, the global error is lower than 10% for increments lower than 40%. Each point is the average of 100 simulations.

We propose to use the vectors $\tilde{v}_i$ to compute a global quantity that determines the amount of change in the position of nodes in the maps obtained by successive rank r TSVD projections. This *global error* (E_{global}) measure is computed by the expression

$$E_{\text{global}} = \frac{\sum_{i=1}^{N_0} \sum_{j=1}^{r} |U_{ij}^z - U_{ij}^0|}{\sum_{i=1}^{N_0} \sum_{j=1}^{r} |U_{ij}^0|}, \tag{3.9}$$

where U^0 represents the truncated left singular vectors of the original network with N_0 nodes; and U^z also represents the truncated left singular vectors, but of the grown network with size $N_z > N_0$.

We have applied this global measure to monitor the evolution of the TSVD stability for growing networks with initial sizes $N_0 = 1000$ and $N_0 = 10000$. Figure 3.5 shows the percentage of relative error with respect to the original network. In the chart, each successive point represents a 5% of nodes addition. Up to a 40% growth the global error remains below 10%, and doubling the network size the average error still remains below 20%. These results show the good reliability of the projection after the growing process.

3.3.2 Measuring local reliability

Though informative, the previous global quantity can overlook changes at the microscopic level. The neighborhood of each node in the $\mathcal{U}_2$ plane could undergo changes in the sequence of computed TSVD projections difficult to be revealed by the global measure defined above. Thus, we propose a measure that reflects these local changes using the distances between nodes in a neighborhood. Instead of defining a sharp border for the neighbors of each node, we propose to use a gaussian neighborhood that weights the distances according to a variance σ.

First, we construct the $N \times N$ matrix of distances between any pair of nodes in the network at stage z as:

$$D_{ij}^z = \sqrt{\sum_{k=1}^{r} \left(U_{ik}^z - U_{jk}^z\right)^2}, \tag{3.10}$$

where U^z represents the truncated left singular vectors of the network. These distances reflect a measure of proximity between nodes, independently on the global positioning in the map. The neighborhood is weighted to prioritize the stability on closer nodes over the distant ones. To this respect, we compute a matrix of weighted distances S^z using a gaussian distribution that establishes a radius of influence according to the equation

$$S_{ij}^z = D_{ij}^z e^{-\frac{{D_{ij}^z}^2}{2\left(\sigma R_i^0\right)^2}}, \tag{3.11}$$

where we have chosen a radius of influence depending on the node. R_i^0 is the module of the projected vector $\tilde{n}_i$ in the initial network, and σ is a constant. This radius of influence, proportional to the distance to the origin, emphasizes nodes with larger R, which are the most connected ones. See [8]. Using different values of σ in the gaussian function we can tune the size of the neighborhood. Figure 3.6 shows, for a network with 1000 nodes, three magnified views of a network projection in $\mathcal{U}_2$ to illustrate the gaussian radius of influence.

Finally, the *local error* (E_{local}) measure of reliability we propose is computed as the relative error:

$$E_{\text{local}} = \frac{\sum_{i=1}^{N_0}\sum_{j=1}^{N_0} |S_{ij}^z - S_{ij}^0|}{\sum_{i=1}^{N_0}\sum_{j=1}^{N_0} |S_{ij}^0|}, \tag{3.12}$$

where S^0 and S^z represent the matrices of weighted distances of the original network with N_0 nodes, and the grown network with size $N_z > N_0$, respectively.

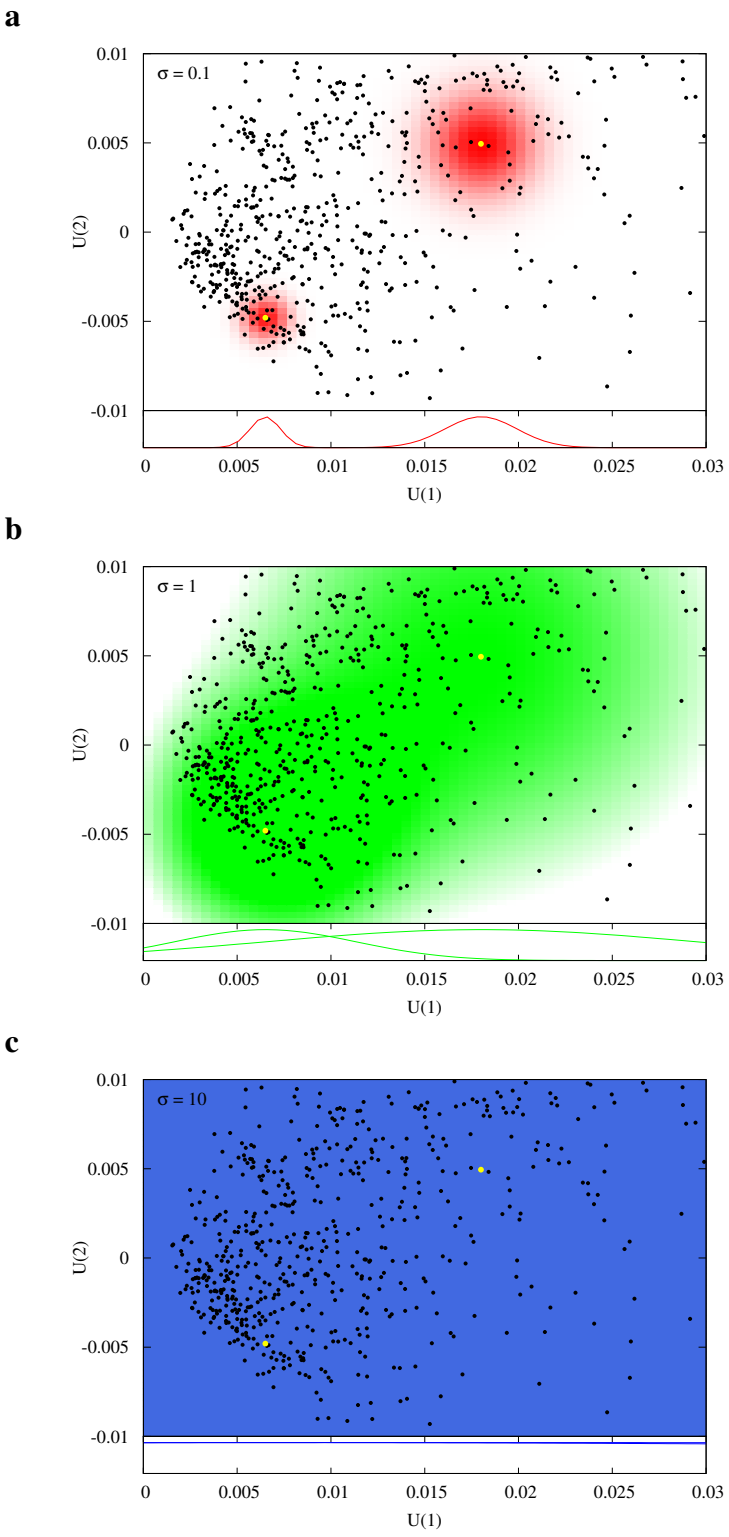


Figure 3.6: Three examples for different values of the radius of influence $\sigma = 0.1$ **(a)**, $\sigma = 1$ **(b)** and $\sigma = 10$ **(c)** in the projection space $\mathcal{U}_2$. At the bottom of each chart, we have plotted for the nodes highlighted in yellow the gaussian curves that multiply the matrix D^z to obtain the matrix of weighted distances S^z.

Figure 3.7 shows the local error measured on two growing networks by increments of 5% of growth. For each network we compute the relative error for $\sigma = 0.1$, 1 and 10. When $\sigma = 0.1$ only the closest neighbors have a significant weight in the measurement of the local error. These low values of σ give the neighbor-wise error a very local sense. On the other hand, when $\sigma = 10$ the gaussian curve becomes flat and the measure is affected by the entire network perturbations, i.e. every node is equally considered as belonging to the neighborhood. Despite this global neighborhood for high σ values, the local error measure represents a relative distance to each node, and we can see that even doubling the network size the average error remains below 0.1%. These very low error rates ensure a good reliability of the projection from a local point of view.

3.4 Conclusions and discussion

In this chapter we have reviewed two mapping mechanisms to extract the structural properties of complex networks. We have also quantified the reliability of SVD-based maps proposed by Arenas et al. [8]. This question is pertinent because standard linear projection techniques, and in particular TSVD, are rooted at the heart of some methodologies which pretend to extract useful and reliable information from dynamic data, i.e. data that is constantly undergoing change.

We have tackled the problem from two complementary points of view. At the large-scale level, we have monitored average changes in nodes' TSVD projections. This means that each node's projection is compared against itself on successive changes.

However, the success in practical applications of TSVD depends mostly on neighborhood stability. In other words, coherence of the output when data has suffered changes relies on the fact that the surroundings of a projected node are similar to those before those changes had happened. From a mathematical point of view, this merely implies that projections change in a coordinated way, such that relative positions are stable. Keeping this in mind, the local measure developed above captures this facet of the problem by comparing not the evolution of a node's position against itself, but rather against the rest of nodes. Furthermore, we have introduced a parameter to weight this variation depending on the distance from the node of interest. This tunable parameter allows for a finer observation of neighborhood stability, ranging from immediate neighborhood measures to far-reaching areas. Note that the local measure is orders of magnitude lower than the global one. This points to the fact that, although the projection changes significantly, displacements in the plane $\mathcal{U}_2$ are similar in magnitude and direction on average. In other words, as the degree of a node in our network grows following the

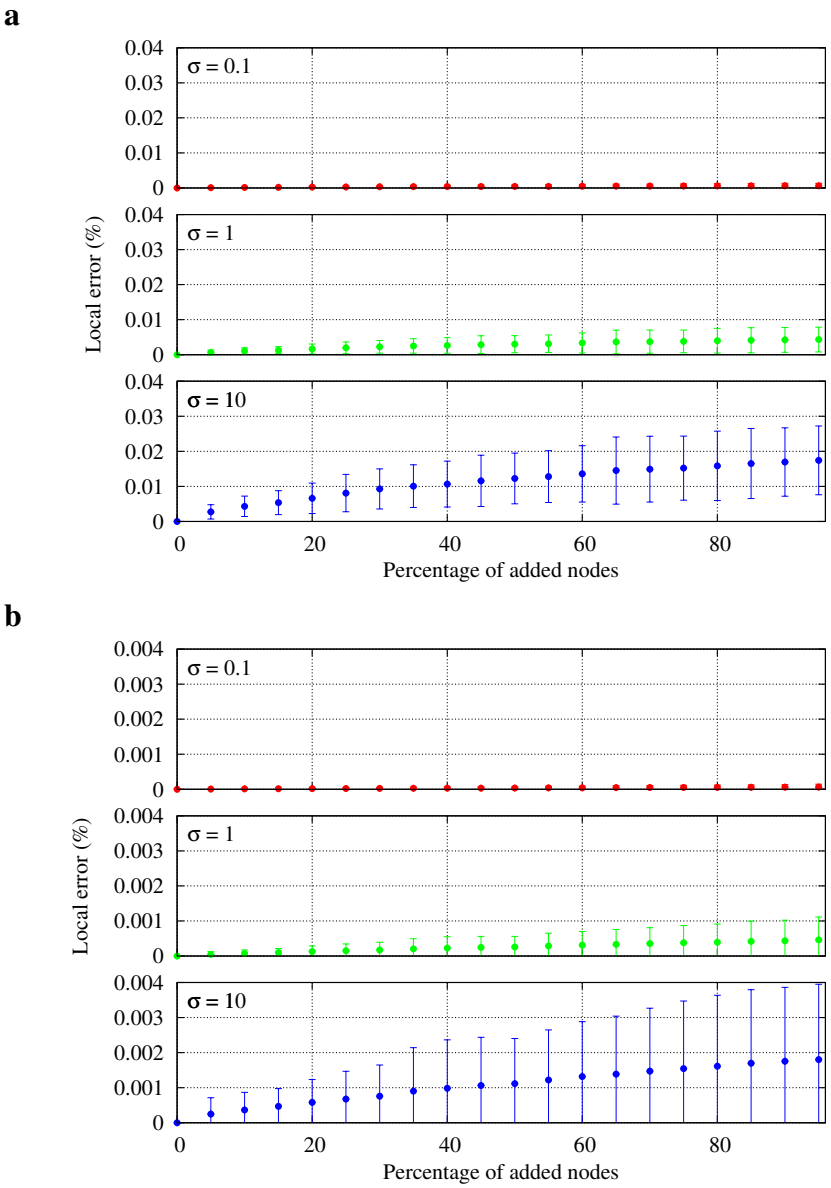


Figure 3.7: Local error on two growing networks with initial size $N_0 = 1000$ **(a)** and $N_0 = 10000$ **(b)**. For each network we have computed the relative error for $\sigma = 0.1$, 1 and 10 by increments of 5% of growth, from top to bottom respectively. Small values of σ yield lower errors, but in all settings the reliability of the projection is high. Each point has been calculated with 100 simulations.

preferential attachment model, it is highly likely that its neighbors also increase their weight staying close together.

The results indicate that TSVD projections are very robust against growing scale-free networks. From a global point of view, an addition of 40% of new data implies only an average change of 10% from initial conditions. Doubling the amount of nodes to a network supposes a modification of 15% in the positions of the set of initial nodes. More importantly, changes at the local level (neighborhood) are close to 0 even in the most demanding case. Such results have been obtained with rather large structures.

It is important to remark that our study focuses on a particular network model (BA) in which time plays an important role: the later a node appears, the lowest its chances to become an important one (a hub). We anticipate that the irruption of important entities at late stages of evolution would surely disrupt TSVD projections in a more significant way. Nonetheless, we stress that growing systems typically develop smoothly, so our conclusions can be safely held. Similarly, Boguñá et al. [32] assert that their mapping is, under the usual pertubations of the Internet topology, essentially static.

Finally, we can briefly relate these results to the original motivation of this chapter, that is, a scenario where the modular structure of networks is taken into account. In that situation, the stability of a TSVD map in the case of network changes is granted given the above reported results. Then, the characterization of the role of nodes and modules in terms of SVD's output can be safely regarded as faithful even in the case of severe changes in the underlying topology.

UNIVERSITAT ROVIRA I VIRGILI
NAVIGABILITY AND SYNCHRONIZATION IN COMPLEX NETWORKS: A COMPUTATIONAL APPROACH
Pau Erola Cañellas

4 Scaling the Internet network: routing without global information

The Internet is increasingly changing the way we do everyday tasks at work, at home, and how we communicate with one another, to the extent that it has become a pervasive and critical infrastructure. The number of connected devices reached 8.7 billion in 2012 [190] and it is expected to continue growing exponentially. However, current Internet protocols require to know the whole network topology and this accelerated growing could jeopardize the Internet scalability.

Here, based on the scale-free nature of the Internet, we will propose a novel greedy routing scheme motivated by Milgram's experiment that showed that, in complex networks, without knowing the global topology, messages can be routed efficiently.

This chapter is organized as follows. Sections 4.1 and 4.2 briefly introduce the Internet network and its scalability issue. Then, in Section 4.3 we review the solutions proposed in the field of compact routing. Section 4.4 present two greedy routing approaches based on complex network theory. Finally, in Section 4.5 we discuss the limitations of these approaches and future directions to explore.

4.1 The Internet as a complex network

The Internet comprises a decentralized and growing collection of more than 30,000 computer networks from all around the world. In its entrails the Internet is made up of networks of routers, each one under the control of a single technical administration called Autonomous Systems (AS). Two ASes are connected if and only if they establish

Number of nodes	23 752
Number of edges	58 416
Average degree of nodes	4.919
Maximum degree of nodes	2 778
Average clustering coefficient	0.360
Power-law exponent	2.18
Diameter	8
Average path length	3.84

Table 4.1: Properties of the Internet ASs network for June 2009 [38].

a business relationship (customer-provider or peer-to-peer relationships), making the Internet a *living self-organized* system.

Knowing the topology of the Internet is useful for various reasons, including the evaluation of efficacy, scalability, robustness and security. However, it is difficult to characterize it because the network is constantly changing and there is no central authority to provide topological information [79]. Still, there have been tremendous efforts to measure and infer [6] this complex structure using reverse engineering tools. University of Oregon's Route Views Project [157] has been storing global BGP routing information since 1997, and other projects such as CAIDA's Skitter [39], Mercator [100] and the community-based DIMES [186], aimed to study the structure of the Internet by continuously collecting traceroutes and probes. Figure 4.1 shows a network of the Internet at router level based on the Route Views table dump archive.

The work of Faloutsos, Faloutsos and Faloutsos [75] was the first to identify the scale-free property of the Internet at inter-domain (or AS) level. The network exhibited a power-law degree distribution with an exponent of $\gamma = 2.1$, and average path length near 3.2 standing out its small-world character [162]. In Table 4.1 we have summarized other statistical properties of a snapshot of the Internet network.

This discovery burst the development of new models of the Internet network. Early, Yook, Jeong and Barabási [215] presented a model based on the preferential attachment mechanism. Although this model reproduced similar networks from the perspective of node degree distribution, their structure was not equal [4]. Improving this model, Zhou and Mondragón [221] introduced the Positive-Feedback Preference model which also accurately reproduces the rich-club connectivity, disassortative mixing and betweenness centrality properties, among others, of the Internet. A different solution was proposed by Boguñá and Díaz-Guilera [182], who proposed a model of interconnected ASs which self-organize and grow driven by competition for resources and adaptation to maintain

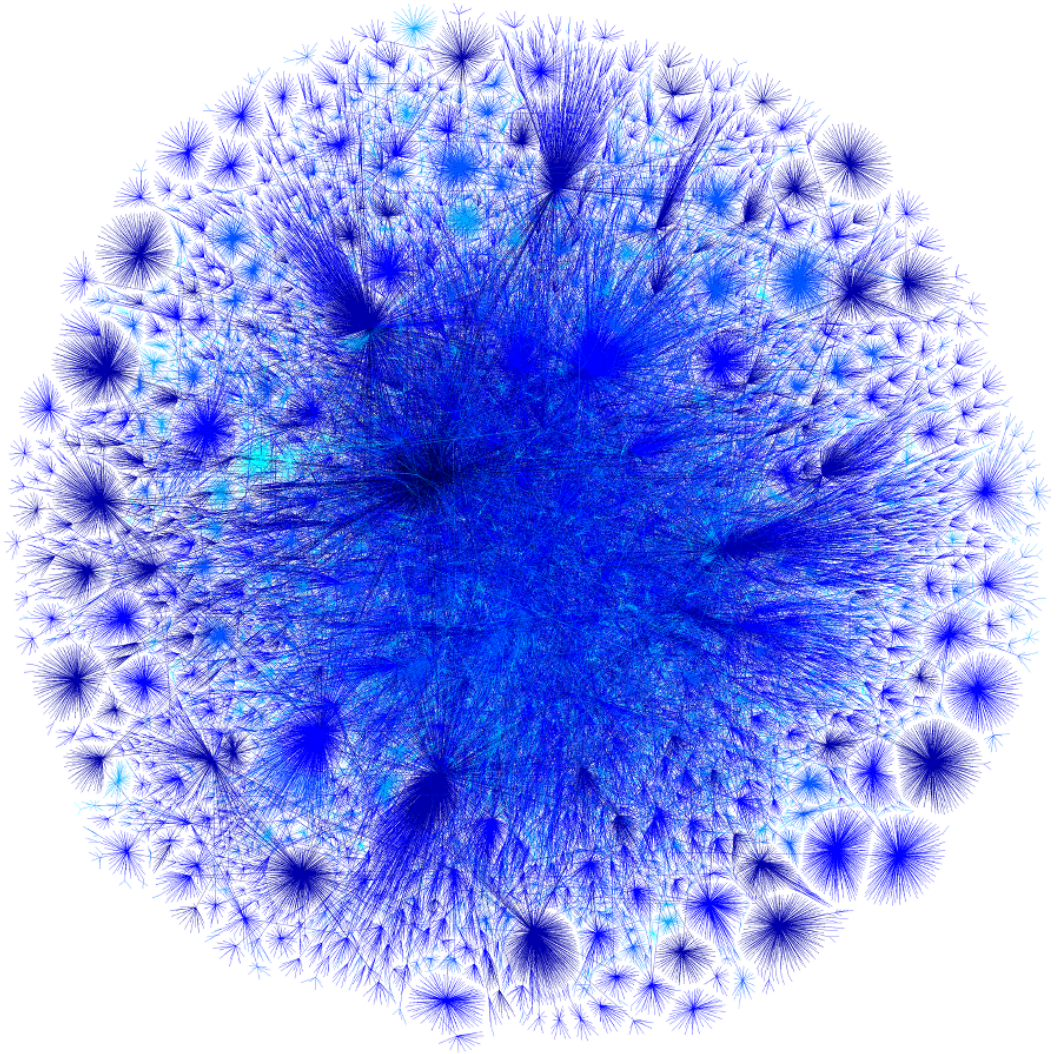


Figure 4.1: Two-dimensional image of a router-level Internet network (2010) created by The Opte Project [135]. Source available under a Creative Commons Attribution-NonCommercial 4.0 International License at http://www.opte.org.

functionality.

Before to conclude this section, we believe it is appropriate to make an aside about the scale-free nature of the Internet. Willinger, Alderson and Doyle pointed out that we must be cautious when using available data [213]. Tools like traceroute have several technical limitations (IP alias resolution, opacity of layers of the TCP/IP stack,...), and as a result high-degree nodes may simply be artifacts that defy common technological and economical sense of the Internet. Deepen this aspect is out of the scope of this thesis and, despite the confusion, we assume the accuracy of the power-law claim, without forgetting that

> "Power laws do have an interesting and possibly even important role to play, but one needs to be very cautious when interpreting them. The most productive use of power laws in the real world will therefore, we believe, come from recognizing their ubiquity (and perhaps exploiting them to simplify or even motivate subsequent analysis) rather than from imbuing them with a vague and mistakenly mystical sense of universality." [198]

4.2 Scalability issues of the Internet

The main function of the Internet is to forward information from an origin host, traversing switches, routers and other network nodes, to reach a destination. To do so it uses the Border Gateway Protocol (BGP) [196], formerly [175], a routing protocol to exchange reachability information between autonomous systems on the Internet, which implements a vector with end-to-end paths to guide the routing of information packets. This vector is called the Routing Table (RT) and represents, in fact, a distributed global view of the network topology. To maintain this consistent map, routers need to exchange information of reachability through the network, and it is currently the root of one of the most challenging problems in network architecture: ensure the scalability of the Internet [136].

This problem is compounded because the Internet is growing abruptly wrapped in nontrivial dynamics. Empirical studies qualified this overall growth as the outcome of a net balance between births and deaths involving large fractions of the system [162]. This growth is estimated to be exponential, and the number of entries in the RT are currently growing at super-linear rate in the inter-domain level. In addition, other dynamics such as failures to aggregate prefixes, address fragmentation, load balancing and multi-homing, are augmenting the RT demands [37]. Withal, this behavior is compromising the BGP

scalability due to technological constrains. RT updates to maintain the information of the shortest paths involves a huge amount of data exchange and significant convergence times of up tens of seconds, hindering the communication process.

The poor scaling properties of this routing scheme has been studied in depth within the context of compact routing. Unfortunately, these studies have concluded that in the presence of topology dynamics, a better scaling on Internet-like topologies is fundamentally impossible: while routing tables can be greatly reduced, the amount of messages per topology change can not grow slower than linearly [120]. This limitation has raised the need to explore new lines of research and, given the scale-free topology of the Internet, complex network theory promises to be the natural framework to analyze and propose solutions.

Next section briefly reviews some designs in the fields of compact routing, and then we present two methodologies based on complex network theory that use navigable maps of the Internet.

4.3 Schemes based on compact routing and network architecture

In the last years, some studies have raised concerns with respect to the scalability of routing tables of the BGP protocol [121]. To reduce the cost of communication networks, we should be concerned about the routing of messages through the network, which ideally should follow a shortest path, and the amount of information required, which should be minimal. Simple solutions can guarantee optimal shortest paths at the expenses of keeping in memory big routing tables, but these solutions are too expensive for large systems. In the research field of compact routing, Peleg and Upfal [163] addressed the trade-off between the average stretch (the average ratio of every path length relative to the shortest path) and the space needed to store RTs for general networks.

Early, Kleinrock and Kamoun [118] presented an alternative routing strategy to reduce the RT length in large networks. Their method intended to create a nested hierarchical structure of clusters (groups of closer nodes according to some measure), where any node only needs to maintain a small amount of information from distant nodes in other clusters, while it maintains complete information about its neighbors in the cluster. With this approach, the authors achieved a substantial size reduction of the RT (from N to $\ln N$), at the expense of increasing the average message length due to extra labels. However, the stretch analysis is satisfactory only under certain topologies, those in which the shortest

path distance between nodes rapidly increases with the network size. Because of this, scale-free networks (under the small-world phenomenon) are not good scenarios for hierarchical routing, suffering severe increments of the paths length. But despite the problem described, hierarchical routing is on the basis of the implanted inter-domain address strategy Classless Inter-Domain Routing (CIDR) [86]. CIDR was introduced in 1993 and allows to group (or aggregate) addresses into blocks using bitwise masks, reducing the number of entries in the RT.

CIDR is an implicit scheme, i.e. the nodes are *a priori* labeled with structural information. The routing process explores this information to choose the neighbor to which a message should be sent [179]. It is also *name-depedent* or *labeled* routing [16], which means that nodes are tagged with topology dependent information identifiers used to guide the packet forwarding. Among the newer name-dependent solutions we found the schemes of Thorup and Zwick [206] and Brady and Cowen [34] interesting, which are highlighted here because of their proved efficiency in scale-free networks.

Thorup and Zwick (TZ) presented a scheme with stretch 3 and RT sizes of $O(n^{1/2}\log^{1/2} n)$-bits, being these results for the worst-case graph. In the case of scale-free networks Krioukov et al. [120] estimated the performance of this routing and found on average stretch of 1.1 and very small, nearly optimal, RT sizes . The basic idea behind TZ is to use a small set of *landmarks* (nodes potentially involved in the process of routing) to guide the navigation. That reveals why we achieve the best possible performance in the TZ scheme: scale-free graphs are optimally structured to exploit high degree nodes (that will turn out to be landmarks) which are very important for finding shortest paths in such networks [117]. In turn, Brady and Cowen (BC) designed a routing scheme for undirected and unweighted graphs with the basic idea that trees cover scale-free graphs with minor deviations. In scale-free networks BC scheme guarantees an average stretch of 1.1 and logarithmic scaling RT sizes of $O(\log^2 n)$.

Alternatively to name-dependent solutions, we find *name-independent* routing schemes [16]. In this variant, nodes may be labeled arbitrarily making routing generally harder: first we need to know the location of the destination, i.e. we need dictionary tables to translate name-independent labels into locators in a name-dependent map. Abraham et al. [1] presented a nearly optimal name-independent routing for undirected graphs. The scheme has stretch 3, and a size upperbound of the RT of $O(\sqrt{n})$ per node, the same upper limits that in TZ. According to these results, the use of name-independent schemes provides no clear advantage over the use of name-dependent routing [120]. Other name-independent routing solutions have been designed specifically to address the Internet routing problems [188]. Particularly noteworthy are the Locator-Identifier

Split (LIS) approaches like LISP [76], ENCAPS [110] and NIMROD [43]. All these LIS proposals separate the identifier and the locator of each node allowing aggressive aggregation techniques. During the communication process, the locators are encapsulated in each packet in a special wrapper and, at inter-domain level, packets are forwarded using only these locators. Krioukov et al. identified the problems that may affect these solutions: they require in general a database to maintain updated locator information, and due to its hierarchical structure, aggressive aggregation is impossible on scale-free topologies.

It must be emphasized that all these proposals are effective for static networks. Under time varying networks, the proposals assume that each change in the network structure generates a new graph and its routing solution must be recalculated. Krioukov et al. reported the pessimistic scenario that the communication cost lower bound for scale-free graphs is $O(n)$. Nevertheless, adaptive solutions have not been studied in depth yet.

4.4 Solutions based on complex network theory

The ever-increasing amount of information and goods transported has aroused great interest in the scientific community to reveal the characteristics of natural and human-made transportation and communication networks. And in particular, great effort has been exerted to understand technological communication networks such as the Internet or the World Wide Web which display both scale-free and small-world features.

Arenas, Danon, Díaz-Guilera and Guimerà [10] studied the congestion in complex communication networks and introduced formalisms to handle it. Several other research were focused in optimizing the communication capacity of the network by considering the effect of network topology on the packet routing [195], minimizing the maximum node betweenness [48, 214], estimating the waiting time along the shortest path [217], and other adaptive strategies [62, 89, 211].

Focused on the structural properties, Adamic, Lukose, Puniyani and Huberman [2] studied the role that high connectivity nodes (hubs) play in the communication process. Hubs are important actors in the routing process that facilitate search and information distribution, specially in large networks. They also introduce several local search strategies that exploit high degree nodes which have costs that scale sub-linearly with the network size. Note however that, in weighted networks, inhomogeneity in the weight of links may destroying the small-world nature of a networks [36].

These routing strategies presented so far assume that we know the global topology of the

network. Nevertheless, the famous Milgram's small-world experiment revealed that in complex networks a message can be routed efficiently between any pair of nodes even without knowing the global topology [3]. Understanding how this process works is a key question still to be answered.

From the perspective of artificial intelligence, the search of a solution in a dataset modeled as a network has been a widely studied problem. Let us explain this with an archetype. There is a classical example in artificial intelligence aimed to find the shortest path between two cities in Romania, from Arad to Bucharest [177]. Let us assume that we know the straight-line distances between all cities of Romania to Bucharest. Having this information, we can use an informed search strategy to find efficiently the shortest path to our destination. In the given example, a successful strategy is to choose the neighbor city (with connection by road) whose distance to destination is shorter. Representing the problem as a network, an informed search method chooses a node with lowest cost based on an heuristic function, e.g. minimize the straight-line distance in the above case. For instance, [132] used a greedy approach where navigators have incomplete knowledge about the network topology but some directional information. Like in this case, if we can build a coordinate system that reveals the network structure in sufficient detail, at each point in our path we can use this map to determine which direction to choose to take us closer to the destination. That is, a greedy algorithm that mimics the efficient routing process of natural networks.

In this section, we present two heuristic approaches that use maps based on the scale-free characteristic of the Internet (see Chapter 3) to propose local alternatives to BGP protocol.

4.4.1 Greedy routing on hyperbolic spaces

In section Section 3.1 we outlined the mapping of networks into hyperbolic geometric spaces. Boguñá, Papadopoulos and Krioukov [32] exploited the property that hyperbolic geometry matches strong heterogeneity (in terms of the power-law degree distribution exponent) and clustering properties of complex network topologies [122] to design a greedy routing algorithm for the Internet.

As anticipated, there is a congruence between the geodesic distances and the shortest paths. Like the trace drawn by the geodesics (see Fig. 3.1), shortest paths tend to originate in the outer hyperbolic disk, then getting closer to the center of the embedded space, and finally coming back to the exterior of the disk to reach the destination. Using this similarity, Boguñá et al. designed a greedy packet forwarding that selects at each step of

the path the neighbour closest to the destination, following the geometric shortest path between points in the space.

On mapped snapshots of the Internet network, this scheme achieved near optimal results because of the high concordance between the Internet structure and underlying space. The average success ratio is 97%, a very high value considering that the nodes only have knowlegde about its neighbors. And the greedy paths are very close to the shortest path with an average stretch of 1.1.

Furthermore, the local nature of the design minimizes the routing communication needs in front of dynamic topology changes, favoring the network scalability. Boguñá et al. showed that its projection successfully overcomes the failures of links and nodes, with minor degradation of the performance of greedy forwarding, and that the addition of new nodes is possible without recalculating the entire map. Also, that the projection of the nodes in the metric space, under normal growing conditions, is very stable in time eliminating the need to recalculate the coordinates.

4.4.2 Local routing in the Internet using TSVD-based map

Kleinberg [117] pointed out that the correlation between the local structure of a network and long-range links provides critical cues for finding paths through the network. We can interpret these terms of *local structure* and *long-range links* as a definition of the modular structure of the network. So we start from the hypotesis that modular structure inherent to real complex networks provides a useful information whose exploitation is competitive with global shortest path strategies. And to exploit this information in a greedy routing, we will make use of the TSVD-based projections described in Section 3.2. Then we start with this, presumably helpful, geometrical projection of the modular structure of our network.

Generally, a greedy routing algorithm will work as follows. Let us assume we want to go from node i to node j, and let N_i refer to the neighbors of node i. Each node $k \in N_i$ is a candidate in the path. In the routing process, for each node $k \in N_i$ we have to compute a metaheuristic cost function and select the candidate that minimizes it in each step. This process is repeated until the destination is reached, the current node i does not find a feasible successor or a time constraint is violated. Loops are not allowed.

In an unstructured network an algorithm based on hubs' transit (i.e. route toward hubs with the hope that hubs will be directly connected to any target destination) will result in a decent routing even improving classical routing techniques [2]. We will use

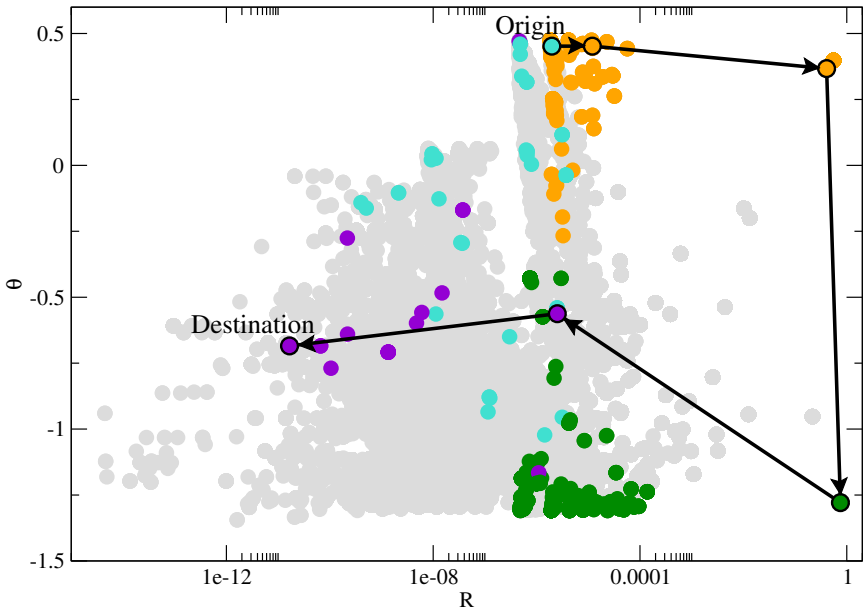


Figure 4.2: Path example on the TSVD-based projection of the Internet network. The color of the nodes represents their community, though some communities have been omitted to simplify the figure. The path begins in a small degree node, which searches the network hubs (high values of R) and gets closer in the angular position θ, to enter the destination community and reach the destination node.

this observation as a first general strategy in our routing algorithm: we will prioritize nodes that could act as hub connectors within the community and eventually with other communities. Hubs, nodes with high values of R, are nodes with many connections that will presumably allow us to lead to the destination quickly reducing the path length. On the other hand, it will also be important to consider those nodes that are similar to our destination, that is, that belong to the same community. Figure 4.2 shows a typical path in the projection plane.

Based on this general idea, we propose to differentiate between two scenarios: when we are into the destination community and when we are not. When our neighbor k belongs to the same community α_j than the destination node j, $k \in \alpha_j$, we are interested in finding nodes with an important weight in the community. In the other case, if $k \notin \alpha_j$, we seek for nodes near to the boundaries of other communities. Algorithm 1 presents the cost functions that reflect both scenarios [69, 70].

When $k \in \alpha_j$ we want to prioritize nodes that are closer to the angular position of the destination node θ_j and have many connections in the destination community. Hence,

Algorithm 1 Greedy routing algorithm.

cost[*] $\leftarrow \infty$
if $\exists\, k \in \alpha_j$ **then**
  for $\forall\, k \in \alpha_j$ **do**
    cost[k] $\leftarrow \dfrac{\lambda + |\Delta\theta_{k\to j}|}{R_{\text{int}_k}}$
  end for
else
  for $\forall\, k$ with $R_{\text{ext}_j} > 0$ **do**
    cost[k] $\leftarrow \dfrac{|\Delta\theta_{k\to j}|}{R_{\text{ext}_k}}$
  end for
end if
return k with lower cost[k]

the cost function acquires the form

$$\frac{\lambda + |\Delta\theta_{k\to j}|}{R_{\text{int}_k}}, \tag{4.1}$$

where $|\Delta\theta_{k\to j}|$ is the angular distance between the neighbor node k and the destination node j. Since $|\Delta\theta_{k\to j}|$ is small for nodes within the same community, we add a shift value λ to reduce its fast annealing.

In the other case, $k \notin \alpha_j$, we consider that optimal nodes are those closer to θ_j and with high intramodular connectivity. Those are the ones that minimize

$$\frac{|\Delta\theta_{k\to j}|}{R_{\text{ext}_k}}\,. \tag{4.2}$$

This heuristic considers explicitly the angular position of nodes that reflects the boundaries of each community and somehow gives directionality to the path. However, due to the information loss produced by rank 2 truncation, it is difficult to assess the uncertainty of this value. Thus, we propose another approach to guide the routing process using only the inter- and intramodular modules, R_{int} and R_{ext} respectively [68]. Note too, that these values have implicit the θ angular value of each node. We present this solution in Algorithm 2.

In the first case of the greedy algorithm, $k \in \alpha_j$, we are interested in finding nodes with an important intramodular contribution, i.e. with many connections within the

Algorithm 2 Simplified version of the greedy routing algorithm.

```
cost[*] ← ∞
if ∃ k ∈ α_j then
  for ∀ k ∈ α_j do
    cost[k] ← 1 / R_int_k
  end for
else
  for ∀ k with R_ext_j > 0 do
    cost[k] ← 1 / R_ext_k
  end for
end if
return k with lower cost[k]
```

community. So we define the cost function

$$\frac{1}{R_{\text{int}_k}}. \tag{4.3}$$

In the other case, if $k \notin \alpha_j$, we seek for nodes with high connectivity to other communities. Thus we find for high values of R_{ext} through

$$\frac{1}{R_{\text{ext}_k}}. \tag{4.4}$$

We evaluated the performance of our algorithms using a snapshot of the Internet network for June 2009 [38]. This dataset is an empirically collected view of the RouteViews BGP table and therefore it's prone to contain errors. For simplicity, we build an undirected and unweighted network. Then we conducted the community analysis using Extremal Optimization algorithm [60, 171]. We applied it recursively to decompose big communities into smaller subcommunities, resulting in 349 communities of sizes ranging between 2 and 546 nodes. Finally, we calculated the matrix C to be projected on a plane using the TSVD.

The simulation consisted in 10^6 paths of randomly selected pairs of origin and destination. Figure 4.3 presents the distribution of path lengths. Both algorithms achieved a success rate of 94% and a median path of 5 steps (shortest path median is 4), that given that we are using only local information, the percentage of success is very good. We will discuss further in the next section about the equality of results between the two algorithms, that

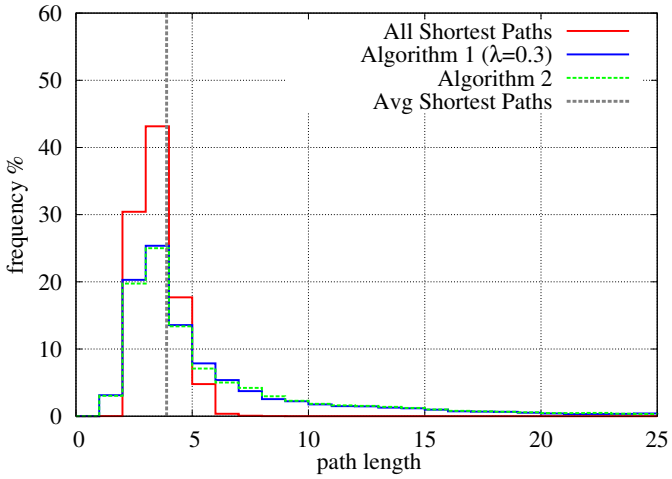


Figure 4.3: Distribution of path lenght of the simulation of 10^6 paths.

puts emphasis on the importance of the degree of nodes.

We attribute the unreachable destinations to nodes that are very far from any hub of the network. Nodes with longer paths are, in most cases, internal nodes connected only with other nodes of small degree. Our projection reflects the boundaries of communities of the network but only provides information about the number of connections a node has in the internal topology of its community. This makes our search strategy to have difficulties to find the pathway to those poorly connected nodes with only intramodular contribution. Because of this, the stretch of our algorithm can get very unfavorable, as is suggested by the long tail of the distribution that shows Figure 4.3. The consideration of customer-provider roles can minimize partially this problem by reducing the possible paths.

In Figure 4.4 we profiled the behavior of our algorithms. We selected paths of lenght 3, 4, and 5, that represent more than the 60% of successful paths. These plots confirmed that our proposal first looks for big hubs up to access the destination community, and within it chooses the nodes with largest internal contribution.

In situations where the data is continuously changing, like the Internet, it is an interesting question whether the TSVD projection of an initial dataset is reliable. In Section 3.3.2 we quantified the reliability of a sequence of computed TSVD projections of growing Barabási-Albert's scale-free networks. In that situation, we proved that the stability of a TSDV map is very high when considering neighborhood stability (note that R_{int} and R_{ext} are relative modules). Hence, under normal topology dynamics (catastrophic events never happened in Internet history) addition and removal of nodes in the network can be done without recalculating the projection.

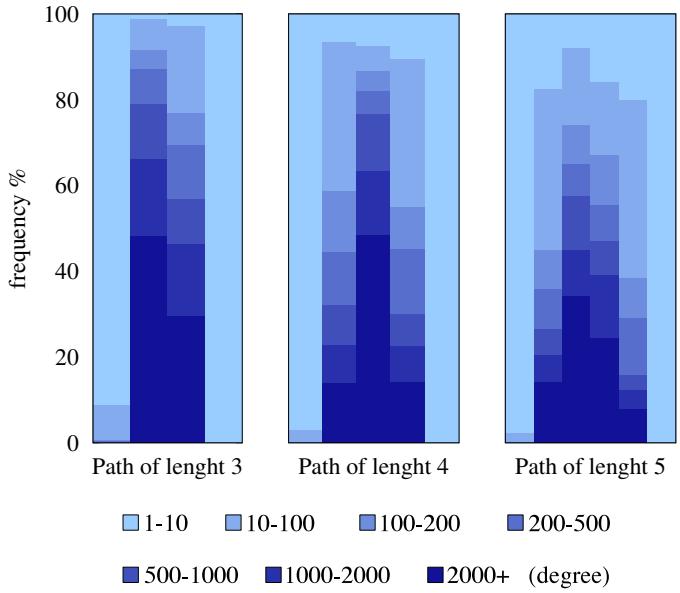


Figure 4.4: Distribution of degree nodes in each step of path. Each column in x axis corresponds to a node (step) in the path.

4.5 Conclusions and discussion

This chapter addressed one of the major problems that compromises the scalability of Internet: its dependence on a global routing table. We have reviewed some of the proposed solutions both in the area of compact routing as well as novel proposals based on complex network theory.

The work by Boguñá et al. [32] proposed a greedy routing algorithm for the Internet based on hyperbolic maps (see Section 3.1). In a Poincare representation, the nodes with high degree go to the center of the space and their angular position is determined by their probability to be connected, thus on a equivalent way to a community structure. This fact establish a similarity between the topological shortest paths and geodesic distances, and the greedy forwarding just needs to follow the geodesics.

Conversely, we have proposed to exploit TSVD-based maps (see Section 3.2) where nodes of the same community are organized around specific singular directions, and hubs have large module values. The main differences with the previous scheme are that Boguñá et al. look for the best projection in terms of the statistical distribution of connections, while TSVD-based maps look for the best projection in terms of the modular structure inherent in the network. But at the end, both methods use the same principle: going through the geodesics of the hyperbolic space from any origen node to any destination node is equivalent to finding the hubs of the network and going closer

to the destination community. Both approaches prove to be competitive and probably complementary.

Interestingly, we have proposed two different sets of cost functions to guide the navigation process that got similar performance. Algorithm 1 prioritizes nodes closer in angular position to the destination node, considering that it is more probable that they have connections to the destination community. Nonetheless, Algorithm 2 just used inter- and intramodular modules. To understand these results, first note that there is a relation between the number of communities in the network and dimensional reduction. The more communities we have, the greater the loss of information and therefore, angular values θ become more uncertain. Secondly, in accordance with [2], this fact highlights the importance of hubs in the routing process. Also in hyperbolic spaces, where the distance between two points is approximately the sum of their radial coordinates, see Eq. (3.1), the angular difference has less influence.

Another recent work that studied the routing process in scale-free networks was introduced by Lattanzi *et al.* [130]. This study focused on social networks and used a model of *Affiliation Networks* that considers the existence of an *interest space* lying underneath. The search in this space can be conducted greedily and the results presented reinforce our assumption that low degree nodes not connected directly to hubs are hard to find, and that large hubs are essential for an efficient routing process.

To all of this must be added that these proposals have proved to be resilient to changes in the network structure due to failures and growth. Therefore, there is no need to recalculate continuously the corresponding projections, obtaining at the same time high performances of the routing algorithms. Nevertheless, profound changes in the core of large hubs may accelerate their obsolescence. For instance, Labovitz et al. [126] have identified a significant evolution of provider interconnection strategies involving a rapid transition to more densely interconnected and less hierarchical inter-domain topology. This particular time varying behavior may affect the success of these new routing schemes and should be analyzed in the future.

UNIVERSITAT ROVIRA I VIRGILI
NAVIGABILITY AND SYNCHRONIZATION IN COMPLEX NETWORKS: A COMPUTATIONAL APPROACH
Pau Erola Cañellas

5 Traffic-aware routing for multiplex networks

When too many packets are present in a part of the network, performance degrades [205]. This is what happened in early 2014, when some US broadband Internet subscribers began to experience a slowing down of their traffic. The problem was identified as a congestion issue along end-to-end paths between Comcast, the largest broadcasting company in the world by revenue, and Netflix, but also affected to many other Internet destinations [78].

From the complex networks perspective, many works have devoted to the study the congestion phenomena in communication systems, and in special the Internet [162]. However, simple single-layer representations may veil some nontrivial properties in systems as heterogeneous as the "network of networks". Models based in new multiplex framework may be fitted candidates to study such systems.

The aim of this chapter is to study an adaptive routing scheme for multiplex communication networks. Section 5.1 will serve as an introduction to the congestion problem. In Section 5.2 we describe our model of multiplex communication networks. Finally, in Section 5.3 we present the results which are discussed so far in Section 5.4.

5.1 Traffic congestion in complex networks

Congestion arises when the load in a network is greater than the resources of a part of the system can handle and then, performance degrades [205]. Macroscopically, it is defined as a phase transition from the free flow to a congested state due to contagious

propagation of congestion among routers [201].

In the Internet, the adopted all-you-can-eat consumption model has been increasing bandwidth demands, and has forced large content providers to build a more densely interconnected and less hierarchical inter-domain topology [125]. For instance, in 1999 Internet service providers were concerned about the increasing demand of multimedia content, and some of them installed networks of computers that bypassed the main traffic routes to avoid critical congestion sites [174]. However, episodes of congestion are still frequent. Recently, a congestion problem along the end-to-end paths between Comcast and Netflix caused delays in the traffic to many Internet destinations. Although the episode was solved with a paid peering arrangement, the reasons that caused this congestion chapter remain unknown [77, 78]. And the recent Open Internet Order does not require their disclosure:

> "We decline at this time to require disclosure of the source, location, timing, or duration of network congestion, noting that congestion may originate beyond the broadband provider's network and the limitations of a broadband provider's knowledge of some of these performance characteristics." [46]

This lack of real data on the global traffic load and the poor fidelity of existing tools to infer congestion data deep into the network using indirect methods, makes difficult identifying points of congestion along end-to-end paths. In the field of complex networks, this has led to the zero-order assumption of using the betweenness centrality as an estimate of the actual load in the network [162]. In addition, due to the absence of a blue print of the whole connectivity and the technical heterogeneity [202], many models on traffic flow have adopted a simplified representation of the system using three dynamical parameters: the rate at which new packets are injected in the system, the routing protocol that describes how traffic is distributed, and a queuing system that represents the limited capacity of the nodes [59].

Previous studies in complex networks have characterized the phase transition point from free flow to a jammed state in simple models of network traffic dynamics [9, 10, 192]. On the other hand, in an empirical study, Fukuda, Takayasu and Takayasu [85] estimated the congestion levels on TCP/IP networks by observing test packet round trip time sequences.

The interplay between the traffic load and the network topology has also attracted a lot of interest. Holme [112] investigated the relationship between centrality measures and traffic density, finding non-trivial dependences. Zao, Lai, Park and Ye [218] analyzed the effects of the injection rate and the capacity of nodes in different topologies, and

weighted networks were studied in [96]. Goh, Kahng and Kim [95] also tested the effects of the degree distribution of the network. In addition, Moreno, Pastor-Satorras, Vázquez and Vespignani [142] studied the tolerance to congestion failures in scale-free networks. Finally, other studies have proposed models and formalisms to obtain optimized network topologies [22, 103]. And Singh and Gupte [189] proposed strategies to relieve congestion in hubs, which proved to be very important in the communication process [2], adjusting their capacity and number of connections.

Other efforts have focused in optimizing the communication capacity of the network with routing schemes that redistribute traffic load avoiding hubs or central nodes to minimze the maximum node betweenness [48, 195, 214, 211]. The use of adaptive models that integrate actual traffic load in the routing algorithm also proved to enhance the communication performance [61, 62, 217]. In the same line, Gavaldà, Duch and Gómez-Gardeñes proposed an adaptive mechanism that modifies the weight of the communication channels to guide the routing process [89].

Multiplex networks may represent a more realistic framework to study the complex mesh of networks through which information flow. These networks present different structural features and have shown some significantly different nontrivial properties from those of isolated networks [28]. For instance, they are at risk of abrupt changes in the structure [172] or can enhance diffusive processes [98]. This is partly due to the way and density of coupling links between layers [87, 107, 131, 143, 161, 185, 204]. It has been observed that, in real-world complex systems, layers are not coupled randomly but correlated [131], which may manifest new dynamical properties.

A few recent works have studied traffic congestion in communication networks [203, 220], but so far, no adaptive routing scheme has been studied. In next sections, we propose to analyze a dynamic routing strategy in multiplex layers with different coupling correlations.

5.2 Traffic model for multiplex communication networks

In communication networks, Zhou, Yan and Lai [220] studied the effect of the number of coupling links in a multilayered system where nodes have heterogeneous transport capacity. More recently, Tan, Wu, Xia and Tse [203] have explored the effect of coupling preference and coupling probability on traffic congestion. They found that when different scale-free networks are interconnected in an assortative way, i.e. nodes are connected to other nodes with a similar degree, the efficiency of the network is improved. Both

studies implemented static routing schemes [214] that minimize the maximum effective betweenness of the network to manage congestion redistributing traffic load in central nodes to other non-central ones. However, these approaches are based on the assumption that the actual traffic load in the network can be estimated *a priori* using betweenness centrality, making them indifferent to changes in traffic distribution. Moreover, their static nature makes them more difficult to maintain under structure dynamics and evolution.

Here, we propose to study packet transportation dynamics using a traffic-aware routing mechanism that selects less crowded paths using ongoing congestion information about their neighbors [61, 62]. To this end we will use multiplex networks and we will assess the effects of different coupling preferences on congestion onset. The ingredients of our model are detailed below.

5.2.1 Coupling complex networks

To uncover the effects of coupling assortativity in the delivery of information, we consider a simple setting with imposed couplings. Each multiplex network consists of two layers, layer α and layer β, of N nodes. We will analyze three different coupling preferences based on degree as proposed by Tan, Xia, Zhan and Jin [204]:

- Assortative coupling. We sort nodes in layer α and layer β by degree, both in ascending order, placing nodes with same degree at random. Then we link the first node in layer α with the first node in layer β, the second node in α with the second node in β, and so on till all nodes are coupled.
- Disassortative coupling. We sort by degree nodes in layer α in ascending order, and nodes in layer β in descending order. Nodes with same degree are placed at random. Then we link the first node in layer α with the first node in layer β, the second node in α with the second node in β, and so on till all nodes are coupled.
- Random coupling. We randomly sort nodes in layer α and layerβ. Then we link the first node in α with the first node in β, the second node in α with the second node in β, and so on till all nodes are coupled.

5.2.2 Traffic dynamics model

Let's assume that each node in our network, in each layer, is endowed with a first-in-first-out (FIFO) queue with infinite storage capacity. We have modeled the traffic dynamics as

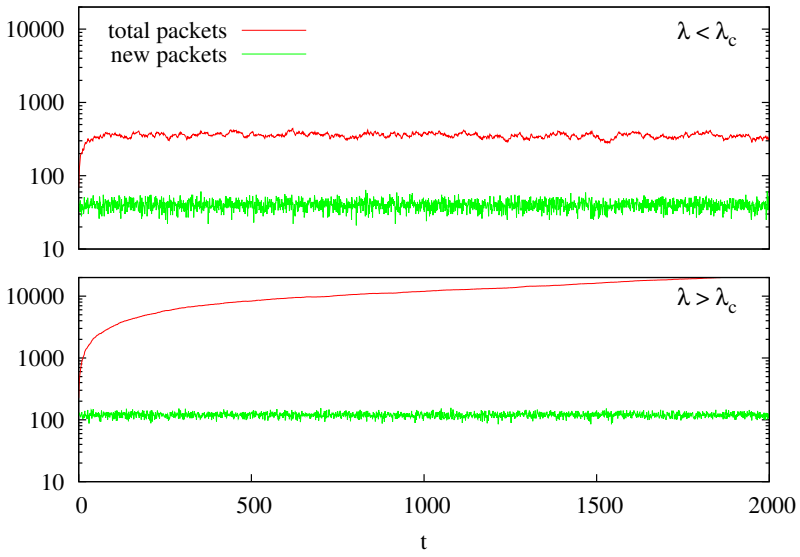


Figure 5.1: The injection rate of packets in the network is controlled by a Poisson process with intensity parameter λ. When λ is low (top) the number of new incoming packets is low and the load in the system is, in average, constant. Conversely, when λ is above a certain threshold (bottom) some nodes are overloaded, and packets, those that have not reached their destination, accumulate in the system.

a continuous time stochastic process. Packets are created at each node i with a randomly chosen destination node j, following a Poisson process with probability

$$P\{N(t+\tau) - N(t) = k\} = \frac{e^{-\lambda\tau}(\lambda\tau)^k}{k!}, \qquad k = 0, 1, \dots, \tag{5.1}$$

where $N(t+\tau) - N(t) = k$ is the number of events that occur in time interval $(t, t+\tau]$, and λ is the rate parameter that controls the number of events per unit time, also known as *intensity*. Therefore, the higher the value of λ the higher the amount of packets generated. Figure 5.1 depicts the effects of λ.

When a new packet is created, it is stored in the queue of node i at layer α selected by the routing heuristic described below, where it waits to be processed. Node i attempts to forward packets to its neighbors according to a Poisson process with rate parameter λ_q, which represents the bandwidth. Without loss of generality, we assume that $\lambda_q = 1$ for all nodes, i.e. nodes can handle in average one packet per unit time.

Then, in turn, the packet is sent to one of the neighbors of node i, chosen by the routing protocol, where it is enqueued again, and so on. Once the packet is processed in the destination node, it disappears from the network.

5.2.3 Routing protocol

We have implemented the routing strategy proposed by Echenique, Gómez-Gardeñes and Moreno [61, 62] that was introduced to take into account the traffic load on the network. It is an heuristic approach that, at each step, chooses the path that minimizes the effective distance δ_{kj} from a neighboring node k of node i to destination node j, defined as

$$\delta_{kj} = hd_{kj} + (1-h)c_k\,, \tag{5.2}$$

where d_{kj} is the shortest path distance between node k and j, and c_k reflects the number of packets in the queue of k. Thus the algorithm considers the overloaded nodes ($c_k > 0$) trying to divert a packet through a less congested path. If there are several paths with the same value of δ_{kj}, one of them is chosen randomly. Note also, that when the limit $h = 1$, the routing behaves like the standard shortest path protocol.

This integration of topological information and traffic information showed to perform better than the standard shortest path protocol, making the load uniform in a larger part of the network. But when congestion arises upon incresing intensity λ, it may appear suddenly as a jamming transition.

Here, we generalize this routing strategy for multiplex networks. In multiplex networks, nodes belong to several different layers simultaneously and they are connected by means of a specific set of edges in each layer [51]. Because of this interconnected structure, we have to take into account the connections within the same layer and the connections to its counterparts in different layers.

Let's assume we are in node i in layer α, and we want to reach the destination node j through the path that minimizes the heuristic function

$$\delta^{\beta}_{kj} = hd^{\beta}_{kj} + (1-h)c^{\beta}_{k}\,, \tag{5.3}$$

where node k in layer β denotes any vertex directly connected to i in layer α. Hence, when $k \neq i$ and $\beta = \alpha$ we are choosing a neighboring node within the same layer, and when $k = i$ and $\beta \neq \alpha$, we are jumping from layer α to layer β without changing node. Note that multiplex architectures do not allow any other possible scenario.

5.3 Simulation analysis

Simulation is the most suitable tool for understanding many of the complexities of Internet topologies and traffic dynamics, including adaptive congestion control [79]. Hence, to quantitatively characterize the transmission capacity of different coupling setups in multiplex networks we have built an event-driven simulator of hot-potato routing networks [31]. The design was in line with the constraints and dynamics described above.

To perform our analysis, we have generated 100 multiplex networks of two layers, where each layer is an unweighted scale-free network (BA) with 1000 nodes. Each multiplex network is coupled in fashion with the three preference schemes describes in 5.2.1. For the sake of simplicity, we assume that the cost of switching between layers is the same as the cost of all intralayer edges. Simulations consist in a warm up period of 10^3 unit time, plus $2 \cdot 10^4$ unit time simulation.

In the rest of this section, first, we examine the actual flow of data to characterize the behavior of the adaptive routing mechanism implemented. Then we fix our attention to the global scale congestion.

5.3.1 Traffic balance between layers

We start our analysis by observing how packets are spread along the network as a function of the injection rate of new packets. To this end we will use three different layers, characterized in Table 5.1, to construct two multiplex networks. The first network will be formed by layers 1 and 2, which show very similar properties. Conversely, the second multiplex network will be formed by layers 1 and 3, having one layer slightly shorter communication paths. We will assess the balance of packets handled by each layer of the network.

In Fig. 5.2 we present the relative number of packets handled by each layer of the two multiplex networks build for, from top to bottom, assortative, disassortative and random couplings. Balance is computed as the difference between the total number of packets in layers α and β (generated, ending or circulating, as appropriate), divided by the total number of packets in layer α.

At first, we can observe that the plots for all three coupling setups have very similar profiles. The left column corresponds to the multiplex network where both layers have very similar properties and the balancing between them is fair. More interesting is the right side case, where we take into account layers with slightly different properties. In

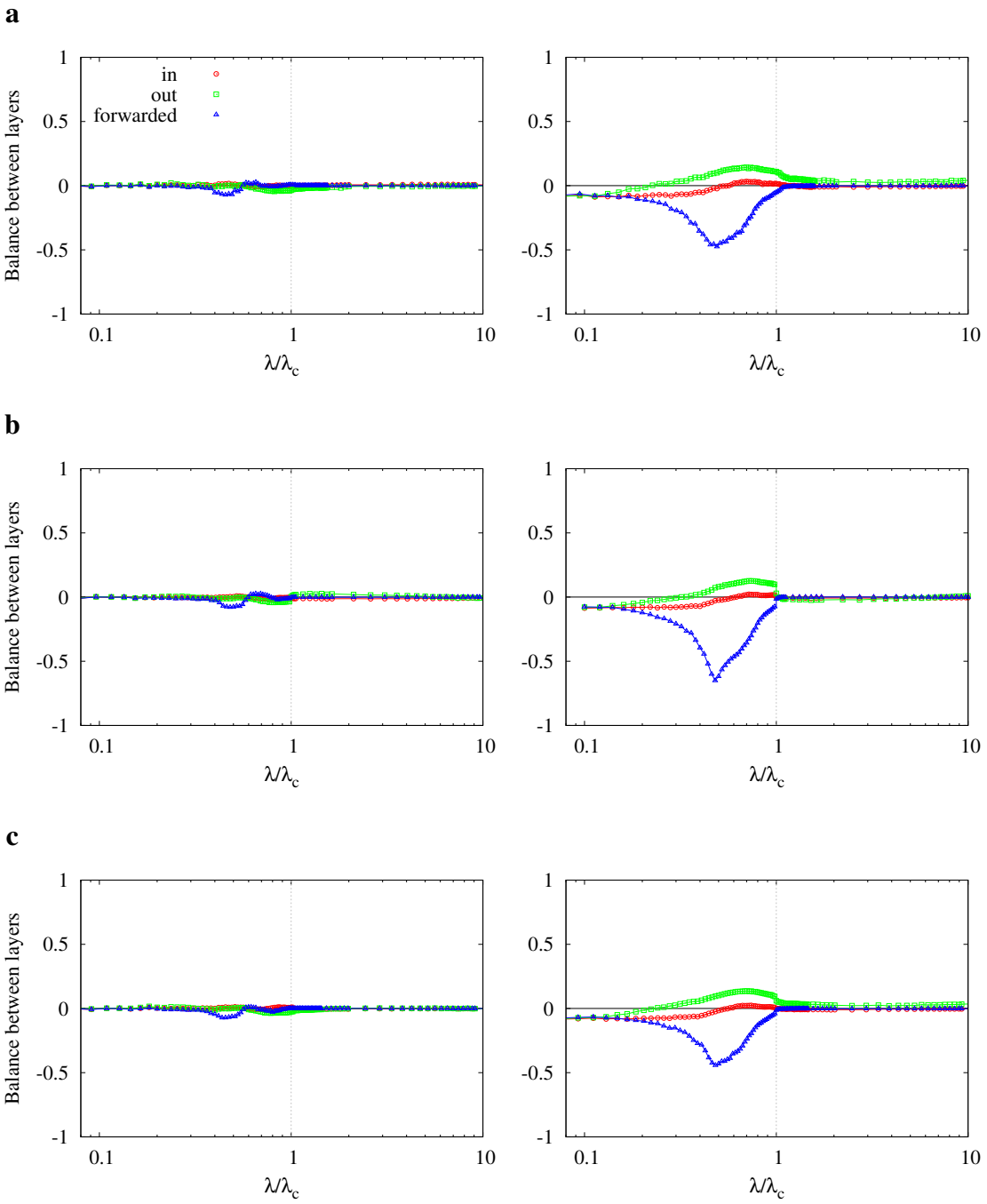


Figure 5.2: Balance of packets handled by each layer of a multiplex network formed by layer 1 and 2 (left side), and another network formed by layers 1 and 3 (right side), coupled in **(a)** assortative, **(b)** dissassortative and **(c)** random means. Results for the traffic-aware routing scheme with $h = 0.9$. Packets are separated as (in) new packets generated, (out) packets that arrive at their destination nodes, and (forwarded) packets dispatched along the path.

	Layer 1	Layer 2	Layer 3
Number of nodes	1 000	1 000	1 000
Number of edges	2 994	2 994	2 994
Average degree	5.98	5.98	5.98
Minimum degree	3	3	3
Maximum degree	86	73	105
Clustering coefficient	2.26×10^{-2}	3.41×10^{-2}	3.92×10^{-2}
Diameter	6	6	6
Average path length	3.51	3.51	3.46
Efficiency	3.01×10^{-1}	3.02×10^{-1}	3.06×10^{-1}

Table 5.1: Properties of three scale-free network layers.

this case, if we look at the number of packets that forwards each node we can differentiate an initial stationary free flow phase, when the generation of packets is very low, and where going through the layer with shorter paths is favored. Gradually, there appear some casual collisions and the adaptive scheme balances the traffic to one of the layers, favoring the one in which the nodes have better connectivity. With the increased traffic, temporary jams turn up, but are subsequently slowly dissipated by the system. At this point, the queues of nodes in both layers are saturated and the load imbalance decreases. Finally, when $\lambda > \lambda_c$, high posting rates produce a jammed state. After congestion onset, balance between layers is zero.

Last, we now look at where new packets are enqueued and in which layer they finish their paths. When the first queues are filled, preferentially those belonging to the layer with better communication, packets are sent to the less saturated layer where they reach the destination node. Similarly, the more congested is a layer, the fewer new packets receives.

5.3.2 Number of jumps between layers

More quantitative information can be gathered by weighting how often packets are sent to a different layer. Figure 5.3 depicts the frequency of jumps with respect to the number of packets forwarded in the network as function of the injection rate.

For the shortest path routing strategy the number of jumps is constant, corresponding only to those changes of layer that are part of a shortest path. Alike, in our adaptive scheme, when the injection rate is low the number of jumps match the shortest path, as there is no need to modify the routes.

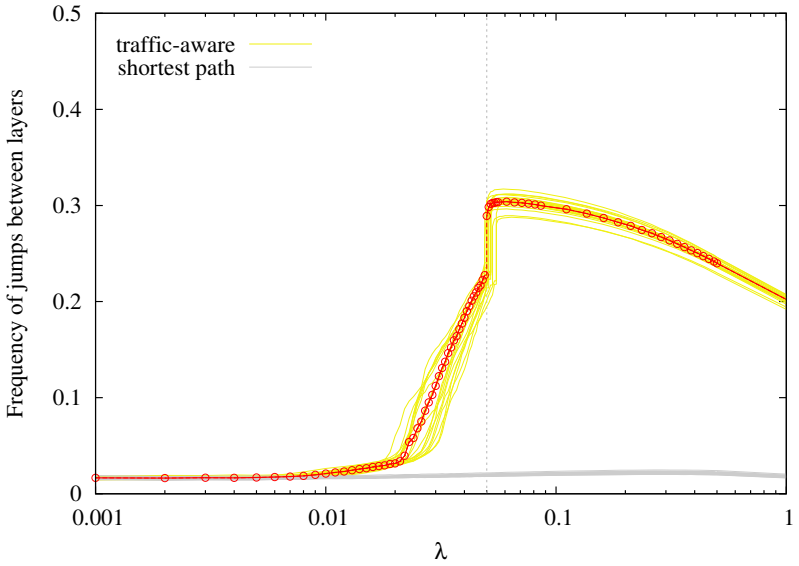


Figure 5.3: Ratio of jumps between layers and the number packets forwarded for different values of injection rate λ. We have illustrated the results for disassortatively coupled networks, and $h = 0.9$, but analogous outcomes are observed for other coupling configurations. For the sake of clarity, one case is highlighted in red.

When the generation of new packets is increased, we appreciate a slight migration of paths to avoid local random congestions. Then, from a certain point, which corresponds to the negative peak observed in Fig. 5.2, the number of jumps is increased to maintain the congestion controlled distributing the load between layers. Finally, when congestion arises, the routing scheme becomes extremely inefficient and the number of jumps tends to the frequency expected in a random walk.

5.3.3 Queue occupancy

It is also important to determine the number of packets waiting to be processed in each queue, given that this microscopic data shows where delays occur in the network.

Figures 5.4, 5.5 and 5.6 show the number of packets in each queue as function of λ, for each different type of coupling. Note that nodes are sorted according to the coupling preferences as described in 5.2.1. These matrices display the instantaneous usage of queues at the end of the simulations. Thus, lighter colors may represent only temporary jams result of traffic fluctuations.

The results show that when $h = 1.0$, i.e. the baseline shortest path strategy, a few number of queues accumulate a high number of packets. This saturation occurs in those nodes

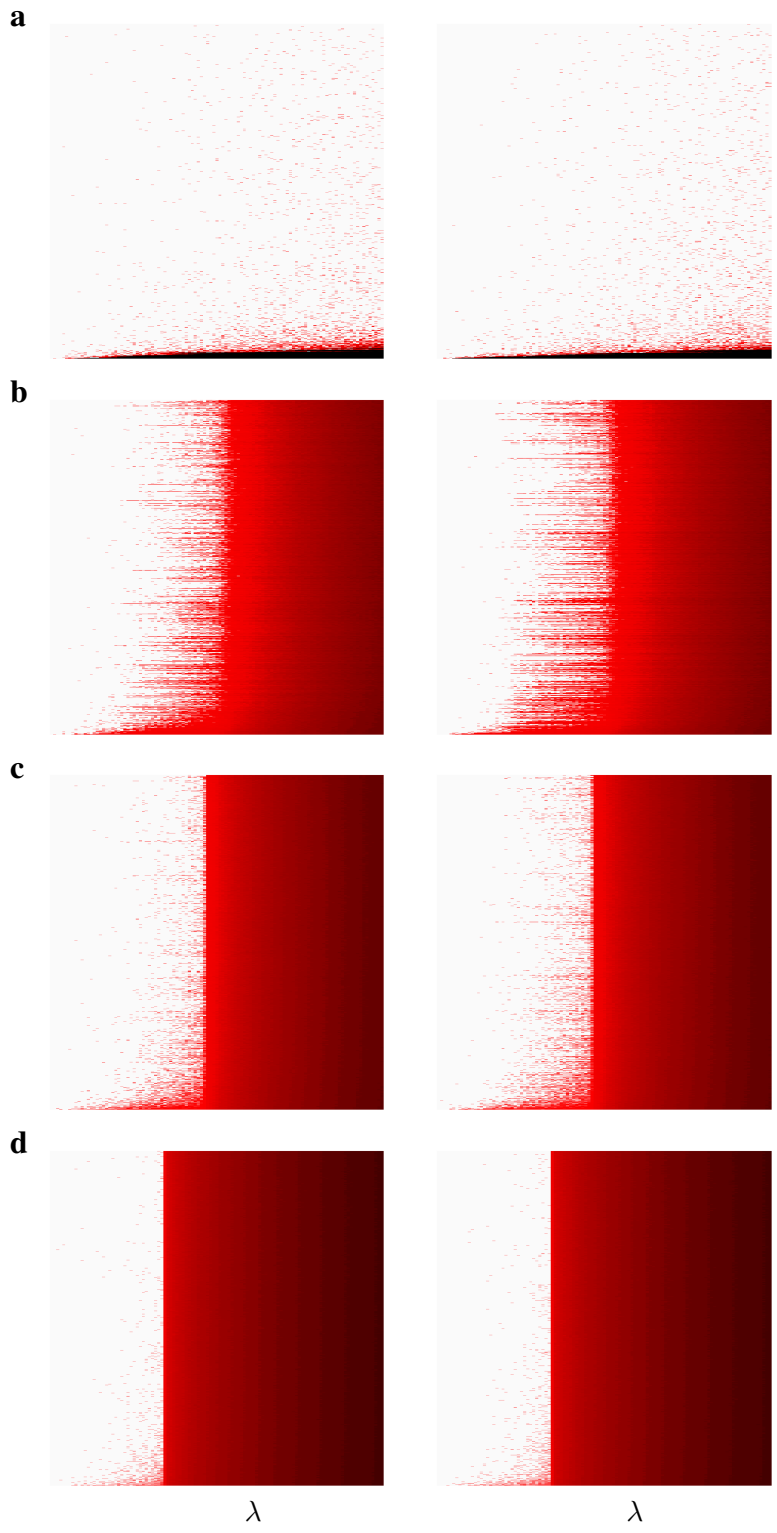


Figure 5.4: Matrices with the number of packets per queue as function of intensity paramter λ in an assortatively coupled multiplex network. Different routing configurations are presented for **(a)** $h = 1.0$, **(b)** $h = 0.9$, **(c)** $h = 0.7$ and **(d)** $h = 0.5$. Each row represents a queue in the first (left) and second (right) layers, and the number of items is graded in logarithmic scale, where darker colors represent a higher number of packets.

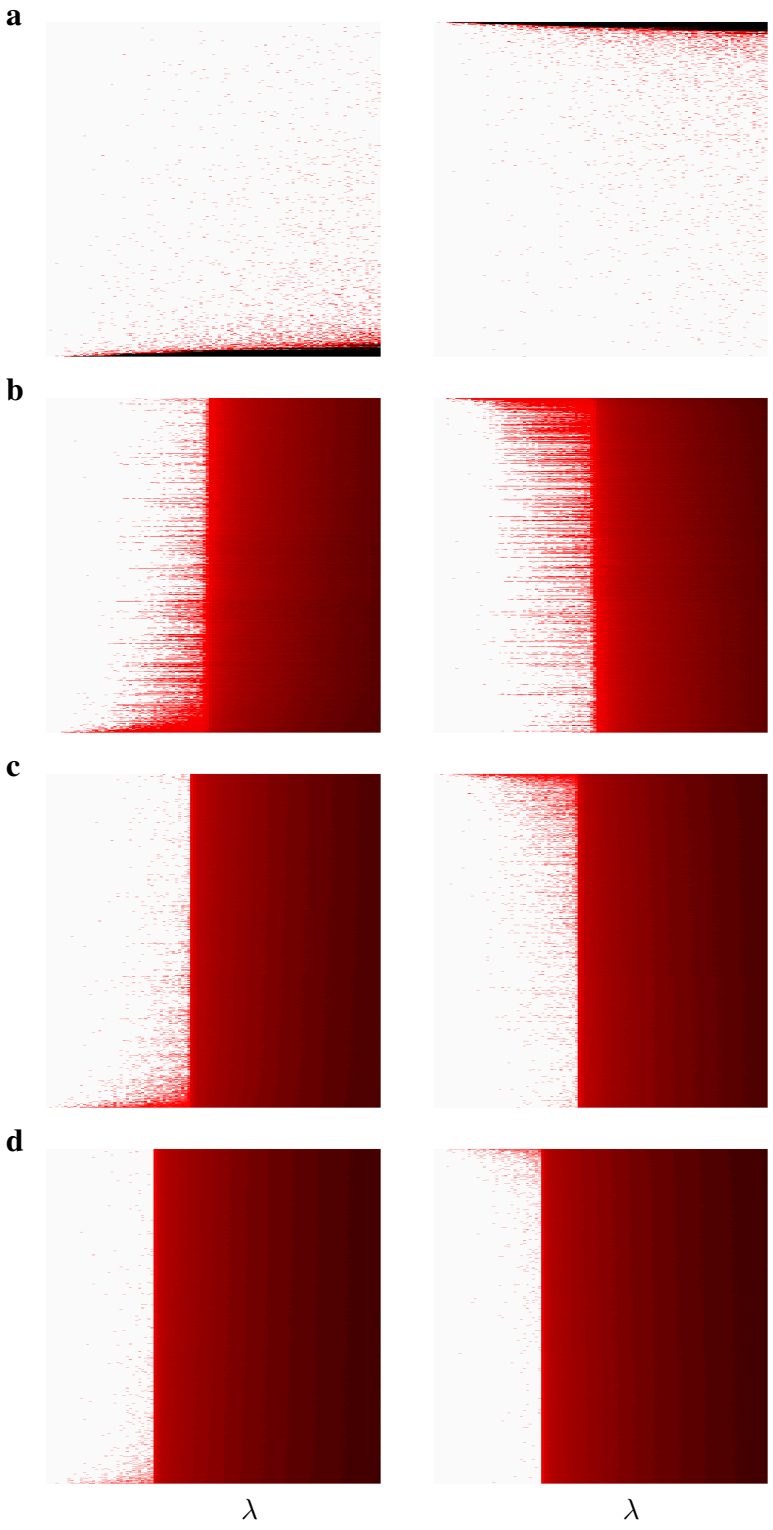


Figure 5.5: Matrices with the number of packets per queue as function of intensity paramter λ in a disassortatively coupled multiplex network. We present different routing configurations: **(a)** $h = 1.0$, **(b)** $h = 0.9$, **(c)** $h = 0.7$ and **(d)** $h = 0.5$. See Fig. 5.4 for details.

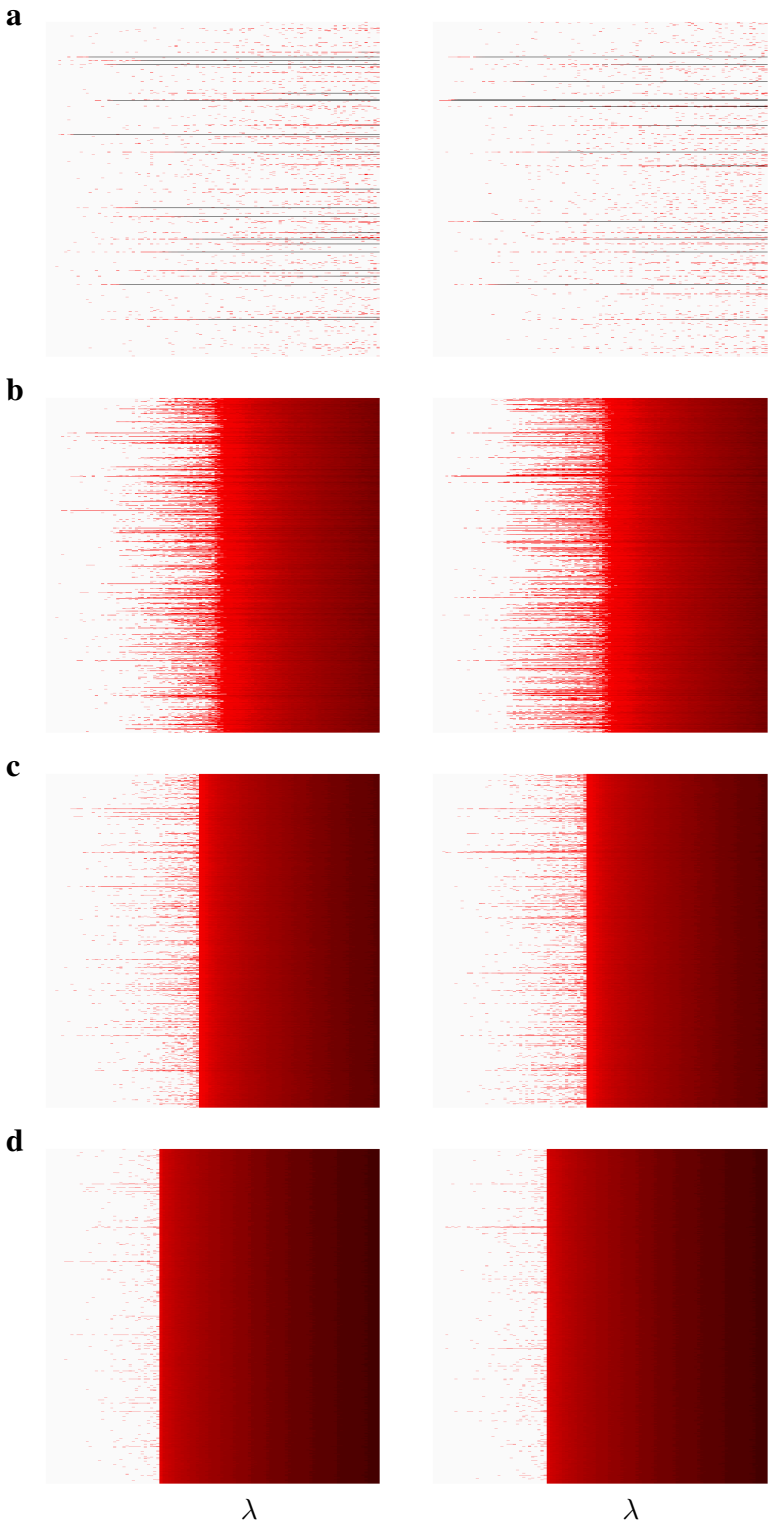


Figure 5.6: Matrices with the number of packets per queue as function of intensity paramter λ in a randomly coupled multiplex network. We present different routing configurations: **(a)** $h = 1.0$, **(b)** $h = 0.9$, **(c)** $h = 0.7$ and **(d)** $h = 0.5$. See Fig. 5.4 for details.

with higher betweenness centrality, which used to be related to the degree. The same behavior is noted in all coupling configurations.

On the contrary, when applying the adaptive algorithm, we can identify three flow regimes [200]. First, when the overall traffic is low all queues are empty. Then, there appears a region where some queues experience small temporary congestions that are slowly dissipated. In this stationary state, bottlenecks appear in nodes with greater centrality and the distribution of the number of packets is expected to be exponential [54]. With higher injection rates packets are redistributed throughout the whole network making the traffic load more uniform, and filling occasionally the queues of non-central nodes. Eventually, when $\lambda > \lambda_c$, packets accumulate promptly in all queues and the entire network, both layers, congests.

When we look at the effects of parameter h in the traffic-aware routing scheme, in general, we observe that for $h = 0.9$, when compared with lower values of h, the critical congestion point occurs later. Note also that the stationary free-flow phase is less evident for lower values of h, denoting that the redistribution of traffic is less effective. In these cases, packets are diverted following the less loaded nodes which increases the effective path length per packet. Hence, traffic is distributed more uniformly and temporary congestions diminish, but the overall performance deteriorates because of the longer distances traveled.

In the same spirit, when comparing different coupling correlations we observe that assortative coupling outperforms disassortative setups. Results of random models are somewhere in between. Having in mind that all nodes have the same capacity, this is attributable to the fact that the balance of packets from hubs to other hubs looks with favor on maintaining effective paths shorter since there is a wider range of routes. Conversely, in disassociative multiplex networks the hubs tend to receive more traffic from their counterparts with smaller degree.

5.3.4 Congestion transition

We end this section having a look at the transition from the free regime to the congested regime at global scale. To this end, we gauged the velocity at which the number of packets accumulate in the system by computing the slope b of $\mathcal{N}(t)$, the total number of packets in the system at time t. Slope b is the regression coefficient obtained using linear

least squares fitting, given by

$$b = \frac{\text{cov}(t, \mathcal{N}(t))}{\sigma_t^2} = \frac{\text{ss}_{t\mathcal{N}(t)}}{\text{ss}_{tt}}\,, \tag{5.4}$$

where $\text{cov}(t, \mathcal{N}(t))$ is the covariance, σ_t^2 the variance, and $\text{ss}_{t\mathcal{N}(t)}$ ss_{tt} are sums of squares. When all the packets are delivered the average slope is 0, while for $\lambda > \lambda_c$ it is larger than 0.

We also calculated the global order parameter ρ, introduced by [11], that refers to the ratio between undelivered and generated packets, i.e. the percentage of unabsorbed packets for unit time. It is computed as

$$\rho = \lim_{t\to\infty} \frac{\mathcal{N}(t+\tau) - \mathcal{N}(t)}{N\lambda\tau} \tag{5.5}$$

where $N\lambda$ is the rate of creation of packets in the system and τ is the observation time.

These results are depicted in Figs. 5.7, 5.8 and 5.9, which correspond to assortative, disassortative and random couplings, respectively. As the analysis of queue occupancy anticipated, regression coefficient b shows that in shortest path strategy congestion appears early and in a progressive way. Notice also that it seems to be immune to variations in the coupling preference.

In contrast, the algorithm that integrates traffic information delays the congestion onset, in accordance with the results reported for single-layer networks [61]. But when traffic increases too much, transition to jammed state may arise abruptly and information packets accumulate faster. It is therefore natural to ask up to what point the adaptive routing improves the shortest path strategy. It is clear that after both curves cross, the inclusion of traffic information is not an advance.

When looking at dynamics see that congestion emerges early for lower values of h. In addition, we appreciate that for higher values of h the region of b under the shortest path curve is longer, whereas for smaller values it is surpassed at lower injections rates. The influence of h on the length of routes described previously explains this performance.

Finally, we also observe that assortative coupling may delay congestion onset and reduces its severity. It is worthy to mention the case for $h = 0.9$, where we do not appreciate an abrupt transition to the jammed phase. We can associate this behavior with the slow gradient of the stationary free-flow phase observed in Fig. 5.4.

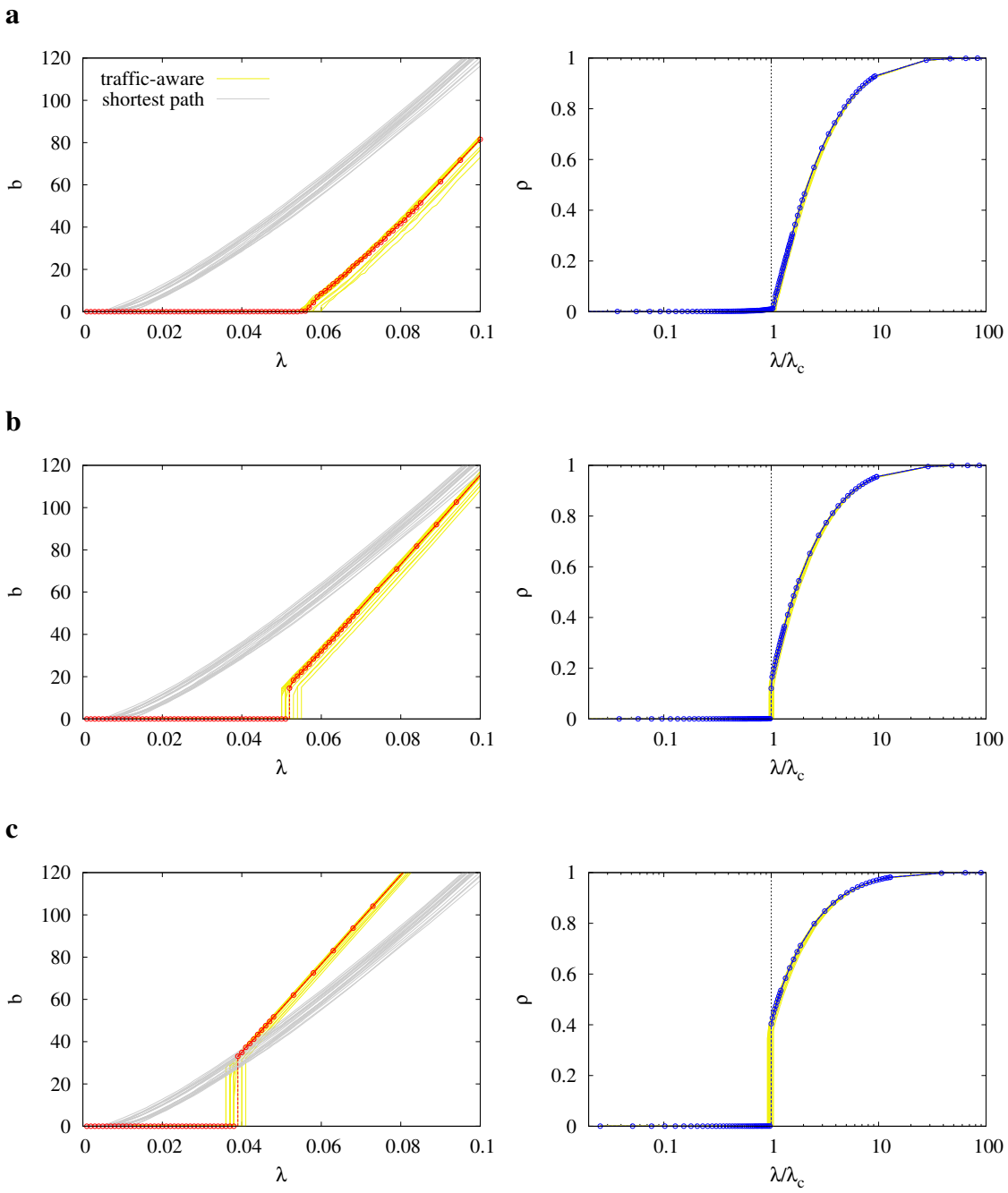


Figure 5.7: Global phase transition from sparse to jam state in multiplex networks with assortative coupling, for different adaptive routing setups: **(a)** $h = 0.9$, **(b)** $h = 0.7$, **(c)** $h = 0.5$. Left side plots depict the regression coefficient b of $\mathcal{N}(t)$ for different generation rates. On the right, we present the classical representation of the global parameter ρ as function of the scaled intensity λ by the critical congestion rate λ_c. We highlight two random cases in red and blue.

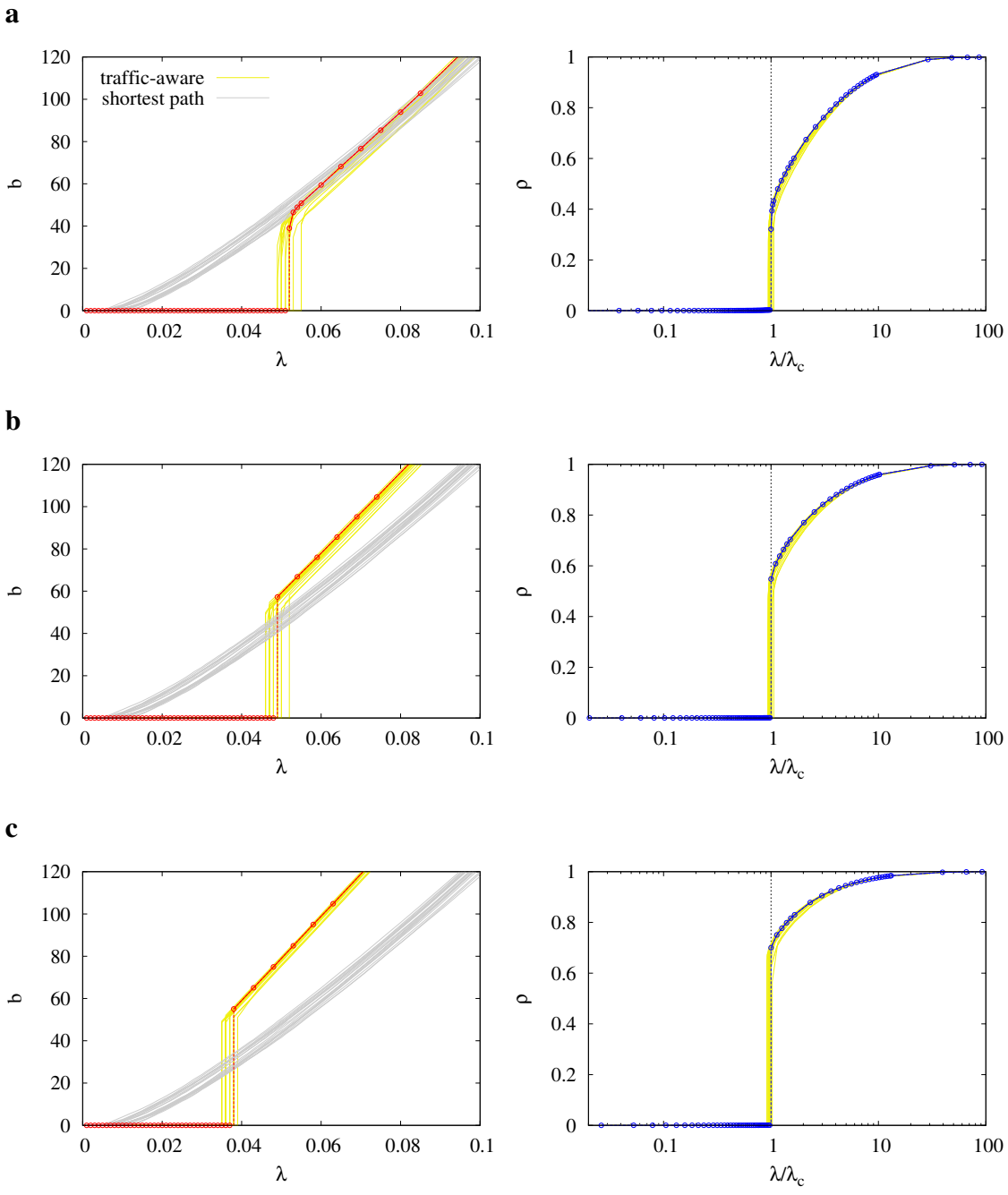


Figure 5.8: Global phase transition from sparse to jam state in multiplex networks with disassortative coupling, for different adaptive routing setups: **(a)** $h = 0.9$, **(b)** $h = 0.7$, **(c)** $h = 0.5$. See details in Fig. 5.7.

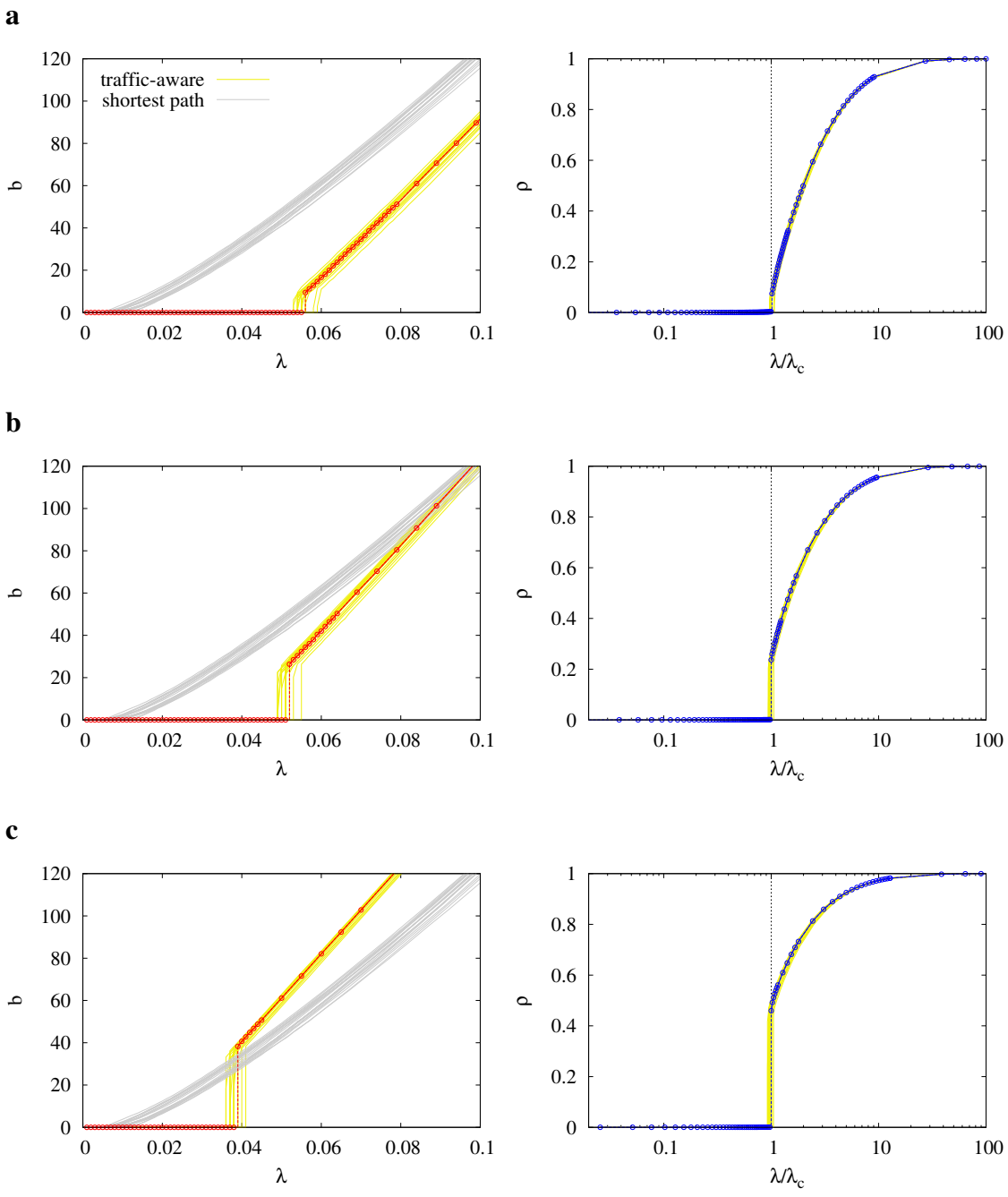


Figure 5.9: Global phase transition from sparse to jam state in multiplex networks with random coupling, for different adaptive routing setups: **(a)** $h = 0.9$, **(b)** $h = 0.7$, **(c)** $h = 0.5$. See details in Fig. 5.7.

5.4 Conclusions and discussion

The growing interest in multiplex networks led us to study in this chapter a traffic-aware routing strategy [61, 62] for multiplex communication networks. To this end, we have constructed multiplex networks with two scale-free layers coupled in assortative, disassortative and random forms. The adopted routing solution requires only information on the traffic status of direct neighbors, making it feasible to be implemented.

The results obtained show that in comparison with the shortest path strategy, that only considers topological information, including traffic information improves the delivery performance and delays the congestion onset. However, the phase transition to the jammed state appears abruptly, and this may not allow the system to self-organize at the edge of criticality, having catastrophic consequences for the network [54]. Adjusting the parameter h, that controls the sensitivity to traffic load, or limiting the second term of Eq. (5.3), may be strategies to overcome this issue.

On the other hand, in single layer networks, star-like structures are particularly vulnerable to congestion due to their centrality [103]. Analogously, it would be reasonable to think that in assortatively coupled networks hubs might become bottlenecks of the network. However, we have shown that our adaptive routing performs better when layers are coupled with assortative preference. This non-trivial effect is in accordance with the results presented in [203, 220]. As Fig. 5.4b suggests, the redistribution of information packets in assortatively coupled networks is better due to the fact that hubs exploit their neighborhood nodes and do not saturate their counterparts.

Our analysis also suggests that information packets migrate to those layers that have better navigability because they are topologically more efficient. At higher injection rates, however, these layers become overloaded and the adaptive routing scheme fairly balances the traffic between both layers. Similar observations have been reported by Solé-Ribalta et al. [193].

Lastly, we want to point out that some authors have stated that real-world multiplex networks tend to be sparse [156]. However, our work assumes that each node of a multiplex network has connections in all layers. Our approach is clearly a simplification, although widely adopted, and may need further research.

UNIVERSITAT ROVIRA I VIRGILI
NAVIGABILITY AND SYNCHRONIZATION IN COMPLEX NETWORKS: A COMPUTATIONAL APPROACH
Pau Erola Cañellas

6 Modeling financial crises: synchronization of integrate-and-fire oscillators

In September 2008, the bankruptcy of Lehman Brothers marked for many the beginning of the current global crisis. Why the weakness of the North American financial system spread world wide? Was it predictable?

Although there is no clear answer to these questions, earlier economic studies had already pointed out the relevance of the structure of the international trade system in understanding how crises spread [93]. Therefore, trade policies and agreements, and advances in transportation and communication systems, that have made the world smaller, could have been important factors in economic perturbations contagion.

Here, we will use these premises to propose a model based on oscillators to assess the effects of the evolution of the international trade network in the spreading of financial crises.

Next section introduces the network of world trade and its properties. Sections 6.2 and 6.3 present the framework of our model, and in Section 6.4 we show the outcome of our simulations. Finally, Section 6.5 provides the conclusions.

6.1 World trade as a complex network

The World Trade Web (WTW) represents the international trade transactions between different countries as a network. This network is formed by N nodes, one for each country, and L links that correspond to trade flows. Each link has an associated weight w_{ij} that quantifies the trade relationship between two countries i and j. Thus, the link

w_{ij} accounts for the exports of node i to node j, and w_{ji} for the imports of i from j.

The first empirical characterization of the WTW as a complex system was presented by Serrano and Boguñá [181]. In an unweighted representation, the WTW exhibited a small world property and showed a scale-free degree distribution, high clustering and degree-degree correlation between different vertices. Garlaschelli and Loffredo [88] studied also the relation between the topological properties and GDP of countries and its temporal dependence.

The weighted representation of the WTW was examined by Fagiolo, Reyes and Schiavo [72, 73, 74]. They found that the WTW is an extremely symmetric network (relationships tend to be reciprocal), yet there exists a core-periphery structure where a group of countries hold more intense trade relationships.

In this chapter we will use six snapshots of the WTW [92] corresponding to years between 1950 and 2000. We will pay attention also to the modular structure of these snapshots. Given that the WTW is a weighted directed network, the mathematical definition of modularity [14, 151] expressed in terms of the weighted adjacency matrix w_{ij} is

$$Q = \frac{1}{2w} \sum_{i,j} \left(w_{ij} - \frac{s_i^{\text{out}}\, s_j^{\text{in}}}{2w} \right) \delta(C_i, C_j)\,, \tag{6.1}$$

where s_i^{in} and s_i^{out} are the input (imports) and output (exports) strengths respectively, the Kronecker delta function $\delta(C_i, C_j)$ takes the values, 1 if node i and j are into the same module, 0 otherwise, and the total strength is $2w = \sum_{i,j} w_{ij}$.

The analysis of the mesoscopic scale of the WTW reveals different topological communities for different snapshots. This is eventually an expected fact because the trading weights have changed over years, and consequently the topology of the WTW. In Fig. 6.1 we show the world-map labeled according to the best partition found for each year (see Appendix for the detailed list of countries in each community). Likewise, Table 6.1 presents the number of communities, their size in number of countries, and their relative strength according to import-export data.

In the following sections we use these data, microscopic and mesoscopic structures, to analyze how the structural evolution can affect the spreading of financial crisis.

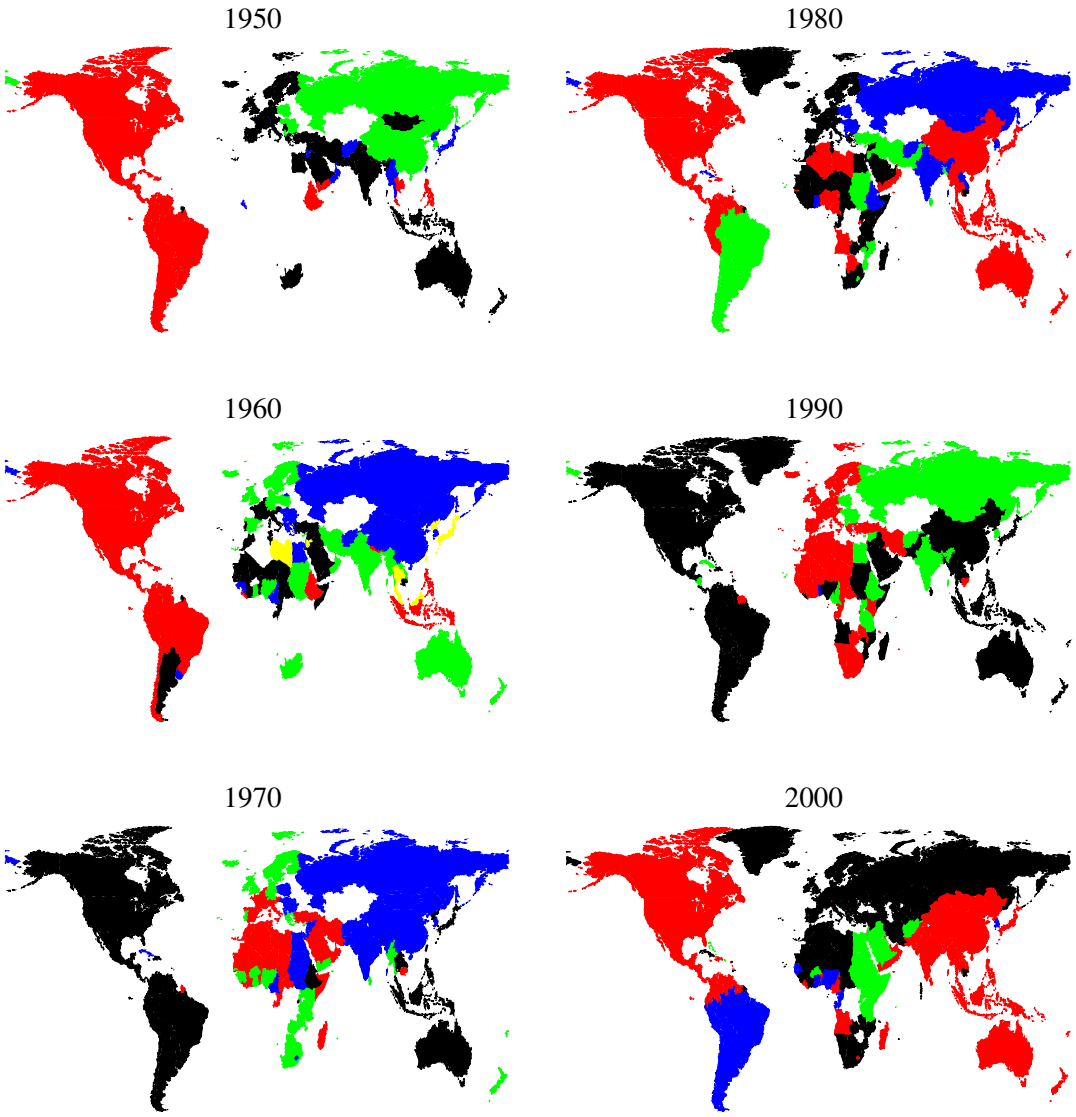


Figure 6.1: Topological communities of the WTW found by modularity maximization [171]. Each color corresponds to a different community. Since the community structure evolves through the years, the color of a certain country may change through time. Regions in white correspond to countries without information in the dataset, or to regions belonging to other countries.

year	N	D	$\langle k \rangle$	$\langle s \rangle$	$Q_{\max}$	M	N_α
1950	83	0.1753	28.75	705.99	0.4519	4	[37, 26, 10, 10]
1960	113	0.1670	37.42	1011.47	0.3312	5	[36, 25, 25, 18, 9]
1970	140	0.2015	56.03	2284.69	0.3375	4	[41, 38, 36, 25]
1980	162	0.2080	66.96	12367.65	0.2669	4	[65, 58, 20, 19]
1990	169	0.2239	75.23	20554.50	0.2763	4	[78, 63, 26, 2]
2000	187	0.2833	105.40	36758.72	0.2585	4	[88, 66, 18, 15]

Table 6.1: Characteristics of the WTW network per year: N, number of nodes (countries); D, link density, i.e. number of edges divided by $N(N-1)$; $\langle k \rangle$, mean degree; $\langle s \rangle$, mean strength; $Q_{\max}$, maximum modularity found; M, number of communities; N_α, sizes of the communities.

6.2 Financial crisis, business cycles and globalization

Let us begin to sharpen the discussion by defining a qualitative feature of economic time series called *business cycle* or *economic cycle*. This term is used to characterize the fluctuations in economic activity experienced by a nation over time, i.e. expansion and recession periods. During this alternation between growth and contraction, a sudden downturn may be indicative of a financial crisis.

The study of financial crises has always attracted a fair amount of interest, but we still know very little about them and seem neither predictable nor avoidable *ex ante*. Post-Keynesian Minsky's financial instability hypothesis [139, 140] asserts that a crisis is triggered when concur falls in economic activity, poor management of bank credit and systemic liquidity requirements. That is to say, it emphasises the destabilising effects of interest rate policy and credit regulation, and defines financial crises as a natural consequence of changes in the economic cycle and the fragility of the structure of debt. In a capitalist economy the financial system swings between robustness and fragility and does not rely upon exogenous shocks to generate business cycles of varying severity.

According to this definition, a global crisis seems to be highly unusual, and will only occur when declining economic cycles of all economies are synchronized. But in spite of this, in September 2008 the bankruptcy of Lehman Brothers spread the North American financial crisis relentlessly world wide. The magnitude of this current economic crisis is hard to explain and we will sustain our work on theories that point to a possible relationship between the globalization process and the synchronization of economic cycles.

Deardorff [55] defines globalization as "the increasing world-wide integration of markets

for goods, services and capital", stressing the idea that globalization is affecting the world as a whole. Global trade and financial flows have been increasing along the last century; but around mid eighties it has been observed an acceleration of this process, which has been identified with the appearance of global trade. Free trade and increased financial flows can set the channels to spill shocks world wide [93].

Several recent studies have analyzed the change of economic cycles on an evolving internationally connected world. Kose, Otrok and Prasad [119] have studied the evolution of business cycles over the globalization period 1985-2005, finding some convergence of business cycle fluctuations among industrial economies and among emerging market economies. Instead, they find little changes in the degree of international synchronization.

There is also evidence of synchronization phenomena of economic cycles in the WTW. Using correlation trade-data analysis, and the substrate of the WTW, Li, Ying Jin and Chen [133] showed the presence of strong correlated synchronized behavior, without imposing any dynamics on the system. He and Deem [108] have also addressed the response of the WTW to recessions. They found a major trend in the reduction of the hierarchy of the network (attributed to globalization) that makes the trade network less robust, i.e. high modularity plays a protective role.

Here we will rely our work in the paradigmatic synchronization of oscillators in complex networks theory, which has attracted lot of attention in the physics literature [12]. Nevertheless, there is no model of economic cycles in the WTW reporting such synchronization phenomenon.

In next section we propose a simple model to analyze how the effects of globalization in the WTW can affect the ability of synchronization of business cycles. To this end, we have chosen six snapshots of this network between 1950 and 2000, some for the pre-globalization period (1950, 1960, 1970 and 1980) and the rest belonging to the globalization period (1990 and 2000). In an increasingly globalized world, we can expect that more tightly coupled countries synchronize quickly. We also analyze how this synchronization is driven, at the mesoscopic scale, by the existence of modules in the network representing stronger associations of trading between certain groups of nations.

6.3 A simple model of economic cycles in the WTW

Our concern is if the emergence of trading relations is conductive to a global financial instability. With this aim, the abstraction we propose is that each country has its own economic cycle represented by an oscillator, which we assume evolves sooner or later

towards a sudden dramatic depression leading to a financial crisis [66, 67]. And the interaction between these oscillators is performed through the commercial channels represented by the links of the WTW network.

To this end, in this study we use the simple Mirollo and Strogatz *Integrate-and-Fire Oscillators* (IFO) model [141]. To mimic a period of expansion up to peak and a deep fall to tough, each oscillator is characterized by a monotonic increasing state variable $x \in [0, 1]$ that evolves according to

$$x = f(\phi) = \frac{1}{b} \ln(1 + (e^b - 1)\phi) \,, \tag{6.2}$$

where $\phi \in [0, 1]$ is a phase variable proportional to time, and b is the dissipation parameter that measures the extent to which x is concave down; when b approaches zero f becomes a linear function. We can calculate ϕ using the inverse function

$$\phi = f^{-1}(x) = \frac{e^{bx} - 1}{e^b - 1} \,. \tag{6.3}$$

As Fig. 6.2 depicts, hen the variable x attains the threshold $x = 1$ it is said to *fire*, and it is instantly reset to zero, after which the cycle repeats. Now let us assume that a node i of our network of oscillators fires. This node, in turn, transmits to all its neighbors j an excitation signal of magnitude $\varepsilon_{ij} > 0$, thus leaving their state variables x_j^+ with values

$$x_j^+ = \begin{cases} x_j + \varepsilon_{ij} & \text{if } x_j + \varepsilon_{ij} < 1, \\ 0 & \text{if } x_j + \varepsilon_{ij} \geqslant 1. \end{cases} \tag{6.4}$$

If $x_j + \varepsilon_{ij} \geqslant 1$, oscillator j is also reset and propagates the fire signal to its neighbors, thus generating cascades of fires. This is, at the moment a nation *fires*, it propagates the problem to other countries through its connection in the WTW by boosting their own evolution to a crisis.

The WTW presents a large diversity in the economic weight of countries and their trade flow. It is reasonable to think that a shock in a small country is not spread to a large country with the same intensity than vice versa. And regardless of country size, the volume of transactions appears to be a factor of the ability of trade channel to synchronize the cycles of two economies. To reflect this dependence we set the excitation signal of node i to its neighbors j as

$$\varepsilon_{ij} = \frac{w_{ji}}{s_j^{\text{out}}} \,, \tag{6.5}$$

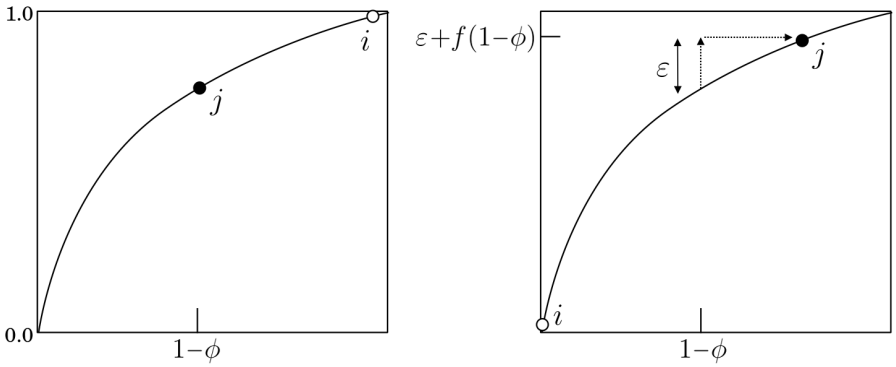


Figure 6.2: A system of two Integrate-and-Fire oscillators. When node i reaches the threshold value (left), it propagates an excitatory signal to its neighbor j and i is reset to zero (right). Adapted from [141].

which is the fraction of total exports of country j going to the firing node i. For the sake of simplicity, we will assume identical oscillators, isolating the effect of synchronization of crises to the dynamics of interactions weighted by the trading exchanges.

6.4 Simulation analysis

The structure of the WTW, a densely connected weighted directed graph, makes certainly difficult to assess *a priori* the global outcome of different economic cycles interacting through pulses. No general theory exists to ascertain the stationary state, if any, of such a network of pulse-coupled oscillators. For this reason we will rely on computer simulations to shed some light on the behavior of this complex system. Previous studies of complex networks of pulse-coupled oscillators [40, 207] have shown the coexistence of synchronized and non synchronized dynamical regimes when the coupling ε is positive (excitatory) in homogeneous and heterogeneous in degree networks. Also recurrent events of synchrony and asynchrony have been reported in complex networks when interactions have delay and refractoriness [176]. Here we expect all this phenomenology to show up in our particular network, with essential differences between the snapshots at different years of the WTW.

For the simulations, we assign an oscillator per country with random initial phase. The system evolves following the dynamics given by Eqs. (6.2) to (6.4), with $b = 3$ and ε as described in Eq. (6.5). The simulation is stopped when 90% of the oscillators are synchronized, and the statistics shown have been obtained after 10^3 realizations.

As anticipated, we will pay also attention to the modular structure of the different

snapshots of the WTW detected using modularity (see Section 6.1) to investigate the effects of globalization on the spreading of financial crises. The modularity analysis revealed different topological communities for the different years, and according to previous studies [13, 216] is reasonable to expect a more or less clustered synchronization behavior.

For the evaluation of the interaction dynamics, we follow the classical analysis of the order parameter r that indicates the degree of synchronization of a system with N oscillators. We recall that in our setup, synchronization will reflect the scope of the crisis at the international level. The order parameter is

$$r = \left| \frac{1}{N} \sum_{j=1}^{N} e^{2\pi i \phi_j} \right| = \sqrt{\left(\frac{1}{N} \sum_{j=1}^{N} \cos(2\pi\phi_j) \right)^2 + \left(\frac{1}{N} \sum_{j=1}^{N} \sin(2\pi\phi_j) \right)^2}, \qquad (6.6)$$

which only depends on the phases ϕ_j of the oscillators. This parameter r is 1 for complete phase synchronization, and close to 0 in the incoherent state.

The rest of this section presents the results of the simulation. First, we examine the dynamics of the IFO model at the global (worldwide) scale, and after, we fix our attention to the evolution of the same observables at the mesoscopic scale.

6.4.1 Synchronization time

The first global aspect we investigate is that of the time needed for the whole network to synchronize. To this end, we track the time needed to synchronize 90% of nodes of the network.

Figure 6.3 presents the cumulative distribution of time up to synchronization. Interestingly, the distribution shows that as the WTW evolves along years, the time for almost global synchronization is faster. Specially interesting is the change of slopes observed between 1960–1970 and 1980. From the data obtained, our analysis suggests that the globalization process should take place within this two decades, producing a clear differentiation between the period before 1950–1960 and the posterior 1990–2000.

6.4.2 Cascades of fires

Following the analysis, we consider the scope of a certain firing event on the rest of the network. We will call it cascade or avalanche.

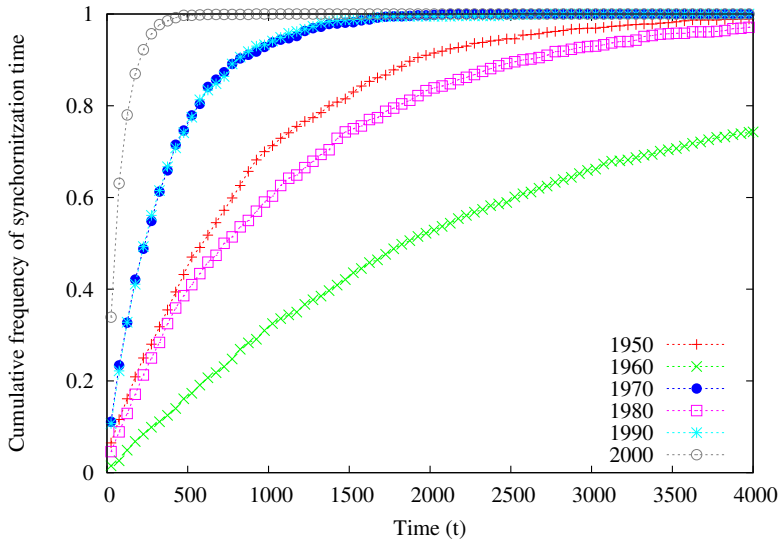


Figure 6.3: Cumulative frequency of synchronization time, i.e. the time needed to synchronize 90% of the nodes, for 10^3 repetitions of the IFO dynamics. The width of the bins in these histograms is of 50 time cycles.

Figure 6.4 shows the probability distribution of sizes of cascades. These plots add more information to the previous discussion: as years go on, the WTW follows a cascade distribution that tends to a power law intermediate regime plus the obvious cutoff due to the finite size of the system. The power law structure of the distribution starts to be significant from 1970. The consequence of this power law is the absence of a characteristic scale of sizes of cascades in the system and implies that, in principle, the assessment of the scope of international crisis becomes unpredictable.

6.4.3 Synchronization boundaries

Having analyzed the global outcome of the IFO dynamics on the WTW, we pay attention to the development of cascades restricted to the topological communities found (see Section 6.1). In Fig. 6.5 we show a raster plot of firings for each country, labeled with a color representing its topological community. Qualitatively, we detect that the communities behave more independently for the years 1950–1960–1970, a pre-globalization period, and more coherently between them for the 1980–1990–2000 periods.

This fact is confirmed in Fig. 6.6, where we show the evolution of the synchronization parameter for each community (named r_α) and the same parameter for the whole network (named r) as a function of time. In this plot it can be observed that before and in 1970 (included) the synchronization parameter r_α presents larger fluctuations than the

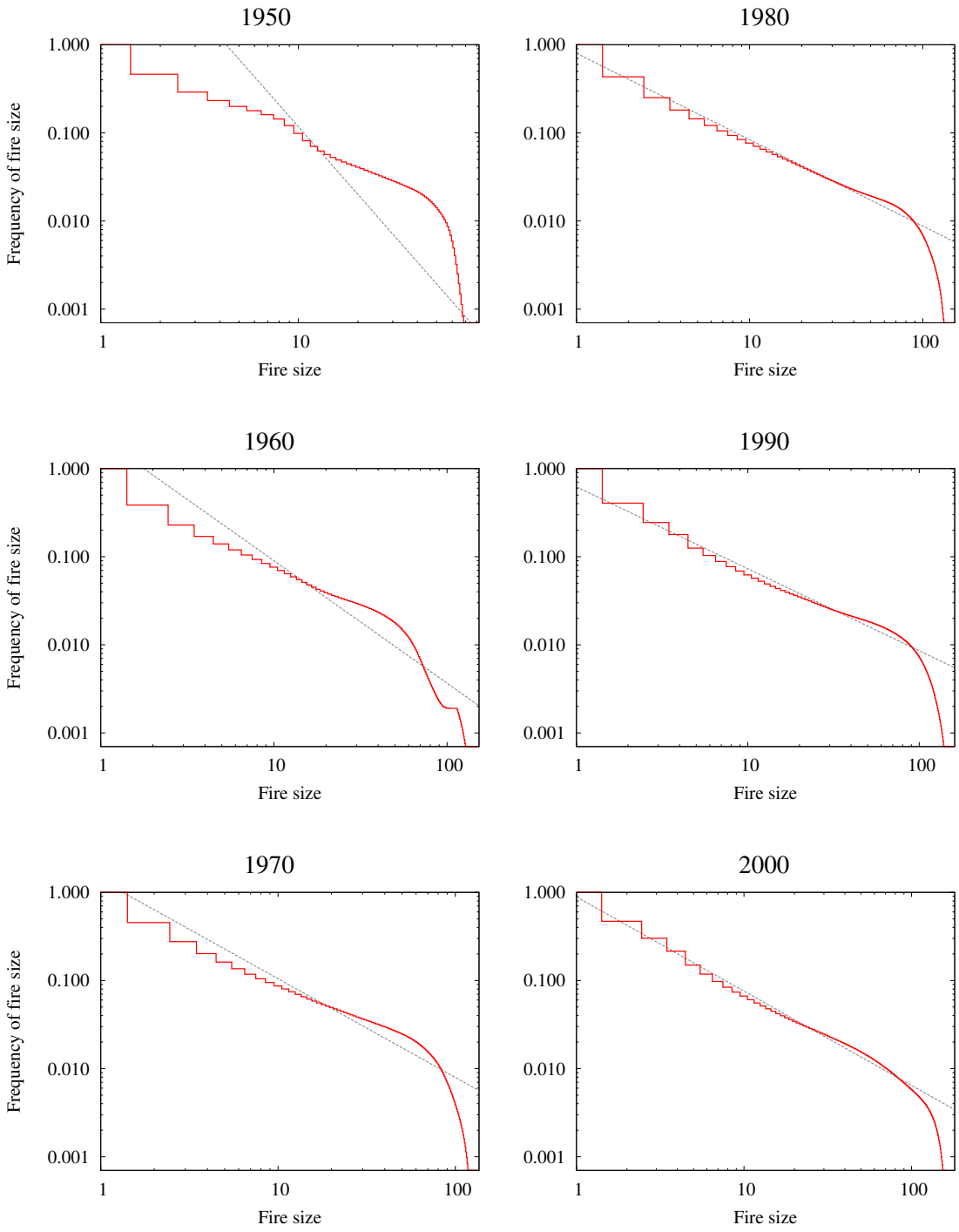


Figure 6.4: Frequency of firing size, i.e. the number of nodes involved in firing cascades, in log-log scale and for 10^3 repetitions of the IFO dynamics. The straight lines are the best power-law fits to the data.

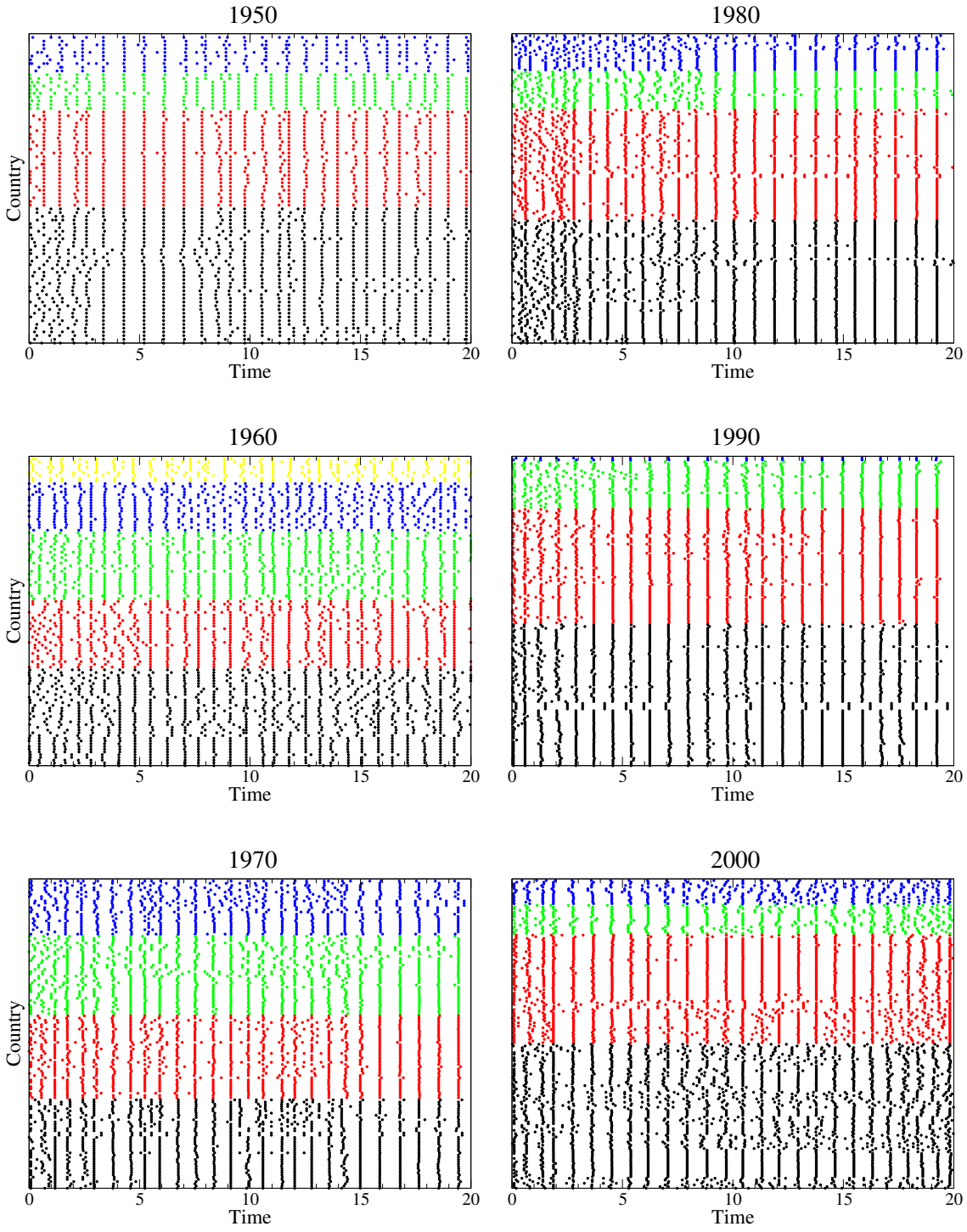


Figure 6.5: Firing times of the nodes in the first 20 time cycles of a single run of the IFO dynamics. Colors correspond to the communities in Fig. 6.1.

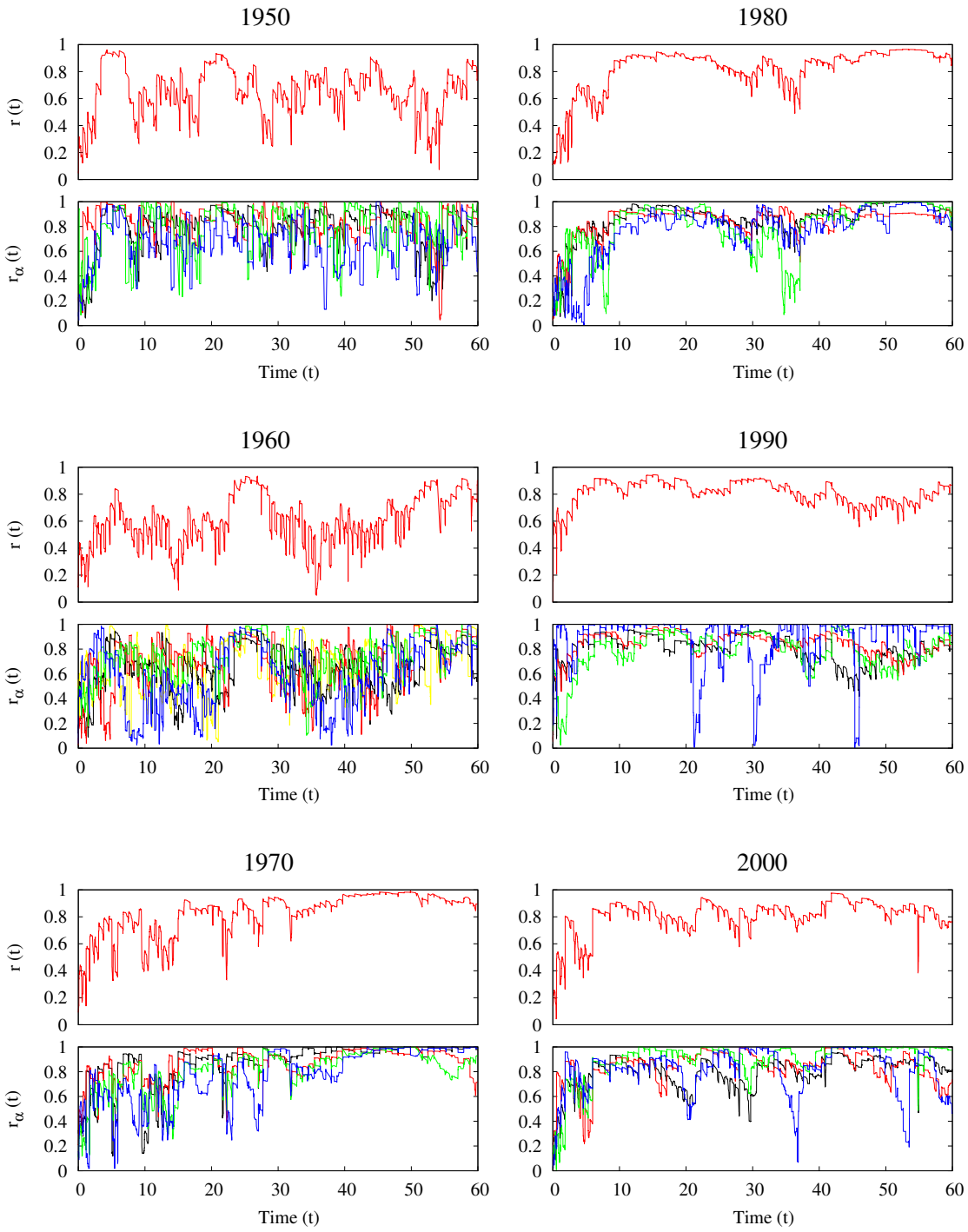


Figure 6.6: Evolution of the global and communities synchronization parameters, r and r_α respectively, in the first 60 time cycles of a single run of the IFO dynamics. Colors in r_α plots correspond to the communities in Fig. 6.1.

global synchronization r. Moreover, it is clear that periods of synchronization and desynchronization of economical crises coexist [40, 207] among different communities at the same time, and with the global measure. After 1970, the tendency is that of having a high level of synchronization in time with a very few desynchronization periods.

Finally in Fig. 6.7, we present the difference of the values of r_α of each community in front of that of all the network, which reveals the discrepancies between the mesoscopic view and the global view. Again, after 1970 the signs of globalization are clear, the dynamic effect of the mesoscopic structure is practically collapsed to the global world scale behavior.

Withal, these observations allow us to conjecture that the effect of topological borders of the communities have almost no effect after globalization emerges.

6.5 Conclusions and discussion

So long, we have presented a simple dynamical model of economic cycles interaction between countries in the WTW network. The model represents each country with an integrate-and-fire oscillator that emit a pulse at the end of the expansion of its economical cycle. We have proposed a longitudinal study to assess the effects of the mesoscopic structure of the WTW as time evolves.

The results presented support the theory of a globalization process emerging in the decade 1970–1980, in accordance with [108], and that the synchronization phenomena after this period accelerates. Then, the WTW evolves towards a more globalized trading system, where the topological mesoscopic structure is less and less representative of particularly different dynamical behavior, being the global phenomenology equivalent to the mesoscopic one, and vice versa. Unfortunately, the more dissipated the community borders are, the harder is to predict the scope of an international crisis.

The extension of the presented model with non-aggregated data, i.e. trade data between countries disaggregated across commodity categories [209], could enlighten novel local properties. In this trend, for instance, Ikeda et al. [113] used coupled Kuramoto oscillators to study the synchronization stability of the Japanese industry sectors and goods markets under shocks. Likewise, the refining of the model presented via data driven approaches or by introducing more heterogeneity on the state of nodes can be a good way to investigate the scope of world crisis.

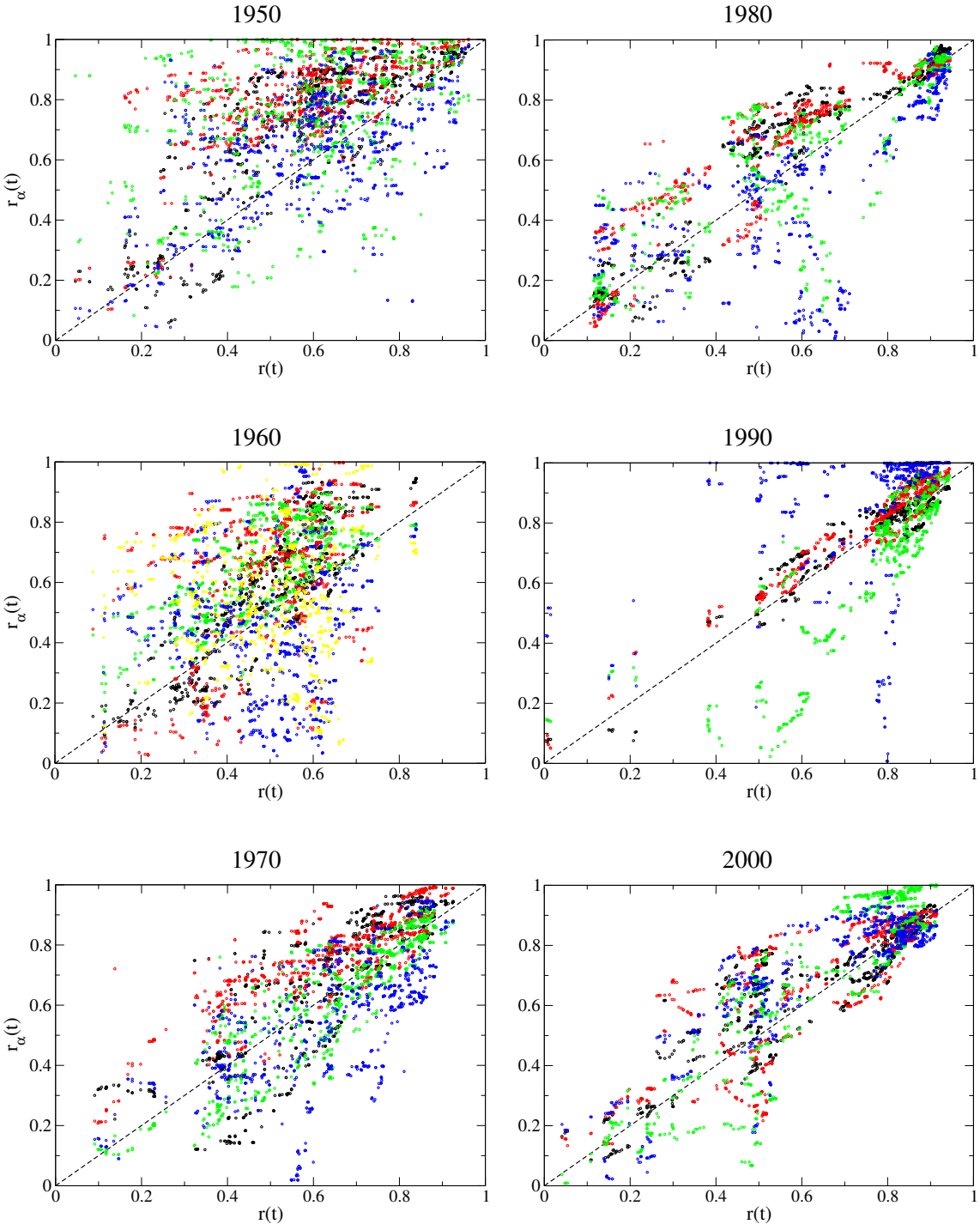


Figure 6.7: Deviation of the synchronization of the communities r_α in front of the global synchronization r, corresponding to the data in Fig. 6.6. Colors correspond to the communities in Fig. 6.7.

7 Conclusions and perspectives

The work developed in this doctoral thesis addresses computationally many issues on large complex systems. The first part of this document was partially motivated by the rapid development of the Internet, from growth to the increasing traffic demand, and turns around the navigability in complex networks. The second part, is a reflection of the actual world wide socio-economical situation and our interest in understanding financial crises, and it has been proposed as a case study of the synchronization phenomena in the global trade market. We have tackled this broad scope of problems with simple approaches that may facilitate knowledge extraction and generalization of the models. The main results that we have obtained are described below.

7.1 Conclusions

In Chapter 3 we have studied the reliability of TSVD projections of complex networks. TSVD has proven very successful in many machine learning applications to reduce the dimensionality of big datasets and, similarly, can be useful to extract information of the structure of large complex networks. However, real natural and artificial networks are constantly undergoing topological changes, hence their projection might become quickly obsolete. To determine the stability of these projections, first, we have proposed two measures that assess the changes between two projections. Then, we have applied these measures to successive TSVD projections of growing scale-free networks. The results indicate that, even in the case of severe changes in the underlying topology, these projections are very robust. This finding establishes TSVD as a reliable technique for applications that manage dynamic networks.

In Chapter 4 we have used the projection technique described in previous chapter to propose a solution to the main scalability issue of the Internet network, i.e. maintaining an updated global routing table. We have presented a novel decentralized routing strategy that uses the projection of the mesoscopic structure of the network as a reference frame, and where each node only needs to sustain information about its neighborhood. We have evaluated our routing scheme on the Internet ASs network achieving good success rates and reasonable path costs, which reinforces the hypothesis that the modular structure and hubs are important actors in the routing process. Hence, TSVD stands as a promising tool to study the transportation phenomena in natural systems.

The work presented in Chapter 5 has been motivated by the emerging interest on multilayer network representations and by the establishment of an all-you-can-eat consumption model in the Internet, that increases bandwidth demands and congestion. Both topics have led us to study an adaptive routing solution for multiplex communication networks. To this end, we have extended a traffic-aware routing strategy [61, 62] to the new multiplex framework, and we have conducted a longitudinal analysis on synthetic networks composed by two scale-free layers coupled with various assortativity models. The simulation results demonstrate that considering traffic load in the information transmission process can delay the congestion onset in comparison with the merely topological shortest path routing. However, the uniform distribution of traffic produces an abrupt phase transition, simultaneously in both layers, from free-flow to a congested state. On the other hand, the assessment of the effects of coupling preference suggests that assortative models should be the best consideration for optimal network designs. Withal, the characterization of multiplex systems is still in early stages, and this work provides a starting point to study adaptive routing models.

Finally, in Chapter 6 we have presented a theoretical work that illustrates how the globalization process of the international trade system may affect the spreading of financial crises. To this end, we have proposed a simple model of networked integrate-and-fire oscillators and we have characterized their synchronization on the evolving WTW network. The results show that a globalization process, starting in the seventies, dilutes the presence of topological borders and accelerates the synchronization, i.e. the spreading of financial crashes. Unfortunately, this outcome allow us to conjecture that, perhaps, crises in a globalized world are unpredictable. To the best of our knowledge, this is one of the few theoretical works done in studying the critical aspects of the evolution of economic and financial systems, and we hope it will serve to broaden the knowledge about financial crises.

7.2 Perspectives

So far, this dissertation represents a merely drop in the ocean of the incredibly vast and rapid field of complex networks. In a time when quintillion bytes of data are created every day, we hope that the computational ideas presented are useful to model and extract insights about real large complex systems. Here we present some general extensions of this work that could be of interest to explore in the future.

Beyond the issue of network navigability, TSVD projections of the mesosocopic structure of the network may be a valuable tool to study the relationships between structure and dynamics of complex systems. And in particular, to deepen our understanding on multiplex networks. Unfortunately, there does not exist a general decomposition method equivalent to SVD for higher-order tensors, and different methods of decomposition should be studied.

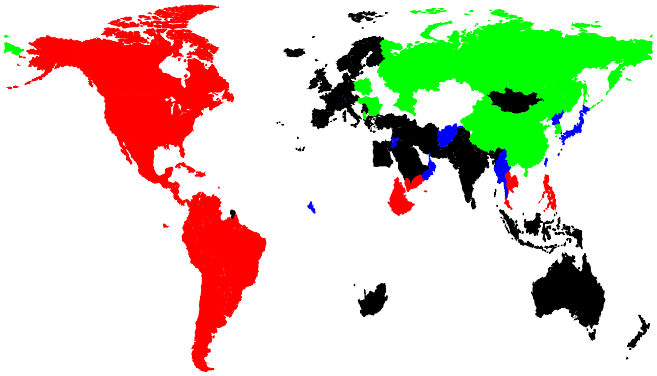
We have shown that the coupling of layers with slightly different structural properties affects the traffic dynamics in multilayer networks. And similarly, one should be cautious when defining the coupling preference. To understand the general consequences of these intralayer and interlayer structures on a particular dynamical process one can do further explorations on combinatorial sets of network models. However, this approach naturally consumes a vast number of computational resources. From our perspective, to better understand dynamical processes future research should be based on real-world multiplex datasets or benchmark models. This issue has to be investigated.

Finally, we have studied the spreading of financial crises using different snapshots of the world trade network, which ignores the time dependence of the trade market. These dependences can be introduced in our model using a time-dependent multiplex network. In this new framework, interlayer links introduce new connectivity patterns that may redefine the borders of our communities and the synchronization behavior, revealing new insights about economical dynamics.

UNIVERSITAT ROVIRA I VIRGILI
NAVIGABILITY AND SYNCHRONIZATION IN COMPLEX NETWORKS: A COMPUTATIONAL APPROACH
Pau Erola Cañellas

Appendix: community structure of the WTW

1950



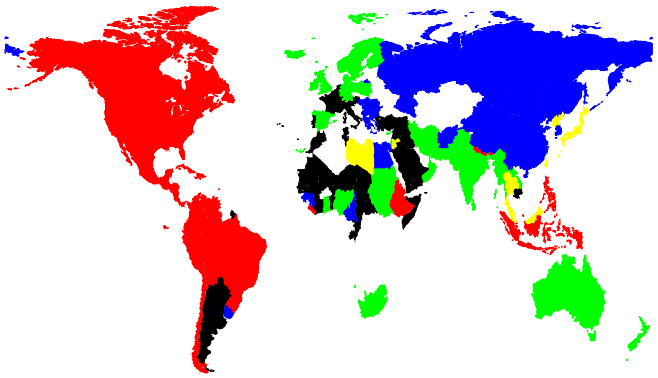
Black United Kingdom, Ireland, Netherlands, Belgium, Luxembourg, France, Switzerland, Spain, Portugal, French Guiana, Austria, Italy, Serbia, Greece, Finland, Sweden, Norway, Denmark, Germany, Iceland, South Africa, Iran, Turkey, Iraq, Egypt, Syria, Lebanon, Israel, Saudi Arabia, Mongolia, India, Bhutan, Pakistan, Sri Lanka, Nepal, Indonesia, Australia, New Zealand

Red United States, Canada, Cuba, Haiti, Dominican Republic, Mexico, Guatemala, Honduras, Santa Lucia, Nicaragua, Costa Rica, Panama, Colombia, Venezuela, Ecuador, Peru, Brazil, Bolivia, Paraguay, Chile, Argentina, Uruguay, Ethiopia, Yemen, Thailand, Philippines

Green Poland, Hungary, Albania, Bulgaria, Romania, Russia, China, South Korea

Blue Liechtenstein, Liberia, Jordan, Oman, Afghanistan, Taiwan, North Korea, Japan, Myanmar

1960



Black Paraguay, Argentina, Netherlands, Belgium, Luxembourg, France, Switzerland, Portugal, French Guiana, Austria, Italy, Mali, Senegal, Benin, Mauritania, Niger, Cote Ivoire, Togo, Cambodia, Central African Republic, Chad, Congo Republic, Somalia, Morocco, Tunisia, Turkey, Iraq, Syria, Lebanon, Saudi Arabia, Yemen

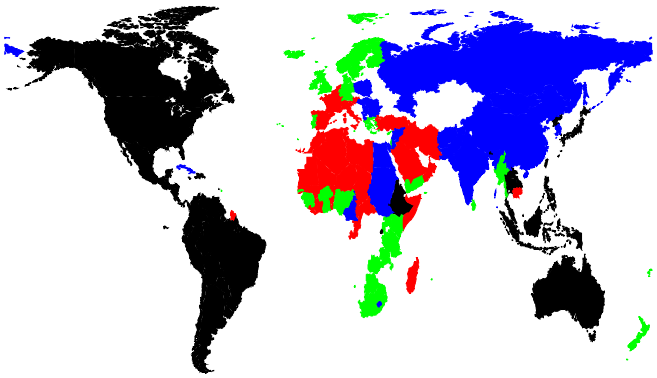
Red United States, Canada, Cuba, Haiti, Dominican Republic, Mexico, Guatemala, Honduras, Santa Lucia, Nicaragua, Costa Rica, Panama, Colombia, Venezuela, Ecuador, Peru, Brazil, Bolivia, Chile, Liberia, Ethiopia, Bhutan, Nepal, Philippines, Indonesia

Green United Kingdom, Ireland, Spain, Poland, Cyprus, Finland, Sweden, Norway, Denmark, Germany, Iceland, Ghana, Nigeria, South Africa, Sudan, Iran, Israel, Oman, India, Pakistan, Myanmar, Sri Lanka, Laos, Australia, New Zealand

Blue Uruguay, Hungary, Albania, Serbia, Greece, Bulgaria, Romania, Russia, Guinea, Egypt, Afghanistan, China, Mongolia, South Korea, Cameroon

Yellow Liechtenstein, Libya, Jordan, Taiwan, North Korea, Japan, Thailand, Malaysia

1970



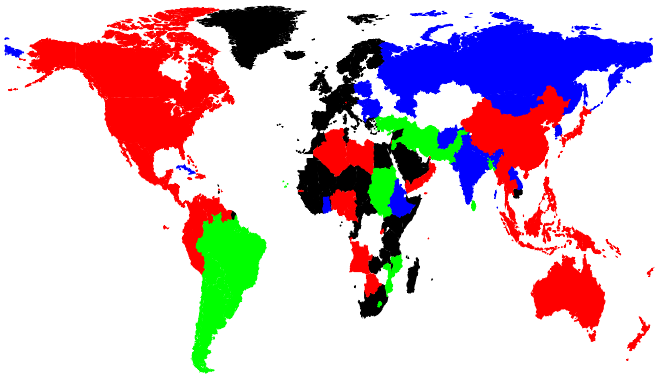
Black United States, Canada, Haiti, Dominican Republic, Jamaica, Trinidad and Tobago, Mexico, Guatemala, Honduras, Santa Lucia, Nicaragua, Costa Rica, Panama, Colombia, Venezuela, Guyana, Ecuador, Peru, Brazil, Bolivia, Paraguay, Chile, Argentina, Uruguay, Liechtenstein, Rwanda, Ethiopia, Taiwan, North Korea, Japan, Bhutan, Thailand, Laos, Malaysia, Singapore, Philippines, Indonesia, Australia, Palestinian Authonomy

Red Netherlands, Belgium, Luxembourg, France, Switzerland, Spain, French Guiana, Austria, Italy, Mali, Senegal, Benin, Mauritania, Niger, Cote Ivoire, Liberia, Togo, Cambodia, Central African Republic, Chad, Congo Republic, Somalia, Morocco, Algeria, Tunisia, Libya, Iran, Turkey, Iraq, Saudi Arabia, Kuwait, Oman, Madagascar

Green Barbados, United Kingdom, Ireland, Portugal, Malta, Greece, Cyprus, Finland, Sweden, Norway, Denmark, Germany, Iceland, Gambia, Guinea, Sierra Leone, Ghana, Nigeria, Uganda, Kenya, Tanzania, Burkina, Zambia, Malawi, South Africa, Botswana, Swaziland, Mauritius, Israel, Yemen, Myanmar, Sri Lanka, New Zealand, Fiji

Blue Cuba, Poland, Hungary, Albania, Serbia, Bulgaria, Romania, Russia, Lesotho, Sudan, Egypt, Syria, Lebanon, Jordan, Afghanistan, China, Mongolia, South Korea, India, Pakistan, Nepal, Cameroon

1980



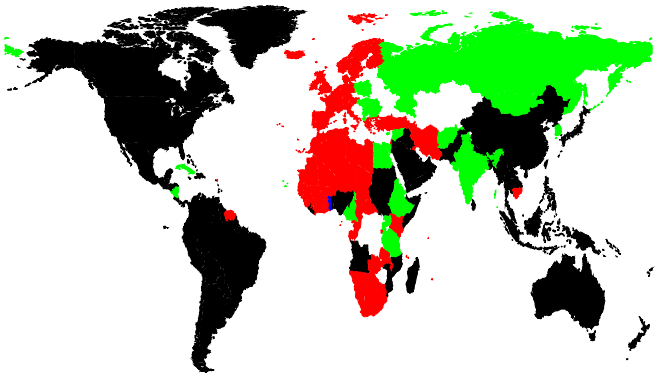
Black Dominica, Greenland, St Vincent and Grenadines, United Kingdom, Ireland, Netherlands, Belgium, Luxembourg, France, Switzerland, Spain, Portugal, French Guiana, Austria, Italy, Malta, Greece, Finland, Sweden, Norway, Denmark, Germany, Iceland, Sao Tome and Principe, Guinea Bissau, Mali, Senegal, Mauritania, Niger, Cote Ivoire, Guinea, Liberia, Sierra Leone, Togo, Cambodia, Central African Republic, Chad, Congo Republic, Uganda, Kenya, Tanzania, Burkina, Somalia, Zambia, Malawi, South Africa, Swaziland, Comoros, Mauritius, Morocco, Tunisia, Egypt, Syria, Lebanon, Israel, Saudi Arabia, Qatar, United Arab Emirates, Madagascar

Red United States, Canada, Haiti, Dominican Republic, Jamaica, Trinidad and Tobago, Barbados, Mexico, Guatemala, Honduras, Santa Lucia, Nicaragua, Costa Rica, Panama, Colombia, Venezuela, Guyana, Suriname, Ecuador, Peru, Liechtenstein, Gambia, Benin, Nigeria, Rwanda, Angola, Botswana, Seychelles, Algeria, Libya, Yemen, Bahamas, Oman, China, Taiwan, North Korea, Japan, Bhutan, Myanmar, Thailand, Cameroon, Malaysia, Singapore, Philippines, Indonesia, Australia, Papua New Guinea, New Zealand, Solomon Islands, Fiji, Palestinian Authonomy

Green Brazil, Bolivia, Paraguay, Chile, Argentina, Uruguay, Cyprus, Cabo Verde, Mozambique, Lesotho, Sudan, Iran, Turkey, Iraq, Jordan, Kuwait, Pakistan, Bangladesh, Sri Lanka

Blue Cuba, Poland, Hungary, Albania, Serbia, Bulgaria, Romania, Russia, Ghana, Ethiopia, Afghanistan, Mongolia, South Korea, India, Nepal, Laos

1990



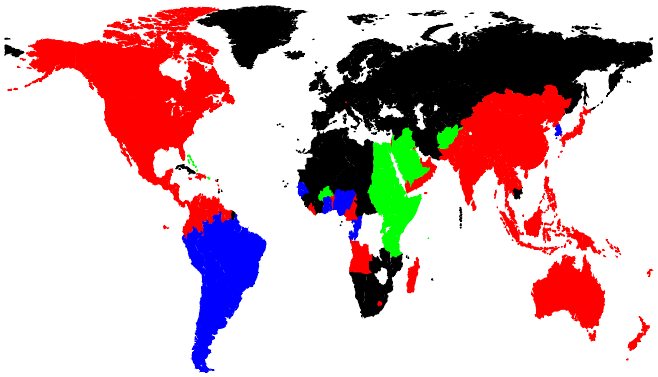
Black United States, Canada, Haiti, Dominican Republic, Jamaica, Trinidad and Tobago, Barbados, Dominica, Greenland, St Vincent and Grenadines, St Kitts and Nevis, Mexico, Belize, Guatemala, Honduras, Santa Lucia, Costa Rica, Panama, Colombia, Venezuela, Guyana, Ecuador, Peru, Brazil, Bolivia, Paraguay, Chile, Argentina, Uruguay, Monaco, Liechtenstein, Benin, Liberia, Nigeria, Somalia, Angola, Mozambique, Lesotho, Sudan, Iraq, Jordan, Saudi Arabia, Yemen, Bahamas, Qatar, United Arab Emirates, Oman, China, Taiwan, North Korea, Japan, Bhutan, Pakistan, Bangladesh, Myanmar, Sri Lanka, Madagascar, Nepal, Thailand, Laos, Malaysia, Singapore, Brunei, Philippines, Indonesia, Australia, Papua New Guinea, New Zealand, Solomon Islands, Fiji, Palestinian Authonomy

Red Antigua and Barbuda, Suriname, United Kingdom, Ireland, Netherlands, Belgium, Luxembourg, France, Switzerland, Spain, Portugal, French Guiana, Austria, Italy, Malta, Greece, Finland, Sweden, Norway, Denmark, Germany, Iceland, Sao Tome and Principe, Guinea Bissau, Gambia, Mali, Senegal, Mauritania, Niger, Cote Ivoire, Guinea, Sierra Leone, Ghana, Cambodia, Central African Republic, Chad, Congo Republic, Kenya, Burkina, Rwanda, Zambia, Malawi, South Africa, Namibia, Botswana, Swaziland, Comoros, Mauritius, Seychelles, Morocco, Algeria, Tunisia, Libya, Iran, Turkey, Israel, Kuwait

Green Cuba, Nicaragua, Poland, Hungary, Albania, Serbia, Cyprus, Bulgaria, Romania, Russia, Cabo Verde, Uganda, Tanzania, Ethiopia, Egypt, Syria, Lebanon, Afghanistan, Mongolia, South Korea, India, Cameroon

Blue Togo, Burkina Faso

2000



Black Cuba, Barbados, Greenland, St Vincent and Grenadines, Antigua and Barbuda, United Kingdom, Ireland, Netherlands, Belgium, Luxembourg, France, Switzerland, Spain, Portugal, French Guiana, Poland, Austria, Hungary, Czech Republic, Slovakia, Italy, Malta, Albania, Macedonia, Croatia, Serbia, Bosnia and Herzegovina, Slovenia, Greece, Cyprus, Bulgaria, Maldives, Romania, Russia, Estonia, Latvia, Lithuania, Ukraine, Belarus, Armenia, Georgia, Azerbaijan, Finland, Sweden, Norway, Denmark, Germany, Iceland, Cabo Verde, Sao Tome and Principe, Mali, Mauritania, Niger, Cote Ivoire, Guinea, Sierra Leone, Cambodia, Central African Republic, Chad, Mozambique, Zambia, Malawi, South Africa, Namibia, Botswana, Swaziland, Comoros, Mauritius, Morocco, Algeria, Tunisia, Libya, Iran, Turkey, Syria, Turkmenistan, Tajikistan, Kyrgyzstan, Uzbekistan, Kazakhstan

Red United States, Canada, Haiti, Dominican Republic, Jamaica, Trinidad and Tobago, Dominica, St Kitts and Nevis, Mexico, Belize, Guatemala, Honduras, Santa Lucia, Nicaragua, Costa Rica, Panama, Colombia, Venezuela, Guyana, Suriname, Ecuador, Monaco, Liechtenstein, Benin, Liberia, Angola, Lesotho, Israel, Yemen, Kuwait, Qatar, United Arab Emirates, Oman, China, Mongolia, Taiwan, North Korea, Japan, India, Bhutan, Pakistan, Bangladesh, Myanmar, Sri Lanka, Madagascar, Nepal, Thailand, Cameroon, Laos, Malaysia, Singapore, Brunei, Philippines, Indonesia, Australia, Papua New Guinea, New Zealand, Solomon Islands, Fiji, Palestinian Authonomy

Green Uganda, Kenya, Tanzania, Burkina, Rwanda, Somalia, Ethiopia, Puerto Rico, Seychelles, Sudan, Iraq, Egypt, Lebanon, Jordan, Saudi Arabia, Bahamas, Afghanistan

Blue Peru, Brazil, Bolivia, Paraguay, Chile, Argentina, Uruguay, Guinea Bissau, Gambia, Senegal, Ghana, Togo, Nigeria, Congo Republic, South Korea

List of publications

Publications within the scope of this thesis

P. Erola, S. Gómez, and A. Arenas. Structural navigability on complex networks. *International Journal of Complex Systems in Science*, 1(1):37–41, 2011.

P. Erola, A. Díaz-Guilera, S. Gómez, and A. Arenas. Trade Synchronization in the World Trade Web. *International Journal of Complex Systems in Science*, 1(2):202–208, 2011.

P. Erola, S. Gómez, and A. Arenas. An Internet Local Routing Approach Based on Network Structural Connectivity. *Proceedings of IEEE GLOBECOM Workshops*, 95–99, 2011.

P. Erola, J. Borge-Holthoefer, S. Gómez, and A. Arenas. Reliability of optimal linear projection of growing scale-free networks. *International Journal of Bifurcation and Chaos*, 22(7), 2012.

P. Erola, A. Díaz-Guilera, S. Gómez, and A. Arenas. Modeling international crisis synchronization in the World Trade Web. *Networks and Heterogeneous Media*, 7(3), 2012.

P. Erola, S. Gómez, and A. Arenas. On the Routability of the Internet. In *Dynamics On and Of Complex Networks*, 2:41–54. Springer, 2013.

P. Erola, S. H. Lee, S. Gómez, A. Arenas, and M. A. Porter. Traffic in multiplex communication networks. *In preparation*.

Other publications

L. Rodríguez-Zamora, X. Iglesias, A. Barrero, D. Chaverri, P. Erola, and F. A. Rodríguez. Physiological Responses in Relation to Performance during Competition in Elite Synchronized Swimmers, *PLoS ONE* 7(11): e49098, 2012.

L. Rodríguez-Zamora, X. Iglesias, A. Barrero, D. Chaverri, A. Irurtia, P. Erola, and F. A. Rodríguez. Perceived Exertion, Time of Immersion and Physiological Correlates in Synchronized Swimming. *International Journal of Sports Medicine*, 35(5): 403-411, 2014.

A. Barrero, D. Chaverri, P. Erola, X. Iglesias, and F. A. Rodríguez. Intensity profile during an ultra-endurance triathlon in relation to testing and performance. *International Journal of Sports Medicine*, 35(14): 1170-1178, 2014.

A. Barrero, P. Erola, R. Bescós. Energy Balance of Triathletes during an Ultra-Endurance Event. *Nutrients*, 7, pp. 209-222, 2015.

Bibliography

[1] I. Abraham, C. Gavoille, D. Malkhi, N. Nisan, and M. Thorup. Compact name-independent routing with minimum stretch. In *Proceedings of the sixteenth annual ACM symposium on Parallelism in algorithms and architectures*, pages 20–24. ACM, 2004. (p. 42)

[2] L. A. Adamic, R. M. Lukose, A. R. Puniyani, and B. A. Huberman. Search in power-law networks. *Physical Review E*, 64(4):046135, 2001. (pp. 43, 45, 51, 55)

[3] R. Albert and A.-L. Barabási. Statistical mechanics of complex networks. *Reviews of Modern Physics*, 74(1):47, 2002. (pp. 9, 44)

[4] D. Alderson, H. Chang, M. Roughan, S. Uhlig, and W. Willinger. The many facets of Internet topology and traffic. *Networks and Heterogeneous Media*, 1(4):569–600, 2006. (p. 38)

[5] G. Alexanderson. Euler and Königsberg's Bridges: a historical view. *Bulletin of the American Mathematical Society*, 43(4):567–573, 2006. (p. 2)

[6] D. G. Andersen, N. Feamster, S. Bauer, and H. Balakrishnan. Topology inference from BGP routing dynamics. In *Proceedings of the 2nd ACM SIGCOMM Workshop on Internet measurment*, pages 243–248. ACM, 2002. (p. 38)

[7] P. W. Anderson. More is different. *Science*, 177(4047):393–396, Aug. 1972. (p. 1)

[8] A. Arenas, J. Borge-Holthoefer, S. Gómez, and G. Zamora-López. Optimal map of the modular structure of complex networks. *New Journal of Physics*, 12(5):053009, 2010. (pp. 21, 24, 28, 31, 33)

[9] A. Arenas, A. Cabrales, A. Díaz-Guilera, R. Guimerà, and F. Vega-Redondo. Search and congestion in complex networks. In *Statistical Mechanics of Complex Networks*, number 625 in Lecture Notes in Physics, pages 175–194. Springer Berlin Heidelberg, 2003. (p. 54)

[10] A. Arenas, L. Danon, A. Díaz-Guilera, and R. Guimerà. Local search with congestion in complex communication networks. In *Computational Science-ICCS 2004*, pages 1078–1085. Springer, 2004. (pp. 43, 54)

[11] A. Arenas, A. Díaz-Guilera, and R. Guimerà. Communication in networks with hierarchical branching. *Physical Review Letters*, 86(14):3196, 2001. (p. 67)

[12] A. Arenas, A. Díaz-Guilera, J. Kurths, Y. Moreno, and C. Zhou. Synchronization in complex networks. *Physics Reports*, 469(3):93–153, 2008. (pp. 5, 9, 77)

[13] A. Arenas, A. Díaz-Guilera, and C. J. Pérez-Vicente. Synchronization reveals topological scales in complex networks. *Physical Review Letters*, 96(11):114102, 2006. (p. 80)

[14] A. Arenas, J. Duch, A. Fernández, and S. Gómez. Size reduction of complex networks preserving modularity. *New Journal of Physics*, 9(6):176, 2007. (p. 74)

[15] A. Arenas, A. Fernández, and S. Gómez. Analysis of the structure of complex networks at different resolution levels. *New Journal of Physics*, 10(5):053039, 2008. (p. 18)

[16] B. Awerbuch, A. Bar-Noy, N. Linial, and D. Peleg. Compact distributed data structures for adaptive routing. In *Proceedings of the twenty-first annual ACM symposium on Theory of computing*, pages 479–489. ACM, 1989. (p. 42)

[17] J. P. Bagrow and E. M. Bollt. Local method for detecting communities. *Physical Review E*, 72(4):046108, 2005. (p. 17)

[18] J. R. Banavar, A. Maritan, and A. Rinaldo. Size and form in efficient transportation networks. *Nature*, 399(6732):130–132, May 1999. (pp. 3, 4)

[19] A. L. Barabási. *Network Science*. http://barabasi.com/networksciencebook/, 2015. (pp. 9, 16)

[20] A.-L. Barabási and R. Albert. Emergence of scaling in random networks. *Science*, 286(5439):509–512, 1999. (pp. 1, 2, 27)

[21] A. Barrat, M. Barthelemy, and A. Vespignani. *Dynamical processes on complex networks*. Cambridge University Press, 2008. (p. 9)

[22] M. Barthélemy and A. Flammini. Optimal traffic networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2006(07):L07002, 2006. (p. 55)

[23] F. Battiston, V. Nicosia, and V. Latora. Structural measures for multiplex networks. *Physical Review E*, 89(3):032804, 2014. (p. 19)

[24] M. Bazzi, M. A. Porter, S. Williams, M. McDonald, D. J. Fenn, and S. D. Howison. Community detection in temporal multilayer networks, and its application to correlation networks. *arXiv:1501.00040 [nlin, physics:physics, q-fin]*, Dec. 2014. (p. 19)

[25] E. T. Bell. Exponential numbers. *American Mathematical Monthly*, pages 411–419, 1934. (p. 18)

[26] L. Bennett, A. Kittas, G. Muirhead, L. G. Papageorgiou, and S. Tsoka. Detection of composite communities in multiplex biological networks. *Scientific reports*, 5, 2015. (p. 19)

[27] M. W. Berry, S. T. Dumais, and G. W. O'Brien. Using linear algebra for intelligent information retrieval. *SIAM review*, 37(4):573–595, 1995. (p. 24)

[28] S. Boccaletti, G. Bianconi, R. Criado, C. I. del Genio, J. Gómez-Gardeñes, M. Romance, I. Sendiña-Nadal, Z. Wang, and M. Zanin. The structure and dynamics of multilayer networks. *Physics Reports*, 544(1):1–122, Nov. 2014. (pp. 18, 55)

[29] S. Boccaletti, V. Latora, and Y. Moreno. *Handbook on biological networks*, volume 10. World Scientific, Dec. 2009. (p. 3)

[30] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D. U. Hwang. Complex networks: structure and dynamics. *Physics Reports*, 424(4–5):175–308, Feb. 2006. (p. 9)

[31] S. P. Boehm and P. Baran. On distributed communications series: II. Digital simulation of hot-potato routing in a broadband distributed communications network. *RAND Memoranda*, (3103), 1964. (p. 59)

[32] M. Boguñá, F. Papadopoulos, and D. Krioukov. Sustaining the Internet with hyperbolic mapping. *Nature Communications*, 1:62, 2010. (pp. 21, 22, 35, 44, 50)

[33] J. Borge-Holthoefer. *Semantic networks and cognitive dynamics*. PhD thesis, Universitat Rovira i Virgili, Jan. 2011. (p. 2)

[34] A. Brady and L. J. Cowen. Compact routing on power law graphs with additive stretch. In *ALENEX*, volume 6, pages 119–128. SIAM, 2006. (p. 42)

[35] U. Brandes, D. Delling, M. Gaertler, R. Gorke, M. Hoefer, Z. Nikoloski, and D. Wagner. On modularity clustering. *Knowledge and Data Engineering, IEEE Transactions on*, 20(2):172–188, 2008. (p. 18)

[36] L. A. Braunstein, S. V. Buldyrev, R. Cohen, S. Havlin, and H. E. Stanley. Optimal paths in disordered complex networks. *Physical review letters*, 91(16):168701, 2003. (p. 43)

[37] T. Bu, L. Gao, and D. Towsley. On characterizing BGP routing table growth. *Computer Networks*, 45(1):45–54, 2004. (p. 40)

[38] CAIDA. AS Relationships. http://www.caida.org/data/as-relationships. (pp. 38, 48)

[39] CAIDA. The CAIDA UCSD macroscopic skitter topology dataset. http://www.caida.org/tools/measurements/skitter. (p. 38)

[40] S. R. Campbell, D. L. Wang, and C. Jayaprakash. Synchrony and desynchrony in integrate-and-fire oscillators. *Neural computation*, 11(7):1595–1619, 1999. (pp. 5, 79, 85)

[41] J. A. Capitán, J. Borge-Holthoefer, S. Gómez, J. Martinez-Romo, L. Araujo, J. A. Cuesta, and A. Arenas. Local-based semantic navigation on a networked representation of information. *PLoS ONE*, 7(8):e43694, Aug. 2012. (p. 21)

[42] A. Capocci, V. D. Servedio, F. Colaiori, L. S. Buriol, D. Donato, S. Leonardi, and G. Caldarelli. Preferential attachment in the growth of social networks: The internet encyclopedia Wikipedia. *Physical Review E*, 74(3):036116, 2006. (p. 27)

[43] I. Castineyra and M. Steenstrup. The Nimrod routing architecture. In *IETF, RFC*. Citeseer, 1992. (p. 43)

[44] D. Chowdhury, L. Santen, and A. Schadschneider. Statistical physics of vehicular traffic and some related systems. *Physics Reports*, 329(4):199–329, 2000. (p. 4)

[45] A. Clauset, M. E. Newman, and C. Moore. Finding community structure in very large networks. *Physical review E*, 70(6):066111, 2004. (p. 18)

[46] F. C. Commission. In the matter of protecting and promoting the open internet. report and order on remand, declaratory ruling and order. *GN Docket No. 14-28*, Mar. 2015. (p. 54)

[47] E. Cozzo, M. Kivelä, M. De Domenico, A. Solé, A. Arenas, S. Gómez, M. A. Porter, and Y. Moreno. Structure of triadic relations in multiplex networks. *New Journal of Physics*, 17(7):073029, July 2015. arXiv: 1307.6780. (p. 19)

[48] B. Danila, Y. Yu, J. A. Marsh, and K. E. Bassler. Optimal transport on complex networks. *Physical Review E*, 74(4):046106, 2006. (pp. 43, 55)

[49] L. Danon, A. Díaz-Guilera, J. Duch, and A. Arenas. Comparing community structure identification. *Journal of Statistical Mechanics: Theory and Experiment*, 2005(09):P09008, 2005. (pp. 17, 18)

[50] M. De Domenico, A. Solé-Ribalta, E. Cozzo, M. Kivelä, Y. Moreno, M. A. Porter, S. Gómez, and A. Arenas. Mathematical formulation of multilayer networks. *Physical Review X*, 3(4):041022, 2013. (p. 19)

[51] M. De Domenico, A. Solé-Ribalta, S. Gómez, and A. Arenas. Navigability of interconnected networks under random failures. *Proceedings of the National Academy of Sciences*, 111(23):8351–8356, June 2014. (p. 58)

[52] M. De Domenico, A. Solé-Ribalta, E. Omodei, S. Gómez, and A. Arenas. Centrality in interconnected multilayer networks. *Supplementary material of Nature Communications*, 6:6868, Apr. 2015. arXiv: 1311.2906. (p. 19)

[53] M. De Domenico, A. Solé-Ribalta, E. Omodei, S. Gómez, and A. Arenas. Ranking in interconnected multilayer networks reveals versatile nodes. *Nature communications*, 6:6868, 2015. (p. 19)

[54] D. De Martino, L. Dall'Asta, G. Bianconi, and M. Marsili. Congestion phenomena on complex networks. *Physical Review E*, 79(1):015101, Jan. 2009. (pp. 4, 66, 71)

[55] A. V. Deardorff. *Glossary of international economics*. World Scientific, 2006. (p. 76)

[56] S. C. Deerwester, S. T. Dumais, T. K. Landauer, G. W. Furnas, and R. A. Harshman. Indexing by latent semantic analysis. *JASIS*, 41(6):391–407, 1990. (p. 24)

[57] L. Donetti and M. A. Muñoz. Detecting network communities: a new systematic and efficient algorithm. *Journal of Statistical Mechanics: Theory and Experiment*, 2004(10):P10012, 2004. (p. 17)

[58] S. N. Dorogovtsev, A. V. Goltsev, and J. F. Mendes. Critical phenomena in complex networks. *Reviews of Modern Physics*, 80(4):1275, 2008. (p. 9)

[59] J. Duch. *Structure and traffic on complex networks*. PhD thesis, Universitat de Barcelona, Apr. 2008. (pp. 4, 54)

[60] J. Duch and A. Arenas. Community detection in complex networks using extremal optimization. *Physical Review E*, 72(2):027104, 2005. (pp. 18, 26, 48)

[61] P. Echenique, J. Gómez-Gardeñes, and Y. Moreno. Improved routing strategies for Internet traffic delivery. *Physical Review E*, 70(5):056105, Nov. 2004. (pp. 4, 55, 56, 58, 67, 71, 88)

[62] P. Echenique, J. Gómez-Gardeñes, and Y. Moreno. Dynamics of jamming transitions in complex networks. *EPL (Europhysics Letters)*, 71(2):325, 2005. (pp. 4, 43, 55, 56, 58, 71, 88)

[63] P. Erdös and A. Rényi. On random graphs I. *Publ. Math. Debrecen*, 6:290–297, 1959. (p. 15)

[64] P. Erdös and A. Rényi. On the evolution of random graphs. *Publ. Math. Inst. Hungar. Acad. Sci*, 5:17–61, 1960. (p. 15)

[65] P. Erola, J. Borge-Holthoefer, S. Gómez, and A. Arenas. Reliability of optimal linear projection of growing scale-free networks. *International Journal of Bifurcation and Chaos*, 22(07), 2012. (p. 27)

[66] P. Erola, A. Díaz-Guilera, S. Gómez, and A. Arenas. Trade synchronization in the world trade web. *Int J Comp Sys Science*, 1(2):202–208, 2011. (p. 78)

[67] P. Erola, A. Díaz-Guilera, S. Gómez, and A. Arenas. Modeling international crisis synchronization in the World Trade Web. *Networks and Heterogeneous Media*, 7(3):385–397, Oct. 2012. (p. 78)

[68] P. Erola, S. Gómez, and A. Arenas. An Internet local routing approach based on network structural connectivity. In *GLOBECOM Workshops (GC Wkshps), 2011 IEEE*, pages 95–99. IEEE, 2011. (p. 47)

[69] P. Erola, S. Gómez, and A. Arenas. Structural navigability on complex networks. *Int J Comp Sys Science*, 1(1):37–41, 2011. (p. 46)

[70] P. Erola, S. Gómez, and A. Arenas. On the routability of the Internet. In *Dynamics On and Of Complex Networks, Volume 2*, pages 41–54. Springer, 2013. (p. 46)

[71] L. Euler. Solutio problematis ad geometriam situs pertinentis. *Commentarii academiae scientiarum Petropolitanae*, 8:128–140, 1741. (p. 3)

[72] G. Fagiolo, J. Reyes, and S. 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. (p. 74)

[73] G. Fagiolo, J. Reyes, and S. Schiavo. World-trade web: topological properties, dynamics, and evolution. *Physical Review E*, 79(3):036115, 2009. (p. 74)

[74] G. Fagiolo, J. Reyes, and S. Schiavo. The evolution of the world trade web: a weighted-network analysis. *Journal of Evolutionary Economics*, 20(4):479–514, 2010. (p. 74)

[75] M. Faloutsos, P. Faloutsos, and C. Faloutsos. On power-law relationships of the Internet topology. *ACM SIGCOMM Computer Communication Review*, 29(4):251–262, Oct. 1999. (p. 38)

[76] D. Farinacci, D. Lewis, D. Meyer, and V. Fuller. The locator/ID separation protocol (LISP). RFC 6830. Technical report, 2013. (p. 43)

[77] N. Feamster. Where is Internet congestion occurring? *Freedom to Tinker*, Apr. 2015. (p. 54)

[78] N. Feamster. Why your Netflix traffic is slow, and why the Open Internet Order won't (necessarily) make it faster. *Freedom to Tinker*, Mar. 2015. (pp. 53, 54)

[79] S. Floyd and V. Paxson. Difficulties in simulating the Internet. *IEEE/ACM Transactions on Networking (TON)*, 9(4):392–403, 2001. (pp. 38, 59)

[80] S. Fortunato. Community detection in graphs. *Physics Reports*, 486(3–5):75–174, Feb. 2010. (p. 17)

[81] S. Fortunato and M. Barthelemy. Resolution limit in community detection. *Proceedings of the National Academy of Sciences*, 104(1):36–41, 2007. (p. 18)

[82] S. Fortunato, V. Latora, and M. Marchiori. Method to find community structures based on information centrality. *Physical Review E*, 70(5):056104, 2004. (p. 17)

[83] M. Franceschetti and R. Meester. *Random networks for communication: from statistical physics to information systems*, volume 24. Cambridge University Press, 2008. (p. 15)

[84] L. C. Freeman. A set of measures of centrality based on betweenness. *Sociometry*, pages 35–41, 1977. (p. 14)

[85] K. Fukuda, H. Takayasu, and M. Takayasu. Spatial and temporal behavior of congestion in Internet traffic. *Fractals*, 7(01):23–31, 1999. (p. 54)

[86] V. Fuller, T. Li, J. Yu, and K. Varadhan. Classless inter-domain routing (CIDR): an address assignment and aggregation strategy. Technical report, 1993. (p. 42)

[87] J. Gao, S. V. Buldyrev, H. E. Stanley, and S. Havlin. Networks formed from interdependent networks. *Nature Physics*, 8(1):40–48, 2012. (p. 55)

[88] D. Garlaschelli and M. I. Loffredo. Structure and evolution of the world trade network. *Physica A: Statistical Mechanics and its Applications*, 355(1):138–144, 2005. (p. 74)

[89] A. Gavaldà, J. Duch, and J. Gómez-Gardeñes. Reciprocal interactions out of congestion-free adaptive networks. *Physical Review E*, 85(2):026112, 2012. (pp. 4, 43, 55)

[90] G. Ghoshal. *Structural and dynamical properties of complex networks*. PhD thesis, University of Michigan, 2009. (p. 2)

[91] M. Girvan and M. E. J. Newman. Community structure in social and biological networks. *Proceedings of the National Academy of Sciences*, 99(12):7821–7826, June 2002. (pp. 3, 17)

[92] K. S. Gleditsch. Expanded trade and GDP data. *Journal of Conflict Resolution*, 46(5):712–724, 2002. (p. 74)

[93] R. Glick and A. K. Rose. Contagion and trade: Why are currency crises regional? *Journal of international Money and Finance*, 18(4):603–617, 1999. (pp. 73, 77)

[94] A. Godoy-Lorite, R. Guimerà, and M. Sales-Pardo. Long-term evolution of techno-social networks: statistical regularities, predictability and stability of social behaviors. *arXiv:1506.01516 [cond-mat, physics:physics]*, June 2015. arXiv: 1506.01516. (p. 18)

[95] K.-I. Goh, B. Kahng, and D. Kim. Universal behavior of load distribution in scale-free networks. *Physical Review Letters*, 87(27):278701, 2001. (pp. 4, 55)

[96] K.-I. Goh, J. D. Noh, B. Kahng, and D. Kim. Load distribution in weighted complex networks. *Physical Review E*, 72(1):017102, 2005. (p. 55)

[97] G. H. Golub and C. F. Van Loan. *Matrix computations*, volume 3. JHU Press, 2012. (pp. 23, 24)

[98] S. Gómez, A. Díaz-Guilera, J. Gómez-Gardeñes, C. J. Pérez-Vicente, Y. Moreno, and A. Arenas. Diffusion dynamics on multiplex networks. *Physical Review Letters*, 110(2):028701, 2013. (p. 55)

[99] J. Gómez-Gardeñes, Y. Moreno, and A. Arenas. Paths to synchronization on complex networks. *Physical Review Letters*, 98(3):034101, Jan. 2007. (p. 5)

[100] R. Govindan and H. Tangmunarunkit. Heuristics for Internet map discovery. In *INFOCOM 2000. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, volume 3, pages 1371–1380. IEEE, 2000. (p. 38)

[101] C. Granell, S. Gómez, and A. Arenas. Hierarchical multiresolution method to overcome the resolution limit in complex networks. *International Journal of Bifurcation and Chaos*, 22(07):1250171, 2012. (p. 18)

[102] R. Guimerà and L. A. N. Amaral. Cartography of complex networks: modules and universal roles. *Journal of Statistical Mechanics: Theory and Experiment*, 2005(02):P02001, 2005. (p. 18)

[103] R. Guimerà, A. Díaz-Guilera, F. Vega-Redondo, A. Cabrales, and A. Arenas. Optimal network topologies for local search with congestion. *Physical Review Letters*, 89(24):248701, Nov. 2002. (pp. 4, 55, 71)

[104] A. Halu, R. J. Mondragón, P. Panzarasa, and G. Bianconi. Multiplex PageRank. *PLoS ONE*, 8(10):e78293, Oct. 2013. (p. 19)

[105] D. Harding and A. Pagan. Synchronization of cycles. *Journal of Econometrics*, 132(1):59–79, 2006. (p. 6)

[106] H. A. Harrington, M. Beguerisse-Díaz, M. ROMBACH, L. M. Keating, and M. A. Porter. Commentary: teach network science to teenagers. *Network Science*, 1(02):226–247, 2013. (p. 16)

[107] S. Havlin, N. A. M. Araujo, S. V. Buldyrev, C. S. Dias, R. Parshani, G. Paul, and H. E. Stanley. Catastrophic cascade of failures in interdependent networks. *arXiv:1012.0206 [cond-mat, physics:physics]*, Dec. 2010. (p. 55)

[108] J. He and M. W. Deem. Structure and response in the world trade network. *Physical Review Letters*, 105(19):198701, 2010. (pp. 77, 85)

[109] D. Helbing. Traffic and related self-driven many-particle systems. *Reviews of Modern Physics*, 73(4):1067–1141, Dec. 2001. (p. 4)

[110] R. Hinden. New scheme for Internet routing and addressing (ENCAPS) for IPNG. RFC 1955. Technical report, 1996. (p. 43)

[111] P. W. Holland and S. Leinhardt. A method for detecting structure in sociometric data. *American Journal of Sociology*, 76(3):492–513, Nov. 1970. (p. 21)

[112] P. Holme. Congestion and centrality in traffic flow on complex networks. *Advances in Complex Systems*, 6(02):163–176, 2003. (p. 54)

[113] Y. Ikeda, H. Aoyama, Y. Fujiwara, H. Iyetomi, K. Ogimoto, W. Souma, and H. Yoshikawa. Coupled oscillator model of the business cycle with fluctuating goods markets. *Progress of Theoretical Physics Supplement*, 194:111–121, May 2012. (p. 85)

[114] J. Imbs. Trade, finance, specialization, and synchronization. *Review of Economics and Statistics*, 86(3):723–734, 2004. (p. 6)

[115] R. Inklaar, R. Jong-A-Pin, and J. De Haan. Trade and business cycle synchronization in OECD countries—A re-examination. *European Economic Review*, 52(4):646–666, 2008. (p. 6)

[116] M. Kivelä, A. Arenas, M. Barthelemy, J. P. Gleeson, Y. Moreno, and M. A. Porter. Multilayer networks. *Journal of Complex Networks*, 2(3):203–271, 2014. (pp. 5, 9, 19)

[117] J. M. Kleinberg. Navigation in a small world. *Nature*, 406(6798):845–845, 2000. (pp. 4, 42, 45)

[118] L. Kleinrock and F. Kamoun. Hierarchical routing for large networks performance evaluation and optimization. *Computer Networks (1976)*, 1(3):155–174, 1977. (p. 41)

[119] M. A. Kose, C. Otrok, and E. Prasad. Global business cycles: convergence or decoupling? *International Economic Review*, 53(2):511–538, 2012. (p. 77)

[120] D. Krioukov, K. C. Claffy, K. Fall, and A. Brady. On compact routing for the Internet. *ACM SIGCOMM Computer Communication Review*, 37(3):41–52, 2007. (pp. 41, 42)

[121] D. Krioukov, K. Fall, and X. Yang. Compact routing on Internet-like graphs. In *INFOCOM 2004. Twenty-Third AnnualJoint Conference of the IEEE Computer and Communications Societies*, volume 1. IEEE, 2004. (p. 41)

[122] D. Krioukov, F. Papadopoulos, M. Kitsak, A. Vahdat, and M. Boguñá. Hyperbolic geometry of complex networks. *Physical Review E*, 82(3):036106, Sept. 2010. (pp. 22, 23, 44)

[123] D. Krioukov, F. Papadopoulos, A. Vahdat, and M. Boguñá. Curvature and temperature of complex networks. *Physical Review E*, 80(3):035101, 2009. (p. 22)

[124] C. Labovitz, A. Ahuja, and F. Jahanian. Experimental study of internet stability and backbone failures. In *Fault-Tolerant Computing, 1999. Digest of Papers. Twenty-Ninth Annual International Symposium on*, pages 278–285. IEEE, 1999. (p. 5)

[125] C. Labovitz, S. Iekel-Johnson, D. McPherson, J. Oberheide, and F. Jahanian. Internet Inter-domain Traffic. In *Proceedings of the ACM SIGCOMM 2010 Conference*, SIGCOMM '10, pages 75–86, New York, NY, USA, 2010. ACM. (p. 54)

[126] C. Labovitz, S. Iekel-Johnson, D. McPherson, J. Oberheide, and F. Jahanian. Internet Inter-domain traffic. *ACM SIGCOMM Computer Communication Review*, 41(4):75–86, 2011. (p. 51)

[127] A. Lancichinetti and S. Fortunato. Limits of modularity maximization in community detection. *Physical Review E*, 84(6):066122, 2011. (p. 18)

[128] T. K. Landauer and S. T. Dumais. A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, 104(2):211, 1997. (p. 24)

[129] T. K. Landauer, P. W. Foltz, and D. Laham. An introduction to latent semantic analysis. *Discourse Processes*, 25(2-3):259–284, 1998. (p. 24)

[130] S. Lattanzi, A. Panconesi, and D. Sivakumar. Milgram-routing in social networks. In *Proceedings of the 20th international conference on World wide web*, pages 725–734. ACM, 2011. (p. 51)

[131] K.-M. Lee, J. Y. Kim, W.-k. Cho, K. I. Goh, and I. M. Kim. Correlated multiplexity and connectivity of multiplex random networks. *New Journal of Physics*, 14(3):033027, 2012. (p. 55)

[132] S. H. Lee and P. Holme. Exploring maps with greedy navigators. *Physical Review Letters*, 108(12), Mar. 2012. (p. 44)

[133] X. Li, Y. Ying Jin, and G. Chen. Complexity and synchronization of the world trade web. *Physica A: Statistical Mechanics and its Applications*, 328(1):287–296, 2003. (p. 77)

[134] P. A. Longley, M. F. Goodchild, D. J. Maguire, and D. W. Rhind. *Geographic information science and systems*. John Wiley & Sons, 2015. (p. 21)

[135] B. Lyon. The Opte Project. http://www.opte.org. (p. 39)

[136] D. Meyer, L. Zhang, and K. Fall. Report from the IAB Workshop on Routing and Addressing. Technical report, 2007. (pp. 4, 40)

[137] C. Mihyeon Jeon and A. Amekudzi. Addressing sustainability in transportation systems: definitions, indicators, and metrics. *Journal of Infrastructure Systems*, 11(1):31–50, 2005. (p. 4)

[138] S. Milgram. The small world problem. *Psychology Today*, 2(1):60–67, 1967. (p. 2)

[139] H. P. Minsky. The financial instability hypothesis. *The Jerome Levy Economics Institute Working Paper*, (74), 1992. (p. 76)

[140] H. P. Minsky and H. Kaufman. *Stabilizing an unstable economy*, volume 1. McGraw-Hill New York, 2008. (p. 76)

[141] R. E. Mirollo and S. H. Strogatz. Synchronization of pulse-coupled biological oscillators. *SIAM Journal on Applied Mathematics*, 50(6):1645–1662, 1990. (pp. 78, 79)

[142] Y. Moreno, R. Pastor-Satorras, A. Vázquez, and A. Vespignani. Critical load and congestion instabilities in scale-free networks. *EPL (Europhysics Letters)*, 62(2):292, 2003. (pp. 4, 55)

[143] R. G. Morris and M. Barthelemy. Transport on coupled spatial networks. *Physical review letters*, 109(12):128703, 2012. (p. 55)

[144] A. E. Motter, C. Zhou, and J. Kurths. Network synchronization, diffusion, and the paradox of heterogeneity. *Physical Review E*, 71(1):016116, Jan. 2005. (p. 5)

[145] P. J. Mucha, T. Richardson, K. Macon, M. A. Porter, and J.-P. Onnela. Community structure in time-dependent, multiscale, and multiplex networks. *Science*, 328(5980):876–878, May 2010. (pp. 18, 19)

[146] M. Newman. *Networks: an introduction*. Oxford University Press, 2010. (pp. 3, 21)

[147] M. Newman, A.-L. Barabasi, and D. J. Watts. *The structure and dynamics of networks*. Princeton University Press, 2006. (p. 9)

[148] M. E. Newman. Scientific collaboration networks I. Network construction and fundamental results. *Physical review E*, 64(1):016131, 2001. (p. 14)

[149] M. E. Newman. Assortative mixing in networks. *Physical review letters*, 89(20):208701, 2002. (p. 12)

[150] M. E. Newman. The structure and function of complex networks. *SIAM review*, 45(2):167–256, 2003. arXiv:cond-mat/0303516. (pp. 2, 9, 11, 12)

[151] M. E. Newman. Analysis of weighted networks. *Physical Review E*, 70(5):056131, 2004. (p. 74)

[152] M. E. Newman. Fast algorithm for detecting community structure in networks. *Physical Review E*, 69(6):066133, 2004. (p. 18)

[153] M. E. Newman. Finding community structure in networks using the eigenvectors of matrices. *Physical Review E*, 74(3):036104, 2006. (p. 17)

[154] M. E. Newman. Modularity and community structure in networks. *Proceedings of the National Academy of Sciences*, 103(23):8577–8582, 2006. (p. 18)

[155] M. E. Newman and M. Girvan. Finding and evaluating community structure in networks. *Physical Review E*, 69(2):026113, 2004. (p. 17)

[156] V. Nicosia and V. Latora. Measuring and modeling correlations in multiplex networks. *Physical Review E*, 92(3):032805, Sept. 2015. (pp. 19, 71)

[157] U. of Oregon's Route Views Project. http://www.routeviews.org. (p. 38)

[158] J.-P. Onnela, J. Saramäki, J. Hyvönen, G. Szabó, D. Lazer, K. Kaski, J. Kertész, and A.-L. Barabási. Structure and tie strengths in mobile communication networks. *Proceedings of the National Academy of Sciences*, 104(18):7332–7336, 2007. (p. 5)

[159] G. Palla, I. J. Farkas, P. Pollner, I. Derenyi, and T. Vicsek. Directed network modules. *New Journal of Physics*, 9(6):186, 2007. (p. 17)

[160] F. Papadopoulos, D. Krioukov, M. Bogua, and A. Vahdat. Greedy forwarding in dynamic scale-free networks embedded in hyperbolic metric spaces. In *INFOCOM, 2010 Proceedings IEEE*, pages 1–9. IEEE, 2010. arXiv:0805.1266. (p. 22)

[161] R. Parshani, C. Rozenblat, D. Ietri, C. Ducruet, and S. Havlin. Inter-similarity between coupled networks. *EPL (Europhysics Letters)*, 92(6):68002, 2010. (p. 55)

[162] R. Pastor-Satorras and A. Vespignani. *Evolution and structure of the Internet: a statistical physics approach*. Cambridge University Press, 2007. (pp. 3, 4, 14, 27, 38, 40, 53, 54)

[163] D. Peleg and E. Upfal. A trade-off between space and efficiency for routing tables. *Journal of the ACM (JACM)*, 36(3):510–530, 1989. (p. 41)

[164] A. Pikovsky, M. G. Rosenblum, and J. Kurths. *Synchronization—a universal concept in nonlinear sciences*. Cambridge University Press, Cambridge, 2001. (p. 5)

[165] P. Pons and M. Latapy. Computing communities in large networks using random walks. In *Computer and Information Sciences-ISCIS 2005*, pages 284–293. Springer, 2005. (p. 17)

[166] M. A. Porter. *A Terse Introduction to Networks*. Springer-Verlag, forthcoming. (pp. 1, 9, 10)

[167] M. A. Porter, J.-P. Onnela, and P. J. Mucha. Communities in Networks. *Notices of the American Mathematical Society*, 56(9):1082–1097, Feb. 2009. arXiv: 0902.3788. (p. 17)

[168] L. Prignano, O. Sagarra, and A. Díaz-Guilera. Tuning synchronization of integrate-and-fire oscillators through mobility. *Physical Review Letters*, 110(11):114101, 2013. (p. 5)

[169] L. Prignano, O. Sagarra, P. M. Gleiser, and A. Díaz-Guilera. Synchronization of moving integrate and fire oscillators. *International Journal of Bifurcation and Chaos*, 22(07), 2012. (p. 5)

[170] J. M. Pujol, J. Béjar, and J. Delgado. Clustering algorithm for determining community structure in large networks. *Physical Review E*, 74(1):016107, 2006. (p. 18)

[171] Radatools. Community detection in complex networks and other tools. http://deim.urv.cat/sergio.gomez/radatools.php. (pp. 26, 48, 75)

[172] F. Radicchi and A. Arenas. Abrupt transition in the structural formation of interconnected networks. *Nature Physics*, 9(11):717–720, 2013. (p. 55)

[173] F. Radicchi, C. Castellano, F. Cecconi, V. Loreto, and D. Parisi. Defining and identifying communities in networks. *Proceedings of the National Academy of Sciences of the United States of America*, 101(9):2658–2663, 2004. (p. 17)

[174] S. Robinson. Multimedia transmissions are driving Internet toward gridlock. *The New York Times*, Aug. 1999. (p. 54)

[175] E. C. Rosen. Exterior gateway protocol (EGP). 1982. (p. 40)

[176] A. Rothkegel and K. Lehnertz. Recurrent events of synchrony in complex networks of pulse-coupled oscillators. *EPL (Europhysics Letters)*, 95(3):38001, 2011. (pp. 5, 79)

[177] S. J. Russell and P. Norvig. *Artificial Intelligence: a modern approach*. Prentice Hall, Englewood Cliffs, N.J, 1st edition edition, Jan. 1995. (p. 44)

[178] G. Sabidussi. The centrality index of a graph. *Psychometrika*, 31(4):581–603, Dec. 1966. (p. 14)

[179] N. Santoro and R. Khatib. Labelling and implicit routing in networks. *The Computer Journal*, 28(1):5–8, 1985. (p. 42)

[180] J. Scott. *Social network analysis*. Sage, 2012. (p. 17)

[181] M. Á. Serrano and M. Boguñá. Topology of the world trade web. *Physical Review E*, 68(1):015101, 2003. (p. 74)

[182] M. Á. Serrano, M. Boguñá, and A. Díaz-Guilera. Modeling the Internet. *The European Physical Journal B-Condensed Matter and Complex Systems*, 50(1-2):249–254, 2006. (p. 38)

[183] M. Á. Serrano, M. Boguñá, and F. Sagués. Uncovering the hidden geometry behind metabolic networks. *Molecular BioSystems*, 8(3):843, 2012. (pp. 21, 22)

[184] M. Á. Serrano, D. Krioukov, and M. Boguñá. Self-similarity of complex networks and hidden metric spaces. *Physical Review Letters*, 100(7):078701, Feb. 2008. (p. 22)

[185] S. Shai and S. Dobson. Effect of resource constraints on intersimilar coupled networks. *Physical Review E*, 86(6):066120, 2012. (p. 55)

[186] Y. Shavitt and E. Shir. DIMES: Let the Internet measure itself. *ACM SIGCOMM Computer Communication Review*, 35(5):71–74, 2005. (p. 38)

[187] Y. Sheffi. *Urban transportation network*. Prentice Hall, 1985. (p. 5)

[188] C. A. Shue and M. Gupta. Packet forwarding: name-based vs. prefix-based. In *IEEE Global Internet Symposium*, pages 73–78, 2007. (p. 42)

[189] B. K. Singh and N. Gupte. Congestion and decongestion in a communication network. *Physical Review E*, 71(5):055103, 2005. (p. 55)

[190] R. Soderbery. How many things are currently connected to the "Internet of Things" (IoT)? *Forbes*, Jan. 2013. (p. 37)

[191] L. Solá, M. Romance, R. Criado, J. Flores, A. G. del Amo, and S. Boccaletti. Eigenvector centrality of nodes in multiplex networks. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 23(3):033131, 2013. (p. 19)

[192] R. V. Solé and S. Valverde. Information transfer and phase transitions in a model of Internet traffic. *Physica A: Statistical Mechanics and its Applications*, 289(3):595–605, 2001. (p. 54)

[193] A. Solé-Ribalta, C. Granell, S. Gómez, and A. Arenas. Information transfer in community structured multiplex networks. *Frontiers in Physics*, 3:61, 2015. (pp. 5, 71)

[194] O. Sporns. *Networks of the Brain*. MIT press, 2011. (p. 3)

[195] S. Sreenivasan, R. Cohen, E. López, Z. Toroczkai, and H. E. Stanley. Structural bottlenecks for communication in networks. *Physical Review E*, 75(3):036105, 2007. (pp. 43, 55)

[196] J. W. Stewart III. *BGP4: inter-domain routing in the Internet*. Addison-Wesley Longman Publishing Co., Inc., 1998. (p. 40)

[197] S. H. Strogatz. Exploring complex networks. *Nature*, 410(6825):268–276, 2001. (p. 3)

[198] M. P. Stumpf and M. A. Porter. Critical truths about power laws. *Science*, 335(6069):665–666, 2012. (p. 40)

[199] B. Tadić and G. J. Rodgers. Packet transport on scale-free networks. *Advances in Complex Systems*, 5(04):445–456, 2002. (p. 4)

[200] B. Tadić, S. Thurner, and G. J. Rodgers. Traffic on complex networks: Towards understanding global statistical properties from microscopic density fluctuations. *Physical Review E*, 69(3):036102, 2004. (p. 66)

[201] M. Takayasu, K. Fukuda, and H. Takayasu. Application of statistical physics to the Internet traffics. *Physica A: Statistical Mechanics and its Applications*, 274(1–2):140–148, Dec. 1999. (p. 54)

[202] M. Takayasu, H. Takayasu, and K. Fukuda. Dynamic phase transition observed in the Internet traffic flow. *Physica A: Statistical Mechanics and its Applications*, 277(1–2):248–255, Mar. 2000. (p. 54)

[203] F. Tan, J. Wu, Y. Xia, and C. K. Tse. Traffic congestion in interconnected complex networks. *Physical Review E*, 89(6), June 2014. (pp. 5, 55, 71)

[204] F. Tan, Y. Xia, W. Zhang, and X. Jin. Cascading failures of loads in interconnected networks under intentional attack. *EPL (Europhysics Letters)*, 102(2):28009, Apr. 2013. (pp. 55, 56)

[205] A. S. Tanenbaum. Computer networks, fourth edition. *ed: Prentice Hall*, 2003. (p. 53)

[206] M. Thorup and U. Zwick. Compact routing schemes. In *Proceedings of the thirteenth annual ACM symposium on Parallel algorithms and architectures*, pages 1–10. ACM, 2001. (p. 42)

[207] M. Timme, F. Wolf, and T. Geisel. Coexistence of regular and irregular dynamics in complex networks of pulse-coupled oscillators. *Physical Review Letters*, 89(25):258701, 2002. (pp. 5, 79, 85)

[208] J. Travers and S. Milgram. An experimental study of the small world problem. *Sociometry*, pages 425–443, 1969. (p. 2)

[209] United Nations. UN Comtrade Database. http://comtrade.un.org. (p. 85)

[210] M. Valencia, J. Martinerie, S. Dupont, and M. Chavez. Dynamic small-world behavior in functional brain networks unveiled by an event-related networks approach. *Physical Review E*, 77(5):050905, 2008. (p. 5)

[211] W.-X. Wang, B.-H. Wang, C.-Y. Yin, Y.-B. Xie, and T. Zhou. Traffic dynamics based on local routing protocol on a scale-free network. *Physical Review E*, 73(2):026111, 2006. (pp. 5, 43, 55)

[212] D. J. Watts and S. H. Strogatz. Collective dynamics of 'small-world'networks. *Nature*, 393(6684):440–442, 1998. (pp. 2, 13, 15)

[213] W. Willinger, D. Alderson, and J. C. Doyle. *Mathematics and the Internet: a source of enormous confusion and great potential*. Defense Technical Information Center, 2009. (p. 40)

[214] G. Yan, T. Zhou, B. Hu, Z.-Q. Fu, and B.-H. Wang. Efficient routing on complex networks. *Physical Review E*, 73(4):046108, 2006. (pp. 43, 55, 56)

[215] S.-H. Yook, H. Jeong, and A.-L. Barabási. Modeling the Internet's large-scale topology. *Proceedings of the National Academy of Sciences*, 99(21):13382–13386, 2002. (p. 38)

[216] L. Zemanová, C. Zhou, and J. Kurths. Structural and functional clusters of complex brain networks. *Physica D: Nonlinear Phenomena*, 224(1):202–212, 2006. (p. 80)

[217] H. Zhang, Z. Liu, M. Tang, and P. M. Hui. An adaptive routing strategy for packet delivery in complex networks. *Physics Letters A*, 364(3):177–182, 2007. (pp. 43, 55)

[218] L. Zhao, Y.-C. Lai, K. Park, and N. Ye. Onset of traffic congestion in complex networks. *Physical Review E*, 71(2):026125, 2005. (p. 54)

[219] H. Zhou and R. Lipowsky. Network brownian motion: A new method to measure vertex-vertex proximity and to identify communities and subcommunities. In *Computational Science-ICCS 2004*, pages 1062–1069. Springer, 2004. (p. 17)

[220] J. Zhou, G. Yan, and C.-H. Lai. Efficient routing on multilayered communication networks. *EPL (Europhysics Letters)*, 102(2):28002, 2013. (pp. 5, 55, 71)

[221] S. Zhou and R. J. Mondragón. Accurately modeling the Internet topology. *Physical Review E*, 70(6):066108, 2004. (p. 38)

[222] V. Zlatić, M. Božičević, H. Štefančić, and M. Domazet. Wikipedias: collaborative web-based encyclopedias as complex networks. *Physical Review E*, 74(1):016115, 2006. (p. 27)