

MAPPING COMPLEXITY THROUGH NETWORK GEOMETRY

From structure to dynamics
of real-world networks

PhD Thesis
Elisenda Ortiz Castillo

Mapping complexity through network geometry: from structure to dynamics of real-world networks

Memòria presentada per optar al grau de doctor
per la Universitat de Barcelona

Programa de doctorat en física

Autora: Elisenda Ortiz Castillo
Directora: Prof. M^a Ángeles Serrano
Tutor: Dr. Giancarlo Franzese



“To M.A. for filling my life with opportunities”

Preface

“Think not of what you see, but what it took to produce what you see.”

BENOÎT MANDELBROT

1 About this thesis

This thesis presents a geometric approach to the study of connectivity patterns and dynamical phenomena in real-world complex networks. The main hypothesis of network geometry states that the architecture of real complex networks has a geometric nature. This has been confirmed during the last decade, with a great number of networks from different domains being successfully mapped into a (hidden) metric space with hyperbolic geometry. Maps have served throughout history as a precise and relevant source of information, at the center of political, economic, and geostrategic decisions. In this thesis, geometric maps of networks play a central role and constitute the underlying thread that weaves the various chapters together. In particular, the work presented in this thesis exploits the current explosion in computing power and the ability of geometric maps to abstract complex topologies in order to unravel fundamental principles underlying the organization and function of real networks.

This thesis starts by introducing complex systems and their mathematical description in terms of networks. In the Preliminary Methods section, we review the \mathbb{S}^1 and \mathbb{H}^2 models, which provide a geometric explanation of networks, and discuss the embedding methods used to produce geometric maps. The next four chapters contain the results of our investigations about the structure of and the dynamics on real complex networks. These are organized so that each chapter covers a different topic, while the focus shifts from structure to dynamics traversing through the combination of both. First, we concentrate on structure and devote to the diagnosis of communities through the development of a specialized geometric configuration model, which provides an alternative to the \mathbb{S}^1 model for generating random versions of geometric networks. Second, we advance our knowledge on structure by providing a thorough analysis of the hierarchical organization of networks while uncovering mechanisms that favor cooperation in social dilemmas susceptible to stratified order. In chapter three, we further enmesh structure and function by coupling opinion dynamics on a network with its metric structure in order to unravel the multiscale anatomy of opinion formation processes in real systems. In chapter four we fully shift our attention to the dynamics of networks and explore the navigability of time-evolving graphs in hyperbolic space. Finally, we present our conclusions followed by a list of publications related to this thesis.

2 Acknowledgements

*“Those swirls in the cream mixing into the coffee? That’s us.
Ephemeral patterns of complexity, riding a wave of increasing entropy from
simple beginnings to a simple end. We should enjoy the ride.”*

SEAN CARROLL

First of all I would like to thank my PhD supervisor for...well, basically everything. For bringing me the opportunity to pursue the path of scientific research, for the guidance she's provided along the way and for pushing me towards excellence everytime I felt unable or unsuitable. I thank her also for being a great role model, for re-directing my creativity when necessary and for working with me under a diversity of circumstances and time-zones. Thank you, this thesis is for sure the result of our joint efforts.

The second person who owns most my gratitude is my colleague Guillermo García-Pérez, for his previous research provided fundamental results and methods upon which I have been able to base my work. I thank him too for being always helpful and a great example to follow. Guille, thanks a lot also for our imaginative conversations and your contagious love for science.

Following, I would like to thank Antoine Allard for the release of "Mercator", which simplified my life orders of magnitude, akin to the wheel for humankind. Next is my collaborator Michele Starnini, who kindly helped improving my research even when he agonized through the process. Additionally, I thank Kolja Kleineberg for always listening to me and providing valuable feedback as well as good advise in terms of PhD life.

I am also greatful to Marián Boguñá for his nice talks which always brought light to the tiniest details of network geometry. I thank as well my other group fellows, Muhua (specially for the nice cover image), Roya and Pedro for their company in seminars and travels. I would also like to express my gratitude to the many ClabB members, who made conferences less scary, seminars more interesting and coffee-breaks and calçotades really enjoyable. Roberta Amato notoriously contributed to all these. A special mention goes to all the people involved in the Complexity72h workshop in Lucca, in particular Chiara Poletto and Eugenio Valdano for supporting our team in the endeavour of developing a draft in 3 days; and to my friends and collaborators Marc Grau and Xavier R. Hoffmann for the excellent time I had (hard-) working with them. I would like to thank again Hoffmann for sharing some beautiful coding skills that facilitated my teaching tasks in Computational Physics.

I am thankful as well for my funding opportunities, both from the James S. McDonnell Foundation and the APIF grant. I also would like to thank Jennifer A. Dunne for her wonderful offer to cover my attendance to the Complex Systems Summer School at Santa Fe Institute and for accepting me for an internship, despite the fact that a pandemic prevented all that. Regarding unfinished experiences, oddly enough, I am grateful too

for the challenge that supposed withdrawing from my initial PhD in New Zealand and the lessons that came with it.

Finally, I would like to deeply thank all my office mates and the Biophysics team for making the long hours in front of the computer more bearable, for their technical and moral support and for the great laughs and occasional beers. Besides, I thank my uni folks Alba, Gemma, Marc and Toni for being an inspiration and motivate me to join the doctor club. Also a big thank you goes to my long-time friends who have politely refrained from asking "aren't you finished yet?" a number of times. Another big thanks is for my friends Mary and Pilar who helped me cope with this thesis in tropical lands.

Last but not least, I thank my mum for her unconditional support, and also my dad who would have loved more than anyone else to see this day come true.

Contents

Preface	iii
1 About this thesis	iii
2 Acknowledgements	iv
1 Introduction	1
1.1 Complexity and network science	1
1.2 From topology to hidden geometry	2
2 Preliminary Methods	7
2.1 Latent metric spaces & geometric models	7
2.1.1 The \mathbb{S}^1 model	8
2.1.2 The \mathbb{H}^2 model	10
2.2 Hyperbolic Maps of real-world networks	12
2.2.1 Maximum likelihood estimation	13
2.2.2 Mercator	15
2.3 Angular heterogeneity	17
2.3.1 Geometric Preferential Attachment	17
2.3.2 Soft Communities in Similarity Space	18
2.3.3 Critical Gap Method	19
3 Geometric randomization	23
3.1 Communities in network maps	23
3.2 The geometric randomization model	24
3.3 Tuning clustering through parameter β	25
3.4 Effects of geometric randomization in real networks	28
3.4.1 Clustering and degree correlations	28
3.4.2 Community structure	31
3.5 Discussion	35
4 Geometric detection of hierarchical backbones	37
4.1 Hierarchical order in networked systems	37
4.2 Definition of hierarchy load of links and nodes	39
4.2.1 Hierarchy load of nodes in terms of angular concentration	42
4.3 Analysis of the hierarchy spectrum	43
4.3.1 Hierarchy load vs geometric communities in synthetic networks	43
4.3.2 Hierarchy spectrum of real networks	43
4.3.3 Hierarchy spectrum of GR network replicas	45

4.4	The similarity filter and hierarchical similarity backbones	46
4.5	Evolution of cooperation in HSBs	52
4.5.1	Evolutionary prisoner's dilemma	52
4.5.2	Game dynamics in HSBs	54
4.6	Discussion	59
5	Multiscale opinion dynamics	61
5.1	Group influence in opinion formation	61
5.2	The multiscale voter model	62
5.3	Simulation of the MVM	64
5.3.1	Consensus in synthetic networks	65
5.3.2	The MVM on empirical networks	68
5.3.3	Opinion domains in the latent space	71
5.4	Discussion	76
6	Navigability of temporal networks in hyperbolic space	79
6.1	Network maps and navigation	79
6.2	Greedy routing on temporal networks	80
6.3	Performance metrics and empirical data sets	81
6.4	Effects of network dynamics on navigability	82
6.5	Optimal activation probability	87
6.6	Heterogeneous activation dynamics	87
6.6.1	Activation of nodes within degree intervals	88
6.6.2	Linear activation depending on degree	90
6.7	Discussion	93
7	Conclusions	95
Resum en català		101
List of publications		107
Publications related to this thesis	107	
Other publications	107	
Bibliography		109
Appendix		123
A Empirical data sets		125

1 Introduction

“I think the next century will be the century of complexity.”

STEPHEN HAWKING

1.1 Complexity and network science

There is a critical difference between complex and complicated. Think, for instance, of the wiring of an airplane. At first, it might be extremely difficult to deduce where every cylinder, cable and bolt goes. However, if one studies the aircraft parts long enough, it will become clear that each one has its own specialized function; and that the plane as a whole is nothing more than the sum of these functions. Complicated systems can be hard to address, but they are ultimately knowable from the study of their parts. Now, add a crew and some passengers to that airplane and, suddenly, the system becomes complex. Not even studying each passenger’s life forever in more and more detail would allow us to achieve certainty about the group behaviour during the flight. The reason is, complex systems involve a massive amount of microscopic components interacting with each other in nonlinear ways. As a whole, a complex system exhibits behaviour at a macroscopic level that is not predictable from the behaviour of its components in isolation. This property, referred to as *emergence* implies the existence of a nontrivial connection between the properties of the system at different scales [1]. It is due to emergence indeed that complex systems are poorly described using a reductionist approach.

Other important properties usually associated with complex systems include feedback loops, memory, adaptation and self-organization. In fact, many real-world systems of very different nature exhibit such properties and hence can be modeled and understood as complex systems. Examples include political organizations, human cultures and languages, national and international economies, stock markets, the Internet, social networks, the global climate, food webs, brains, physiological systems, and even gene regulatory networks within a single cell; essentially, complex networks are everywhere. Given the important role that complex systems play in our daily life, in science and society, their mathematical description and eventually prediction and control have become the topic of an independent research area: network science.

Emerging as a separate discipline at the dawn of the 21st century, network science exploits the fact that behind each discrete complex system there is an intricate network, where nodes represent the system’s elements and links encode the pairwise interactions between them. A key discovery of network science is that the architecture of networks arising in completely different domains are remarkably similar to each other. Conse-

quently, they can be studied using a common set of mathematical tools. Graph theory — a rich field of mathematics founded by the brilliant mathematician Leonhard Euler in 1736 [2] — originally provided the mathematical structures used to model pairwise relations between objects. Beyond that, network science distinguishes from graph theory in its empirical nature and purpose, which sets the focus on data and practicality. Network science does not aim at solely developing abstract mathematical tools to describe a certain network property, instead such tools are usually judged by the insights they offer about a real system's functioning. In addition to the graph formalism acquired from mathematics, network science has adopted the conceptual framework to deal with randomness and seek universal organizing principles from statistical physics [3]. In addition, the exceptional amount of data available nowadays has imprinted the field with a strong computational character, actively borrowing algorithms, database management and data mining techniques from computer science arenas.

In the end, networks permeate science, technology, business and nature to an extent that is well beyond the evident upon a casual inspection. For this reason, in order to understand complex systems, it is mandatory to develop a deep understanding of the networks behind them.

1.2 From topology to hidden geometry

The simplest kind of network representation of a complex system consists of a set of N nodes (or vertices) joined by E links (or edges) and is described in terms of its adjacency matrix A , whose elements a_{ij} express the connections between every pair of nodes i and j . In particular, simple network representations have symmetric adjacency matrices, where $a_{ij} = a_{ji} = 1$ indicates nodes i and j are connected and $a_{ij} = a_{ji} = 0$ otherwise. Other kinds of networks exist, like directed ones, where the directionality of the interactions matters and is captured by non-symmetric adjacency matrices ($a_{ij} \neq a_{ji}$). There are also weighted networks, where the adjacency matrix elements are positive real numbers, usually denoted as w_{ij} instead of a_{ij} , which inform the intensity of the interactions. The mathematical characterization of any network topology implies the description of how links and nodes are set up to relate to each other. To that end, in the following we introduce useful concepts for undirected unweighted graphs which are the most prevalent along this thesis.

A relevant information that can be obtained from the adjacency matrix is the degree $k_i = \sum_{j=1}^N a_{ij}$, which tells the number of neighbors of a node i . When looking at very large systems, like the World Wide Web or some social networks, an appropriate description is better done by means of other statistical measures such as the degree distribution $P(k)$. The degree distribution is defined as the probability for a randomly chosen node to have k neighbors and it allows to classify networks into homogeneous or heterogeneous. The first ones exhibit a fast decaying tail, well described for instance by a Poisson distribution, and hence their average degree, $\langle k \rangle = \sum_k kP(k)$, corresponds to the typical value in the system. This is not the case for real networks which are found

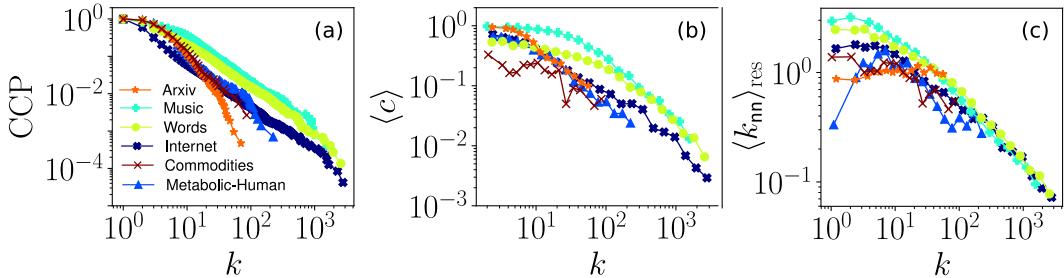


Figure 1.1: Topological properties of several real networks. (a) **Complementary cumulative degree distribution**, $CCP(k) = \sum_{k'=k}^{\infty} P(k')$, representing the probability for a randomly chosen node in the network to have a degree larger than or equal to k , which is less noisy than $P(k)$ (b), **Clustering spectrum**, $\langle c \rangle(k)$, that is, average clustering computed over nodes of the same degree. (c) **Normalised average nearest neighbors degree** of all nodes in every degree class, $\langle k_{nn} \rangle_{\text{res}} = \langle k_{nn}(k) \rangle \langle k \rangle / \langle k^2 \rangle$. The normalization factor $\langle k \rangle / \langle k^2 \rangle$ leaves this curve fluctuating around 1 for uncorrelated networks. This property is used to measure assortativity, which indicates degree-degree correlations. For instance, a network is termed disassortative when low-degree nodes tend to connect to high-degree nodes and vice-versa. Although degree-degree correlations are not universal, in the sense that networks can be assortative, disassortative or even uncorrelated (eg. Arxiv), it is an important structural feature used to characterize real topologies. Descriptions of all real data sets can be found in Appendix.

to be rather heterogeneous, meaning most nodes have small degrees but there is still a non-negligible probability of finding nodes with extremely large degree values. The latter are commonly referred to as hubs. In consequence, real networks usually exhibit heavy tail degree distributions that can be approximated by a power-law

$$P(k) \sim k^{-\gamma} \quad (1.1)$$

with an exponent $\gamma \in (2, 3)$. This means that, in the thermodynamic limit $N \rightarrow \infty$, the fluctuations around the average degree are unbounded, $\langle k^2 \rangle \rightarrow \infty$, and lack any intrinsic scale. In turn, this means that despite the average degree is well defined and finite, it is not a characteristic scale of the system. For this reason, real networks are customarily said to be *scale-free networks*. This is the case despite any real network has a finite size N which ultimately implies a finite $\langle k^2 \rangle$. Beyond its importance in shaping the connectivity structure of real networks, exponent γ has been found to have profound implications on the dynamical processes running on top of them [4,5]. Notwithstanding, in real systems the power-law distribution of degrees tends to coexist with some corrective function [6], yielding power-laws with exponential cutoffs, stretched exponentials, logarithmic corrections and so on. The debate [7] on scale-freeness of real networks still flares up occasionally, largely due to lack of precision of language in the relatively young field of network science. Most scientists though agree that knowledge of whether or not a distribution is heavy-tailed is far more important than whether it can be fit using a pure

power-law [8]. Another agreed important characteristic, yet imprecise in finite systems, is *sparsity*. Real networks are said to be sparse, meaning the network's expected average degree $\langle k \rangle = 2E/N$ is small as compared to the system size. A more formal sense of sparsity exists in the case of infinite network models, where a synthetic graph is called dense or sparse depending on whether the average degree scales linearly or sublinearly with N .

Some networks also present degree-degree correlations, which characterize the relationship between degrees of connected nodes. For instance, technological systems are usually disassortative [9], which means that low-degree nodes tend to connect to high-degree nodes and vice-versa. Social systems, on the other hand, usually exhibit the opposite assortative behaviour [10], according to which there is a positive correlation between the degree of a node and that of its neighbours meaning that high-degree nodes tend to connect to high-degree neighbors. These correlations are typically measured in terms of the average degree of the neighbours of nodes of a given degree, as in Fig. 1.2c. Although degree-degree correlations are not universal, in the sense that networks can be assortative, disassortative or even uncorrelated, it is an important structural feature used to characterize real topologies.

Beyond the degree distribution, a network can also be described by the structure of the neighborhood of a node. Many natural networks exhibit triangles arising in the neighborhood of a given node due to transitive connections. For example, in a social environment this may mean that a friend of a friend is also likely to become your own friend; leaving the three of you connected in a triangular fashion. This property, named clustering, can be accounted for through the *clustering coefficient of a node*,

$$c_i = \frac{2n_i}{k_i(k_i - 1)} \quad (1.2)$$

which measures the relative number of connections among the neighbors of i , or equivalently the probability for two randomly chosen neighbors of i to be connected. The factor $\frac{1}{2}k_i(k_i - 1)$ accounts for the maximum possible number of links among the neighbors of i while n_i stands for the existing number of connections between neighbors of i . Besides, when doing network analysis, a common representation of clustering consists of its spectrum, which shows the average clustering of nodes within a degree class (ie. a set of nodes with equal degree). As shown in Fig. 1.1(b), typically, real networks display high levels of clustering which decay for increasingly large degrees. In the next chapter we will discuss in detail the key role that clustering plays in connecting the topology and geometry of networks.

Related to the high levels of clustering observed in real networks, is the presence of densely connected groups of nodes, or modules, which confer the networks a *community structure* [11]. These tightly knit groups of nodes show only sparse connections with other modules, thus rendering a partition of the network. The diverse practical

applications and informative nature of such divisions, have made the identification and characterization of communities the subject of intense interdisciplinary research over the past decade [12–15]. For instance, community structure is a well-known property of social networks, where it's been long known to affect critical dynamical processes such as epidemic spreading. Consequently, it has become essential to incorporate community structure into the modeling and design of any current epidemic control strategies [16–18].

Along with its clustering coefficient and its degree distribution, one of the most robust measures of network topology is the average shortest path length, $\langle \ell \rangle$. In a network, the length of a path is the number of edges that the path between two nodes contains. Let l_{ij} be the length of the shortest path between nodes i and j so that we define the average shortest path length as

$$\langle \ell \rangle = \frac{1}{N(N-1)} \sum_{i \neq j} l_{ij}. \quad (1.3)$$

In real networks the average distances computed over all pairs of nodes are extremely short and scale as $\langle \ell \rangle \sim \ln N$ or even slower. This property, first discovered in off-line social networks in the 1960's during Milgram's famous experiment [19], is known as the *small-world property* [20]. Furthermore, the presence of hubs (large degree nodes) in real networks has been found to add a significant number of shortcuts which can dramatically reduce the pathlengths, yielding to a scaling behaviour of $\langle \ell \rangle \sim \ln \ln N$. When this is the case, scale-free networks are said to be *ultra-small* [21].

As argued thus far, graphs of real networks are imprinted with some universal features such as scale-freeness, high clustering, the small-world property, community structure, and other well important ones, like *self-similarity* [22, 23] – a characteristic expressing invariance under rescaling, typically associated with certain geometric objects known as *fractals* [24]. In fact, the definition of self-similarity and scale invariance in complex networks was revised and greatly upgraded when topology finally met hidden geometry [25, 26].

The fact that real networks share such common properties despite their entirely different origins suggests the existence of general principles governing their generation and evolution. Several random network models have been proposed in order to unveil such mechanisms, providing different rules that dictate the probability with which the events that generate a network occur. Among the most notable are the: *i*) Erdős–Rényi model [27], which was the first to produce small-world networks but couldn't capture scale-freeness and high clustering levels *ii*) Barabási–Albert model [28], which produced scale-free, small-world networks, however, with vanishing clustering *iii*) Watts–Strogatz model [20], which generates small-world networks with high clustering but homogeneous degrees.

Remarkably, many models for the structure of complex networks have been proposed to date [29–34] but the difficulty in reproducing simultaneously all the mentioned topo-

logical features has remained a constant. The challenge is in part attributable to the fact that it is hard to reconcile the roles of long and short range connections, that is, the small-world property with the presence of strong clustering. To address this issue topological distances were suggested, however, they present a serious problem to distinguish close from far away due to the small-world (or ultra small world) property. As a consequence, not a wide range of length scales is available to help inferring an accurate picture of the system. The discovery that clustering in the topology reflects the triangle inequality of a metric space draw a powerful connection between topology and geometry, suggesting the existence of better suited distances in a geometric domain. It was in fact the first evidence that complex networks may possess some nontrivial geometric properties and, therefore, could possibly be seen as something more than pure topological objects. In fact, many networks are embedded in metric spaces where geometric distances add valuable information. Some spaces are explicit [35] —like in airport networks [36,37], power grids, or urban networks—, whereas some are hidden, yet intimately connected to the topological structure [25,38–40], see Fig. 2.1.

About twelve years ago, the identification of a hidden, or latent, metric space underlying network topologies opened the path to the development of geometric network models. Such models demonstrated to be finally able of explaining, simultaneously, all the most relevant and ubiquitous features observed in real-world systems. These results gave birth to the field of Network Geometry, one of the fundamental research areas within network science today [41].

In the following section, we review the core models of network geometry that constitute the foundations of this thesis and detail the mapping procedures to obtain network representations in a hidden metric space. The subsequent chapters introduce our novel contributions, demonstrating how to apply the network geometric framework to investigate salient structural and dynamical features of real complex networks. In particular, we first cover new aspects of their community and hierarchical organization and secondly provide original insights concerning cooperation dynamics, opinion formation and the navigability of time-evolving networks. Lastly, we conclude with a discussion and summary of the many ways in which our results spark progress in the field of network science.

2 Preliminary Methods

“The greatest strategy is doomed if it’s implemented badly.”

BERNHARD RIEMANN

2.1 Latent metric spaces & geometric models

This section presents a brief introduction to the latent, or hidden, metric space approach to complex networks. Because this thesis relies mainly on the widely used one-dimensional \mathbb{S}^1 model [25] as well as on its purely geometric counterpart, the \mathbb{H}^2 model [42, 43], here we review these two essential models along with their close relationship. Information about other useful versions of these geometric models which are capable of reproducing communities appears in section 2.3 in this chapter.

The first remarkable observation that led to the development of the geometric paradigm was the identification of clustering as the key connection between complex networks and an underlying hidden geometry. Indeed, the triangle inequality in a metric space induces clustering in the structure of the graph, as illustrated in Fig. 2.1. The reason for this is built in the similarity between two nodes and the fact that it shows transitivity. That is, if two nodes A and C are both similar to a third node B, then A and C should also be similar to one another. Consequently, if we consider the space of node’s intrinsic properties a metric space, the distance d between two points becomes a measure of how dissimilar they are so that

$$d_{AC} \leq d_{AB} + d_{BC}, \quad (2.1)$$

meaning the dissimilarity between nodes A and C is bounded by the sum of the dissimilarities of each A and C to B. Hidden metric space models encode both the effect of a similarity distance and the effect of node’s popularity (parallel to their degree) in the connection probability between nodes. This two major forces, allow to generate scale-free, small-world, highly clustered, sparse graphs where the heterogeneity in the degree distribution can be controlled independently of the level of clustering.

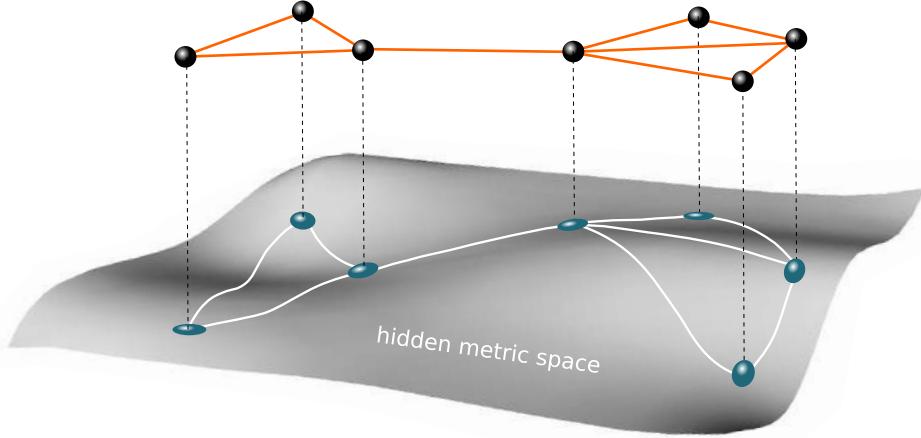


Figure 2.1: Visual representation of the relationship between network topology and hidden geometry. Nodes are scattered in an arbitrary hidden metric space, illustrated by an undulating surface. Distance among nodes is given by the length of the geodesics on the surface, drawn as white lines, so that nodes lying sufficiently nearby become connected. Such connections, give rise to the topology depicted on top with orange straight lines, where triangles naturally emerge as a consequence of the closeness between nodes in the underlying metric space.

2.1.1 The \mathbb{S}^1 model

The \mathbb{S}^1 model is a mixed model, in the sense that it combines a metric component and a topological component. The metric component represents the similarity space, which a priori can be any compact homogeneous manifold of any curvature and dimension D [44]. From a theoretical point of view one dimension suffices in order to generate realistic networks and, in particular, it has been shown that the maximum clustering coefficient obtained from a geometric model decreases as D increases, see supplementary material in Ref. [26]. This suggests that real-world networks, which typically exhibit strong clustering, must belong to the low-dimensionality regime. Taking these facts into consideration, here we restrict ourselves to a model with a single dimension in spherical coordinates.

In the one-dimensional \mathbb{S}^1 model, a node i is assigned two hidden variables: a hidden degree κ_i , quantifying its popularity, and an angular position θ_i in a one-dimensional sphere (or circle) abstracting the similarity space, where angular distances between nodes are a proxy for their similarity. The radius of the circle is adjusted to $R_{\mathbb{S}^1} = N/2\pi$, where N is the number of nodes, so that the density is set to 1 without loss of generality. The connection probability between any pair of nodes takes the form of a gravity law, whose magnitude increases with the product of the hidden degrees (i.e., their combined popularities), and decreases with the (angular) distance between the two nodes. In other words, more similar nodes are angularly closer, hence they are more likely to connect whereas not so similar pairs have high probability of connecting only when they are

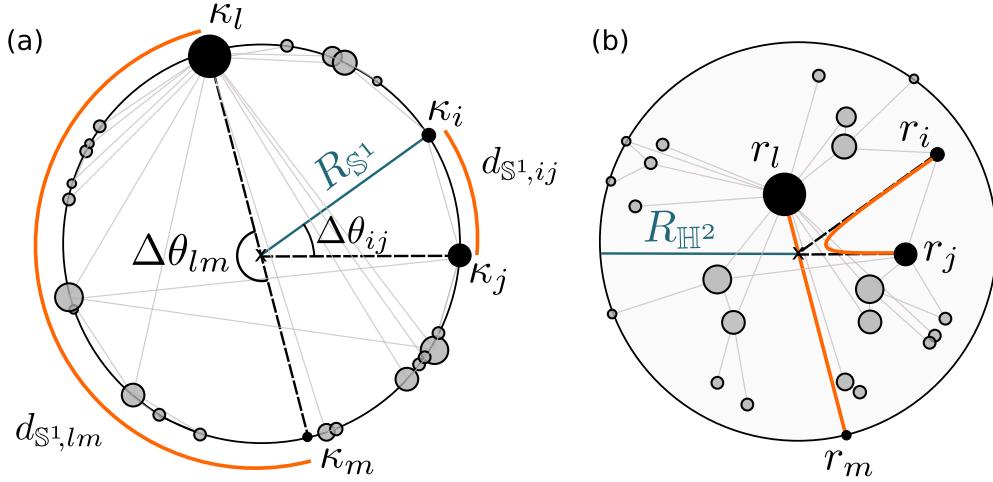


Figure 2.2: (a) S^1 model representation in which the angular distance between two nodes, d_{S^1} , is given by the arc length along the circle of radius R_{S^1} . Nodes with higher hidden degrees, κ , are represented by larger dots. (b) H^2 model representation of the network in (a). The figure corresponds to the native representation of hyperbolic space, where nodes are positioned at radial coordinates r inside a hyperbolic disc of radius R_{H^2} . The higher the degree of a node the closer it lies to the disc center. Orange lines indicate the geodesics (or shortest path curves) between two nodes and showcase the distortion of space in the hyperbolic plane.

popular. Specifically, nodes i and j are connected with probability

$$p_{ij} = \frac{1}{1 + \chi_{ij}^\beta} = \frac{1}{1 + \left(\frac{d_{S^1}}{\mu \kappa_i \kappa_j}\right)^\beta} \quad (2.2)$$

where $\mu > 0$ controls the average degree of the network, β controls its level of clustering, $d_{S^1} = \Delta\theta_{ij} R_{S^1}$, and $\Delta\theta_{ij} = \min(|\theta_i - \theta_j|, 2\pi - |\theta_i - \theta_j|)$ is the angular distance between nodes i and j . A priori, the functional form of the connection probability could be any integrable function of argument $f(\chi_{ij})$. However, the Fermi-Dirac form in Eq. 2.2 is the only possible choice [44] that defines maximally random ensembles of graphs that are simultaneously heterogeneous, clustered, small-worlds, maximally degree-degree uncorrelated and sparse (ie. the average degree is independent of system size). Besides, there are no constraints on the distribution of hidden variables κ and θ . In fact, the angular distribution, $\rho(\theta)$, could be non-homogeneous and both variables could be even correlated. This is an important observation because such angular inhomogeneities or correlations can explain the emergence of communities and other non-trivial topological patterns [38, 45, 46]. Nevertheless, our interest in real-world networks dictates that we choose degree distributions that are power-law $\rho(\kappa) \sim \kappa^{-\gamma}$ with a cut-off κ_c . In fact, the S^1 model can generate networks with any target degree distribution. This is possible because – in the thermodynamic limit – the expected degree of a node is proportional to its hidden degree, $\bar{k}_i(\kappa_i) \propto \kappa_i$, which indeed justifies the name of hidden degree. This

result also implies that parameter β has full control of the level of clustering without affecting the degree distribution. Precisely, it can be shown that the model undergoes a structural phase transition at $\beta = 1$ so that for $\beta < 1$ networks are unclustered in the thermodynamic limit whereas for $\beta > 1$ the ensemble generates networks with finite clustering. In fact, the S^1 model generates networks that are scale-free (if hidden degrees are power-law distributed), sparse, highly clustered and small-world if $2 < \gamma < 3$ or $1 < \beta < 2$, see Ref. [25] and the connectivity phase diagram in Ref. [26]. For further analytical details regarding the properties of the model we refer the reader to Refs. [25, 43, 47].

2.1.2 The \mathbb{H}^2 model

Previously, we noticed that in the S^1 model node's expected degrees can be treated as masses yielding a connection probability akin to Newton's gravitation law. Nonetheless, Einstein claimed that gravity can be described in exclusively geometric terms once we consider the space is curved, that is non-Euclidean. In analogy, hyperbolic geometry naturally emerges as the effective geometry of the S^1 model, [42, 43] so that real networks can be described in a purely geometric fashion.

Hyperbolic space is a metric space with constant negative curvature [48]. Visualizing the hyperbolic plane, \mathbb{H}^2 , can be unintuitive as it cannot be fully embedded isometrically in the 3D-Euclidean space as a surface. Consequently, many representations exist. In this thesis we use the two-dimensional hyperboloid model, where points in \mathbb{H}^2 (of constant curvature $K = -1$) are characterized by polar coordinates (r, θ) and the metric tensor reads

$$ds^2 = dr^2 + \sinh^2 r d\theta^2 . \quad (2.3)$$

A quick comparison with the metric of the familiar 2D-Euclidean plane $ds^2 = dr^2 + r^2 d\theta^2$ reveals how much faster the perimeter of hyperbolic circles grows in contrast to that of Euclidean ones. Beyond doubt, the main property of hyperbolic geometry is the exponential expansion of space, as illustrated in Fig. 2.3, which is analogous to the growth behaviour in trees. The area $A(r)$ of a two-dimensional hyperbolic disc of radius r grows with r as $A(r) \sim e^r$. Henceforth, a uniform density of nodes in a hyperbolic space appears as exponentially growing with distance r from the origin. This also means that for a given node, the number of other nodes at reach will scale as the area of the hyperbolic disc centered at this node, which means exponentially. It is this exponential scaling indeed that sustains the small-world property.

The hyperbolic distance, $d_{\mathbb{H}^2}$, between two points at radial coordinates r_i and r_j separated by an angular distance $\Delta\theta_{ij}$, can be computed using the hyperbolic law of cosines

$$\cosh d_{\mathbb{H}^2} = \cosh r_i \cosh r_j - \sinh r_i \sinh r_j \cos \Delta\theta_{ij} , \quad (2.4)$$



Figure 2.3: Exponential expansion of hyperbolic space is illustrated by the exponentially growing number of flowers on the hyperbolic disc. All flowers are of same hyperbolic size and contained in tiles made of regular hyperbolic polygons. The Poincaré tool developed by Malin Christersson [49] is used to construct the tessellation of the hyperbolic plane in the Poincaré disc model with the Schläfli symbol $\{7,4\}$, by rendering an image of a Plumeria flower floating in tropical waters.

and from this expression another useful approximation for the hyperbolic distance can be derived, which reads

$$\tilde{d}_{\mathbb{H}^2} \simeq r_i + r_j + 2\ln \frac{\Delta\theta_{ij}}{2}. \quad (2.5)$$

As we shall see in the following, it is Eq. 2.5 in fact which allows for the existence of a (quasi-) isomorphism between the \mathbb{S}^1 and \mathbb{H}^2 models. In particular, in order to establish a mapping between the \mathbb{S}^1 model and the \mathbb{H}^2 model, we keep the angular coordinates as in the \mathbb{S}^1 model, but transform the hidden degrees into radial coordinates according to

$$\kappa_i \mapsto r_i = R_{\mathbb{H}^2} - 2\ln \frac{\kappa_i}{\kappa_0}, \quad (2.6)$$

where the radius of the two-dimensional hyperbolic disc containing all nodes is

$$R_{\mathbb{H}^2} = 2\ln \frac{N}{\mu\pi\kappa_0^2}. \quad (2.7)$$

In terms of these new quantities we see that $\kappa \sim e^{-(r-R_{\mathbb{H}^2})/2}$, which means in the hyperbolic version, nodes with larger radial coordinates are located towards the edge of the hyperbolic disc and show lower expected degrees, whereas nodes near the origin tend to have very high degrees. It is the combination of exponentially decreasing average degrees and exponentially increasing node density which explains the emergence of scale-free degree distributions. For instance, when hidden degrees are power-law distributed according to the probability density

$$\rho(\kappa) = (\gamma - 1)\kappa_0^{\gamma-1}\kappa^{-\gamma}; \quad \kappa > \kappa_0 = \frac{\gamma - 2}{\gamma - 1}\langle k \rangle; \quad \gamma > 2, \quad (2.8)$$

then the mapping in Eq. 2.6 makes the radial coordinates to be distributed as

$$\rho(r) = \alpha \frac{\sinh(\alpha r)}{\cosh(\alpha R_{\mathbb{H}^2}) - 1} \approx \alpha e^{\alpha(r-R_{\mathbb{H}^2})} \quad (2.9)$$

for large $R_{\mathbb{H}^2}$. As a result, a power-law degree distribution with exponent $\gamma = 2\alpha + 1$ holds for $\alpha \geq 1/2$ in the hyperbolic plane. Under these circumstances, the node density in the hyperbolic plane is quasi-uniform, which means the node's angular coordinates $\{\theta_i\}$ are distributed uniformly while the radial coordinates $\{r_i\}$ follow Eq. 2.9. Further substituting Eqs. 2.6 and 2.7, into Eq. 2.2 yields to the expression for the connection probability of the \mathbb{H}^2 model

$$p_{ij} = \frac{1}{1 + e^{\frac{\beta}{2}(\tilde{d}_{\mathbb{H}^2} - R_{\mathbb{H}^2})}} \quad (2.10)$$

where $\tilde{d}_{\mathbb{H}^2}$ is, as discussed above, a very good approximation of the hyperbolic distance between two points with coordinates (r_i, θ_i) and (r_j, θ_j) in the hyperbolic disc with curvature $K = -1$.

By mapping every node's hidden degree onto a radial coordinate as dictated by Eqs. 2.6 and 2.7, the similarity and popularity attributes of nodes in the \mathbb{S}^1 model fuse into a single hyperbolic distance in the \mathbb{H}^2 which is responsible for controlling the nodes' probability of connection. Because the approximation $\tilde{d}_{\mathbb{H}^2} \simeq d_{\mathbb{H}^2}$ holds, we refer to Eqs. 2.6 and 2.7 with the word "quasi—" isomorphism. However, it is arguable [41] that the mapping becomes even exact in the thermodynamic limit. Notwithstanding, the powerful quasi-isomorphism existing between the two models guarantees we can use them both interchangeably depending on the application at hand.

2.2 Hyperbolic Maps of real-world networks

As we have seen, latent metric space models bring the possibility to generate realistic complex networks by coupling their topologies to an underlying geometry through a probabilistic connectivity law. Accordingly, a very natural question arises: can we revert the process? That is, for a given real-world network, can we infer the hidden

variables and model parameters that best characterize the observed topology? To do so is indeed possible and known as *embedding* the network into the hidden metric space. Essentially, producing an embedding is equivalent to uncovering a meaningful map of the real system which is depicted in a two-dimensional hyperbolic plane. These hyperbolic maps provide an accurate way of visualizing, storing, and communicating information, help to recognize locational distributions and spatial patterns and relationships, and allow us to track dynamical processes that operate through space at different length scales.

The exercise of embedding real-world networks in hyperbolic space poses some conceptual and technical challenges, therefore, several algorithms have been conceived up to date to solve this problem. In the embedding methods used throughout this thesis, the basic idea is to find the set of hidden variables and model parameters that maximize the likelihood for the observed network to be generated by the model. For this purpose, embedding methods either use maximum likelihood estimation (MLE) techniques [38, 50–52], machine learning [53, 54], or a combination of both [55, 56]. MLE techniques require finding the coordinates of every node in the latent geometry that maximize a likelihood function: a task that, in general, is NP-hard and consequently must rely on heuristics to obtain a solution. Accordingly, MLE methods can be slow, and their accuracy dependent upon the chosen heuristic and the underlying theoretical model. In contrast, machine learning techniques can fasten the embedding process while making it less model dependent. Despite the great contribution of machine learning, however, these techniques alone do not yet outperform MLE existing methods in terms of accuracy, where accuracy refers to the correct inference of angular coordinates as well as other model parameters, including hidden degrees. At present, it is the balance between the two approaches which seems to yield the best results.

Graph embedding is indeed an active field of research providing new procedures every year. Because our application purposes require the embeddings to be faithful maps, in this thesis we adopted at each moment the most novel yet reliable algorithm available. This led us to rely on two different methods during the development of this thesis: the first algorithm, based on MLE, was applied to obtain the embeddings used in Chapter 6; the second, including also machine learning techniques, provided the embeddings in Chapters 3,4 and 5. In the following, we proceed to briefly describe the main insides and steps of each algorithm.

2.2.1 Maximum likelihood estimation

Creating meaningful geometric maps of large real networks has only become possible due to the large amount of empirical data available nowadays and the current explosion in computing power. The algorithm we outline here, was proposed in Ref. [38] with the goal to embed large networks ($N \lesssim 10^5$ nodes) such as the Internet AS network [57], using the \mathbb{S}^1 model and the MLE method. More precisely, the maps are inferred by finding the hidden degree and angular position of each node $\{\kappa_i, \theta_i\}_{i=1,\dots,N}$ that maximize the likelihood that the structure of the network was generated by the \mathbb{S}^1 model, where the

likelihood \mathcal{L} is evaluated as

$$\mathcal{L}(\{a_{ij}\}|\{k_i, \theta_i\}, \mathbb{S}^1) = \prod_{i>j} (p_{ij})^{a_{ij}} (1 - p_{ij})^{1-a_{ij}}. \quad (2.11)$$

Here, p_{ij} is the connection probability given by Eq. 2.2, and a_{ij} are the entries of the adjacency matrix of the network, where $a_{ij} = a_{ji} = 1$ if nodes i and j are connected and 0 otherwise. Note that a_{ij} depends on the network to be embedded, whereas p_{ij} depends on both $\{\kappa_i, \theta_i\}$ and the global parameters of the model: $R = N/2\pi$, μ which is related to the average degree, and β accounting for the clustering. For this reason, the traditional approach to tackle the maximization of the likelihood is also two-stepped.

First, the values of the global parameters are inferred and, secondly, these parameters are used to find the hidden degrees and angular coordinates maximizing Eq. 2.11. In particular, in a network generated by the \mathbb{S}^1 model (and specially for sparse ones), some nodes may end up not being connected and, consequently, being discarded from the resulting network. This makes the number of nodes N in the model to relate to the actual number of nodes N_{obs} in an instance of a network as

$$N_{\text{obs}} = N(1 - P(0)) \quad (2.12)$$

where $P(0)$ stands for the probability of any node to have degree zero. Moreover, the model relies on a finite sequence of N hidden degrees generated from a power-law distribution with exponent γ and mean $\langle \kappa \rangle$, which will have an upper cutoff κ_c that in turn implies

$$\langle \kappa(N) \rangle = \alpha \langle \kappa \rangle < \langle \kappa \rangle \quad (2.13)$$

where $\alpha \rightarrow 1$ as $N \rightarrow \infty$ when $\kappa_c \rightarrow \infty$. In order to estimate the global parameters, the values of $P(0)$ and α need to be found as the solution of a system of equations, see Ref. [38], involving γ , k_{obs}^{\max} , N_{obs} , and $\langle k \rangle_{\text{obs}}$ — which are all measurable from the real network topology. With $P(0)$ and α one can compute N and $\langle k \rangle$, and finally obtain the estimated values of R and μ . The last global parameter β , is chosen such that the clustering coefficient in the model matches the observed in the real network.

Once the global parameters have been inferred, the next step is to proceed to the maximization of Eq. 2.11 with respect to the hidden degrees and angular coordinates. The maximization with respect to the hidden degrees can be done analytically, see Ref. [38] yielding to

$$\kappa = \max \left\{ \frac{\gamma - 2}{\gamma - 1} \langle k \rangle, \frac{1}{\alpha} \left(k - \frac{\gamma}{\beta} \right) \right\}. \quad (2.14)$$

On the contrary, the computation of the angular coordinates maximizing the likelihood function can only be achieved numerically. This computation is expensive due to the large size of the space of all possible angular configurations. Since a brute force approach is prohibitive, it becomes mandatory to use heuristic optimization techniques that exploit the properties of the model. In scale-free networks, subgraphs of high degree nodes have

a higher internal average degree as compared to the complete network. This allows to define an onion-like structure that guides the maximization process. According to a sequence of degrees such $k_1 > k_2 > \dots > k_m$ the network is divided into a nested set of subgraphs $\{G(k_l)\}_{l=1,\dots,m}$. Subsequently, each subgraph $G(k_l)$ is embedded by computing only the likelihood of pairs of nodes i and j belonging to $G(k_l)$. That is, for every node i in $G(k_l)$ we find the angular position θ_i maximizing its local log-likelihood

$$\ln\mathcal{L}_i = \sum_{j \in G(k_{l-1})} a_{ij} \ln(p_{ij}) + \sum_{j \in G(k_{l-1})} (1 - a_{ij}) \ln(1 - p_{ij}) . \quad (2.15)$$

Once all nodes are assigned a θ_i coordinate, a fraction are selected at random and their angular coordinates re-adjusted by further maximizing the local log-likelihoods, now incorporating the information of the recently assigned angular positions of the nodes in the subgraph. This step is repeated until the log-likelihoods converge and we proceed to embed the next subgraph. After conducting the process for all subgraphs the complete embedding of the network is achieved.

Noteworthy, this algorithm can be used to determine the coordinates of newly added nodes without modifying those of previously embedded ones, which is specially interesting in the case of growing systems, for instance like the Internet. Moreover, the heuristics in which the algorithm relies allow for the embedding times to reasonably scale as $\mathcal{O}(N^2)$ with system size. Lastly, notice that node coordinates inferred in the \mathbb{S}^1 formulation are immediately transferable to the hyperbolic plane by converting the hidden degrees to radial coordinates using Eq. 2.6.

2.2.2 Mercator

Mercator is an algorithm, presented as a ready-to-use C++ code [58], that mixes the best of the MLE and machine-learning approaches to discover accurate embeddings in the hyperbolic plane of (large) real networks in $\mathcal{O}(N^2)$ time. The final result is the most solid embedding method currently available in the literature.

Machine learning techniques, such as Laplacian eigenmaps (LE) [54, 56], provide fast and promising methods to aid the embedding process. The LE method is a dimensionality reduction mechanism [59], designed to find dimensional reductions of a set of points embedded in \mathbb{R}^n to an arbitrary dimensional space \mathbb{R}^m with $m < n$. The LE method, however, requires the definition of Euclidean distances between nodes in \mathbb{R}^n , yet we know the geometry of real complex networks is better described by hyperbolic space. Because machine-learning embeddings using LE reductions need to be performed in Euclidean geometry, these methods can only be used to infer the angular coordinates corresponding to the similarity sub-space, while radial coordinates have to be inferred using a geometric model. Consequently, in reality both MLE and machine-learning methods are sensitive to the model used to describe the networks.

The initial step of Mercator consists in providing a set of hidden variables $\{\kappa_i\}$ such that the expected degree of each node in the \mathbb{S}^1 model matches the observed degree in the original network and, at the same time, the theoretical clustering coefficient matches the observed in the real network. This first step, therefore, sets the values of the global parameters β , μ and the hidden degrees while assuming an homogeneous distribution of angular coordinates. We refer the reader to the Appendix in Ref. [58] for further details regarding this initial assignment of hidden variables.

The second step is to find an angular ordering of the nodes in the \mathbb{S}^1 circle to act as a base from where to assign the nodes' precise angular coordinates. The angular ordering is found using LE in a similar fashion as in Ref. [54] and gives angular coordinates in \mathbb{R}^2 for a network whose structure can be modeled by the \mathbb{S}^1 model. That is, given a set of points $\{\mathbf{x}_i\}_{i=1,\dots,N}$ in \mathbb{R}^n with Euclidean metric, LE is used to find a mapping of these points $\{\mathbf{x}_i \mapsto \mathbf{y}_i \in \mathbb{R}^m\}$ with $m < n$ that minimizes the loss function

$$\epsilon = \sum_{i,j} |\mathbf{y}_i - \mathbf{y}_j|^2 \omega(|\mathbf{x}_i - \mathbf{x}_j|^2) , \quad (2.16)$$

where $|\mathbf{y}_i - \mathbf{y}_j|$ is the Euclidean distance between points i and j in \mathbb{R}^m and $\omega(\cdot)$ is a decreasing weight function of the distance between the same points in \mathbb{R}^n . Intuitively, the loss function in Eq. 2.16 increases if we place pairs of points far apart in \mathbb{R}^m when they were originally close in \mathbb{R}^n , thus providing a faithful representation of the data in the reduced space. The weight function $\omega(\cdot)$ is taken to be proportional to the adjacency matrix so that it only takes values different from zero if nodes i and j are connected. Such weight values associated to connected nodes are estimated as the nodes' expected chord length distances, d_{ij} , in the \mathbb{S}^1 model

$$d_{ij} = 2 \sin \frac{\langle \Delta \theta_{ij} \rangle}{2} \quad (2.17)$$

where $\langle \Delta \theta_{ij} \rangle$ is the expected angular separation between connected nodes i and j in the \mathbb{S}^1 model. The set of coordinates $\{\theta_i\}$ that minimize the loss function ϵ is obtained as the solution of a generalized eigenvalue problem with the Laplacian matrix, for which very fast algorithms exist if the network is sparse [60]. In this way, a sorted list of nodes based on their angular position is obtained and the angular separation between each consecutive pair of nodes is computed. The angular gap between two consecutive nodes is compared to their expected gap given by the \mathbb{S}^1 model and in the end the biggest separation is selected for each pair so that a final set of angular coordinates $\{\theta_i\}$ is obtained.

The above steps yield a first (and fast) embedding that uses only information about pairs of connected neighbors and is referred as the "fast mode Mercator". However, this embedding can be further improved in the "refined Mercator" version by using it as the starting point in a MLE optimization. This optimization is based on the \mathbb{S}^1 model, which indeed uses information from both connected and not-connected pairs of nodes (see Eq. 2.15), therefore increasing the accuracy of the inferred coordinates. Moreover,

in the refined version it is also possible to find better estimates of the hidden variables κ_i using the newly inferred angular positions. Although this last step is optional, it generally leads to substantial improvements of the final embedding.

The Mercator embedding algorithm used throughout this thesis is open source. The code is currently available from <https://github.com/networkgeometry/mercator>.

2.3 Angular heterogeneity

In this section we address a key aspect that differentiates real-network embeddings from the synthetic geometric networks generated by the $\mathbb{S}^1 / \mathbb{H}^2$ models. This crucial discriminating factor is angular heterogeneity. Angular coordinates in the $\mathbb{S}^1 / \mathbb{H}^2$ models are distributed homogeneously whereas in real network maps we observe (see Fig.2.4) that nodes populate some angular regions more densely than others. These crammed regions, which appear separated by angular gaps in the hyperbolic disc representation, are known as *geometric* or *soft communities*. The areas where nodes more profusely connect with each other in fact reveal clusters of similar nodes and thus encode the community structure of the network in geometric terms. It is interesting then to introduce geometric models that can naturally recreate soft communities in their generative processes, as well as detection methods specifically designed to identify community structure in geometric networks.

In the following we review the Geometric Preferential Attachment (GPA) and the Soft Communities in Similarity Space (SCSS) models, which account for the emergence of soft communities by introducing correlations between hidden degrees and angular coordinates. The two models differ in that the GPA model is a growing model while the SCSS consists of a static formulation of the \mathbb{S}^1 that generates soft communities through a predefined angular distribution. Both models are used in Chapter 3 and just the SCSS again in Chapter 4. After examining the GPA and SCSS models, we review the Critical Gap Method (CGM), a geometric community detection mechanism used in chapters 3,4 and 5.

2.3.1 Geometric Preferential Attachment

The Geometric Preferential Attachment (GPA) [61] model is a generative mechanism explaining scale-free degree distributions, strong clustering and community structure in growing networks. The model is a generalization of the Popularity vs. Similarity Optimization (PSO) [62] geometric model, which is a growing version of the \mathbb{H}^2 explaining preferential attachment as an emergent property. Remarkably, in the GPA model community structure emerges from the growth dynamics of the network.

The growing procedure in the hyperbolic plane operated by the GPA makes the angular coordinates and hidden degrees of different nodes to be correlated. This is in contrast

to the bare \mathbb{S}^1 model, where node's hidden variables are withdrawn independently from some joint distribution $\rho(\kappa, \theta)$ that factorizes. During the GPA growth process, the network is initially empty and then new nodes t start to appear one at a time so that the older the node is, the higher its degree. When a new node t is added to the system, the probability for it to be placed at polar coordinate θ_t depends on the current node density around this angular position. In particular, the higher the node density the more likely for the node to materialize at (r_t, θ_t) . This implies the angular coordinate of a node depends on the angular coordinates of all previous nodes with higher degrees. Consequently, the heterogeneity in angular node density must be controlled at time 0, by some initial attractiveness of different angular regions. This initial attractiveness, is abstracted and regulated by parameter Λ . Specifically, heterogeneity of node density is a decreasing function of Λ such that for $\Lambda \rightarrow \infty$ we recover a homogeneous distribution. In that latter case θ_t of a new node t would be just sampled uniformly at random on $[0, 2\pi]$, making the GPA equivalent to the standard \mathbb{S}^1 model.

Since the connection probability of the GPA model does not explicitly depend on Λ , neither does the degree distribution or the clustering coefficient. Therefore, both topological properties can be tuned independently of community structure emerging from the growing model. See proof in Methods section of Ref. [61]. Ultimately, the GPA produces synthetic networks similar in many ways to real-world ones essentially thanks to the preferential attachment of nodes to the latent network geometry.

2.3.2 Soft Communities in Similarity Space

Analogously to the GPA, the Soft Communities in Similarity Space (SCSS) [63] model generates networks with targeted topological features and soft communities, that is, inhomogeneous angular distributions. Considering heterogeneous angular coordinates requires hidden degrees to be adjusted to depend on the angles. The SCSS provides a general algorithm to numerically find the corrected hidden degrees yielding the desired observed degree distribution for any inhomogeneous angular distribution of nodes. Despite this static model can be fed with any angular distribution, a sequential assignment of angles obtained from the GPA model is frequently used as it emerges from a preferential attachment process in similarity space, which seems to be a plausible explanation for the emergence of communities in real systems. The key advantage of using the SCSS model versus the GPA is that, whereas the level of clustering of the GPA model is the maximum possible by construction, the SCSS model allows to tune the average clustering coefficient of the generated networks as a function of parameter β .

In practice, the SCSS model takes an heterogeneous angular distribution and adjusts it to an independent power-law degree distribution $P(k) \sim k^{-\gamma}$ and a tunable level of mean clustering $\langle c \rangle$ by conducting the following steps:

- a)** Generate a set of N target degrees k^{tar} from a power-law distribution with exponent γ . Order the target degrees such that $k_1^{\text{tar}} > k_2^{\text{tar}} > \dots > k_N^{\text{tar}}$.

b) Assign to every node i a hidden degree κ_i , where i corresponds to the labelling in step 1, initially set to $\kappa_i = k_i^{\text{tar}}$.

c) Repeat N times:

1. Choose at random some node i .
2. Compute the expected degree \bar{k}_i of node i as

$$\bar{k}_i = \sum_{j \neq i} \frac{1}{1 + \left(\frac{d_{ij}}{\mu \kappa_i \kappa_j} \right)^\beta}, \quad (2.18)$$

where $\mu = \frac{\beta \sin\left(\frac{\pi}{\beta}\right)}{2\pi \langle \kappa \rangle}$ as in the \mathbb{S}^1 model.

3. Correct the value of κ_i so that the expected degree \bar{k}_i matches the target degree k_i^{tar} . One method is to reset $|\kappa_i + (k_i^{\text{tar}} - \bar{k}_i)\delta| \rightarrow \kappa_i$, where δ is a random variable withdrawn from the uniform distribution $U(0, 0.1)$. The random fluctuations of δ prevent the system from getting stuck at local optima.

d) Compute all relative deviations

$$\epsilon_i = \frac{|k_i^{\text{tar}} - \bar{k}_i|}{k_i^{\text{tar}}}. \quad (2.19)$$

If $\max\{\epsilon_i\}_i < \eta$, where η is a tolerance set to $\eta = 10^{-2}$, continue to e). Otherwise, repeat step 2 in c).

e) Lastly, generate the network by connecting every pair of nodes using the \mathbb{S}^1 model connection probability (Eq. 2.2).

The SCSS model demonstrates how the \mathbb{S}^1 model can be used beyond the simplifying premises of independent hidden variables and strict power-law degree distributions to generate realistic networks. As a result, the SCSS allows to generate complex networks with soft communities with any given heterogeneous angular distribution.

2.3.3 Critical Gap Method

The Critical Gap Method (CGM) is a community detection method that uses the geometric framework to find a partition of the network into modules. This method is useful to discover geometric (or soft) communities [61] in both real and synthetic networks with heterogeneous angular distributions. In the geometric context, a community is defined as a group of nodes in the similarity space separated from the rest by two angular gaps that mark the boundaries of the community, see Fig. 2.4. Hence, community boundaries are naturally established through the notion of *critical gap*. The critical gap refers to the particular angular distance, $\Delta\theta_c$, that separates consecutive nodes such that no other nodes exist in between. Therefore, the critical gap distance can be used to divide a network into angular sectors separated by angular gaps that satisfy $\Delta\theta \geq \Delta\theta_c$. Different

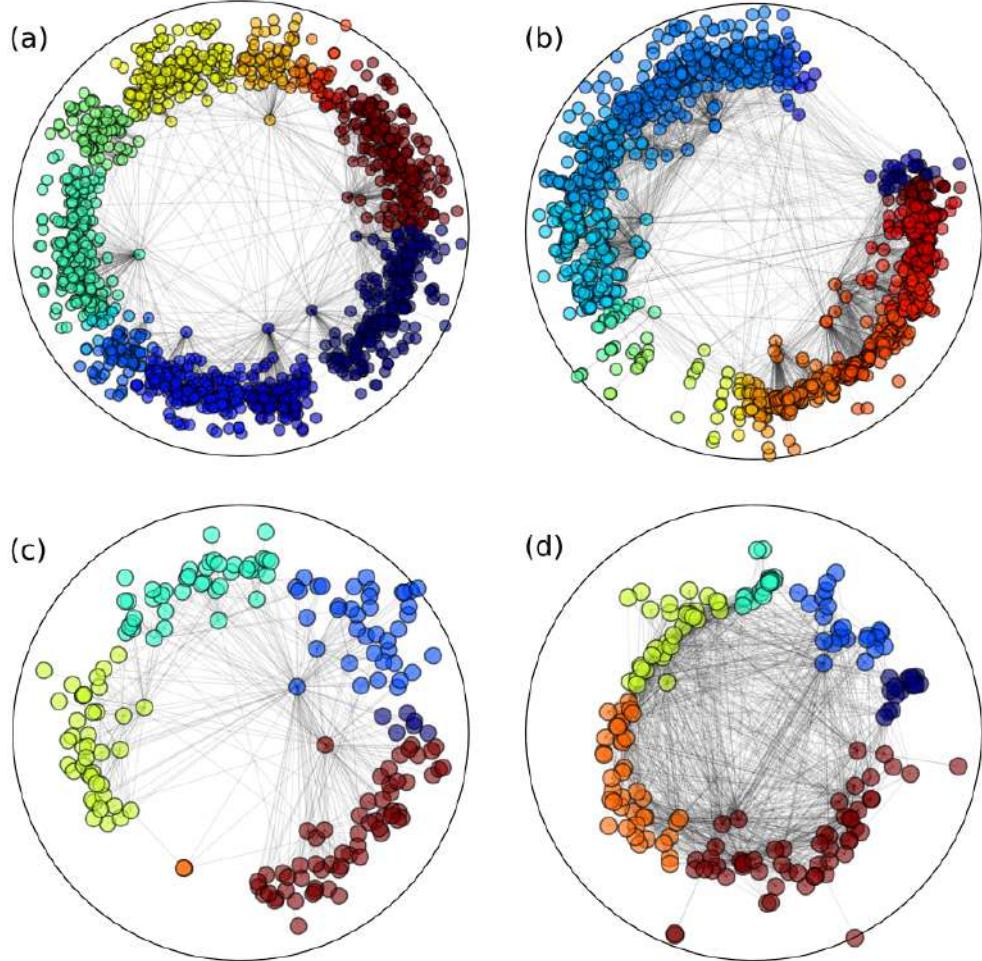


Figure 2.4: Hyperbolic maps of (a) a synthetic network obtained using the SCSS model with $N = 1000$, $\beta = 2.2$, $\gamma = 1.8$ and heterogeneous angular distribution obtained from the GPA model using initial attractiveness $\Lambda = 0.1$. (b) Metabolic Human (c) World Trade Web (d) and Enron Social networks. For further details of each real data set please refer to Appendix. In all hyperbolic maps, nodes of same color belong to the same geometric community as detected from the partition discovered by the CGM algorithm.

partitions of the network are obtained for different $\Delta\theta_c$, so it is essential to establish a criteria to determine the critical gap value.

Most frequently, the CGM refers to the algorithm that determines $\Delta\theta_c$ by maximizing the modularity $Q \in [-1, 1]$ of the obtained partition of the network. Modularity [16] is a standard measure used in community detection, which compares the fraction of links inside communities with the expected fraction for a random distribution of edges. If Q equals 0 nodes are randomly assigned into communities. The steps followed by the CGM are:

- Set a variable q to store the highest value of Q and initialize $q = -1$.
- Increase $\Delta\theta_c$ from 0 to 1 and for each value:
 1. Connect all pairs of nodes (i, j) whose normalized angular distance $\Delta\theta_{ij}^* \in [0, 1]$ satisfies $\Delta\theta_{ij}^* \leq \Delta\theta_c$.
 2. Find all connected components in the resulting network and assign all nodes in the same connected component to the same community.
 3. Compute the modularity Q of the partition. If $Q > q$ set $Q \rightarrow q$ and keep the partition.

At the end of the process, the critical gap $\Delta\theta_c$ obtained yields the partition of the network with highest modularity. In fact, the method finds a modularity comparable to that of other community detection mechanisms in the literature but frequently with higher resolution [46], that meaning, the method usually identifies a larger number of smaller and less discernible communities.

Another criteria used to determine the critical gap is to calculate $\Delta\theta_c$ as the expected value of the largest gap between two nodes when the angular coordinates are distributed uniformly at random, that is: $\Delta\theta_c = \frac{2\pi H_N}{N} \approx 2\pi \ln(N)/N$, where H_N is the N^{th} harmonic number, see Ref. [61]. This choice of the critical gap is commonly known as the Geometric CGM (G-CGM) to distinguish it from the critical gap method that maximizes modularity, which is usually referred just as CGM or Topological CGM (T-CGM). The G-CGM has as well a remarkable accuracy in detecting fine communities and provides significant levels of modularity.

3 Geometric randomization

*“When you create something new, you’re breaking tradition
– which is an act of defiance.”*
STEVEN STROGATZ

This chapter was – with some small changes – published in “New Journal of Physics” by IOP Publishing Ltd. under the title “Geometric randomization of real networks with prescribed degree sequence” [64]. A preprint version is also available at [65].

3.1 Communities in network maps

A natural starting point in the process of understanding complex systems through the use of geometric network maps is to learn to accurately interpret the explicit information they contain. A notable characteristic of the angular coordinates obtained from the embedding of real networks is non-uniformity. As shown in section 2.3, real hyperbolic maps usually display angular regions where the density of nodes is higher than average followed by large empty gaps separating such regions, see Fig. 2.4. The reason for the appearance of these denser areas (known as soft or geometric communities) is that, within any of them, the link density is higher than average. This is, nodes in such regions tend to connect with each other abundantly, thus uncovering the community structure of the network.

Geometric communities have proved to be very meaningful and highly congruent with metadata not contained in the topology of the networks. For instance, they are able to reveal geographic correlations in the Internet at the autonomous systems level [38], trade-communities consistent with geopolitical blocks [46], and expose functional neuroanatomical modules in the human brain [66]. Consequently, the angular heterogeneity of the similarity coordinate featured in geometric maps is key in explaining the anatomy of real networks. This observation opens the door to the use of geometric models with homogeneous similarity distribution as null models for the investigation of the community organization and other structural properties of real networks.

In this chapter, we present a rewiring procedure [67] for the randomization of complex networks with geometric structure that can be used as a benchmark for the analysis of soft communities. The geometric randomization (GR) model, as we named it, is akin to a geometric version of the configuration model, where in our case, the networks with fixed degrees in the ensemble present clustering but otherwise are maximally random. The

GR model exactly preserves the degree sequence of the input network while completely randomizes the angular coordinates of the nodes, and maximizes the congruency with the underlying geometry. Such randomization of the similarity coordinate supports the use of the GR as a null model for the analysis of the topological properties of real networks. The main goal of the geometric randomization model is that of allowing us to compare a real network topology with an analogous counterpart, where only the original heterogeneous angular organization has been discounted. For that purpose, the GR model assumes the same form of the connection probability as in the \mathbb{S}^1 model, and a uniform distribution for the similarity coordinate. Additionally, and in contrast to the \mathbb{S}^1 model, it is fit with a given degree sequence. Gainfully, the use of prescribed degrees allows to skip the delicate task of estimating hidden degree variables from real data in the process of adjusting the \mathbb{S}^1 model to observations. This attribute is remarkably helpful, for instance, in the analysis of features which are specially sensitive to fluctuations of the degree cutoff, like the behavior of dynamical processes such as epidemic spreading or synchronization, or for high-fidelity reproduction of real network topologies.

3.2 The geometric randomization model

The geometric randomization model makes sense on networks with geometric structure—small-worlds with heterogeneous degrees and high levels of clustering—where nodes have an observed degree and exist in a similarity space. The similarity space is taken to be a circle, as in the $\mathbb{S}^1 / \mathbb{H}^2$ models. As reviewed in Chapter 2, in the \mathbb{S}^1 model every node i is characterized by a popularity-similarity pair (κ_i, θ_i) , where κ_i is the node's hidden degree (expected to be proportional to the observed degree k_i) and θ_i its angular or similarity coordinate. In principle, geometric randomizations of networks can be obtained using directly the \mathbb{S}^1 model with parameters γ , β and μ —controlling the exponent of the power-law hidden degree distribution, the clustering coefficient, and the average degree, respectively—estimated from the empirical network. This alternative however, requires the explicit estimation of the hidden degree sequence $P(\kappa)$ or of the exponent of the hidden degree distribution, and, thus, it may introduce undesired fluctuations in the degree cutoff of synthetic networks which can induce relevant differences between the topological properties of real and \mathbb{S}^1 generated replicas.

In the GR model, instead, only angular coordinates are assigned to the nodes, chosen uniformly at random from $[0, 2\pi)$. The network is then rewired in order to maximize the likelihood that the new topology is generated by the \mathbb{S}^1 model while preserving the observed degrees, and thus the total number of edges E . The rewiring procedure is conducted by executing the Metropolis-Hastings algorithm, aimed at finding the network connectivity (i.e. the adjacency matrix a_{ij}) that maximizes the likelihood function

$$\mathcal{L} = \prod_{i < j} p(\kappa_i, \kappa_j, \Delta\theta_{ij})^{a_{ij}} \left[(1 - p(\kappa_i, \kappa_j, \Delta\theta_{ij}))^{1-a_{ij}} \right], \quad (3.1)$$

where $\Delta\theta_{ij}$ stands for the angular distance between nodes i and j , and the \mathbb{S}^1 connection

probability $p(\kappa_i, \kappa_j, \Delta\theta_{ij})$ reads

$$p(\kappa_i, \kappa_j, \Delta\theta_{ij}) = \frac{1}{1 + \left(\frac{\Delta\theta_{ij}R}{\mu\kappa_i\kappa_j}\right)^\beta} = \frac{1}{1 + \chi_{ij}^\beta} = \tilde{p}(\chi_{ij}). \quad (3.2)$$

Parameter μ depends on the observed average degree $\langle k \rangle$ of the network, and R is the radius of the circle, adjusted to have a density of nodes equal to 1.

The algorithm proceeds by repeating the following steps:

- Compute the current likelihood \mathcal{L}_c
- Two links, between nodes i and j , and between nodes l and m , are randomly chosen and swapped: the new links connect nodes i and m , and nodes l and j avoiding self-loops and multiple connections.
- Compute the new likelihood \mathcal{L}_n
- If $\mathcal{L}_n > \mathcal{L}_c$ accept the link swap
- Otherwise, if $\mathcal{L}_n < \mathcal{L}_c$ accept the link swap with probability $\mathcal{L}_n/\mathcal{L}_c$

The rewiring algorithm is terminated after a number E^2 of edges are chosen to be swapped, ensuring that the likelihood has reached a plateau, meaning that the chances that the new rewired network is a geometric one generated by the \mathbb{S}^1 model are maximum. Notice that at the end of the rewiring procedure the degrees of the nodes have not changed but the resulting network might not be connected. Since the hidden degrees are kept constant (independently of their values), the probability of swapping links between nodes i and j and between nodes l and m simply reads

$$\mathcal{L}_n/\mathcal{L}_c = \left(\frac{\Delta\theta_{ij} \Delta\theta_{lm}}{\Delta\theta_{il} \Delta\theta_{jm}} \right)^\beta. \quad (3.3)$$

Therefore, the GR model does not actually require to estimate the hidden degrees of the nodes because they do not enter in any step of the algorithm. In contrast, the GR model simply needs to assign uniformly distributed angular coordinates and give a value for the clustering parameter β . We cover this last step in more detail in the following.

3.3 Tuning clustering through parameter β

In order to randomize a network using the GR model one simply needs to read the degrees in the network and fix parameter β , which controls the level of clustering in the network. As explained in Chapter 2.1, clustering is a signature of the metricity of geometric networks and gives the connection between the observed topology and the

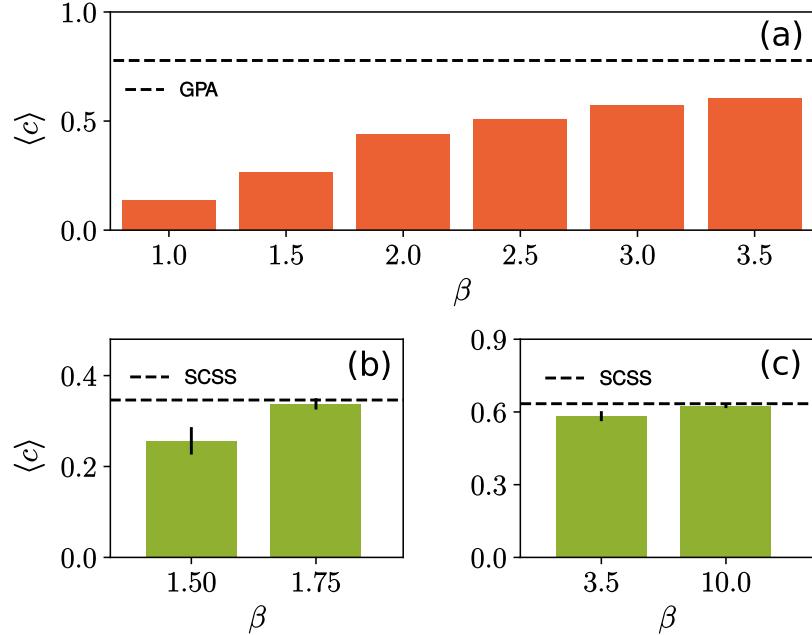


Figure 3.1: (a) Average clustering $\langle c \rangle$ of a network generated by the GPA model (dashed line) and rewired versions (orange) obtained by applying the GR model with different values of β . All networks have size $N = 10^3$, exponent of the degree distribution $\gamma = 2.5$, number of links per node $m = 4$, and the initial attractiveness $\Lambda = 0.1$. (b) Average clustering of two networks generated with the SCSS model (dashed line) with attractiveness $\Lambda = 0.1$, $N = 10^3$, exponent of the degree distribution $\gamma = 2.5$ and $\beta_0 = 1.5$ in (b) and $\beta_0 = 3.5$ in (c). Green bars indicate the $\langle c \rangle$ of networks obtained by applying the GR with β_0 and with β , respectively.

underlying metric space, as a reflection of the triangle inequality.

Note that the value of β affects the probability to accept a link swap (see Eq. 3.3) so it determines the final network's structure. We address the role of β by applying the GR model to synthetic networks generated by the Geometric Preferential Attachment (GPA) model [61] and the soft communities in similarity space (SCSS) model [63]. Both models are intended to produce synthetic networks with tunable community structure, the first implementing a growing strategy and the second a static formulation. See section 2.3 for a comprehensive review.

Essentially, the GPA model generates geometric networks with soft-communities using a growing preferential attachment mechanism in the hyperbolic plane. The probability of connection depends on parameter Λ controlling the initial attractiveness of the different angular regions, such that the heterogeneity of the angular coordinate is a decreasing function of Λ , with $\Lambda \rightarrow \infty$ recovering the homogeneous distribution. Notice that the degree distribution and the clustering coefficient in networks generated by the GPA model are independent of Λ and, in particular, one can decide the $\langle k \rangle$ as well as the exponent

Data set	Nodes	N	γ	β_0	β	$\langle k \rangle$	$\langle c \rangle$	D_{KS}
Enron E.	Email threads	33696	2.14	2.70	2.60	10.73	0.71	0.027
Commodities	US industrial sectors	374	2.50	1.06	1.25	5.83	0.22	0.144
Metabolic H.	Metabolites	1436	2.60	2.13	2.50	6.57	0.54	0.092
Words	Words	7377	2.25	1.01	1.00	11.98	0.47	0.116
Internet	Autonomous Systems	23748	2.16	1.88	2.20	4.92	0.61	0.123
Music	Chords	2476	2.27	2.50	2.65	16.66	0.82	0.072

Table 3.1: Properties of the data sets under consideration: N , size of the network; γ , exponent of the power-law fitting the degree distribution, $P(k) \sim k^{-\gamma}$; parameter β_0 estimated from the embedding of the real network; parameter β that preserves the level of clustering in the GR network; $\langle k \rangle$, average degree; and the D score (95% CI) of the KS test performed between the $P(\theta)$ distributions of the original networks and networks obtained by applying the GR model.

of the power-law degree distribution. However, $\beta \rightarrow \infty$ by construction and, thus, the level of clustering is always the maximum possible. In contrast, the SCSS model consists of a static formulation of the \mathbb{S}^1 model for the generation of soft communities that allows to change the generated level of clustering as a function of β .

Fig. 3.1(a) shows the average clustering coefficient $\langle c \rangle$ of a GPA network compared with the randomizations obtained after applying the GR model using different values of β . As expected, the average clustering of the rewired networks strongly depends on the value of β : the lower β , the lower $\langle c \rangle$ in the resulting network. A level of clustering similar to GPA values can be obtained in GR networks by using large values of β , such as $\beta = 10$. In Fig. 3.1(b)-(c), we report the average clustering coefficient obtained by applying the GR model to synthetic networks generated with the SCSS model. The SCSS networks are produced using two different generating values, referred as β_0 . Fig. 3.1(b)-(c) show that it is possible to fine tune the value of β used by the GR networks so that they reproduce the same average clustering $\langle c \rangle$ as the original networks. If the generation value β_0 is used for the rewiring, the level of clustering in the GR instances does not reach that in the original networks and remains smaller. This observation can be understood by noticing the following two points. First, for SCSS networks the $\langle c \rangle$ is independent of the level of angular clusterization, so any two SCSS networks with equal β_0 and the same distribution of hidden degrees, $P(\kappa)$, will have equal $\langle c \rangle$. Second, a GR instance of a SCSS network obtained using β_0 would be one with homogeneous $P(\theta)$ and the same observed degree distribution $P(k)$ as in the SCSS network. That is, if $P(k) = P(\kappa)$ exactly, then the average clustering $\langle c \rangle$ reached by the GR instance with β_0 would need to match that of the SCSS network. Since we do not observe this matching in Fig. 3.1(b)-(c), we conclude it is due to differences between the distribution of observed and hidden degrees of the SCSS network.

3.4 Effects of geometric randomization in real networks

In the following, we apply the GR model to real networks in order to understand to which extent a network's heterogeneous angular distribution can affect key structural attributes such as its degree-degree correlations or clustering. We consider six empirical networks from different domains: the US commodities network (Commodities), the network of chords transitions in western popular music (Music), the one-mode projection onto metabolites of the human metabolic network at the cell level (Metabolic Human), the word adjacency network in Darwin's book *On the Origin of Species* (Words), the email communication network within the Enron company (Enron Email), and the Internet at the autonomous system level (Internet). See Table 3.1 for further details and Appendix A for detailed descriptions of the data sets.

As described in the previous section, β is the only free parameter of the model, and can be used to tune the clustering coefficient. In the following, we will show results by using a value of β ensuring that the average clustering coefficient of the rewired network is equal to that of the real one. Another possible choice for β is the value estimated when embedding the real network into the underlying metric space (using Mercator embedding tool), which we indicate as β_0 in Table 3.1. The embedding method estimates the coordinates of the nodes in the underlying geometry by maximizing the likelihood that the observed topology has been produced by the model. In the process, β_0 is estimated such that the expected clustering coefficient of the embedded network matches the observed clustering coefficient of the network topology. As explained in the previous section for synthetic networks, using β_0 as the input in GR does not produce in general rewired networks with the same average clustering $\langle c \rangle$ as the original networks. For real networks, the two values of β are very similar but not always identical, see Table 3.1. The small difference is related to the fact that, for some real networks, the GR model cannot adjust simultaneously the empirical connection probability and the observed clustering using a single value of β , see Fig. 3.2.

3.4.1 Clustering and degree correlations

Figure 3.3 shows the average clustering $\langle c \rangle$ of the empirical networks under consideration as compared to the randomized versions obtained by the GR model. We consider both values β and β_0 (the corresponding networks are indicated by GR and GR_0 , respectively), and we include also a comparison with real network replicas generated by the S^1 model. To produce replicas of the real networks using the S^1 model, we extracted the parameters from the empirical networks, namely the size N and the exponent γ of the degree distribution, and used the exponent β_0 given by the embedding of the network into the hyperbolic disc. In order to generate the hidden degree sequence $P(k)$ we adjusted the parameter μ to obtain the observed average degree $\langle k \rangle$ reported in Table 3.1.

As expected, GR networks show an average clustering practically identical to that of the original data, while GR_0 networks present mild deviations, and differences are

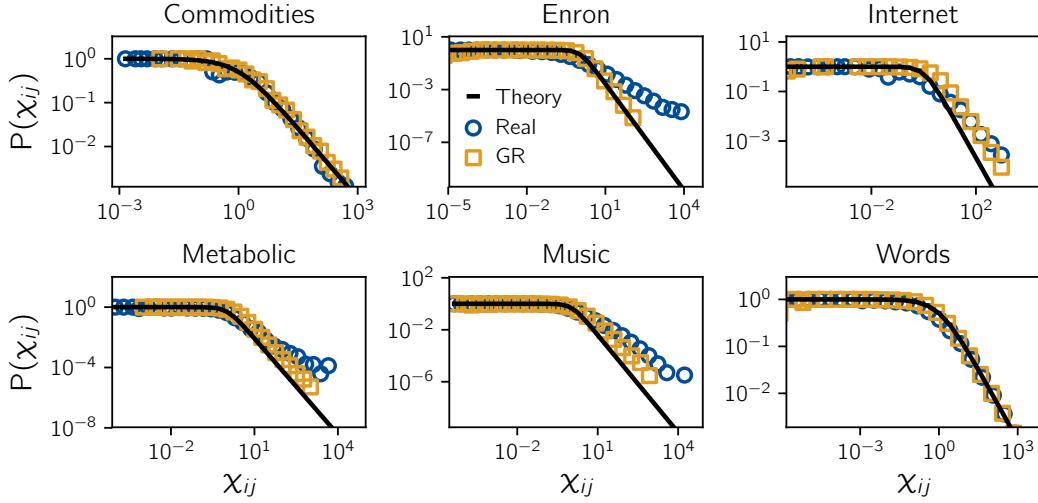


Figure 3.2: Empirical connection probability for original (blue dots) and GR (orange dots) networks. Fraction of connected pairs of nodes as a function of $\chi_{ij} = \Delta\theta_{ij}R/(\mu\kappa_i\kappa_j)$. The black line shows the theoretical curve, Eq. 3.2.

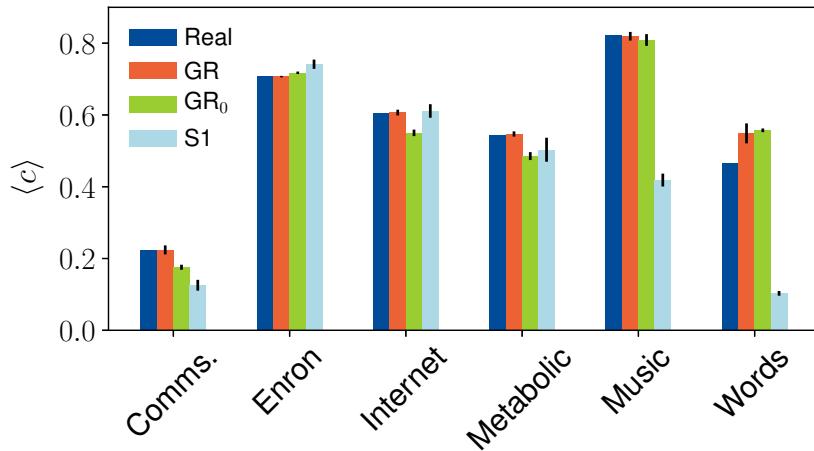


Figure 3.3: Average clustering $\langle c \rangle$ of empirical networks (blue), networks obtained from the GR (red) and S1 (light blue) models. GR networks obtained with β_0 (green) are indicated as GR₀. Error bars are calculated over 10 realizations of the GR and \mathbb{S}^1 models.

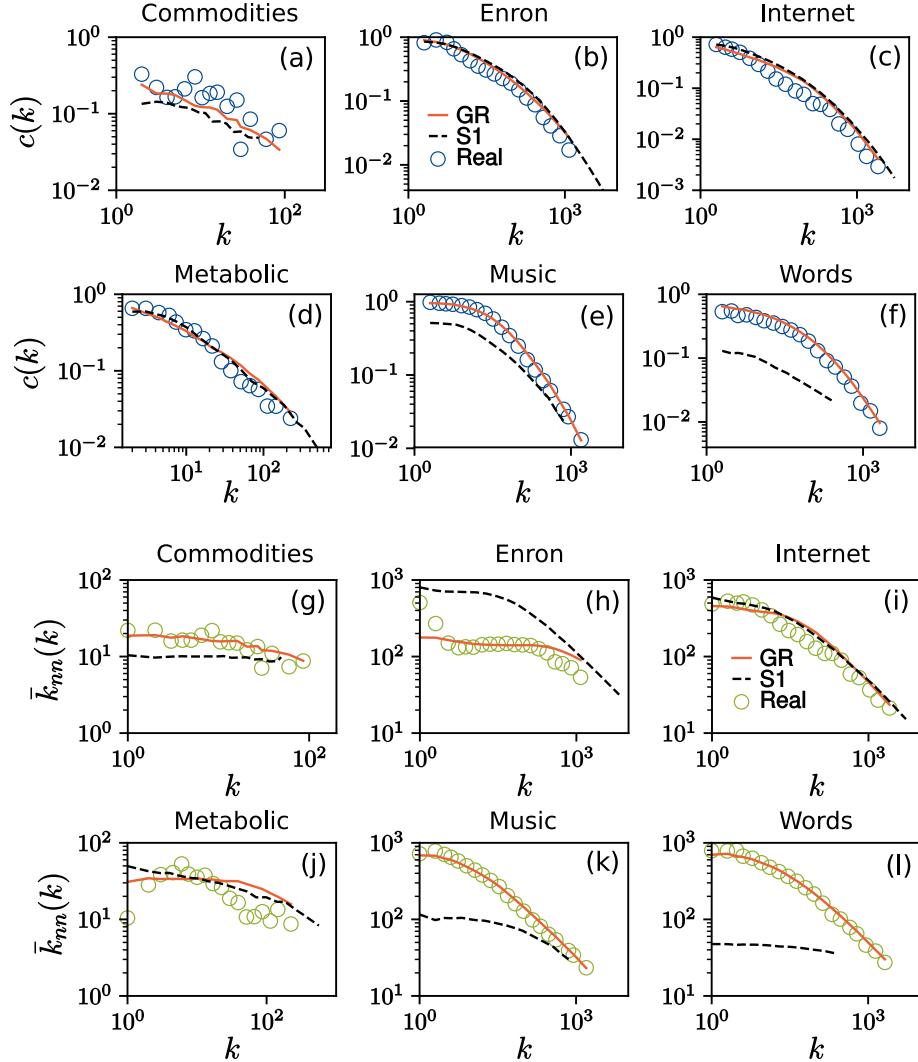


Figure 3.4: (a)-(f) Clustering, $c(k)$, and (g)-(l) average degree of nearest neighbors, $\bar{k}_{nn}(k)$, as a function of the degree, for empirical networks (dots), and networks obtained from the GR (continuous orange line) and \mathbb{S}^1 (black dashed line) models.

usually more important for \mathbb{S}^1 networks due to deviations in the obtained degrees. One exception to the preservation of clustering in GR instances is the Words data set. This empirical network has a β_0 extremely close to the minimal threshold of $\beta_0 = 1$ defined in hidden metric space network models. The β value necessary to ensure that the GR network has the same level of clustering as the empirical one cannot be achieved since it would need to be lower than 1. In general, an embedding value of $\beta_0 \simeq 1$ suggests that clustering is due to finite size effects, since $\beta_0 = 1$ marks a transition point of the geometric network models in the thermodynamic limit between a phase with vanishing clustering ($\beta_0 < 1$) and a phase with clustering.

Graphs on the top row of Fig. 3.4 show the clustering spectrum $c(k)$ for empirical networks and networks obtained by the GR and \mathbb{S}^1 models. In all cases, the functional form of $c(k)$ is similar, a decreasing function of k with a broad tail. The clustering spectrum of the GR networks is always very close to the original data, while the \mathbb{S}^1 networks present important departures in some systems, as a result of the lack of preservation of the empirical degrees. This is especially evident for the \mathbb{S}^1 versions of the Music and Words networks, with the clustering spectrum much lower than that of the original data.

On the other hand, the real networks under consideration are generally disassortative, as revealed by the decreasing form of the average degree of nearest neighbors, $\bar{k}_{nn}(k)$ function, Fig. 3.4(g)-(l). Internet, Music and Words show a decay with power law form, while other data sets show milder degree correlations. In all cases, GR networks have $\bar{k}_{nn}(k)$ distributions very similar to the original data, while \mathbb{S}^1 networks exhibit strong deviations, with the exception of the Internet.

3.4.2 Community structure

So far, GR randomized versions of real and synthetic geometric networks seem to be able to preserve topological features beyond the degree distribution, including clustering and the average nearest neighbors degree. However, the GR randomization homogenizes the distribution of nodes in the similarity space, while nodes in real networks are typically heterogeneously distributed, as they are more concentrated in some specific regions [46, 68]. As previously discussed, this denotes the presence of geometric communities of similar nodes [61].

The polar plots in Fig. 3.5 show the representations of the empirical networks embedded in the hyperbolic plane, with coordinates (r, θ) . One can clearly see that the angular coordinates θ are heterogeneously distributed in $[0, 2\pi]$. A different perspective is shown by the corresponding histograms in Fig. 3.5, displaying the probability density function $P(\theta)$ of the similarity coordinate of the nodes for the six empirical networks. The heterogeneity of the angular coordinate can be quantified by performing a Kolmogorov-Smirnov (KS) test between the probability density functions $P(\theta)$ and $P_{GR}(\theta)$. The KS statistic measures the difference between two probability distributions, and it is defined as the maximum difference between the values of the distributions $P(\theta)$ and $P_{GR}(\theta)$. The larger

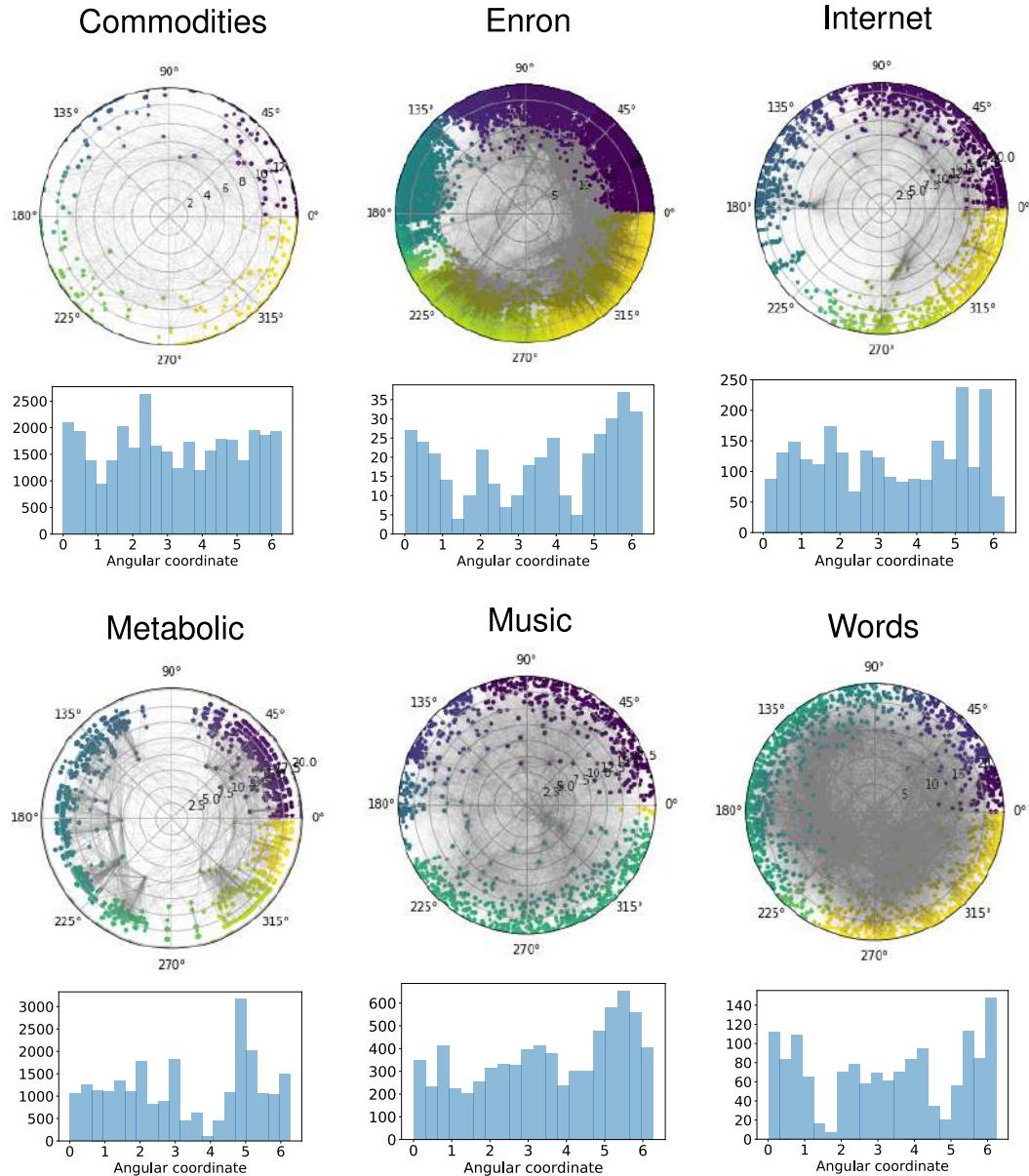


Figure 3.5: Polar plots: Empirical networks embedded in the hyperbolic disk. Distinct communities are indicated by different colors. **Histograms:** Probability distribution of the angular coordinate, $P(\theta)$, of the empirical networks.

the KS score, the more heterogeneous the angular distribution. Thus, it can be used to discard the null hypothesis that the empirical $P(\theta)$ and synthetic $P_{GR}(\theta)$ samples (with uniform distribution by construction) present the same angular distribution. The KS distance D_{KS} for empirical networks under consideration is reported in Table 3.1. One can see that the null hypothesis is strongly rejected for all real networks.

As explained in Chapter 2.3, soft communities in the geometric domain can be detected using geometric methods. We use the definition of soft communities given in [61], where they are defined as a group of nodes in similarity space separated from the rest by two angular gaps that exceed a certain critical value, $\Delta\theta_c$. Here, the critical gap $\Delta\theta_c$ is calculated as the expected value of the largest gap between two nodes when the angular coordinates are distributed uniformly at random: $\Delta\theta_c \simeq 2\pi \ln(N)/N$. In the polar plots displayed in Fig. 3.5, we highlight the soft community deterministic partition detected by the Geometric CGM (G-CGM) in the real networks using different colors.

Next, we compare the community structure of the real networks with their randomized counterparts. To quantify their topological community structure, we apply the widely used Louvain method [69], aimed at maximizing the modularity $Q \in [-1, 1]$, that compares the fraction of links inside communities with the expected fraction for a random distribution of edges with the same node degree distribution as the given network. Interestingly, Fig. 3.6(a) shows that in real networks, albeit the Louvain method identifies topological communities with higher modularity, the soft communities discovered by the G-CGM display large Q values, in some cases (e.g. Metabolic or Music data sets) comparable to the modularities given by the purely topological LM.

This picture is completely different for GR networks, reported in Fig. 3.6(b). GR networks show strong community organization at the topological level, resulting in large values of Q as measured by the Louvain method, which is induced by structural constraints imposed by the geometric models [70]. However, as expected, the critical gap does not detect soft communities, as demonstrated by the non-significant values of the modularity, compatible with zero, over different realizations of the randomization process.

Finally, we study in more detail the relationship between soft communities and topological ones by comparing the partition obtained by the Louvain method with the partition generated by the critical gap. The overlap between the two partitions can be quantified by the normalized mutual information (NMI) [71], which gives a value of 0 when there is no overlap between the partitions (no mutual information) and 1 otherwise. Fig. 3.6(c) shows that the overlap between geometric and topological communities is quite large for real networks, specially for Metabolic and Internet data sets, meaning that communities identified by purely (deterministic) geometric methods are meaningful, though subject to the degree of congruency of the real network with the hidden metric space. On the contrary, Fig. 3.6(c) shows that the overlap between soft and topologi-

cal communities in GR networks is very low due to the complete randomization of the angular coordinate operated by GR.

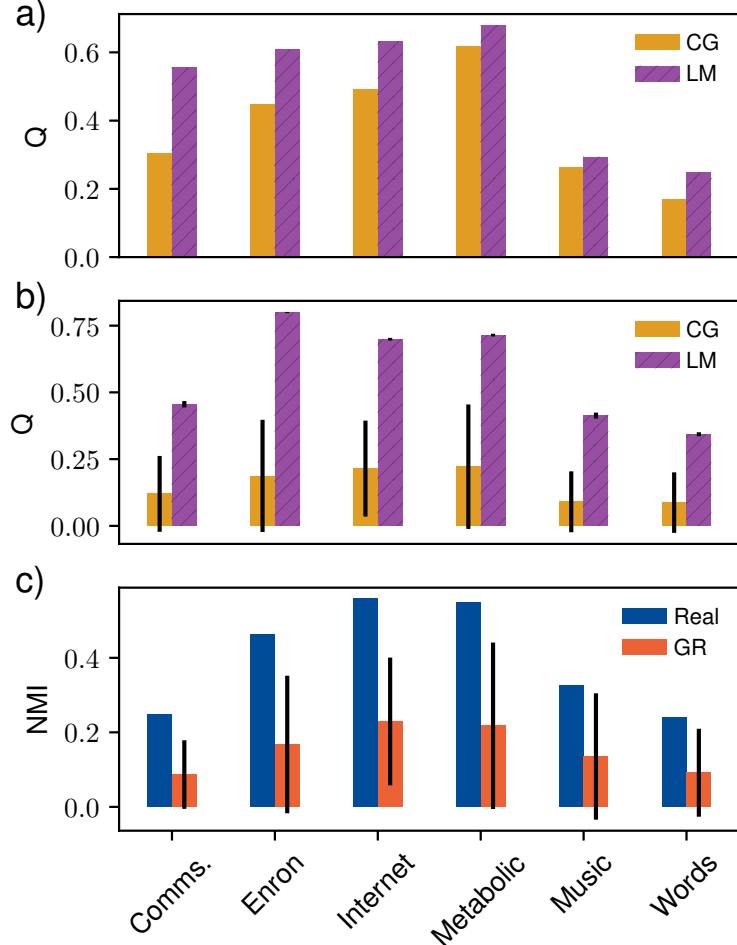


Figure 3.6: (a)-(b) **Modularity** Q as detected by the Lovain method (purple) and the critical gap (yellow), for real networks in plot (a) and GR replicas in plot (b). Error bars in plot (b) are obtained using 10 realizations of the GR model. (c) **Normalized mutual information** between the partition detected by the Louvain and the critical gap methods, for empirical (blue) and GR (red) networks. Error bars are obtained using 10 realizations of the GR model.

3.5 Discussion

In this chapter we have demonstrated how to build a rewiring mechanism based in the S^1 model that allows to replicate geometric networks preserving exactly key topological features such as the original degree sequence, and at the same time completely homogenizing their angular distribution in order to isolate any effects coming from the presence of angular heterogeneities and thus correlations in similarity values.

The rewiring process preserving degrees in the geometric randomization of real networks gives an alternative to their replication using directly the popularity-similarity model as a topology generator. The GR offers the advantage of avoiding the delicate task of estimating the hidden degree distribution, and it can be especially useful in problems responsive to fluctuations of the degree cutoff, like the behavior of some dynamical processes including epidemic spreading processes [72, 73].

As a model, the GR depends on a single parameter controlling the level of clustering in the resulting networks, so that the clustering coefficient of real networks can be chosen to be replicated or not. Interestingly, the discrepancies between hidden and observed degrees in embedded networks have an effect on the clustering level achieved by the GR. In particular, the parameter value suggested by the embedding of the original data is, in general, not far but not totally coincident with the needed value for replicating the clustering coefficient of the original network. Our results also indicate that, in some networks, degree-degree correlations can only be replicated by the geometric network models if the observed degrees are preserved.

As a null model, the GR can be used to investigate the relevance of geometric communities in real networks. Taken together, our results indicate that in the real networks analyzed here geometric communities are always meaningful. At the same time, topological communities, like those detected in GR networks, are not always reliable and can be a result of constraints induced by the underlying geometric architecture. The fact that an underlying geometric organization imposes structural constraints on complex networks, which are strong enough for recreating detectable topological communities even in the absence of geometric ones, is an interesting subject by itself and will be investigated in future work.

The practice of testing hypotheses against a properly specified control case is at the heart of the scientific method. In fact, it is essential to carefully elucidate the structural details that influence most heavily the functioning of real-world networks. Nonetheless, null models for networks that incorporate geometric information are scarce and mainly focused on spatial networks [74–76]. The relevance of this contribution is to fill this gap, facilitating a novel geometric approach capable of unravelling the impact of distinctive organizational patterns found in real-world networks. We further demonstrate the applicability and usage of the geometric randomization model in Chapter 4 and prove how

it can strengthen our knowledge about the intimate relationship between the dynamics and structure of complex networks. In particular, in the next chapter the GR is used to address another distinctive feature of geometric maps of real-world networks: hierarchy.

4 Geometric detection of hierarchical backbones

*“Authority systems must be based on people arranged in a hierarchy.
Thus the critical question in determining control is, Who is over whom?”*

STANLEY MILGRAM

This chapter was – with some small changes – published in “Physical Review Research” by the American Physical Society under the title “Geometric detection of hierarchical backbones in real networks” [77]. A preprint version is also available at [78].

In Chapter 3 we showed that an enhanced comprehension of community structure in real-world networks can be achieved from the study of their topology when coupled to the corresponding underlying geometry. In this chapter we advance our analysis of network’s structure by focusing on the hierarchical nature of complex systems. We also introduce a dynamic process that highlights the relevance of hierarchies in shaping cooperation in structured populations.

4.1 Hierarchical order in networked systems

Many real systems display a hierarchical organization [79], colloquially known as pecking order, where higher status members dominate over lower-graded ones, according to a certain measure of power, wealth, importance or influence. Examples are ubiquitous in living systems, including molecular regulators governing gene expression [80], animal communities of eusocial insects [81], dominant-subordinate relationships in mammals [82], and different structures—companies, political parties, courts, military, organized religion, etc.—in human society [83]. Additionally, hierarchical organization can be found in nonliving systems such as computer generated imagery (CGI) [84], grammatical theory of language [85], or the structure of a musical composition [86]. Hierarchies are, thus, ubiquitous, and shape more easily controllable structures [87] that can emerge as the result of opposing forces, such as competition between individuals [88] or a combination of cooperation and imitation strategies [89].

Complex networked systems, however, present important challenges when it comes to the detection of hierarchies due to the lack of a unique and unambiguous stratification scheme, possibly including nestedness or layered structure. This is caused by the small world property and the presence of a large number of cycles of different lengths in their topologies, in particular clustering [90], so that networks’ organization deviates

strongly from treelike. In directed networks, link directionality can be exploited to ease the problem and hierarchical order can be detected using, for instance, penalty function minimization strategies [91–94]. Nevertheless, most frequently the only meaningful or available representation of a complex system is an undirected graph. Within this architecture, a hierarchy is typically defined as a ranking where the status of a node becomes determined by some heterogeneous topological property, for instance degree [95] or some other centrality measure [87,96]. However, other attributes shape as well the hierarchical structure of real networks, like similarity between nodes [25]. Clearly, the control exerted by a higher status node over a lower status one will be more effective when there exists closeness or affinity between them. Conversely, the strength of hierarchical relations gets dissolved as nodes loose their proximity and become dissimilar.

In this chapter, we integrate degree rank, or popularity, and similarity between nodes in an enriched interpretation of hierarchy, valid for real networks with undirected links. For this purpose, we capture network architecture using the geometric network maps and models explained in Chapter 2, which naturally encode popularity and similarity attributes of nodes as coordinates in an underlying metric space. Exploiting the geometric approach, we are able to characterize the individual contribution of each node and each link to the hierarchical structure of a network. Moreover, we exploit the great heterogeneity found in the hierarchy load of links to propose a filtering method, the *similarity filter*, that offers a practical procedure to extract a hierarchical similarity backbone (HSB) of a network. The obtained similarity backbones contain the links that dominate the hierarchical organization of the network.

Hierarchies are tightly related to power structures and, consequently, have been a historical source of developing social dilemmas. For this reason, in this chapter we lastly explore the role of hierarchical similarity backbones in a dynamical process with intriguing behavior in networked systems: the evolution of cooperation. Evolutionary dynamical processes have been argued for years to be sensitive to the hierarchical organization of complex architectures [97, 98]. In particular, despite the evolution of cooperation has been studied in different fields [99–102], it is not well understood yet in scale-free networks [103–107] or in real networks [108], which additionally present high level of clustering and finite size effects. In fact, only recently some mechanisms have been proposed to aid cooperation in real networks [109], based precisely in the geometric approach followed in this contribution. Implementing an evolutionary prisoner’s dilemma game [110–113] on the real networks under study, we illustrate the relevance of hierarchical similarity backbones, discovering that HSBs tend to achieve final states of greater cooperation as compared to the corresponding original real networks for equivalent initial conditions. As we shall see in the following, similarity backbones are capable of capturing the hierarchical links of real networks that better support cooperative behavior in evolutionary dynamics.

4.2 Definition of hierarchy load of links and nodes

We base our definitions of hierarchy on geometric maps of real networks obtained from the isomorphic geometric network models, $\mathbb{S}^1 / \mathbb{H}^2$, discussed in Chapter 2, and additionally use the \mathbb{S}^1 and GR as null models. We characterize the local contribution of a link or a node to the hierarchical structure of a network by measuring its *hierarchy load*, which depends on status, similarity, and the reference provided by a null model to discount the effects of random fluctuations. We use the \mathbb{S}^1 as a reference since it is the maximum entropy model for geometric networks with heterogeneous degrees [44]. It provides expectations for the distribution of hierarchy loads in a pure random assignment of angular positions of nodes given a degree distribution and a level of clustering, so that anomalous fluctuations can be detected.

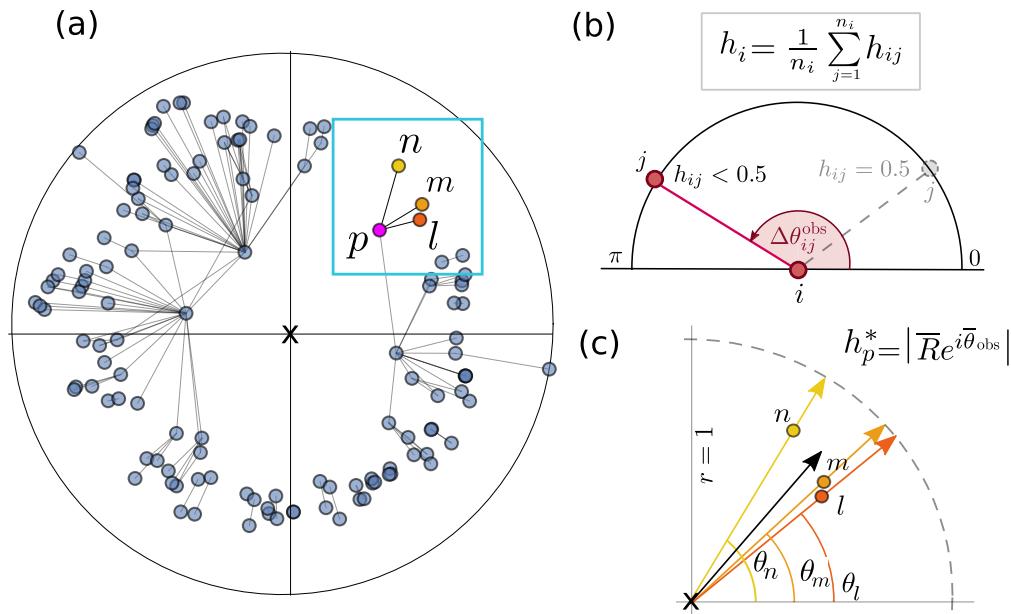


Figure 4.1: Hierarchy load in geometric networks. (a) Illustration of a geometric network in the hyperbolic disc. Node p and its subordinates l, m, n are highlighted inside a blue square. Notice node p has a fourth neighbor which is not a subordinate since it lies at a smaller radial position from the center of the disc and thus has higher degree than node p . (b) Hierarchy load of a link, h_{ij} , between a generic node i and a lower-status neighbor j located at angular distance $\Delta\theta_{ij}^{\text{obs}}$. Notice in this example $h_{ij} < 0.5$ as j is located at an angular distance greater than the expected from the null model, which appears depicted in grey. The hierarchy load of node i , h_i , is obtained from averaging all link hierarchy loads corresponding to this node. (c) Hierarchy load of node p from panel A as the circular mean of the angular coordinates of lower-status neighbors l, m, n . Mean resultant vector is depicted by a black arrow of $|\bar{R}| < 1$. Vectors added in the calculation of h_p^* (see Eq.4.6) appear as arrows in yellow, orange and red inside a circle of unit radius.

Given a geometric network embedding, where nodes have coordinates $\{\kappa_i^{\text{obs}}, \theta_i^{\text{obs}}\}$, we consider that popularity, corresponding to the hidden degree or, equivalently, to the radial position in the hyperbolic plane, is a measure of status. This means that a node i has a lower-status neighbor j when $\kappa_j^{\text{obs}} < \kappa_i^{\text{obs}}$. Similarity between the two nodes is represented by their angular separation in the network map, $\Delta\theta_{ij}^{\text{obs}} = \min(|\theta_i^{\text{obs}} - \theta_j^{\text{obs}}|, 2\pi - |\theta_i^{\text{obs}} - \theta_j^{\text{obs}}|)$.

First, we define the hierarchy load h_{ij} of a link between node i and its lower-status neighbor j as the probability of obtaining a similarity distance between them in the \mathbb{S}^1 model greater than the one observed in the map

$$h_{ij} = P(\Delta\theta_{ij} > \Delta\theta_{ij}^{\text{obs}}). \quad (4.1)$$

The rationale behind this definition is that a high probability in the null model for the angular separation to be larger than the observed, indicates that in reality the nodes are closer in similarity space than expected, hence signaling a highly hierarchical connection. Being a probability, h_{ij} is bounded in the interval $[0,1]$, while the angular separation between nodes $\Delta\theta_{ij}^{\text{obs}} \in [0, \pi]$. Furthermore, Eq. (4.1) can be computed analytically giving an expression depending on the coordinates of the nodes in the embedding as we proceed to demonstrate.

To derive a closed expression for the hierarchy load of an observed link, h_{ij} , given the angular distance between the two nodes observed in the embedding, $\Delta\theta_{ij}^{\text{obs}}$, we first need to write down an expression for the probability distribution of the angular separation between two nodes in the \mathbb{S}^1 model *conditioned to the fact that they are connected*. Using Bayes' rule, we see that

$$p(\Delta\theta_{ij}|a_{ij} = 1) = \frac{p(a_{ij} = 1|\Delta\theta_{ij}) p(\Delta\theta_{ij})}{p(a_{ij} = 1)}, \quad (4.2)$$

where a_{ij} is the adjacency matrix element corresponding to the two nodes. In the above expression, $p(a_{ij} = 1|\Delta\theta_{ij})$ is the connection probability between nodes in the \mathbb{S}^1 model, Eq. (2.2), while the distribution of angular distances is simply $p(\Delta\theta_{ij}) = 1/\pi$, given that angular coordinates are homogeneously distributed in the \mathbb{S}^1 model. The denominator can be obtained by direct integration, $p(a_{ij} = 1) = \int_0^\pi p(a_{ij} = 1|\Delta\theta_{ij}) p(\Delta\theta_{ij}) d\Delta\theta_{ij}$. Using the definition of h_{ij} , we obtain

$$\begin{aligned}
 h_{ij} &= P(\Delta\theta_{ij} > \Delta\theta_{ij}^{\text{obs}}) \\
 &= 1 - \int_0^{\Delta\theta_{ij}^{\text{obs}}} p(\Delta\theta_{ij} | a_{ij} = 1) d\Delta\theta_{ij} \\
 &= 1 - \frac{\int_0^{\Delta\theta_{ij}^{\text{obs}}} p(a_{ij} = 1 | \Delta\theta_{ij}) p(\Delta\theta_{ij}) d\Delta\theta_{ij}}{\int_0^\pi p(a_{ij} = 1 | \Delta\theta_{ij}) p(\Delta\theta_{ij}) d\Delta\theta_{ij}} \\
 &= 1 - \frac{\Delta\theta_{ij}^{\text{obs}} {}_2F_1\left(1, \frac{1}{\beta}; 1 + \frac{1}{\beta}; -\left(\frac{R\Delta\theta_{ij}^{\text{obs}}}{\mu\kappa_i^{\text{obs}}\kappa_j^{\text{obs}}}\right)^\beta\right)}{\pi {}_2F_1\left(1, \frac{1}{\beta}; 1 + \frac{1}{\beta}; -\left(\frac{R\pi}{\mu\kappa_i^{\text{obs}}\kappa_j^{\text{obs}}}\right)^\beta\right)}, \tag{4.3}
 \end{aligned}$$

where ${}_2F_1(a, b; c; z)$ is the hypergeometric function. Equation (4.3) yields the hierarchy load of a given link in terms of the angular separation and product of hidden degrees of the nodes at both ends of the edge, as well as of the global parameters R , μ , and β .

Importantly, let us also show that the expected value of h_{ij} for any link in a network generated by the \mathbb{S}^1 is $\langle h_{ij} \rangle_{\mathbb{S}^1} = 1/2$. In this case, we only need to notice that, the angular separation between the nodes at the ends of the edge in the resulting network, $\Delta\theta_{ij}^{\text{obs}}$, is itself a random variable with distribution $\rho(\Delta\theta_{ij}^{\text{obs}})$ given by Eq. (4.2), that is, $\rho(\Delta\theta_{ij}^{\text{obs}}) = p(\Delta\theta_{ij}^{\text{obs}} | a_{ij} = 1)$ and, therefore, the expected value of h_{ij} is

$$\begin{aligned}
 \langle h_{ij} \rangle_{\mathbb{S}^1} &= \left\langle P(\Delta\theta_{ij} > \Delta\theta_{ij}^{\text{obs}}) \right\rangle_{\mathbb{S}^1} \\
 &= \int_0^\pi \rho(\Delta\theta_{ij}^{\text{obs}}) P(\Delta\theta_{ij} > \Delta\theta_{ij}^{\text{obs}}) d\Delta\theta_{ij}^{\text{obs}} \\
 &= \int_0^\pi \rho(\Delta\theta_{ij}^{\text{obs}}) \int_{\Delta\theta_{ij}^{\text{obs}}}^\pi \rho(\Delta\theta_{ij}) d\Delta\theta_{ij} d\Delta\theta_{ij}^{\text{obs}} = \frac{1}{2}. \tag{4.4}
 \end{aligned}$$

In the last step, we have used the fact that $\rho(z)$ is normalized, $\int_0^\pi \rho(z) dz = 1$.

In a synthetic network generated with the \mathbb{S}^1 model, the expected value of h_{ij} for any $\{i, j\}$ is $h_{ij} = 1/2$ because in that case the observed angular distance $\Delta\theta_{ij}^{\text{obs}}$ is generated by the model and, hence, it is a random variable distributed according to the same distribution as $\Delta\theta_{ij}$. In other words, Eq. (4.1) reduces to the probability for two equally distributed variables a and b to fulfill $a > b$. As a consequence, Eq. (4.1) has the clear advantage of being size independent, in the sense that the value $h_{ij} = 1/2$ defines the

\mathbb{S}^1 model hierarchy baseline for any N , therefore allowing us to compare the hierarchy structure of networks of different size.

Link hierarchy loads inform how substantial is the contribution of a link towards the hierarchy by comparison with the above mentioned reference level ($h_{ij} = 1/2$) provided by the \mathbb{S}^1 null model. Accordingly, when the link hierarchy load is higher than the reference, $h_{ij} > 1/2$, nodes i and j are closer than expected in angular distance, and so they are more similar and their relationship more hierarchical. In contrast, when the probability is low, $h_{ij} < 1/2$, i and j are more dissimilar than expected by the null model, meaning that they lie farther away in the angular space and, thus, their relationship is less hierarchical. In fact, $h_{ij} = 0$ if $\Delta\theta_{ij}^{\text{obs}} = \pi$, while $h_{ij} = 1$ if $\Delta\theta_{ij}^{\text{obs}} = 0$.

Finally, within the same framework we also define a measure of hierarchy load for nodes as

$$h_i = \frac{1}{n_i} \sum_{j=1}^{n_i} h_{ij}, \quad (4.5)$$

where the sum runs over the n_i lower-status neighbours j of node i satisfying $\kappa_j < \kappa_i$.

4.2.1 Hierarchy load of nodes in terms of angular concentration

Alternatively, one can also measure the hierarchy load of a node i as the angular concentration of its n_i lower-status neighbors by computing the *circular mean* of their angles $\{\theta_j\}_{j=1,\dots,n_i}$. This method is different but very similar to the node hierarchy measure used in [46] for the analysis of the World Trade Web. When lower-status neighbors are concentrated in a narrow angular sector it means that node i has a tendency to establish links with more similar nodes, hence node i is contributing towards a more hierarchical structure and thus carrying a higher hierarchy load h_i . Conversely, when the lower-status neighbors of i are distributed in a very broad angular sector, the hierarchy load of node i is low, indicating that it is able to establish links with more dissimilar lower-status nodes, thus supporting a flatter organization.

The hierarchy load of a node in terms of angular concentration of its n lower-status neighbors is computed as the length of the mean resultant vector,

$$h^* = \left| \frac{1}{n} \sum_{j=1}^n e^{i\theta_j^{\text{obs}}} \right| = \left| \overline{R} e^{i\bar{\theta}^{\text{obs}}} \right| = \overline{R}, \quad (4.6)$$

see Fig. 1(c). The modulus of the circular mean vector $\overline{R} \in [0, 1]$ is a good proxy of angular confinement since it is 0 for angles pointing in opposite directions and it becomes 1 when the angles are totally aligned. The measure is simple enough so that it generalizes well to networks of very different domains as long as they admit a geometric interpretation. It is worth mentioning that, since the average angular separation between nodes decreases with the network size N , this quantity increases with network size. This should be taken into account when comparing h^* measurements among networks of different size.

4.3 Analysis of the hierarchy spectrum

4.3.1 Hierarchy load vs geometric communities in synthetic networks

To understand how the global distribution of angles, and, in particular, the existence of geometric communities, could affect the spectrum of hierarchy loads, we consider synthetic networks with controllable angular concentration of nodes generated using the soft communities in similarity space (SCSS) model (see section 2.3.2). The SCSS model enables to create scale-free networks with high clustering while controlling for the global heterogeneity of the distribution of angular coordinates. In particular we feed the SCSS with an heterogeneous angular distribution generated following the GPA model (see section 2.3.1), where parameter Λ regulates the level of inhomogeneity such that for $\Lambda \rightarrow \infty$ we recover an homogeneous angular distribution.

We show the spectra of hierarchy loads h_i , Eq. 4.5, and h_i^* , Eq. 4.6, in Fig. 4.2(a) and Fig. 4.2(b) for synthetic networks with very different geometric community strengths ($\Lambda = 0.01$ and $\Lambda = 10.0$). The spectra of hierarchy loads is measured by averaging the node hierarchy loads over degree classes. The two measures display different results, but in both cases the global angular heterogeneity has a minor effect in shaping the hierarchy loads of nodes. Therefore, the spectrum of hierarchy loads of nodes is not merely a measurement of geometric community structure. As expected, h_i spectra are flat and lie around the average hierarchy load $\langle h \rangle = \sum_{i=1}^N h_i/N$ of the whole network. At large Λ values, $\langle h \rangle$ tends to 0.5 because, by construction, SCSS networks recover the \mathbb{S}^1 model in this limit. Heterogeneous angular distributions in the limit of small Λ values reduce the average angular distance between nodes in the network, and as a consequence the average hierarchy load of the network is slightly above 0.5.

This effect is also evident in the spectra of hierarchy loads h_i^* , where $\langle h \rangle^*$ is lower for more homogeneous angular distributions ($\Lambda = 10.00$, $\langle h \rangle^* = 0.89 \pm 0.10$), as compared with networks with more heterogeneous distributions that present higher values ($\Lambda = 0.01$, $\langle h \rangle^* = 0.95 \pm 0.06$). In this case, the average hierarchy level of the synthetic networks is rather high ($\langle h \rangle^* \approx 0.9$), basically due to sustained large values of hierarchy load for a wide range of node degrees.

4.3.2 Hierarchy spectrum of real networks

We measured and compared the hierarchy spectrum of several real networks from different domains: the social network of Enron employees connected through email (Enron Social), the Internet at the AS level (Internet), the one-mode projection onto metabolites of the human metabolic network at the cell level (Metabolic Human) and the network of chord transitions in western popular music (Music). For detailed information regarding the data sets see Table 4.1 and Appendix A.

Real networks show a variety of profiles, see blue curves in Fig. 2(c)-(f). They also present in all cases an average hierarchy load $\langle h \rangle$ below the reference of 1/2. In particular,

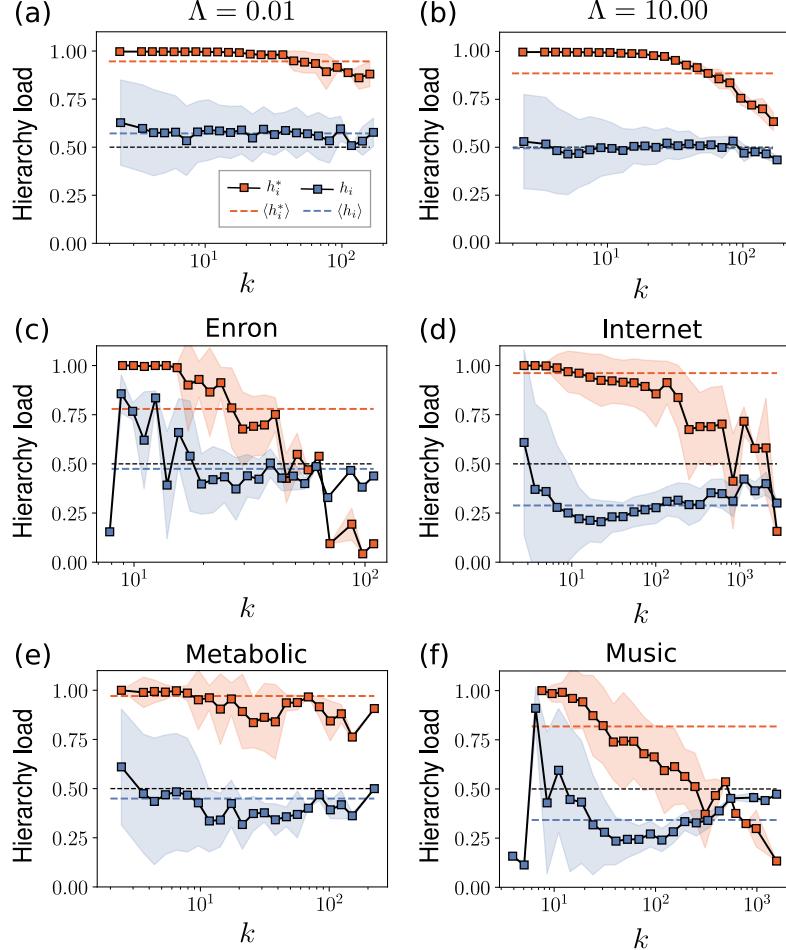


Figure 4.2: Hierarchy load spectra of synthetic and real networks. (a) and (b) Hierarchy load spectra for synthetic SCSS networks of size $N = 1000$ nodes, generated with power-law degree distribution exponent $\gamma = 2.50$, clustering parameter $\beta = 2.50$ and variable attractiveness Λ . Results are averaged over 10 network realizations for each choice of Λ . (c)-(f) Hierarchy load spectra of 4 real networks. In all plots blue curves correspond to h_i in Eq. (4.5) computed from the link hierarchy loads while red curves correspond to h_i^* in Eq. (4.6) as given by the circular mean resultant vector. Dashed blue and dashed red lines indicate the corresponding average hierarchy load values for the whole network. Dashed black lines provide the reference level of the model, placed at hierarchy load $1/2$.

Data set	Nodes	N	E	β	k_{\max}	$\langle k \rangle$	$\langle c \rangle$	$\langle h \rangle$	$\langle h \rangle^*$
Enron S.	Employees	182	2097	1.99	109	23.04	0.50	0.47 ± 0.19	0.77 ± 0.25
Internet	A.S.	23752	58416	1.91	2778	4.92	0.61	0.29 ± 0.20	0.96 ± 0.10
Metabolic H.	Metabolites	1436	4718	2.15	224	6.57	0.54	0.45 ± 0.28	0.97 ± 0.09
Music	Chords	2476	20624	2.30	1566	16.66	0.82	0.34 ± 0.25	0.81 ± 0.23

Table 4.1: Properties of the data sets under consideration: N , size of the network; E , number of edges; parameter β estimated from the embedding of the real network; k_{\max} , highest degree; $\langle k \rangle$, average degree; and $\langle c \rangle$, average clustering coefficient; $\langle h \rangle$, average hierarchy load; $\langle h \rangle^*$, alternative average hierarchy load in terms of angular concentration.

Enron shows the highest average hierarchy load at the node level, $\langle h \rangle = 0.47 \pm 0.19$, followed by Metabolic, Music and, lastly the Internet with $\langle h \rangle = 0.29 \pm 0.20$. Variations in $\langle h \rangle$ across networks conform to their distinct spectra. For instance, whereas Music and Internet networks show noticeable fluctuations in h_i across degree classes, these are milder for Enron and Metabolic networks. In general, however, all networks show a tendency for the lowest degree nodes to have the highest hierarchy loads and for the highest degree nodes, or hubs, to approach $h_i = 1/2$. This last observation may be attributable to great heterogeneity in the hierarchy load of links h_{ij} and the fact that hubs present numerous connections with lower status nodes which are averaged when computing h_i .

Regarding the hierarchy load of nodes in terms of angular concentration, the Internet and Metabolic are significantly more hierarchical ($\langle h \rangle^* \approx 1$) than Music and Enron. The particular trend followed by each profile has its roots in each network's specific degree sequence, whereas angular communities show again minor influence as revealed by the resemblance between the spectra of the 4 real networks with the spectra of their corresponding geometrically randomized counterparts. In the following we discuss this observation in more detail.

4.3.3 Hierarchy spectrum of GR network replicas

As explained in Chapter 3, the randomized counterparts consist in replicas of the real networks where the distribution of angular coordinates is homogenized, thus eliminating geometric community structure, while the rest of properties are preserved and the replica network remains maximally geometric. In Fig. 4.2(c)-(f) we showed the spectrum of hierarchy loads of nodes in terms of angular concentration, h_i^* , for 4 real networks. Here, we provide the same metric for angularly randomized versions of such real networks, obtained using the Geometric Randomization model from Chapter 3. By comparing the spectra in Fig. 4.2(c)-(f) and Fig. 4.3 one can notice that the complete homogenization of the angular coordinates of nodes (and thus the full elimination of geometric communities) does not translate into radical changes in the hierarchy load profiles. Instead, in Fig. 4.3 we observe for each of the randomized networks that the inverse correlation with the

degree of the hierarchy loads is compatible with that of the original network, and that the average hierarchy load $\langle h \rangle^*$ values of the GR networks are not remarkably different to that of the real ones. Moreover, we note that the hierarchy loads of nodes tend to show strong heterogeneity and an inverse correlation with the degree, so that more popular nodes (higher k) contribute more to dilute the hierarchical structure by connecting to less affine lower-popularity neighbors. This trend is remarkably pronounced for the Internet, Enron, and Music, while heterogeneity is less pronounced in Metabolic. In fact, this happens because the Metabolic network presents more modular hubs, less prone to connect with dissimilar nodes due to an unusually marked partition of the similarity space (see Fig. 2.4(b) in Chapter 2).

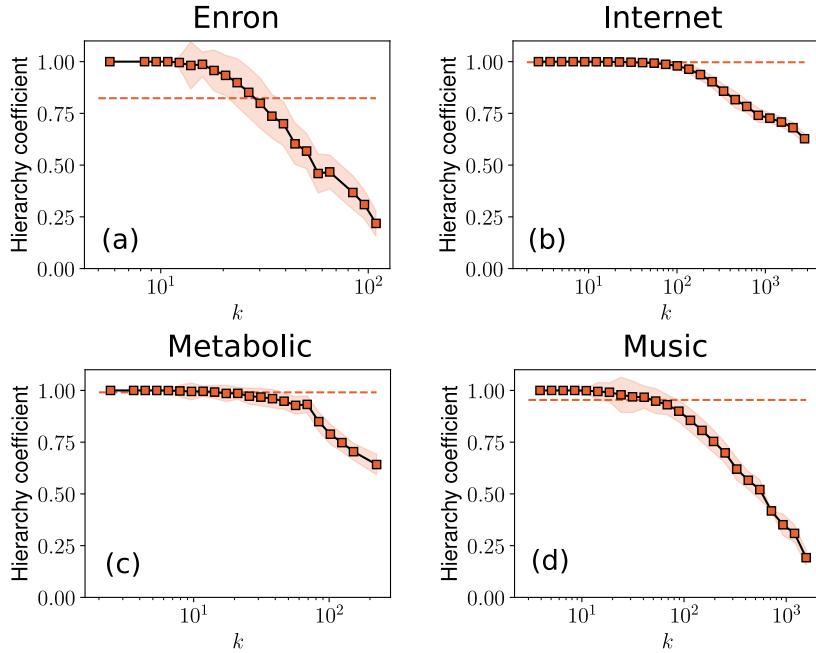


Figure 4.3: Hierarchy load spectrums for Geometric Randomization (GR) model replicas of each of the 4 real networks under study. Each spectrum is the average over 10 independent realizations. Dashed lines indicate the average hierarchy load of the network, obtained as $\langle h \rangle^* = N^{-1} \sum_{i=1}^N h_i^*$, yielding to the following values for each network: $\langle h \rangle_{\text{Enron}}^* = 0.82 \pm 0.19$, $\langle h \rangle_{\text{Internet}}^* = 0.99 \pm 0.02$, $\langle h \rangle_{\text{Metabolic}}^* = 0.99 \pm 0.04$ and $\langle h \rangle_{\text{Music}}^* = 0.95 \pm 0.11$.

4.4 The similarity filter and hierarchical similarity backbones

Strong heterogeneity is found in real networks if we analyze the contribution of lower-popularity neighbors to the hierarchy load of nodes, see Fig. 4.4. The link hierarchy load contributions involving low popularity neighbors are represented in Fig. 4.4 by dots color-coded to the blue end of the scale. We find those dots covering almost all

the area of the plots, thus indicating that these link hierarchy loads span the entire range of h_{ij} , both when the link is shared with another node of low degree or with a node of high degree (right end of the x-axis). This observation reveals that the vast majority of link hierarchy loads in the network are distributed rather heterogeneously. We can take advantage of this heterogeneity to filter out the connections that dominate the hierarchical organization of the network in terms of popularity and similarity, what we name the hierarchical similarity backbone (HSB). A hierarchical similarity backbone contains the links that represent statistically significant contributions with respect to the null hypothesis given by the \mathbb{S}^1 model, that is the only model able to produce maximum entropy ensembles constrained by power-law degree distributions and clustering and without nonstructural degree correlations [44].

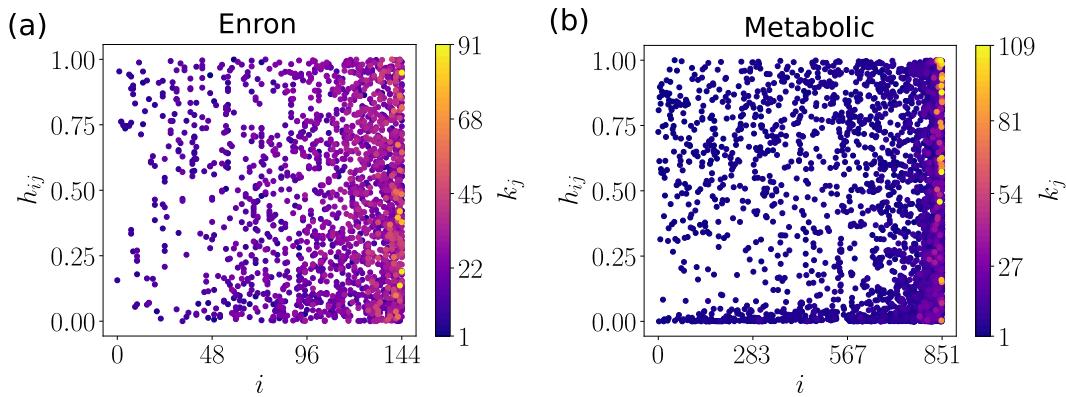


Figure 4.4: Link hierarchy loads of real networks. (a) and (b) Link hierarchy loads of the Enron and Metabolic networks, respectively. Each dot indicates the value of h_{ij} of a link established between node i and one of its lower-status neighbors j with hidden degree κ_j indicated by the color code. Notice node labels in the x-axis are sorted by increasing hidden degree κ_i , so that more popular i nodes appear to the right.

The link hierarchy load in Eq.(4.1) measures the probability under the null hypothesis that the similarity distance between a node and a lower-status neighbor is larger than the observed in the embedding of the network, what is known as p-value in statistical inference. By imposing a significance level α , the links that carry a hierarchy load that are not compatible with the random angular distribution of angles in the \mathbb{S}^1 model, and reject the null hypothesis, can be filtered out. A hierarchical similarity backbone is then obtained by preserving all the links that satisfy the criterion $h_{ij} \geq \alpha$, while discounting the rest. As we increase the significance level $\alpha \in [0, 1]$, the filter progressively focuses on more relevant links to obtain a sequence of nested subgraphs, each with a more strict condition for a link to belong to the HSB of the network, see an illustration in Fig. 4.5(a). Noteworthy, since the similarity filter is applied to the links, nodes of any degree may find a place in a very hierarchical backbone if they are found to have significantly strong hierarchical connections.

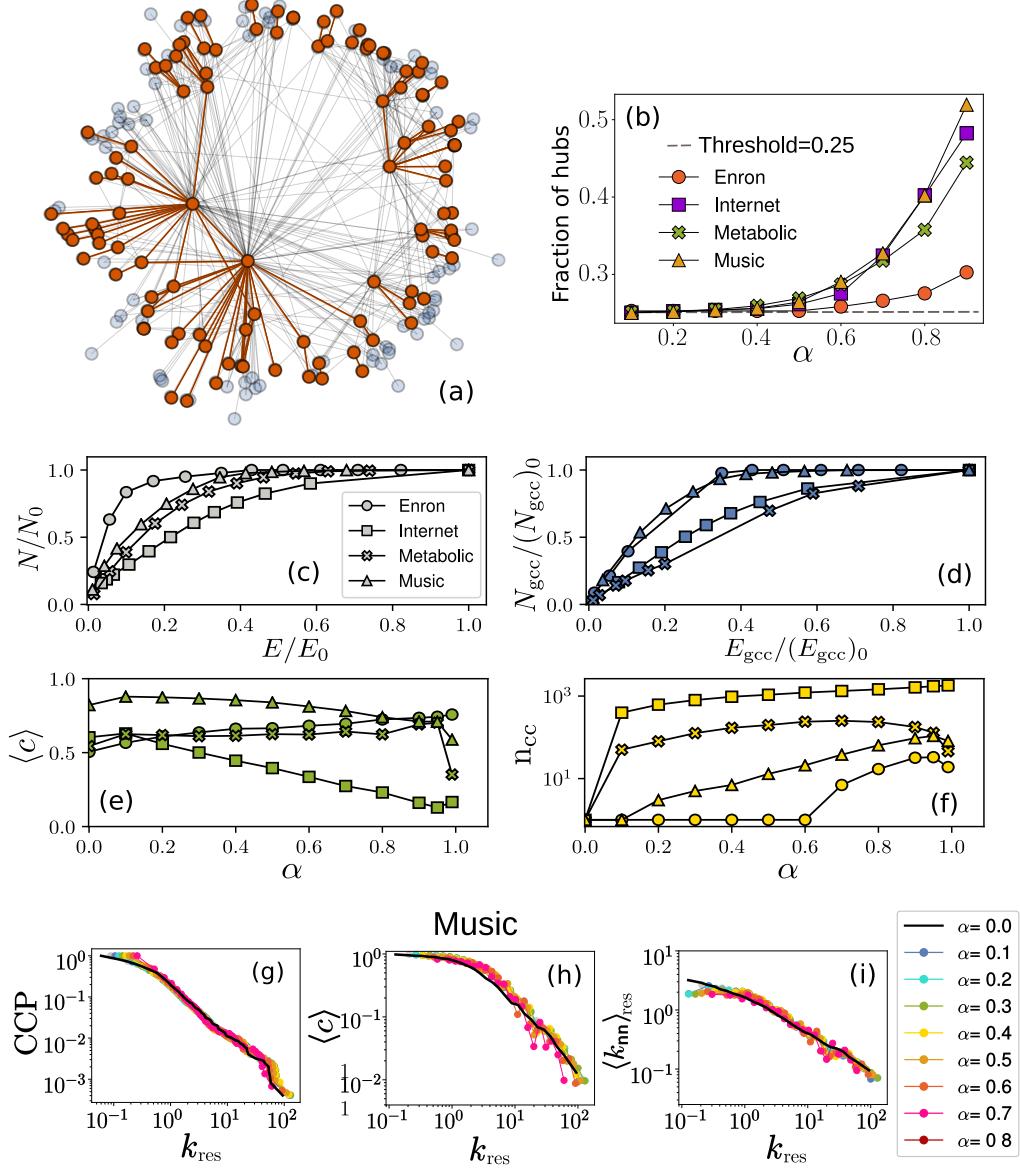


Figure 4.5: Hierarchical similarity backbones. (a) A hierarchical similarity backbone of the World Trade Web [46] (in red) filtered with $\alpha = 0.4$, on top of the complete network (pale blue). (b) Fraction of nodes considered hubs, for threshold value $\tau = 0.25$, found in backbones obtained with increasing α . (c-f) Plots show in order: the relative number of nodes in the backbone, against the relative number of edges in the backbone (subindex 0 refers to the complete network); the relative number of nodes in the backbone giant connected component (gcc), against the relative number of edges in the backbone gcc; mean clustering coefficient of the backbone for increasing α ; number of disconnected components against increasing values of the significance level α (g-i) Topological features of the HSBs of the Music network, obtained for α 's from 0.1 to 0.8. Value $\alpha = 0.0$ corresponds to the original unaltered network, whereas in the most restrictive HSB ($\alpha = 0.8$) 0.60% of the nodes and 14% of the links remain. From left to right: complementary cumulative probability distribution (CCP) of rescaled degrees, $k_{\text{res}} = k/\langle k \rangle$, degree dependent clustering coefficient $\langle c \rangle$ over rescaled-degree classes, normalised average nearest neighbor degree $\langle k_{\text{nn}} \rangle_{\text{res}} = \langle k_{\text{nn}}(k_{\text{res}}) \rangle / \langle k \rangle / \langle k^2 \rangle$.

First, in Fig. 4.5(b) we checked the composition of the HSBs by inspecting the participation of hubs in the HSBs architecture. For this purpose, we sort the nodes in the network from highest degree to lowest and tag as “hubs” all nodes lying within a top slice of the list, delimited by a threshold value τ . For instance, when $\tau = 0.25$, the top 25% of the nodes in the ranked list are considered as hubs. Subsequently, we keep track of the proportion of such high-degree nodes in every hierarchical backbone for increasing α . Figure 4.5(b) demonstrates that, even when considering that a large fraction of the network ($\tau = 0.25$) are nodes of *high* degree, in fact these nodes only represent, at best, half of the backbone composition (see Internet, Metabolic and Music in Fig. 4.5(b) for $\alpha = 0.9$). In Fig. 4.6 we show analogous plots to Fig. 4.5(b) for a wider range of threshold values $\tau \in [0.05 – 0.30]$ providing further evidence that, while more restrictive HSBs become enriched with hubs, the similarity backbones still present an assorted composition in terms of node degrees.

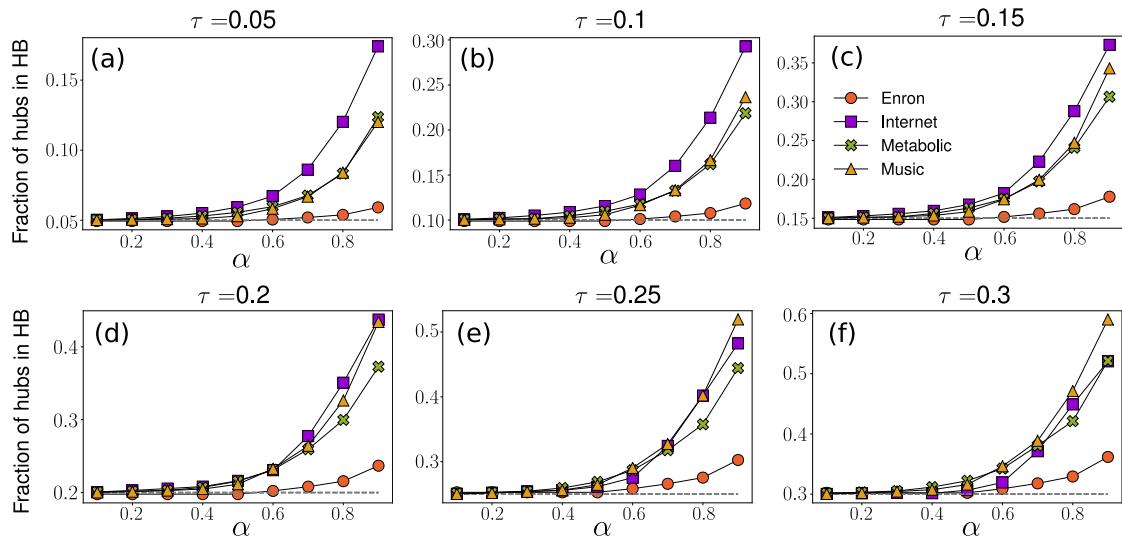


Figure 4.6: Fraction of nodes considered hubs for threshold values in range $\tau = [0.05 – 0.30]$, found in backbones obtained with increasing α for the real networks Enron, Internet, Metabolic and Music.

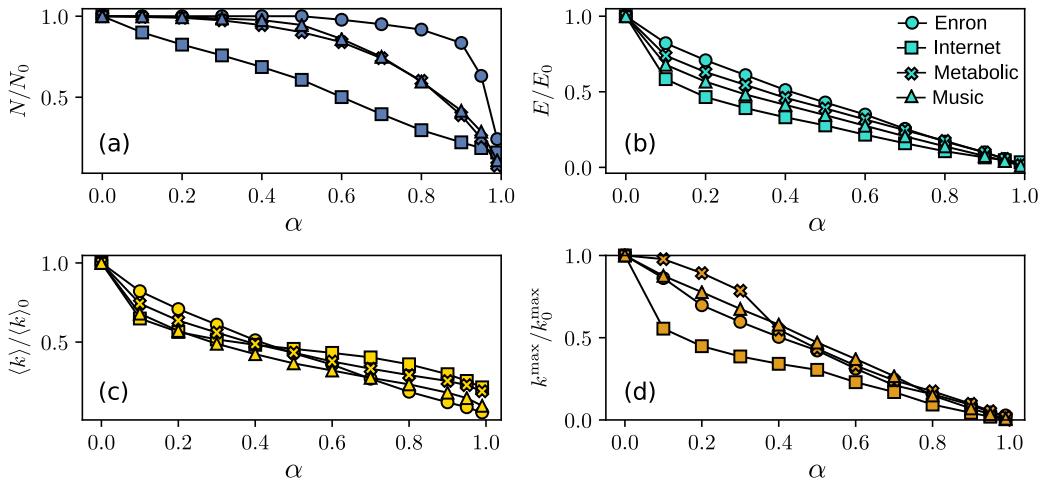


Figure 4.7: Topological properties of hierarchical backbones of the four real networks against increasingly restrictive filtering parameter (α) values. Plots from (a) to (d) show, for increasing values of α : the normalised number of nodes in the backbone; the normalised number of edges in the backbone; normalised average degree of the backbone and the HSB normalised maximum degree.

Secondly, we tested the performance of the similarity filter by examining HSBs topologies obtained with different α , for the four real networks considered in this work. Figure 4.5(c) shows that, for all real networks, low values of α reduce the number of links drastically while most of the nodes are preserved in the backbone. Notice that α increases from right to left in Figs. 4.5(c)-(d). For instance, when filtered with $\alpha = 0.25$, the Internet, Metabolic and Music HSBs contain a proportion of edges that is already less than half of the original, so $E/E_0 < 0.5$. In contrast, the proportion of nodes in the same backbones remains very high, the lowest case being the Internet, but with still 85% of the nodes. The results in Fig. 4.5(d) show similar behaviour for the reduction in nodes and edges of the giant connected components (gcc 's) of the backbones, for the 4 networks under study. Only the decay in number of nodes in the gcc of the backbones tends to be less abrupt than in Fig. 4.5(c) and start sooner, at smaller values of the filtering parameter α . Moreover, we observe in Fig. 4.5(e) that the mean clustering coefficient of the filtered similarity backbones does not have strong fluctuations and varies little with α , with the exception of the Internet which shows a clear decreasing trend. In Fig. 4.7 we cover further topological metrics of the HSBs like the average and maximum normalised degrees and showcase them against the α parameter controlling the filtering procedure. As expected, both $\langle k \rangle / k_0$ and k^{\max} / k_0 decrease with α for all networks.

Finally, we find that topological features —degree distribution, clustering coefficient and average nearest neighbors degree— of the original network are preserved in the subgraphs as we increase α and the backbone is progressively restricted to exceedingly hierarchical links, see Fig. 4.5(g)-(i) and Fig. 4.8. Only for very high values of the

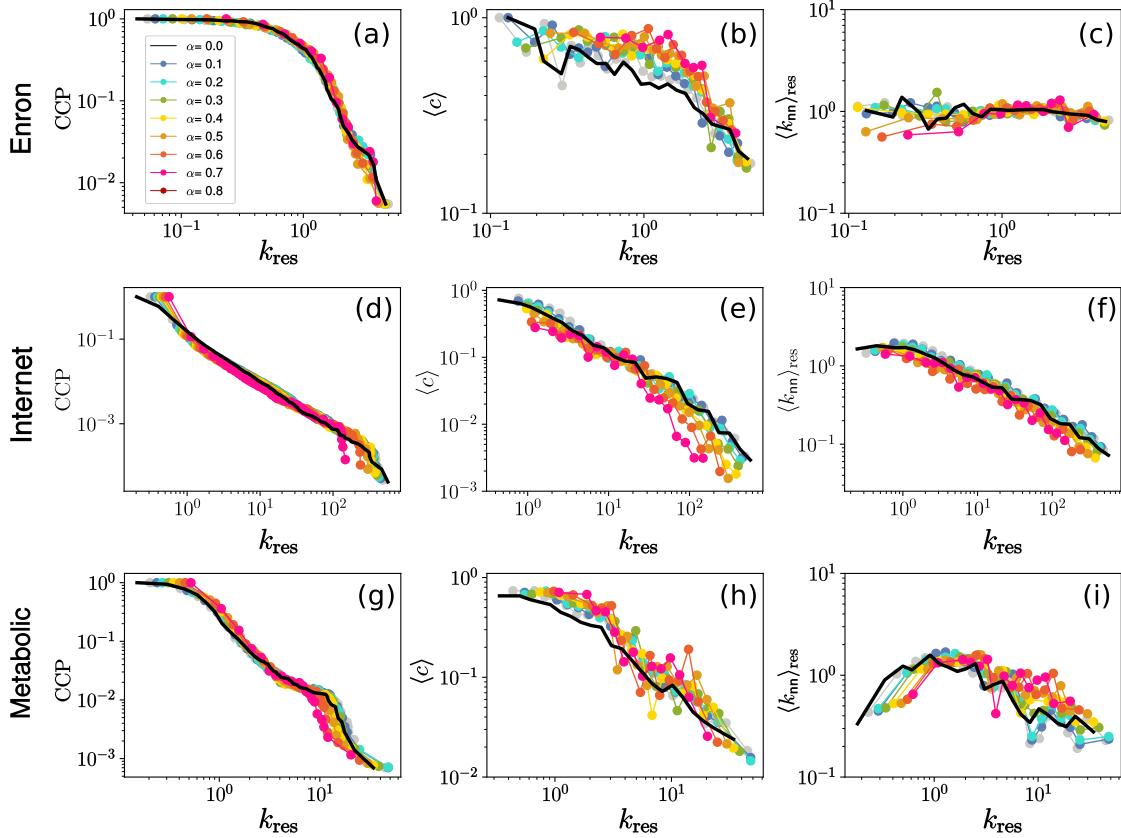


Figure 4.8: Topological features of the hierarchical backbones of (a-c) Enron, (d-f) Internet, (g-i) Metabolic, obtained for $\alpha \in [0.1 - 0.8]$. In each case, value $\alpha = 0.0$ corresponds to the original network. The properties in each row, from left to right, are: complementary cumulative distribution of rescaled degrees ($k_{\text{res}} = k/\langle k \rangle$), degree dependent clustering coefficient over rescaled-degree classes, normalised average nearest neighbor degree $\langle k_{\text{nn}} \rangle_{\text{res}} = \langle k_{\text{nn}}(k_{\text{res}}) \rangle / \langle k \rangle / \langle k^2 \rangle$.

significance level, $\alpha \gtrsim 0.8$, when not only the number of links but also the number of nodes is strongly reduced, the measured topological properties start to deviate from the original curves. This suggests some grade of self-similarity across the sequence of HSBs.

4.5 Evolution of cooperation in HSBs

As an illustration of the importance of HSBs, we study a dynamical process with intriguing behavior in networked systems: the evolution of cooperation. The evolution of cooperation has been extensively studied, yet it is not well understood in scale-free real networks with high levels of clustering and finite size effects. Only recently some mechanisms, precisely following the geometric approach, have been proposed to aid cooperation in real networks [109]. Here, we show that, counterintuitively, our HSBs capture the links of real networks that better support cooperative behavior in evolutionary dynamics.

4.5.1 Evolutionary prisoner's dilemma

Over the last decade research on the prisoner's dilemma game, both analytical and computational, has been extended to various networks such as regular lattices [114], cycles [115], small-worlds [116], scale-free networks [117, 118] and even empirical complex networks [119]. The main conclusion that emerges from previous studies is that graph topology plays a determinant role in the evolutionary prisoner's dilemma game allowing cooperation to be maintained under suitable circumstances [120], which is never possible in a well-mixed population. Despite the large number of studies, surprisingly little is known about how the fundamental properties of the network structure –mean degree, heterogeneity in degree distribution, clustering and assortativity – exactly affect cooperation. What little is known in this direction is almost completely confined to the effect of the average degree [121, 122] of the networks. Here we show that HSBs directly expose the network sub-structures, or backbones, that better support cooperation in real complex networks.

For that, we consider the evolutionary prisoner's dilemma game consisting of two players deciding to cooperate or defect with one another, and gaining a specific reward depending on the four possible outcomes of such interaction. More accurately, the evolutionary prisoner's dilemma game [110], conducted on a network, considers that individual nodes playing with their contacts choose to either cooperate (C) or defect (D) every turn. The choice of strategies of the two interacting agents leads to specific payoffs, summarized by the payoff-matrix

$$A = \begin{array}{c|cc} & C & D \\ \hline C & R & S \\ D & T & P \end{array}. \quad (4.7)$$

That is, if both players cooperate, they both receive the *reward* R for cooperating. If both players defect, they both receive the *punishment* payoff P . Lastly, if one of them defects while the other agent cooperates, the defector receives the *temptation* payoff T , while the cooperator receives the "sucker's" payoff, S . In order for the game to be recognized as a prisoners's dilemma the condition $T > R > P > S$ must apply. Different ordinalities of the parameters define further classes of games [112]. In this chapter, the prisoner's dilemma is in general defined with parameter values: $T = 1.5, R = 1, P = 0, S = -0.5$. Additional parameter values are explored in Fig. 4.11.

Network	Filtering method	Threshold	N_{gcc}	E_{gcc}
Enron S.	none	—	182	2097
Enron S.	HSB	$\alpha^o=0.66$	159	587
Enron S.	HSB	$\alpha^D=0.68$	136	498
Enron S.	RLR	$LR^o=1510$	175.30 ± 1.42	587
Enron S.	RLR	$LR^D=1599$	174.30 ± 2.45	498
Internet	none	—	23748	58414
Internet	HSB	$\alpha^o=0.18$	18536	27374
Internet	HSB	$\alpha^D=0.27$	16691	22970
Internet	RLR	$LR^o=31040$	15901.20 ± 36.06	27374
Internet	RLR	$LR^D=35444$	14172.80 ± 40.27	22970
Metabolic H.	none	—	1436	4718
Metabolic H.	HSB	$\alpha^o=0.21$	1157	2717
Metabolic H.	HSB	$\alpha^D=0.25$	1069	2470
Metabolic H.	RLR	$LR^o=2001$	1195.60 ± 14.21	2717
Metabolic	RLR	$LR^D=2248$	1147.10 ± 13.35	2470
Music	none	—	2476	20624
Music	HSB	$\alpha^o=0.53$	2252	6617
Music	HSB	$\alpha^D=0.59$	2111	5774
Music	RLR	$LR^o=14007$	1905.60 ± 12.04	6617
Music	RLR	$LR^D=14850$	1824.60 ± 21.49	5774

Table 4.2: Backbones used to study evolutionary game dynamics. HSB stands for hierarchical similarity backbone, meaning these backbones are obtained using the similarity filter with specific α . RLR stands for random link removal, thus this method gives random surrogates where a specific number of links have been removed, indicated by LR (links removed). The results in right column of Fig. 4.9 and in Fig. 4.10 are the average of 10 random surrogate realizations. Because of this, and since by construction we fix the number of edges in the random surrogates, only their number of nodes can fluctuate and hence display an error interval.

The game proceeds in successive rounds. After each round, the strategies (C or D) of all nodes are updated synchronously according to the outcome of the imitation dynamics [110, 123, 124] that outlines the evolutionary mechanism. This means, during a single round, each individual node i collects payoffs given by Eq. 4.7 from the interactions with all its neighbors and obtains an accumulated payoff π_i . All players chose then between their old strategy and the strategy of a randomly picked up neighbor j . In this way, node i will adopt j 's strategy with a probability that depends upon the difference between collected payoffs of both nodes ($\pi_i - \pi_j$) as

$$p_{i \rightarrow j} = \frac{1}{1 + e^{(\pi_i - \pi_j)/\tau}} , \quad (4.8)$$

which reflects the popular tendency of individuals to copy more successful neighbors. Such updating rule is a common choice in evolutionary dynamics [125], known as the Fermi rule since it is based in the Fermi distribution from Statistical Mechanics. The variable $(\frac{1}{\tau})$, which in Physics stands for the inverse temperature, can be interpreted here as the intensity of the selection. That is, parameter $(\frac{1}{\tau})$ (which we set to 0.5) controls the noise added to the decision-making process of, otherwise, perfectly rational players. After the simultaneous update of strategy of all nodes, their accumulated payoffs are reset and a new round begins.

4.5.2 Game dynamics in HSBs

We simulated the dynamics on the four real systems analyzed in this chapter and on two different HSBs for each of them (with α values and corresponding sizes in number of nodes and links reported in Table 4.2). The results are provided in the first and second columns of Fig. 4.9. Notice that the similarity backbones are always selected so that they face a considerable reduction in the number of links while their number of nodes does not decay drastically. That is, for a given real network, the HSBs where we run the game dynamics lie along the slope change part of the blue curves in right column of Fig. 4.9, and are identified by blue symbols. We use random surrogates to discern whether the results of the dynamics on HSBs are due to their hierarchical nature. The surrogate backbones are obtained by removing a number of links at random so that they have exactly the same number of links as the corresponding HSBs (see matching fraction of edges between red and blue symbols in Fig. 4.9, right column). As a result, the fraction of remaining nodes in a given HSB may be higher or lower than in the analogous random surrogate. Regardless of the situation, however, the dynamics results stay consistent, see blue curves in left and middle columns in Fig. 4.9. Note that the fraction of nodes in the random surrogates is still very high even when the number of links has been greatly reduced, akin to the case of similarity backbones. This was expected from the reported robustness of scale-free networks to random removals [126, 127]. In Fig. 4.10 one can compare the main topological features of the random surrogates with the original real network for the four data sets under study.

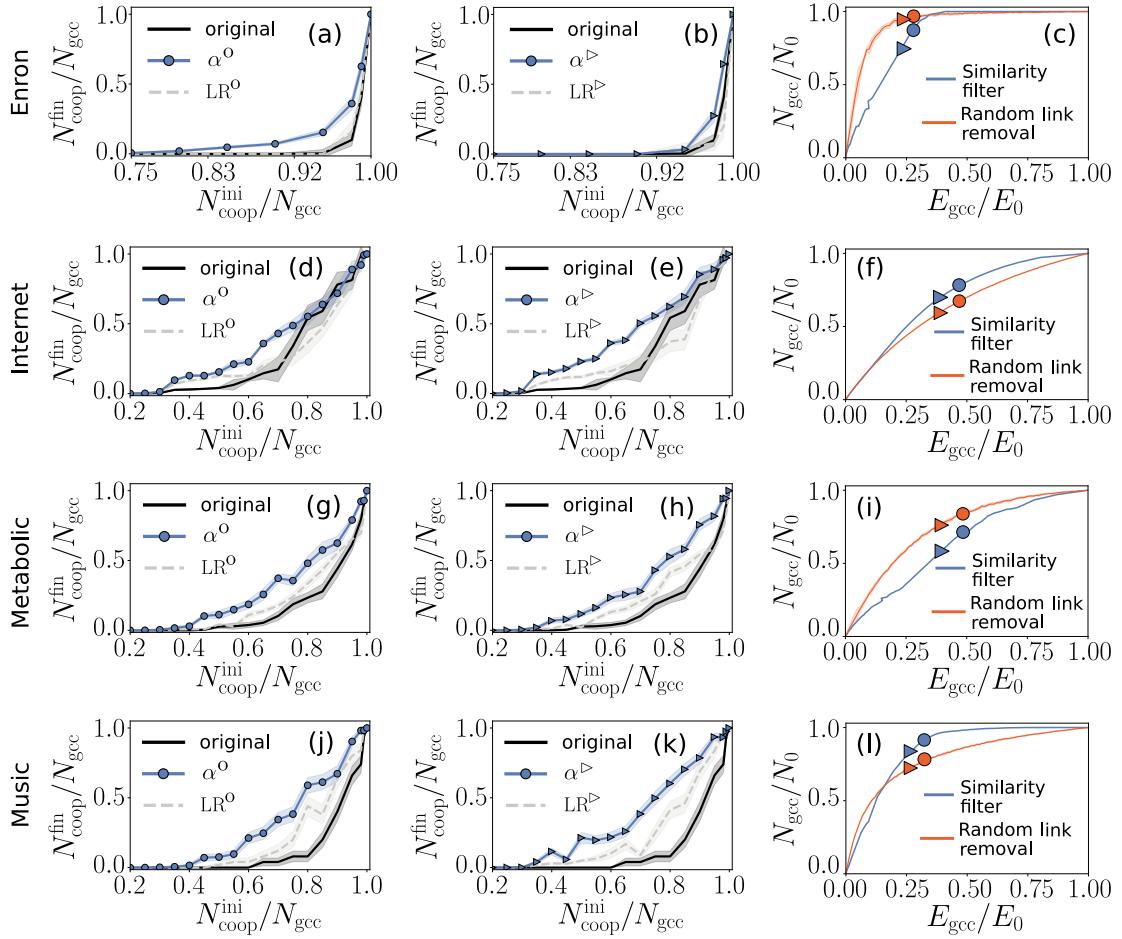


Figure 4.9: Evolutionary game dynamics in HSBs of real networks. Left and middle columns show the fraction of final cooperators against fraction of initial cooperators for similarity backbones of Enron, Internet, Music and Metabolic networks. For every network we show two plots, each with the results of an HSB filtered with either α^o or α^D (see numeric values in Table 4.2). For an HSB (blue curve with symbols), the corresponding random surrogate appears as a grey dashed line together with the original network in black line for reference. Rightmost column shows the relative number of nodes against relative number of edges in the *gcc* of similarity backbones, and their corresponding surrogates, for the four networks under study. Blue lines correspond to backbones obtained using the similarity filter and red lines to surrogates obtained by random link removal. Blue circle and triangle symbols highlight the fraction of nodes and edges of the 2 similarity backbones filtered with α^o and α^D , respectively. The same information is featured by red symbols for the random surrogates.

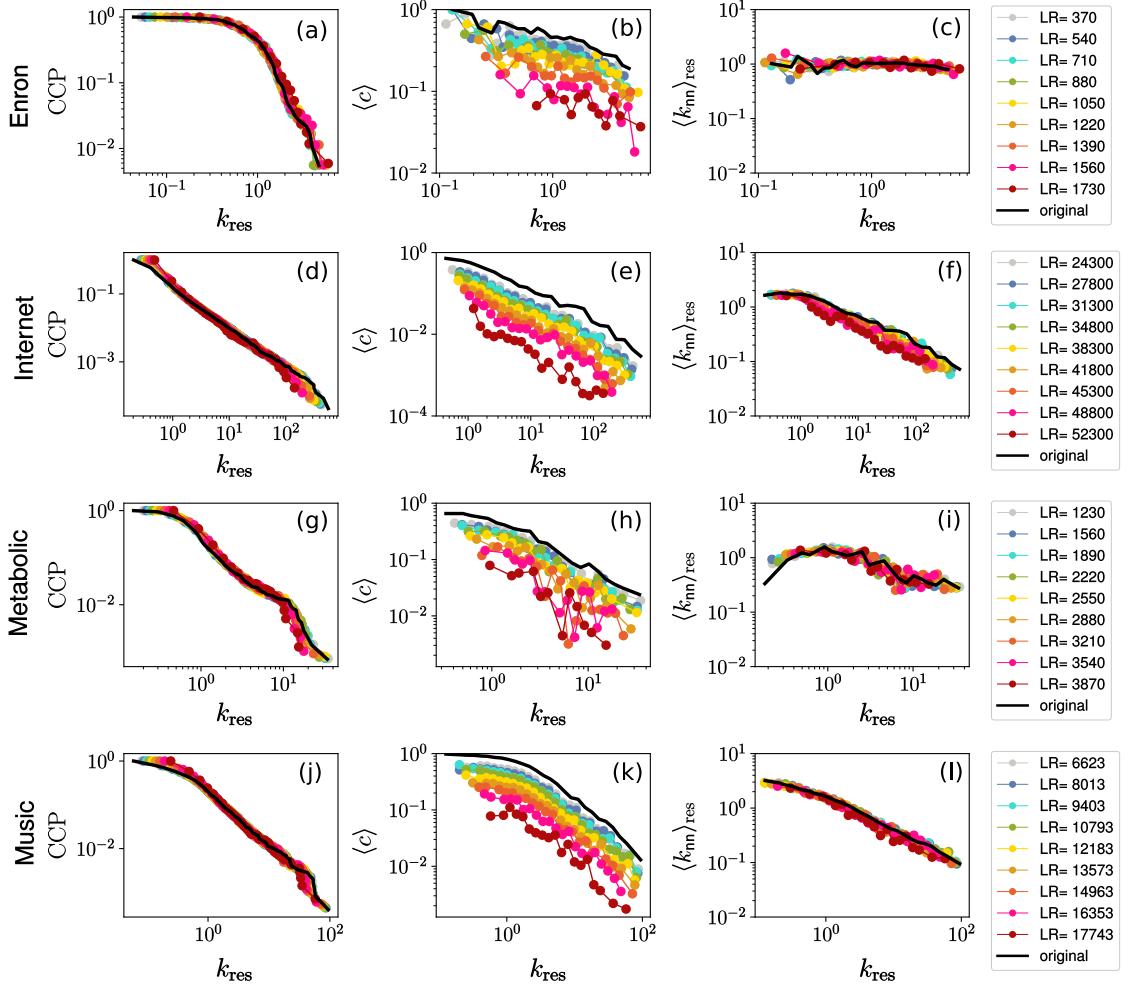


Figure 4.10: Topological features of random surrogates and original network of (a-c) Enron, (d-f) Internet, (g-i) Metabolic (j-l) Music. In each case, value LR indicates the number of links that have been removed at random to produce the surrogate backbone. The properties in each row, from left to right, are: complementary cummulative distribution of rescaled degrees $k_{res} = k/\langle k \rangle$, degree dependent clustering coefficient over rescaled-degree classes, normalised average nearest neighbor degree $\langle k_{nn} \rangle_{res} = \langle k_{nn}(k_{res}) \rangle \langle k \rangle / \langle k^2 \rangle$.

The evolutionary dynamics are initiated by distributing a proportion of initial cooperators uniformly at random among the nodes of a network. For each of the three graphs (original network and the two similarity backbones with α^o and α^p), we vary the proportion of initial cooperators and quantify the level of cooperation achieved in the network at the end of the dynamics by measuring the fraction of final cooperators N_{coop}^{fin}/N_{gcc} after 10^5 rounds of the game. Notice no node alters its strategy while playing with its neighbors during an individual round. The fraction of final cooperators N_{coop}^{fin}/N_{gcc} is averaged over 100 realizations in all showcased curves in Fig. 4.9 showing the results for the dynamics on HSBs. Notice the system can reach a quenched state before the maximum number of rounds is achieved if all agents become either cooperators or defectors. In that situation the imitation mechanism does not induce any further evolution and the dynamics becomes effectively frozen.

The results in Fig. 4.9 (left and middle columns) show that the real networks have a tendency for their hierarchical similarity backbones to display final cooperation levels equal or greater than the achieved in the original network for equal proportions of initial cooperators, despite their radically reduced number of links. We expect the dynamics not to work as well as they would in moving the system towards full cooperation when the number of agents to convince is still large but the nodes play the game within small groups of neighbors, as it happens in the backbones. This is because the probability of adopting a neighbor's strategy, $P_{i \rightarrow j}$, drives the evolution towards consensus only when it reflects the tendency to copy a more successful neighbor. This mechanism may be compromised by large fluctuations in the difference of nodes' collected payoffs ($\pi_i - \pi_j$), when these payoffs come from just few interactions with a small number of neighbors. However, the enhanced final cooperation displayed by the HSBs is observed in general for all networks in Fig. 4.9, and specially for Metabolic and Music whose HSBs curves are visibly above the curves for the original networks (in black) for a wide range of initial conditions, $N_{coop}^{ini}/N_{gcc} \in (0.4 - 0.9)$. For instance, the HSB with $\alpha^p = 0.59$ of the Music network (see blue symbols curve in Fig. 4.9(k)) has 73% less links than the original network while preserving 83% of nodes, and still sustains up to ≈ 8 times more final cooperation than the original network, for a fraction of initial cooperators of 0.8. To further ensure that the enhanced cooperation actually stems from the categorical structure of the backbones one should compare an HSB curve with that of its corresponding random surrogate. By doing so, we observe that the random surrogates happen to reproduce closer the cooperative behaviour of the original network, that is LR curves in Fig. 4.9 follow the profile of the original network instead of appreciably deviating upwards, thus revealing that the surrogates do not provide a substantial gain in cooperation as opposite to HSBs. In general, the surrogates also require a higher proportion of initial cooperators, around $N_{coop}^{ini}/N_{gcc} \gtrsim 0.6$, to produce any sizable increase in final cooperation with respect to the original network. This indicates that the internal hierarchical organization of the HSBs is key to sustain enhanced cooperation. Actually, the similarity filter preferentially removes links with lower hierarchy load, usu-

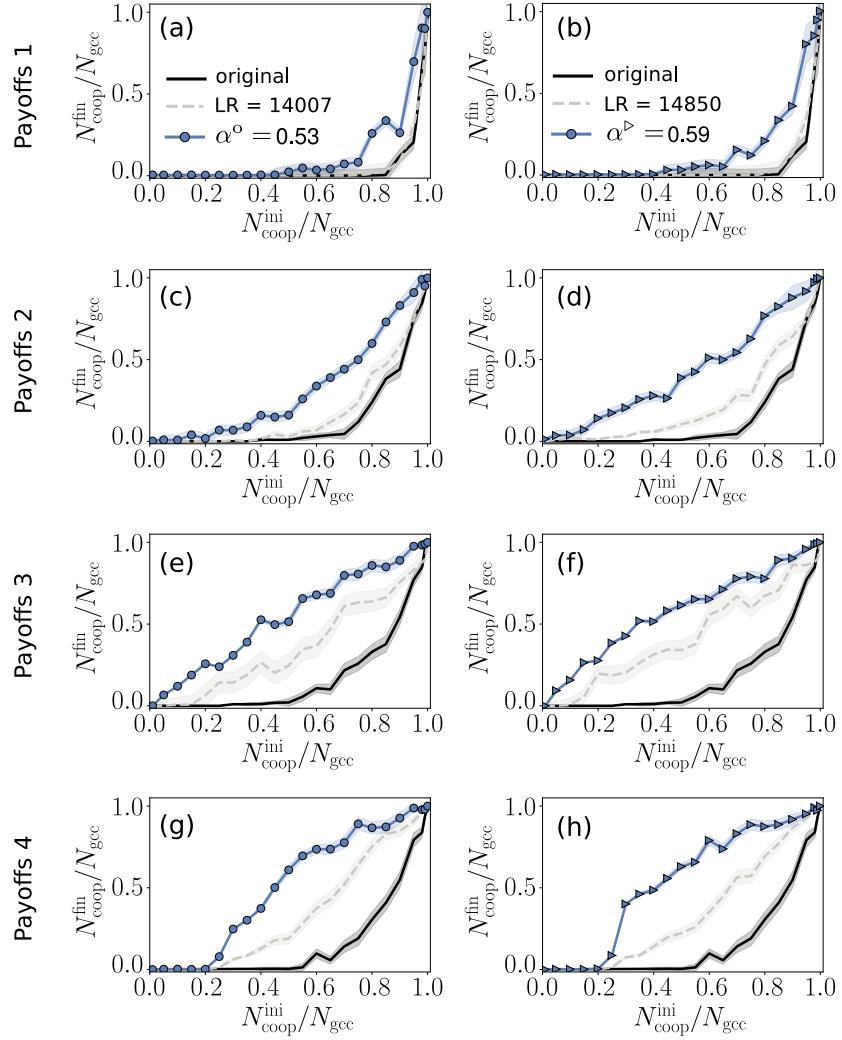


Figure 4.11: Fraction of final cooperators against fraction of initial cooperators for similarity backbones and random surrogates of Music network under different payoffs. The results of the dynamics in each plot are obtained using the following payoff values for the evolutionary prisoner's dilemma game: **(a-b)** $T = 1.0, R = 0.5, P = -0.5, S = -1.0$ **(c-d)** $T = 1.5, R = 1.0, P = 0.5, S = 0.0$ **(e-f)** $T = 2.0, R = 1.5, P = 0.5, S = 0.0$ **(g-h)** $T = 2.0, R = 1.5, P = 0.0, S = -0.5$. Results for the original network appear for reference, indicated by a black line. HSB filtered with $\alpha = 0.53$ are shown as blue curves with circles whereas HSBs filtered with $\alpha = 0.59$ correspond to blue curves with triangles in the second row. Results of random surrogates are displayed as grey dashed lines.

ally consisting of long-range connections established by high degree nodes, whereas the random removal makes no distinction. Given the scale-freeness of real networks, deleting a long-range link at random is less likely due to their scarcity. Therefore, during the dynamics, similarity backbones may develop clusters of same-strategy nodes that are more stable through the evolutionary process than those found in random surrogates, the reason being the former are less exposed to distant contacts belonging to clusters of opposite strategy. This means the hierarchical structure of HSBs enables a better shielding for the groups of cooperators in the shape of metric clusters [109], which in turn can explain the increased cooperation levels found in similarity backbones.

To additionally validate our results we choose to explore four more different combinations of payoff values for the Music network in Fig. 4.11. We observe that modifying the payoffs produces the same qualitative results as discussed above, with HSBs curves clearly surpassing the original network and evidencing that similarity backbones can reach superior cooperation.

4.6 Discussion

The existence of a metric space underlying complex networks allows us to provide an enriched interpretation of hierarchy that integrates two dimensions: popularity, or degree rank, and similarity between nodes, thus overcoming the problem of detecting hierarchies in the presence of clustering and the small world effect. The metric approach enables a unified framework to define the hierarchy loads of nodes and links in undirected networks.

We characterized the local contribution of a link or a node to the hierarchical structure of a network by measuring its *hierarchy load*, which depends on status, similarity, and the reference provided by the \mathbb{S}^1 used as a null model. In particular, we defined the link hierarchy load as the probability that a node and its lower-status neighbor are angularly closer in the network embedding than predicted by the \mathbb{S}^1 model. From this definition, the hierarchy load of a node can be readily computed within the same framework as the sum of link hierarchies of its lower-status neighbors. Alternatively, we present a formulation of node hierarchy load based on a measure of angular concentration of lower-status neighbors, given by the modulus of their circular mean in the similarity subspace.

Interestingly, the spectra of hierarchy loads of real networks revealed that, in general, these networks are less hierarchical than the reference provided by the maximum entropy null model and show greater variation in the hierarchy load of nodes across degree classes. Particularly, the lowest degree nodes typically contribute more towards the hierarchical structure, although their fluctuations are remarkable. Moreover, we introduced the *similarity filter*, a link pruning method which exploits the heterogeneity found in the hierarchy load of links. The filter extracts the connections that dominate the hierarchical structure of networks in terms of popularity and similarity, providing what we name hierarchical similarity backbones (HSB). The analysis of such backbones uncovered that, strikingly, the similarity filter is able to preserve network topological features at all scales

while discarding a large number of links. Accordingly, from a fundamental point of view, hierarchical backbones could help provide new insight about the percolation properties of highly stratified real networks, aiding control of cascading failures, as well as have the potential to become a standard methodology for the detection, visualization and inspection of hierarchical clusters [128] in machine learning and data science environments.

From a practical point of view, the similarity filter has proven to be an exceptional tool to unravel the backbone that sustains enhanced cooperation in social dilemmas on structured populations. This is in line with previous simulations of prisoners dilemma type dynamics on adaptive networks, showing that cooperation combined with imitation can lead to a hierarchical structure [89]. When this dynamics is played on heterogeneous contact networks with underlying metric structure, the evolution of cooperation leads to the formation of clusters of cooperators in the similarity subspace [109]. In the presence of these clusters, heterogeneity in the degrees was nevertheless found to hinder cooperation. Those findings reveal a tension between the popularity and similarity dimensions in evolutionary dynamics modelling social dilemmas. Our findings here solve this opposition by identifying the similarity backbones composed of significant links that are simultaneously hierarchical in terms of popularity and similarity, and which are expressly relevant in supporting and fostering cooperation.

Lastly, the methods developed in this contribution can be used to study the hierarchical nature of complex networks of any domain as long as they admit a geometric representation. The detection of hierarchical similarity backbones could for instance help in designing controllability of gene regulatory networks, improve communicability in information systems and infrastructures or assess robustness to species loss in ecological networks. Other possibilities for our framework include its extension to multiplex networks, opening promising future lines of research.

In the next chapter we will further bind structure and dynamics to address another relevant process closely related to the evolution of cooperation investigated here. This is the dynamics of opinion formation and, hence, the process of achieving consensus in structured populations. We will demonstrate how latent geometry becomes a powerful benchmark that naturally allows to incorporate in the dynamics, group-level information encoded in multiple scales of the network architecture.

5 Multiscale opinion dynamics

“A large part of our attitude toward things is conditioned by opinions and emotions which we unconsciously absorb.”

ALBERT EINSTEIN

This chapter – with some small changes – can be found as a preprint at arXiv repository under the title “Multiscale opinion dynamics on real networks” [129].

In Chapter 4 we showed that the geometric approach can help gaining knowledge about the hierarchical structure of real networks as well as shine light on puzzling processes running on top of them such as cooperation in social dilemmas. In this chapter we move one step forward and merge information from the metric structure of the networks directly into a paradigmatic dynamical process for opinion formation, the voter model. The result is a multiscale voter model, able to address opinion dynamics in real systems where group influence takes place at different granularities.

5.1 Group influence in opinion formation

Opinion dynamics can be modeled using systems of interacting agents, which continuously intermix their states based on established rules, in order to investigate the spreading of attitudes, beliefs, or sentiments in society. In this field, the Voter Model (VM) is an archetypal stochastic nonequilibrium model that gives a standard framework for studying imitation as the underlying mechanism of opinion formation. The initial appeal of the VM came from its exact solvability on regular lattices [130, 131]. Later, it was also extended to complex networks, whose non-regular topologies have nontrivial effects on the dynamics. In particular, the small-world property [20] extremely reduces the time to reach agreement in finite systems [132, 133]. Moreover, the heterogeneous distributions of the number of neighbors observed in real networks also promote quick consensus [134]. Conversely, in real life scenarios we rarely find that a large group of individuals easily comes to a consensus on sensitive topics. This dichotomy between eventual consensus in the VM on complex networks and the common experience of opinion diversity in society has motivated generalizations of the VM that include more realistic features such as zealots [135, 136], bounded confidence [137], noise [138] or memory effects [139].

In this chapter, this contradiction is addressed by introducing the Multiscale Voter Model (MVM), which assumes that opinion formation of an individual may be affected by the viewpoint of groups at different granularities—family, neighborhood, political party,

country—to which she belongs. Despite group-level information is known to affect behavioural responses in human [140] and even animal social networks [141], few models account for it. Some do for instance, through multiplex network representations [142,143] or by coupling individual information exchange with external information fields [144]. The MVM is built on network geometric models $\mathbb{S}^1 / \mathbb{H}^2$ (see Preliminary Methods), and relies on a latent metric space representation to account for homophilic influences at different scales. It interpolates in a natural way between states that reach consensus fast, as in the VM in small-world networks, and frozen disordered states typical of lattices, going through competition between opinion domains.

In this chapter numerical simulations are used to investigate the evolution of MVM dynamics in synthetic geometric networks and in several real data sets. The simulated dynamics uncover a transition between a final stage with mixed binary opinions to one of full consensus depending on the composition and size of the groups as well as the strength of influence. Strikingly, our results support that groups made of more diverse (or less affine) nodes, as captured by the latent geometry of the networks, promote eventual consensus. On the contrary, when the groups reflect similarity between nodes metastable clusters of same opinion emerge which hinder the transition to general agreement. Lastly, we studied in detail the phenomenon of opinion segregation and discovered that opinion clusters admit a visual interpretation as spatial patterns in the underlying hyperbolic space, leading to new insights in the dynamics of opinion polarization.

5.2 The multiscale voter model

The MVM runs on networks where nodes belong to meaningful groups at a coarse-grained scale. In the MVM, each node i represents a voter or individual who holds one of two possible opinions $s_i = \{-1, 1\}$. Individuals interact by copying the states of randomly chosen neighbors, as in the standard VM. To avoid biases induced by heterogeneous degrees, we implemented a *link update* dynamics [145]. A link of the network is selected uniformly at random and the roles of each node at both ends of the link, as copycat or imitated, are assigned with equal probability. In contrast to the standard VM, under the MVM a selected node does not copy the state of its neighbor straight away, but with a certain probability tuned by the influence of the node's group.

To incorporate the group-level influence on node i we divide the network into groups. These can be constructed in different ways as we will show later. We measure the opinion of group ν_i , to which node i belongs, as the average $s_{\nu_i} = \sum_{(l \neq i) \in \nu} s_l / (r - 1)$, where we have excluded node i and r indicates the total number of nodes in the group. Notice that s_{ν_i} is continuous in $[-1, 1]$ while individual states of nodes are binary. Next, we introduce a distance in opinion, $d_{i,\nu_i} = |s_i - s_{\nu_i}| \in [0, 2]$, between node i and its own group ν_i that serves to weight the probability that node i has of adopting the state of

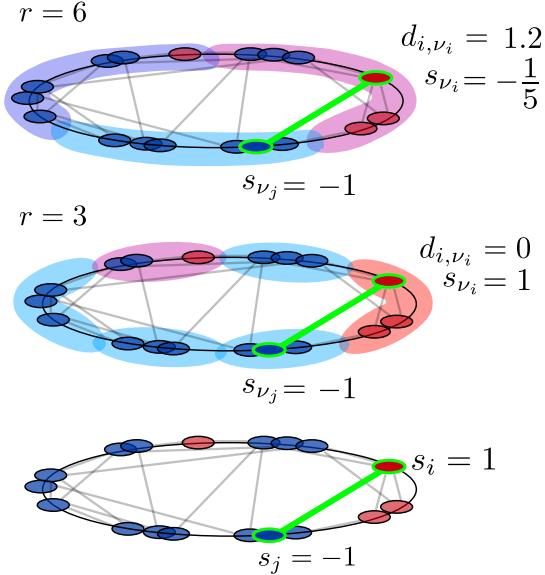


Figure 5.1: Graphic illustration of the MVM model on a toy geometric network. At the bottom, we highlight in green the link between a copycat node i with opinion $s_i = 1$ and its neighbor j with opposite opinion $s_j = -1$. In the middle, we construct similarity groups of size $r = 3$. Notice, when N/r is odd the last group to be build may contain less than r nodes. Groups are colored according to their opinion, ranging from -1 (all nodes blue) to 1 (all red) going through purple (mixed composition). At this point node i is completely aligned with the opinion of its $r = 3$ group so the probability to copy j is low. At the top, similarity groups include $r = 6$ nodes per group. At this scale, distance in opinion d_{i,ν_i} between node i and its group ν_i increases and so does the probability to copy j .

one of its neighbors j selected at random. This probability reads

$$P_{i \rightarrow j} = \frac{1}{1 + e^{\frac{1}{\tau}(1-d_{i,\nu_i})}} , \quad (5.1)$$

where parameter $\tau \in [0, \infty)$ controls the strength of group influence. The larger the value of τ , the less important the influence of the group is, and vice versa. We deliberately choose a Fermi-like probability function, which is a popular updating protocol in evolutionary game theory (EGT) and its stochasticity is similar to the real-world human decision-making process [146–149]. The probability in Eq. (5.1) reflects the popular tendency of individuals to refrain from adopting behaviours that contradict their environment or group norm. $P_{i \rightarrow j}$ is symmetric around $d_{i,\nu_i} = 1$, so that if a node is very aligned with its environment ($d_{i,\nu_i} \rightarrow 0$) it has less probability of copying node j ; and vice versa, if node i is not aligned with its group ($d_{i,\nu_i} \rightarrow 2$) the probability to copy j increases. When $\tau \rightarrow 0$ the probability of copying the state of the neighbor tends

to a step-function, which leads the system to frozen disordered states. On the other end, when $\tau \rightarrow \infty$ the MVM becomes the voter model with a rescaled activation rate that will make the process to evolve slower but eventually to reach consensus. The case where $0 < \tau < \infty$ is akin to introducing heterogeneous and dynamic activation rates dependant upon the states of node i and its group, and leads to system organization into competition between opinion domains.

In the following, we investigate three different prescriptions for designing the groups.

- First, we build groups at random, which serves as a null model that helps to discover unexpected behaviors.
- Second, groups are dictated by proximity in geometric network maps, allowing to explore the role of homophily in enhancing or hindering consensus. Similarity groups are formed by dividing the similarity circle of the S^1 model [25] in angular sectors with r consecutive nodes (see Fig. 5.1). As discussed in Chapter 2, in this space, every node i has a popularity-similarity pair of coordinates (κ_i, θ_i) and, most importantly, the more similar two nodes are, the shorter is their angular separation $\Delta\theta_{ij} = \min(|\theta_i - \theta_j|, 2\pi - |\theta_i - \theta_j|)$. This coarse-graining procedure for making the groups is also at the core of the geometric renormalization formalism for networks introduced in [26]. For consistency, we always start making the groups from the node with θ_{\min} but this is an arbitrary choice due to the rotational symmetry of the circle. Additionally, we exploit the H^2 model, the purely geometric isomorphic formulation of the S^1 model, to visualize the networks in a 2-dimensional hyperbolic disc when appropriate.
- Finally, groups are based on geometric communities [63], which tests the influence of the network's intrinsic modular organization into the opinion formation process. Geometric communities are essentially similarity groups of different sizes partitioned to maximize modularity. Further details regarding community-based groups are discussed in the following.

5.3 Simulation of the MVM

The MVM dynamics is simulated in real and synthetic networks starting from a random uniform distribution of states, where half of the nodes are in state $s = 1$. Despite network's finite size effects will always lead the dynamics to an absorbing consensus state, some realizations can be extremely long-lived. For this reason, we set a cutoff time t_c , so that a simulation is stopped when all nodes are in the same state or when time reaches t_c . At each simulation step, time is advanced by $\Delta t = (1/N)$ so that a node is updated once per unit time on average. The MVM algorithm can be implemented as:

1. Initialize opinions of all nodes uniformly at random. Here we chose $\rho_{s=1}(t = 0) = 0.5$.

2. Select a link of the network at random and, with equal probability, assign the copycat role i to a node at one end, and the neighbor role j to the other.
3. Change node i opinion to s_j with $P_{i \rightarrow j}$ given by Eq. 5.1.
4. Advance time as $t' = t + (1/N)$ and repeat steps from 2-4 until all nodes share same opinion, or $t = t_c$.

We measure the level of consensus in the network at termination time $t_{\max} \leq t_c$ as $\langle \text{Cons} \rangle = \langle |\rho(t_{\max}) - 0.5| / 0.5 \rangle$, where the average is over 100 independent realizations and ρ indicates the density of nodes in state $s = 1$. The lowest level of consensus, $\langle \text{Cons} \rangle = 0$, happens when half of the network is in one state and the other half in the opposite. We compute consensus fluctuations i.e., the susceptibility, as $\chi = (\langle \text{Cons}^2 \rangle - \langle \text{Cons} \rangle^2) / \langle \text{Cons} \rangle$, where the normalization factor is chosen following Ref. [150]. Finally, we evaluate the survival probability, S , measuring the fraction of realizations that have not reached consensus up to $t_{\max} = t_c$. S gives insight about how individual realizations contribute to the average level of consensus.

5.3.1 Consensus in synthetic networks

Simulations on \mathbb{S}^1 synthetic networks with realistic parameters allow us to estimate the impact of specific topological features in the final stage of the dynamics and better gauge the expected results for real networks. We generated synthetic networks [25] using the \mathbb{S}^1 model with sizes $N = 1000, 5000$, average degree $\langle k \rangle = 10$ and several γ, β parameter values, allowing to control for the scale-freeness of the degree distribution and the mean clustering coefficient, respectively. We show extended results for one of the networks in Fig. 5.2 and report additional results in Figs. 5.3.

Figures 5.2(a)-(b) show consensus heatmaps displaying the average level of consensus reached as a function of group size and influence strength within $r \in [2, 50]$ and $\tau \in [10^{-5}, 1.0]$ respectively. Figure 5.2(a) corresponds to similarity-groups and shows three different regions indicating low (yellow), intermediate (purple) and high levels of consensus (dark blue) at the final stage. In contrast, mainly two separate areas of no-consensus and full unanimity emerge for the dynamics using random groups, see Figs. 5.2(b). In the later case, the transition between the two different final outcomes of the dynamics happens at a value of $\tau_{\text{crit}} \lesssim 0.15$. On the contrary, when groups are based on similarity the same transition occurs more gradually and for a range of τ values that is particular to the network and dependant on the group scale r , see Figs 5.2(a).

This last observation becomes more evident in Figs. 5.2(c)-(d) where we show $\langle \text{Cons} \rangle$ against τ for several r 's. While for the random-group prescription we see (Fig. 5.2(d)) that the order parameter rises from 0 to 1 abruptly around $\tau_{\text{crit}} \simeq 0.15$, for similarity groups the slope change is smoother and τ_{crit} decreases as groups get smaller. For small similarity groups and low τ the dynamics is rather static because the impact of opinion alignment with the group becomes remarkable and, consequently, nodes are not prone

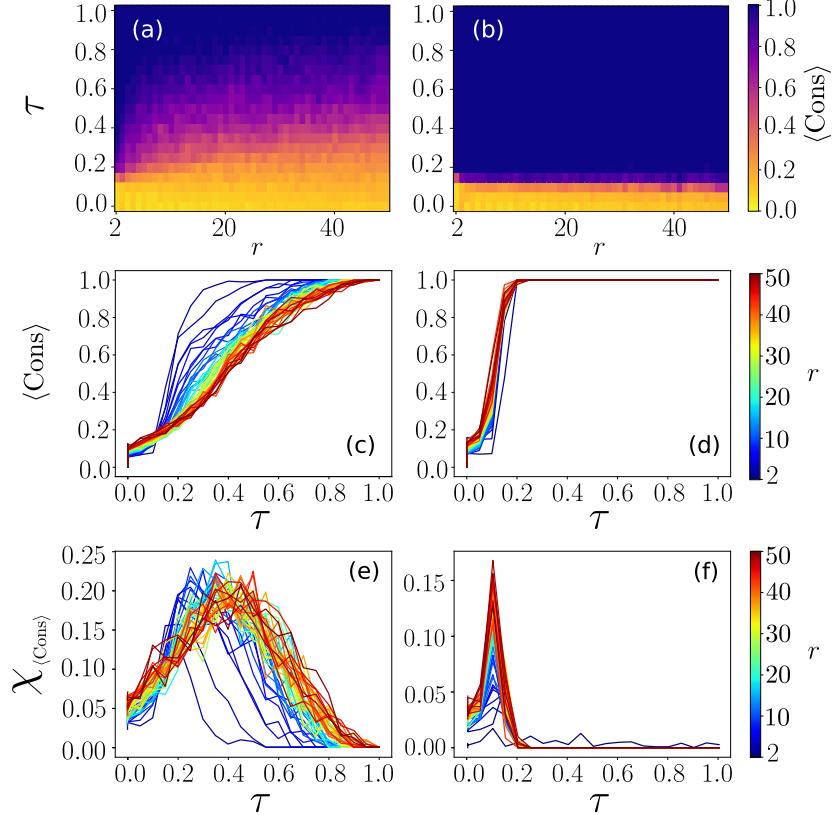


Figure 5.2: Consensus level in a \mathbb{S}^1 synthetic network of $N = 1000$ nodes, $\gamma = 2.5$, $\beta = 2.5$ and $\langle k \rangle = 10$. (a)-(b) MVM consensus heatmaps obtained using similarity groups and random groups respectively. Average consensus level over 100 realizations is color coded and displayed for a region of the parameter space confined in $r \in [2, 50]$ and $\tau \in [10^{-5}, 1.0]$. (c)-(d) Average consensus against strength of group-influence τ , for several group sizes r which appear color-coded from blue to red. (e)-(f) Fluctuations of average consensus level against strength of group-influence τ for several color-coded group sizes r .

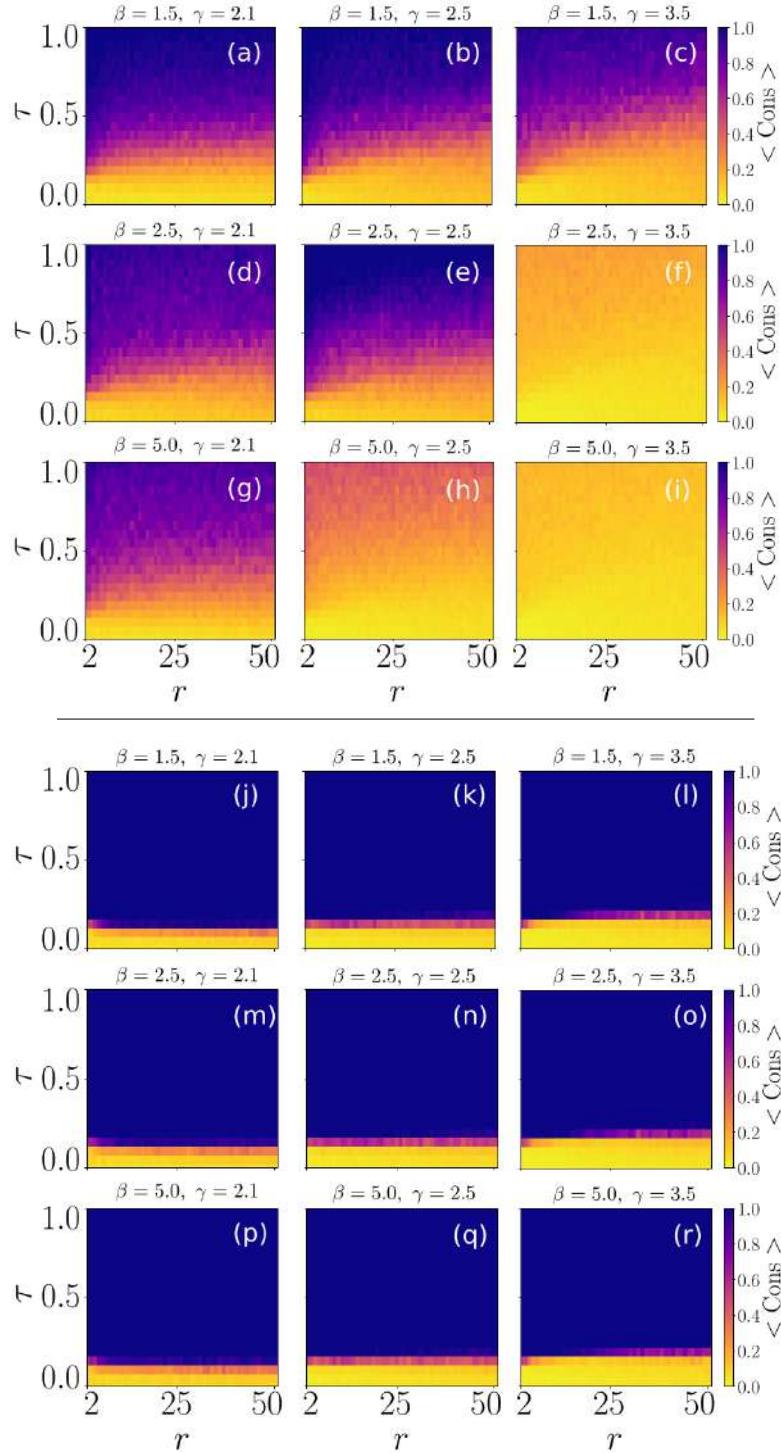


Figure 5.3: Consensus heatmaps (a)-(i) obtained using similarity groups (j)-(r) obtained using random groups, for synthetic \mathbb{S}^1 networks of $N = 5000$ nodes, $\langle k \rangle = 10$ and several (β, γ) . Strength of group-influence τ range is $[10^{-5}, 1.0]$ and group sizes $r \in [2, 50]$. Average of consensus level is over 100 independent realizations.

Data set	Nodes	N	E	β	k_{\max}	$\langle k \rangle$	$\langle c \rangle$	n_c
Dolphins	Dolphins	149	1553	25.34	69	20.84	0.80	4
NZ MPs	Politicians	94	662	10.39	35	14.09	0.78	9
Facebook	Friends	762	16651	1.60	248	43.70	0.43	9
WTW	Countries	189	550	2.12	110	5.82	0.63	6

Table 5.1: Properties of the data sets under consideration: N , size of the network; E , number of edges; parameter β estimated from the embedding of the real network; k_{\max} , highest degree; $\langle k \rangle$, average degree; $\langle c \rangle$, average clustering coefficient; and n_c , number of communities detected by the CGM method.

to copy different opinions from their neighbors. In that case, just a minor increase in τ can lessen the strength of group influence enough to suddenly move the system from its arrested state to a fast evolution towards consensus. Oppositely, for larger similarity groups (large r) there is a high rate of opinion exchange, which allows to sustain more intermediate levels of global consensus until t_{\max} for a wider range of τ values. In Figs. 5.2(e)-(f) we show the susceptibility of $\langle \text{Cons} \rangle$, χ , versus τ which further evidences the aforementioned transitions. In the case of similarity groups in Fig. 5.2(e) we observe fluctuations in the level of consensus rising around $\tau_{\text{crit}}(r)$, with higher peaks visible for small r centered around $\tau \approx 0.3$, and peaks for larger r 's around $\tau \approx 0.45$. In contrast, for random groups Fig. 5.2(f) reveals very pronounced peaks focalized at a single $\tau_{\text{crit}} \simeq 0.15$, with highest maxima obtained for larger r values.

The same differences observed between consensus heatmaps that use similarity groups or random groups are as well visible for the rest of synthetic networks produced with different β and γ values, see heatmaps in Fig. 5.3. Indeed, Figure 5.3 reveals that as the degree distribution of the networks becomes more homogeneous, the more the full consensus area of the heatmaps is reduced. Also, the dependence of τ_{crit} on r develops a more visible positive linear trend. As for the role of the mean clustering, from Fig. 5.3 we notice that changing β has a milder effect than changing γ but still, as the networks become more clusterized (higher β), the average consensus level detected in the whole parameter space diminishes. All things considered, what we realize is that as network topology becomes more homogeneous and clustered (increasing γ and β) the harder it becomes to achieve unanimity.

5.3.2 The MVM on empirical networks

We also applied the MVM to real networks, in particular, we considered four empirical data sets from different domains: the one-mode projection onto Members of Parliament of the political debate network of the 48th New Zealand Parliament [151] (NZ-MPs), a facebook friendship network among Caltech students [152] (Facebook), a social proximity network of bottlenose dolphins [153] (Dolphins) and the World Trade Web (WTW) [46]. For more information see Table 5.1 and Appendix. Notice that we selected networks

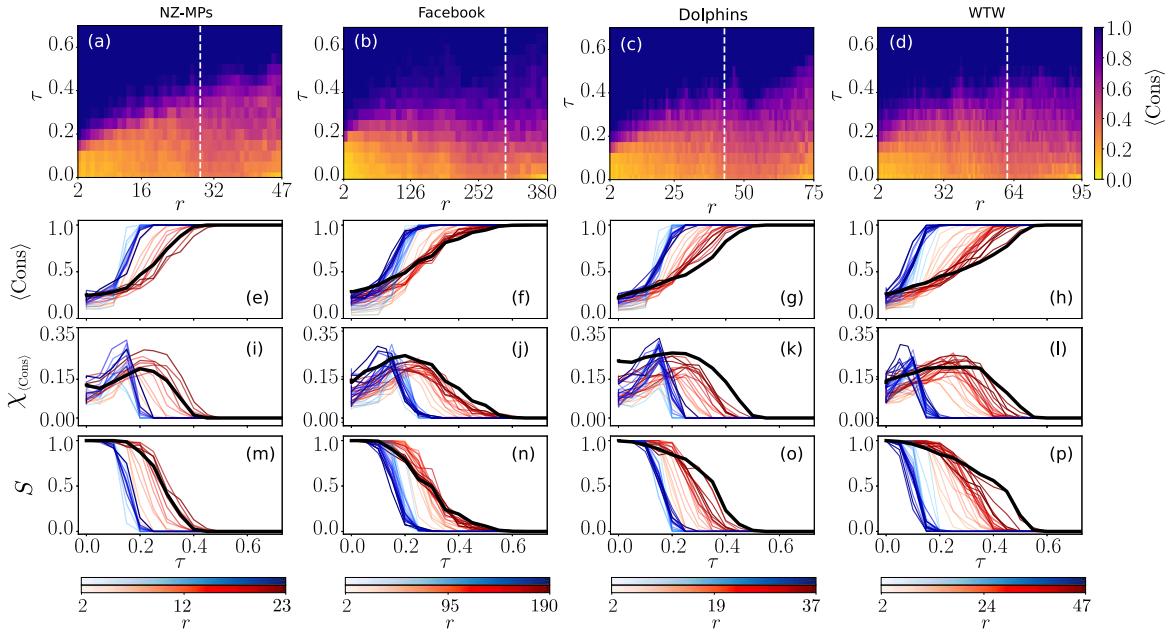


Figure 5.4: Consensus level in real networks. (a)-(d) Consensus heatmaps of four real networks obtained using similarity groups. Average consensus level is computed over 100 realizations and color coded from low (yellow) to high (dark blue). $\langle \text{Cons} \rangle$ is displayed in a region of the parameter space confined in $r \in [2, N/4]$ and $\tau \in [10^{-5}, 0.7]$. A white dashed line across the heatmap denotes the size r of the largest community detected via the CGM. In plots (e)-(p) red curves correspond to similarity groups and blue ones to random groups. Darker tones indicate larger group sizes in both cases. Solid black lines denote the case when groups correspond to communities. (e)-(h) Average consensus $\langle \text{Cons} \rangle$ against strength of group-influence τ , for several group sizes r . (i)-(l) Fluctuations of average consensus level against strength of group-influence τ , for several group sizes r . (m)-(p) Survival probability against strength of group-influence τ for several group sizes r .

where groups of nodes could potentially find a natural interpretation. For instance, in the WTW groups may denote geo-politically affine blocks, whereas in the NZ-MPs network groups could parallel political parties.

First, in figures 5.4(a)-(d) we show the consensus heatmaps for the MVM dynamics under a similarity group prescription. Analogous plots for the control case of randomized groups appear in Fig. 5.5(a)-(d). Since each real network has a particular size (see Tab. 5.1), r_{\max} is chosen as the group size that divides the network in two equal portions. As predicted, the heatmaps of all real networks show three areas of low, moderate and high level of final consensus, with predominant intermediate unanimity $\langle \text{Cons} \rangle \approx 0.5$ in the region of large r for a wide $\tau \in [0.15 - 0.5]$ regime. In contrast, for low r the level of $\langle \text{Cons} \rangle$ detected along the τ axis is rather bipolar. This is specially noticeable for all

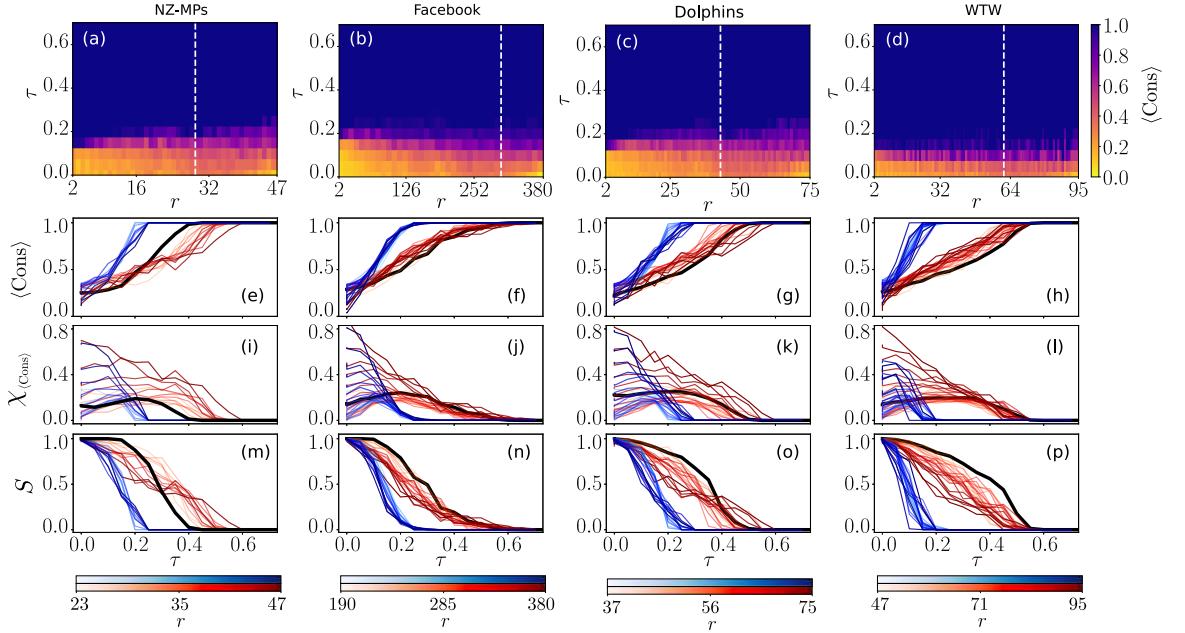


Figure 5.5: Consensus level in real networks. (a)-(d) Consensus heatmaps of four real networks obtained using random groups. Average consensus level is computed over 100 realizations and color coded from low (yellow) to high (dark blue). $\langle \text{Cons} \rangle$ is displayed in a region of the parameter space confined in $r \in [2, N/4]$ and $\tau \in [10^{-5}, 0.7]$. A white dashed line across the heatmap denotes the size r of the largest community detected via the CGM. In plots (e)-(h) red curves correspond to similarity groups and blue ones to random groups. Darker tones indicate larger group sizes. Solid black lines denote the case when groups correspond to communities. Notice r scale compared to Fig. 5.4. (e)-(h) Average consensus $\langle \text{Cons} \rangle$ against strength of group-influence τ , for several group sizes r . (i)-(l) Fluctuations of average consensus level against strength of group-influence τ , for several group sizes r . (m)-(p) Survival probability against strength of group-influence τ for several group sizes r .

real networks over the same parameter region ($r < 20$, $\tau \in [0.25 - 0.30]$). This indicates that, in real networks, group influence at smaller scales dictates more drastically whether the system evolves towards global agreement or not. Furthermore, when groups are not made of similar individuals but at random, all real networks heatmaps (see Fig. 5.5) display a neat transition centered at $\tau_{\text{crit}} \approx 0.15$ independent of scale r . This means mixed opinion configurations are invariably less stable over time when groups don't capture actual similarities.

In Figs. 5.4(e)-(h) we show $\langle \text{Cons} \rangle$ against τ for each real data set. We compare the results of the dynamics when considering similarity groups (red curves) and random groups (blue curves). Increasing r 's appear in darker tones for both. As anticipated, the four data sets show a more abrupt increase in $\langle \text{Cons} \rangle$ from 0 to 1 for random groups

(blue) than for similarity groups (red). Red slopes show significant variation with r for all data sets. Next, in figures 5.4(i)-(l) we analyze the fluctuations of the average final level of consensus, χ , against τ . For all real networks, blue curves indicating random groups show that fluctuations have a maxima around $\tau_{\text{crit}}(r) < 0.2$. A similar behaviour is found for the red curves (similarity groups), this time with flatter peaks displaying lower maxima at $\tau_{\text{crit}}(r) > 0.2$ in all cases. In the bottom row of Fig. 5.4 we present the survival probability S , indicating the fraction of independent realizations of the dynamics that live until the established cutoff time t_c . We observe a decreasing trend with τ for both, similarity groups (in red) and random groups (blue) in all networks. However, the decay is more abrupt for the blue curves, pointing out that before $\tau_{\text{crit}}(r)$ the networks are in very low agreement configurations, but as soon as $\tau > \tau_{\text{crit}}(r)$ most runs reach full consensus before the finishing time of the simulation. The red curves show indeed two distinct decreasing rates for small and large r . Simulations with smaller r in Figs. 5.4(m)-(p) decay fast and continuously to 0 while the process is slower and less monotonous for larger r 's in Figs. 5.5(m)-(p).

Lastly, we present the results of simulating the MVM dynamics with groups corresponding to the geometric communities detected via the Critical Gap Method (CGM), previously explained in Preliminary Methods. As discussed in Chapter 2.3, real network embeddings present heterogeneous similarity distributions where clusters of nodes concentrating nearby in the angular space reveal meaningful communities. The number of communities detected for each network appears in Tab. 5.1. Notice that for each particular network, communities, and therefore groups, now contain a different number of nodes. For instance, the size of the largest community in each of our data sets are: $r_c^{\text{MPs}} = 29$, $r_c^{\text{Dolph}} = 43$, $r_c^{\text{WTW}} = 61$ and $r_c^{\text{Fb}} = 302$. In figures 5.4(a)-(d) we indicate r_c of each network with a white color dashed line across the heatmap. Results of running the dynamics using the geometric communities as groups are reported on figures 5.4(e)-(p) with black color solid lines. Interestingly, when examining $\langle \text{Cons} \rangle$, χ and S , we identify a pattern that holds across networks. This is, the results for simulations using community groups follow the trend of the results for similarity-group curves of $r = r_c$. This pattern repeats for all metrics and real networks studied in Figs. 5.4 and 5.5. This indicates that indeed the largest community of the network effectively rules the evolution and eventual outcome of the MVM dynamics.

5.3.3 Opinion domains in the latent space

So far we have noticed that similarity groups in structured populations have a determining role in shaping the temporal evolution of opinions influenced by the environment. In particular, slowing down the process of achieving global consensus. Following, we show that MVM dynamics in fact triggers the formation of meta-stable clusters of homogeneous opinion that prevent the system from rapidly collapsing into consensus. Here we focus our study on two of the networks: NZ-MPs and WTW.

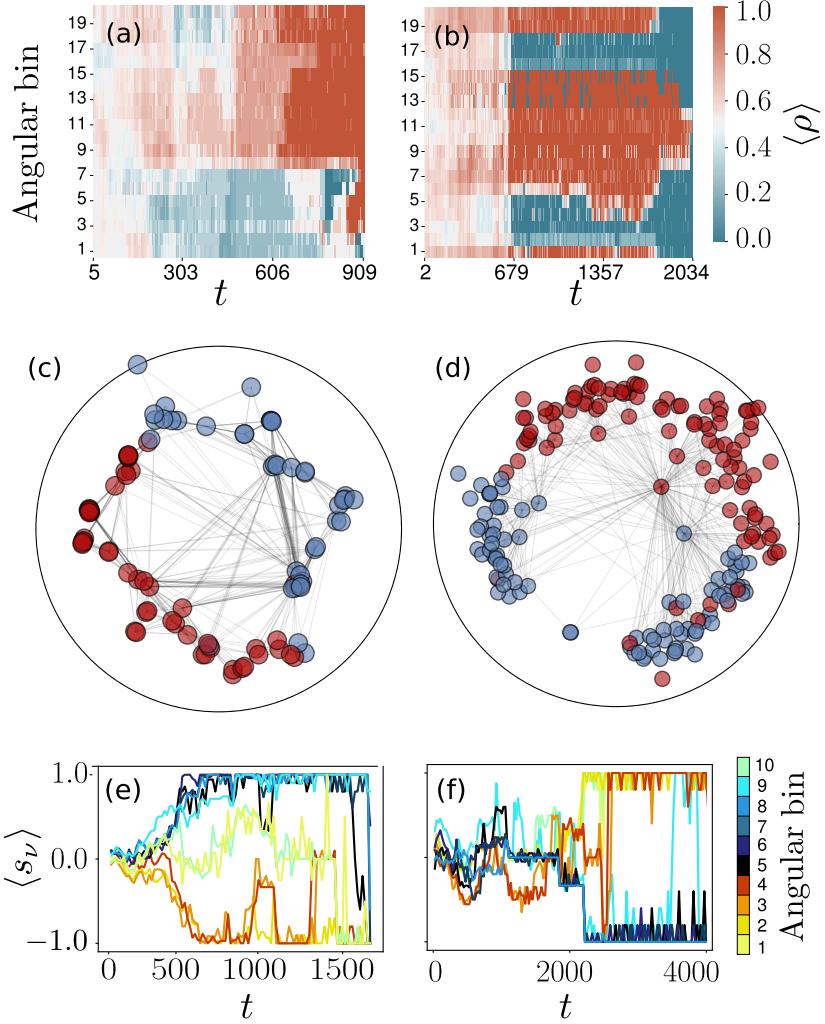


Figure 5.6: Geometric opinion domains of the NZ-MPs (left column) and the WTW (right column) networks. (a)-(b) Time evolution of average density $\langle \rho \rangle$ of nodes in state $s = 1$ in 20 equidistant bins of the angular coordinate θ . The simulations use $\rho_{\text{ini}} = 0.5$, $r = 10$, and $\tau = 0.45$. Notice in (a) mainly two different domains (red vs blue) arise over time while in (b) there is an intermediate time interval where two blue clusters are separated by 2 red domains. (c)-(d) Hyperbolic maps showing a time stamp of the MVM dynamics using similarity groups. We show $t_{\text{MPs}} = 851.05$ and $t_{\text{WTW}} = 1269.84$, and for both networks $r = 10$ and $\tau = 0.45$. Two angular domains of different opinion are visible with nodes in state $s = 1$ depicted in red and nodes in state $s = -1$ in blue. (e)-(f) Average group opinion s_ν over time for 10 similarity groups. The average is over 50 realizations where the random initial condition is fixed. Color code indicates the angular sector assigned to each group. Note that these angular bins are not equidistant (ie. of equal arc-lengths) since they are assigned one per similarity group, and the latter must contain a fixed number of nodes despite the angular heterogeneity of the network.

In Figs.5.6(a)-(b) we show the evolution of the average density $\langle \rho \rangle$ of nodes in state $s = 1$, in different equidistant angular bins of the similarity space. For both networks, we observe that initially the two opinions of nodes, $s = \{-1, 1\}$, are equally spread across all the angular space (see unsaturated color indicating $\langle \rho \rangle = 0.5$), in accordance with our initial condition. As time advances, Fig. 5.6(a) shows for the NZ-MPs network an increase in $\langle \rho \rangle$ (red) in upper bins (between 9 and 20), while the opposite opinion spreads in the rest of the angular space. Notice continuous boundary conditions make bin 1 and 20 contiguous. For the WTW Fig. 5.6 shows a rising $\langle \rho \rangle$ concentrating around central bins 6-15 and also at consecutive bins 1-20. In this case we find two clusters of $s = 1$ (red) that remain for most of the simulation separated by two other blue clusters of predominant opinion $s = -1$. Eventually, the blue clusters join ends and state $s = -1$ wins and $\langle \rho \rangle$ drops to 0. In figure Fig. 5.7 we extend the results for the WTW for various group sizes r while fixing $\tau = 0.6$. We observe that for all r 's, as long as similarity groups apply, opposite opinions tend to evolve in separate angular regions until one colonizes the space of the other or a big fluctuation drives the system to sudden full consensus. In general, it is difficult to observe more than one angular domain per opinion. Only for very particular values of the parameters we observe a maximum of two well-defined clusters per opinion sustained over time, such as discussed in Fig. 5.6(b).

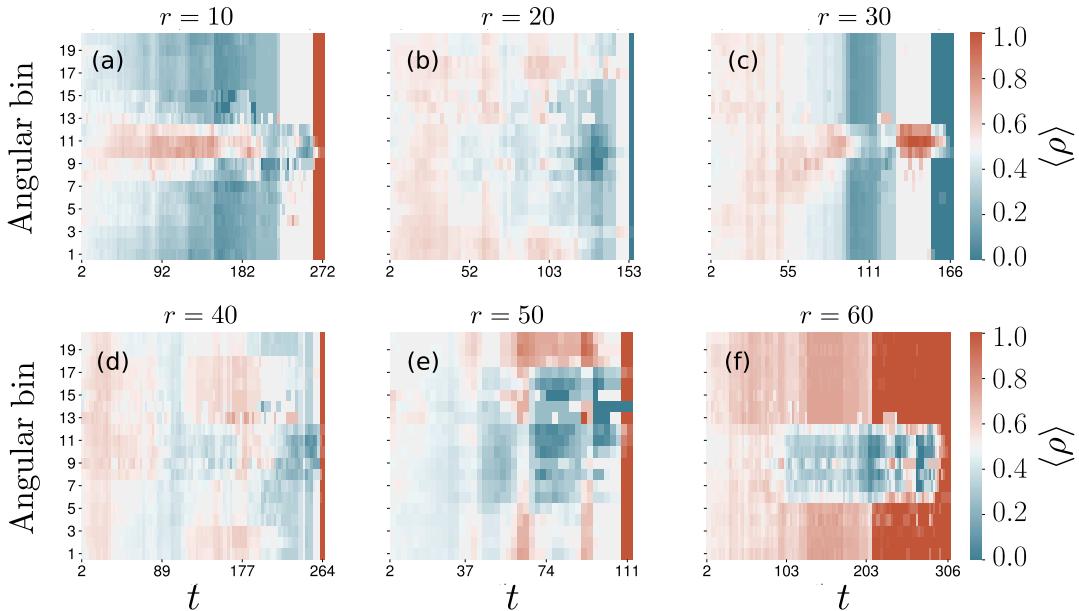


Figure 5.7: (a)-(f) Time evolution of average density $\langle \rho \rangle$ of nodes in state $s = 1$ in 20 equidistant bins of the angular coordinate θ of the WTW network. The MVM dynamics uses similarity groups, $\tau = 0.60$ and group sizes vary from $r = 10$ to $r = 60$.

Moreover, in figures 5.8 and 5.9 we expand the range of τ 's used to simulate the evolution of $\langle \rho \rangle$ for a fixed group size $r = 10$, for both NZ-MPs and WTW networks. In this case we provide also analogous plots using the random group prescription. Importantly, we assert that random groups do not display geometric domains over time. Instead, a higher (red) or lower (blue) $\langle \rho \rangle$ may alternate at some points in time but always spreading homogeneously along all angular bins. Besides, all MVM simulations using random groups span significantly shorter times on average than the ones using similarity groups. We thus confirm that similarity between nodes is key in sustaining metastable opinion clusters.

In Figs. 5.6(c)-(d) the hyperbolic maps showcase the spatial distribution of opinions in the networks at particular points in time for a single realization of the dynamics. Similarly to the case of domain formation in lattice topologies, we visualize the emergence of the discussed clusters of equal opinion along the angular dimension of the underlying hyperbolic space. Furthermore, we provide two animations showing the time evolution of node states in hyperbolic maps of the WTW, for the MVM dynamics under similarity and random groups respectively. These are available from the author upon request. The animation using similarity groups clearly features two spatial clusters that are sustained over time. We notice also that nodes alternating state most frequently are positioned at the borders of the two adjoining opinion domains. On the contrary, the animation using random groups does not exhibit any opinion segregation of node states along the circle. In this case, the dynamics evolves over time without nodes being more active in any particular region of the angular space.

Lastly, in Figs 5.6(e)-(f) we show the temporal evolution of the average group's opinion $\langle s_\nu \rangle \in [-1, 1]$. We divide both networks in 10 similarity groups and for visualization purposes assign an angular sector of the circle to each group. Fig. 5.6(e) shows the NZ-MPs network developing a gradual opinion radicalization from an early start. However, at least two groups in consecutive bins 1 and 10 are able to hold a neutral viewpoint $s_\nu \approx 0.0$. For the WTW we find an initial phase where groups are not opinionated at all and a second ($t > 2000$) where they rapidly adopt polarized postures. In both networks we observe a remarkable correlation between angular sector proximity of groups and the establishment of similar extreme opinions.

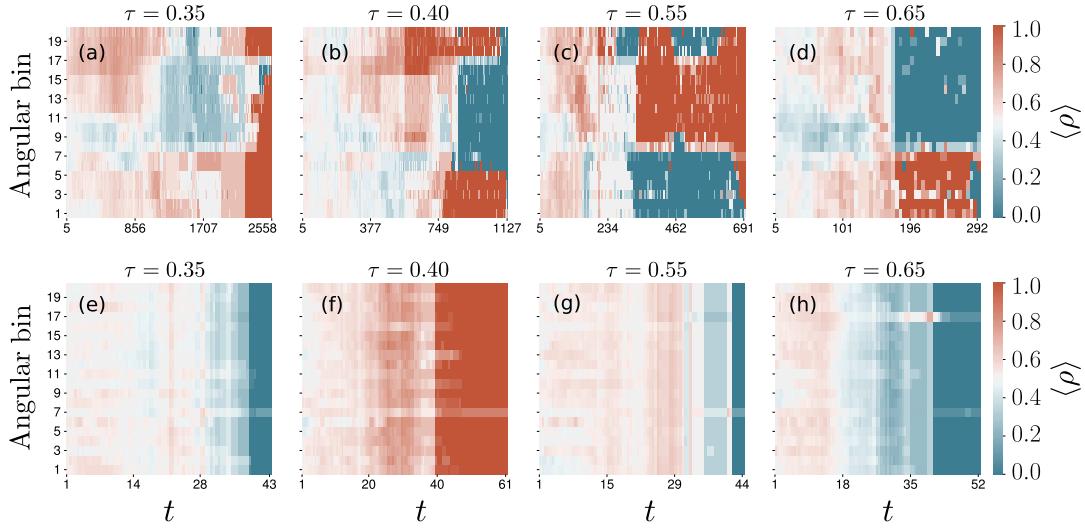


Figure 5.8: Time evolution of average density $\langle \rho \rangle$ of nodes in state $s = 1$ in 20 equidistant bins of the angular coordinate θ . (a)-(d) The MVM dynamics use similarity groups of size $r = 10$ and in (e)-(h) random groups or the same size. Notice the selected τ values span before and after the transition value τ_{crit} separating a no-consensus from full consensus final stages for the NZ-MPs network at $r = 10$.

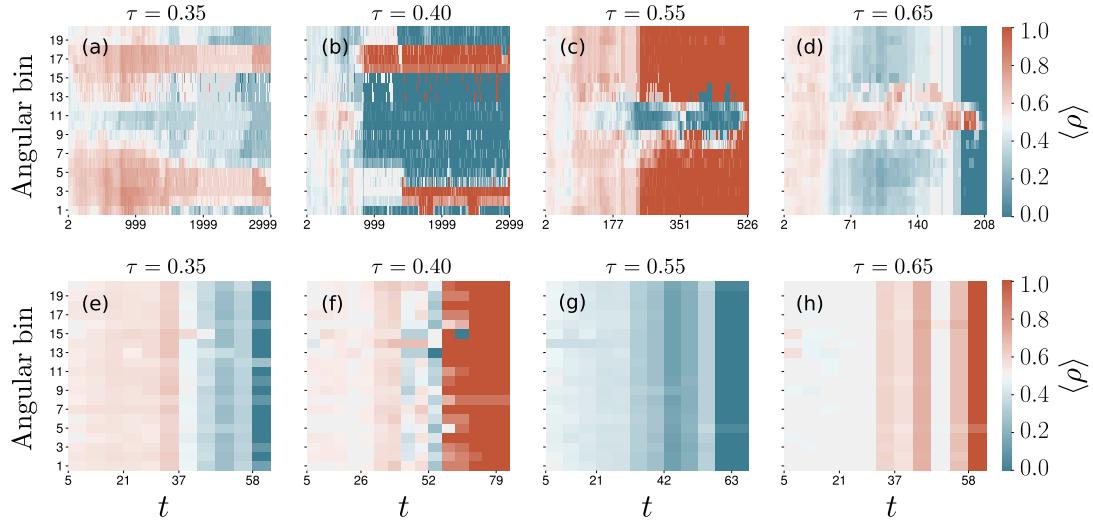


Figure 5.9: Time evolution of average density $\langle \rho \rangle$ of nodes in state $s = 1$ in 20 equidistant bins of the angular coordinate θ . (a)-(d) The MVM dynamics use similarity groups of size $r = 10$ and in (e)-(h) random groups or the same size. Notice the selected τ values span before and after the transition value τ_{crit} separating a no-consensus from full consensus final stages for the WTW network at $r = 10$.

5.4 Discussion

In this chapter we analized the evolution of consensus in real networks under a Multi-scale Voter Model (MVM) dynamics, where the probability that a node copies the state of a neighbor is influenced by the aggregated opinion of all nodes in its group. For that purpose, we constructed the groups based on geometric maps of networks, which are able to capture similarities between nodes and community structure. In that way, the probability of copying a neighbor is weighted by the distance in opinion between the copycat node and its own group, so that the less aligned is the node's opinion with the opinion of its environment the higher the chances it will copy the viewpoint of some neighbor. Therefore, the MVM probability reflects the popular tendency of individuals to refrain from adopting behaviours that contradict their group norm.

We have shown that the MVM dynamics on real and synthetic networks presents a transition between a final stage of mixed binary opinions and one of full consensus, which depends on the scale of the groups when those are made of similar nodes. Specifically, larger groups can sustain for longer mixed opinion configurations while small groups of very affine nodes yield to either very low or very high levels of consensus. This behaviour can be understood in terms of the more strict influence that very single-minded groups convey during the process of opinon formation and can explain why it is rarely observed that big groups easily come to a full consensus in the real world.

Moreover, we have shown that beyond group scale and strength of influence, group composition radically affects the outcome of opinion dynamics. Although large differences in backgrounds and perspectives might be expected to contribute to conflict and gridlock, we did not find this to be the case. In fact, we found that when groups included nodes with meaningful similarities the dynamics typically survived for longer without reaching global agreement. On the contrary, when groups where randomized and similarities dissolved the opposite was true. This indicates that, group diversity can help promote global agreement by reducing friction between sectors of like-minded individuals that pull in opposite directions. Indeed, real observations support that interdisciplinary teams can operate smoothly and reach high performance even in the abscense of a single leader [154].

Finally, we demonstrated that under certain conditions the MVM dynamics on real networks leads to the formation of stable domains of similar opinion, which can hinder the evolution towards consensus. These domains create visible spatial patterns in the angular dimension of the hyperbolic embeddings of the networks akin to the formation of spatial domains in Euclidean lattices. In this sense, the MVM represents a major step forward in the understanding of opinion segregation in real heterogeneous networked systems. Furthermore, our multiscale framework for opinion dynamics can be easily extended in many directions. For instance, one could study the influence of the group of the neighbor-node whose strategy is to be copied. This could help understand how social acceptance is modified depending on the backup tribe of the influencer. Another

possibility is to include multiscale zealots – groups that never change state –, which for example can mimic political parties with stringent ideologies. Another interesting alternative is to study the model including multiple discrete opinions. The extensions that our framework allows open up promising future lines of research.

Additionally, one should not forget that personal influence encoded in social networks is in constant evolution since the networks of connections between individuals themselves change over time. Characterizing and modeling time-varying networks is still an open and active area of research where much work is still needed to fully understand the mechanisms that perpetually reshape heterogeneous ties, and also to determine their effect on other dynamical phenomena. In the next chapter, we will delve in this direction and take into account the time-varying nature of real networks to investigate their effect on the paramount process of navigation.

6 Navigability of temporal networks in hyperbolic space

“The fundamental problem of communication is that of reproducing at one point either exactly or approximately a message selected at another point.”

Claude Shannon

This chapter was – with some small changes – published in “Scientific Reports” by the Nature Publishing Group under the title “Navigability of temporal networks in hyperbolic space” [155]. A preprint version is available at [156].

In the previous chapter, we showed how the static geometric maps of the networks can be effectively coupled to dynamical processes, adding valuable information to the modeling of intricate phenomena such as opinion formation in real topologies. In this chapter, we move one step forward and reframe our geometric maps as dynamic, so to model time-evolving topologies and investigate their role on a crucial network function: navigation.

6.1 Network maps and navigation.

Transfer of information, mass, or energy is a key function in many natural and artificial complex systems, ranging from gene-regulatory networks [157] and the brain [158] to online and offline social networks [159]. In fact, the embedding of real networks in hyperbolic space to produce network maps is a valuable resource for efficient network navigation [160, 161] and therefore transportation. This is especially important for some technological systems like, for instance, the Internet where the routing of information between Autonomous Systems relies on the Border Gateway Protocol (BGP), which presents scalability problems due to the dynamic and growing nature of the Internet. Luckily, such scalability issues are overcome when using an alternative hidden-geometric protocol, known as *greedy routing* [162] where dynamic routing tables are replaced by network maps. In the information forwarding process described by the greedy-routing, every information packet is equipped with the hyperbolic coordinates of the destination node, so that when an intermediate node along the path needs to decide which of its neighbours is optimal, it simply chooses the closest one to the destination in hyperbolic space. With this elegant procedure, the routing in network maps is successful with most greedy routing packets reaching destinations, and the paths traced are congruent with the geodesics in the underlying hyperbolic space, which results in the topological length

of the paths being almost minimal.

The advances in the understanding of the navigability of complex networks are typically framed within the traditional approach taking the structure of networks as static. However, this assumption has been recently challenged by the empirical observation of a temporal dimension in many natural and social systems [163–166], demonstrating that nodes and edges switch on and off with several time scales. The empirical analysis of such temporal networks has unveiled new statistical properties, such as a heavy-tailed distribution of inter-event times denoting intermittent functioning, also known as burstiness [167], or the heterogeneous distribution of activity in social interactions [168]. Besides, time-respecting paths [169] – where the next link traversed is always required to be active after the current one – play a crucial role in slowing down or speeding up the spreading of information or diseases [170], and certainly affect the message routing throughout the network. Although navigation is expected to be substantially different in temporal networks than in static ones, few empirical or theoretical works have been devoted to study the impact of the temporal dimension on the navigability of complex systems [162, 171, 172].

For this purpose, in the following we introduce a hybrid model, where nodes embedded in hyperbolic space are subject to a stochastic activation-inactivation dynamics simulating the temporal evolution of the networks. Surprisingly, we find that such dynamics enhances navigability with respect to the static case. Furthermore, we show that some real networks are ultranavigable, meaning that they remain highly navigable even when the network topology is strongly dynamic. Because our model considers static reconstructions of real networks and a simple node activation-inactivation dynamics, this allows us to control for the maximum duration of the routing process, as well as to discard peculiar features of specific real evolving systems, such as circadian rhythm [173]. In fact, the activation dynamics can be interpreted as temporal failures of nodes due to random unknown events, or noise. Our approach suggests a new greedy routing protocol in static networks, that combines standard greedy routing and a simulated activation dynamics, which can boost the navigability of some real networks, at the expense of elongated paths.

6.2 Greedy routing on temporal networks

Information packets, or other assets, are transferred in a network from a source node to a destination one by following greedy routing in hyperbolic space. To implement this forwarding mechanism, we consider a network embedded in a two-dimensional hyperbolic plane of constant negative curvature where each node i has polar coordinates (r_i, θ_i) . The routing algorithm requires that there is only one packet per source-destination pair, that each node knows its coordinates, the coordinates of its neighbors in the network, and the coordinates of the destination node. Then, the node holding the packet will transfer it to its neighbor with the smallest hyperbolic distance to the destination node.

We take the hyperbolic embedding of the largest connected component [174] of each real complex network considered here, that we refer as the static map $\mathcal{M}(G_0, S)$, where G_0 stands for the static graph and S is the underlying metric space where the nodes have permanent coordinates. Next, we generate several synthetic temporal networks by applying a Poissonian activation-inactivation dynamics on its nodes. We consider that nodes can be in an active state, being able to receive and forward information, or in an inactive state, in which case they cannot receive neither forward information packets. At each time step t , each node i is active with probability a_i . Thus, at each time step t , a graph G_t is defined, in which only active nodes and the links between them are present. The sequence of graphs $\mathcal{G} = \{G_t\}_{t=1,2,\dots,T}$ constitutes a synthetic temporal network of length (duration) T . The activation probabilities control the link density of the temporal networks, affecting the probability of a message being sent. For instance, in the case of a constant activation probability set equal for all nodes, $a_i = a$, each graph G_t has an expected average degree equal to $\bar{k}_t = a \bar{k}$, where \bar{k} is the average degree of the original static network.

Therefore, the greedy routing acts on a temporal map $\mathcal{M}(\mathcal{G}, S)$ depending on the temporal network \mathcal{G} and the underlying hyperbolic space S . The greedy forwarding algorithm is implemented sequentially on the temporal map $\mathcal{M}(\mathcal{G}, S)$, so there is one attempt to forward the information packet for each time step t . At time t , the node holding the information packet tries to forward it to its neighbor with the lowest distance to the final destination. If the neighbor is active at time t , then it receives the packet. Otherwise, the packet remains at the holding node. The model with $a = 1$ corresponds to greedy routing on the original static network, with all nodes active at all times, for a number of steps equal to T . Therefore, the network's duration T can be interpreted as the maximum lifetime of information packets. In this scenario, a greedy path is successful when a packet reaches its destination in a time $t \leq T$, and unsuccessful otherwise. In the limit of $T \rightarrow \infty$, all packets in a time-varying network are expected to be able to reach their destination because the number of different paths that can be realized by greedy routing on the temporal networks grows with T . This is in contrast with static networks, where some packets can be trapped at specific regions.

We run numerical simulations for different network's duration T , taking a number of random source–destination pairs which is the minimum between 10^5 and $N(N - 1)/2$, where N is the number of nodes of the network. In numerical experiments varying the activation probability, the random subset of source–destination node pairs is kept the same, while it is changed when varying T .

6.3 Performance metrics and empirical data sets

We first consider a constant activation probability set equal for all nodes, $a_i = a$. In this case, the model is characterized by two parameters, the activation probability a , which controls the activation dynamics, and the network's duration T , which represents

the maximum lifetime of information packets. We evaluate the performance of greedy routing on the temporal map by measuring two main quantities: the *success ratio* p_s , defined as the fraction of packets that successfully reach their destination within a time T over the total number of source-destination pairs considered; and the *average topological stretch* \bar{s} of successful greedy paths, where the stretch is defined as the ratio between the hop-length of a greedy path and the shortest path between the corresponding source and destination nodes. The stretch tells us how much the successful greedy paths are longer with respect to the shortest ones. We also provide results for the average geometric stretch \bar{s}_g , which is defined analogously to \bar{s} but considering the hyperbolic lengths of greedy and shortest paths; and the average coverage Cov , which informs of the average number of different visited nodes against the average number of nodes that compound a successful path.

To assess the performance of greedy routing in time-evolving real networks we use five different real scale-free networks from different domains. In particular, the datasets considered are: the Arxiv coauthorship network, the US commodities network, the Metabolic network of bacterium E. Coli, the Internet at the AS level and, finally, the World Trade Web (WTW). Detailed descriptions of these data sets can be found in Appendix A. In Table 6.1, we report the values of the mentioned performance metrics and the main topological properties of the five data sets. Notice that we used the giant connected component of the networks in all cases.

Data set	Nodes	N	E	$\langle k \rangle$	k_{\max}	γ	p_s	\bar{s}
Arxiv	Authors	2121	5473	5.16	70	2.86	0.24	1.14
Commodities	US industrial sectors	374	1090	5.83	86	2.61	0.57	1.19
Metabolic E.C.	Metabolites	1008	3285	6.51	143	2.53	0.77	1.17
Internet	Autonomous Systems	23748	58414	4.92	2778	2.10	0.97	1.11
WTW	Countries	189	550	5.82	110	2.22	0.98	1.04

Table 6.1: Topological properties and navigation performance values of five real static maps $\mathcal{M}(G_0, S)$. From left to right: number of nodes, number of edges, average degree, maximum degree, exponent of the power-law degree distribution, success ratio and average topological stretch.

6.4 Effects of network dynamics on navigability

The success ratio p_s is a key parameter in determining the navigability of complex networks. A large success ratio, close to $p_s \sim 1$, means that almost all nodes can be reached by a message sent by any other node. On the contrary, if p_s is small, information can not be successfully transmitted from most nodes. Fig. 6.1, top row, shows the fraction of successful paths p_s as a function of the network duration T , for different values of the activation probability a . The success ratio varies considerably across different static

networks ($a = 1$), ranging from very low success ratio for the Arxiv, to $p_s \sim 1$ for the WTW and the Internet, which indicates a better congruence of these systems with their underlying geometry. Remarkably, for sufficiently large T , the success ratio in all temporal networks ($a < 1$) under consideration is always larger than the one achieved on their static counterparts (dashed line, top row Fig. 6.1 and Table 6.1). This effect is particularly evident for the cases where p_s in the static map is rather low, such as for the Arxiv network, where the success ratio increases from $p_s = 0.24$ for $a = 1$ to $p_s = 0.90$ for $a \sim 0.2$. Nonetheless, when the static success ratio is high (e.g. Internet), p_s on the temporal maps increases too. As expected, p_s is a growing function of the network duration T : the larger the maximum lifetime of the packets, the higher p_s . In the limit of $T \rightarrow \infty$, p_s is expected to reach its maximum since, for any pair of nodes, all different paths between them will be available at some time, ensuring that a successful one will certainly arise. This implies that the success ratio always increases with T , although the growing rate can be extremely slow for very large T . Oppositely, in the routing on static networks, p_s does not vary with T because no new paths are added by increasing the lifetime of information packets.

Our results show that, surprisingly, it is more efficient to have some (or even a great number of) nodes inactive than having all nodes active and contributing to the routing process. The reason for this behavior is rooted in the fact that, with $a = 1$, some packets might get stuck into *topological traps*. From the greedy routing definition, indeed, it is clear that if a packet comes back to a node twice, it will come back again, and the loop would continue forever with the packet never reaching its destination.

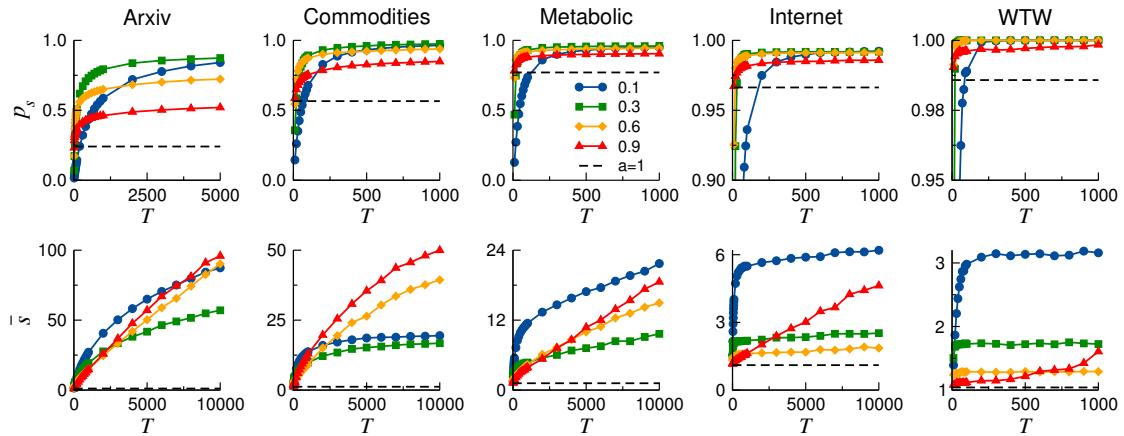


Figure 6.1: Success ratio p_s (top row) and average stretch \bar{s} (bottom row) as a function of T , for different values of the activation probability a , in five real networks. The success ratio and the average stretch in the static map, $a = 1$, appear plotted with dashed lines.

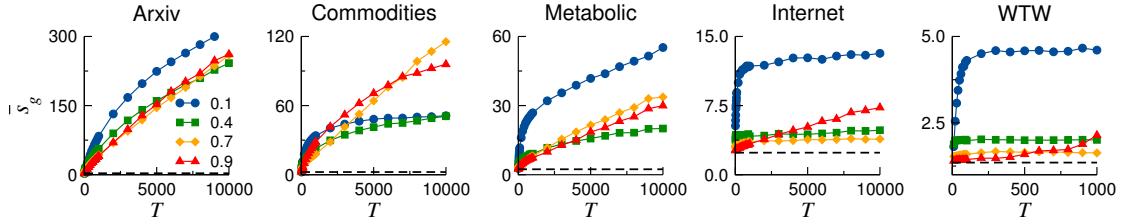


Figure 6.2: Average geometric stretch \bar{s}_g as a function of T , for different values of the activation probability a , in five temporal maps $\mathcal{M}(\mathcal{G}, S)$. The \bar{s}_g of successful greedy paths, is defined as the ratio between the cumulative hyperbolic distance from the hop-length of greedy paths and the corresponding geodesics. Notice that nodes that occupy the exact same position in the hyperbolic disc were not considered as possible source–destination pairs since the geodesic connecting them is of length 0, thus making \bar{s}_g to diverge. Average geometric stretch for $a = 1$, corresponding to greedy routing in the static maps, is plotted as a dashed line and takes the following values for each data set: $\bar{s}_g^{\text{Arxiv}} = 3.10$, $\bar{s}_g^{\text{Comm.}} = 2.33$, $\bar{s}_g^{\text{Metab.}} = 2.38$, $\bar{s}_g^{\text{Internet}} = 2.39$ and $\bar{s}_g^{\text{WTW}} = 1.35$.

To understand this mechanism, consider a node i sending a packet to his neighbor j , because j is the closest (among all i 's neighbors) to destination node k . If during the next time step, node i turns out to be the closest node to destination k among j 's neighbors, then the packet will return to i .

As long as no topological change takes place in the network, this process will repeat endlessly. Any cycle involving a packet coming back to a node twice constitutes a topological trap, see Fig. 6.3. In contrast, if $a < 1$, the topology of the network changes at each time step, hence the packet is able to escape any topological traps it may encounter along the route and eventually reach its destination. Nevertheless, the new successful path followed by the packet will deviate from the geodesic connecting the two nodes in the hyperbolic plane, thus the path length will necessarily be longer than the shortest.

The average topological stretch \bar{s} , defined as the ratio between the hop-length of greedy paths and the corresponding shortest paths in the network, is also a measure of navigation efficiency. From its definition, it holds that $\bar{s} \geq 1$. A small stretch, $\bar{s} \gtrsim 1$, indicates that most packets follow a route very close to the shortest one, while if $\bar{s} \gg 1$, paths are much longer. Fig. 6.1, bottom row, shows the average stretch \bar{s} as a function of the network's duration T , for different values of the activation probability a . As for the success ratio, \bar{s} is also an increasing function of T . Indeed, the larger the duration T , the lengthier the paths that become successful, and these very long paths increase the average stretch. It is important to note that the shortest paths between two nodes in $\mathcal{M}(\mathcal{G}, S)$ may be much longer than the shortest path in the corresponding static maps because time-respecting paths require links to be active at time t in order to traverse them. This is particularly true for very sparse temporal networks, i.e. with low activation probability. Therefore, \bar{s} is always greater in temporal maps than in

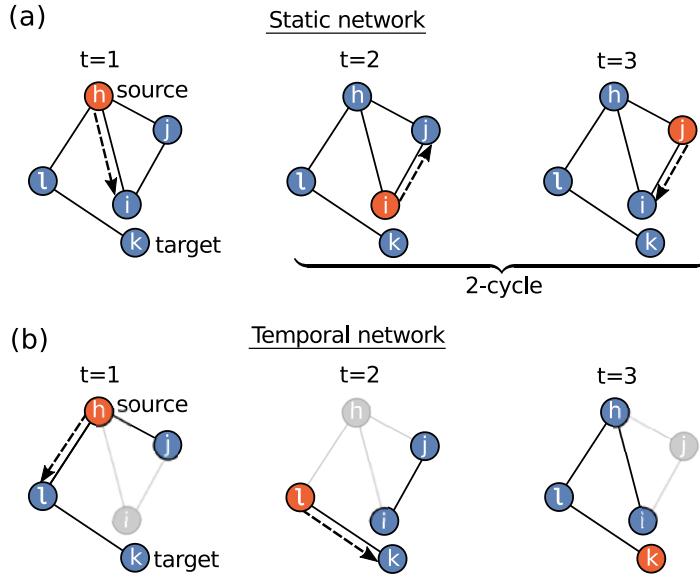


Figure 6.3: Representation of a topological trap, in Euclidean space. Greedy routing demands sending the information packet always to the neighbor closest to destination. Consequently, in **a)** the packet never stops jumping between nodes i and j and is captured by the topological trap, while in **b)** the inactivation of i enables the packet to follow an alternative route through l and successfully reach destination node k .

the corresponding static ones, as shown in Fig. 6.1, demonstrating that the activation dynamics is responsible for creating lengthier successful paths. This is also true (see Fig. 6.2) for the average geometric stretch denoting the elongation of hyperbolic greedy paths. This effect is clearly visible in those data sets where p_s in the static networks is low, such as the Arxiv or Commodities. In these networks, the large increase in the success ratio due to the activation dynamics comes with a large growth in the average stretch. The probability of finding much more successful tracks is increased at the cost of choosing longer paths. On the contrary, if p_s in the static maps is high, such as for the Internet or the WTW, \bar{s} shows a small increase in the temporal maps. These different profiles correspond to the different geometricity of the considered networks. In fact, the less congruent topology and geometry are, the larger the number of topological traps present in $\mathcal{M}(\mathcal{G}, S)$ and the larger the potential increase in success. Temporal maps with limited congruency, such as Arxiv or Commodities, show the larger gains in success overcoming traps at the expenses of a notable increase in \bar{s} , and longer durations T . On the contrary, networks with a conspicuous latent geometry, like the Internet and the WTW, are not characterized by a large number of topological traps, hence \bar{s} does not rise as much.

Interestingly, different effects are obtained on the success ratio and the stretch depending on the activation probability. The lowest values of \bar{s} are found for intermediate

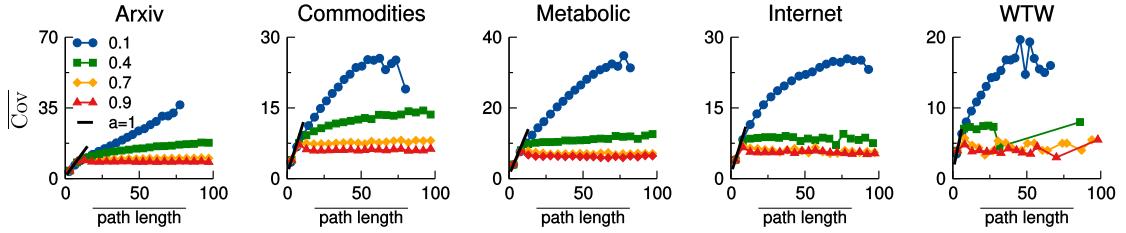


Figure 6.4: Average coverage $\overline{\text{Cov}}$ as a function of the average path length of successful paths for different values of the activation probability in five temporal maps $\mathcal{M}(\mathcal{G}, S)$. The black solid lines correspond to $a = 1$, which indicates the greedy routing is performed on static maps. For the static references, we observe straight lines of slope 1 because in the static case, a path is always declared unsuccessful if a node is tried to be visited twice, hence, the average path length always coincides with the number of different visited nodes producing unitary slopes. Nonetheless, the solid lines inform us about the maximum average path length that one can expect from a network only due to its own structure.

values of a , while p_s generally increases as the activation probability decreases, down to a value for which the network becomes too inactive, and then p_s becomes lower again. In most networks, p_s remains almost unchanged if the activation probability is set equal to $a = 0.3$ or to $a = 0.1$ in the limit of large T , while \bar{s} significantly increases if the activation decreases from $a = 0.3$ to $a = 0.1$, specially for the Internet and the WTW. Conversely, choosing $a = 0.6$, p_s grows from 0.96 in the static case to 0.99 for the Internet and from 0.99 to 0.997 in the WTW, but the stretch increases very little from $\bar{s} = 1.11$ to $\bar{s} \approx 1.76$ for the Internet, and from $\bar{s} = 1.04$ to $\bar{s} \approx 1.29$ for the WTW. This indicates it may exist an optimal activation probability that maximizes the increase in the success ratio and minimizes the increase in the stretch.

Lastly, we examine the average coverage $\overline{\text{Cov}}$, informing the mean number of *different* visited nodes, against the average number of nodes that constitute a successful path. The $\overline{\text{Cov}}$ measures how much the information packets travel around the network and hence the metric space. Figure 6.4 shows that for low activation probability ($a = 0.1$), an information packet jumps among a great number of different nodes, whereas for high activation ($a = 0.9$) it only visits a few. By lowering the activity of the nodes, the information packet is forced to jump often to any available position, hence we find that the longer the path the higher the coverage. On the contrary, when most nodes remain almost invariably active, the longer path lengths displaying low $\overline{\text{Cov}}$ confirm the idea that the packet is found moving in cycles, failing to escape a topological trap.

6.5 Optimal activation probability

Top row of Fig. 6.5 shows \bar{s} as a function of the activation probability a , for several values of the network's duration T . Interestingly, the average stretch is not a strictly decreasing function of a , but it reaches a minimum for some intermediate value. On one hand, when the activation probability is very small, the stretch is typically large because of the lack of available active neighbors. The packet will usually remain in the holding node or it will be transferred erratically, resulting in an increase of \bar{s} . On the other hand, if $a \lesssim 1$, the topology of $\mathcal{M}(\mathcal{G}, S)$ is similar to the static one, so the packet tends to fall into the same topological traps spending a long time moving in cycles (thus increasing the stretch) before it succeeds to escape the loop. Remarkably, the minimum of \bar{s} is reached for some optimal value of the activation probability, $a = a_O$.

This feature is addressed in more detail in the bottom row of Fig. 6.5, which shows the ratio between the success ratio and the average stretch (p_s/\bar{s}) as a function of the activation probability a . Since $p_s \leq 1$ and $\bar{s} \geq 1$ by definition, perfect navigability is reached when $p_s = \bar{s} = 1$, and thus $p_s/\bar{s} = 1$. The ratio p_s/\bar{s} represents a measure of the trade-off between the increase in both the success ratio and the stretch. The larger the ratio, the more efficient the navigation.

For each network under consideration, it exists an optimal value a_O of the activation probability that maximizes the trade-off between success ratio and stretch. For the Arxiv and the Commodities, the curves of the ratio p_s/\bar{s} as a function of a depend on the duration T , with larger p_s/\bar{s} for smaller T , while for the Internet and the WTW, these curves are independent of T and collapse. Fig. 6.5 shows that the WTW combines the largest success ratio with the smallest stretch, followed by the Internet, Metabolic, Commodities and the Arxiv networks.

However, it is important to remark that the p_s/\bar{s} ratio is always higher for the static maps than for the temporal ones. For instance, the static value for the Internet is $p_s/\bar{s} = 0.87$ (see Tab. 6.1), while in the temporal network it does not exceed 0.70. This is due to the fact that, in temporal maps the large gain in success, which is bounded with a top value of 1, necessarily comes with an increase in stretch, which in fact can be quite limited, as for the Internet or the WTW, but is unbounded.

6.6 Heterogeneous activation dynamics

In this section, we analyse how navigation is affected by an activation probability which varies across nodes. We do it in two different fashions: *i*) using a constant activation probability $a < 1$ only for nodes whose degree belongs to a certain interval, and *ii*) using an activation probability linearly dependent on nodes' degrees.

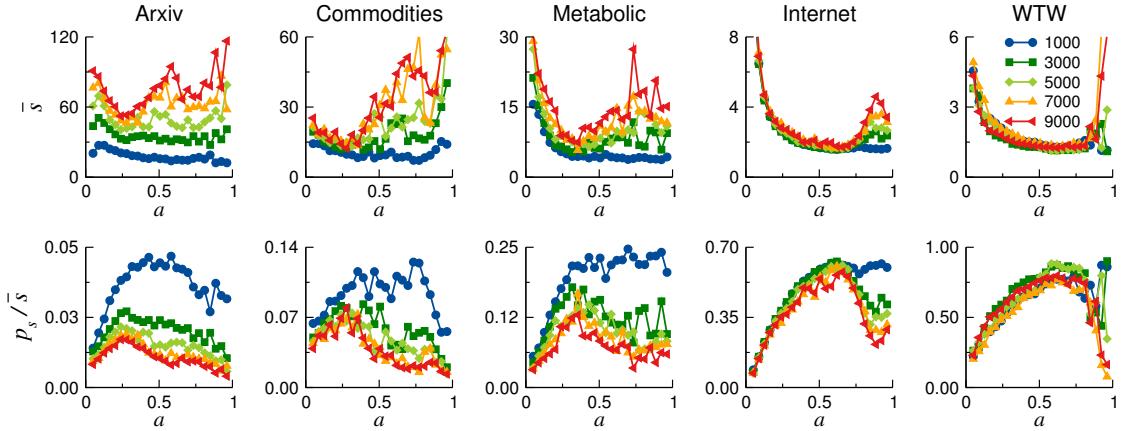


Figure 6.5: Average stretch, \bar{s} (top row), and success ratio divided by the average stretch, p_s/\bar{s} (bottom row) as a function of a . Each curve corresponds to a different value of the duration T . Notice the rightmost point is not 1 (the static reference) but 0.96. The optimal activation a_O of each network is observed as a maximum in the plots of the bottom row. The approximate values of a_O for each network are: $a_O^{\text{Arx}} \approx 0.25$, $a_O^{\text{Com}} \approx 0.25$, $a_O^{\text{Met}} \approx 0.33$, $a_O^{\text{Int}} \approx 0.60$ and $a_O^{\text{WTW}} \approx 0.67$.

6.6.1 Activation of nodes within degree intervals

Here, the random activation dynamics is targeted to subsets of equal number of nodes with degrees in a certain range of values. We measure the success ratio when only one of these subsets of nodes is randomly activated-inactivated with constant a and the rest of the network remains active. To implement this prescription, we order all nodes in a network from highest to lowest degree and divide this sorted list in segments of same number of nodes. The node bins are then labeled using the average degree \bar{k} of the nodes belonging to that bin. The size of the bins has been set to $\xi = 5\%$ of the total number of nodes N .

This method aims at identifying which degree intervals have a major contribution to the destruction of topological traps, and hence specially boost the success. We find that all temporal maps $\mathcal{M}(\mathcal{G}, S)$ experience a sudden increase in p_s when \bar{k} is at its maximum, see Fig. 6.6. In scale-free networks, node degrees are distributed as a power-law $p(k) \sim k^{-\gamma}$ (the γ values for the considered networks are reported in Table 6.1). This means that the interval \bar{k}_{\max} contains not only the biggest hub but also several densely connected nodes. Therefore, our results imply that switching on and off nodes with low degree has a limited effect on the efficiency of navigation. In fact, the higher the degree of a node the more radical the changes it can induce in the direction of a greedy path.

When we construct the bin with temporal behaviour by sampling the nodes uniformly at random from any part of the degree spectrum, so that the bin is approximately

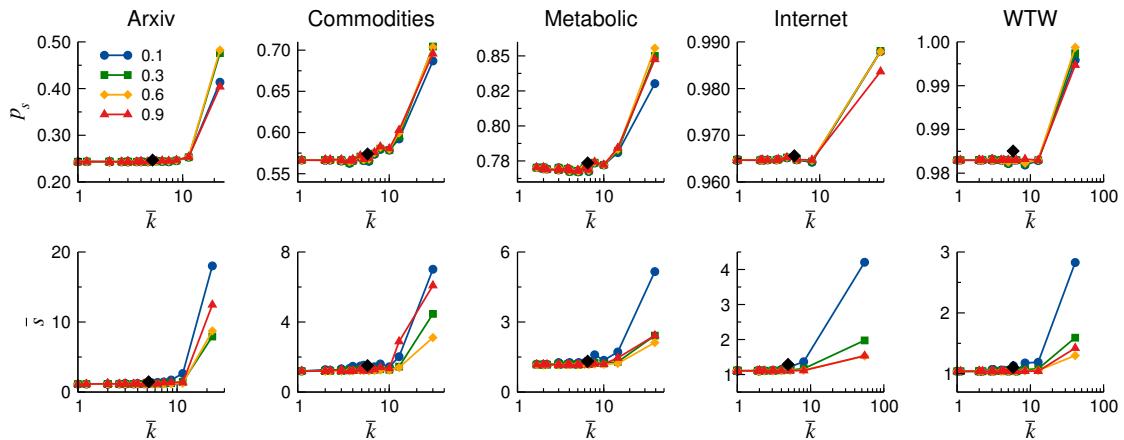


Figure 6.6: Success ratio p_s (top row) and average stretch \bar{s} (bottom row) as a function of the average degree \bar{k} of node bins with temporal behaviour, for different values of the activation probability a . The success ratio is measured in a fully active network, where only nodes from a single bin of size $\xi = 5\%$ can activate with probability $a < 1$. The black diamonds correspond to the case of selecting the number of nodes that activate with $a < 1$ uniformly at random from the whole network and averaging over 10^3 realizations. In this later case, the results for different a 's overlap, so for clarity the black diamonds are displayed only for $a = 0.1$.

characterized by the \bar{k} of the entire network, we find a low p_s close to the static and similar to that obtained for bins of low degree nodes. Therefore, not all nodes are equally able of beating topological traps, and the increase in success mostly relies on activation dynamics affecting densely connected nodes. Also, notice that activating with $a < 1$ some randomly selected nodes is not equivalent to activating the whole network with the corresponding average probability $\bar{a} = 1 - \xi(1 - a)$. This is due to the fact that, while navigating the network, some nodes are visited more often than others. If all visited nodes can activate with some $\bar{a} < 1$, the actual noise affecting the network becomes effectively greater than in our implementation.

When information packets are able to escape cycles, and the success rises due to the emergence of new (lengthier) successful paths, the average stretch increases too. We corroborate this statement in the bottom row of Fig. 6.6, which shows that the maximum stretch precisely occurs at \bar{k}_{\max} in all networks. As expected, the highest \bar{s} is always found for the minimum activation $a = 0.1$, which corresponds to the situation where packets find most nodes along their routes to be inactive so they are constantly redirected, thus producing long greedy paths. Interestingly, at \bar{k}_{\max} the \bar{s} for $a = 0.9$ considerably varies across networks, with the Arxiv exhibiting the highest value, then Commodities, Metabolic, Internet and finally the WTW. This supports the idea that more congruent networks possess less topological traps. In general it is also satisfied that a values around a_O display lower \bar{s} at \bar{k}_{\max} .

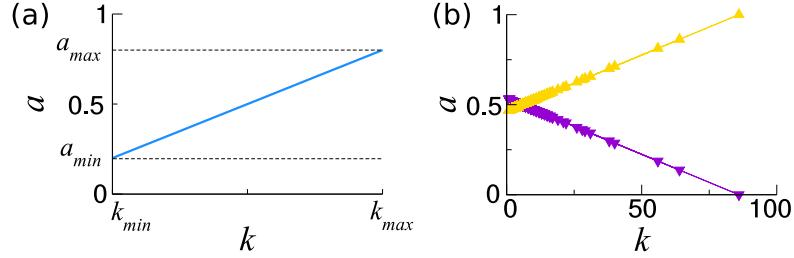


Figure 6.7: (a) In blue, example of a linear activation probability. (b) Example, from Commodities network, of the two curves $a(k)$, used for linearly activating the nodes when $\bar{a} = 0.5$. In yellow $a(k)$ with $b > 0$ and in purple with $b < 0$. The minimum activation probability is explicitly set to not to be exactly 0 but of order 10^{-3} ; symmetrically the maximum activation corresponds to $a \approx 1 - 10^{-3}$.

6.6.2 Linear activation depending on degree

Here, we study the navigability of temporal maps in which the activation probability a_i of a node depends linearly on its degrees k_i , such as

$$a_i(k_i) = bk_i + c. \quad (6.1)$$

For $b > 0$, the activation probability is proportional to the node's degree so the larger the degree the more active is the node; if $b < 0$, the activation probability is inversely proportional to k , and if $b = 0$ we recover the case of constant activation probability.

The average activation probability \bar{a} of the whole network, $\bar{a} = N^{-1} \sum_{i=1}^N a_i(k)$, varies depending on the choices of the coefficients b and c . Given that, we set the average activation probability \bar{a} as an independent parameter, and choose c so that

$$\begin{aligned} \bar{a} &= \frac{1}{N} \sum_{i=1}^N (bk_i + c) \\ c &= \bar{a} - \frac{b}{N} \sum_{i=1}^N k_i \\ c &= \bar{a} - b\bar{k} \end{aligned} \quad . \quad (6.2)$$

The constraints for the coefficient b arise from the network's minimum and maximum degrees, k_{\min} and k_{\max} , which relate to the maximum, a_{\max} , and minimum, a_{\min} , values of the activation probability through the equation of a line, see Fig. 6.7(a). Considering k_{\min} and k_{\max} are fixed attributes of the network, one must satisfy

$$\begin{aligned} \bullet \underline{b > 0}: \quad a_{\max} &= bk_{\max} + c & a_{\min} &= bk_{\min} + c \\ &= bk_{\max} + \bar{a} - b\bar{k} & &= \bar{a} + b(k_{\min} - \bar{k}) \\ &= \bar{a} + b(k_{\max} - \bar{k}) \end{aligned} \quad (6.3)$$

$$\bullet \underline{b < 0}: \quad a_{\max} = \bar{a} - |b|(k_{\min} - \bar{k}); \quad a_{\min} = \bar{a} - |b|(k_{\max} - \bar{k}) \quad (6.4)$$

Taking into account that $a(k)$ is a probability and that therefore it must satisfy $a_{\max} \leq 1$ and $a_{\min} \geq 0$, yields

$$\begin{aligned} \bullet \quad b > 0: \quad & \bar{a} + b(k_{\max} - \bar{k}) \leq 1 \quad \bar{a} + b(k_{\min} - \bar{k}) \geq 0 \\ & b \leq \frac{1 - \bar{a}}{(k_{\max} - \bar{k})} \quad b \leq -\frac{\bar{a}}{(k_{\min} - \bar{k})} \end{aligned} \quad (6.5)$$

$$\begin{aligned} \bullet \quad b < 0: \quad & \bar{a} - |b|(k_{\min} - \bar{k}) \leq 1 \quad \bar{a} - |b|(k_{\max} - \bar{k}) \geq 0 \\ & |b| \leq -\frac{(1 - \bar{a})}{(k_{\min} - \bar{k})} \quad |b| \leq \frac{\bar{a}}{(k_{\max} - \bar{k})} \end{aligned} \quad (6.6)$$

From the above conditions we obtain that b must satisfy

$$b \leq \min \left\{ \frac{1 - \bar{a}}{(k_{\max} - \bar{k})}, \frac{-\bar{a}}{(k_{\min} - \bar{k})} \right\} \quad \text{if } b > 0 \quad (6.7)$$

$$|b| \leq \min \left\{ \frac{-(1 - \bar{a})}{(k_{\min} - \bar{k})}, \frac{\bar{a}}{(k_{\max} - \bar{k})} \right\} \quad \text{if } b < 0. \quad (6.8)$$

Therefore, we can select any b values that satisfy the above inequalities, and afterwards, the coefficient c is readily found from Eq. 6.2. For each network under consideration, we choose two values of the coefficient b (one positive and one negative) that ensure the highest heterogeneity in the activation probability, without completely inactivating any of the nodes, i.e. $a(k_{\min}) > 0$ and $a(k_{\max}) > 0$, see Fig. 6.7(b). In particular, we set the minimum activation probability for the main hub to be $a_{\min} = 10^{-3}$, and symmetrically the maximum to $a_{\max} = 1 - 10^{-3}$.

Figure 6.8 shows the effects of an heterogeneous activation probability of nodes, both proportionally ($b > 0$) and inversely proportional ($b < 0$) to their degree k , compared with constant activation probability ($b = 0$), for $\bar{a} = 0.5$. We observe that when the activation probability is proportional to k , all $\mathcal{M}(\mathcal{G}, S)$ tend to exhibit lower success ratios than in the case with the same \bar{a} but constant activation probability.

This effect can be understood by considering that highly connected nodes are visited more often than the rest during the routing process. Consequently, the system exhibits an effective \bar{a} higher than 0.5, which induces lower p_s values (closer to the static reference) as shown in previous results. The same reasoning explains the observed behaviour of the stretch \bar{s} in bottom row of Fig. 6.8. For all networks and when $b > 0$, the tendency of \bar{s} is comparable to that found for $\bar{a} > 0.5$ in the constant activation case, see bottom row Fig. 6.1. This feature is specially noticeable for the Internet, where \bar{s} growth is similar to the obtained for $\bar{a} = 0.9$.

On the opposite situation, for $b < 0$, when the activation probability gets lower as nodes become more connected, the reverse phenomenon occurs. Figure 6.8 (bottom row)

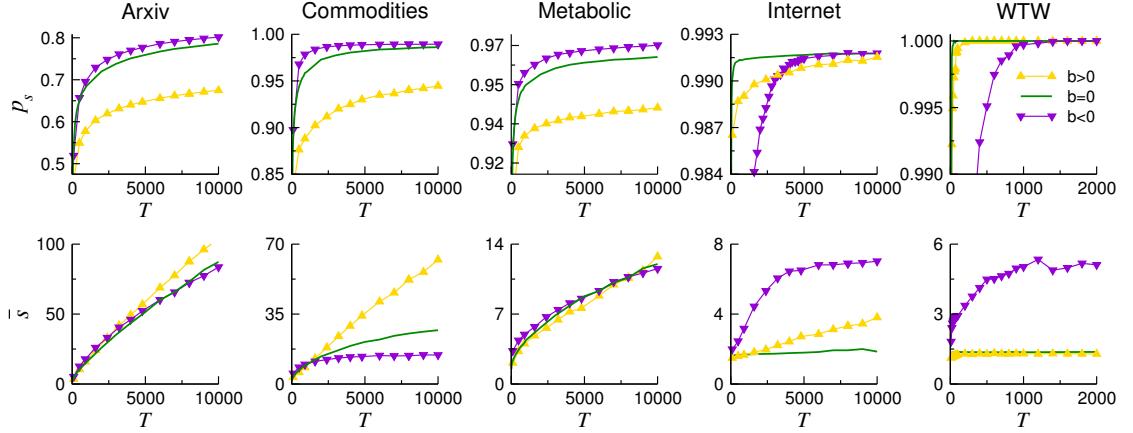


Figure 6.8: Success ratio p_s (top row) and average stretch \bar{s} (bottom row) as a function of T for $\bar{a} = 0.5$. Solid lines designate constant activation of nodes, while symbols indicate that nodes activate linearly and proportionally to their degree ($b > 0$), or inversely proportional ($b < 0$). The minimum and maximum activation probabilities allowed are 10^{-3} and $1 - 10^{-3}$ respectively. As a reference, the values of p_s and \bar{s} for the static networks, $a = 1$, are given in Table 6.1.

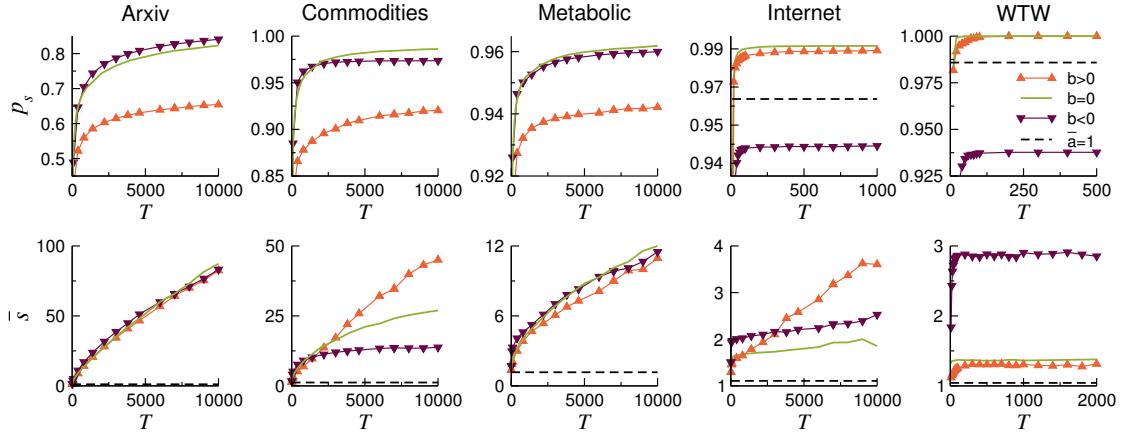


Figure 6.9: Success ratio p_s (top row) and average stretch \bar{s} (bottom row) as a function of T , for $\bar{a} = 0.5$, in five temporal maps $M(G, S)$. In this case, the minimum activation of main hub is allowed to be null and the maximum exactly 1. Solid lines designate constant activation of nodes, while symbols indicate nodes linearly activate following Eq. 6.1; proportionally to their k ($b > 0$), or inversely proportional ($b < 0$). In dashed line, p_s and \bar{s} corresponding to greedy routing in static maps $M(G_0, S)$. Notice the low p_s and \bar{s} achieved by the Internet and the WTW when $b < 0$ due to the main hub being always totally inactive during the routing process.

shows that for all $\mathcal{M}(\mathcal{G}, S)$, the \bar{s} resembles that found for low constant \bar{a} 's. Moreover, a small increase in p_s with respect to $b = 0$ is visible for the Arxiv, the Commodities, and the Metabolic networks in top row of Fig. 6.8. The cause is an effective overall activation $\bar{a} < 0.5$. For the Internet, p_s still grows higher for $b < 0$ than for $b = 0$, though it is barely noticeable and it requires large T . For the WTW, p_s reaches 1 for any b so we can not observe an incremented p_s . However, we note that p_s growth for the WTW is delayed when $b < 0$, just as happens for the Internet. In these two last networks, the influence of the largest hub on the routing performance is remarkable due to their strong hierarchical nature. For this reason, the time T needed for achieving $p_s \sim 1$ is noticeably enlarged in the event that the main hub is poorly activated. An extreme example of this can be observed in Fig. 6.9 when we allow the minimum activation probability to be exactly 0, thus inactivating permanently the main hub during the routing process. In this case, the Internet and WTW networks in particular, show success ratios that even saturate to values $p_s < 1$ regardless of the increasing time T .

6.7 Discussion

In this chapter, our work sets a first attempt to measure the effects of temporal dynamics on the navigability of real networks. Navigability is a primary function in many complex networks that, as we have shown, can be strongly affected by temporal alterations in the activity of nodes.

The interplay between the activation dynamics generating the temporal networks and the greedy routing process, indeed, yields a rich phenomenology. The activation process can be understood as the result of random events, like service failures, or, alternatively, it could be thought as part of a local information transfer protocol applied by the node holding the packet in a static network, so to boost the success of the routing operation at a limited cost. Our results show that, surprisingly, temporal maps can be navigated more efficiently than the corresponding static ones, even though the number of simultaneously available paths to transfer the packets is greatly reduced. Interestingly, the number of successful paths, in which the packet reaches its destination, is increased due to the activation dynamics. This increase in the success ratio p_s comes at the cost of a growth in the stretch \bar{s} , meaning that longer paths are required to successfully deliver the packet. However, the ratio between the success and the stretch, p_s/\bar{s} , shows a non-trivial behavior as a function of the activation probability, unveiling the existence of an optimal value which maximizes the increase in the success and at the same time minimizes the increase in the stretch.

More realistic forms of the activation probability, i.e., when the dynamics only affects a subset of nodes or when the activation is correlated with the degree, show similar results. This analysis uncovered the role of highly connected nodes in the routing process, which are mainly responsible for the larger success ratio achieved in temporal maps. Contrary to expectations, our findings suggest that it is possible to improve the routing performance by switching on and off the hubs of the network more often than the

rest of the nodes. Finally, the navigability of some real networks, like the Internet and the WTW, remains extremely high in the temporal maps. In fact, time-varying effects increase even more the high success rate associated to the static maps, at the cost of a very small increase in the stretch, slowly growing with T , a feature that we name ultranavigability. Even more, temporal changes in the structure of these networks increase the success even if the activation probability is very low. At the same time, the high routing success observed in these networks could be due in part to temporal behavior in the system, although this possibility has not been acknowledged before and all the merit for their navigability properties has been accredited to their static architecture.

It has been increasingly recognized, that networks are dynamic entities that evolve in time, with connections being established and terminated for different reasons. This study paves the way towards a better understanding of the role of the network's temporal dimension in navigation processes through a geometric framework. In particular, it provides hints for developing better routing strategies that exploit the dynamic nature of network topologies. Further research may consider more sophisticated generative models of temporal networks, that may, for instance, incorporate bursty dynamics of links or nodes.

In perspective, the hidden metric space approach has proved useful even for dealing with specially challenging aspects of complex networks, such as the interplay between their functioning and the dynamics of their architecture. At this point, it is clear that the geometric framework offers great advantages and powerful tools to explore both structural elements and dynamical processes on real networks in a unified fashion. Next, in the Conclusions chapter of this thesis we will review the impact we achieved through the network geometric approach and acknowledge potential research venues yet to explore.

7 Conclusions

“Logic is the foundation of the certainty of all the knowledge we acquire.”

LEONHARD EULER

Complex systems, which involve a massive amount of components interacting in nontrivial ways and exhibiting properties such as emergence, feedback-loops and self-organization, permeate our world to an unthinkable extent. In order to understand such pervasive systems in nature, science, technology and business one has to abandon unsuitable reductionist approaches and develop a deep understanding of the networks of interactions that sustain them. In this respect, network science provided the key discovery that the architecture of networks arising in completely different domains are in fact remarkably similar to each other, and thus can be modeled using the same set of mathematical tools. Certain prominent regularities discovered across real-world networks include scale-free degree distributions, sparseness, high clustering, small-worldness and community structure. The fact that real networks share such common properties despite their entirely different origins indicates the existence of general principles governing the generation and evolution of their topologies. In that vein, about twelve years ago, the discovery that clustering in the topology reflects the triangle inequality in a hidden metric space draw a powerful connection between topology and geometry implying that complex networks may be forged through principles beyond the purely topological. The identification of a hidden, or latent, metric space underlying network topologies opened the path to the development of geometric network models that have successfully explained simultaneously all the most ubiquitous features observed in real-world systems.

Nowadays, the field of network geometry encompasses all approaches within the complex networks community that rely on geometric descriptions of network's connectivity. This includes for instance, simplicial complexes [175–177], where models are based on simplices of different dimensions expressing interactions among nodes, as well as the notoriously important $\mathbb{S}^1 / \mathbb{H}^2$ formalism assuming a hidden metric space underlying the structure of complex networks. It is this last two isomorphic geometric models, the \mathbb{S}^1 and the \mathbb{H}^2 that constitute the building blocks of this thesis. The powerful isomorphism existing between the two models guarantees equivalent descriptions of the same ensembles so that we can use them both interchangeably. Both models allow to generate scale-free, small-world, highly clustered, sparse graphs where the heterogeneity in the degree distribution can be controlled independently of the level of clustering. Beyond the remarkable generative capabilities of the $\mathbb{S}^1 / \mathbb{H}^2$ models, they can also be used to infer the hidden metric space of real-world networks through the process of network embedding. The embedding process yields powerful maps with applications ranging from

efficient information routing to community detection. In this thesis, geometric maps play a central role and constitute the underlying thread that weaves the various chapters together. In particular, along the chapters of this thesis we have tried to answer two main research questions. The first: is there a limit to the applicability of geometric maps? Also rephrased as: is the formalism better suited to deal with structural or dynamical problems? And the second question: what specific network properties can we elucidate that are essential to understand real-world systems? In this regard, we chose to tackle four different subjects that we considered of upmost importance, namely: community structure, hierarchical order, multiscale organization and temporal evolution. By studying these, we aimed at accomplishing the main goal of this thesis, which is to exploit the full potential of geometric maps in order to uncover fundamental principles of function and organization of real-world networks.

The ability of geometric maps to abstract complex topologies is specially advantageous since it exposes the community structure of embedded networks as heterogeneous distributions of the angular coordinate. In fact, communities detected in the maps frequently correlate with available metadata. For instance, for the Internet we find that Autonomous Systems close in terms of angular distance tend to belong to the same country. In the case of the brain, geometric communities relate to neuroanatomical modules and for the World Trade Web countries that are either geographically or politically close occupy the same angular sectors. The heterogeneity of the angular distribution of real networks allows us to explore their anatomy – structural properties – using geometric models with homogeneous similarity distributions. This is what we accomplished in Chapter 3, where we designed the Geometric Randomization model (GR), consisting of a rewiring procedure that homogenizes the angular coordinates of the networks while exactly preserves their degree distribution and maximizes their congruency with the hidden metric space. In fact, the GR is akin to a geometric version of the configuration model, where in our case the networks in the ensemble with fixed degrees have clustering but otherwise are maximally random. The GR effectively allows us to compare a real network topology with an analogous counterpart where only the (geometric-) community organization has been discounted. Besides, the use of prescribed degree sequences in the GR proves specially useful as it avoids the estimation of hidden degree variables in order to produce network replicas. This allows for high fidelity reproduction of specific topological characteristics of real networks, which in turn helps in the analysis of features susceptible to degree cutoff fluctuations, like epidemic spreading or synchronization.

Implementing the GR as a null model for community analysis we discovered that geometric communities are invariably meaningful, while in contrast, topological ones are not always reliable and can be detected even in randomized networks where we geometrically deleted communities in similarity space. Unraveling the structural constraints that recreate topological communities despite the absence of geometric ones is a relevant subject for future investigations and opens the door to the development of community detection methods that can bridge the gap between topology and geometry. All things

considered, the GR provides a novel methodology to elucidate structural details that can influence the functioning of networks of any domain that are equipped with geometric information. The GR applicability is not limited to the analysis of community structure, but the model can aid disclosing other distinctive patterns such as hierarchical organization.

Precisely, we used the GR again in Chapter 4 where our focus shifted to the modeling and understanding of real network hierarchies, where these hierarchies incorporate not only ranking but also similarity information. In this chapter, we provided a new geometric definition of hierarchy that overcomes the important challenge of finding an unambiguous stratification scheme for undirected complex networks, where the small world property and the presence of clustering cause graphs to strongly deviate from tree-like. In our enriched interpretation of hierarchy, the control exerted by a higher status individual over a lower status one – status being abstracted by the hidden degree – is stronger when there exists closeness between the two parties, which translates in a small similarity distance in the metric space. Conversely, the strength of hierarchical relations gets dissolved as individuals loose their proximity and become dissimilar. In social terms, one can think for instance of the relationship between the employee that plays golf with the boss compared to that who doesn't. Clearly, the business hierarchy is prone to reflect interactions based on affinity.

Given our augmented definition, the geometric $\mathbb{S}^1 / \mathbb{H}^2$ models stand as a natural framework to exploit the node's attributes of popularity and similarity in order to characterize the contribution of each node and each link to the hierarchical structure of a network. Moreover, our approach allows to exploit the great heterogeneity found in the hierarchy load of links to build a filtering method, the *similarity filter*, that offers a procedure to extract hierarchical similarity backbones (HSBs) that contain the links dominating the hierarchical organization of the network.

We applied our geometric approach to the detection of similarity backbones of real networks from different domains, and found that the backbones preserve local topological features at all scales. This indicates scale invariance in the nested hierarchy of backbones, and therefore it means that the networks exhibit a certain degree of self-similarity. Our HSBs manifest a self-similarity that is not just focalized in a single topological attribute, like the degree distribution, but it has a broader scope applying as well to the clustering and nearest neighbors spectrums. The self-similarity found through the process operated by our similarity filter is thus comparable in nature to that observed through geometric renormalization [26]. In turn, that means we uncovered a global structural property of hierarchical complex networks.

Finally, we explored the role of HSBs in an evolutionary dynamical process which historically has been argued to be sensitive to the hierarchical organization of complex architectures [97, 98]. Implementing an evolutionary prisoner's dilemma game [110–113] on the real networks under study, we discovered that HSBs indeed capture the links that

better support cooperative behavior in evolutionary dynamics. This opens an interesting venue for future work, which is the study of other social dilemma dynamics such as the snowdrift or the stag-hunt games in order discover under which conditions (if any) HSBs do support cooperation. Another engaging possibility is that of using HSBs as minimal sub-structures able to reproduce diffusion properties of a larger network. This could yield to future important applications in resource management for instance. In fact, our work demonstrates the applicability of geometric maps to the study of hierarchies in complex networks of any domain, that meaning the detection of HSBs can help improving from the controlability of technological infrastructures to the robustness to species loss in ecological networks.

Popular mechanisms that often give rise to hierarchies, like competition and imitation between individuals, also guide opinion formation processes in complex networks which strongly impact the functioning of real systems of different domains. In Chapter 5, we introduced a Multiscale Voter Model (MVM) that uses geometric maps to incorporate group influence at different granularities into the opinion formation of an individual, hence adding a realistic but typically neglected component to the process of opinion adoption.

In particular, the probability of copying a neighbor in the MVM is weighted by the distance in opinion between the copycat node and its own group, so that the less aligned is the node's opinion with the opinion of its environment the higher the chances it will copy the viewpoint of some neighbor. Consequently, the MVM reflects the popular tendency of individuals to refrain from adopting behaviours that contradict their group norm. In turn, our results show that this mechanism helps explaining why it can be hard to come to a full consensus even in small-world heterogeneous networks which usually promote fast agreement under the standard voter model.

Specifically, the MVM dynamics on synthetic and on real —political networks, international trade networks, and more— uncovered a transition between a final stage of mixed binary opinions and one of full consensus. This transition is a function of the scale of the groups when those are made of similar nodes and similarity is expressed through the angular separation between them in the \mathbb{S}^1 circle. Our model also revealed that larger groups can sustain for longer mixed opinion configurations while small groups of very affine nodes yield to either very low or very high levels of consensus. This behaviour explains why big groups do not reach unanimity easily, in accordance with real world observations. Remarkably, our numerical simulations point that groups made of more diverse (or less similar) nodes promote eventual consensus and, on the contrary, very affine groups yield to metastable clusters of same opinion which hinder the transition to general agreement.

Interestingly, we found that these clusters create visible spatial patterns in the angular dimension of the hyperbolic embeddings of the networks, akin to the formation of spatial domains in Euclidean lattices. In this sense, the MVM represents a major step forward

in the understanding and visualization of opinion polarization in structured populations. Furthermore, our multiscale framework is appealing for its simplicity and versatility, allowing for many readily implementable extensions. These include for example, studying the influence of the group of the neighbor-node whose strategy is to be copied to address social acceptance. Other possibilities are adding multiscale zealots or including multiple discrete opinions. The MVM can also find application informing conflict resolution in important areas such as Environmental policy where typically groups of different backgrounds, like politicians and scientists, need to come to agreement. Finally, another great extension is to take into account the time-varying nature of network ties which will make the groups and their influence to change over time.

Determining the effect that time-varying topologies have in dynamical phenomena is not an easy endeavour since the modeling of real networks as time-varying is itself an open area of research. To meet this challenge, in Chapter 6 we set the first attempt to measure the effects of temporal dynamics on the navigability of real networks. Navigation is indeed a fundamental function of many natural and artificial complex networks, which exist and evolve to support the transfer of information, mass, or energy. In our work, the temporal component of the networks is modeled through an activation dynamics, which can be understood as the result of random events, like service failures, or, alternatively, as part of a local transfer protocol that boosts the success of the routing operation. This activation process yields to temporal geometric maps which can be navigated using the standard greedy routing protocol in hyperbolic space. Surprisingly, we found that temporal maps can be navigated more efficiently than static ones, even though the number of simultaneously available paths to transfer the packets is greatly reduced. Interestingly, the number of successful paths, in which the packet reaches its destination, is increased due to the activation dynamics although at the cost of elongated routes. Remarkably, we uncovered the existence of an optimal activation value which maximizes the increase in the number of successful paths and at the same time minimizes their stretch.

Additionally, in our study we tested different forms of the activation probability like when the dynamics only affects a subset of nodes or when the activation is correlated with the degree. This analysis revealed that, contrary to expectations, it is possible to improve the routing performance by switching on and off the most connected nodes of the network, or hubs, more often than the rest. Finally, our results uncovered that some real networks are ultranavigable and their navigability is not disrupted even when the network structure is extremely unsteady. Taken together, our findings have important implications for the design and evaluation of efficient routing protocols that account for the temporal nature of real complex networks. Indeed further realism could easily be added to our framework by considering an activation dynamics that includes burstiness of links or nodes. Alternatively, future research may aim at determining the optimal activation probability in networks undergoing other kinds of dynamics such as the spreading of opinions or diseases. This option would allow for an enhanced controllability of unquestionably crucial processes.

Along this thesis we have based all our work in geometric maps complying with the main hypothesis of network geometry which states that the architecture of real complex networks has a geometric origin. This conjecture would have stood as a mere hypothesis if it were not for the considerable amount of evidence supporting the existence of hidden metric spaces accumulated over the years. It is worth emphasizing the numerous applications that, in particular the $\mathbb{S}^1 / \mathbb{H}^2$ formalism, has generated ranging from link prediction methods [178], to renormalization procedures [26], going through the analysis of correlations in multiplexes [179] and unraveling weighted representations of real networks [180]. The future of the geometric approach to network science remains full of opportunities, some awaiting in the quantum realm, others in the extension of paradigms like renormalization to weighted networks, or even just exploring the intricated dynamics of co-evolution of human behaviour and the ecological world. The many insights yet to uncover are only bounded by the imagination of those who venture to exploit the full potential of network geometry.

Resum en català

Els sistemes complexos són aquells que involucren un nombre elevat de components que interaccionen de manera no trivial entre ells, mostrant l'aparició de fenòmens emergents i propietats com l'auto-organització o la retroalimentació. Trobem sistemes complexos de tota mena a la natura, en ciència, tecnologia i fins i tot en el món empresarial, tots ells esperant a ser entesos des d'una lent no reduccionista, basada en l'estudi de les xarxes que els defineixen. En aquest aspecte, el camp de la ciència de xarxes va revolucionar l'estudi dels sistemes complexos. Va revelar que l'arquitectura de les xarxes complexes, malgrat aquestes siguin de dominis molt diferents, comparteixen grans similituds estructurals de manera que la seva modelització és possible mitjançant el mateix conjunt d'eines matemàtiques. Algunes d'aquestes regularitats observades en diverses xarxes reals inclouen: distribucions de grau sense escala, la presència d'un gran nombre de triangles (o *clustering*), la propietat de món petit (*small-worldness*) o l'estructura de comunitats. El fet que les xarxes reals comparteixin aquests trets independentment de la seva naturalesa, indica l'existència de principis fonamentals que governen l'origen i evolució de la topologia de les xarxes. En aquesta línia, al voltant de dotze anys enrere, el descobriment que els triangles mesurats a la topologia d'una xarxa reflecteixen la desigualtat triangular en un espai mètric ocult, va permetre establir una relació entre topologia i geometria sense precedents, que permetria entendre les xarxes com quelcom més que simples objectes topològics. D'aquesta manera, el descobriment dels espais mètrics ocults va portar al desenvolupament de models geomètrics de xarxes, que permetrien crear mapes en un espai hiperbòlic subjacent i finalment explicar totes les propietats més importants observades en les xarxes reals de forma simultània.

A dia d'avui, el camp de la geometria de xarxes engloba tots els enfocaments dins la comunitat de ciència de xarxes que es basen en l'existència d'espais mètrics subjacents per representar la connectivitat. Això inclou per exemple, els models de símplies complexos [175–177] o l'importantíssim formalisme $\mathbb{S}^1 / \mathbb{H}^2$ [25, 42, 43]. Són aquests dos darrers models geomètrics, $\mathbb{S}^1 / \mathbb{H}^2$ i l'isomorfisme que entre ells aplica, que constitueixen els blocs fonamentals d'aquesta tesi. El model \mathbb{S}^1 és un model mixte, en el sentit que combina una component mètrica que representa el sub-espai de similitud entre nodes i una component topològica per expressar la seva popularitat o grau. Concretament, el model unidimensional \mathbb{S}^1 que utilitza aquesta tesi, assumeix el cercle com la geometria del sub-espai mètric, de manera que la similitud entre nodes correspon a la seva distància angular. En el cas del model \mathbb{H}^2 , els nodes es caracteritzen exclusivament mitjançant coordenades en un pla hiperbòlic, de manera que la popularitat i similitud dels nodes queda codificada únicament en la distància (hiperbòlica) que els separa. Aquestes dues forces, popularitat i similitud, determinen la probabilitat de connexió dels nodes, que

prèn la forma d'una llei de gravitació, on la magnitud decreix amb la distància angular o hiperbòlica dependent del model. L'isomorfisme que existeix entre el model S^1 i l' \mathbb{H}^2 garanteix descripcions equivalents de les xarxes complexes, permetent intercanviar-los i explotar un o l'altre dependent del problema o aplicació. Aquests models, a més a més de generar xarxes semblant a les reals, poden utilitzar-se per inferir les coordenades en l'espai mètric ocult de les pròpies xarxes reals. Aquest procés (coneugut com *embedding*) porta a descobrir representacions de les xarxes reals en forma de mapes geomètrics que poden ser utilitzats en aplicacions diverses, des del transport eficient d'informació, fins la detecció de comunitats.

En particular, l'estructura de comunitats és fàcilment observable en els mapes geomètrics en forma de distribucions angulars heterogènies. Això permet explorar amb detall l'anatomia de les xarxes reals mitjançant models geomètrics que, per contra, tinguin distribucions angulars uniformes. D'això mateix tracta el capítol 3, on hem proposat un model geomètric de randomització de xarxes (Geometric Randomization) on les coordenades angulars s'uniformitzen mentre que les connexions entre nodes es redibuiixen garantint que la xarxa preservi la màxima congruència possible amb l'espai mètric i els graus dels nodes en la xarxa original. El model GR ens permet doncs, comparar la topologia d'una xarxa real amb una d'anàloga on només les comunitats geomètriques han estat esborrades. El model no modifica en cap cas la distribució de grau de la xarxa i per tant esdevé molt útil a l'hora de produir rèpliques fidedignes o d'abordar problemes sensibles a les fluctuacions del grau mínim, com ara la propagació d'epidèmies o la sincronització. L'ús del GR com a model nul desvela també una discrepància important entre les comunitats geomètriques i les topològiques, on les darreres no sempre tenen significat real degut a que poden ser detectades malgrat haver esborrat del mapa geomètric les comunitats de la xarxa. L'estudi de les restriccions estructurals que dónen lloc a aquesta discrepància és un tema rellevant per futures investigacions. Tenint en compte tot l'esmentat, el model GR constitueix una eina potent a l'hora d'elucidar detalls estructurals que poden influir el funcionament de les xarxes reals que admeten una representació espacial. Tal com hem explorat també en el capítol 4, la utilitat del GR no es limita a estudiar l'impacte de l'estructura de comunitats, sino que també permet analitzar altres patrons d'organització com ara les jerarquies.

En el capítol 4, hem donat una nova definició de jerarquia en termes geomètrics que supera el repte d'establir un esquema d'estratificació adequat per xarxes reals on les connexions no contenen informació sobre la direccionalitat de les interaccions; i on a més a més la propietat de món petit i l'elevat nombre de triangles fan que els grafs no presentin una organització esglaonada òbvia. La nova definició de jerarquia té en compte que el control exercit per un individu de més alt estatus sobre un subordinat és més rellevant quant més propera és la seva relació, és a dir, quan existeix certa similitud entre ambdós. Al contrari, la jerarquia és menys pronunciada quan el node de més estatus i el de menys són individus distants. Les coordenades de popularitat i similitud del model S^1 , esdevenen doncs el marc idòni per a la nova definició de jerarquia que permet

caracteritzar la contribució a la jerarquia global de cada connexió i de cada node de la xarxa. Aquest enfocament ens ha permès també explotar l'heterogeneitat de la jerarquia de les connexions per definir un sistema de filtrat, capaç d'estreure sub-grafs on les connexions presents dónen una jerarquia cada cop més marcada. Aquests sub-grafs, que anomenem *hierarchical similarity backbones* (*HSBs*), preserven les propietats topològiques de les xarxes reals a totes les escales i a més a més tenen un rol crucial en processos dinàmics susceptibles a l'estructura jeràrquica, com és el cas de l'evolució de dilemes socials. Concretament, trobem que els *HSBs* contenen les connexions de la xarxa que millor sostenen la cooperació en tals processos dinàmics. En aquest capítol hem demonstrat l'aplicabilitat dels mapes geomètrics en l'anàlisi de l'organització jeràrquica de xarxes de qualsevol domini, de tal forma que el seu potencial abarca temes tan distants com el control d'infraestructures tecnològiques fins la prevenció de l'extinció d'espècies en xarxes ecològiques.

Alguns dels mecanismes populars que dónen lloc a l'aparició de jerarquies, com ara la competició i la imitació entre individus, també controlen els processos de formació d'opinió, influint fortament en el funcionament de sistemes reals clau com les organitzacions polítiques o les xarxes de comerç internacional. En el capítol 5, s'introduceix un nou model, *Multiscale Voter Model* (*MVM*), que utilitza els mapes geomètrics per incorporar informació sobre l'opinió grupal de diferents escales en el procés de formació d'opinió individual; afegint així un nivell més de realisme que típicament altres models no aconsegueixen incorporar. Concretament, en el model *MVM* la probabilitat de què un individu de la xarxa adopti l'opinió d'un veí ve pesada segons la distància d'opinió entre el propi individu i el seu grup, tal que si l'individu està molt alineat amb el grup al qual pertany, la probabilitat de copiar el veí és poca i viceversa. D'aquesta manera, el *MVM* reflecteix la tendència popular dels individus de no contradir en excés la norma grupal. Els nostres resultats mostren que aquest mecanisme explica la dificultat d'arribar a consens observada en el món real, inclús en xarxes on la propietat de món petit i l'heterogeneitat de grau típicament afavoreixen una ràpida homogeneització de l'opinió. La dinàmica del *MVM*, en xarxes tant reals com sintètiques, mostra un transició entre un estat final on es mantenen opinions oposades en el sistema, i una altra fase de total unanimitat. La transició depèn de l'escala dels grups i de si els grups contenen nodes que comparteixen similituds, és a dir són propers en el cercle del model \mathbb{S}^1 . El *MVM* revela a més a més que quan es consideren grups grans el sistema sosté configuracions d'opinió mixtes durant temps llargs, mentre que per grups petits i molt afins de seguida obtenim, o bé un molt baix nivell de consens o la total unanimitat. Aquest comportament indica que els grups que incorporen més diversitat (nodes menys similars) promouen el consens, mentre grups molt afins formen dominis metaestables d'igual opinió que dificulten la transició cap a l'acord global. Curiosament, trobem que aquests dominis es poden visualitzar en la dimensió angular del pla hiperbòlic, en analogia als patrons espaials que típicament es formen en quadràtules Euclidianes. En aquest sentit, el *MVM* representa un pas important versus l'enteniment profund de la polarització d'opinions en societats estructurades. L'enfocament multiescala admet, a més a més, diverses modificacions di-

rectament implementables, com ara l'inclusió de grups d'opinió inmutable o l'extensió a múltiples opinions discretes. Un altra possibilitat remarcable és la d'incorporar la dimensió temporal de les connexions de les xarxes, fent que l'influència dels grups fluctuï amb el temps tal com passa en el món real.

Determinar l'efecte que l'evolució temporal de la topologia imparteix en altres processos dinàmics que tenen lloc a la xarxa és una tasca difícil, especialment degut a què la modelització de xarxes temporals és encara un camp d'investigació actiu. Per tal d'afrontar aquest repte, en el capítol 6 s'estableix un primer intent de mesurar l'efecte que la dinàmica temporal té en la navegació de xarxes reals. Per fer-ho, hem modelitzat la component temporal a través d'una dinàmica d'activació dels nodes, que es pot entendre o bé com events aleatòris causats per interrupcions en el sistema, o al contrari, com a part d'un protocol local pensat per millorar l'eficiència del traspàs d'informació. Aquest procés d'activació es tradueix en mapes geomètrics temporals que podem navegar utilitzant protocols estàndards com l'anomenat *greedy routing* en el pla hiperbòlic [162]. Sorprendentment, trobem que els mapes temporals es poden navegar més eficientment que els estàtics malgrat el nombre de rutes disponibles simultàniament per transferir els paquets d'informació es redueix en gran mesura. Curiosament, el nombre de rutes exitoses en les quals el paquet arriba a la destinació creix degut a la dinàmica d'activació temporal, encara que amb el cost afegit d'allargar la seva longitud respecte el cas estàtic. Notablement, trobem que existeix una activació òptima per a la qual es dóna el màxim nombre de rutes exitoses alhora que aquestes s'allarguen mínimament. En aquests capítol s'exploren també altres formes funcionals de la probabilitat d'activació, per exemple quan aquesta correlaciona amb el grau dels nodes. Així, descobrim que contraintuitivament és possible millorar la navegabilitat de les xarxes apagant i encenent els nodes més altament connectats amb una freqüència major que la resta. Els nostres resultats indiquen també que algunes xarxes són ultra-navegables i no perdren aquesta funcionalitat encara que la topologia sigui altament inestable. Considerant tot l'esmentat, els nostres descobriments revelen mecanismes clau a tenir en compte en el disseny i avaluació de protocols de transferència d'informació que incorporin la naturalesa temporal de les xarxes complexes reals.

Al llarg d'aquesta tesi el nostre treball s'ha basat en els mapes geomètrics, els quals obereixen la hipòtesi que l'arquitectura de les xarxes complexes reals es pot explicar a través de la geometria dels espais mètrics subjacents. Aquesta idea hauria romàs com a hipòtesis de no ser per tota l'evidència acumulada al llarg dels anys que soporta l'existència dels espais mètrics ocults. Concretament, val la pena destacar el formalisme S^1 / \mathbb{H}^2 i la diversitat d'aplicacions que ha generat, des de mecanismes de predicció de connexions [178], passant per procediments de renormalització [26], mètodes per l'anàlisis de correlacions en multiplexes [179] o representacions pesades de xarxes reals [180]. El futur del paradigma geomètric presenta encara un munt d'oportunitats, algunes en el reialme de la quàntica, d'altres en l'extensió de marcs teòrics com la renormalització de xarxes pesades, o fins i tot en l'estudi de dinàmiques de co-evolució entre el comporta-

ment humà i el món ecològic. L'última frontera és la imaginació d'aquells que vulguin submergir-se en el món de la geometria de xarxes.

List of publications

Publications related to this thesis

Parts of this thesis were published in the following papers/preprints:

- ◊ *Navigability of temporal networks in hyperbolic space*
E. Ortiz, M. Starnini, M.Á Serrano
Scientific Reports **7** (1), 1-9 (2017)
- ◊ *Geometric randomization of real networks with prescribed degree sequence*
M. Starnini, E. Ortiz, M.Á Serrano
New Journal of Physics **21** (5), 053039 (2019)
- ◊ *Geometric detection of hierarchical backbones in real networks*
E. Ortiz, G. García-Pérez, M.Á Serrano
Physical Review Research **2** (3), 033519 (2020)
- ◊ *Multiscale opinion dynamics on real networks*
E. Ortiz, M.Á Serrano
Preprint, arXiv:2107.06656 (2021)

Other publications

- ◊ *Look who's talking: Two-mode networks as representations of a topic model of New Zealand parliamentary speeches*
B. Curran, K. Higham, E. Ortiz, D. Vasques Filho
PLOS One **13** (6), e0199072 (2018)
- ◊ *Dynamics of new strain emergence on a temporal network*
S. Chakraborty, X.R. Hoffmann, M.G. Leguia, F. Nolet, E. Ortiz, O. Prunas, L. Zavojanni, E. Valdano, C. Poletto
Preprint, arXiv:1805.04343 (2018)

Bibliography

- [1] Y. Bar-Yam. *Dynamics of Complex Systems*. Addison-Wesley, Massachusetts, 1997.
- [2] L. Euler. Solutio problematis ad geometriam situs pertinentis. *Comment Acad Sci U Petrop*, 8:128–40, 1736.
- [3] G. Cimini, T. Squartini, F. Saracco, D. Garlaschelli, A. Gabrielli, and G. Caldarelli. The statistical physics of real-world networks. *Nat Rev Phys*, 1:58–71, 2019.
- [4] Y. Moreno and A.F Pacheco. Synchronization of kuramoto oscillators in scale-free networks. *Europhysics Lett*, 68:603–609, 2004.
- [5] C. Castellano and R. Pastor-Satorras. Competing activation mechanisms in epidemics on networks. *Sci Rep*, 2:371, 2012.
- [6] A.-L. Barabási. *Network Science*. Cambridge University Press, 2017.
- [7] P. Holme. Rare and everywhere: Perspectives on scale-free networks. *Nat Commun*, 10:1016, 2019.
- [8] M.P.H. Stumpf and M.A. Porter. Critical truths about power laws. *Science*, 335:665–666, 2012.
- [9] S. Maslov, K. Sneppen, and A. Zaliznyak. Detection of topological patterns in complex networks: Correlation profile of the internet. *Physica A*, 333:1–4, 2004.
- [10] L.D.F. Costa, F.A. Rodrigues, G. Travieso, and P.R.V. Boas. Characterization of complex networks: A survey of measurements. *Adv Phys*, 56:78, 2005.
- [11] M.A. Porter, J.-P. Onnela, and P.J. Mucha. Communities in networks. *Notices Amer Math Soc*, 56:1082–1097, 2009.
- [12] H. Cherifi, G. Palla, B.K. Szymanski, and X. Lu. On community structure in complex networks: challenges and opportunities. *Appl Netw Sci*, 4:117, 2019.
- [13] S. Fortunato. Community detection in graphs. *Phys Rep*, 3–5:75–174, 486.
- [14] L. Danon, A. Díaz-Guilera, J. Duch, and A. Arenas. Comparing community structure identification. *J Stat Mech*, page P09008, 2005.
- [15] M. Girvan and M.E.J. Newman. Community structure in social and biological networks. *Proc Natl Acad Sci USA*, 12:7821–7826, 2002.

Bibliography

- [16] M.E.J. Newman. Modularity and community structure in networks. *PNAS*, 103:8577–8582, 2006.
- [17] M. Salathé and J.H. Jones. Dynamics and control of diseases in networks with community structure. *PLOS Comput Biol*, 4:1000736, 2010.
- [18] K. Gong, M. Tang, P.M. Hui, H.F. Zhang, D. Younghae, and Y.-C. Lai. An efficient immunization strategy for community networks. *PLOS One*, 12:83489, 2013.
- [19] J. Travers and S. Milgram. An experimental study of the small world problem. *Sociometry*, 4:425–443, 1969.
- [20] D.J. Watts and S.H. Strogatz. Collective dynamics of “small-world” networks. *Nature*, 393:440–442, 1998.
- [21] R. Cohen and S. Havlin. Scale-free networks are ultrasmall. *Phys Rev Lett*, 5:79, 2003.
- [22] C. Song, S. Havlin, and H.A. Makse. Self-similarity of complex networks. *Nature*, 433:392–395, 2005.
- [23] L.K. Gallos, C. Song, and H.A. Makse. A review of fractality and self-similarity in complex networks. *Physica A*, 386:686–691, 2007.
- [24] B.B. Mandelbrot. *The fractal geometry of nature*. W.H. Freeman, 1982.
- [25] M.Á. Serrano, D. Krioukov, and M. Boguñá. Self-similarity of complex networks and hidden metric spaces. *Phys Rev Lett*, 100:07870, 2008.
- [26] G. García-Pérez, M. Boguñá, and M.Á. Serrano. Multiscale unfolding of real networks by geometric renormalization. *Nature Phys*, 14:583–589, 2018.
- [27] P. Erdős and A. Rényi. On random graphs. *Publicationes Mathematicae*, 6:290–297, 1959.
- [28] A.L. Barabási and R. Albert. Emergence of scaling in random networks. *Science*, 15:509–512, 1999.
- [29] G. Bianconi and A.L. Barabási. Competition and multiscaling in evolving networks. *Europhys Lett*, 54:436–442, 2001.
- [30] E.A. Bender and E. Canfield. The asymptotic number of labeled graphs with given degree sequences. *J Comb Theory A*, 24:296–307, 1978.
- [31] B.A. Bollobás. Probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *Eur J Combin*, 1:311–316, 1980.
- [32] J.K. Harris. *An Introduction to Exponential Random Graph Modeling*. SAGE Publications, 2013.

- [33] C. Lee and D.J. Wilkinson. A review of stochastic block models and extensions for graph clustering. *Appl Netw Sci*, 4:122, 2019.
- [34] P. van der Hoorn, G. Lippner, and D. Krioukov. Sparse maximum-entropy random graphs with a given power-law degree distribution. *J Stat Phys*, 173:806–844, 2018.
- [35] M. Barthélemy. Spatial networks. *Physics Reports*, 1-3:1–101, 2011.
- [36] A. Barrat, M. Barthélemy, R. Pastor-Satorras, and A. Vespignani. The architecture of complex weighted networks. *Proc Natl Acad Sci USA*, 11:3747–3752, 2004.
- [37] R. Guimerà, S. Mossa, A. Turtschi, and L.A.N. Amaral. The worldwide air transportation network: Anomalous centrality, community structure, and cities. *Proc Natl Acad Sci USA*, 10:7794–7799, 2005.
- [38] M. Boguñá, F. Papadopoulos, and D. Krioukov. Sustaining the internet with hyperbolic mapping. *Nat Commun*, 1:62, 2010.
- [39] D. Krioukov, F. Papadopoulos, M. Kitsak, A. Vahdat, and M. Boguñá. Hyperbolic geometry of complex networks. *Phys Rev E*, 82:036106, 2010.
- [40] D. Krioukov, F. Papadopoulos, A. Vahdat, and M. Boguñá. Curvature and temperature of complex networks. *Phys Rev E*, 80:035101, 2009.
- [41] M. Boguñá, I. Bonamassa, M. De Domenico, S. Havlin, D. Krioukov, and M.Á. Serrano. Network geometry. *Nat Rev Phys*, 3:114–135, 2021.
- [42] D. Krioukov, F. Papadopoulos, A. Vahdat, and M. Boguñá. Curvature and temperature of complex networks. *Phys Rev E*, 80:035101, 2009.
- [43] D. Krioukov, F. Papadopoulos, M. Kitsak, A. Vahdat, and M. Boguñá. Hyperbolic geometry of complex networks. *Phys Rev E*, 82:036106, 2010.
- [44] M. Boguñá, D. Krioukov, P. Almagro, and M.Á. Serrano. Small worlds and clustering in spatial networks. *Phys Rev Res*, 2:02304, 2020.
- [45] A. Allard and M.Á. Serrano. Navigable maps of structural brain networks across species. *PLOS Comput Biol*, 16:1–20, 2020.
- [46] G. García-Pérez, M. Boguñá, A. Allard, and M.Á. Serrano. The hidden hyperbolic geometry of international trade: World trade atlas 1870–2013. *Sci Rep*, 6:33441, 2016.
- [47] E. Candellero and N. Fountoulakis. Clustering and the hyperbolic geometry of complex networks. *Internet Mathematics*, 1-2:2–53, 2016.
- [48] J. Cannon, W. Floyd, R. Kenyon, and W. Parry. *Flavors of Geometry*. Chapter Hyperbolic Geometry, MSRI, Berkeley, 1997.

- [49] M. Christersson. Hyperbolic tiling tool. <http://www.malinc.se/m/ImageTiling.php>, 2019.
- [50] F. Papadopoulos, C. Psomas, and D. Krioukov. Network mapping by replaying hyperbolic growth. *IEEE/ACM Transactions on Networking*, 23:198–211, 2015.
- [51] F. Papadopoulos, R. Aldecoa, and D. Krioukov. Network geometry inference using common neighbors. *Phys Rev E*, 92:022807, 2015.
- [52] T. Blasius, T. Friedrich, A. Krohmer, and S. Laue. Efficient embedding of scale-free graphs in the hyperbolic plane. *IEEE/ACM Transactions on Networking*, 26:920, 2018.
- [53] G. Alanis-Lobato, P. Mier, and M. Andrade-Navarro. Efficient embedding of complex networks to hyperbolic space via their laplacian. *Sci Rep*, 6:30108, 2016.
- [54] A. Muscoloni, J.M. Thomas, S. Ciucci, G. Bianconi, and C.V. Cannistraci. Machine learning meets complex networks via coalescent embedding in the hyperbolic space. *Nat Commun*, 8:1615, 2017.
- [55] A. Muscoloni and C.V. Cannistraci. Minimum curvilinear automata with similarity attachment for network embedding and link prediction in the hyperbolic space. In *arXiv:1802.01183v1*, 2018.
- [56] G. Alanis-Lobato, P. Mier, and M.A. Andrade-Navarro. Manifold learning and maximum likelihood estimation for hyperbolic network embedding. *Appl Netw Sci*, 1:10, 2016.
- [57] K. Claffy, Y. Hyun, K. Keys, M. Fomenkov, and D. Krioukov. Internet mapping: from art to science. In *CATCH '09, Proc. Int. Conf. on Cybersecurity Applications and Technology Conference for Homeland Security*, page 205–211. IEEE Computer Society, 2009.
- [58] G. García-Pérez, A. Allard, M.Á. Serrano, and M. Boguñá. Mercator: uncovering faithful hyperbolic embeddings of complex networks. *New J Phys*, 21:123033, 2019.
- [59] M. Belkin and P. Niyogi. Laplacian eigenmaps and spectral techniques for embedding and clustering. In *NIPS'01, Proc. Int. Conf. on Neural Information Processing Systems: Natural and Synthetic*, page 585–91. Cambridge, MA. MIT Press., 2002.
- [60] R.B. Lehoucq and D.C. Sorensen. Deflation techniques for an implicitly restarted arnoldi iteration. *SIAM J Matrix Anal Appl*, 17:789, 1996.
- [61] K. Zuev, M. Boguñá, G. Bianconi, and D. Krioukov. Emergence of soft communities from geometric preferential attachment. *Sci Rep*, 5:9421, 2015.
- [62] F. Papadopoulos, M. Kitsak, M.Á. Serrano, M. Boguñá, and D. Krioukov. Popularity versus similarity in growing networks. *Nature*, 489:537–540, 2012.

- [63] G. García-Pérez, M.Á. Serrano, and M. Boguñá. Soft communities in similarity space. *J Stat Phys*, 173:775–78, 2018.
- [64] M. Starnini, E. Ortiz, and M.Á. Serrano. Geometric randomization of real networks with prescribed degree sequence. *New J Phys*, 21:053039, 2019.
- [65] M. Starnini, E. Ortiz, and M.Á. Serrano. Geometric randomization of real networks with prescribed degree sequence. In *arXiv:1812.06471v1*, 2019.
- [66] M. Zheng, A. Allard, P. Hagmann, Y. Alemán-Gómez, and M.Á. Serrano. Geometric renormalization unravels self-similarity of the multiscale human connectome. *PNAS*, 117:20244–20253, 2020.
- [67] S. Maslov and K. Sneppen. Specificity and stability in topology of protein networks. *Science*, 296:910–913, 2002.
- [68] M.Á. Serrano, M. Boguñá, and F. Sagués. Uncovering the hidden geometry behind metabolic networks. *Mol Biosyst*, 8:843–850, 2012.
- [69] V.D. Blondel, J-L. Guillaume, R. Lambiotte, and É. Lefebvre. Fast unfolding of communities in large networks. *J Stat Mech*, 10:P10008, 2008.
- [70] A. Faqeeh, S. Osat, and F. Radicchi. Characterizing the analogy between hyperbolic embedding and community structure of complex networks. *Phys Rev Lett*, 121:098301, 2018.
- [71] T.M. Cover and J.A. Thomas. *Elements of Information Theory*. Wiley-Interscience, New York, 1991.
- [72] M.E.J. Newman. The spread of epidemic disease on networks. *Phys Rev E*, 66:016128, 2002.
- [73] C. Castellano and R. Pastor-Satorras. Thresholds for epidemic spreading in networks. *Phys Rev Lett*, 105:218701, 2010.
- [74] M. Wiedermann, J. Donges, J. Kurths, and R. Donner. Spatial network surrogates for disentangling complex system structure from spatial embedding of nodes. *Phys Rev E*, 93:042308, 2016.
- [75] M. Sarzynska, E.A. Leicht, G. Chowell, and M.A. Porter. Null models for community detection in spatially embedded temporal networks. *J Complex Netw*, 4:363–406, 2015.
- [76] P. Expert, T.S. Evans, V.D. Blondel, and R. Lambiotte. Uncovering space-independent communities in spatial networks. *PNAS*, 19:7663–7668, 2011.
- [77] E. Ortiz, G. García-Pérez, and M.Á. Serrano. Geometric detection of hierarchical backbones in real networks. *Phys Rev Research*, 2:033519, 2020.

- [78] E. Ortiz, G. García-Pérez, and M.Á Serrano. Geometric detection of hierarchical backbones in real networks. In *arXiv:2006.03207v3*, 2020.
- [79] B. Corominas-Murtra, J. Goñi, R.V. Solé, and C. Rodríguez-Caso. On the origins of hierarchy in complex networks. *PNAS*, 110, 2013.
- [80] D. Erwin and E. Davidson. The evolution of hierarchical gene regulatory networks. *Nat Rev Genet*, 10:141–148, 2009.
- [81] B. Johnson and J.R. Carey. Hierarchy and connectedness as determinants of health and longevity in social insects. in "Sociality, Hierarchy, Health: Comparative Biodemography: A Collection of Papers", National Academies Press, pages 269–294, 2014.
- [82] A. Samuels, J.B. Silk, and P. Rodman. Changes in the dominance rank and reproductive behavior of male bonnet macaques (*macaca radiate*). *Anim Behav*, 4:994–1003, 1984.
- [83] P. Saunders. *Social Class and Stratification*. Society now. Routledge, 1990.
- [84] M.A. De Loura. *Game Programming Gems 2*. Charles River Media programming. Charles River Media, 2001.
- [85] D.J. Allerton. *Essentials of Grammatical Theory: A Consensus View of Syntax and Morphology*. Routledge & Kegan Paul, 1979.
- [86] M. Yeston. *The Stratification of Musical Rhythm*. Yale University Press, 1976.
- [87] E. Mones, L. Vicsek, and T. Vicsek. Hierarchy measure for complex networks. *PLOS One*, 7, 2012.
- [88] G. Theraulaz, E. Bonabeau, and J.-L. Deneubourg. Self-organization of hierarchies in animal societies: the case of the primitively eusocial wasppolistes dominuluschrist. *J Theor Biol*, 174:313–323, 1995.
- [89] V.M. Eguíluz, M.G. Zimmermann, C.J. Cela-Conde, and M. San-Miguel. Cooperation and the emergence of role differentiation in the dynamics of social networks. *Am J Sociol*, 110:977–1008, 2005.
- [90] E. Ravasz and A.L. Barabási. Hierarchical organization in complex networks. *Phys Rev E*, 67, 2003.
- [91] L. da Fontoura Costa. The hierarchical backbone of complex networks. *Phys Rev Lett*, 93, 2004.
- [92] T. Nikolaj. Hierarchies in directed networks. In *IEEE International Conference on Data Mining*, pages 991–996, 2015.

- [93] E. Letizia, P. Barucca, and F. Lillo. Resolution of ranking hierarchies in directed networks. *PLOS One*, 13, 2018.
- [94] T. Nikolaj. Dynamic hierarchies in temporal directed networks. *Machine Learning and Knowledge Discovery in Databases, ECML PKDD 2018, Lecture Notes in Computer Science*, 11052, 2019.
- [95] A. Trusina, S. Maslov, P. Minnhagen, and K. Sneppen. Hierarchy measures in complex networks. *Phys Rev Lett*, 92, 2004.
- [96] N. Perra and S. Fortunato. Spectral centrality measures in complex networks. *Phys Rev E*, 78:036107, 2008.
- [97] Herbert A. Simon. The architecture of complexity. volume 106, pages 467–482, 1962.
- [98] P.E. Agre. Hierarchy and history in simon’s “architecture of complexity”. *J Learn Sci*, 3:413–426, 2003.
- [99] R.M. Axelrod. *The Evolution of Cooperation*. Basic Books, New York, 1984.
- [100] M.A. Nowak and R. Highfield. *SuperCooperators: Altruism, Evolution, and Why We Need Each Other to Succeed*. Free Press, New York, 2011.
- [101] P.G. Higgs and N. Lehman. The rna world: molecular cooperation at the origins of life. *Nat Rev Genet*, 16:7–17, 2015.
- [102] M.A. Cant. The role of threats in animal cooperation. *Proc R Soc B*, 278, 2010.
- [103] S. Assenza, J. Gómez-Gardeñes, and V. Latora. Enhancement of cooperation in highly clustered scale-free networks. *Phys Rev E*, 78:017101, 2008.
- [104] F.C. Santos, J.F. Rodrigues, and J.M. Pacheco. Graph topology plays a determinant role in the evolution of cooperation. *Proc Biol Sci*, 273:1582, 2006.
- [105] A. Szolnoki, M. Perc, and Z. Danku. Towards effective payoffs in the prisoner’s dilemma game on scale-free networks. *Phys A*, 387:2075–208, 2008.
- [106] F.C. Santos, J.M. Pacheco, and T. Lenaerts. Evolutionary dynamics of social dilemmas in structured heterogeneous populations. *PNAS*, 103:3490–3494, 2006.
- [107] J. Gómez-Gardeñes, M. Campillo, L.M. Floría, and Y. Moreno. Dynamical organization of cooperation in complex topologies. *Phys Rev Lett*, 98:108103, 2007.
- [108] C. Gracia-Lázaro, A. Ferrer, G. Ruiz, A. Tarancón, J.A. Cuesta, A. Sánchez, and Y. Moreno. Heterogeneous networks do not promote cooperation when humans play a prisoner’s dilemma. *PNAS*, 32:109, 2012.
- [109] K.K. Kleineberg. Metric clusters in evolutionary games on scale-free networks. *Nat Commun*, 8, 2017.

Bibliography

- [110] G. Szabó and G. Fáth. Evolutionary games on graphs. *Phys Rep*, 446:97–216, 2007.
- [111] G. Szabó and C. Töke. Evolutionary prisoner’s dilemma game on a square lattice. *Phys Rev E*, 58:69–73, 1998.
- [112] M.A. Nowak. *Evolutionary Dynamics: Exploring the Equations of Life*. Harvard University Press, Cambridge, M.A., 2006.
- [113] J.W. Weibull. *Evolutionary Game Theory*. MIT Press, Cambridge, M.A., 1998.
- [114] M. Nakamaru, H. Matsuda, and Y. Iwasa. The evolution of cooperation in a lattice-structured population. *J Theor Biol*, 184:65–81, 1997.
- [115] H. Ohtsuki and M.A. Nowak. Evolutionary games on cycles. *Proc R Soc Lond B*, 273:2249–2256, 2006.
- [116] G. Abramson and M. Kuperman. Social games in a social network. *Phys Rev E*, 63:030901, 2001.
- [117] Santos F.C and J.M. Pacheco. Scale-free networks provide a unifying framework for the emergence of cooperation. *Phys Rev Lett*, 95:098104, 2005.
- [118] D. Yang, J.W. Shuai, H. Lin, and C. Wu. Individual’s strategy characterized by local topology conditions in prisoner’s dilemma on scale-free networks. *Physica A*, 388:2750–2756, 2009.
- [119] S. Iyer and T. Killingback. Evolution of cooperation in social dilemmas on complex networks. *PLOS Comput. Biol.*, 12:1–25, 2016.
- [120] H. Ohtsuki, C. Hauert, E. Lieberman, and M.A. Nowak. A simple rule for the evolution of cooperation on graphs and social networks. *Nature*, 441:502–505, 2006.
- [121] M. Ifti, T. Killingback, and M. Doebeli. Effects of neighbourhood size and connectivity on spatial continuous prisoner’s dilemma. *J Theor Biol*, 231:97–106, 2004.
- [122] C.L. Tang, W.X. Wang, X. Wu, and B.H. Wang. Effects of average degree on cooperation in networked evolutionary games. *Eur Phys J B*, 53:411–415, 2006.
- [123] R. Cressman and Y. Tao. The replicator equation and other game dynamics. *Proc R Soc Lond*, 111:10810–10817, 2014.
- [124] D. Helbing. A stochastic behavioral model and a microscopic foundation of evolutionary game theory. *Theor Decis*, 40:149–179, 1996.
- [125] J. Poncela. *Evolutionary Games in Complex Topologies: Interplay Between Structure and Dynamics*. Springer Science and Business Media, Heidelberg, Germany, 2012.

- [126] R. Cohen, K. Erez, D. ben Avraham, and S. Havlin. Resilience of the internet to random breakdowns. *Phys Rev Lett*, 85:4626–4628, 2000.
- [127] D.S. Callaway, M.E.J. Newman, S.H. Strogatz, and D.J. Watts. Network robustness and fragility: Percolation on random graphs. *Phys Rev Lett*, 85:5468–5471, 2000.
- [128] F. Nielsen. *Introduction to HPC with MPI for Data Science*. Chapter 8: hierarchical clustering, Springer, 2016.
- [129] E. Ortiz and M.Á Serrano. Multiscale opinion dynamics on real networks. In *arXiv:2107.06656*, 2021.
- [130] P.L. Krapivsky. Kinetics of monomer-monomer surface catalytic reactions. *Phys Rev A*, 45:1067, 1992.
- [131] L. Frachebourg and L. Krapivsky. Exact results for kinetics of catalytic reactions. *Phys Rev E*, 53, 1996.
- [132] C. Castellano, D. Vilone, and A. Vespignani. Incomplete ordering of the voter model on small-world networks. *EPL*, 63:153, 2003.
- [133] D. Vilone and C. Castellano. Solution of voter model dynamics on annealed small-world networks. *Phys Rev E*, 69:016109, 2004.
- [134] V. Sood, T. Antal, and S. Redner. Voter models on heterogeneous networks. *Phys Rev E*, 77:041121, 2008.
- [135] M. Mobilia. Does a single zealot affect an infinite group of voters? *Phys Rev Lett*, 91:028701, 2003.
- [136] M. Mobilia, A. Petersen, and S. Redner. On the role of zealotry in the voter model. *J Stat Mech*, page P08029, 2007.
- [137] R. Hegselmann and U. Krause. Opinion dynamics and bounded confidence models, analysis and simulation. *J Artif Soc Soc Simul*, 5, 2002.
- [138] N.G.F. Medeiros, A.T.C. Silva, and F.G.B. Moreira. Domain motion in the voter model with noise. *Phys Rev E*, 73:046120, 2006.
- [139] L. Dall’Asta and C. Castellano. Effective surface-tension in the noise-reduced voter model. *EPL*, 77:60005, 2007.
- [140] F. Baumann, P. Lorenz-Spreen, I.M. Sokolov, and M. Starnini. Modeling echo chambers and polarization dynamics in social networks. *Phys Rev Lett*, 124:048301, 2020.
- [141] E.A. Hobson, D. Mønster, and S. DeDeo. Aggression heuristics underlie animal dominance hierarchies and provide evidence of group-level social information. *PNAS*, 9:118, 2021.

Bibliography

- [142] M. Diakonova, V. Nicosia, V. Latora, and M. San-Miguel. Irreducibility of multilayer network dynamics: the case of the voter model. *New J Phys*, 18:023010, 2016.
- [143] N.E. Amato, R. Kouvaris, M. San-Miguel, and A. Díaz-Guilera. Opinion competition dynamics on multiplex networks. *New J Phys*, 19:123019, 2017.
- [144] D. Tsarev, A. Trofimova, A. Alodjants, and A. Khrennikov. Phase transitions, collective emotions and decision-making problem in heterogeneous social systems. *Sci Rep*, 9:18039, 2019.
- [145] K. Suchecki, V.M. Eguíluz, and M. San-Miguel. Conservation laws for the voter model in complex networks. *EPL*, 69:228, 2005.
- [146] P. Matjaž. Phase transitions in models of human cooperation. *Phys Lett A*, 380:2803–2808, 2016.
- [147] J. Tanimoto. *Evolutionary Games with Sociophysics: Analysis of Traffic Flow and Epidemics*. Evolutionary economics and social complexity science. Springer, 2018.
- [148] A. Traulsen, M.A. Nowak, and J.M. Pacheco. Stochastic dynamics of invasion and fixation. *Phys Rev E*, 74:011909, 2006.
- [149] L.E. Blume. The statistical mechanics of strategic interaction. *Games and Economic Behavior*, 5(3):387–424, 1993.
- [150] P. Colomer-de Simón and M. Boguñá. Double percolation phase transition in clustered complex networks. *PRX*, 4:041020, 2014.
- [151] B. Curran, K. Higham, E. Ortiz, and D. Vasques-Filho. Look who's talking: Two-mode networks as representations of a topic model of new zealand parliamentary speeches. *PLOS One*, 13:e0199072, 2018.
- [152] A.L. Traud, E.D Kelsic, P.J. Mucha, and M.A. Porter. Comparing community structure to characteristics in online collegiate social networks. *SIAM Rev.*, 53(3):526–543, 2011.
- [153] S. Gazda, S. Iyer, T. Killingback, R. Connor, and S. Brault. The importance of delineating networks by activity type in bottlenose dolphins (*tursiops truncatus*) in cedar key, florida. *R Soc Open Sci*, 2:140263, 2015.
- [154] J. Balagna, C.R. Williams, J. Wang, S. Burch, E. Dalton, J. Kirchick, and P. Sosa. Consensus-driven approach for decision-making in diverse groups. *AJPH*, 110:5, 2020.
- [155] E. Ortiz, M. Starnini, and M.Á. Serrano. Navigability of temporal networks in hyperbolic space. *Sci Rep*, 7:15054, 2017.

- [156] E. Ortiz, M. Starnini, and M.Á. Serrano. Navigability of temporal networks in hyperbolic space. In *arXiv:1709.02623v1*, 2017.
- [157] G. Tkačik, C.G. Callan, and W. Bialek. Information flow and optimization in transcriptional regulation. *PNAS*, 105:12265–12270, 2008.
- [158] D. Papo, J.M. Buldú, S. Boccaletti, and E.T. Bullmore. Complex network theory and the brain. *Philos Trans R Soc B*, 369:1653, 2014.
- [159] J. Zhang. Greedy forwarding for mobile social networks embedded in hyperbolic spaces. In *ACM SIGCOMM 2013 Conference, Hong-Kong, China*, 2013.
- [160] M. Boguñá, D. Krioukov, and K.C. Claffy. Navigability of complex networks. *Nat Phys*, 16:116, 2009.
- [161] M. Boguñá and D. Krioukov. Navigating ultrasmall worlds in ultrashort time. *Phys Rev Lett*, 102:058701, 2009.
- [162] F. Papadopoulos, D. Krioukov, M. Boguñá, and A. Vahdat. Greedy forwarding in dynamic scale-free networks embedded in hyperbolic metric spaces. *INFOCOM*, 16, 2010.
- [163] A. Corral. Long-term clustering, scaling, and universality in the temporal occurrence of earthquakes. *Phys Rev Lett*, 92:108501, 2004.
- [164] W.H. Thompson, P. Brantefors, and P. Fransson. From static to temporal network theory - applications to functional brain connectivity. In *bioRxiv*, 2016.
- [165] K.-I. Goh and A.L. Barabási. Burstiness and memory in complex systems. *Euro-physics Letters*, 81:48002, 2008.
- [166] H.H. Jo, R.K. Pan, and K. Kaski. Emergence of bursts and communities in evolving weighted networks. *PLOS One*, 6:e22687, 2011.
- [167] A.L. Barabási. The origin of bursts and heavy tails in human dynamics. *Nature*, 435:207–211, 2005.
- [168] N. Perra, B. Gonçalves, R. Pastor-Satorras, and A. Vespignani. Activity driven modeling of time varying networks. *Sci Rep*, 2:469, 2012.
- [169] P. Holme and J. Saramäki. *Temporal networks*. Springer, Berlin, 2013.
- [170] I. Scholtes, N. Wider, R. Pfitzner, A. Garas, C.J. Tessone, and F. Schweitzer. Causality-driven slow-down and speed-up of diffusion in non-markovian temporal networks. *Nat Commun*, 5, 2014.
- [171] B. George, S. Kim, and S. Shekhar. Spatio-temporal network databases and routing algorithms: A summary of results. In *SSTD'07, Proceedings of the 10th International Conference on Advances in Spatial and Temporal Databases*, Springer-Verlag, Berlin, 2007.

Bibliography

- [172] C. Kirst, M. Timme, and D. Battaglia. Dynamic information routing in complex networks. *Nat Commun*, 7:11061, 2016.
- [173] H.H. Jo, M. Karsai, J. Kertész, and K. Kaski. Circadian pattern and burstiness in mobile phone communication. *New J Phys*, 14, 2012.
- [174] M. Newman. *Networks: An Introduction*. Oxford Scholarship. Oxford University Press, 2010.
- [175] Z. Wu, G. Menichetti, C. Rahmede, and G. Bianconi. Emergent complex network geometry. *Sci Rep*, 5:10073, 2015.
- [176] G. Bianconi and C. Rahmede. Complex quantum network manifolds in dimension $d > 2$ are scale-free. *Sci Rep*, 5:13979, 2015.
- [177] G. Bianconi and C. Rahmede. Network geometry with flavor: From complexity to quantum geometry. *Phys Rev E*, 93:032315, 2016.
- [178] G. García-Pérez, R.G. Aliakbarisani, and M.Á. Serrano. Precision as a measure of predictability of missing links in real networks. *Phys Rev E*, 101:052318, 2020.
- [179] K. Kleineberg, M. Boguñá, and M.Á. Serrano. Hidden geometric correlations in real multiplex networks. *Nat Phys*, 12:1076–1081, 2016.
- [180] A. Allard, M.Á. Serrano, G. García-Pérez, and M. Boguñá. The geometric nature of weights in real complex networks. *Nat Commun*, 8:14103, 2017.
- [181] M. Domenico, A. Lancichinetti, A. Arenas, and M. Rosvall. Identifying modular flows on multilayer networks reveals highly overlapping organization in interconnected systems. *Phys Rev X*, 5, 2015.
- [182] R.A. Rossi and N.K. Ahmed. The network data repository with interactive graph analytics and visualization. In *AAAI*, 2015.
- [183] B. Klimt and Y. Yang. Introducing the enron corpus. In *CEAS*, 2004.
- [184] J. Leskovec, K.J. Lang, A. Dasgupta, and M.W. Mahoney. Community structure in large networks: natural cluster sizes and the absence of large well defined clusters. *Internet Mathematics*, 6:29, 2009.
- [185] C.E. Priebe, J.M. Conroy, D.J. Marchette, and Y. Park. Scan statistics on enron graphs. *Comput Math Organ Theory*, 11:229–247, 2005.
- [186] J. Schellenberger, J.O. Park, T.M. Conrad, and B.Ø. Palsson. Bigg: a biochemical genetic and genomic knowledgebase of large scale metabolic reconstructions. <http://bigg.ucsd.edu/>, 2010.
- [187] J. Serrà, A. Corral, M. Boguñá, M. Haro, and Arcos J.L. Measuring the evolution of contemporary western popular music. *Sci Rep*, 2:521, 2012.

- [188] D. Grady, C. Thiemann, and D. Brockmann. Robust classification of salient links in complex networks. *Nat Commun*, 3:864, 2012.
- [189] R. Milo, S. Itzkovitz, N. Kashtan, R. Levitt, I. Shen-Orr, I. Ayzenshtat, M. Sheffer, and U. Alon. Superfamilies of evolved and designed networks. *Science*, 303:1538, 2004.

Appendix

A Empirical data sets

In the following, we describe the real-world networks studied in this thesis and provide references to their corresponding sources. The data sets are presented in alphabetic order and a table summarizing their main topological properties is included at the end.

- **Arxiv.** This network is a graph representing co-authorship of papers [181], elaborated from data of the free scientific repository arXiv. The nodes are authors which are connected if they have co-authored a paper belonging to category “Disordered Systems and Neural Networks” (cond-mat.disnn). The data considers only papers from up to May 2014, with the word “networks” in the title or abstract. The network used in this thesis comes from Ref. [179].
- **Dolphins.** This is a social network of bottlenose dolphins from Cedar Key, Florida. Each link between two dolphins represents spatial proximity as surveyed in prey capture events or persistent incidents of prey searching. The original network is from Ref. [153] but we use the curated version provided in repository [182].
- **Enron Email.** This is the network of email messaging activity within the Enron company as structured in email threads in Refs. [183, 184]. Emails were considered to be in the same thread if they contained the same words in their subjects and were among the same users. We used the network topology provided in Ref. [26].
- **Enron Social.** The network represents social connections among employees of the Enron company. Edges are established between email addresses of employees that shared correspondence. We use the processed dataset provided in Ref. [185] which includes metadata about the organizational roles of 130 users. The data set is available at <http://www.cis.jhu.edu/~parky/Enron/>.
- **Facebook.** An online social network extracted from Facebook where nodes consist of people (Caltech university students) with edges representing friendship ties. We use the network as provided in Ref. [152].
- **Internet A.S.** The IPv6 Autonomous Systems (AS) Internet topology was extracted from the data collected in June 2009 by the Archipelago project developed by CAIDA [57], and reconstructed as a network in Ref. [38]. The connections are not physical but logical, representing Autonomous Systems (AS) relationships. An AS is a part of the Internet infrastructure administrated by a single company or organization. Pairs of ASs peer to exchange traffic. These peering relationships in the AS topology are represented as links between AS nodes. CAIDA’s IPv6 data

sets provide regular snapshots of AS links derived from ongoing traceroute-based IP-level topology measurements.

- **Metabolic E. Coli.**. This network is a one-mode projection of metabolites of the bipartite metabolic network of the bacterium E. Coli. In this representation, two metabolites are connected if they participate in the same biochemical reaction. In the hyperbolic map of the network distance indicates biochemical affinity. In this thesis we use the data originally extracted from the BiGG database [186] and reconstructed in [68] as a spatially embedded network.
- **Metabolic Human.** This network is the one-mode projection of metabolites of the bipartite metabolic network of human cell metabolisms. Links between metabolites are established when they participate in the same biochemical reaction. The network was first reconstructed and spatially embedded in Ref. [68].
- **Music.** In this network nodes are chords—sets of musical notes played in a single beat and links represent observed transitions among them. The construction process of the network is described in Ref. [187]. In this thesis we use the topology of a sparser and undirected version of such network as reconstructed, and spatially embedded, in Ref. [26].
- **NZ-MPs.** This network is the one-mode projection onto Members of Parliament (MPs) of the bipartite political network of New Zealand parliamentary speeches of 48th parliament (years 2005-2008). Links between MPs are established when they participated in discussions about the same topic. The network was first reconstructed in Ref. [151].
- **US Commodities.** This network [188] is a representation of the flows of services and goods exchanged (in USD) between industrial sectors of the United States during year 2007. The network was obtained from Ref. [180].
- **Words.** This is the network of adjacency between words in the book *The Origin of Species* by Darwin, obtained from Ref. [189]. We use the spatially embedded version presented in Ref. [26].
- **World Trade Web (WTW).** This network consists of significant bilateral trade exchanges between countries. A link represents an active trade channel between two countries in the world. We use the network corresponding to the most recently curated data (year 2013), as given in Ref. [46]. The WTW is constructed from public databases – such as the International Monetary Fund (IMF), the United Nations Commodity Trade Statistics Database or The Correlates of War Project (COW)– which report import and export exchanges of merchandises between pairs of countries in the world.

Data set	Type	Nodes	N_{gcc}	E_{gcc}	k_{max}	$\langle k \rangle$	$\langle c \rangle$
ArXiv	Social	Authors	2121	5473	70	5.16	0.76
Dolphins	Animal Social	Dolphins	149	1553	69	20.84	0.80
Enron Email	Communication	Email threads	33696	180811	1383	10.73	0.71
Enron Social	Social	Employees	182	2097	109	23.04	0.50
Facebook	Online Social	Friends	762	16651	248	43.70	0.43
Internet AS	Tech	A.S.	23748	58414	2778	4.92	0.61
Metabolic E.Coli	Bio/Cell	Metabolites	1008	3285	143	6.51	0.48
Metabolic Human	Bio/Cell	Metabolites	1436	4718	224	6.57	0.54
Music	Language	Chords	2476	20624	1566	16.66	0.82
NZ-MPs	Political	Politicians	94	662	35	14.09	0.78
US Commodities	Economic	US indust. sect.	374	1090	86	5.83	0.22
Words	Language	Words	7377	44205	2568	11.98	0.47
World Trade Web	Economic	Countries	189	550	110	5.82	0.63

Table A.1: Main topological properties of the empirical data sets used in this thesis: N_{gcc} , number of nodes in the giant connected component of the network; E number of edges in the giant connected component; k_{max} , highest degree; $\langle k \rangle$, average degree and $\langle c \rangle$, mean clustering coefficient. We report the number of nodes and edges of the gcc 's of networks because in all chapters we mainly work with their giant connected components.