UNIVERSITAT DE BARCELONA

A geometric approach to the structure of complex networks

Guillermo García Pérez

PhD Thesis

A geometric approach to the structure of complex networks

Author

Guillermo García Pérez

Advisors

Prof. María Ángeles Serrano

Dr. Marián Boguñá

Universitat de Barcelona

A geometric approach to the structure of complex networks

MEMÒRIA PRESENTADA PER OPTAR AL GRAU DE DOCTOR PER LA UNIVERSITAT DE BARCELONA

PROGRAMA DE DOCTORAT EN FÍSICA

Autor

GUILLERMO GARCÍA PÉREZ

Directors

PROF. MARÍA ÁNGELES SERRANO

DR. MARIÁN BOGUÑÁ

Tutor

DR. GIANCARLO FRANZESE

UNIVERSITAT DE BARCELONA

A mis padres

Acknowledgements

I have many reasons to be grateful to my advisors, M. Ángeles Serrano and Marián Boguñá: for the trust you put in me with a fellowship, giving me the opportunity to start a scientific career; for all the hours and selfless effort that you have dedicated to patiently guide and mentor me, always finding the time to discuss with me, and for everything that I have learnt from your expertise as a result; for the comfortable and active working environment—in which we have had more research ideas than we could possibly tackle—that has motivated me throughout these years; and for being always supportive when times were difficult, either personally or professionally. Thank you, I am very lucky to have studied my PhD with you, and I hope that we continue to work together.

I am also very thankful to Antoine Allard, with whom I have collaborated much these years. Antoine, it has been great working with you, and thanks for everything you have taught me. I also want to thank Francesco, Michele, Muhua, Nikos and Roya for the good time working together.

I would like to thank Profs. Sabrina Maniscalco and Jyrki Piilo for allowing me to stay at the Turku Quantum Technology group as a visiting PhD student. I am also grateful to them, as well as to Johannes Nokkala, for our collaboration: I really enjoyed it. I had a great time in Finland, and for that I must thank everyone at the UTU Theoretical Physics Department.

Thanks also to my colleagues at UB for the good times and the interesting discussions. I specially thank Elisenda, Jan, Michele, Oleguer, Pol and Xavi for their valuable advice and comments. I also acknowledge Albert Díaz for managing the ClabB as well as many group activities.

It has been quite stimulating to witness how Quadrívium, which started as a few of us giving some outreach talks every now and then, has grown during these years. I must acknowledge Blai G., Blai P., Gian, Matt and Nahuel for it.

I am deeply thankful to Marta, probably the person who knows me best, for always helping me to pursue my goals, even in the hardest moments. I also want to thank my friends, especially Elisa, for the scientific conversations and for their help whenever I needed it during this period.

Finally, I want to express my gratitude to my family, especially to my parents Margarita Pérez and Mario García, who encouraged me to return to my studies many years ago and unconditionally supported me along the path. I owe you many things, but I take this occasion to thank you for helping me to find out what I really wanted to do and to make it possible.

Contents

CHAPTER 1

Introduction

1.1 Complex networks

The last three centuries have witnessed an explosive development of scientific knowledge. In 1687, Sir Isaac Newton published the *Principia* [120], a book in which he introduced the fundamental laws that govern the motion of bodies. His important discoveries allowed, for the first time in human history, to explain and predict in precise mathematical terms the motion of any body using three simple principles. Even more strikingly, not only did Newton explain how bodies move when forces act upon them, but also how bodies exert gravitational forces one onto another. He even devised a beautifully simple thought experiment some years later [121], depicted in Fig. 1.1, which illustrates how his discoveries unify the motion of celestial bodies, ever before regarded as divine entities, and the motion of regular, mundane earthly bodies. Newton imagined a cannon, the power of which can be regulated, located at the summit of a mountain; as the velocity with which the cannonballs are shot parallel to the ground increases, the trajectories described by them become further-reaching until, for some critical initial speed, they orbit the planet, just like the Moon orbits the Earth. In other words, the orbit of the Moon around our planet was no longer to be considered different in nature from that of an everyday object falling to the ground.

Not surprisingly, the great success of being able to reduce so many seemingly diverse phenomena to a few principles paved the way to *reductionism*, which has ever since been doubtlessly fruitful in physics. This approach is guided by the idea that it is sufficient to understand the elementary processes taking part in a system in order to explain its observed behaviour. Nevertheless, and despite the long list of achievements of reductionism—and the ones still to come—, it might not always lead to a complete understanding of Nature. While it is certainly true that many systems can be analyzed in terms of the basic rules governing each of their parts—like, for instance, the macroscopic behaviour of an ideal gas can be derived from the assumption that its constituent atoms obey Newton's laws of motion—, that might be possible for some particular systems only. For example, it is hard to imagine how one could explain human behaviour from first physical principles, even though all processes taking place in a human body are driven by them.

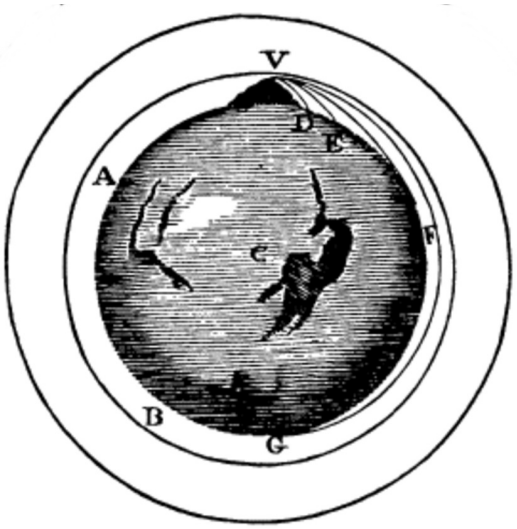


Figure 1.1: **Netwon's cannonball.** Thought experiment devised by Newton to illustrate how an attractive force between massive objects explains both the falling of objects and planetary orbits. The cannonballs are shot parallel to the ground from the summit of the mountain. In the absence of air, they become satellites provided sufficient initial velocity.

In the systems that we call *complex*, which exhibit behaviours that are not expected from the driving principles applicable to their elements, the interaction patterns between its constituent units are—at least partly—responsible for these emergent phenomena and, therefore, they need to be accounted for as well. In other words, besides knowing the laws governing each part of the system, it is also imperative to know how these different parts interact, which forces us to drop the reductionist approach and turn towards a holistic description. This is where complex networks become essential.

Complex networks [118, 50, 57] represent the interaction patterns of complex systems in terms of graphs, mathematical objects conceived by the great mathematician Leonhard Euler [62]. They are composed of nodes, or vertices, representing distinct parts of a system, and edges, or links, representing their pairwise interactions. Clearly, such an abstract description can be applied to a vast variety of systems, and they are indeed used in many different fields of knowledge, like sociology, biology, technology, mathematics or physics, to name a few. For instance, social networks usually represent interactions of some sort—like friendship or collaboration—between people, whereas biological networks range from descriptions of metabolisms to interactions between neurons. In any case, the generality of the concept of network and the importance of considering interactions in many fields make complex networks an appealing approach to

many sciences which, in turn, makes network science one of the most transversal lines of research nowadays.

During the last 20 years, since the first papers on network theory were published, many real networks have been measured, which has allowed researchers to characterize their topologies. Remarkably, although they are far from being regular lattice-like structures or purely random and unstructured graphs, it has become clear that there are some ubiquitous features exhibited by networks from completely different domains [40, 166].

1.1.1 The universal properties of real networks

Complex networks are usually classified according to the nature of their links. Networks are called *simple* if their links are binary, that is, if they do not have any attributes other than their existence, whereas they can be *weighted* if there is an intensity assigned to every link, or *directed* if the links have a directionality. Mathematically, a network is usually represented in terms of its adjacency matrix A, whose elements $a_{ij} \equiv (A)_{ij}$ contain the information regarding the connections between every pair of nodes i and j. Hence, for simple networks, $a_{ij} = a_{ji} = 1$ if both nodes are connected and $a_{ij} = a_{ji} = 0$ otherwise. In directed networks, however, the adjacency matrix is not symmetric, while in weighted networks, the adjacency matrix elements are positive real numbers. In the latter case, it is frequent to use the notation ω_{ij} instead of a_{ij}.

In this thesis, we mainly focus on simple and weighted networks, so we restrict our account of network properties in this section to these two classes of graphs. We first review the main topological features, which are only related to the existence of links and, secondly, the weighted properties, which are defined for weighted networks only.

1.1.1.1 Topological properties

The first feature of real networks to be identified was degree heterogeneity [56, 11, 63, 87]. The degree of a node i, typically referred to as k_i, is the number of links reaching it, that is, its number of neighbours. In real networks, this quantity is very heterogeneously distributed, since most nodes have very low degree, while a very small amount of nodes are connected to a macroscopic fraction of the system. More precisely, the degree distribution, defined as the probability for a randomly chosen node in the network to have k neighbours, scales as a power-law

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

where the exponent γ, which measures degree heterogeneity, is usually found to be between 2 and 3 (see Fig. 1.2**a**). Therefore, all moments higher than 1 in

Eq. (1.1) diverge as the size of the network $N \to \infty$, which implies that, even if the average degree is well defined (finite), it does not represent a measure of typical degrees in the network, since fluctuations are very large (infinite in the thermodynamic limit). This is the reason why real-world networks are called *scale-free networks*. Real networks are also *sparse*, meaning that their average degrees are small as compared to the system size[1].

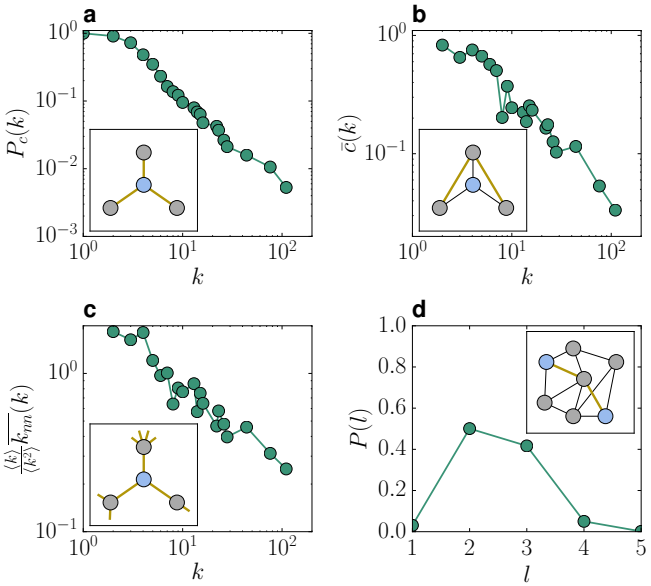


Figure 1.2: **Topological properties of a real network.** The green curves represent different features of the World Trade Web network in 2013 (see Chapter 3 for details). The insets illustrate the different quantities under scrutiny. **a**, Complementary cumulative degree distribution $P_c(k) = \sum_{k'=k}^{\infty} P(k')$ 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, that is, average clustering computed over nodes of the same degree. As an example, the blue node in the sketch has $c = 2/3$. **c**, Normalized average nearest neighbours degree averaged over all nodes in every degree class. The normalization factor $\langle k^2 \rangle / \langle k \rangle$ leaves this curve fluctuating around one for uncorrelated networks. In this case, however, the network shows a clear disassortativity. The blue node in the inset has average nearest neighbours degree 10/3. **d**, Distribution of shortest path lengths. Despite the size of the network, $N = 189$ nodes, the most likely distance is $l = 2$, whereas very few pairs are separated by more than 4 hops. In the sketch, the two blue nodes lay at distance $l = 2$.

[1]The definition of sparseness is more precise in the context of network models, in which network size N can be controlled. We say that a model generates sparse graphs if their average degree is independent of N.

Beyond the degrees of single nodes, some structures involving more nodes are also commonly found in real networks. In particular, the presence of triangles, or transitive connections, is much higher in real networks than predicted by most network models [111], as we will see in the next section. Triangles can be quantified in networks through the so-called *clustering coefficient*. Given a node i with degree $k_i > 1$, its clustering coefficient is defined as the probability for two randomly chosen neighbours of i to be connected. This quantity is computed as

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

where n_i is the number of connections among neighbours of i and the factor $\frac{1}{2}k_i\,(k_i-1)$ accounts for the maximum possible number of links among the neighbours of i. Clustering is often represented by its spectrum, that is, the average clustering of nodes of a given degree (see Fig. 1.2**b**).

Global regularities are also observed in real topologies. Given a pair of nodes i and j in a graph, the length of the shortest path, that is, the shortest sequence of nodes with i and j as endpoints such that every pair of consecutive elements are connected (if there is any), defines a distance between the nodes. Indeed, this distance fulfils all the properties of a metric space: it is positively defined, equal to zero if and only if $j = i$, symmetric and it satisfies the triangle inequality,

$$d_{ij} \leq d_{il} + d_{lj}, \forall\, l. \tag{1.3}$$

This last property can be explained as follows. Since by definition no path can be shorter than the shortest path, the path resulting from the concatenation of the shortest paths from i to l and from l to j, the length of which is given by the right-hand side of the above inequality, cannot be shorter than d_{ij}. In real networks, the average distances computed over all pairs of nodes are extremely short and scale as $\bar{d} \sim \log N$ with the system size. This is called the *small-world* property, and it was experimentally discovered on social networks in the 1960's by Milgram [163]. Later on, it was also found that scale-free networks are *ultra-small* [49], with a scaling given by $\bar{d} \sim \log\log N$. In both cases, the diameter of the network D, defined as the longest shortest path, scales as $D \sim \log N$, which implies that the fluctuations of topological distances are small as well. In Fig. 1.2**d**, we show their distribution in the World Trade Web.

Some networks also present degree-degree correlations. For instance, technological systems are usually disassortative [107], 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 [53], according to which there is a positive correlation between the degree of a node and that of its neighbours. These correlations are typically measured in terms of the average degree of the neighbours of nodes of a given degree, as in Fig. 1.2**c**. 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.

Most real networks also present a clear community structure, which means that nodes can be partitioned into groups—communities—such that the density of connections among nodes within the same community is higher than among nodes belonging to different communities [117, 137, 14]. Identifying the optimal partition into communities for a given network, i. e. the one that maximizes the link density within communities, has been a very active topic of research in network science.

It is worth pointing out that these are not the only structural features which have been recognized in real-world networks. For instance, they have also been studied in terms of motifs [111], quantifying the density of small subgraphs in them, of k-core decompositions [10], assessing the structure of the subgraphs in which all nodes have degree larger or equal to k, and even in terms of betweenness, measuring the number of shortest paths passing through a node or a link. Interestingly, all these properties seem to be reproduced by random networks as long as they have the same degree distribution, degree-degree correlations and clustering coefficient [126].

1.1.1.2 Weighted structure

In weighted networks, the local weighted structure of nodes is encoded in their *strength*, defined as the sum of the weights of all links reaching it. In some sense, it is the weighted equivalent to the topological degree and, as their topological counterparts, strengths are also power-law distributed in real networks,

$$P(s) \sim s^{-\xi}. \tag{1.4}$$

Furthermore, another feature broadly observed in real datasets is a non-linear relation between strengths and degrees [20, 136, 22],

$$\bar{s}(k) \sim k^{\eta}, \tag{1.5}$$

where $\bar{s}(k)$ is the average strength of nodes with degree k and η is usually larger than 1. Notice that this observation is not compatible with the trivial coupling between weights and topology, which leads to

$$\bar{s}(k) = \langle \omega \rangle k, \tag{1.6}$$

where $\langle \omega \rangle$ is the average weight per link.

1.1.2 Network models

The fact that many real-world networks from utterly different origins share such important similarities suggests that there might be some general mechanisms driving their formation and evolution. In network science, a great deal of effort has been devoted to unveiling and identifying what those principles are by proposing network models.

A network model is a set of rules defining the generation process of a graph; the goal is, therefore, to find a model with a minimal set of ingredients capable of reproducing all the aforementioned network properties: scale-free degree distribution, sparseness, high clustering, small-world property and community structure. Indeed, such model would allow us to answer the question of why those features are ubiquitous.

It is important to note that there is a random component in most real networks. This is a rather expected statement if we think in social terms, which are probably more intuitive to us; even if the decision to establish a relationship of some sort with someone is not purely random, there is randomness, at least, in who we have the chance to establish a connection with. Consequently, if we aim at modelling complex networks, it seems reasonable to adopt the approach of designing random models.

In a random model, the set of rules dictates the probability with which the events that generate the network happen, rather than specifying which particular events happen. Therefore, a given random model with a given initial condition or model parameters does not generate a single network, but a whole ensemble of networks, each of them with some probability, as opposed to a deterministic model in which the resulting network would be uniquely determined.

In addition to the considerations regarding the randomness of the approach, it is also important to distinguish between *growing* and *static* models. A growing model aims at explaining how complex networks grow in time, that is, how their topological structures emerge during the network-formation process. Typically, in growing models, nodes are added sequentially and the model specifies with what probability they connect with already existing nodes. Static models, on the other hand, focus on defining realistic ensembles of networks with a fixed set of nodes. In the rest of this section, we will briefly review several random models of particular importance in the development of network theory.

The first random model of networks was the *Erdős-Rényi model* [61, 60], which generates completely random graphs. In this static model, for a given size N, E edges are randomly and uniformly distributed among the $N(N-1)/2$ pairs of nodes. There is also a less constrained version of this model in which the number of links E is not fixed, but every pair of nodes has a probability p of being connected (the same for all pairs), which fixes the average number of

edges $\langle E \rangle$ [76]. In any of the two versions, the resulting graphs are small-world, but their degrees are homogeneously distributed and the clustering coefficient vanishes as $N \to \infty$ if the average degree $\langle k \rangle$ is fixed. These two properties are easy to understand in the second and less restricted version. Since all links are identically and independently distributed, degrees follow a Binomial distribution with mean $p(N-1)$. As for the clustering, notice that the probability for a pair of neighbours of some node to be connected is simply p, which scales as $p \sim (N-1)^{-1}$ for fixed $p(N-1)$. Hence, this model shows that real-world networks are not purely random, but some underlying mechanisms must be responsible for degree heterogeneity and clustering.

The explanation of degree heterogeneity came by the hand of what is now commonly known as the *preferential attachment* mechanism, encoded in the *Barabási-Albert model* [17]. In this growing network model, nodes are added sequentially to the network and they establish a fixed number of connections m with the already existing nodes. However, instead of choosing them with equal probability, they tend to connect to higher-degree nodes with higher likelihood. This simple rule, which implies that popularity plays a central role in the network-formation process, indeed results in power-law degree distributions. Nevertheless, this model presents a null clustering coefficient in the thermodynamic limit, so preferential attachment cannot be the only mechanism taking part in the evolution of real networks.

Although the Barabási-Albert model explains scale-freeness, it is not the only model capable of generating power-law degree distributions. There is yet another very important model in network science that allows to generate networks with a predefined target degree sequence but which are random in all other respects: the *Configuration Model* [37, 24]. In this static model, every node i is assigned k_i stubs, or half-links, and the network is constructed by randomly selecting pairs of stubs to form the links (avoiding pairs of stubs belonging to the same node or to two already connected nodes to be chosen in order not to generate networks with self-loops or multiple connections). Clearly, in the resulting network, every node has degree k_i by construction. As in the case of the *Erdős-Rényi model*, there is also a less constrained version, also known as the Soft Configuration Model, in which every node is assigned an expected degree k_i and every pair of nodes is connected with a probability proportional to the product of them, $p_{ij} \propto k_i k_j$. As in the case of the last two models, the Configuration Model also has a vanishing clustering coefficient for large networks.

Finally, the small-world property was also explained in the early days of network theory by means of the *Watts-Strogatz model* [166]. In this model, N nodes are placed on a ring and connected to their k nearest neighbours, hence forming a regular ring-like lattice. Then, every link is visited and, with probability p, it is rewired, that is, one of its endpoints is moved to another randomly chosen node

in the network. With this simple rewiring mechanism, which generates shortcuts across the system, the model undergoes a transition from the non-small-world to the small-world regime as p is increased. Thus, this model identified randomness as an important element determining the structural properties of real networks. Contrarily to the models introduced so far, the Watts-Strogatz model does exhibit strong clustering that remains from the original lattice-like graph. However, degrees are homogeneous.

As for weighted networks, many models have been proposed to explain their observed properties. For instance, growing network models [27, 21, 174, 99, 13, 176, 165, 103] focus on identifying the basic rules driving the growth of weighted networks, usually in the form of a weighted preferential attachment. The maximum-entropy class of models [108, 70, 71, 144, 145], on the other hand, aims at describing real weighted structures in terms of maximally random graphs with a minimal set of constraints. Nevertheless, none of these models is able to reproduce simultaneously the topology and weighted structure of real weighted complex networks.

1.2 Hidden metric spaces

As we have seen in the previous section, although the canonical models in network science reveal some of the main mechanisms shaping the topology and weighted structure of real networks, they are not capable of incorporating all the ingredients in a simple set of rules. Furthermore, none of them provides a simple ontological explanation for the origin of clustering, in the same terms as preferential attachment provides a reasonable explanation for degree heterogeneity.

Ten years ago, the idea of similarity between nodes as a plausible source of clustering was introduced by Serrano, Krioukov and Boguñá [152]. Remarkably, their considerations opened the path to a geometric understanding of complex networks that, besides yielding models capable of reproducing the observed features of real-world networks [152, 94], provides us with new tools and techniques for their analysis and efficient navigation. This section presents a brief introduction to the *hidden metric spaces* approach to complex networks. After discussing the basic concepts, we introduce the different static and growing geometric models and discuss how these allow to construct maps of real networks that can be used to analyze them from a novel perspective as well as to navigate them efficiently.

1.2.1 Similarity, geometry, and clustering

As we saw when we introduced the preferential attachment mechanism, scale-freeness emerges from the predilection of new nodes entering the system to connect to already existing ones based on their degrees, i. e. on their popularities. Nevertheless, it seems reasonable to assume that popularity might not be the only driving force in the connection-formation process. We could indeed presume that the different parts conforming the system, the nodes, are not all equal to one another, but that they are equipped with some intrinsic properties other than popularity, and that the similarity between these intrinsic properties also plays a role in the probability for two nodes to interact. This simple yet reasonable assumption is suitable to be described in geometric terms, as we shall see.

To understand why geometry might indeed be adequate for the description of the similarity between nodes, we should first recognize a very important property of similarity: transitivity. 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. Now, suppose that the space of the nodes' intrinsic properties is a metric space, such that the distance between two points in the space is a measure of their dissimilarity, that is, the closer two points lay in the space, the more similar the intrinsic properties represented by those points are. Then, the transitivity of similarity is automatically incorporated in the structure of the metric space as a result of the triangle inequality

$$d_{AC} \leq d_{AB} + d_{BC}, \tag{1.7}$$

which, in terms of similarity, admits the following reading: the dissimilarity between nodes A and C is bounded by the sum of the dissimilarities of both nodes to node B.

In our network-modelling context, the transitivity of similarity—or, equivalently, of similarity-distance closeness—represents a natural source of clustering. If similarity takes part in the probability for two nodes to connect, then two nodes A and C near in similarity space to node B should be connected to B with high probability. Since, in light of Eq. (1.7), A and C must also be near in similarity space, they should have a high probability of being connected as well, thus yielding a non-zero clustering coefficient for node B. Figure 1.3 illustrates how triangles emerge among nearby nodes in the underlying space.

1.2.2 Geometric network models

In order to explore the geometric origin of clustering in real networks in precise mathematical terms, we need network models based on underlying similarity spaces. In this section, we introduce the main geometric network models proposed during the last years.

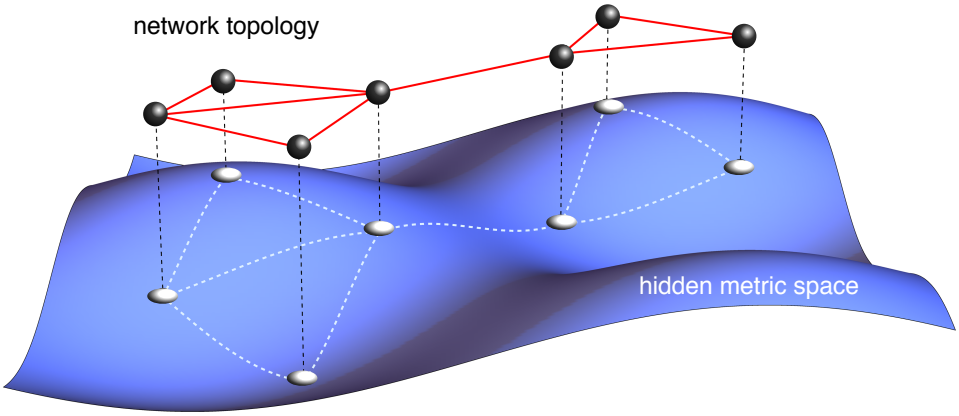


Figure 1.3: **Illustration of the connection between network topology and hidden metric spaces.** The blue surface represents an arbitrary metric space on which several nodes are scattered. The distance among them is given by the length of the geodesics, or shortest curves, drawn as light dashed lines on the surface. Connections are established among nearby pairs only, giving rise to resulting topology depicted in the upper layer. Notice that topological triangles appear naturally as a consequence of the closeness in the hidden metric space.

The first model based on an underlying geometry to be proposed was the $\mathbb{S}^1$ *model*[2] [152]. In this model, every node is assigned two hidden variables, a hidden degree and a coordinate on a circle abstracting the simplest geometry for a similarity space. The connection probability between two nodes then depends on the ratio between their distance along the circle—their similarity—and the product of their hidden degrees—their popularities. Therefore, this model resembles a gravity law for connectivity, in the sense that the likelihood for two nodes to be connected increases with the product of their "masses", the hidden degrees, and decreases with the distance among them. This simple model generates very realistic topologies regardless of the specific form of the connection probability function, as long as it is a sufficiently fast decaying function of the aforementioned ratio. Moreover, every node's degree becomes proportional to its hidden degree, which confers the model a high level of control over the degree distribution. As we will see when we discuss the mathematical details of the model, two additional global parameters enable to control the average degree and clustering of the resulting networks as well. It is worth mentioning that

[2] In fact, it was proposed as a model in arbitrary dimensions, although it has ever since particularized to its one-dimensional version for simplicity. We explore the general version in detail in Section 2.6.

the model has also been extended to bipartite networks in which one mode exhibits a heterogeneous degree distribution while, in the other mode, degrees are homogeneous [155].

Interestingly, it was later discovered that hyperbolic geometry emerges from the $\mathbb{S}^1$ model [93, 94]. By mapping every node's hidden degree into a radial coordinate appropriately, with high-degree nodes closer to the origin of the coordinate system than low-degree nodes, they all lay on a two-dimensional finite disk, and the connection probability becomes a function of the distance separating them only; this distance, however, is not the familiar Euclidean distance but, instead, that of the negatively curved homogeneous hyperbolic space. Therefore, the $\mathbb{S}^1$ model is isomorphic to a purely geometric model, the $\mathbb{H}^2$ *model*, in which nodes are sprinkled on a hyperbolic disk and then connected with a probability that decreases with the distance among every pair of them. It is important to note that, in this purely geometric representation, geometry does not only encode similarity distances, but it includes the popularity dimension (degrees) as well in a sole distance metric.

Both the $\mathbb{S}^1$ and the $\mathbb{H}^2$ models will be explained in detail in Chapter 2. However, it might be beneficial to give some intuition about hyperbolic geometry and its interplay with complex topologies at this point. Hyperbolic space is a metric space with constant negative curvature [45]. Unfortunately, contrarily to its positively curved counterpart—the two-dimensional sphere—the two-dimensional hyperbolic plane cannot be completely embedded isometrically as a surface in three-dimensional Euclidean space, which hinders its visualization and makes it less intuitive. Nevertheless, we can still grasp some intuition about this geometry by writing down the metric tensor explicitly in the coordinate system of the so-called native representation, or polar coordinate system, of hyperbolic space, which will be the only one used in this thesis. In these coordinates, every point in the space is given a radial coordinate standing for the distance to the (arbitrary) origin of the space, as well as an angular coordinate between zero and 2π. The metric then reads

$$\mathrm{d}s^2 = \mathrm{d}r^2 + \sinh^2 r \mathrm{d}\theta^2. \tag{1.8}$$

By comparing it to the metric of the familiar Euclidean two-dimensional plane in polar coordinates, $\mathrm{d}s^2 = \mathrm{d}r^2 + r^2\mathrm{d}\theta^2$, we can immediately notice that the perimeters of hyperbolic circles grow much faster with their radii than Euclidean circles. In particular, they grow exponentially, with a scaling given by the hyperbolic sine in Eq. (1.8). Naturally, the areas of disks also scale exponentially with their radii.

Now, in the $\mathbb{H}^2$ model, nodes are sprinkled with a fixed (and quasi-uniform) density, as will be explained in detail in the next chapter. This means that the number of nodes at distance smaller than r from a given node is proportional to the area of the hyperbolic disk of radius r centred at the node and, hence,

scales exponentially with r; this is fully consistent with the small-world property, as a result of which the number of nodes at a given topological distance d from any node scales exponentially with d. Besides explaining small-worldness, clustering is immediately encoded in the triangle inequality, so all the discussion presented in Section 1.2.1 can be applied here as well. Finally, we can give an intuitive interpretation of degree heterogeneity in this context. As we will see in the next chapter, nodes tend to establish connections with other nodes at distance smaller than the radius of the hyperbolic disk $\hat{R}$ (see Eq. (2.12) for details). Consequently, a node with very small radial coordinate is close to the point laying at distance smaller than $\hat{R}$ from all the nodes in the system: the origin. Therefore, nodes near the origin tend to have very high degrees. Nodes in the periphery, on the other hand, can only connect to other nodes that are either near the origin or peripheral but very close angularly. This last point can be seen from the metric: in the periphery of the disk, a very small angular separation becomes a very large hyperbolic distance. Therefore, nodes at large radial coordinates can only connect to a much more restricted set of nodes.

This thesis strongly relies on these two static geometric models, and the isomorphism between them is repeatedly used throughout. Nevertheless, there are also two growing geometric models which we present here for completeness.

The *Popularity versus Similarity Optimization model* (PSO) [133] is based on the following assumptions. First, the popularity of a node at a given growth-time is determined by its age, that is, how long ago it appeared into the system; hence, older nodes should receive more connections from newcomer nodes. In other words, there is a preferential attachment process. Second, newly added nodes do not choose their neighbours according to their popularity only, but they also take similarity, abstracted as an angular coordinate, into consideration, henceforth generating clustering in the resulting networks. In the PSO model, when a node enters the system, it connects with a fixed number m of nodes that are maximally popular or similar to itself, where m regulates the average degree. Remarkably, by mapping each age to a radial coordinate appropriately, the model can be interpreted as nodes connecting to the m closest nodes in hyperbolic space. Furthermore, allowing the popularity of nodes to fade away over time confers the model control over degree heterogeneity, whereas the level of clustering can be regulated by adding some randomness to the selection of neighbours.

As we discuss in the next section, when we unveil the hidden similarity space coordinates of nodes in real networks through a process that we call embedding, we observe a non-uniform angular distribution of nodes, a phenomenon intimately entangled with the community structure of the network. To account for these geometric communities, a model based on the PSO was proposed: *the Geometric Preferential Attachment model* (GPA) [177]. In the GPA, new nodes do

not appear with uniform angular density on similarity space; instead, the probability for a new node to appear at a given coordinate increases with the amount of nodes in the vicinity of such point. Hence, there is a geometric preferential attachment in similarity space that provides a plausible explanation for the community structure of real networks: the intrinsic properties of nodes, defining their position in similarity space, are influenced by the properties of already existing nodes, as they tend to mimic features which are frequent in the system.

Notice that the $\mathbb{S}^1$ model shares some similarities with the gravity-law approach to trade in economics [162, 26, 158] to predict trade volumes, whereas the $\mathbb{H}^2$ model resembles latent space models [109, 110], conceived to explain homophily and similarity in social systems.

1.2.3 Embedding of real networks and navigability

One of the most interesting possibilities that geometric models offer is the embedding of real-world networks into hidden metric spaces. To that end, we consider the following problem: given a network, and assuming its geometric origin, can we find its *embedding* on a hidden metric space, that is, can we infer the hidden variables and model parameters that best describe the observed topology? This question has been proved to have a positive answer in general, although it poses some technical challenges. In this section, we briefly discuss the most commonly adopted methods to tackle this problem and elaborate on the possibilities that the resulting maps of real systems open.

A very widely used method to address the problem of geometric-information inference is likelihood maximization [35, 155, 2, 30]. 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. This is indeed the general scheme followed to obtain all the embeddings in this thesis, so the procedures will be detailed in Chapter 2. However, it is worth mentioning that this embedding scheme is not representative of all currently available methods. Several new approaches to tackle this problem apply machine learning algorithms to find the angular coordinates, disregarding the likelihood completely [8, 115]. Some other methods follow a pure likelihood-maximization strategy making use of the growing models instead of the static versions [134].

What makes it worth overcoming the technical difficulties of network embedding is applying the technique to real-world networks. If we adhere to the hypothesis that similarity between nodes influences the observed topologies of real networks, we can interpret the coordinates of nodes obtained from the embedding as their positions in the underlying similarity space. In other words, embedding a network onto a metric space provides us with a map of the system's similarity space, which would otherwise remain hidden. A scep-

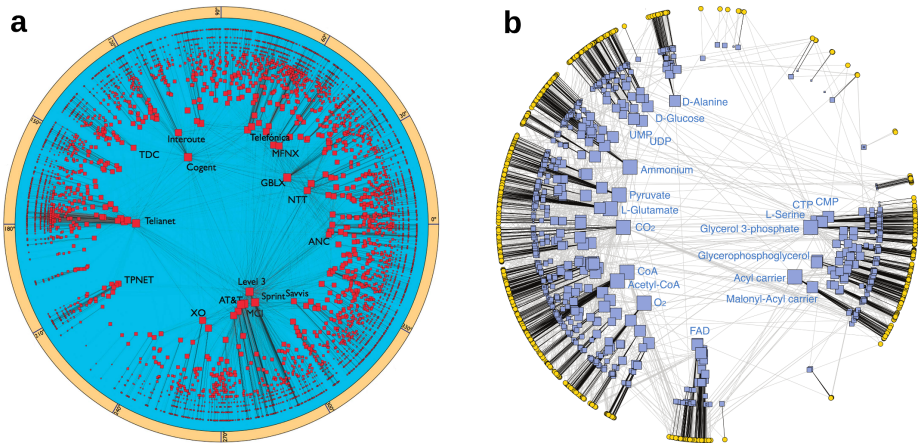


Figure 1.4: **Hyperbolic representations of two real-network embeddings.** Nodes are represented by squares and links by lines between them. Every node is assigned a radial coordinate representing its popularity, or degree, such that high-degree nodes are near the centre (see Section 2.3 for further details). Angular coordinates account for the positions of nodes in similarity space, so the boundary of the disk is the $\mathbb{S}^1$ model metric space. Notice that both angular distributions are far from being homogeneous, as they exhibit large gaps and highly populated angular regions. **a**, The Internet Autonomous Systems network. Every node is an Autonomous System, a subnetwork in the physical network of the Internet, taking part in the information-routing process. The embedding reveals a clear correlation between nodes and their geographical location [35]. **b**, Metabolic network of the *Escherichia coli*. Every node is a metabolite and the connections among them represent their participation into common reactions. In this case, angular inhomogeneities are closely related to metabolic pathways [155].

tical reader, however, might argue that maximizing the likelihood for *any* network, including graphs without a geometric origin, would yield a set of hidden coordinates, which in turn implies that the coordinates obtained from the embedding of real networks could be meaningless, in the sense that they need not have a similarity-coordinate interpretation. However, during the last ten years, much evidence supporting precisely the opposite claim has been accumulated [35, 155, 82, 2, 3, 4], including many of the results in this thesis.

A notable characteristic of the angular coordinates obtained from the embedding or real networks is non-uniformity; we usually find angular regions where the density of nodes is higher than average followed by large empty gaps separating these regions (see Fig. 1.4). The reason for the appearance of these denser areas is that, within any of them, the link density is higher than average, so they are caused by the community structure of the network. In order

for the model to encode this higher link density, the average angular distance between nodes within each community needs to be shorter, thus forming these angular correlations in the embedding. Moreover, the resulting angular distribution usually has a very evident interpretation in terms of similarity. For instance, in the embedding of the Internet [35], the angular communities were found to be highly correlated with geographic localization, while in the geometric analysis of the Human and *E. Coli* metabolisms [155], the pathways could be clearly identified in different angular sectors. Notice, nevertheless, that the embedding process is completely blind to this information, as it simply maximizes the likelihood function based on the adjacency-matrix elements.

In addition to revealing the hidden similarity and, consequently, the community structure of networks at a glance, the embedding of real networks is a valuable resource for efficient network navigation as well [36, 33]. This is especially important for some technological systems, the Internet being a paradigmatic example, in which finding an effective protocol to route information across the network is crucial. To put the problem into context, let us consider how information travels through the Internet nowadays. When we want to communicate with a distant computer, information is sent in the form of packets on which the destination node is indicated with an identifier: the Internet Protocol (IP) address. When any intermediate node receives the packet, it must decide to which of its neighbours the packet should be forwarded so that it eventually reaches the target IP. Moreover, the routing decision should ideally minimize the number of hops the packet travels. This is a difficult problem that, in principle, requires shortest paths to be computed and, in addition, all nodes to be aware of which is the best next-neighbour for every possible destination. In the case of the Internet, the fact that it is not static because it is subject to failures of nodes or links, and also because it is a growing system constantly incorporating new nodes, makes the problem even more involved. Indeed, protocols like the Border Gateway Protocol (BGP), which is currently used to exchange routing information among Autonomous Systems on the Internet, might present scalability issues.

An alternative protocol that could overcome possible scalability problems is *greedy routing* [131]. In this hidden-geometry information-routing protocol, every node must be given a position in the hyperbolic plane in accordance with the embedding of the network in the $\mathbb{H}^2$ model. In the information-forwarding process, 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 paths traced by the greedy-routing packets are congruent with the geodesics in the underlying hyperbolic manifold, which results in the topological length of the

paths being almost minimal. However, the main drawback of the protocol is that it does not guarantee the success of the routing for all pairs of nodes. In some cases, the packet can get trapped in a loop, hence never reaching the target. This forces us to consider the performance of the navigation process in terms of the success rate, i. e. the fraction of pairs of nodes that can communicate via greedy routing. Although overcoming this issue still calls for further investigation, the advantages of the approach are remarkable: every node only needs to be provided with local information—the coordinates of its neighbours—and, yet, the protocol is robust against sudden changes of the network. In Ref. [35], the authors simulated this dynamical process on the real topology of the Internet and measured its performance when a fraction of nodes fails. Surprisingly, the protocol is almost fully functional even when a large fraction of the network is suddenly removed. They also showed that the whole set of coordinates needs not be recomputed every time a new node enters the system. It is sufficient to assign to newcomers a hyperbolic position based on a local optimization of the likelihood, that is, constrained to the fixed coordinates of all other nodes. Surprisingly, it has even been shown that, in temporal networks in which nodes activate and deactivate dynamically, navigability is enhanced with respect to the static situation [128].

Beyond the obvious technical advantages of greedy routing, the fact that the protocol is generally capable of finding nearly shortest paths in real networks (not just technological ones) can be interpreted as a confirmation of the meaningfulness of the embedding coordinates.

1.3 Outline of the thesis

This thesis starts by reviewing the $\mathbb{S}^1$ and $\mathbb{H}^2$ models, which are central in this work, in Chapter 2. We provide a description of their mathematical details, as well as of the isomorphism connecting both. Moreover, we discuss the embedding methods used in the thesis. We also present some results regarding the extension of the $\mathbb{S}^1$ model to the heterogeneous angular distributions regime, with a method allowing to construct networks with geometric communities. This chapter also explores the idea of higher-dimensional similarity spaces. We review in detail the general D-dimensional $\mathbb{S}^D$ model along with the upper-bound on the number of similarity-space dimensions that this model implies for real, highly clustered, complex networks.

In Chapter 3, we turn our attention to the study of the evolution of the World Trade Web, the network of international trade between countries in the world, from the hidden metric spaces perspective. After reconstructing yearly world trade networks covering a time span of 14 decades, we embed them into hy-

perbolic space according to the $\mathbb{H}^2$ model with an embedding algorithm specially devised for this particular system. The maps resulting from the embedding, which can be examined using our online interactive video tool [69], open the path to a novel understanding of the interactions between world economies based on their otherwise hidden similarities. Our analysis of *The World Trade Atlas 1870-2013*, as we have named the collection of maps, reveals that there are three main forces shaping the world in economical terms: globalization, localization and hierarchization. From the complex networks point of view, our results also show that network embeddings enable to define a high-resolution community-detection method, as well as a geometry-based measure of the level of hierarchization of real systems. Clearly, all these results add further evidence to the existence of hidden similarity spaces underlying real networks.

The results presented in Chapter 4 also support the meaningfulness of hidden similarity spaces, although from a different and perhaps more abstract perspective: they reveal a previously unknown symmetry of real complex networks. As we explain in the chapter, we hypothesize that similarity distances could be a better source of length scales in networks than topological ones, which, although they define a metric space, lack variability due to the small-world property. Based on this assumption, we define a Geometric Renormalization Group transformation that reveals that real-world networks are self-similar with respect to similarity-space scale transformations, as all their topological features are preserved in the renormalization process. Interestingly, the results on real systems match the predictions of the $\mathbb{S}^1$ model accurately. This Geometric Renormalization Group also brings immediate applications. One of the main contributions of the chapter is the development of a method that allows to construct smaller-scale replicas of real networks that, aside from the number of nodes, present similar topological properties as their original full-size counterparts. The resemblance between the replicas and the original networks is so significant that the dynamical processes running on them behave almost identically. Furthermore, the ability to observe real complex networks at different resolutions is also exploited in the chapter to devise a multiscale greedy-routing algorithm that outperforms the single-scale one.

All the results in the first chapters are restricted to binary networks. In Chapter 5, however, we explore whether the structure of weighted networks also finds an explanation in terms of underlying hidden metric spaces, and we show that similarity space has a strong influence on the weighted structure of real networks as well. In particular, we confirm the conjecture that, if there is a relation between the distance separating two nodes in the underlying similarity space of a real system and the resulting weight of their interaction, then there should be an observable correlation between weights and edge multiplicities (number of triangles sharing an edge), simply because nearby nodes in the metric space

tend to have more common neighbours. Given this positive result, we devise a geometric model for weighted networks. This model is proven to be an exceptional tool to understand the weighted structure of real systems, as it is able to reproduce not only weighted features like the weight and strength distributions, but also the non-trivial behaviour of the Triangle Inequality Curve and, simultaneously, the basic topological properties. We also present a method to infer the model parameters for real weighted networks, with which the model performs extremely well at reproducing all the observed structural features of real systems.

CHAPTER 2

Static geometric models and embeddings of networks

2.1 Similarity space as a sphere

The idea of a geometric origin of clustering as a result of the transitivity of similarity relies on the triangle inequality, a defining property of *any* metric space. This raises the question of what geometry should be used to model real networks. Since we have no prior information about real similarity spaces, it seems reasonable to choose a highly symmetric geometry in order to avoid any unnecessary biases, which leads us to consider homogeneous and isotropic spaces. Furthermore, since we are interested in modelling finite systems, it is convenient to focus on manifolds that are compact as well. These considerations suggest using spherical similarity spaces.

Regarding the dimensionality of the space, it is also unknown in general for real networks. It seems clear that there might be more than one single property determining the resemblance between nodes. However, as we show in this chapter, the dimension of similarity space must be upper-bounded for networks with high clustering coefficient, like real-world networks. Thus, although real similarity spaces might have dimensions higher than one, these cannot be much higher, so assuming them to be one-dimensional results in a good description of real systems. Indeed, most studies in the field have been restricted to one single similarity dimension for simplicity. From the theoretical point of view, one dimension suffices in order to generate realistic networks and, therefore, to explain typical topological characteristics. In terms of applications, which usually require embedding networks on similarity space, considering more than one dimension has been, at least until recently, intractable in practice [115].

This thesis relies mainly on the widely used one-dimensional $\mathbb{S}^1$ model as well as on its purely geometric counterpart, the $\mathbb{H}^2$ model, defined in hyperbolic space. Therefore, in this chapter we first review these two models along with the close connection between them. We also present our results concerning their generalization to networks with angular communities and, finally, we explore the question of the dimension of real similarity spaces through the D-dimensional $\mathbb{S}^D$ model.

2.2 The $\mathbb{S}^1$ model

The first geometric model to be proposed, the $\mathbb{S}^1$ model, is the simplest one capable of generating realistic topologies [152]. It assumes a very elementary similarity metric space, a one-dimensional sphere or circle $\mathbb{S}^1$. In the first step of the model, we scatter N nodes with uniform density on a circle, so every node i is assigned a hidden angular coordinate θ_i between zero and 2π. The radius of the circle R is set to $R = \frac{N}{2\pi}$, which makes the density of nodes equal to 1. Secondly, we also assign to every node another variable, a hidden degree κ_i, randomly from an arbitrary distribution $\rho_\kappa(\kappa_i)$ (see Fig. 2.1). Given that we are interested in real-world networks, we will usually consider power-law distributions $\rho_\kappa(\kappa_i) \sim \kappa_i^{-\gamma}$, possibly with a cut-off κ_c. Notice that all hidden degrees and hidden angles are independent random variables.

Once all hidden variables have been assigned, we connect every pair of nodes i and j with independent probabilities

$$p\left(\kappa_i, \kappa_j, d_{ij}\right) = p\left(\chi\right), \chi = \frac{d_{ij}}{\mu \kappa_i \kappa_j}, \tag{2.1}$$

where the function $p\left(\chi\right)$ is any integrable decreasing function taking values between 0 and 1. In Eq. (2.1), $d_{ij} = R\Delta\theta_{ij}$ represents the metric distance between the nodes, with $\Delta\theta_{ij}$ standing for the angular distance separating the two nodes, $\Delta\theta_{ij} = \min\left(\left|\theta_i - \theta_j\right|, 2\pi - \left|\theta_i - \theta_j\right|\right)$. The particular dependence of χ on distances and hidden degrees confers the model the form of a gravity law: connection probabilities increase with the product of hidden degrees (masses) and decrease with distance. The free parameter $\mu > 0$ regulates the contribution of the product of hidden degrees in the denominator, hence controlling the average degree of the resulting networks. This can be seen in more detail by computing the expected degree of a node with hidden degree κ_i [152, 34]. To that end, consider the probability for i to connect to a randomly drawn node j, that is, whose angular distribution $\rho_\theta\left(\theta_j\right)$ is $1/(2\pi)$ and whose hidden degree is assigned from $\rho_\kappa\left(\kappa_j\right)$,

$$\rho\left(a_{ij} = 1 | \kappa_i\right) = \int p\left(\kappa_i, \kappa_j, d_{ij}\right) \rho_\theta\left(\theta_j\right) \rho_\kappa\left(\kappa_j\right) \mathrm{d}\kappa_j \mathrm{d}\theta_j. \tag{2.2}$$

Notice that this probability cannot depend on the specific angular coordinate of node i due to the homogeneity of the space, so we can place the origin of angular coordinates at node i's position without loss of generality. This allows us to compute the above integral over θ_j between 0 and 2π as twice the integral between 0 and π, where now $d_{ij} = R\theta_j$. Moreover, since there are $N-1$ ($\approx N$ for large N) independent potential neighbours in the system,

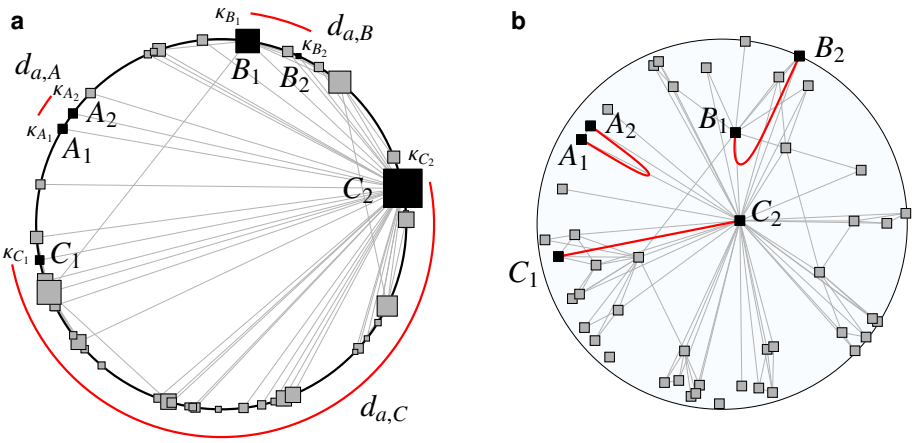


Figure 2.1: **Gravity model and hyperbolic representation.** Three different pairs of nodes $A_1 - A_2$, $B_1 - B_2$ and $C_1 - C_2$ have been highlighted. **a**, $\mathbb{S}^1$ model in which the angular distance, d_a, is given by the angular separation along the circle. The size of a node is proportional to the logarithm of its hidden degree κ. **b**, Hyperbolic space representation of the model in **a**. The three different pairs of nodes have been chosen such that they are separated by the same hyperbolic distance, $\hat{R}$. The red lines represent the geodesics (shortest curves) between the nodes. The figure corresponds to the native representation, so distances to the center of the disk are not distorted, whereas the perimeters of circles centred at the origin are actually much longer than they appear to be in the image, an effect that increases with radius (see the discussion about hyperbolic geometry in Section 1.2.2 for details), and which explains the shape of the geodesics: to minimize distance, one should move angularly at small radial coordinates. As an example, nodes $A_1 - A_2$ with small degree and very low angular separation can be very far apart in hyperbolic space. Notice that, although nodes are equally sized, the ones with higher degree (larger size in **a**) are positioned closer to the centre in the hyperbolic plane.

$$\begin{aligned}\bar{k}_i(\kappa_i) &= (N-1)\,\rho\left(a_{ij}=1|\kappa_i\right) \approx \frac{N}{\pi}\int_{\kappa_0}^{\kappa_c}\int_{0}^{\pi} p\left(\frac{R\theta_j}{\mu\kappa_i\kappa_j}\right)\rho_\kappa(\kappa_j)\mathrm{d}\kappa_j\mathrm{d}\theta_j \\ &= 2\mu\kappa_i\int_{\kappa_0}^{\kappa_c}\kappa_j\rho_\kappa(\kappa_j)\int_{0}^{N/(2\mu\kappa_i\kappa_j)} p\left(\chi\right)\mathrm{d}\chi\mathrm{d}\kappa_j \xrightarrow{N\to\infty} 2\mu\kappa_i\langle\kappa\rangle\int_{0}^{\infty} p\left(\chi\right)\mathrm{d}\chi.\end{aligned} \tag{2.3}$$

Interestingly, we see that the expected degree of a node is proportional to its hidden degree regardless of the specific form of the function $p\left(\chi\right)$. Moreover, by

setting $\mu = 1/\left(2\langle\kappa\rangle \int_0^\infty p\left(\chi\right) \mathrm{d}\chi\right)$, we have $\bar{k}_i(\kappa_i) = \kappa_i$. This is one of the most important properties of the model: since the resulting degrees of nodes are proportional to their hidden degrees, we can generate networks with any target degree distribution. This property is also preserved in higher dimensions, so we present a thorough general discussion of the behaviour of degrees in Section 2.6.

Since degrees can be controlled in the model rather independently from the particular functional form of $p\left(\chi\right)$, this function can be exploited to regulate the dependence of the connection probabilities on the underlying metric distances which, according to our previous discussions, should influence the resulting clustering coefficient [152, 94]. It is customary to define

$$p\left(\chi\right) = \frac{1}{1+\chi^\beta}, \tag{2.4}$$

since it casts the ensemble of networks generated by the model into exponential random networks [94]: networks that are maximally random given the constraints imposed by the model parameters. The new global parameter β controls the coupling of the model to the underlying metric space. For very large values of β, the connection probability becomes strongly dependent on the metric distances d_{ij}, whereas lower β allows for spurious long-range connections to appear. As a consequence, this parameter tunes the resulting clustering coefficient in the model, which increases with increasing β. In addition, at $\beta = 1$ the model undergoes a phase transition and clustering vanishes for lower values of β. With this choice of connection probability, μ is given by

$$\mu = \frac{\beta \sin\frac{\pi}{\beta}}{2\pi\langle\kappa\rangle}, \tag{2.5}$$

where, in light of the equality between hidden and expected degrees, $\langle\kappa\rangle$ can be substituted by the expected average degree.

Despite the simplicity of this model, the networks that it generates exhibit realistic topologies [152, 94, 35, 155, 1, 2, 3, 4]: they are scale-free (if hidden degrees are power-law distributed), sparse, highly clustered and small-world if $\gamma < 3$ or $\beta < 2$ or both, as we will see in Chapter 4. Since the analytical study of these other properties is rather involved, we do not include it here, but we refer the reader to Refs. [152, 94, 44]. In Fig. 2.2, we show the degree distributions and clustering spectra of networks generated by the $\mathbb{S}^1$ model.

The model can also be modified to generate networks with community structure. As we discussed when we introduced the embedding of networks, the community structure of a network is related to angular correlations between nodes. In particular, communities are represented as densely populated angular regions in the underlying space. Interestingly, the model can be adapted accordingly so that these angular heterogeneities emerge in the first step of the process; we explore this point in Section 2.5 of this chapter.

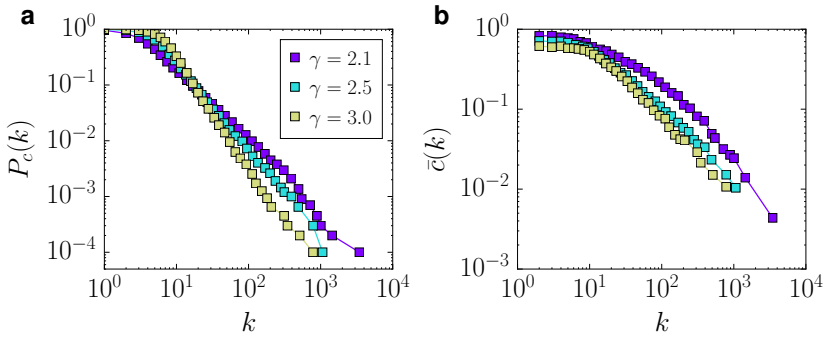


Figure 2.2: **Topological properties of synthetic networks.** The plots show the topological properties of three networks with $N = 10^4$ nodes, $\langle k \rangle = 10$ and different power-law degree-distribution exponents generated with the $\mathbb{S}^1$ model with $\beta = 2.5$. **a**, Complementary cumulative degree distributions. **b**, Clustering spectra. Notice that all three networks exhibit a high average clustering coefficient (higher than 0.5 in all cases).

2.3 The $\mathbb{H}^2$ model

In the $\mathbb{S}^1$ model, geometry explicitly represents resemblance between nodes, while the popularity dimension is encoded into the non-geometric hidden degrees. However, the disentanglement of popularity from geometry is only apparent; this is implied by the fact that the model is isomorphic to a different geometric model, the $\mathbb{H}^2$, in which the geometry of the space encodes both dimensions, popularity and similarity, and where the connection probability becomes a function of the distance in the space only [93, 94, 35, 155, 2, 4]. In this section, we show how to construct a purely geometric representation from the $\mathbb{S}^1$ model, which defines the $\mathbb{H}^2$ model, and its connection to hyperbolic geometry.

2.3.1 Isomorphism between the $\mathbb{S}^1$ and the $\mathbb{H}^2$ models

Consider some given realization of the $\mathbb{S}^1$ model. We can assign a new hidden variable r_i to every node as a function of its hidden degree κ_i according to

$$r_i = \hat{R} - 2\ln\frac{\kappa_i}{\kappa_0}, \tag{2.6}$$

where κ_0 is the smallest hidden degree in the system and

$$\hat{R} = 2\ln\frac{N}{\mu\pi\kappa_0^2}. \tag{2.7}$$

In terms of these new quantities, the connection probability in Eq. (2.1) becomes

$$p\left(\frac{R\Delta\theta_{ij}}{\mu\kappa_i\kappa_j}\right) = p\left(e^{\frac{1}{2}(x_{ij}-\hat{R})}\right), \tag{2.8}$$

where the only pairwise-dependent term is

$$x_{ij} = r_i + r_j + 2\ln\frac{\Delta\theta_{ij}}{2}. \tag{2.9}$$

Strikingly, hyperbolic geometry emerges from this relations; the expression in Eq. (2.9) is, approximately, the distance between two points with radial coordinates r_i and r_j and angular separation $\Delta\theta_{ij}$ in the hyperbolic plane with curvature $K=-1$ (in its native representation), the exact expression for which is given by the hyperbolic law of cosines [94, 7],

$$\hat{d}_{ij} = \operatorname{acosh}\left(\cosh r_i \cosh r_j - \sinh r_i \sinh r_j \cos\Delta\theta_{ij}\right). \tag{2.10}$$

In other words, by mapping every node's hidden degree onto a radial coordinate according to Eqs. (2.6) and (2.7), all nodes lay on a disk of radius $\hat{R}$ and their connection probabilities only depend on the hyperbolic distances separating them (see Fig. 2.1). This is, fundamentally, the $\mathbb{H}^2$ model [94], and Eqs. (2.6) and (2.7) constitute the quasi-isomorphism (the "quasi" here stands for the fact that Eq. (2.9) is not exactly equal to Eq. (2.10)[1]) between this model and the $\mathbb{S}^1$ [152], which we will invoke repeatedly throughout this thesis.

2.3.2 Specifications

The $\mathbb{H}^2$ model [94] generates realistic complex networks by scattering N nodes on a hyperbolic disk of finite radius $\hat{R}$ with quasi-uniform density. The density of nodes is not completely uniform in hyperbolic space, but it is given a radial dependence in order to tune the level of degree heterogeneity in the resulting graph. Indeed, according to Eq. (2.6), a power-law hidden-degree distribution implies an exponential distribution of radial coordinates. Therefore, the radial density is chosen to be

$$\rho(r) = \alpha\frac{\sinh\alpha r}{\cosh\alpha\hat{R}-1} \approx \alpha e^{\alpha(r-\hat{R})}, \tag{2.11}$$

for which the resulting degrees are power-law distributed with exponent γ given by $\gamma = 2\alpha+1$ if $\alpha \geq 1/2$ or $\gamma = 2$ otherwise. Once all nodes have been sprinkled, every pair of nodes is connected with the probability given by Eq. (2.8), where

[1] It is important to mention that the goodness of the approximation increases with the radial coordinates r_i and r_j of the two nodes [94, 7] and, henceforth, with increasing network size N.

the exact formula for the hyperbolic distances, Eq. (2.10), can be used instead of its approximate version, Eq. (2.9).

Particularizing now to the connection probability given in Eq. (2.4), two nodes connect as a function of the hyperbolic distance between them according to

$$p\left(\hat{d}_{ij}\right) = \frac{1}{1+e^{\frac{\beta}{2}\left(\hat{d}_{ij}-\hat{R}\right)}}, \tag{2.12}$$

which has a nice statistical-physics interpretation: it represents the occupation probability of a fermionic state of energy $\hat{d}_{ij}/2$ in a grand-canonical ensemble at inverse temperature β and with chemical potential $\hat{R}/2$ [93, 94]. This interpretation is further explored in Chapter 4. For completeness, let us finally express the hyperbolic radius in Eq. (2.7) as a function of the target average degree and power-law exponent γ for this form of the connection probability. Using Eq. (2.5) and the equivalence between hidden and expected degrees, we find

$$\hat{R} = 2\ln\left[\frac{2N}{\beta\sin\frac{\pi}{\beta}\langle k\rangle}\left(\frac{\gamma-1}{\gamma-2}\right)^2\right]. \tag{2.13}$$

2.4 Embedding methods

As we explained in Chapter 1, the $\mathbb{S}^1$ and $\mathbb{H}^2$ models can be used to embed networks in hidden metric spaces, that is, to find a set of values for the global parameters and coordinates that maximize the congruency between the network and the model. In this section, we present two different methods. The first one, which was proposed in Ref. [35], can be used to embed large networks ($N \lesssim 10^5$ nodes) using the $\mathbb{S}^1$ model. The second one, originally introduced in Ref. [2], is designed to embed sequences of small networks through the $\mathbb{H}^2$ model such that the coordinates of two consecutive networks in the sequence are correlated.

Both embedding methods rely on likelihood maximization. The main idea is as follows. Given a set of hidden variables for every node in the network that we want to embed, and given some values of the global parameters as well, all connection probabilities p_{ij} in the model are specified. Since the model assigns every link independently from the rest, the likelihood $\mathscr{L}$ for the observed network to be generated by the model is given by

$$\mathscr{L} = \prod_{i<j} p_{ij}^{a_{ij}}\left(1-p_{ij}\right)^{1-a_{ij}}, \tag{2.14}$$

where $\{a_{ij}\}$ is the adjacency matrix, with entries $a_{ij}=1$ if nodes i and j are connected and $a_{ij}=0$ otherwise. Note that a_{ij} depends on the network to be embedded, whereas p_{ij} depends on the coordinates (and hidden degrees in the case of the $\mathbb{S}^1$ model) to be inferred. Moreover, the connection probability,

Eq. (2.4) or (2.12), depends also on the global parameters of the model. The traditional approach to tackle the maximization of the likelihood is therefore two-stepped. We first infer the values of the global parameters and, with these, we proceed to find the coordinates maximizing Eq. (2.14). The two algorithms contained in this section follow this approach. Moreover, they assume a power-law degree distribution.

2.4.1 Algorithm for large networks

The algorithm presented in this section was originally designed to embed the Internet AS network [35]. It is based on several heuristics conferring the method two important properties. On the one hand, the embedding time scales as $\mathcal{O}\left(N^2\right)$ with the system size, which enables to embed large graphs. On the other hand, the method can be used to determine the coordinates of newly added nodes to the network with no need for the modification of the coordinates of previously embedded nodes; this is particularly relevant for the Internet network, which is a growing system.

2.4.1.1 Estimation of the global parameters

This algorithm embeds networks on the $\mathbb{S}^1$ model, which has three global parameters: $R = N/(2\pi)$, μ, related to the average degree by Eq. (2.5), and β, which determines the clustering in the model. The values of the parameters N and $\langle k \rangle$, however, need to be estimated carefully for two main reasons. The first one is that, when a synthetic network is generated with the model, some of its nodes may have no connections (which is more likely to occur for very sparse graphs) and we do not consider them as belonging to the system. Hence, the target N and $\langle k \rangle$ in the model can be different from the observed ones. In particular, the observed system size N_{obs} is related to the number of nodes in the model N through the relation

$$N_{\text{obs}} = N\left(1 - P\left(0\right)\right), \tag{2.15}$$

where $P\left(0\right)$ is the probability for a randomly chosen node to have degree zero. The second reason is a finite-size effect. If we generate a finite sequence of N hidden degrees from a power-law distribution with mean $\langle \kappa \rangle$, the sequence will have an upper cut-off κ_c as a consequence of which the resulting average will be

$$\langle \kappa\left(N\right) \rangle = \alpha \langle \kappa \rangle < \langle \kappa \rangle, \tag{2.16}$$

where $\alpha \to 1$ as $N \to \infty$, since $\kappa_c \to \infty$. Consequently, we cannot simply use the observed values for the number of nodes and average degree to determine the model parameters. Instead, the *hidden* values of these parameters need to

be determined from observed quantities. The precise details of the method to obtain them are rather involved, so we refer the reader to the Supplementary Information in Ref. [35] for a detailed description. The main idea is to measure the value of the power-law exponent γ by direct inspection (or from the slope of the cumulative distribution in a log-log plot, as we explain in the next subsection) and the maximum observed degree $k_{\text{obs}}^{\max}$. With these two quantities, along with N_{obs} and $\langle k \rangle_{\text{obs}}$, we can write a system of equations involving all the aforementioned variables, the solution of which gives $P(0)$ and α. These two unknowns finally allow us to obtain the hidden values of N and $\langle k \rangle$ with which R and μ can be computed.

The value of β can be chosen such that the clustering coefficient in the model matches the observed value in the real network. Notice, however, that in Ref. [35], the aim was to obtain an efficient embedding for network navigation. Consequently, the comparison of the clustering coefficient was used to determine a range for possible values of β, whereas the precise value was chosen to be the one maximizing the success rate of navigation on synthetic networks.

2.4.1.2 Inference of hidden degrees and angular coordinates

With the inferred values of the global parameters, we can proceed to the maximization of Eq. (2.14) with respect to hidden degrees and angular coordinates. As shown in Ref. [35], if we assume a power-law hidden-degree distribution, the maximization with respect to hidden degrees can be carried out analytically. As a result, a node with observed degree k must be assigned a hidden degree κ according to

$$\kappa = \max\left(\frac{\gamma-2}{\gamma-1}\langle k \rangle, \frac{1}{\alpha}\left(k - \frac{\gamma}{\beta}\right)\right), \tag{2.17}$$

where the $1/\alpha$ factor accounts for the previously discussed finite-size corrections. Moreover, notice that the maximum function guarantees that no hidden degree is smaller than the lower cut-off $\kappa_0 = \frac{\gamma-2}{\gamma-1}\langle k \rangle$ of a power-law distribution with mean $\langle k \rangle$.

Unfortunately, the angular coordinates maximizing the likelihood function cannot be derived analytically. Indeed, it is even difficult to obtain them numerically, mainly because, for large networks, the space of all possible angular configurations is too large to be exhaustively explored, so the problem requires some heuristics. This algorithm drives the system towards the solution by dividing the network into nested subgraphs and embedding them sequentially, so that the inferred coordinates of a subgraph become the starting point of the next one. We detail the method in what follows.

The first step is to divide the network into a set of m subgraphs $\mathscr{G}(k_i)$, where $\mathscr{G}(k_i)$ is the subgraph formed by the nodes that have degrees larger than or equal

to k_i in the network and the connections among them. The sequence of degrees $k_1 > k_2 > \ldots > k_m$ thus yields a sequence of such subgraphs fulfilling $\mathcal{G}(k_1) \subset \mathcal{G}(k_2) \subset \ldots \subset \mathcal{G}(k_m)$. Next, we remove all connections among the nodes in $\mathcal{G}(k_1)$ and assign random angular positions to them. The reason to do so is that, since nodes with degree higher than $\sim \sqrt{N}$ are almost surely connected in a scale-free network, direct maximization of the likelihood for this subgraph would result in the collapse of all its nodes into a single point. We then proceed to embed each of the subsequent subgraphs $\mathcal{G}(k_l)$ sequentially. When we embed a subgraph $\mathcal{G}(k_l)$, we only consider the likelihood of the active nodes, that is, we compute Eq. (2.14) only for pairs of nodes i and j belonging to $\mathcal{G}(k_l)$. The steps to embed every subgraph $l = 2, \ldots, m$ are:

1. **Add new nodes:** For every node i in $\mathcal{G}(k_l)$ which has not been previously embedded, and whose number of connections with nodes in $\mathcal{G}(k_{l-1})$ is larger than 1, find the angular position θ_i maximizing its local log-likelihood,

$$\ln \mathscr{L}_i = \sum_{j \in \mathcal{G}(k_{l-1})} a_{ij} \ln p_{ij} + \sum_{j \in \mathcal{G}(k_{l-1})} \left(1 - a_{ij}\right) \ln\left(1 - p_{ij}\right). \tag{2.18}$$

   The maximization of the local likelihood, that is, of every node i, is a very interesting insight of this algorithm. Moreover, it is particularly relevant for the Internet, since every newcomer node could find its own position in the embedding without knowledge of the whole network topology. In terms of the global likelihood, the relation $\ln \mathscr{L} = 1/2 \sum_i \ln \mathscr{L}_i$ shows that increasing local log-likelihoods results in a net increase of the global likelihood as well (notice that maximizing the logarithm of Eq. (2.14) is equivalent to maximizing Eq. (2.14) itself. Yet, it is numerically more convenient to work with the logarithms of the connection probabilities).

2. **Adjust angular coordinates:** If l is smaller than or equal to some predefined l_{critic},

   i. Let $\mathcal{G}_{\text{active}}(k_l)$ be the subgraph of $\mathcal{G}(k_l)$ given by the nodes which have been already assigned an angular coordinate and the connections among them.

   ii. Choose a node i randomly from $\mathcal{G}_{\text{active}}(k_l)$ and find its angular coordinate θ_i maximizing

$$\ln \mathscr{L}_i = \sum_{j \in \mathcal{G}_{\text{active}}(k_l)} a_{ij} \ln p_{ij} + \sum_{j \in \mathcal{G}_{\text{active}}(k_l)} \left(1 - a_{ij}\right) \ln\left(1 - p_{ij}\right). \tag{2.19}$$

   Repeat this step until likelihoods converge.

The condition in step 2, that is, adjusting the coordinates only if $l \leq l_{\text{critic}}$, speeds the process noticeably. Typically, we set l_{critic} such that step 2 is carried out only for the first 1000 nodes, approximately. Moreover, in order to minimize computations, we do not calculate the local log-likelihood over all possible angular coordinates to maximize it. Instead, we first find the angular coordinate maximizing the likelihood over links only, that is,

$$\ln \mathscr{L}'_i = \sum a_{ij} \ln p_{ij}, \tag{2.20}$$

which approximates $\ln \mathscr{L}_i$ while requiring only $\mathscr{O}(k_i)$ computational steps. Once this coordinate θ'_i is known, we search for the optimal θ_i in the neighbourhood of θ'_i, this time considering the contribution of non-links into the likelihood as well. Finally, nodes of degree 1 and 2 in the network are given the same angular coordinate as its neighbour (the highest-degree one for degree-two nodes).

2.4.2 Algorithm for sequences of small networks

Most embedding algorithms [35, 155, 134, 8, 115, 30], including the one presented in the previous subsection, are designed to embed large networks. We now present a method to embed sequences of small networks (with less than 200 nodes), which we will denote as $\mathscr{G}_1, \ldots, \mathscr{G}_n$, using the $\mathbb{H}^2$ model, and such that the correlations between the coordinates found for the different networks in the sequence are least affected by the random component in the embedding process. This algorithm [2] will be used to embed yearly snapshots of the international trade system in Chapter 3.

2.4.2.1 Estimation of the global parameters

For every network $\mathscr{G}_i$ in the sequence, the value of the exponent γ can be estimated directly from the degree distribution measured on the network. We use a linear fit to the curve $\log P_c(k)$ versus $\log k$, where $P_c(k) = \sum_{k'=k}^{\infty} P(k')$ is the complementary cumulative degree distribution giving the probability for a randomly chosen node in the network to have degree greater or equal to k, less noisy than $P(k)$. For a power-law degree distribution with exponent γ, $P_c(k) \sim k^{1-\gamma}$, so the curve becomes the straight line $\log P_c(k) = (1-\gamma)\log k + c$, which allows to obtain γ from its slope. For such small systems, we constrain the values of $\gamma \geq 2.15$.

The values of the parameters N and $\langle k \rangle$, which are affected by finite-size effects and can be estimated analytically for large networks, as explained in the last subsection, need to be inferred using a different method in this case; instead, we propose to obtain the values of the three parameters N, $\langle k \rangle$ and β by comparing the networks to be embedded with synthetic networks generated with the model using different parameter values. We detail the procedure next:

1. Measure γ, N_{obs}, $\langle k\rangle_{\text{obs}}$ and $\bar{c}_{\text{obs}}$, i. e. the exponent of the degree distribution, the number of nodes, average degree and average clustering of the observed network to be embedded.

2. Initialize the target variables N_{targ}, $\langle k\rangle_{\text{targ}}$ and β_{targ} using some values as initial guess. In our case, we use $N_{\text{targ}} = N_{\text{obs}} + 10$, $\langle k\rangle_{\text{targ}} = 4\langle k\rangle_{\text{obs}}$ and $\beta_{\text{targ}} = 5.5$.

3. Generate a synthetic network using the target variables and the observed value of γ.

4. Measure N_{synth}, $\langle k\rangle_{\text{synth}}$ and $\bar{c}_{\text{synth}}$ of the synthetic network.

5. If $|N_{\text{synth}} - N_{\text{obs}}| < \varepsilon_N \wedge |\langle k\rangle_{\text{synth}} - \langle k\rangle_{\text{obs}}| < \varepsilon_{\langle k\rangle} \wedge |\bar{c}_{\text{synth}} - \bar{c}_{\text{obs}}| < \varepsilon_{\bar{c}}$, the synthetic network can be considered similar to the observed one, so the values N_{targ}, $\langle k\rangle_{\text{targ}}$ and β_{targ} used to generate it are good estimations to be used in Eq. (2.14). Keep them and stop. If the condition is not fulfilled, go to step 6.

6. If $N_{\text{synth}} < N_{\text{obs}} \Rightarrow N_{\text{targ}} + \delta_N \to N_{\text{targ}}$. Else $N_{\text{targ}} - \delta_N \to N_{\text{targ}}$.
   If $\langle k\rangle_{\text{synth}} < \langle k\rangle_{\text{obs}} \Rightarrow \langle k\rangle_{\text{targ}} + \delta_{\langle k\rangle} \to \langle k\rangle_{\text{targ}}$. Else $\langle k\rangle_{\text{targ}} - \delta_{\langle k\rangle} \to \langle k\rangle_{\text{targ}}$.
   If $\bar{c}_{\text{synth}} < \bar{c}_{\text{obs}} \Rightarrow \beta_{\text{targ}} + \delta_\beta \to \beta_{\text{targ}}$. Else $\beta_{\text{targ}} - \delta_\beta \to \beta_{\text{targ}}$.
   Go to step 3.

The implementation of this algorithm gives good estimations for the global parameters so that the resulting synthetic networks mimic the real ones. We use as final values the averages over 20 estimations. Moreover, we set $\varepsilon_N = 1$, $\varepsilon_{\langle k\rangle} = 0.1\langle k\rangle_{\text{obs}}$ and $\varepsilon_{\bar{c}} = 0.05$. δ_N is set to 1, while $\delta_{\langle k\rangle}$ and δ_β are withdrawn from the uniform distribution $U(0, 0.2)$ in every iteration of step 6.

Once N, $\langle k\rangle$, β and γ are known, the radius $\hat{R}$ of the hyperbolic disk can be computed using Eq. (2.13).

2.4.2.2 Inference of hyperbolic coordinates

Once the global parameters are known, Eq. (2.14) can be maximized with respect to the coordinates. We use an adapted version of the Metropolis-Hastings algorithm. Since we are not embedding a single network, but a sequence of them, we take as initial coordinates of a given node the ones obtained in the embedding of the previous network; it is therefore convenient to embed the sequence starting from the largest network and then the subsequent graphs in decreasing-size order to maximize the number of nodes with non-random initial coordinates. To avoid numerical issues, we maximize the logarithm of $\mathscr{L}$, which we will denote

as log-likelihood $\ln \mathscr{L}$, instead of $\mathscr{L}$, as with the previous algorithm. The steps to embed the network $\mathscr{G}_m$ are:

1. **Initial positions:** for every node i in the current network:

   i. If i belongs to the previous embedded network $\mathscr{G}_{m-1}$, set (r_i, θ_i) to the values of the $\mathscr{G}_{m-1}$ embedding (constrain r_i so that it is not larger than $\hat{R}$). Else, choose r_i randomly from $U(0, \hat{R})$ and θ_i from $U(0, 2\pi)$.

   ii. Set $(\tilde{r}_i, \tilde{\theta}_i) = (r_i, \theta_i)$. These variables will record the maximum-likelihood configuration.

2. **Compute the initial value of the log-likelihood** $\ln \mathscr{L}$ using Eq. (2.14). Set $\ln \tilde{\mathscr{L}} = \ln \mathscr{L}$. This variable will record the highest value of the log-likelihood.

3. **Maximize likelihood:** Repeat $S \times N^2$ times (typically, $S \approx 30$ is enough to achieve convergence):

   i. Let us relabel the most persistent hub in the sequence as node 1 without loss of generality. Choose a node i at random from the list $[2, N]$; notice that we never select node 1 in order to keep its angular coordinate unchanged. The reason to do so is that the likelihood $\mathscr{L}$ is invariant with respect to global rotations of the system. Hence, not varying the angular coordinate of one of the nodes anchors the embedding by breaking this symmetry.

   Propose a new value for θ_i from $U(0, 2\pi)$.

   Compute the log-likelihood increment. Notice that only $N-1$ steps are required, as opposed to calculating the new log-likelihood as in step 2.

   If $\Delta \ln \mathscr{L} \geq 0$ accept the change of θ_i. Otherwise, accept the change with probability $e^{\Delta \ln \mathscr{L}} = \mathscr{L}_{\text{new}} / \mathscr{L}_{\text{old}} < 1$.

   If the change is accepted, give $\ln \mathscr{L}$ the new value $\ln \mathscr{L} + \Delta \ln \mathscr{L} \to \ln \mathscr{L}$. If $\ln \mathscr{L} > \ln \tilde{\mathscr{L}}$ set $(\tilde{r}_i, \tilde{\theta}_i) = (r_i, \theta_i)$ for all i and $\ln \tilde{\mathscr{L}} = \ln \mathscr{L}$.

   ii. Choose a node i at random from the list $[1, N]$.

   Propose a new value for r_i from $U(0, \hat{R})$.

   Compute the log-likelihood increment.

   If $\Delta \ln \mathscr{L} \geq 0$ accept the change of r_i. Otherwise, accept the change with probability $e^{\Delta \ln \mathscr{L}} = \mathscr{L}_{\text{new}} / \mathscr{L}_{\text{old}} < 1$.

   If the change is accepted, give $\ln \mathscr{L}$ the new value $\ln \mathscr{L} + \Delta \ln \mathscr{L} \to \ln \mathscr{L}$. If $\ln \mathscr{L} > \ln \tilde{\mathscr{L}}$ set $(\tilde{r}_i, \tilde{\theta}_i) = (r_i, \theta_i)$ for all i and $\ln \tilde{\mathscr{L}} = \ln \mathscr{L}$.

For such small sizes, the embedding process does not take much time. Hence, for every network $\mathscr{G}_i$, we repeat the whole algorithm several times (for instance, 10) and keep the configuration with the maximum likelihood. We then run step 3 starting with such initial configuration but accepting changes if and only if $\Delta \ln \mathscr{L} > 0$, hence leading to a strict maximization of the likelihood function.

2.5 Beyond angular homogeneity

In the $\mathbb{S}^1$ and $\mathbb{H}^2$ models, as we have presented them so far, angular coordinates are distributed homogeneously. However, in real-network embeddings, we observe what was coined as *soft communities*, angular regions more densely populated than others. Real networks can then be partitioned into groups such that the average angular distances between nodes belonging to the same soft community are smaller than the corresponding expected average in the bare $\mathbb{S}^1/\mathbb{H}^2$ models. This, in turn, implies that the average angular distance between nodes belonging to different groups must be larger than expected in the homogeneous version of the model and, as a result, soft communities are separated by large angular gaps in the underlying metric space [155]. Interestingly, the topological community structure of the network [64], as defined for instance in [117, 137, 14], is highly correlated with such soft communities. Indeed, partitioning the network using the Critical Gap Method (CGM) algorithm, a geometric method that maximizes modularity considering the largest gaps between consecutive nodes along the circle as community boundaries [2], results in modularities comparable to those of other community detection methods currently available in the literature, but with higher resolution. We explore this point in detail in Chapter 3.

In this section, we address the question of whether the two models introduced in the previous sections can generate networks with given target topological features and soft communities, that is, *inhomogeneous* angular distributions. All the results included in this section were published in Ref. [1]. Although we present our results for the $\mathbb{S}^1$ model, they can be easily applied to the $\mathbb{H}^2$ using the isomorphism between them. Notice that, to derive the fundamental equivalence between hidden and expected degrees, Eq. (2.3), we assumed angular homogeneity, so considering heterogeneous angular distributions requires some level of adjustment over the hidden degrees. Therefore, we propose a general algorithm to find the corrected hidden-degree distribution yielding the desired degree distribution for *any* angular distribution of nodes. Even though the method is general enough to be applied on different angular distributions, in this section we illustrate our results on the particular angular distribution of the

Geometric Preferential Attachment (GPA) model [177]. This model, briefly introduced in Chapter 1, is a generalized version of the growing geometric model Popularity vs. Similarity Optimization model [133] in which soft communities emerge from the growth dynamics of the network without altering topological properties like the degree distribution or the clustering spectrum of the resulting network. The main reason why we choose to use this angular distribution in this section is that the angular distribution from Ref. [177] is not an arbitrary choice, but it rather emerges from a preferential attachment process in similarity space that seems to be a plausible explanation for the nature of communities in real systems. However, we would like to emphasize that our model is not a generalization of the GPA model, as it is valid for other angular distributions as well—for instance, a predefined angular density, as in Ref. [114]. Furthermore, our model is static; even though we use the sequential assignment of angles from the GPA, the connections are not established sequentially and, in fact, the size of the resulting network needs to be determined from the beginning.

The GPA is a growing model in which the angular coordinates and hidden degrees of different nodes are correlated. In contrast, in the bare $\mathbb{S}^1$ model, hidden degrees and similarity coordinates are typically assumed to be uncorrelated, so every node's hidden variables are withdrawn independently from some joint distribution $\rho(\kappa,\theta)$ that factorizes [152, 94]. In the GPA growth process, the degree of a node is determined by its age—the older the node is, the higher its degree. Moreover, 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 number of nodes $s < t$ at angular distance $\Delta\theta_{st} < 2/(s^{\frac{1}{\gamma-1}} t^{\frac{\gamma-2}{\gamma-1}})$, where γ is the exponent of the power-law degree distribution (we explain this dependence in the next subsection). This implies a very particular dependence between similarity coordinates and degrees: the angular coordinate of a node must depend on the angular coordinates of all nodes with higher degree.

Following a similar strategy to design a version of the $\mathbb{S}^1$ model able to generate soft communities, we must consider the impact of the degrees of the nodes in their implicit ordering in the circle. A potential way of action would proceed by first assigning a hidden degree κ_i from a power-law distribution $\rho_\kappa(\kappa_i) \sim \kappa_i^{-\gamma}$ to every node, ordering the nodes according to their hidden degrees, and reproducing the angular preferential attachment from the GPA model with that particular ordering. At the end of the process, we would obtain a set of N nodes with hidden degrees power-law distributed with exponent γ and the same angular distribution as the GPA model for that value of γ. However, the problem is not that simple and if we then connect every pair of nodes with the probabilities given by Eq. (2.4), degrees and hidden degrees would not be proportional. The reason for such deviation from the usual behaviour of the $\mathbb{S}^1$ model is that a ho-

mogeneous angular distribution is required for the proportionality between hidden and observed degrees, which is not fulfilled here by construction. Indeed, if two nodes with the same hidden degree κ have local geometric neighbourhoods with different densities, they will have different degrees simply because the distances d_{ij} in the connection probability, Eq. (2.4), are smaller on average in the higher-density neighbourhood. The solution is, therefore, that hidden degrees must depend on the spatial distribution of nodes as well. In the following subsection, we address this issue. We explore the inhomogeneous-similarity regime of the $\mathbb{S}^1$ model and show that it is capable of generating networks with power-law degree distributions, high clustering and soft communities.

2.5.1 Geometric Preferential Attachment in the $\mathbb{S}^1$ model

From the previous discussion, we see that hidden degrees and angles should be considerably entangled in the modelling of geometric networks with soft communities. In this context, the soft-communities $\mathbb{S}^1$ model requires the following steps:

1. *Assigning angular coordinates:* Angular coordinates are assigned according to the Geometric Preferential Attachment [177]. First, a label $i = 1,\dots,N$ is assigned to every node. Then, for every i from 1 to N:

   a. Sample i candidate angular positions $\phi_l, l = 1,\dots,i$ for node i from $U(0,2\pi)$.

   b. For every candidate position, define its attractiveness $A(\phi_l)$ as the number of nodes—belonging to the set of nodes with an already defined angular position, that is, with $s < i$—at angular distance $\Delta\theta_{ls} < 2/(s^{\frac{1}{\gamma-1}} i^{\frac{\gamma-2}{\gamma-1}})$, where γ is the exponent of the target power-law degree distribution.

   c. Assign to node i the angular coordinate of position l, i. e. set $\theta_i = \phi_l$, with probability
   $$\Pi(\phi_l) = \frac{A(\phi_l) + \Lambda}{\sum_{n=1}^{i} \left(A(\phi_n) + \Lambda\right)}. \tag{2.21}$$
   The initial attractiveness $\Lambda \geq 0$ is a parameter that sets the strength of the geometric preferential attachment. For very high values of Λ, all candidate angles become equally likely, so the resulting angular distribution is homogeneous as in the standard $\mathbb{S}^1$ model (see Fig. 2.3).

   This process generates a distribution of nodes in the circle analogous to the angular distribution of the GPA model. However, notice that, in the

GPA model, connections are established *at the same time* as positions are decided, whereas in the former steps, no connections have yet been made.

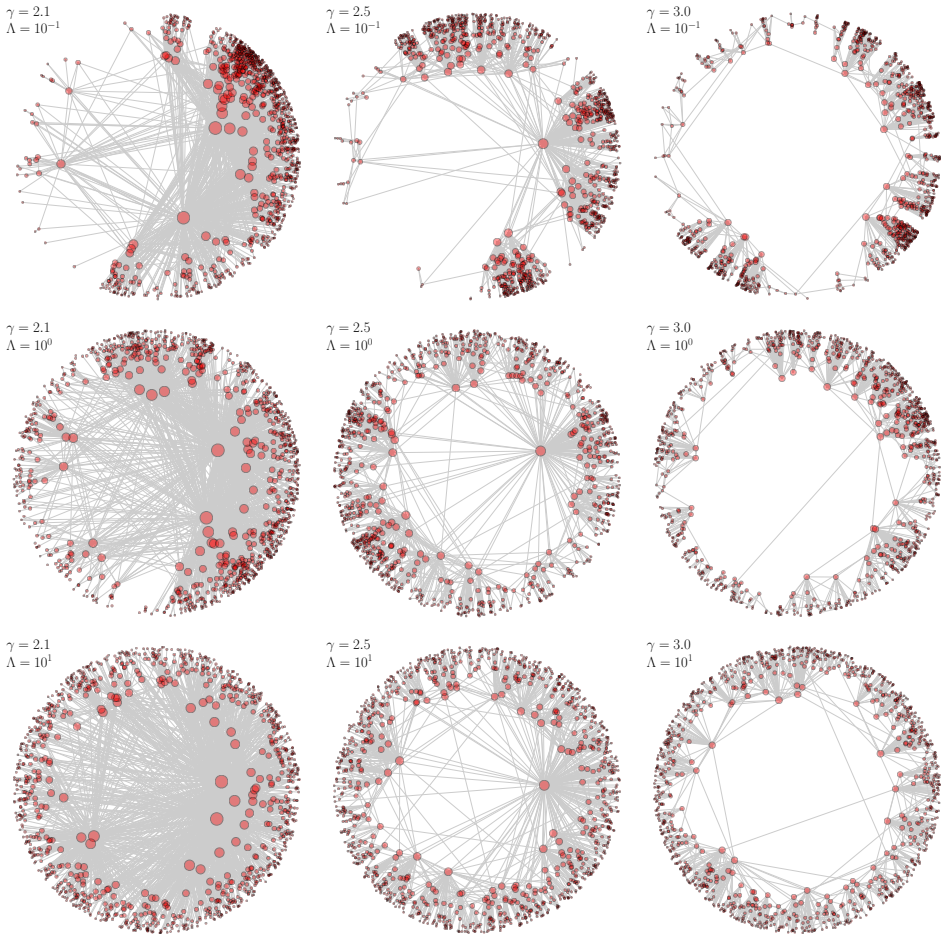


Figure 2.3: **Geometric layout of the networks generated by the soft-communities $\mathbb{S}^1$ model with Geometric Preferential Attachment.** In all cases, $N = 1000$ and $\beta = 2.5$. Every column corresponds to a value of γ and every row to a value of Λ. As in Ref. [177], soft communities emerge for low values of the initial attractiveness Λ. In order to clarify the figure, every node's target degree is represented as a radial coordinate $r_i = \hat{R} - 2\ln k_i^{\text{tar}}/k_N^{\text{tar}}$, where k_N^{tar} is the smallest target degree and $\hat{R} = 2\ln\left(N/(\pi\mu(k_N^{\text{tar}})^2)\right)$. When using the hidden degrees instead of the target degrees, this mapping constitutes the isomorphism between the $\mathbb{S}^1$ model and the $\mathbb{H}^2$ model in hyperbolic space [94, 35], Eqs. (2.6) and (2.7).

2. *Assigning hidden degrees:* Once every node has a defined angular position, we need to determine their hidden degrees such that the resulting observed degrees, that is, after the connections have been actually established, are power-law distributed with exponent γ. As mentioned earlier in this section, we must take into account that the spatial distribution is inhomogeneous (especially for low values of Λ), so that every hidden degree κ_i must be corrected accordingly. We propose the following method:

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

   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:

      i. Choose some node i randomly.

      ii. Compute the expected degree $\bar{k}_i$ of node i as

      $$\bar{k}_i = \sum_{j \neq i} \frac{1}{1 + \left(\frac{d_{ij}}{\mu \kappa_i \kappa_j}\right)^{\beta}}, \tag{2.22}$$

      where μ is given by Eq. (2.5). This might seem contradictory, since the result in Eq. (2.5) was obtained under the assumption of homogeneous angular distribution. However, notice that μ can be modified leaving all connection probabilities unchanged as long as hidden degrees are modified accordingly[2].

      iii. Correct the value of κ_i so that the expected degree $\bar{k}_i$ matches the target degree k_i^{tar}. We propose to reset $\left|\kappa_i + \left(k_i^{\text{tar}} - \bar{k}_i\right)\delta\right| \to \kappa_i$, where δ is a random variable withdrawn from the uniform distribution $U(0, 0.1)$. The random fluctuations of δ prevents the system from getting stuck at local optima. Other numerical methods could be used with the same end.

   d. Compute all relative deviations

   $$\varepsilon_i = \frac{\left|k_i^{\text{tar}} - \bar{k}_i\right|}{k_i^{\text{tar}}}. \tag{2.23}$$

   If $\max\{\varepsilon_i\}_i < \eta$, where η is a tolerance which we set to $\eta = 10^{-2}$, continue to step 3. Otherwise, repeat step 2c.

[2] For instance, the change $\mu' = C\mu$ and $\kappa_i' = \frac{1}{\sqrt{C}}\kappa_i$ yields $\mu' \kappa_i' \kappa_j' = \mu \kappa_i \kappa_j$.

3. *Establishing connections with the* $\mathbb{S}^1$ *model:* In the last step, we simply connect every pair of nodes according to Eq. (2.4). Since step 2 assigns a hidden degree to every node such that its expected degree matches its target degree, the resulting observed degrees in the network must be similar to the target degrees as well.

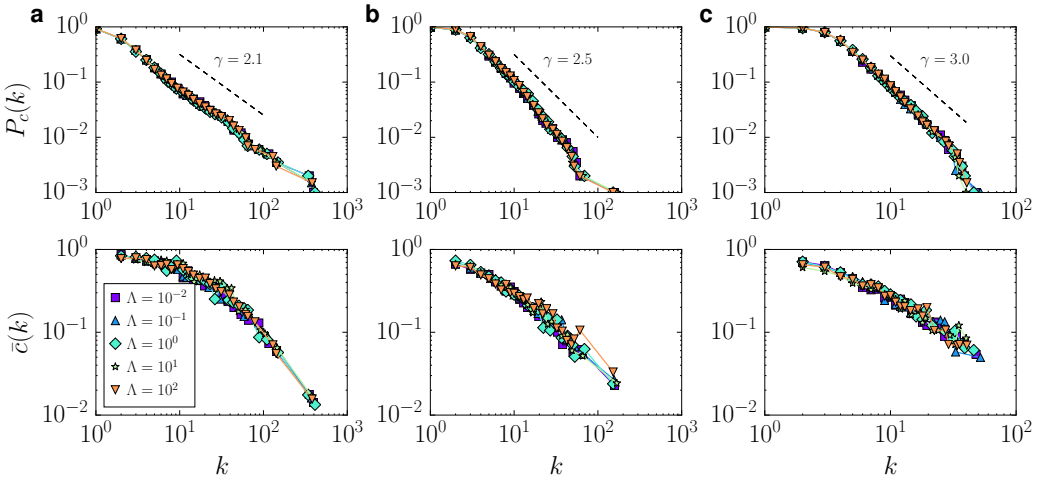


Figure 2.4: **Topological properties of the networks generated by the $\mathbb{S}^1$ model with Geometric Preferential Attachment.** All networks have been generated with $N = 1000$ and $\beta = 2.5$. Every color corresponds to a different value of the initial attractiveness Λ. Every column shows networks with different degree heterogeneities: $\gamma = 2.1$ (**a**), $\gamma = 2.5$ (**b**), and $\gamma = 3.0$ (**c**). The top row shows the complementary cumulative degree distributions. Black dashed lines indicate the scaling corresponding to each γ. In the bottom row, the mean local clustering spectra $\bar{c}(k)$ are drawn. To avoid fluctuations in the target degrees, all networks with the same exponent γ have been generated with the same target degree sequence $\{k_i^{\text{tar}}\}$. Despite their angular distributions being completely different (Fig. 2.3), their topologies are extremely similar.

Figure 2.3 shows the networks generated by the model for different values of γ and Λ. As in Ref. [177], the angular distribution has an evident soft-community structure for low values of Λ, whereas for high values of the initial attractiveness, the angular density resembles that of the homogeneous $\mathbb{S}^1$ model. Despite the considerable differences in the similarity distances between nodes for different values of Λ, the displayed networks are extremely similar from a topological perspective (see Fig. 2.4), with almost undistinguishable degree distributions and clustering spectra. Notice that step 2 in the algorithm above, corresponding to

the assignment of degrees, is not specifically designed for the GPA angular distribution[3]. In principle, it should be valid for other distributions as well. Therefore, the $\mathbb{S}^1$ model can adopt heterogeneous angular distributions as well, although, in general, we might need to correct every node's hidden degree numerically.

2.6 Higher-dimensional similarity space

Even though one single dimension yields a very powerful model, as we have discussed throughout these first two chapters, assessing the dimensionality of real systems and embedding them accordingly is a very relevant and challenging open problem in network geometry that could improve our understanding of complex networks considerably. Interestingly, the $\mathbb{S}^1$ model was originally introduced as the particular one-dimensional instance of the more general D-dimensional $\mathbb{S}^D$ model, in which similarity space is a D-sphere [152]. In this section, we study this model in detail and discuss what we can learn from it about the dimensionality of real similarity spaces.

2.6.1 The $\mathbb{S}^D$ model

In the $\mathbb{S}^D$ model [152], N nodes are scattered with uniform density on the surface of a D-dimensional sphere. As in the case of the $\mathbb{S}^1$ model, we fix the density of nodes to one, which implies that the D-dimensional area satisfies $A_D = N$. Every node i is also given a hidden degree κ_i from an arbitrary distribution $\rho_\kappa(\kappa_i)$ and every pair of nodes i and j is connected with probability [152, 36, 33]

$$p(\kappa_i, \kappa_j, d_{ij}) = p(\chi), \chi = \frac{d_{ij}}{(\mu\kappa_i\kappa_j)^{1/D}}, \tag{2.24}$$

where d_{ij} is the arc-length distance separating the two nodes on the sphere and $p(\chi)$ is a decreasing function taking values between 0 and 1.

Let us now study how degrees behave in the model. Although the following results were previously published [152, 36, 33, 4], we present an alternative derivation here. We start by taking the perspective of a node i, with hidden degree κ, laying somewhere on the D-sphere, and consider the probability for it to connect to a randomly chosen node in the system. Due to the homogeneity of the space and of the distribution of nodes, this probability cannot depend on the specific position of node i, but only on its hidden degree κ, so we may denote it by $\tilde{p}(\kappa)$. The reason for us to focus our attention on this probability is

[3] The ordering of the target degrees might not be necessary in a more general situation where, for instance, hidden degrees are not correlated with angles.

that, since all hidden variables—as well as connections—are assigned independently in the model, the degree of node i can be computed as the number of successful random events, each with probability $\tilde{p}(\kappa)$, out of the $N-1$ potential links reaching the node. Now, to compute it, we need to integrate the connection probability, Eq. (2.24), over all possible values of the hidden degree κ' of the randomly chosen node as well as of the distance among the nodes, weighted by their corresponding distribution functions[4], that is,

$$\tilde{p}(\kappa) = \int_{\kappa_0}^{\kappa_c} \int_{0}^{\pi R} p\left(\kappa, \kappa', d\right) \rho_{\kappa}\left(\kappa'\right) \rho_{d}(d) \, \mathrm{d}\kappa' \mathrm{d}d. \tag{2.25}$$

In the above expression, R is the radius of the D-sphere which, according to the unitary density of nodes, scales as $R \sim N^{1/D}$. The distance distribution among two randomly sprinkled points in the space, $\rho_d(d)$, can be written as

$$\rho_{d}(d) = \frac{A_{D-1}(d)}{A_D}, \tag{2.26}$$

where $A_{D-1}(d)$ is the $(D-1)$-dimensional area of the region of points laying at distance d from node i on the D-sphere[5]. We can give Eq. (2.26) the following interpretation. Since all volume elements on the space have the same probability of being occupied by the randomly chosen node, the probability for it to be at distance d from i is simply proportional to the amount of volume elements at such distance from i. In any case, this distribution encodes the only dependence on the geometry of the space in Eq. (2.25).

To proceed further with the calculation of $\tilde{p}(\kappa)$, some approximations are in order. In particular, we note that, since $p\left(\chi\right)$ is a decreasing function of d, the only contribution to the integral comes from the region of small distances. Therefore, for a sufficiently large system size N, the integrand is non-negligible on a region of the D-sphere in which the curvature of the space can be neglected, as it is locally Euclidean. Under these conditions, $A_{D-1}(d)$ can be approximated as the area of a $(D-1)$-dimensional sphere of radius d embedded in D-dimensional Euclidean space, that is,

$$A_{D-1}(d) \approx \frac{2\pi^{D/2}}{\Gamma\left(\frac{D}{2}\right)} d^{D-1}. \tag{2.27}$$

[4]This can be derived as follows. According to our definition of $\tilde{p}(\kappa)$, it represents $\rho(a_{\kappa\kappa'}|\kappa)$, the probability distribution for the adjacency-matrix element $a_{\kappa\kappa'}$ conditional to the hidden degree of node i, evaluated at $a_{\kappa\kappa'} = 1$, that is, $\tilde{p}(\kappa) = \rho(a_{\kappa\kappa'} = 1|\kappa)$. This conditional probability can be obtained by marginalizing and applying Bayes' rule appropriately to the joint probability distribution for all the random variables involved among the two nodes, $\rho\left(a_{\kappa\kappa'}, \kappa, \kappa', d\right)$.

[5]The exact expression for the distribution of distances is derived in Section 2.8.1. However, since it is not explicitly used in this derivation, we prefer not to include it here.

Furthermore, since we are assuming the integrand to become very small for large distances d, integrating distances to ∞ instead of to πR should give a very good approximation of $\tilde{p}(\kappa)$. Hence, we can finally write

$$\begin{aligned}\tilde{p}(\kappa) &\approx \frac{1}{A_D}\frac{2\pi^{D/2}}{\Gamma\left(\frac{D}{2}\right)}\int_{\kappa_0}^{\kappa_c}\int_0^{\infty} p\left(\kappa,\kappa',d\right)\rho_\kappa\left(\kappa'\right)d^{D-1}\mathrm{d}\kappa'\mathrm{d}d \\ &= \frac{1}{A_D}\frac{2\pi^{D/2}}{\Gamma\left(\frac{D}{2}\right)}\mu\kappa\int_{\kappa_0}^{\kappa_c}\kappa'\rho_\kappa\left(\kappa'\right)\mathrm{d}\kappa'\int_0^{\infty}\chi^{D-1}p\left(\chi\right)\mathrm{d}\chi \\ &= \frac{1}{A_D}\frac{2\pi^{D/2}}{\Gamma\left(\frac{D}{2}\right)}\mu\kappa\langle\kappa\rangle\int_0^{\infty}\chi^{D-1}p\left(\chi\right)\mathrm{d}\chi,\end{aligned} \tag{2.28}$$

where the remaining integral depends on the details of the connection probability function only.

We can now characterize the degree distribution for node i, to which we may refer as $\rho(k|\kappa)$. As we mentioned earlier, this quantity is given by the number of successful events out of the $N-1$ potential connections, each with probability $\tilde{p}(\kappa)$, so we can immediately see that the degree of i follows a Binomial distribution,

$$\rho(k|\kappa) = \binom{N-1}{k}\tilde{p}(\kappa)^k\left(1-\tilde{p}(\kappa)\right)^{N-1-k}, \tag{2.29}$$

with expected value given by

$$\bar{k}(\kappa) = (N-1)\tilde{p}(\kappa) \approx \frac{2\pi^{D/2}}{\Gamma\left(\frac{D}{2}\right)}\mu\kappa\langle\kappa\rangle\int_0^{\infty}\chi^{D-1}p\left(\chi\right)\mathrm{d}\chi, \tag{2.30}$$

where we have further approximated $(N-1)/N \approx 1$. Since μ is a free parameter, we can set it to

$$\mu = \frac{\Gamma\left(\frac{D}{2}\right)}{2\pi^{D/2}\langle\kappa\rangle\int_0^{\infty}\chi^{D-1}p\left(\chi\right)\mathrm{d}\chi}, \tag{2.31}$$

for which $\bar{k}(\kappa) = \kappa$ and, therefore, $\langle k\rangle = \langle\kappa\rangle$. This is a critical property of the model. Regardless of our choice of connection probability $p\left(\chi\right)$ (provided that the above integral converges), we can find a value of μ allowing us to control the expected degree of every node individually. Furthermore, notice that, according to the Poisson limit theorem, Eq. (2.29) can be approximated by a Poisson distribution [34],

$$\rho(k|\kappa) = e^{-\kappa}\frac{\kappa^k}{k!}, \tag{2.32}$$

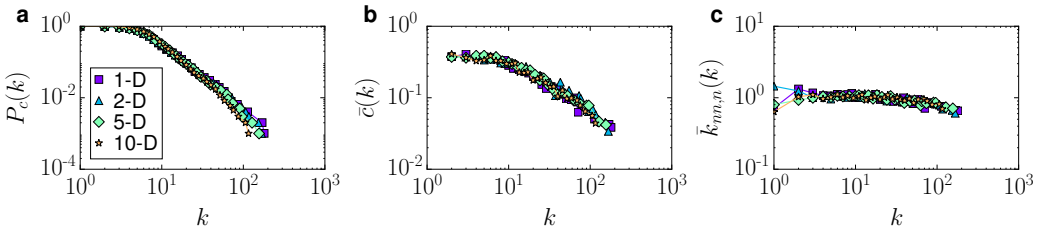


Figure 2.5: **Single realizations of the $\mathbb{S}^D$ model.** Topological properties of networks generated with the model in different dimensions. All four networks share the same hidden-degree sequence, power-law distributed with exponent $\gamma = 2.5$, while the global parameters μ and β have been adjusted so that they all have the same average degrees $\langle k \rangle = 10$ and clustering coefficients $\bar{c} = 0.32$. **a**, Complementary cumulative degree distributions. **b**, Degree-dependent clustering spectra. **c**, Normalized average nearest neighbours degrees.

which, in turn, implies that the fluctuations of the degree of node i are given by $\mathrm{Var}(k) = \kappa$. Thus, the relative fluctuations scale as $\sqrt{\mathrm{Var}(k)}/\kappa = 1/\sqrt{\kappa}$, so they vanish for high-degree nodes. Finally, we can compute the resulting degree distribution as

$$\begin{aligned} P(k) &= \int \rho(k,\kappa)\,\mathrm{d}\kappa = \int \rho(k|\kappa)\,\rho_\kappa(\kappa)\,\mathrm{d}\kappa \\ &= \int e^{-\kappa}\frac{\kappa^k}{k!}\rho_\kappa(\kappa)\,\mathrm{d}\kappa, \end{aligned} \tag{2.33}$$

which evidences its close relation with the hidden-degree distribution $\rho_\kappa(\kappa)$; for very high values of k, $\rho(k|\kappa)$, as a function of κ, becomes highly peaked around $\kappa = k$, so the degree distribution behaves as the hidden degree distribution asymptotically.

Although other properties of the model—like clustering or degree-degree correlations—are difficult to study analytically, they also exhibit similar behaviours at different dimensions (yet, there is a dimension-dependent limit in the maximum clustering attainable, which will be discussed in the next subsection). This can be seen in Fig. 2.5, where we show several features of four networks with dimensions between 1 and 10. They have been generated using the particular connection probability in Eq. (2.4).

2.6.2 Clustering and dimensionality

The $\mathbb{S}^D$ model reveals that it is also possible to obtain realistic networks from higher-dimensional similarity spaces so, in principle, nothing prevents real-world networks from having underlying similarity spaces with $D > 1$. It is there-

fore reasonable to ask whether a one-dimensional similarity-space representation of real-world networks suffices. Although this question still requires further research, in this subsection we will discuss two important points that shed some light to it: first, the high clustering coefficient observed in networks poses an upper limit on the similarity-space dimension [3]. Second, for low-dimensional similarity-space networks, the one-dimensional embedding provides a faithful representation.

The first statement is a consequence of the following observation. As the dimension of a D-sphere increases, the distances between points randomly scattered on its surface become more homogeneous, as shown in Fig. 2.6**a** (see also Eq. (2.39)). As a consequence, the distances in the connection probability eventually become a constant, hence making the model degree-dependent only, which yields a zero-clustering behaviour in the thermodynamic limit. This suggests that the maximum clustering coefficient attainable decreases with the dimension of the underlying space which, in turn, implies that the dimension of the similarity space of a given network with clustering is upper-bounded. To quantify this effect, we have computed the mean local clustering of synthetic networks with $\beta = \infty$—for which the clustering is maximal, since the connection probability Eq. (2.4) becomes a step function—, as a function of the dimension for different values of γ. As shown in Fig. 2.6**b**, for a typical clustering level $\bar{c} > 0.5$, the dimension of similarity space can be, *at most*, $D \approx 10$ for $\gamma = 2.1$ ($D \approx 4$ for $\gamma = 3$). Hence, real similarity-spaces must be low-dimensional.

As for the second point, in order to assess the possible effect of the dimensionality reduction, we have embedded the four networks presented in Fig. 2.5 on the $\mathbb{S}^1$ model and measured the quality of the mapping by comparing the original connection probabilities to their counterparts obtained from the embedding (see Fig. 2.6**c**). Although there is some distortion due to the projection onto a lower-dimensional space, the effect on the connection probabilities is mild. Consequently, the one-dimensional embedding is able to recover most of the original geometric information. This is further supported by the fact that a renormalization procedure yields almost indistinguishable flows when applied to all four networks, as we will show in Chapter 4.

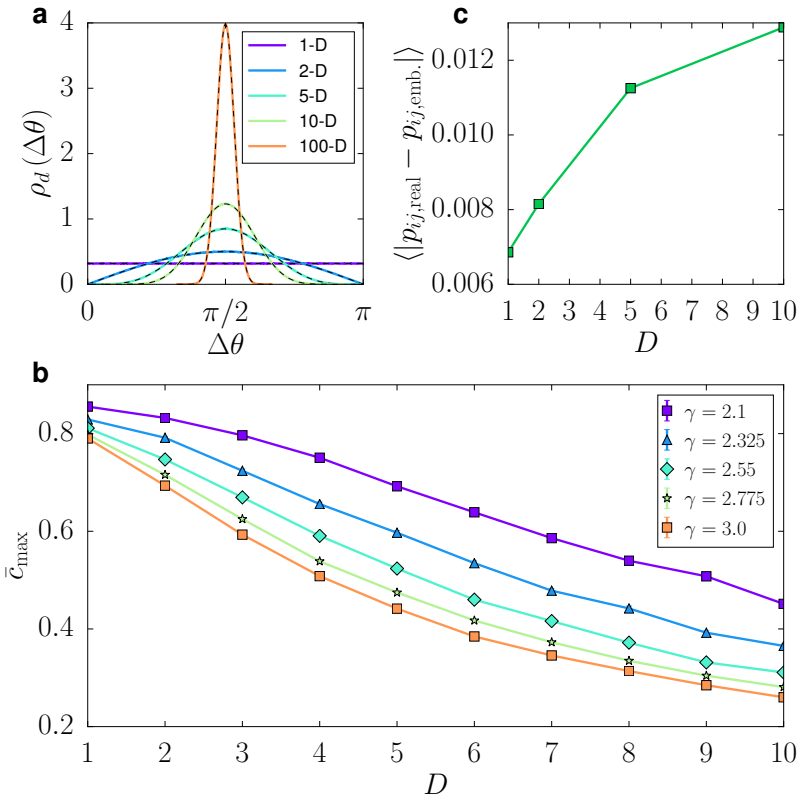


Figure 2.6: **Clustering and dimensionality reduction.** **a**, Distances on the D-sphere. Every color curve corresponds to the distribution of angular distances between $N = 10^4$ randomly scattered points on a D-dimensional sphere. Black dashed lines stand for the corresponding theoretic result, Eq. (2.39). As the dimension increases, the distances become more homogeneously distributed around the mean, $\langle\Delta\theta\rangle = \pi/2$, hence decreasing the clustering in the model. **b**, Clustering coefficient at $\beta = \infty$ as a function of the dimension and γ. The average clustering corresponds to the mean local clustering averaged over 100 realizations on networks of size $N = 1000$; the errorbars, which are smaller than symbol sizes, represent the standard error of the measurement. The value of μ has been adjusted so that the observed average degree is $\langle k\rangle = 10.0 \pm 0.1$ in all networks. **c**, Effects of dimensionality reduction measured through the absolute differences between the real connection probabilities $p_{ij,\mathrm{real}}$, that is, the ones with which the networks were generated, and the embedding connection probabilities $p_{ij,\mathrm{emb.}}$ averaged over all pairs of nodes. Notice that, even in the case of original dimension $D = 1$, the embedding process introduces some distortion to the obtained probabilities. For higher dimensions, the effect of the actual geometric distortion is added; yet, the average error in the connection probabilities remains small.

2.7 Discussion

There is abundant evidence of the geometric origin of many properties of complex networks, not only regarding their topology [35, 155, 82], but also their weighted organization, as we present in Chapter 5. The field of network geometry has therefore attracted much attention recently, and the $\mathbb{S}^1/\mathbb{H}^2$ models stand as cornerstones in this area or research. On the one hand, they provide an intuitive and plausible explanation for clustering in real networks by introducing the concept of similarity space. Indeed, it has even been shown that the presence of clustering implies a geometric origin of sparse networks [95]. On the other hand, they allow to build geometric maps of real networks by embedding them. These maps are remarkably meaningful, to the extent of predicting symmetries in real systems, as will be explored in Chapter 4. In addition, they are very useful; they can be used to navigate the network efficiently [35], to detect communities [2] or even to construct smaller-scale replicas of real networks for efficiently testing dynamics on real networks [3]. The last two points will be explored further in this thesis.

So far, these two models have only been studied under several simplifying premises, like power-law degree distributions or independent hidden variables. Yet, they have been able to explain many observed phenomena in complex networks. However, they can be exploited beyond these assumptions, since the correlation between hidden degrees and angles might clarify many more topological features of real-world networks. The results in Section 2.5 open the path towards such line of study by showing that the model *does not require* those simplifying assumptions. As a result, we can now generate complex networks with soft communities for any given angular distribution; although we have used as particular example the GPA angular distributions, our method is valid for other angular distributions as well.

Finally, we also include a description of the higher-dimensional $\mathbb{S}^D$ model, which shows that higher-dimensional similarity spaces can generate realistic networks as well. Although using this model to embed real networks is currently beyond our technical capabilities, a very important conclusion can be drawn from its analysis. As the dimension of the underlying space increases, the maximum level of clustering attainable decreases, so the similarity spaces of highly clustered networks must be low-dimensional. This further explains the generally observed goodness of real-world network embeddings on a single similarity space dimension, as the distortion inflicted in the embedding of a low-dimensional network is limited.

2.8 Additional information

2.8.1 The distribution of distances on a D-sphere

In this section we present a derivation of the exact distribution of distances between randomly sprinkled points on a D-sphere which, to the best of our knowledge, has not been published in the complex networks literature.

The distribution can be obtained by immersing the sphere in $\mathbb{R}^{D+1}$. In this space, the D-sphere is defined as the set of points $\mathbf{x} = (x_1, \ldots, x_{D+1})$ satisfying

$$\sum_{i=1}^{D+1} x_i^2 = R^2. \tag{2.34}$$

We are now interested in the volume of the region of points on the sphere laying at distance d from a given point $\mathbf{x}^0$. Since distances on the sphere are measured as arc-lengths, hence proportional to the radius R, the distance between two points is given by $R\Delta\theta$, where $\Delta\theta$ is the angle between the two positioning vectors in $\mathbb{R}^{D+1}$. Therefore, any point $\mathbf{x}$ on the sphere at distance d from $\mathbf{x}^0$ must satisfy

$$\mathbf{x}\cdot\mathbf{x}^0 = R^2 \cos\frac{d}{R} \tag{2.35}$$

Without loss of generality, let $\mathbf{x}^0$ be given by $x_i^0 = 0, i \leq D$ and $x_{D+1}^0 = R$, so that Eq. (2.35) determines the $(D+1)$-th component of $\mathbf{x}$,

$$x_{D+1} = R\cos\frac{d}{R}. \tag{2.36}$$

With this constraint, Eq. (2.34) becomes

$$\sum_{i=1}^{D} x_i^2 = R^2 - R^2\cos^2\frac{d}{R} = R^2\sin^2\frac{d}{R}. \tag{2.37}$$

Interestingly, this is the condition fulfilled by the points of a $(D-1)$-sphere of radius $R\sin\frac{d}{R}$. With this observation, we can straightforwardly write

$$A_{D-1}(d) = \frac{2\pi^{D/2}}{\Gamma\left(\frac{D}{2}\right)} R^{D-1}\sin^{D-1}\frac{d}{R}. \tag{2.38}$$

Clearly, this result is consistent with our Euclidean approximation, as it yields Eq. (2.27) for small d/R. Furthermore, we can also see that the distribution of distances, Eq. (2.26), is

$$\rho_d(d) = \frac{\Gamma\left(\frac{D+1}{2}\right)\sin^{D-1}\left(\frac{d}{R}\right)}{\Gamma\left(\frac{D}{2}\right)\sqrt{\pi}R}, \tag{2.39}$$

which becomes increasingly peaked around $d = \pi R/2$ as $D \to \infty$, hence hindering the formation of triangles in the network.

CHAPTER 3

The World Trade Atlas 1870–2013

This chapter addresses the study of the evolution of the international trade system from the perspective of network geometry. It reviews in detail our approach to explain the existence of significant trade interactions between pairs of countries in the world based on our network gravity model: the $\mathbb{S}^1$. We also describe the pure geometric $\mathbb{H}^2$ counterpart in hyperbolic space using the isomorphism introduced in Chapter 2. All the results presented in this chapter correspond to Ref. [2].

3.1 Beyond detached bilateral flows and geographic distance

International trade annually moves trillions of dollars and represents one of the main evolving networks of interactions between countries in the world. Beyond the exchange of goods and services, it interplays with other economic, political and social systems and sustains the propagation of different processes, both beneficial or harming, such as innovations or crises. All these interactions shape the complex architecture of the World Trade Web (WTW), which presents the typical features of complex networks discussed in the introduction of the thesis, such as the small-world property, heterogeneous distributions of the number of trade partners, high levels of triadic closure relationships, community structure and non-trivial global distribution of flows [149, 72, 151, 147, 154].

However, many studies still focus on bilateral relationships to understand the international trade system. Bilateral flows have traditionally been modelled using the gravity model of trade which, in analogy to Newton's law of gravitation, accurately predicts that the volume of trade exchanged between two countries increases with their economic sizes and decreases with their geographical separation [162, 26, 158]. The precision of the gravity model improves when it is supplemented with other factors, such as colony-colonizer relationships, a shared common language, or the effects of political borders and a common currency [66, 12, 25]. Despite the success of the gravity model at replicating trade volumes, it performs very poorly at predicting the existence of trade channels between pairs of countries [59]; an obvious limitation that prevents it from explaining the striking regularities observed in the complex architecture of the

WTW. One of the reasons for this flaw is that the gravity model focuses on detached bilateral relationships and so overlooks multilateral trade resistance and other network effects. Another drawback is that geography is not the only factor that defines distance in international trade. Distance matters but along many other dimensions.

Moreover, trade, transport, and communication have evolved boosted by waves of globalization, as periods when a complex series of closely intertwined changes have dramatically increased the interactions and interdependencies between a growing number of people and human organizations in disparate geographic regions. As a result, the world economies exhibit nowadays a high level of global interdependency. Therefore, the analysis of international trade requires a global perspective, taking into account every country and its trade relationships regardless of its population or wealth, which can be based on the representation of the international trade system as a complex network of interactions. Beyond rich and informative visual representations, the complex network approach is producing a breakthrough in our understanding of the large scale structure, behaviour and evolution of the international trade system [149, 72, 151, 147, 154, 9, 2].

The World Trade Web is the representation of the international trade system as a complex network [149] in which countries are represented as nodes and commercial relations between them as links. This level of description goes beyond the country-specific and bilateral analyses often used in traditional economic studies of trade. Yearly snapshots of the WTW, reconstructed from empirical data, feature robust non-trivial signatures typical of complex networks across domains, which help in explaining emergent properties and systemic effects like, for instance, unexpected economic influences of one country on another even if their trade exchange is absent or not significant [151]. Next, we revise the most basic topological properties of the WTW as a complex network, and discuss how the complex network perspective can be reconciled with the role of distance in international trade by means of the geometric models introduced in Chapter 2.

By embedding the yearly snapshots of the WTW on the $\mathbb{S}^1$ model, we obtain a description of the international trade system in terms of a network gravity model representing the existence of significant trade channels among countries. This model relies on economic sizes, which stand for popularity in this particular context, and on effective similarity distances that incorporate the different dimensions affecting international trade—including but not limited to geography—implicitly encoded on the complex patterns of trade interactions. The isomorphism with the $\mathbb{H}^2$ model brings us a pure geometric description of the WTW in hyperbolic trade space. In this representation, economic sizes and geographical distances are combined into a sole distance metric as a proxy of

aggregate trade barriers, such that the closer countries are in hyperbolic trade space the greater their chance of being connected. We estimate trade distances from empirical data to represent the WTW through World Trade Maps. The maps are annual and cover a time span of fourteen decades. The collection as a whole is referred to as the World Trade Atlas 1870-2013, and can be examined using our online interactive video tool [69].

Beyond the obvious advantages of visualization, the World Trade Atlas increases significantly our understanding of the long-run evolution of the international trade system and helps us to address a number of important and challenging questions. In particular, which is the role of distance in trade, not merely in geographical terms, and how far or close have countries been through recent history? Which role plays each country in the maps and how has this role evolved in time? Are Preferential Trade Arrangements (PTAs) consistent with natural communities as measured by trade distance? Has the formation of PTAs led to lower or higher barriers to trade inside blocs? Are historical, economic, and geographical divisions becoming increasingly irrelevant? The answers lead to a clear picture in which the evolution of the the international trade system has been shaped by three simultaneous forces: globalization, hierarchization, and localization. We find that the international trade system is becoming more heterogeneous and hierarchical than ever before and, at the same time, more interconnected and clustered into natural trade communities not fully consistent with Preferential Trade Agreements.

3.2 The World Trade Web as an undirected complex network

The WTW is customarily constructed from public databases reporting import and export exchanges of merchandises between pairs of countries in the world. Exports render an inflow of money into a country and an outflow of goods, while imports generate an outflow of money and an inflow of goods. By selecting and combining import and export information between every pair of countries, different kind of complex network representations with different levels of detail and information content can be constructed, being the most basic one that in which a link between a pair of countries stands for the existence of a trade channel, independently of its directionality or intensity. More sophisticated accounts can be constructed if the directionality of the bilateral flows or the amounts exchanged are considered [151, 147, 154]

However, signatures of the complex organization of the WTW are obscured by the dramatical increase in the density of connections, in the total amount traded in the world, and in the relative average geographic length of trade chan-

nels in the last decades, which has augmented the entanglement of the network despite the growth in the number of countries N in the international trade arena, see Fig. 3.1. The progressive crumbling of the Imperial Colonies during the sixties and seventies and the former Soviet Union contributed to this fact. The average number of trade partners per country $\langle k \rangle$ increases at a faster rate than that of the number of countries N, Fig. 3.1**a**. As a consequence, the density of connections $\langle k \rangle / (N-1)$ has boosted from 0.2 to 0.7 in the period under study. The temporal evolution of the total international trade—measured as a percentage of the sum of all national gross domestic products (GDPs)—presents a similar trend, with a major reversal in the interwar periods, Fig. 3.1**b**. At the same time, the average geographic length of trade channels has been growing as compared to the average distance among all pairs of countries in the system, Fig. 3.1**c**, proving that trade channels have been progressively established with more geographically distant partners.

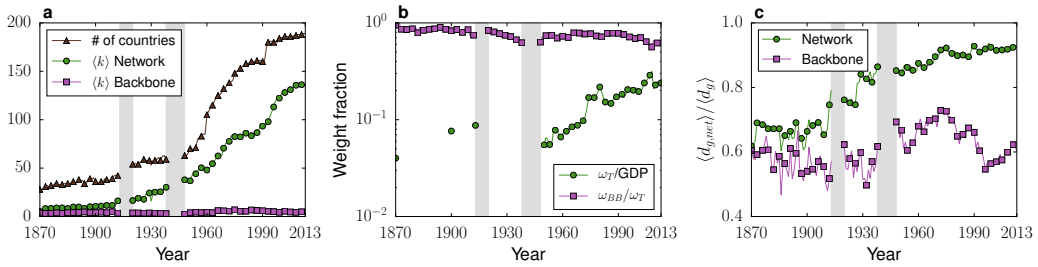


Figure 3.1: **Time evolution of scalar variables in the World Trade Web. a**, Number of countries, and average number of trade partners per country, $\langle k \rangle$, in the WTW and the corresponding backbone. **b**, Ratio between total trade ω_T and world's GDP. The sparsity of points before the WWII is due to lack of GDP data. The plot also shows the fraction of total weight that is retained in the backbone (70% ~ 80%). **c**, Ratio between the average geographic distance of trade channels $\langle d_{g,net} \rangle$ and the average distance among pairs of countries in the world $\langle d_g \rangle$ in the WTW and the corresponding backbone.

In parallel, world trade networks display, especially after World War I (WWI), strong global and local heterogeneity which can be exploited to filter out a sparse subnetwork representing the relevant structure that remains after eliminating the contingent interactions which overshadow the information contained in the system [153]. These backbones, see details in next subsection, retain most of the countries but the number of trade channels is drastically reduced to the statistically significant ones. At the same time, pivotal features denoting the complexity of the international trade system are clearly exposed. The topology of WTW backbones is small-world, so that each pair of countries is very close to

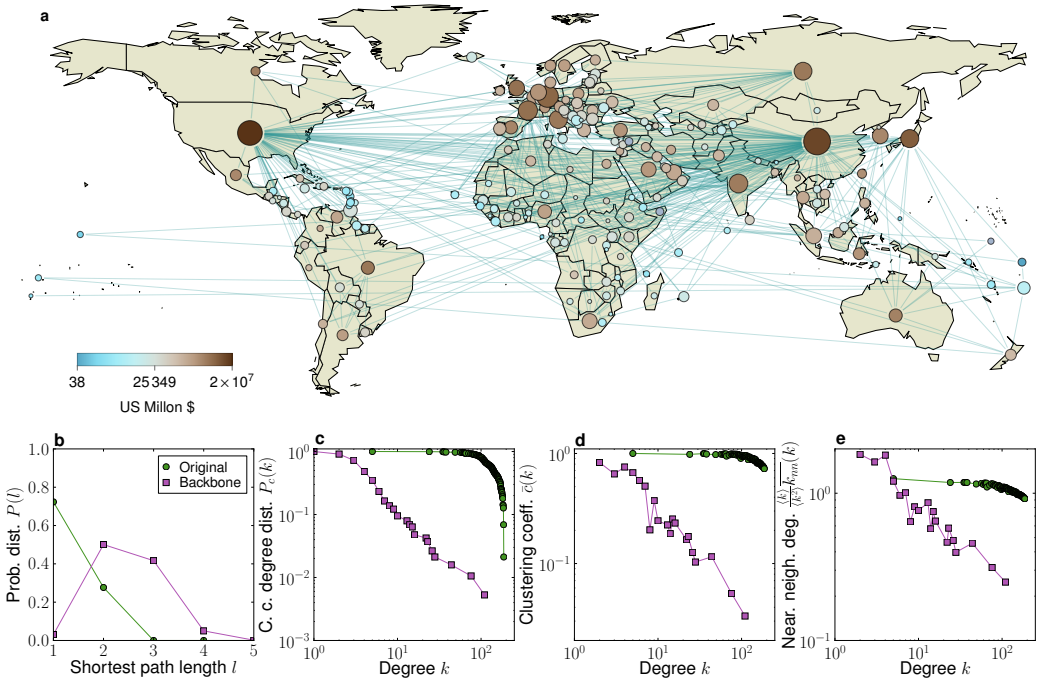


Figure 3.2: **Backbone of the World Trade Web in 2013. a**, Nodes represent countries and their sizes are proportional to the logarithm of the number of trade partners in the backbone. Nodes are colored according to the logarithm of GDP values. Topological properties: **b**, Distribution of shortest path lengths. **c**, Complementary cumulative degree distribution. **d**, Average degree-dependent clustering coefficient. **e**, Normalized average degree of nearest neighbors.

each other in topological distance, with an average value very similar to the one expected for a random network of the same size and average degree; it shows a quite neat scale-free distribution of the number of trading partners of countries; it presents disassortative degree-degree correlations, meaning that highly interconnected economies tend to connect to poorly connected ones; and it displays a high level of clustering—much greater than the value corresponding to a random network of the same size—that decreases as a function of degree indicating that countries trading with a well connected one are poorly interconnected among them. This result, together with the scaling law for two-point correlations, reveals a hierarchical architecture of highly interconnected countries that belong to influential areas that, in turn, connect to other influential areas through hubs. The hierarchical structure of a complex network can also be detected through another important degree correlation measure, the rich club coefficient, which measures the tendency of progressively higher con-

nected nodes to club forming elites [51] or to polarize in competing blocks [148]. The WTW shows a very clear and sharp rich-multipolarization phenomenon with hub countries sharing on average less trade among them than expected in an unbiased situation—they form fully connected subsets in the weighted representation—and polarized into two tightly connected blocks: the United States and Japan and China, against France, Germany, and the United Kingdom, with the separation between the two blocks exposed by a reduced exchange of merchandise as compared to an unbiased situation [148]. All these properties make the World Trade Web a complex network, which is far from being well described through a classical portrait based on detached bilateral relationships. We show measurements of some of these features for the WTW in the year 2013 in Fig. 3.2.

As a consequence, backbones give a very convenient representation of the most significant trade channels and allow to mine and map the international trade system. In the following subsection, we explain in detail how to reconstruct the networks from empirical data and to extract their backbones.

3.2.1 Data and network reconstruction

In order to produce World Trade Maps of the international trade system, the first step is to construct a WTW for every year in which the links between the countries carry a weight ω_{ij} equal to the sum of the trade flows from i to j and from j to i, i. e. $\omega_{ij} = \omega_{i\to j} + \omega_{j\to i}$, where the arrow represents the direction of the flow of goods. Hence, the resulting networks are weighted and undirected. Bilateral trade data reported by countries can be used. Due to a reliability problem, we avoid to work with export data whenever possible. More accurate reported import values of bilateral trade data can be used instead, since an export from one country to another is simultaneously an import from the partner; see details in next subsection.

3.2.1.1 Empirical data

Bilateral trade data to reconstruct weighted versions of the WTW can be found in the websites of the International Monetary Fund (IMF) [86] and the United Nations Commodity Trade Statistics Database [164], among others. More conveniently, datasets with curated versions are also available [77, 19]. We used the dataset provided in The Correlates of War Project (COW) [55], with curated bilateral trade flows from 1870 to 2009 [18] based on values reported by each country to the IMF. Due to some detected flaws in the data for the last years—like the absence of trade data for Germany after 2006—, we restricted the use of this dataset, referred to as DBI, to the period from 1870 to 1996 (with the only modi-

fication of the removal of the flows between Haiti and United Kingdom in 1877, which seem to be outliers). We complemented it with more recent data from the IMF and followed the same curation steps to produce a version with world bilateral trade data from 1997 to 2013, referred to as DBII [2].

Different issues have to be taken into account in the curation process [19]. The first step is to build up a correspondence between IMF and COW codes for countries and to extend the COW's state list from 2011 to 2013. In most cases, the matching is very clear, but there are some special cases. A given country may appear under different names at the IMF. For instance, in the last years some countries still report trade with the U.S.S.R. A different problem is the case of China, whose trade data does not include the one for the two special administrative regions Hong Kong and Macao, which have been added to China. Also, many entries in the COW state list can correspond to the same country, like"Yugoslavia", "Yugoslavia not specified", "Serbia and Montenegro", "Serbia and Montenegro not specified", "Serbia, Republic of", "Montenegro" and "Kosovo" before its independence. Also, several countries report trade with pairs of countries, like "Belgium-Luxembourg not specified". Following the same steps as the COW Project, we have split every such flow into two, proportionally to the GDP of Belgium and Luxembourg the corresponding year (the GDP data of both countries for this operation was retrieved from the World Bank [167]). Finally, there are no trade reports from Taiwan in the IMF database. We have obtained its data from the ROC's Bureau of Foreign Trade [139], matching their country list with the one from the COW Project.

Once the correspondence between codes is established, trade flows can be assigned to links between countries from IMF data. Four numbers are reported for every pair of countries i and j: $\omega^i_{i\to j}$, $\omega^j_{i\to j}$, $\omega^i_{j\to i}$ and $\omega^j_{j\to i}$, where the upper index represents the reporting country. The importer's value is generally more reliable than the exporter's due to the taxes to which imports are subject, so we used the values $\omega_{ij} = \omega^j_{i\to j} + \omega^i_{j\to i}$ whenever available, substituting importers' by exporters' values $\omega^i_{i\to j}$ or $\omega^j_{j\to i}$ otherwise. To check the consistency between DBI and DBII, trade volumes for 1997 in the two databases were compared, see Fig. 3.3. Indeed, the difference is smaller than 10^{-2} US Million \$ in the 96.1% of cases, which indicates a very good agreement. The minimal observed discrepancies are due to rounding criteria in some cases—we round to the second decimal place if the flow is larger than 1 US Million \$—, and to the absence of some importers' reports when DBI was constructed—, as the differences between the importers' and the exporters' reported values can be very large (for instance, $\omega^{SLV}_{SLV\to ZIM} \approx 160000$ US \$ while $\omega^{ZIM}_{SLV\to ZIM} \approx 15$ US \$). Nonetheless, as stated above, these cases are very rare.

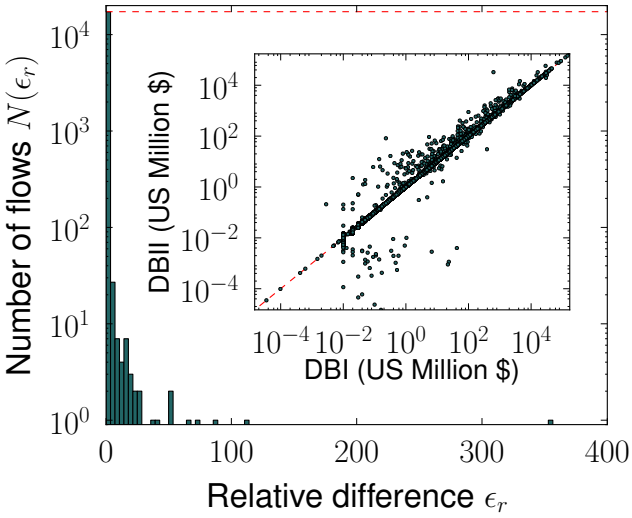


Figure 3.3: **Consistency between DBI and DBII in 1997**. Histogram of relative differences between trade volumes in the two databases. The red horizontal line indicates the total number of flows. **Inset**, Each point corresponds to the value of a given trade volume in both databases. Most points lay on the diagonal.

We also collected GDP data obtained from three different sources: The Maddison Project [38] from 1870 to 1949, Gleditsch's GDP data [78] from 1950 to 2011, and data from the World Bank [167] and the CIA World Factbook [47] for the years 2012 and 2013. As in the case of trade data, country codes need to be matched with those in the COW's state list. Moreover, since trade is reported in current US Million $, we have converted the real GDP values into current GDP values using the deflator calculated by comparing Maddison's data with data from [88].

Finally, data for formal Preferential Trade Agreements between countries was obtained from the World Trade Organization (WTO) [170].

3.2.1.2 Extracting backbones of significant trade channels

The high density of connections observed in the WTW reconstructions, which increases with the passage of time, obscures patterns and regularities within the networks and so prevents the inference of trade distances. At the same time, the networks display, especially after World War I (WWI), a strong heterogeneity in the global distributions of the number of trading partners, the total trade per country and bilateral flows. An increasing heterogeneity is also present at the local level in the distribution of flows between the neighbours of each single

country, which we quantify using the disparity measure $k_i Y_i(k_i)$ [151, 84, 85] for each country i,

$$k_i Y_i(k_i) = k_i \sum_j \left(\frac{\omega_{ij}}{s_i} \right)^2 , \tag{3.1}$$

where ω_{ij} is the total flow between countries i and j and $s_i = \sum_j \omega_{ij}$ is the strength (aggregated trade) of country i. If country i distributes its trade homogeneously between its trade partners $k_i Y_i(k_i) = 1$, whereas in the opposite case—if all its trade is concentrated on a single link—$k_i Y_i(k_i) = k_i$. As shown in Fig. 3.4 for a few selected years, the disparity of all countries in the WTW has been increasing. This indicates that, as countries gained trade partners, they only intensified a few of their trade connections.

These heterogeneities can be exploited to filter out, for each year, a sparse subnetwork representing the relevant structure that remains after eliminating the contingent interactions which overshadow the information contained in the system [153]. The disparity filter proceeds as follows. For each trade channel of a given country i, we compute the probability α_{ij} that the link takes a normalized weight larger than the corresponding observed value ω_{ij}/s_i according to a null model in which countries assign their aggregated trade among their partners randomly from a uniform distribution. By imposing a significance level α, we can determine the statistical significance of a given trade channel by comparing the obtained p-value α_{ij} to α. Therefore, if $\alpha_{ij} > \alpha$, the flow through that trade channel can be considered compatible with a random distribution (with the chosen significance level α) and is thus discarded. The statistically relevant channels are then those that satisfy

$$\alpha_{ij} = 1 - (k-1) \int_0^{\omega_{ij}/s_i} (1-x)^{k-2} dx < \alpha, \tag{3.2}$$

for at least one of the two countries i and j [153].

By applying this selection rule to all the links in the network we filter a subgraph containing, in general, less links and nodes depending on the value of the significance level α. In order to find an optimal value of α, it is convenient to plot the fraction of remaining nodes in the backbone N_{BB}/N versus the fraction of remaining links L_{BB}/L for different values of α (insets in Fig. 3.4). As the filter becomes more restrictive, the number of links decreases dramatically while keeping almost all nodes until a certain critical point α_c after which the number of nodes starts a steep decay. This behaviour is due to the fact that the disparity filter for $\alpha = \alpha_c$ is able to select the minimal set of significant links that preserve the global connectivity, referred to as the backbone. Geometrically, the value α_c defining the backbone can be identified with the value of the significance level

whose corresponding point in the $N_{BB}/N - L_{BB}/L$ plane maximizes the vertical distance to the diagonal. Notice that, on the diagonal, $N_{BB}/N = L_{BB}/L$, so the average degree remains constant ($\langle k\rangle_{BB} = 2L_{BB}/N_{BB} = 2L/N = \langle k\rangle$). Consequently, setting $\alpha = \alpha_c$, where α_c is chosen according to the geometric rule stated above, minimizes the average degree of the backbone $\langle k\rangle_{BB}$. This offers a consistent and systematic criterion to extract the backbones of the WTW at different years. The values of α_c for each year can be found in Table S4 in [69].

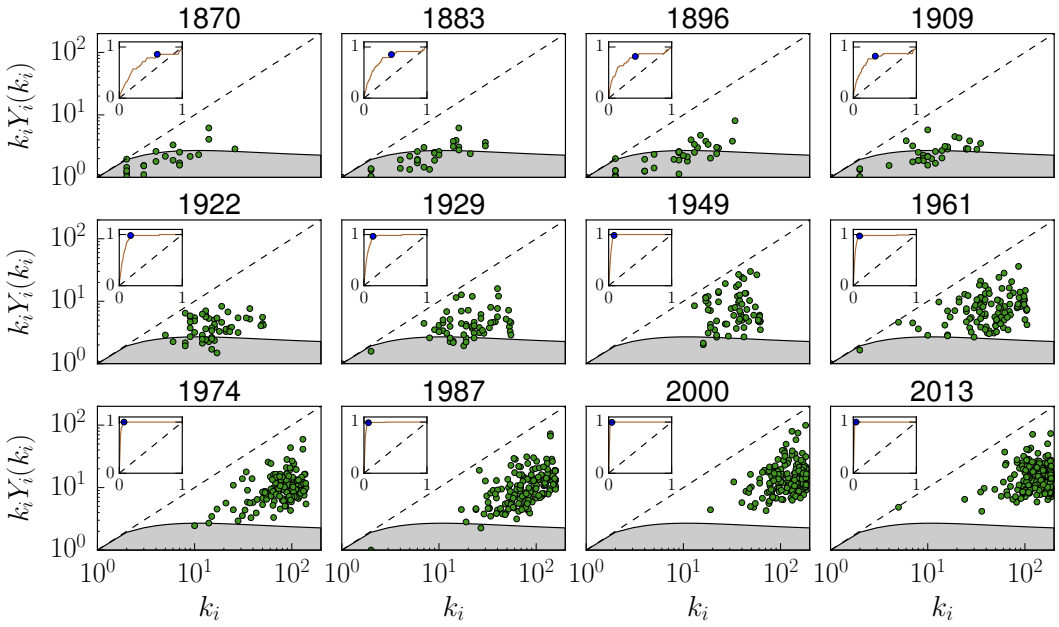


Figure 3.4: **Disparity in trade volumes for all countries of the WTW in different years**. The disparity is calculated according to Eq. (3.1) and is represented as a function of the degree k_i of countries. In each plot, each point corresponds to a different country and the gray area corresponds to the average given by the null model of the backbone filtering method (see text) plus 2 standard deviations. **Insets**, Curves representing the fraction of nodes N_{BB}/N vs. the fraction of links L_{BB}/L in the backbones parametrized by α. The blue circles indicate the values that correspond to the backbone with $\alpha = \alpha_c$ which maximizes the vertical distance to the diagonal (constrained by $N_{BB}/N \geq 0.8$ ($N_{BB}/N \geq 0.7$ in 1894) so that most countries remain in the system). Moreover, in order to avoid fluctuations in the topology between years, the 1975 backbone was selected with a higher link fraction than the given by the geometric rule.

Apart from preserving almost all countries, backbones show an average degree almost constant in time, Fig. 3.1**a** and capture most of the total trade, Fig. 3.1**b**. Backbones expose pivotal features of the international trade system

as a complex network: the small-world property, heterogeneous degree distributions, degree-degree correlations, and clustering, Fig. 3.2**b-e**. Interestingly, the correlation between the degrees of countries within the backbones and their GDPs is always extremely high, Fig. 3.5**a**, meaning that these backbones of world trade networks reveal economic size—in terms of GDP—as an underlying variable. The value is well above the correlation observed in the original networks, so that the backbones stand as a more meaningful network representation of the international trade system from an economic point of view.

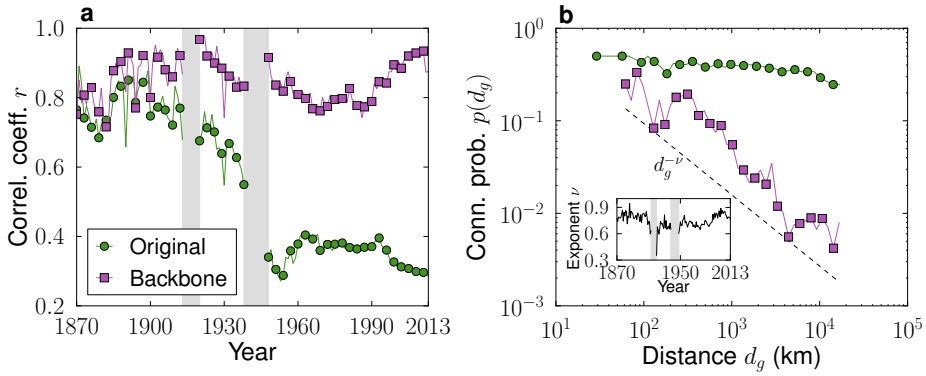


Figure 3.5: **Comparison of different statistics in the WTW and the corresponding backbone. a**, Time evolution of the Pearson correlation coefficient between the number of trade channels of a country and its GDP. **b**, Connection probability for trade channels as a function of geographic distance in 2013. **Inset**, Time evolution of the characteristic power-law exponent of that probability in the WTW backbone.

Backbones differ from the unfiltered networks in the relation between geographic distance and trade connections. Overall, the geographic length of connections relative to the geographic distance between all the countries (located at the coordinates of their capitals) shows a sustained increase in the unfiltered network, see Fig. 3.1**c**. This means that trade connections become less dependent on distance over time. However, the significant trade channels within the backbone are comparatively shorter over the whole period, which means that, despite globalization, countries are more likely to choose their more significant trade partners from among their closest neighbours. The connection probability is almost independent of distance in the unfiltered networks, but it shows a clear power-law decay with a stable exponent in the case of the backbones, Fig. 3.5**b**, showing that geographic distance plays an important role for the most significant trade channels. Strikingly, the filter that produces the backbones is blind to geography.

3.3 Mapping the World Trade Web

3.3.1 A gravity model for trade channels

WTW backbones are suitable to be mapped onto an underlying trade space through the $\mathbb{S}^1$ model. The likelihood for two countries to become connected is then represented in terms of a gravity law depending on economic size and effective trade distances, as in Eq. (2.4) in Chapter 2. In this chapter, however, we use a slightly different definition of the $\mathbb{S}^1$ model in which similarity space is not proportional to the system size, but has unitary radius. The reason for this choice is that we are interested in comparing the trade spaces of networks with different number of countries. The connection probability between two countries i and j separated by an angular distance $d_{a,ij} = \min\left(\left|\Delta\theta_{a,ij}\right|, 2\pi - \left|\Delta\theta_{a,ij}\right|\right)$ then reads

$$p_{ij} = \frac{1}{1+\left(\frac{d_{a,ij}}{\mu\kappa_i\kappa_j}\right)^{\beta}}. \tag{3.3}$$

In this particular context, hidden degrees κ_i can be interpreted as economic sizes [152]. These economic sizes are strongly correlated with degrees in the backbone and, therefore, with the corresponding GDPs, Fig. 3.5**a**. Distances $d_{a,ij}$ incorporate the different factors that shape the complex architecture of the international trade system—not only geography—implicitly encoded in its connectivity pattern. The parameters μ and β govern the average degree and the clustering, respectively. As introduced in Chapter 2, β is an elasticity measure with respect to trade distances; it calibrates the coupling between the topology and the underlying metric space. Hence, the proposed connection probability resembles Newton's law of gravitation and, therefore, the classical gravity model predicting the volume of bilateral trade flows. Notice, however, that p_{ij} is not used to predict the volume of trade but the existence of a significant trade channel; information which has to be provided *a priori* in classical gravity models of trade, since they reproduce the existence of a link very badly [59].

The isomorphism with the $\mathbb{H}^2$ model needs to be modified a little to account for the new gauge of non-unitary angular density. However, notice that this modification is equivalent to a rescaling of μ in the familiar $\mathbb{S}^1$ model connection probability, Eq. (2.4), so the mapping of variables κ_i to radial coordinates r_i, Eqs. (2.6) and (2.7), becomes

$$r_i = R - 2\ln\left[\frac{\kappa_i}{\kappa_0}\right] \quad \text{with} \quad R = 2\ln\left[\frac{2}{\mu\kappa_0^2}\right], \tag{3.4}$$

where κ_0 is the economic size of the smallest economy and R is the radius of the hyperbolic disk onto which nodes are mapped. With this new definition of R, all

the results concerning the properties of the isomorphism presented in Chapter 2 are valid in this case as well. In particular, the connection probability, Eq. (3.3) can then be rewritten as

$$p_{ij} = \frac{1}{1 + e^{\frac{\beta}{2}(x_{ij} - R)}} \tag{3.5}$$

where $x_{ij} = r_i + r_j + 2\ln\frac{d_{a,ij}}{2} \approx d_{h,ij}$ approximates the hyperbolic distance $d_{h,ij}$ between two points at radial coordinates r_i and r_j and angular separation $d_{a,ij}$. The exact hyperbolic distance can be computed with the hyperbolic law of cosines, Eq. (2.10), which we include here for completeness,

$$\cosh d_{h,ij} = \cosh r_i \cosh r_j - \sinh r_i \sinh r_j \cos d_{a,ij}. \tag{3.6}$$

The hyperbolic representation condensates in a pure geometric framework the properties of the entire system—economic size and effective distance in the gravity model Eq. (3.3)—; that is, it allows to draw genuine maps of the trade system where different parts can be compared on an equal footing. More specifically, due to the inverse relation between economic size κ_i and hyperbolic radial coordinate r_i, large economies with high degrees are located close to the center of the disk, whereas small economies are placed near its boundary; the distance between nodes with the same angular separation increases exponentially with their hyperbolic radial coordinates.

World Trade Maps are constructed by embedding WTW backbones in hyperbolic space. As discussed in the previous chapters, the embedding assigns coordinates in the underlying space to each country in a backbone such that trade and effective distances can be defined for every pair; this is achieved by maximizing the likelihood for the backbone to be reproduced by the model. Inferred angular distances represent a measure of trade likelihood, except for the (economic) sizes of the countries. This means that two small countries need to be close in terms of angular distance to increase their chance of becoming connected. The inferred hyperbolic distance, however, incorporates size effects; that is, two countries can be close in hyperbolic space just because of their size. In particular, large economies are in general closer to the rest of the world than smaller economies.

The embeddings are obtained using the algorithm presented in Section 2.4.2 starting the sequence with the backbone of year 2013, the largest network, and then embedding the rest of them in inverse chronological order. It is also important to point out that the likelihood function has a very complex landscape with many degeneracies, that is, configurations with similar likelihood. Usually, the difference between these configurations is the angular inversion of a whole community. Such a change mildly affects a small fraction of the pairs of distances $d_{h,ij}$ only and, therefore, the likelihood takes a similar value. In order to obtain statistically reliable distances between countries, we average them

(both hyperbolic and angular) and connection probabilities, as well as their fluctuations, over 100 embeddings starting with completely random positions, see Fig. 3.6.

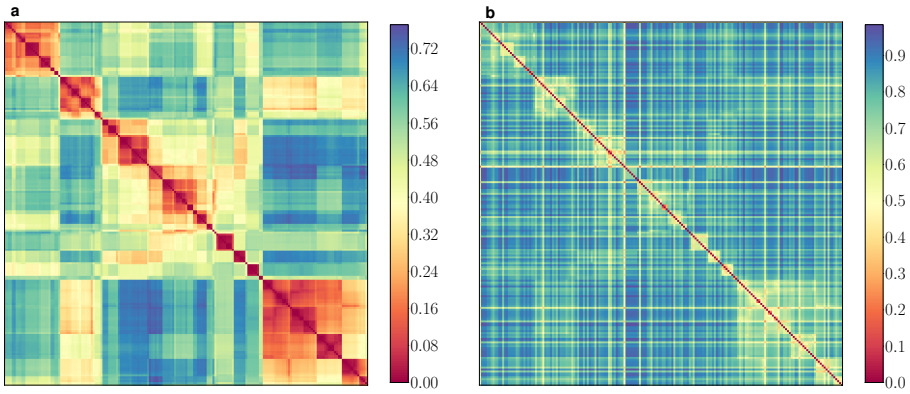


Figure 3.6: **Distance matrices between countries in the WTW backbone in 2013. a**, Average normalized angular distances $\langle \tilde{d}_{a,ij} \rangle$. Nodes are ordered such that nearest neighbours are consecutive. The dark squares in the matrix reveal the community structure of the backbone. **b**, Average normalized hyperbolic distances $\langle \tilde{d}_{h,ij} \rangle$. The ordering of the countries is the same in both matrices.

3.3.2 Hyperbolic maps of WTW backbones

The application of the embedding method to the backbones of annual WTWs from 1870 to 2013 produces 129 World Trade Maps that conform the World Trade Atlas 1870-2013 [2] (WTA 1870-2013). Single embeddings for each year are useful for visualization purposes (see the online interactive video tool [69]), Fig. 3.7, while averaged distances, Fig. 3.6, are more suitable for computations.

3.3.2.1 Pairwise quality tests

Fig. 3.8**a-d** demonstrates the power of the gravity model of trade channels to reproduce the topological properties of WTW backbones. This is due, in particular, to the excellent agreement between the empirical and the theoretical probability of connection Eq. (3.5) as a function of the effective distance between countries rescaled by economic size, Fig. 3.8**e-h**. To measure the empirical connection probability, the effective distances $\frac{d_{a,ij}}{\mu\kappa_i\kappa_j}$ are binned into M subintervals and, for every bin, the fraction of connected pairs out of all pairs whose effective distances fall into the bin is counted. The expected degrees $\langle k_i \rangle$ of countries, defined as $\langle k_i \rangle = \sum_{j \neq i} p_{ij}$ with p_{ij} given by Eq. (3.5) and the distance calculated

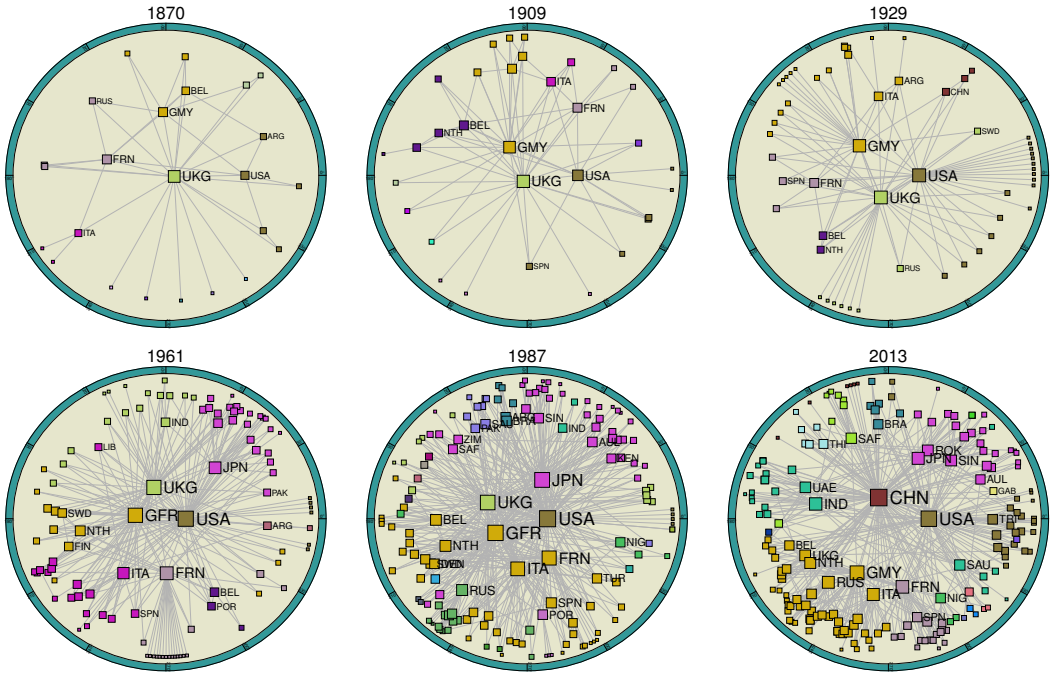


Figure 3.7: **World Trade Maps for different years.** Representation on the hyperbolic plane. The radial and angular coordinates of countries determine the trade distances between them trough Eq. (3.6). Symbol sizes are proportional to the logarithm of the degrees of the countries and colors represent communities revealed by the CGM presented in Section 3.4.2. See the online video tool [69] for the country associated with each acronym.

using the coordinates of the embedding and Eq. (3.6), match almost perfectly the observed values k_i measured in the backbones, see Fig. 3.8**i-l**, in accordance with MLE for expected degrees [35].

A further demonstration of the quality of the embeddings is given by the temporal evolution of the area under the receiver operating characteristic (ROC) curve or AUC statistics measuring the ability of the model to predict real trade channels. Once a network is embedded, Eq. (3.5) assigns a connection probability p_{ij} to each pair of nodes, independently of whether they are connected or not in the backbone. The ROC curve is produced by computing the *true positive rate* (TPR)—the fraction of existing links predicted by the model—as a function of the *false positive rate* (FPR)—the fraction of links that do not exist in the network but have been predicted to exist by the model—parametrized by the thresholds defined by all possible links. Notice that a perfect prediction would correspond to TPR=1 and FPR=0. The area under the ROC curve (AUC) provides a scalar summarizing the result. The best possible result corresponds to AUC=1; its value is

well above 0.975 in all embeddings of the WTA 1870-2013. Fig. 3.8**m** displays the time evolution of the AUC statistic, always almost 1 and, in the inset, the ROC curve for the year 2013.

Finally, a key observation is that the correlation between embedding distances and geographic distances, as shown in Fig. 3.8**n**, is significant over the whole period, but far from one; meaning that trade distances encode more than purely geographical information.

3.3.2.2 Global quality tests

The tests presented so far to assess the quality of the embeddings only take into account pairwise information. We complement the validation by measuring the efficiency of navigation in the embeddings. As we discussed in Chapter 1, the geometrical information provided by the embedding into hyperbolic space can be used to navigate the network efficiently by following the very simple rule of greedy forwarding: to go from a source node i to a target node j, hop to the source's neighbour that is closest (according to Eq. (3.6)) to destination j, and iterate [35]. Following this procedure, networks which are well described by the embedding model sustain a high amount of successful routes in a number of hops not much larger than the topological shortest path length, i. e. the minimum number of hops actually required to go from i to j. This is due to the fact that shortest paths follow closely the geodesics in the underlying space.

We measure the success rate—the fraction of successful paths via greedy forwarding—and the average stretch—the average over all successful routings of the quotient between the number of hops required by greedy forwarding and the topological shortest path length. Notice that, for every pair of countries, both directions must be checked, since the process is not necessarily symmetric. We also compare the results with the ones measured on the same networks with the same radial coordinates but with randomized angular coordinates. The results are displayed in Fig. 3.8**o-p**. As the number of countries grows, the success rate for the randomized embeddings decreases, reaching values as low as 0.5, whereas it takes an almost constant value of 1 along the whole period for the embeddings of the backbones. The stretch increases slightly for the last years in both cases but remains very close to one. However, notice that it is always closer to 1 for the real embedding even though it is computed over a much larger set of pairs; indeed, the stretch is defined for successful greedy paths only (about 50 % in the randomized case in the last years).

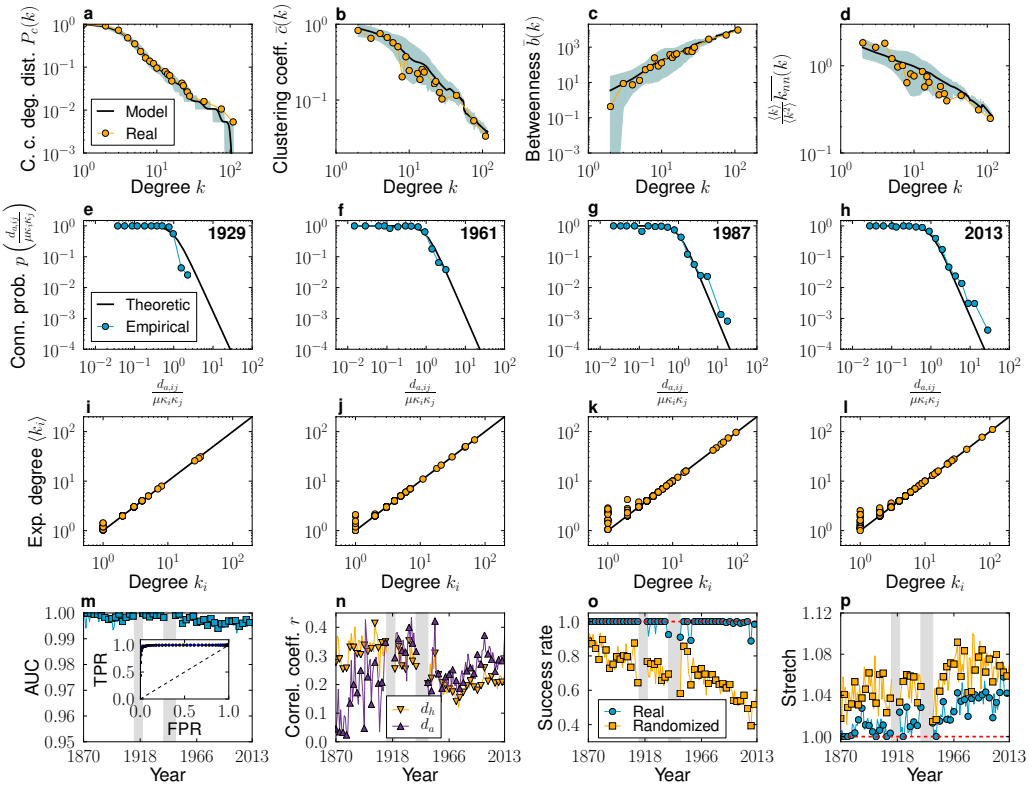


Figure 3.8: **Quality of the embeddings. a-d**, Comparison of the topology of the WTW backbone in 2013 with 1000 synthetic networks generated by the gravity model of trade channels using the coordinates of the corresponding embedding. **a**, Complementary cumulative degree distribution. **b**, Degree-dependent clustering coefficient. **c**, Country betweenness centrality. **d**, Average normalized nearest-neighbor degree. The thick black line gives the average over realizations and the geen areas indicate the 99% confidence interval. **e-h**, Empirical connection probability as a function of effective distance rescaled by economic size and measured using the average distances in the corresponding year. The empirical probability is computed from Eq. (3.3). **i-l**, Expected degree $\langle k_i \rangle$ of node i versus its real degree k_i. **m**, Temporal evolution of the area under the ROC curve, or AUC statistic. **Inset**, ROC curve of the model in 2013. **n**, Historical evolution of the Pearson correlation coefficient between hyperbolic and angular distances, d_h and d_a, and geographical distance d_g. **o**, Evolution of the success rate of the greedy forwarding process. **p**, Average stretch of successful greedy routing paths for every year.

3.4 Trade since the 19th century

An interactive tool displaying The World Trade Atlas 1870-2013 can be accessed at [69]. It allows us to visualize the evolution of the international trade system over fourteen decades. During the 19th century, the atlas shows a sparsely populated trade space with the hegemonic UK at its core; and it reflects the rise of Germany, France and the USA as they move towards the center prior to WWI. During the interwar period, the atlas presents a central triangle formed by the UK, the USA and Germany; with the USA then progressively becoming the new hegemonic economy during the second part of the 20th century. In the 1960s, decolonization introduces many new countries and at the same time the upper intermediate region becomes more populated by actors such as France, Italy, Japan, and the Netherlands; together with, in less prominent positions, India, Russia, Spain and Belgium, among others. In the 1990s, the European Community starts its process of construction, with Germany, Italy and France coming closer; while the Soviet Bloc remains very close to Russia. During the last few years, China has moved towards a more central position, as a new pretender to superpower status; while the USA, the European Community and Japan have moved to less central positions following the decline of their relative dominance.

3.4.1 Hierarchies

The evolution of the international trade system at the large scale has been shaped by three simultaneous forces: globalization—mentioned above—, localization—discussed in the next subsection—, and hierarchization. The impact of the first two has already been discussed from many perspectives [96, 98, 16, 68, 46, 67, 160, 74]. Conversely, the role and implications of hierarchization at the system level have not been recognized before. However, its effects in the WTMs are keenly discernible. The first sign is the radial stratification of economies according to their size, with small economies—low-degree countries—always located at the periphery of the disk and large economies—with high degrees—settling at its center. A rough measure of the hierarchical position of a country is then given by its radial coordinate, and so by its degree, as compared to the radius, R, of the trade space, which measures the hyperbolicity of the whole system or, equivalently, its departure from flatness, Fig. 3.9**a**. R grows until reaching a stationary value at the end of the interwar period, around which it fluctuates thereafter. Interestingly, the opposite behaviour is observed for the elasticity parameter, β, shown in Fig. 3.9**c**, informing about the coupling between the topology of the network and the underlying metric space. Larger values of R imply that differences in trade distances are more important for peripheral nodes; that is, they connect to each other less frequently.

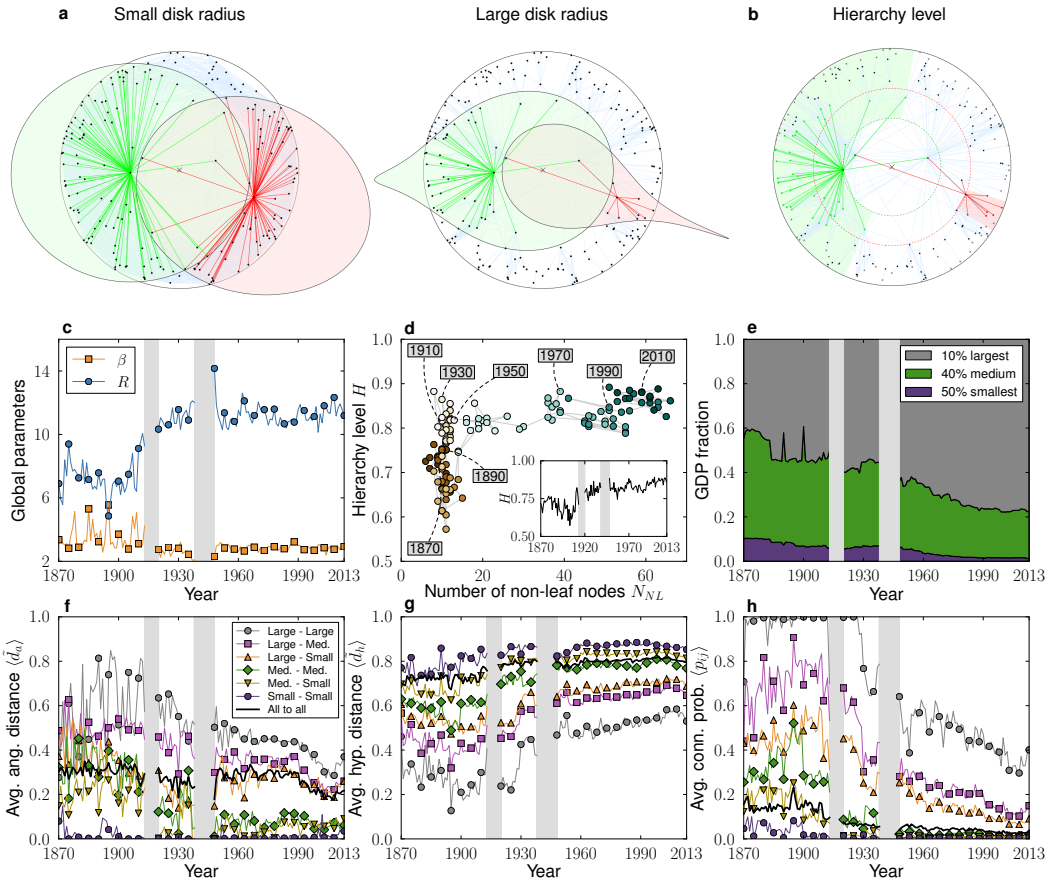


Figure 3.9: **Hierarchy in World Trade Maps.** **a**, Two hyperbolic disks with different radii (rescaled). Within each disk, we draw two circles of the same radius but centered at different locations. For small radius, the two inner circles appear to be similar to Euclidean disks; whereas for large radius, the inner circles are strongly distorted. **b**, Sketch illustrating the definition of the hierarchy level, H, of two different nodes. Leaf nodes appear in gray. **c**, Evolution of the radius, R, of the hyperbolic disk and of the elasticity parameter, β. **d**, Evolution of the hierarchy level, H, as a function of the number of non-leaf countries and of time. **Inset**, Time evolution of H. Note that the number of non-leaf countries has grown (with some fluctuations) over time, so that the increase of N_{NL} along the x-axis is roughly chronological. **e**, Evolution of the fraction of total GDP accumulated by the top 10%, 40%, and 50% of countries by economic size. **f–g–h**, Evolution of: the average normalized angular distance (only computed over connected pairs of countries); the average normalized hyperbolic distance; and the average connection probability between countries within the same group and in the different groups defined in **e**.

A more refined measure is given by the level of hierarchy of the network as a scalar, H, based on angular distances. We define a non-leaf country i as that having at least one trading partner j whose economy is smaller, i.e, with a radial coordinate $r_j > r_i$. For each non-leaf country i with radius r_i, we measure the average angular separation $\overline{d_{a,i}}$ with its trading partners for which $r > r_i$, and define its hierarchical level as $h_i = 1 - 2\overline{d_{a,i}}/\pi$, Fig. 3.9**b**. If the neighbours have exactly the same angular coordinates as i, $h_i = 1$; whereas $h_i = 0$ for random angular positions. The global hierarchy level, H, is obtained by averaging h_i over all non-leaf countries. The evolution of H as a function of the number of non-leaf countries, Fig. 3.9**d**, is revealing. Even if the number of non-leaf countries has increased noticeably in the last decades, H has not decreased but has increased. This indicates an expansion of the depth of the hierarchy, from a quasi star-like organization before WWI to a deeper hierarchical structure after WWII, with more countries at intermediate layers acting as local economic hubs. In the long run, H increases from below 0.7 in the 19th century to very close to 0.9 in 2013, see inset of Fig. 3.9**d**. This represents a substantial increase and situates the system at extremely high levels of hierarchy.

The economic significance of this effect can be further explored by analyzing the evolution of trade distances between countries classified according to their economic size. We rank countries in decreasing order of GDP and divide them into three groups: the top 10%, the middle 40%, and the bottom 50%, Fig. 3.9**e**. Figures 3.9**f**, **g** and **h** show the evolution of: the average angular distance (between connected pairs); the average hyperbolic distance; and the average connection probability, for all pairs of countries within the same group and in the different groups. Hyperbolic and angular distances are normalized by the diameter of the space, $2R$ and π, respectively, so that different years are comparable. There is a sharp and persistent stratification according to economic group, denoting a positive correlation between trade distance, or connection probability, and economic status. The average angular distance strongly correlates with economic size. Interestingly, we observe the opposite behaviour for the average hyperbolic distance: it is anti-correlated with economic size. This reversal reflects the different interpretations of the two trade distance measures: the latter governs the formation of trade channels, and the former is an indicator of clusterization. The hyperbolic distance between economic groups has been steadily increasing since the 1960's; and this increase has been more pronounced for the large economies. In contrast, the angular distance between groups, fluctuating until the end of WWII, has remained very low or decreased over time. This denotes a trend towards concentration of countries in specific regions of the trade space.

Placed in historical perspective, these results indicate that small countries at the bottom of the hierarchy are far from the rest of the world in terms of trade

distance. They therefore encounter greater difficulties establishing trade channels with other countries, except for the largest economies at the top of the hierarchy, which have more chances of becoming connected worldwide. At the same time, the increase in hyperbolic distance may be a consequence of the increase in hierarchy and of market competition effects; while the decrease in angular distance indicates the formation of better-defined trading blocs that form communities in the trade space.

3.4.2 Detecting communities in WTMs: the CGM method

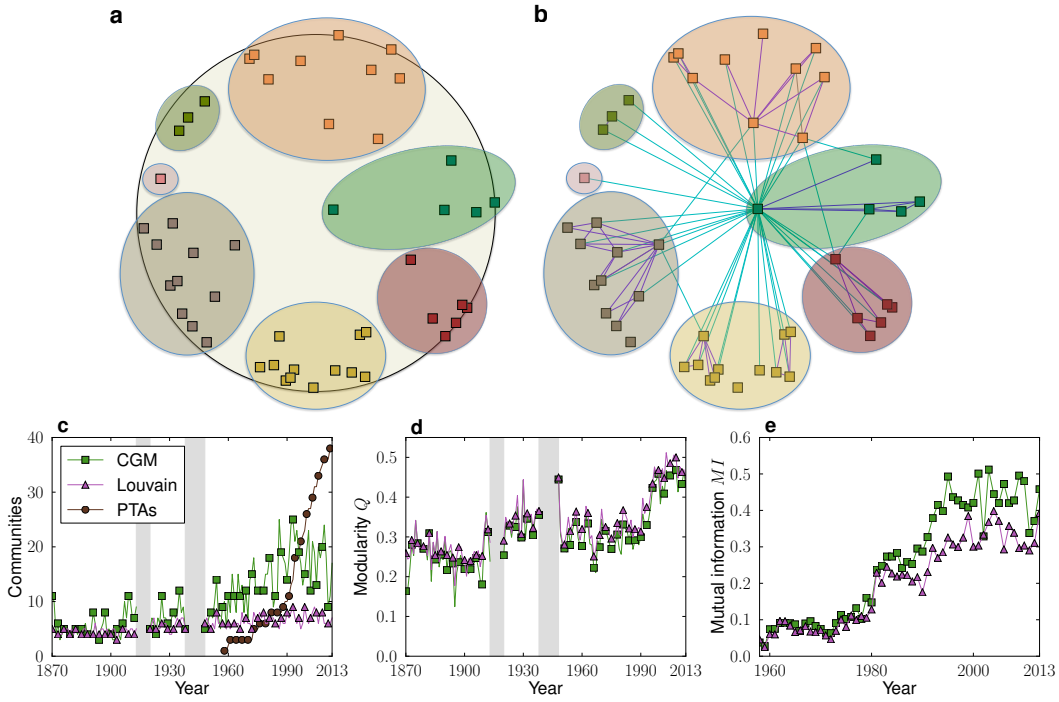


Figure 3.10: **CGM space-based communities versus Preferential Trade Agreements.** **a**, Sketch illustrating the CGM method. Nodes in the hyperbolic plane are partitioned into different communities separated by a large enough angular gaps, hence depending on the geometric information only. **b**, Once the partition is defined, the modularity Q is computed by comparing the number of links within the communities (purple links) to the number of links between nodes in different communities (green links). **c**, Evolution of the number of communities detected by the Louvain method and the CGM, along with the evolution of the number of PTAs. **d**, Evolution of the modularity, Q, for the two community detection methods. **e**, Mutual information between the CGM or Louvain communities and PTAs.

Hub countries near the center of the hyperbolic disk are close to many other countries in the system and therefore have a large number of connections, whereas countries at the periphery are close to other small countries only if separated by a small angular distance. Therefore, to partition the network into communities with similar countries, we must take into account their angular coordinates. Indeed, the $\mathbb{S}^1$ model assumes a uniform angular density of nodes in the underlying circle, so the clusterization of countries around certain angular coordinates reveals regions in which they tend to connect with each other profusely. The Critical Gap Method (CGM) relies on this idea to detect communities by finding angular regions that are densely populated. To that end, it maximizes modularity Q [116], the standard measure for the quality of the partition in community detection algorithms applied to complex networks. This quantity, which is bounded $Q \in [-1,1]$, compares the fraction of links inside communities with the expected fraction in the Configuration Model with the observed degrees, and is equal to 0 if nodes are randomly assigned into communities. It is defined as follows. Suppose that a network has been partitioned into communities, such that every node i belongs to a community c_i. Then, the modularity of the partition is given by

$$Q = \frac{1}{2L}\sum_{i,j}\left(a_{ij} - \frac{k_i k_j}{2L}\right)\delta\left(c_i, c_j\right), \tag{3.7}$$

where L is the total number of links in the network and $\delta\left(c_i, c_j\right) = 1$ if $c_i = c_j$ and $\delta\left(c_i, c_j\right) = 0$ otherwise. To define a partition, a critical angular gap separating countries in different communities needs to be specified. Once a partition of countries in the international trade system is defined on the basis of their localization in a WTM and the critical gap, its modularity is evaluated according to the topology of the corresponding backbone and the partition with the highest modularity is chosen. In detail, the CGM method for the detection of spatially based communities in WTMs reads:

1. Set the value of $\tilde{Q} = -1$. This variable will encode the highest value of Q.
2. Increase the value of the critical gap g_c sequentially from 0 to 1. For every value:
    i. Connect all pairs of nodes (i, j) whose normalized average angular distance $\tilde{d}_{a,ij} \leq g_c$.
    ii. Find all the connected components in the resulting graph.
    iii. Assign all nodes in the same connected component to the same community.
    iv. Compute the modularity Q of the partition. If $Q > \tilde{Q}$ set $Q \to \tilde{Q}$ and keep the partition.

At the end of the process, the partition with the highest modularity is kept. Notice that this method is equivalent to considering all gaps larger than g_c between consecutive nodes in the circle as community boundaries; however, this algorithm works also using distances instead of coordinates. This is useful because averaged distances are more reliable than single embeddings. Figures 3.10**a-b** illustrate the process of community detection using the CGM method. The evolution of the number of CGM communities is shown in Fig. 3.10**c**, as compared to those obtained by applying the Louvain algorithm [31], an efficient method for greedy optimization of modularity, on the topology of the backbones. In terms of modularity, Fig. 3.10**d**, the two methods coincide to an extremely high degree. However, the number of communities discovered by CGM is typically twice that of the Louvain method, since Louvain modules may integrate smaller CGM communities. Both modularity and the number of communities increase over time, reaching a maximum at the beginning of the last economic crisis around 2007, with a minor downturn afterwards.

3.4.3 WTM communities versus PTAs

This tendency towards localism in trade space seems consistent with the proliferation, since the late 1980's, of PTAs as formal trading blocs [97, 65], Fig. 3.10**c**. We use the WTO's Regional Trade Agreements Information System (RTA-IS) [170] to list all plurilateral PTAs in force in 2013 and compare them with natural CGM trade communities. To measure their similarity, we use the normalized mutual information [54] (MI). The normalized mutual information of two partitions X and Y is defined as

$$MI = \frac{\sum_{y\in Y}\sum_{x\in X} p(x,y)\log\left(\frac{p(x,y)}{p(x)p(y)}\right)}{\max\{H(X),H(Y)\}}, \tag{3.8}$$

where

$$H(X) = -\sum_{x\in X} p(x)\log p(x) \tag{3.9}$$

and

$$H(Y) = -\sum_{y\in Y} p(y)\log p(y). \tag{3.10}$$

In the above formulas, $p(x,y)$ is the probability for a randomly chosen node to belong to community x in partition X and to community y in partition Y. The marginal probability $p(x)$ ($p(y)$) is the probability for a randomly chosen node to belong to community x (y) in X (Y). Thus, if two partitions X and Y are equal, $p(x,y) = p(x) = p(y)$ and $MI = 1$. On the other hand, if the two partitions are uncorrelated, $p(x,y) = p(x)p(y)$, the logarithm in Eq. (3.8) is null and, hence,

$MI = 0$. Over the entire period, natural communities are noticeably more congruent with PTAs than those identified by the Louvain method, Fig. 3.10**e**. However, the overlap between PTAs and natural communities is not perfect. This poses a question regarding the progression of barriers to trade within each PTA, which can be assessed from the evolution of the average normalized angular and hyperbolic distances—$\langle \tilde{d}_a \rangle$ and $\langle \tilde{d}_h \rangle$, respectively—between its members. We focus on PTAs with at least ten years of history—31 trading blocs in 2013—for which we find different patterns of evolution, see Fig. 3.11.

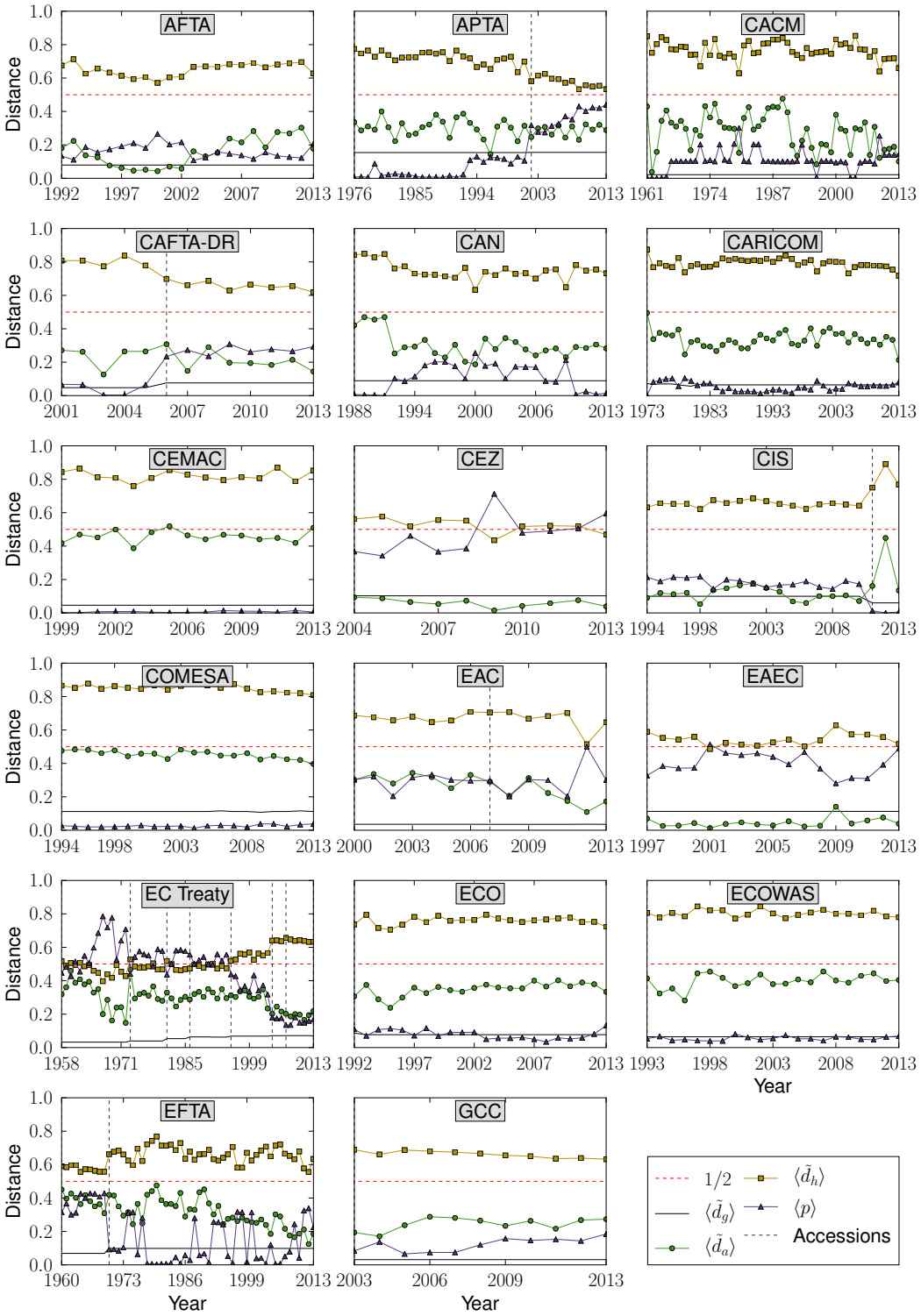
Both average normalized angular and hyperbolic distances remain stable over time in about half (17) of the PTAs, with $\langle \tilde{d}_a \rangle$ typically close but above the geographic value—North American Free Trade Agreement NAFTA, Central American Common Market CACM, Caribbean Community and Common Market CARICOM, Pacific Island Countries Trade Agreement PICTA, Economic Cooperation Organization ECO, and PTAs in Africa, like SADC, WAEMU, CEMAC, ECOWAS, and in Latin America LAIA. In some blocs, $\langle \tilde{d}_a \rangle$ and $\langle \tilde{d}_g \rangle$ are extremely congruent, specifically in PTAs with a worldwide composition—Protocol on Trade Negotiations PTN and Global System of Trade Preferences among Developing Countries GSTP—; or, at the other extreme, with a strong geographical orientation—Southern Common Market MERCOSUR and Commonwealth of Independent States CIS. Strikingly, the value of $\langle \tilde{d}_a \rangle$ is found to be below the geographical average in PTAs interconnecting Russia and the republics of the former Soviet Union—Russian Federation - Belarus - Kazakhstan, Eurasian Economic Community EAEC, Common Economic Zone CEZ—, which denotes communities with reduced trade barriers, even below levels expected given their geographical dispersion.

Distances in the rest of the PTAs show mixed progressions. An interesting case is the Asia-Pacific Trade Agreement (APTA), with $\langle \tilde{d}_h \rangle$ steadily decreasing since its inception, even after the accession of China in 2002; whereas $\langle \tilde{d}_a \rangle$ has remained fairly stable. This implies a significant increase in the internal connectivity of the PTA, which may be related to the increasing economic size of their members. In other PTAs both the average normalized angular and hyperbolic distances show a very mild decrease—Pan-Arab Free Trade Area PAFTA, Dominican Republic - Central America - United States Free Trade Agreement CAFTA-DR, Andean Community CAN, Southern African Customs Union SACU, Common Market for Eastern and Southern Africa COMESA, East African Community EAC. Finally, both distances present a slight but clear increase in some PTAs, denoting a trend towards trade diversion among their members, as in the South Pacific Regional Trade and Economic Cooperation Agreement (SPARTECA), mainly formed of small islands in the Pacific Ocean.

The European Community (EC) Treaty is the largest and most complex of all the PTAs, accounting for 23.6% of world GDP in 2013. Its evolution has

moved through three different stages. The first lasted from its inception until 1972, just before the accession of Denmark, Ireland, and the UK. During this period, both angular and hyperbolic distances reduced significantly and the internal connectivity increased, which indeed indicates an important reduction of barriers to trade. The aforementioned accessions in 1973 reversed the trend, which indicates that the new members were not natural partners of the treaty before they entered the agreement. During the second stage, 1973–2003, many medium/large economies joined, but internal distances remained stable since these new countries were, *de facto*, natural trade partners of the treaty. Between 2003 and 2004, another transition took place with the accession of a large number of small economies, mainly in Eastern Europe. During the transition, $\langle \tilde{d}_a \rangle$ decreased significantly at the same time as $\langle \tilde{d}_h \rangle$ increased, and so the internal connectivity decreased. Before joining the EC, these newly adjoined countries already had members of the EC as their main trading partners. After joining the EC, these new countries kept the same main trading partners but, due to their small size, did not increase trade with them. So, the effect has been augmented internal barriers to trade in the last decade, as the increased localization due to the addition of small economies could not be compensated for due to the heterogeneity in economic sizes of the members.

In summary, the clear movement towards localism as a tendency to concentrate within natural trade communities only finds a moderate overlap with formal PTAs. Indeed, PTAs have not necessarily reduced barriers to trade between their members, as measured by hyperbolic and angular distances.


AFTA
APTA
CACM
CAFTA-DR
CAN
CARICOM
CEMAC
CEZ
CIS
COMESA
EAC
EAEC
EC Treaty
ECO
ECOWAS
EFTA
GCC
Distance
Year
1/2
$\langle \tilde{d}_h \rangle$
$\langle \tilde{d}_g \rangle$
$\langle p \rangle$
$\langle \tilde{d}_a \rangle$
Accessions

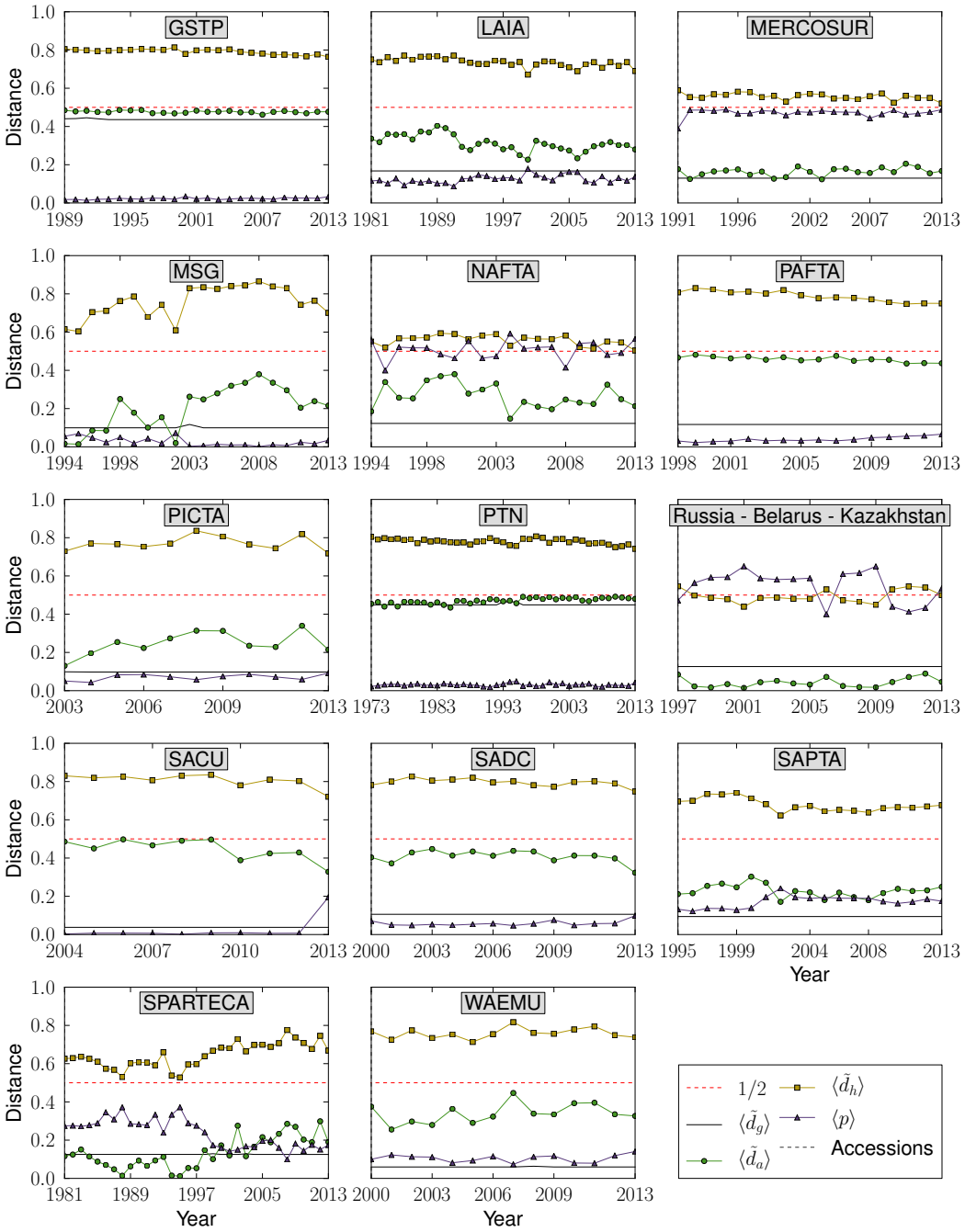


Figure 3.11: **Evolution of effective and trade distances between countries in every PTA.** Evolution of the average normalized distances (geographical $\tilde{d}_g$, angular $\tilde{d}_a$ and hyperbolic $\tilde{d}_h$) and average connection probabilities. The vertical dashed lines represent new accessions into the PTA. The red dashed lines correspond to the $1/2$ value. Since distances are normalized with respect to the maximum possible distance in each case, it represents the average separation for both $\tilde{d}_g$ and $\tilde{d}_a$ for a uniform random distribution of points and the radius R in the hyperbolic case.

3.5 Discussion

In terms of trade, two waves of globalization are identified in recent history corresponding to processes of decolonization and breaking of technical barriers inducing downturns in costs and time expenditures. The first wave is roughly identified from 1870 until the beginning of World War I and was related to lowered costs for transportation of materials and goods triggered by the Industrial Revolution, with steam power encouraging the expansion of railroad networks and oceanic routes and the telegraph connecting the two sides of the Atlantic. The second wave, from 1960 to the present, is intimately related to the ease of the exchange of information and ideas facilitated by the Information Technology Revolution, which is causing communication costs to drop dramatically at the same time that information management capabilities are exploding.

As a result of these globalization processes, the large-scale organization of the world economies exhibits nowadays a high level of local heterogeneity and of global interdependency at the same time. In this scenario, the relevance of trade goes beyond a mere exchange of goods and services. On the one hand, commercial trade flows are indeed highly correlated with other types of cross-country economic interactions—flow of services, financial assets, workers, etc.—and so stand as a good indicator for more general economic relations. On the other, and leaving aside technological, cultural and other non-economic social aspects that interplay with trade, feedback mechanisms operate between international trade and other economic variables such as production, investment, debt, or currency prices. Trade plays a central role as one of the most important interaction channels between states. It can act, for instance, as a substrate for the transmission of economic policies, cycles, and shocks like the 1997 Asiatic crisis, which shows how economic perturbations originated in a single country can somehow propagate globally in the world. In a broad sense, this implies that the collapse of the barriers in human interactions allows for incidences in one part of the system to affect all the rest.

Therefore, it seems natural to analyze the international trade system from a global perspective taking into account every country and its trade relationships regardless of its size or wealth. This is in contrast to traditional analyses of international commerce that have been based on bilateral interactions. However, the entangled connectivity characterizing international trade points out to a complex system, whose properties depend on its global organization. The complex features of the international trade system are well reproduced at the level of trade channels by a network gravity model based on a hidden metric space of trade distances as aggregated measures that bring together and integrate the different dimensions—economic, administrative and political, geographical, and cultural—that shape world trade networks. Trade distances determine the like-

lihood of connections such that two countries close in the natural geometrical embedding space of WTW networks, hyperbolic space, have more chance to interact. The gravity model plus the embedding method conform the framework used to construct the World Trade Atlas 1870-2013 [2, 69] that summarizes international trade history between these dates. Most importantly, as stated before by others in an interesting discussion of globalization [67, 160, 74]: the world is not flat. We claim, on scientific grounds, that it is hyperbolic.

In contrast to the widespread perception that globalization has led to a decrease in the importance of distance, we observe that countries preferentially select their significant trade partners from geographically close neighbours, in line with general statistics [75]. According to the World Trade Atlas 1870-2013, the role of trade distance has not decreased but increased over time over the last 14 decades driven by two main forces. First, we report an increased hierarchical organization related to a persistent stratification by economic size, so that not all global trade market competitors have equal opportunities. It was reported previously that elasticity with respect to distance of bilateral trade between high-income countries fell in the period 1962-1996 [41], with no trend for the group of low-income countries. The portrait that emerges from our maps is more dire. Differences in trade distances are becoming more important, particularly for small economies at the bottom of the hierarchy, which are moving away from the rest of the world. The small economies are encountering more difficulties in establishing trade channels, except for those with the largest economies at the top of the hierarchy, which have more chances of becoming connected worldwide.

Second, we observe a movement towards localism—already encountered before in the context of economic geography [96, 68]—as a tendency to concentrate within natural trade communities. Interestingly, this trend seems to have been reverted since 2009, maybe as a consequence of the fast rising of China as a new commercial power and due to the effects of the economic crisis. Despite the proliferation of PTAs as formal trading blocs, we only found a moderate overlap with natural communities. Indeed, PTAs have not necessarily reduced barriers to trade between their members, as measured by hyperbolic and angular distances. These results reveal PTAs as tools that may serve purposes other than trade in economics or politics, since their ambiguous consequences on the creation or steering of trade depend upon several other conditions [97, 142]. In our framework, we observe that the localization effect is entangled with that of hierarchization; that is, with the formation of intermediate hubs that dominate well-delimited angular regions as the number of layers in the hierarchy grows. Both effects, as two sides of the same coin, may have been exacerbated by trade liberalization policies with uneven effects among non-equals [43].

Our framework can be understood as a new generation of gravity models. Classical gravity models predict the volume of flows between elements but cannot explain the observed topology of the interactions among them, as shown in works for the World Trade Web [59]. Our contribution overcomes this limitation and offers a gravity model that can reproduce the existence of significant interactions. In perspective, the hidden metric space hypothesis and the maps of real complex systems that it enables will lead to a deeper understanding of the interplay between the structure and function of real networks, and will offer insights on the impact they have on the dynamical processes they support and on their own evolutionary dynamics. Our discussion has focused on revealing, through a single historical picture, globalization, hierarchization, and localization as the main forces shaping the trade space, which far from being flat is hyperbolic. The World Trade Atlas 1870-2013 can help to shed light on a number of other questions based on trade distances, for instance, those regarding the specific composition of natural communities in trade space, or it can facilitate a new approach to the analysis of the relation between trade and other economic factors.

CHAPTER 4

Geometric renormalization of complex networks

In the previous chapter, we used the models introduced in Chapter 2 to embed a real system, the network of international trade, to analyze its evolution from a geometric perspective. In this chapter, we show that the congruency between real systems and the $\mathbb{S}^1$ and $\mathbb{H}^2$ models can be exploited further, as it opens the door to a proper geometric definition of self-similarity and scale invariance, as well as to an unfolding of the different scales present in the structure of real networks. The results presented in this chapter correspond to Ref. [3].

4.1 The problem of scales in complex networks

The definition of self-similarity and scale-invariance [106, 159] in complex networks has been limited by the lack of a valid source of geometric length scale transformations. Previous efforts to study these symmetries are based on topology and include spectral coarse-graining [73], or box-covering procedures based on shortest path lengths between nodes [156, 80, 157, 91, 138, 143]. However, the collection of shortest paths is a poor source of length-based scaling factors in networks due to the small-world [166] or even ultrasmall-world [49] property, as a consequence of which the distribution of topological distances is very narrow. Other studies have faced the multiscale structure of network models in a somewhat more geometric way [119, 32], but their findings cannot be directly applied to real-world networks.

Naturally, the range of similarity distances—much broader than in the topological case—allows a reservoir of distance scales so that we can borrow concepts and techniques from the renormalization group in statistical physics [101, 168]. By recursive averaging over short-distance degrees of freedom, renormalization has successfully explained, for instance, the universality properties of critical behavior in phase transitions [169]. In the next section, we will introduce a geometric renormalization group for complex networks (RGN) acting on the similarity space uncovered by the embedding of networks onto the $\mathbb{S}^1$ model. The method, inspired by the block spin renormalization group devised by L. P. Kadanoff [101], coarse-grains nearby nodes into supernodes and defines a new

map which progressively selects long-range connections. In the remaining of the chapter, we will show that real-world networks are self-similar with respect to such scale transformation, and then explain the result in the context of the model. Finally, we will present two important applications of the framework.

4.2 The Geometric Renormalization Group transformation

The renormalization transformation is defined on the basis of the similarity dimension represented by the angular coordinate of the nodes. We present here the formulation for the $\mathbb{S}^1$ model, since it makes the similarity dimension explicit and is mathematically more tractable. Nevertheless, the corresponding $\mathbb{H}^2$ formulation is completely equivalent due to the isomorphism between both models. In order to zoom out, the transformation changes the minimum length scale from that of the original network to a larger value. It proceeds by, first, defining non-overlapping blocks of consecutive nodes of size r along the circle and, second, by coarse-graining the blocks into supernodes. Each supernode is then placed within the angular region defined by the corresponding block, with the only restriction that the order of nodes is preserved. All the links between some node in one supernode and some node in the other, if any, are renormalized into a single link between the two supernodes. This operation can be iterated starting from the original network at $l = 0$. Finally, the set of renormalized network layers l, each r^l times smaller than the original one, forms a multiscale shell of the network. Figure 4.1 illustrates the process.

It is important to note that the geometric renormalization transformation has abelian semigroup structure with respect to the composition, meaning that a certain number of iterations of a given resolution are equivalent to a single transformation of higher resolution. For instance, in Fig. 4.1, the same transformation with $r = 4$ leads from $l = 0$ to $l = 2$ in a single step. This is a crucial property for the scale observation to be well defined, since it would otherwise depend on the steps followed to reach it. Hence, it is worth proving that our transformation is a semigroup. To that end, we need to see that node i (where now indices run from 0 to $N-1$ for simplicity) is mapped to the same supernode whether we apply the coarse-graining with $r = r_1$ first and then a second time with $r = r_2$ or just once with $r = r_1 r_2$. In the first case, the step with $r = r_1$ maps i to supernode $m = \lfloor i/r_1 \rfloor$ (where $\lfloor x \rfloor$ represents the integer part of x), and then m is mapped to $s_1 = \lfloor m/r_2 \rfloor = \lfloor \lfloor i/r_1 \rfloor /r_2 \rfloor$ in the second step. In the second case, i is mapped to supernode $s_2 = \lfloor i/(r_1 r_2) \rfloor$. Notice that $s_2 = \lfloor (i/r_1)/r_2 \rfloor = \lfloor (\lfloor i/r_1 \rfloor + \alpha)/r_2 \rfloor = \lfloor (m+\alpha)/r_2 \rfloor$, where $\alpha = (i \mod r_1)/r_1 < 1$.

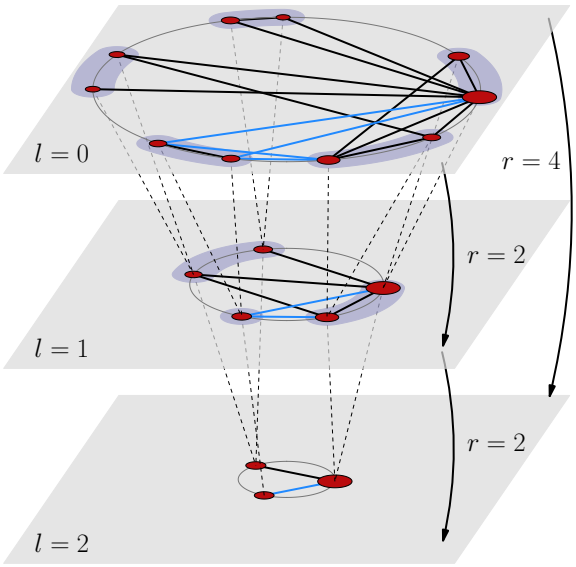


Figure 4.1: **Geometric renormalization transformation for complex networks.** Each layer is obtained after a renormalization step with resolution r starting from the original network in $l = 0$. Each node i in red is placed at an angular position $\theta_i^{(l)}$ on the $\mathbb{S}^1$ circle and has a size proportional to the logarithm of its hidden degree $\kappa_i^{(l)}$. Straight solid lines represent the links in each layer. Coarse-graining blocks correspond to the blue shadowed areas, and dashed lines connect nodes to their supernodes in layer $l+1$. Two supernodes in layer $l+1$ are connected if and only if, in layer l, some node in one supernode is connected to some node in the other (blue links give an example). The RGN transformations are valid for uneven supernode sizes as well; one could divide the circle in equally sized sectors of a certain arc-length such that they contain on average a constant number of nodes. Whenever the number of nodes is not divisible by r, the last supernode in a layer contains less than r nodes, as in the example at $l = 1$. The set of transformations parametrized by r does not include an inverse element to reverse the process.

Now,

$$\frac{m+\alpha}{r_2} = \frac{m}{r_2} + \frac{\alpha}{r_2} = \left\lfloor \frac{m}{r_2} \right\rfloor + \frac{m \mod r_2}{r_2} + \frac{\alpha}{r_2}, \tag{4.1}$$

so

$$\left\lfloor \frac{m+\alpha}{r_2} \right\rfloor = \left\lfloor \frac{m}{r_2} \right\rfloor \Leftrightarrow \frac{m \mod r_2}{r_2} + \frac{\alpha}{r_2} < 1, \tag{4.2}$$

which is always fulfilled since $\alpha < 1$ and

$$\frac{m \mod r_2 + \alpha}{r_2} \leq \frac{r_2 - 1 + \alpha}{r_2} = 1 + \frac{\alpha - 1}{r_2} < 1. \tag{4.3}$$

Thus, $s_1 = s_2$, and node i is mapped to the same supernode in both cases. It follows immediately from this result that both processes yield the same final link structure. Henceforth, the transformation is a well-defined semigroup.

4.3 Evidence of geometric scaling in real networks

The transformation introduced in the previous section can be applied to any network embedded in one-dimensional similarity space. To assess how real-world networks behave under this process, we employ it on eight different real scale-free networks from very different domains: technology, transportation, biology, social and languages. In particular, the datasets considered are:

- The Internet at the Autonomous Systems level. The data was collected by the Cooperative Association for Internet Data Analysis (CAIDA) [48] and corresponds to mid 2009.

- The Airports network. It was obtained from Refs. [125, 100]. Directed links represent flights by airlines. We consider the undirected version obtained by keeping bidirectional edges only.

- The one-mode projection onto metabolites of the human metabolic network at the cell level, as used in Ref. [155].

- The human HI-II-14 interactome. This proteome network was obtained from Ref. [140]. We removed self-loops.

- The Music network. Nodes are chords—sets of musical notes played in a single beat—and connections represent observed transitions among them in a set of songs, see Ref. [146]. The original network is weighted, directed and very dense. Hence, we applied the disparity filter [153], as in Chapter 3, with $\alpha = 0.01$ to obtain a sparser network. Finally, we kept bidirectional edges only to construct the undirected network.

- The Words network is the network of adjacency between words in Darwin's book "The Origin of Species", from Ref. [112].

- The network of email communication within the Enron company, obtained from Refs. [92, 102].

- The Drosophila network is a functional connectome within the *Drosophila melanogaster* optic medulla related to the motion detection circuit, from Ref. [161].

Name	Type	Nodes	N	γ	β	$\langle k \rangle$	$\langle c \rangle$
Internet	Technological	Autonomous systems	23748	2.17	1.44	4.92	0.61
Metabolic	Bio/Cell	Metabolites	1436	2.6	1.3	6.57	0.54
Music	Language	Chords	2476	2.27	1.1	16.66	0.82
Airports	Transportation	World airports	3397	1.88	1.7	11.32	0.63
Proteome	Bio/Cell	Proteins	4100	2.25	1.001	6.52	0.09
Words	Language	Words	7377	2.25	1.01	11.99	0.47
Enron	Social/Email	Employees	33696	2.66	1.2	10.73	0.70
Drosophila	Bio/Connectome	Neurons	1770	1.91	1.125	10.01	0.34

Table 4.1: **Overview of the considered real-world networks.** Topological details of the datasets used.

In all cases, we only considered the largest connected components. Table 4.1 contains their main topological metrics.

First, their geometric maps are obtained by embedding the nodes in the underlying geometry using the method introduced in Section 2.4.1. Second, we apply the coarse-graining by defining blocks of size $r = 2$, and iterate the process. In the limit $N \to \infty$, where N is the number of nodes, the RGN can be applied up to any desired scale of observation, whereas it is bounded to $\mathcal{O}(\log N)$ iterations in finite systems.

The resulting topological features of the renormalized networks are shown in Figs. 4.2 and 4.3. Strikingly, we observe that the degree distributions, degree-degree correlations and clustering spectra are preserved when they are represented as a function of rescaled degrees $k^{(l)}_{res} = k^{(l)}/\langle k^{(l)} \rangle$. The reason for using this new quantity is that the average degree changes along the RG flow, so the rescaling allows us to compare all layers on equal footing. Furthermore, by comparing the partitions into communities of the renormalized layers with the partitions of the original layers, we observe that the community structures also show self-similar behaviour (see details in Figs. 4.2 and 4.3). This last property suggests a new and efficient multiscale community detection algorithm for complex networks [14, 141, 6].

Interestingly, this surprising new symmetry of real-world networks, namely scale-invariance with respect to scale transformations on a hidden dimension, is actually expected if we assume the nature of real networks to be well described by the $\mathbb{S}^1$ model. In the next section, we explore this claim in detail, including the most important analytical results and their derivations.

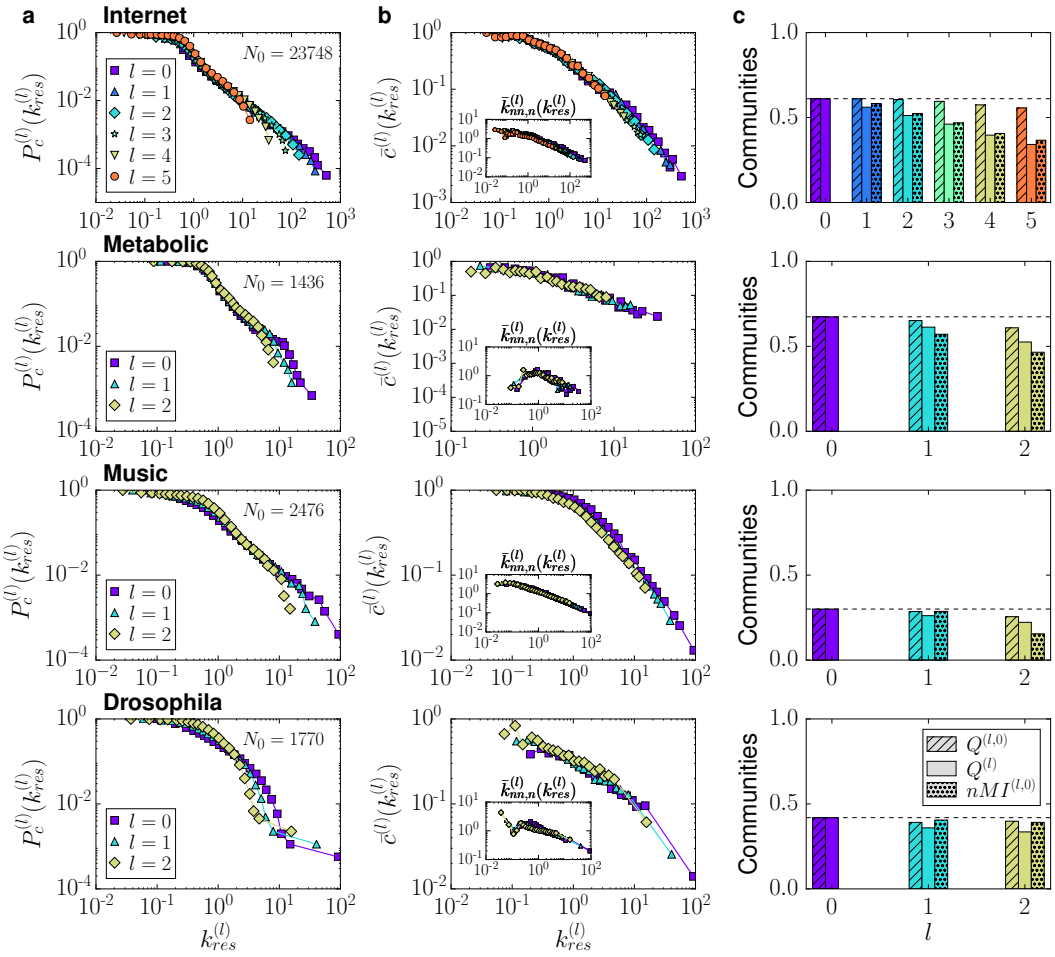


Figure 4.2: **Self-similarity of real networks along the RGN flow.** Each column shows the RGN flow with $r = 2$ of different topological features of all the networks studied. **a**, Complementary cumulative distribution of rescaled degrees $k_{res}^{(l)} = k^{(l)}/\langle k^{(l)}\rangle$. **b**, Degree-dependent clustering coefficient over rescaled-degree classes. Degree-degree correlations, as measured by the normalized average nearest neighbour degree $\bar{k}_{nn,n}(k_{res}^{(l)}) = \bar{k}_{nn}(k_{res}^{(l)})\langle k^{(l)}\rangle/\langle (k^{(l)})^2\rangle$, are shown in the insets. **c**, To asses how the community structure of the network changes with the RGN flow, we obtained a partition into communities of every layer l, $P^{(l)}$, using the Louvain method [31]; $Q^{(l)}$ represents their modularities. We also defined the partition induced by $P^{(l)}$ on the original network, $P^{(l,0)}$, obtained by considering that two nodes i and j of the original network are in the same community in $P^{(l,0)}$ if and only if the supernodes of i and j in layer l belong to the same community in $P^{(l)}$. The bars also show the modularity $Q^{(l,0)}$ of $P^{(l,0)}$ and the normalized mutual information $nMI^{(l,0)}$ between both partitions $P^{(0)}$ and $P^{(l,0)}$.

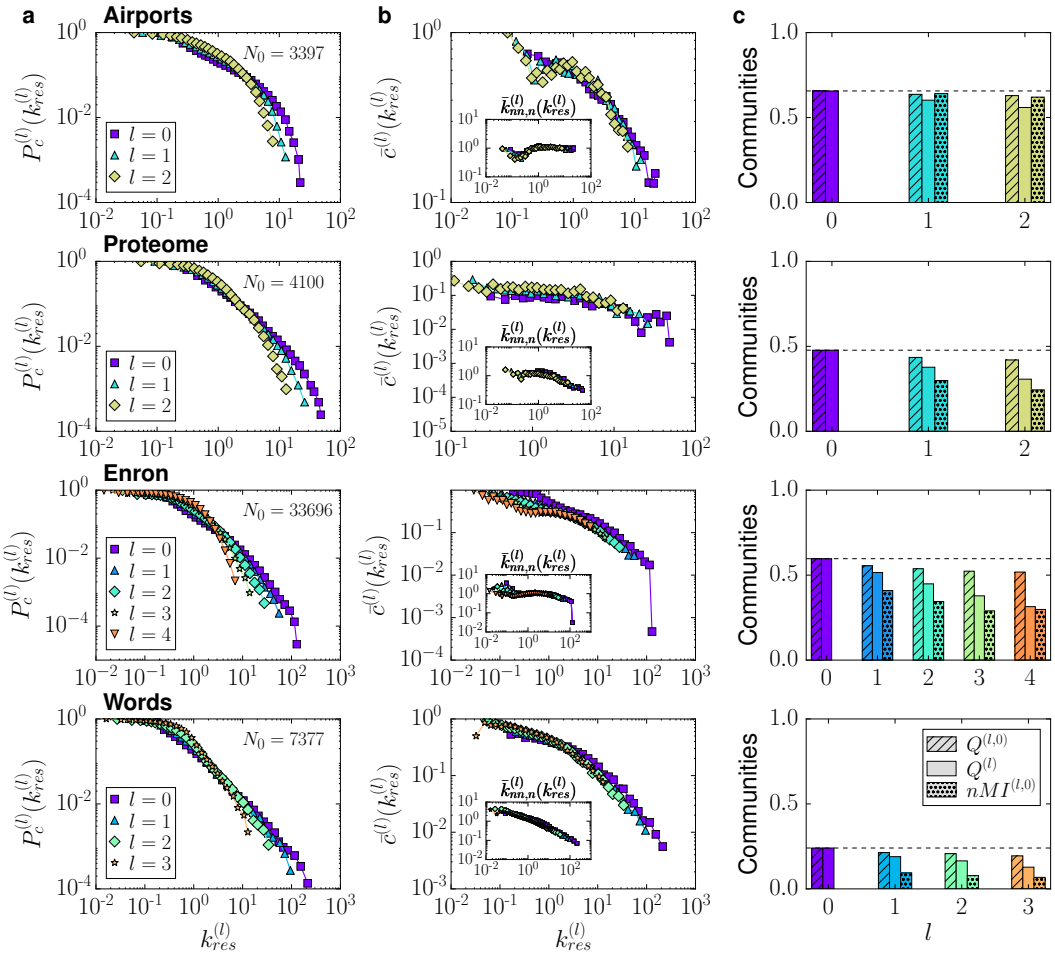


Figure 4.3: **Self-similarity of real networks along the RGN flow.** This figure shows the same results as Fig. 4.2 for the remaining datasets.

4.4 Expected scaling in geometric models

A previously unknown feature of the $\mathbb{S}^1/\mathbb{H}^2$ models is that they are renormalizable with respect to the transformation introduced in this chapter. This means that, if a real scale-free network is compatible with the models and admits a good embedding, as it is the case for the real networks analyzed in this chapter, the model will be able to predict its self-similarity and geometric scaling.

In the remaining of this more technical section, we demonstrate the renormalizability of the $\mathbb{S}^1$ model and derive the corresponding RG equations. The main goal is to show that, by assigning the appropriate hidden degrees and angles (and by rescaling the global parameters as well), the renormalized layers remain well embedded, that is, connected with the corresponding model probabilities. These transformations allow us to calculate the topological properties of the renormalized networks and therefore conclude that they exhibit the same symmetry as real-world networks.

4.4.1 Transforming the parameters of the model

In order to simplify the notation in the forthcoming derivations, all unprimed quantities will refer to layer l, whereas primed ones will correspond to layer $l+1$. Moreover, we will constrain the discussion to the particular case in which all supernodes contain the same number of nodes r for simplicity, although the following calculations are also valid for supernodes of different sizes.

Consider the probability p'_{ij} for two supernodes i and j in layer $l+1$ to be connected, which is given by the probability for at least one link between a pair of the nodes within the supernodes in layer l to exist,

$$p'_{ij} = 1 - \prod_{e=1}^{r^2} \left(1 - p_e\right), \tag{4.4}$$

where the notation has been further simplified; index e runs over all pairs of nodes (m,n) with m in supernode i and n in supernode j. The term p_e is the probability for m and n to be connected in layer l,

$$p_e = \frac{1}{1+\left(\frac{R\Delta\theta_e}{\mu(\kappa_m\kappa_n)_e}\right)^{\beta}}. \tag{4.5}$$

By rearranging Eq. (4.4), we see that it takes the same functional form as Eq. (4.5),

$$p'_{ij} = 1 - \prod_{e=1}^{r^2} \frac{1}{1+\left(\frac{R\Delta\theta_e}{\mu(\kappa_m\kappa_n)_e}\right)^{-\beta}} = 1 - \frac{1}{\prod_{e=1}^{r^2} 1+\left(\frac{\mu(\kappa_m\kappa_n)_e}{R\Delta\theta_e}\right)^{\beta}} = 1 - \frac{1}{1+\Phi'_{ij}} = \frac{1}{1+\frac{1}{\Phi'_{ij}}} \tag{4.6}$$

with

$$\Phi'_{ij} = \sum_{e=1}^{r^2} \left(\frac{\mu(\kappa_m \kappa_n)_e}{R\Delta\theta_e} \right)^\beta + \sum_{e=1}^{r^2-1} \sum_{f=e+1}^{r^2} \left(\frac{\mu(\kappa_m \kappa_n)_e}{R\Delta\theta_e} \right)^\beta \left(\frac{\mu(\kappa_m \kappa_n)_f}{R\Delta\theta_f} \right)^\beta + \dots \tag{4.7}$$

Since the angular distance between the nodes inside each block is generally smaller than the distance between i and j, all the $\Delta\theta_e$ are approximately equal ($\Delta\theta_e \approx \Delta\theta$), so we can write

$$\Phi'_{ij} \approx \left(\frac{\mu}{R\Delta\theta} \right)^\beta \sum_{e=1}^{r^2} (\kappa_m \kappa_n)_e^\beta + \left(\frac{\mu}{R\Delta\theta} \right)^{2\beta} \sum_{e=1}^{r^2-1} \sum_{f=e+1}^{r^2} (\kappa_m \kappa_n)_e^\beta (\kappa_m \kappa_n)_f^\beta + \dots \tag{4.8}$$

In the $\mathbb{S}^1$ model, R is proportional to N, whereas μ is a constant independent of the size of the system. Therefore, $\frac{\mu}{R} \ll 1$, so the first term leads Eq. (4.8) in most cases. Consequently,

$$\Phi'_{ij} \approx \left(\frac{\mu}{R\Delta\theta} \right)^\beta \sum_{e=1}^{r^2} (\kappa_m \kappa_n)_e^\beta . \tag{4.9}$$

Introducing this result into Eq. (4.6),

$$p'_{ij} \approx \frac{1}{1 + \left(\frac{R\Delta\theta}{\mu} \right)^\beta \frac{1}{\sum_{e=1}^{r^2} (\kappa_m \kappa_n)_e^\beta}}, \tag{4.10}$$

we see that, in order for the resulting expression to be congruent with the model, we need a set of equations that transform the parameters according to

$$\left(\frac{R\Delta\theta}{\mu} \right)^\beta \frac{1}{\sum_{e=1}^{r^2} (\kappa_m \kappa_n)_e^\beta} = \left(\frac{R'\Delta\theta'_{ij}}{\mu' \kappa'_i \kappa'_j} \right)^{\beta'} . \tag{4.11}$$

Let us now assume that the angular coordinate of a supernode is some generalized center of mass of the nodes it integrates, so the separation between the two renormalized nodes $\Delta\theta'_{ij}$ is approximately equal to the angular separation between the nodes that belong to different blocks, i.e. $\Delta\theta'_{ij} \approx \Delta\theta$. This implies $\beta' = \beta$. To preserve the angular density in the renormalized layer, the radius should be rescaled as $R' = \frac{R}{r}$. Setting $\mu' = \frac{\mu}{r}$, Eq. (4.11) further requires

$$\left(\kappa'_i \kappa'_j \right)^\beta = \sum_{e=1}^{r^2} (\kappa_m \kappa_n)_e^\beta , \tag{4.12}$$

which is fulfilled if every supernode is given a hidden degree as a function of the hidden degrees of the nodes it contains according to

$$\kappa'_i = \left(\sum_{j=1}^{r} \kappa_j^\beta \right)^{1/\beta} . \tag{4.13}$$

The transformation of hidden degrees preserves the semigroup structure exactly, since

$$\left(\kappa_i''\right)_r = \left(\sum_{j=1}^{r}\left(\kappa_j'\right)^\beta\right)^{1/\beta} = \left(\sum_{j=1}^{r}\sum_{k=1}^{r}\kappa_{j,k}^\beta\right)^{1/\beta} = \left(\kappa_i'\right)_{r^2}. \tag{4.14}$$

We should require the transformation of angles to preserve it as well. This can be achieved using, for instance, the following generalized center of mass with a similar functional form as Eq. (4.13)

$$\theta_i' = \left(\frac{\sum_{j=1}^{r}\left(\theta_j\kappa_j\right)^\beta}{\sum_{j=1}^{r}\kappa_j^\beta}\right)^{1/\beta}, \tag{4.15}$$

given that

$$\begin{aligned}\left(\theta_i''\right)_r &= \frac{1}{\left(\kappa_i''\right)_r}\left(\sum_{j=1}^{r}\left(\theta_j'\kappa_j'\right)^\beta\right)^{1/\beta} \\ &= \frac{1}{\left(\kappa_i'\right)_{r^2}}\left(\sum_{j=1}^{r}\left(\kappa_j'\right)^\beta\frac{1}{\left(\kappa_j'\right)^\beta}\sum_{k=1}^{r}\left(\theta_{j,k}\kappa_{j,k}\right)^\beta\right)^{1/\beta} = \left(\theta_i'\right)_{r^2}.\end{aligned} \tag{4.16}$$

Since the networks remain congruent with the $\mathbb{S}^1$ model, hidden degrees κ' remain proportional to observed degrees k', which allows us to explore the degree distribution of the renormalized layers analytically. As we will now show, if the original distribution of hidden degrees is a power law with characteristic exponent γ, the renormalized distribution is also a power law with the same exponent asymptotically, as long as $(\gamma-1)/2<\beta$. The only difference is the average degree, which scales as

$$\langle k'\rangle = r^\nu\langle k\rangle, \tag{4.17}$$

where the value of ν, to the calculation of which we will devote the rest of the section, reveals the large-scale structure of the network. Interestingly, the global parameter controlling the clustering coefficient, β, does not change along the flow, which explains the self-similarity of the clustering spectra. Notice also that any transformation for the angles mapping supernodes within the boundaries set by the nodes it contains, like Eq. (4.15), preserves the ordering of nodes and the heterogeneity in their angular density and, as a consequence, the community structure is preserved in the flow [35, 155, 177], as shown in Fig. 4.4**b**. The model is therefore renormalizable, and RGN realizations at any scale belong to the same ensemble with a different average degree, which should be rescaled to produce self-similar replicas.

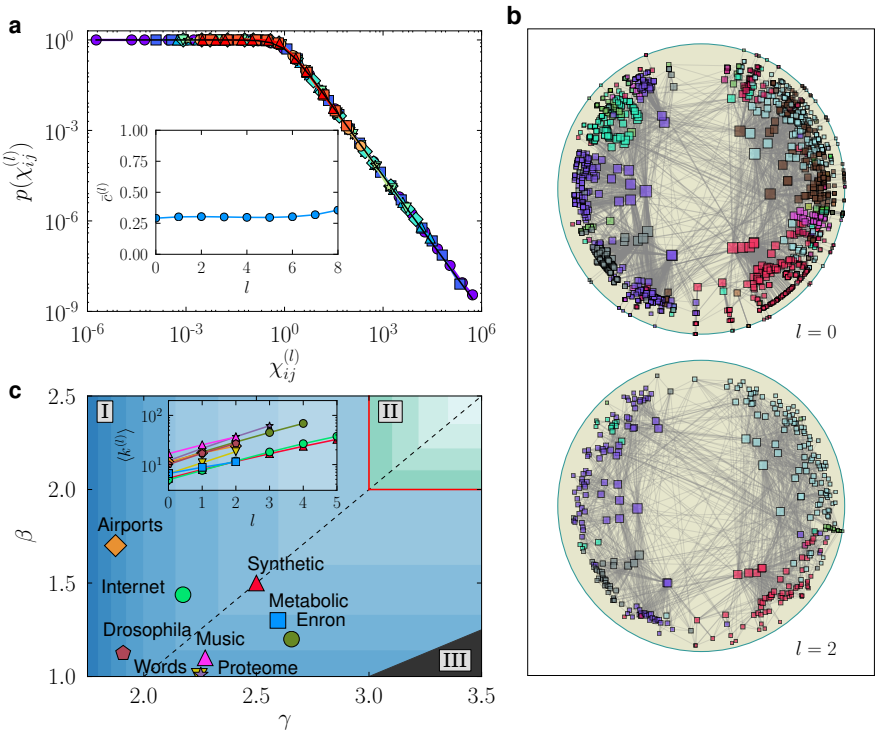


Figure 4.4: **RGN flow of synthetic and real networks. a**, Empirical connection probability in a synthetic $\mathbb{S}^1$ network. Fraction of connected pairs of nodes as a function of $\chi_{ij}^{(l)} = R^{(l)}\Delta\theta_{ij}^{(l)}/(\mu^{(l)}\kappa_i^{(l)}\kappa_j^{(l)})$ in the renormalized layers, from $l = 0$ to $l = 8$, and $r = 2$. The original synthetic network has $N \sim 225000$ nodes, $\gamma = 2.5$ and $\beta = 1.5$. The black solid line shows the theoretic curve Eq. (2.4). The inset shows the invariance of the mean local clustering along the RGN flow. **b**, Hyperbolic embedding of the metabolic network (top) and its renormalized layer $l = 2$ (bottom). The colors of the nodes correspond to the community structure detected by the Louvain algorithm. Notice that the renormalized network preserves the original community structure despite being four times smaller. **c**, Real networks in the connectivity phase diagram. The synthetic network above is also shown. Darker blue (green) in the shaded areas represent higher values of the exponent ν. The dashed line separates the γ-dominated region from the β-dominated region. In phase I, $\nu > 0$ and the network flows towards a fully connected graph. In phase II, $\nu < 0$ and the network flows towards a one-dimensional ring. The red thick line indicates $\nu = 0$ and, hence, the transition between the small-world and non-small-world phases. In region III, the degree distribution loses its scale-freeness along the flow. The inset shows the exponential increase of the average degree of the renormalized real networks $\langle k^{(l)} \rangle$ with respect to l.

4.4.2 Flow of the average degree

To assess how the average degree flows with the transformation, we can take advantage of the congruency of the renormalized networks with the underlying model, as explained before. According to Eq. (2.3), we can express the average degree in the $\mathbb{S}^1$ model, in the same simplified notation of primed and unprimed variables, as

$$\langle k' \rangle = C_0 \mu' \langle \kappa' \rangle^2, \tag{4.18}$$

where C_0 is a constant encoding all parameters that do not change as we renormalize. Since we know how μ transforms, we need to compute $\langle \kappa' \rangle$, where κ' is given by Eq. (4.13) and the original distribution of hidden degrees is assumed to be a power-law,

$$\rho(\kappa) = \frac{1-\gamma}{\kappa_c^{1-\gamma} - \kappa_0^{1-\gamma}} \kappa^{-\gamma}, \quad \kappa \in [\kappa_0, \kappa_c]. \tag{4.19}$$

Given the specific form of the transformation of hidden degrees, Eq. (4.13), it is rather advantageous to define $z \equiv \kappa^\beta$ and find their distribution, $\rho_z(z)$, which is also a power-law,

$$\rho_z(z) = \frac{1-\gamma}{\beta\left(\kappa_c^{1-\gamma} - \kappa_0^{1-\gamma}\right)} z^{-\eta}, \tag{4.20}$$

with $\eta = \frac{\gamma-1}{\beta} + 1$. Since the renormalized variables $z' \equiv \sum_r z = \kappa'^\beta$ are given by a simple sum of their unprimed counterparts, it will soon be convenient to compute the Laplace transform of Eq. (4.20), $\hat{\rho}_z(s)$. Let us first simplify the problem by taking the limit $N \to \infty$, for which $\kappa_c \to \infty$, to obtain

$$\hat{\rho}_z(s) = \int_{\kappa_0^\beta}^{\kappa_c^\beta} \rho_z(z) e^{-sz} \mathrm{d}z \xrightarrow{\kappa_c \to \infty} C s^{\eta-1} \Gamma\left(1-\eta, s z_0\right), \quad C = \frac{\gamma-1}{\beta \kappa_0^{1-\gamma}}, \tag{4.21}$$

where $\Gamma(a,b)$ is the incomplete gamma function,

$$\Gamma(a,b) = \int_b^\infty t^{a-1} e^{-t} \mathrm{d}t. \tag{4.22}$$

Using now the convolution theorem, we can obtain the Laplace transform of the variable z',

$$\hat{\rho}_{z'}(s) = \left[C s^{\eta-1} \Gamma\left(1-\eta, s z_0\right)\right]^r. \tag{4.23}$$

From this function, $\langle\kappa'\rangle$ can be computed as the $1/\beta$-th moment of z', that is, $\langle\kappa'\rangle = \langle z'^{1/\beta}\rangle$. Although this path indeed leads to the exact solution for $r = 2$ (see Section 4.8.1), the result is very convoluted. We thus need to find a simpler approximate solution to study the limit $r \to \infty$. For that purpose, it is useful to note that, if $\gamma < 2\beta + 1$, then $\eta < 3$, so we can expect z' (and, consequently, κ') to be approximately power-law distributed as well because the central limit theorem does not apply (the opposite case corresponds to phase III in Fig. 4.4**c**) [79]. Under this assumption, the distribution of z' is given by

$$\rho_{z'}(z') = C' z'^{-\eta}, \quad C' = \frac{\gamma - 1}{\beta \kappa_0'^{1-\gamma}}, \tag{4.24}$$

which makes calculations amenable. We can now study two different cases separately:

i. $1 < \eta < 2$: In this case, we determine the value of C' and, with it, $\langle\kappa'\rangle = \kappa_0' \frac{\gamma-1}{\gamma-2}$. If the assumption in Eq. (4.24) is correct, its Laplace transform $\hat{\rho}_{z'}(s)$ must behave as [83]

$$\hat{\rho}_{z'}(s) = 1 + C' s^{\eta-1} \Gamma(1-\eta), \quad s \to 0^+. \tag{4.25}$$

According to Eq. (4.23),

$$\begin{aligned}
\hat{\rho}_{z'}(s) &= \left[C s^{\eta-1} \Gamma\left(1-\eta, s z_0\right) \right]^r \\
&= \left[C s^{\eta-1} \Gamma\left(1-\eta\right) \left(1 - (s z_0)^{1-\eta} e^{-s z_0} \sum_{n=0}^{\infty} \frac{(s z_0)^n}{\Gamma(2-\eta+n)} \right) \right]^r \\
&= \left[C \Gamma\left(1-\eta\right) \left(s^{\eta-1} - z_0^{1-\eta} e^{-s z_0} \sum_{n=0}^{\infty} \frac{(s z_0)^n}{\Gamma(2-\eta+n)} \right) \right]^r \\
&\xrightarrow{s \to 0^+} \left[C \Gamma\left(1-\eta\right) \left(s^{\eta-1} - z_0^{1-\eta} \sum_{n=0}^{\infty} \frac{(s z_0)^n}{\Gamma(2-\eta+n)} \right) \right]^r .
\end{aligned} \tag{4.26}$$

In the above expression, we see that the term that does not depend on s is given by the product of the r terms with $n = 0$, whereas the term of order $s^{\eta-1}$ is given by the sum of the r products of $s^{\eta-1}$ with the remaining $r - 1$

terms with $n=0$. Thus, we find

$$\begin{aligned}
\hat{\rho}_{z'}(s) &\xrightarrow{s\to 0^+} \left[C\Gamma\left(1-\eta\right)\left(s^{\eta-1}-z_0^{1-\eta}\sum_{n=0}^{\infty}\frac{(sz_0)^n}{\Gamma(2-\eta+n)}\right)\right]^r \\
&\approx C^r\Gamma^r\left(1-\eta\right)\left[(-1)^r\frac{z_0^{r(1-\eta)}}{\Gamma^r(2-\eta)}+r(-1)^{r-1}s^{\eta-1}\frac{z_0^{(r-1)(1-\eta)}}{\Gamma^{r-1}(2-\eta)}\right] \\
&= \left(\frac{\gamma-1}{\beta\kappa_0^{1-\gamma}}\right)^r\Gamma^r\left(1-\eta\right) \\
&\times\left[(-1)^r\frac{z_0^{r(1-\eta)}}{(1-\eta)^r\Gamma^r(1-\eta)}+r(-1)^{r-1}s^{\eta-1}\frac{z_0^{(r-1)(1-\eta)}}{(1-\eta)^{r-1}\Gamma^{r-1}(1-\eta)}\right] \\
&= \left(\frac{\eta-1}{z_0^{1-\eta}}\right)^r\left[\frac{z_0^{r(1-\eta)}}{(\eta-1)^r}+rs^{\eta-1}\Gamma\left(1-\eta\right)\frac{z_0^{(r-1)(1-\eta)}}{(\eta-1)^{r-1}}\right] \\
&= 1+r\frac{\eta-1}{z_0^{1-\eta}}s^{\eta-1}\Gamma\left(1-\eta\right).
\end{aligned} \tag{4.27}$$

We can now identify C' as

$$C'=\frac{\gamma-1}{\beta\kappa_0'^{1-\gamma}}=r\frac{\eta-1}{z_0^{1-\eta}}=r\frac{\gamma-1}{\beta\kappa_0^{1-\gamma}}, \tag{4.28}$$

which implies

$$\kappa_0'=r^{\frac{1}{\gamma-1}}\kappa_0 \tag{4.29}$$

and, finally,

$$\langle\kappa'\rangle=r^{\frac{1}{\gamma-1}}\langle\kappa\rangle. \tag{4.30}$$

Plugging this result into Eq. (4.18),

$$\langle k'\rangle=C_0\frac{\mu}{r}r^{\frac{2}{\gamma-1}}\langle\kappa\rangle^2=r^{\frac{2}{\gamma-1}-1}\langle k\rangle\xrightarrow{r\to\infty}\begin{cases}\infty & \gamma<3\\ cte. & \gamma=3\\ 0 & \gamma>3\end{cases} \tag{4.31}$$

Thus, for $1<\eta<2$, the value of ν depends on degree heterogeneity only. We will refer to this as the γ-dominated region.

ii. $2<\eta<3$: This case is much simpler, since $\langle z\rangle$ and hence $\langle z'\rangle$ are finite and can be easily computed. Indeed, given that $\langle z'\rangle=r\langle z\rangle$, we see that

$$\begin{aligned}
\langle\kappa'\rangle &= \frac{\gamma-1}{\gamma-2}\kappa_0'=\frac{\gamma-1}{\gamma-2}\left(z_0'\right)^{1/\beta}=\frac{\gamma-1}{\gamma-2}\left(\frac{\eta-2}{\eta-1}\langle z'\rangle\right)^{1/\beta}=\frac{\gamma-1}{\gamma-2}\left(\frac{\eta-2}{\eta-1}r\langle z\rangle\right)^{1/\beta} \\
&= \frac{\gamma-1}{\gamma-2}\left(r\frac{\eta-2}{\eta-1}\frac{\eta-1}{\eta-2}z_0\right)^{1/\beta}=\frac{\gamma-1}{\gamma-2}\left(r\kappa_0^{\beta}\right)^{1/\beta}=r^{1/\beta}\langle\kappa\rangle.
\end{aligned} \tag{4.32}$$

This result and Eq. (4.18) together imply

$$\langle k' \rangle = C_0 \frac{\mu}{r} r^{2/\beta} \langle \kappa \rangle^2 = r^{2/\beta - 1} \langle k \rangle \xrightarrow{r \to \infty} \begin{cases} \infty & \beta < 2 \\ cte. & \beta = 2 \\ 0 & \beta > 2 \end{cases} \tag{4.33}$$

In this other case, ν is a function of the coupling to the metric space, so we will denote this as the β-dominated region.

Both solutions, Eqs. (4.31) and (4.33), are equivalent at $\eta = 2$, since

$$\eta = 2 \Rightarrow \frac{\gamma - 1}{\beta} = 1 \Rightarrow \beta = \gamma - 1. \tag{4.34}$$

4.4.3 Connectivity phase diagram

In terms of Eq. (4.17), we see that, if $\gamma < 3$ or $\beta < 2$ (phase I in Fig. 4.4**c**), then $\nu > 0$ and the model flows towards a highly connected graph; the average degree is preserved if $\gamma = 3$ and $\beta \geq 2$ or $\beta = 2$ and $\gamma \geq 3$, which indicates that the network is at the edge of the transition between the small-world and non-small-world phases; and $\nu < 0$ if $\gamma > 3$ and $\beta > 2$, causing the RGN flow to produce sparser networks approaching a unidimensional ring structure as a fixed point (phase II in Fig. 4.4**c**). In this case, the renormalized layers eventually lose the small-world property. Indeed, this result implies that, when the system is viewed from a sufficiently large scale of observation, all short-cuts have been integrated within the supernodes, so the average shortest path length is proportional to the system size. Finally, if $\beta < (\gamma - 1)/2$, the degree distribution becomes increasingly homogeneous as $r \to \infty$, that is, the approximation in Eq. (4.24) fails as a consequence of the central limit theorem (phase III in Fig. 4.4**c**), revealing that degree heterogeneity is only present at short scales.

In Fig. 4.4**c**, the real networks considered are displayed in the connectivity phase diagram. All of them lay in the region of small-world networks. Furthermore, all of them, except the Internet, the Airports, and the Drosophila networks, belong to the β-dominated region. The inset shows the behaviour of the average degree of each layer l, $\langle k^{(l)} \rangle$; as predicted, it grows exponentially in all cases.

Finally, the RGN transformation can be reformulated for the model in D dimensions. We have recalculated the connectivity phase diagram from Fig. 4.4**c**, obtaining qualitatively the same transitions and phases, including region III. We have tested the renormalization transformation using the one-dimensional embedding of networks generated in higher dimensions for which the clustering is realistic, that is, $D \lesssim 10$, and found the same results as in the $D = 1$ case (see Section 4.8.2). The agreement is explained by the fact that the one-dimensional embedding provides a faithful representation for low-dimensional similarity-space networks, as we discussed in Chapter 2.

4.5 Applications

We propose next two specific applications. The first one, the production of downscaled network replicas, singles out a specific scale, while the second one, a multiscale navigation protocol, exploits multiple scales simultaneously.

4.5.1 Smaller-scale replicas for efficient dynamics simulation

Being able to downscale the topology of large real-world complex networks to produce smaller high-fidelity replicas can find useful applications, for instance, as an alternative or guidance to sampling methods in large-scale simulations and, in networked communication systems like the Internet, as a reduced testbed to analyze the performance of new routing protocols [130, 132, 172, 173]. Downscaled network replicas can also be used to perform finite size scaling of critical phenomena in real networks, so that critical exponents could be evaluated starting from a single size instance network. However, the success of such programs is based upon the quality of the downscaled versions, that should reproduce not only local properties of the original network but also its mesoscopic structure. We now present a method for their construction which exploits the fact that, under renormalization, a scale-free network remains self-similar and congruent with the underlying geometric model in all the self-similarity range of the multilayer shell.

The idea is to single out a specific scale after a certain number of renormalization steps. Typically, the renormalized average degree of real networks increases in the flow (see inset in Fig. 4.4**c**) so, in order to reduce the average degree in a renormalized network to the level of the original one, we apply a pruning of links using the underlying metric model. Since the networks are congruent with the model in all layers, the main idea is to decrease the value of $\mu^{(l)}$ to a new value $\mu^{(l)}_{\text{new}}$—which implies that the connection probability of every pair of nodes (i,j), $p^{(l)}_{ij}$, decreases to $p^{(l)}_{ij,\text{new}}$. We then prune the existing links by keeping them with probability

$$q^{(l)}_{ij} = \frac{p^{(l)}_{ij,\text{new}}}{p^{(l)}_{ij}}. \tag{4.35}$$

Therefore, the probability for a link to exist in the pruned network is,

$$P\left\{a^{(l)}_{ij,\text{new}} = 1\right\} = p^{(l)}_{ij} q^{(l)}_{ij} = p^{(l)}_{ij,\text{new}}, \tag{4.36}$$

whereas the probability for it not to exist is

$$P\left\{a^{(l)}_{ij,\text{new}} = 0\right\} = 1 - p^{(l)}_{ij} + p^{(l)}_{ij}\left(1 - q^{(l)}_{ij}\right) = 1 - p^{(l)}_{ij,\text{new}}, \tag{4.37}$$

that is, the pruned network has a lower average degree and is also congruent with the underlying metric space model with the new value of $\mu^{(l)}_{\text{new}}$. Hence, we only need to find the right value of $\mu^{(l)}_{\text{new}}$ so that $\langle k^{(l)}_{\text{new}}\rangle = \langle k^{(0)}\rangle$. In the thermodynamic limit, the average degree of an $\mathbb{S}^1$ network is proportional to μ, so we could simply set

$$\mu^{(l)}_{\text{new}} = \frac{\langle k^{(0)}\rangle}{\langle k^{(l)}\rangle}\mu^{(l)}. \tag{4.38}$$

However, since we consider real-world networks, finite-size effects play an important role, so we need to correct the value of $\mu^{(l)}_{\text{new}}$ in Eq. (4.38). To this end, we use a correcting factor c, initially set to $c = 1$, and use $\mu^{(l)}_{\text{new}} = c\frac{\langle k^{(0)}\rangle}{\langle k^{(l)}\rangle}\mu^{(l)}$; for every value of c, we prune the layer-l network. If $\langle k^{(l)}_{\text{new}}\rangle > \langle k^{(0)}\rangle$, we give c the new value $c - 0.1u \to c$, where u is a random variable uniformly distributed between 0 and 1. Similarly, if $\langle k^{(l)}_{\text{new}}\rangle < \langle k^{(0)}\rangle$, $c + 0.1u \to c$. The process ends when $|\langle k^{(l)}_{\text{new}}\rangle - \langle k^{(0)}\rangle|$ is below a given threshold (in our case, we set it to 0.1).

To illustrate the high-fidelity that downscaled network replicas can achieve, we use them to reproduce the behaviour of dynamical processes in real networks. We selected three different dynamical processes, the classic ferromagnetic Ising model, the susceptible-infected-susceptible (SIS) epidemic spreading model, and the Kuramoto model of synchronization.

The Ising model is an equilibrium model of interacting spins [58]. Every node i is assigned a variable s_i with two possible values $s_i = \pm 1$, and the energy of the system is, in the absence of external field, given by the Hamiltonian

$$\mathcal{H} = -\sum_{i<j} J_{ij} a_{ij} s_i s_j, \tag{4.39}$$

where a_{ij} are the elements of the adjacency matrix and J_{ij} are coupling constants which we set to one. We start from an initial condition with $s_i = 1$ for all i and explore the ensemble of configurations using the Metropolis-Hastings algorithm: we randomly select one node and propose a change in its spin, $-s_i \to s_i$. If $\Delta\mathcal{H} \leq 0$, we accept the change; otherwise, we accept it with probability $e^{-\Delta\mathcal{H}/T}$, where T is the temperature acting as a control parameter. The order parameter is the absolute magnetization per spin $|m|$, where $m = \frac{1}{N}\sum_i s_i$; if all spins point in the same direction, $|m| = 1$, whereas $|m| = 0$ if half the spins point in each direction.

In the SIS dynamical model of epidemic spreading [135], every node i can present two states at a given time t, susceptible ($n_i(t) = 0$) or infected ($n_i(t) = 1$). Both infection and recovery are Poisson processes. An infected node recovers with rate 1, whereas infected nodes infect their susceptible neighbours at rate λ. We simulate this process using the continuous-time Gillespie algorithm with all nodes initially infected. The order parameter is the prevalence or fraction of infected nodes $\rho(t) = \frac{1}{N}\sum_i n_i(t)$.

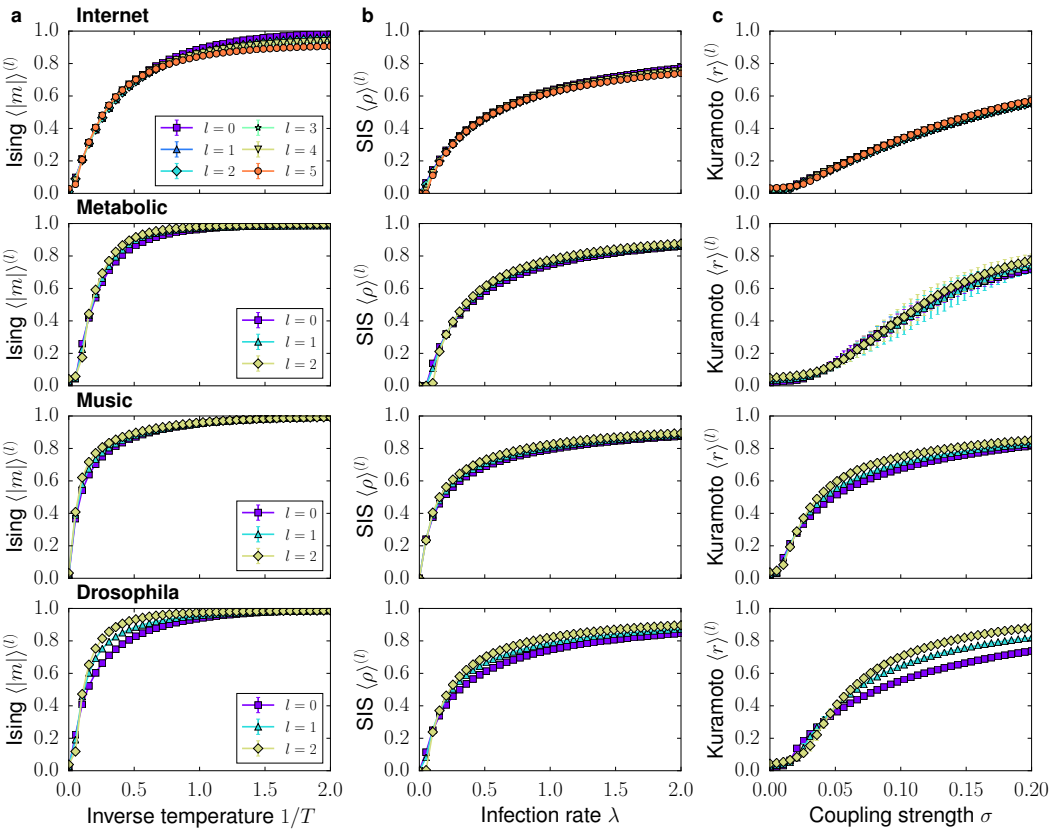


Figure 4.5: **Dynamics on the downscaled replicas.** Each column shows the order parameters versus the control parameters of different dynamical processes on the original and downscaled replicas of the real networks considered with $r = 2$, that is, every value of l identifies a network 2^l times smaller than the original one. All points show the results averaged over 100 simulations. Error bars indicate the standard deviations of the order parameters. In every realization, we compute an average of the order parameter in the stationary state. In the case of the SIS model, the single-realization mean of prevalence values is weighted by time. **a**, Magnetization $\langle |m| \rangle^{(l)}$ of the Ising model as a function of the inverse temperature $1/T$. **b**, Prevalence $\langle \rho \rangle^{(l)}$ of the SIS model as a function of the infection rate λ. **c**, Coherence $\langle r \rangle^{(l)}$ of the Kuramoto model as a function of the coupling strength σ. In all cases, the curves of the smaller-scale replicas are extremely similar to the results obtained on the original networks.

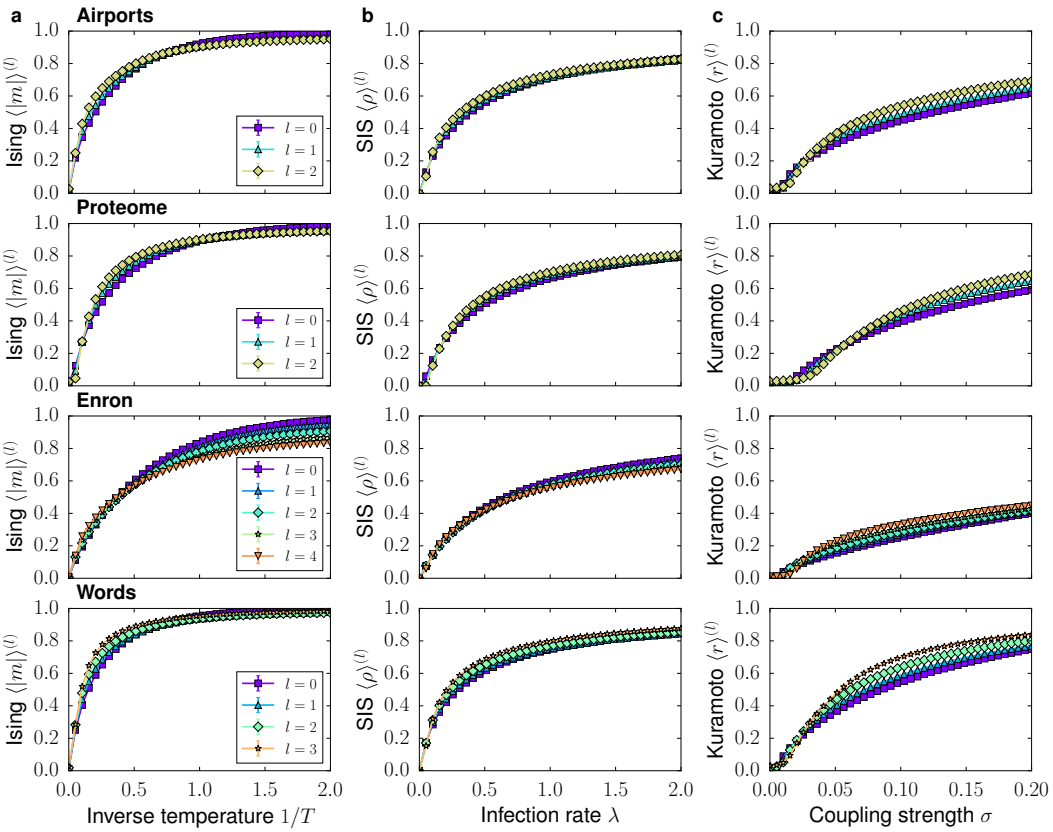


Figure 4.6: **Dynamics on the downscaled replicas.** This figure shows the same results as Fig. 4.5 for the remaining datasets.

The Kuramoto model is a dynamical model for coupled oscillators. Every node i is described by a natural frequency ω_i and a time-dependent phase $\theta_i(t)$. A node's phase evolves according to

$$\dot{\theta}_i = \omega_i + \sigma \sum_{i<j} a_{ij} \sin\left(\theta_j(t) - \theta_i(t)\right), \tag{4.40}$$

where a_{ij} are the adjacency matrix elements and σ is the coupling strength. We integrate the equations of motion using Heun's method. Initially, the phases $\theta_i(0)$ and the frequencies ω_i are randomly drawn from the uniform distributions $U(-\pi,\pi)$ and $U(-1/2,1/2)$ respectively, as in Ref. [113]. The order parameter $r(t) = \frac{1}{N}\left|\sum_i e^{i\theta_i(t)}\right|$ measures the phase coherence of the set of nodes; if all nodes oscillate in phase, $r(t) = 1$, whereas $r(t) \to 0$ if nodes oscillate in a disordered manner.

We test these dynamics in all the self-similar network layers of the real networks analyzed in this chapter. Results are shown in Figs. 4.5 and 4.6. Quite remarkably, for all dynamics and all networks, we observe very similar results between the original and downscaled replicas at all scales. This is particularly interesting as these dynamics have a strong dependence on the mesoscale structure of the underlying networks.

4.5.2 A multiscale navigation protocol in hyperbolic space

Next, we introduce a new multiscale navigation (MN) protocol for networks embedded in hyperbolic space, which improves single-layer results [35]. To this end, we exploit the quasi-isomorphism between the $\mathbb{S}^1$ model and the $\mathbb{H}^2$ model in hyperbolic space [93, 94] to produce a purely geometric representation of the multiscale shell.

The MN protocol is based on greedy routing. However, instead of using the geometric information in the physical layer only, nodes use multiscale information as well. By doing so, in some cases in which the packet being forwarded would get trapped in a loop using the single-layer protocol, the MN protocol can find alternative paths by taking advantage of greedy forwarding in the coarse-grained layers, which usually has a higher success rate. When node i needs to send a packet to some destination node j, node i performs a virtual greedy forwarding step in the highest possible layer to find which supernode should be next in the greedy path. Based on this, node i then forwards the packet to its physical neighbour in the real network that guarantees that it will eventually reach such supernode. The process is depicted in Fig. 4.7**a**. To ensure navigation inside supernodes, we require an extra condition in the renormalization process and only consider blocks of connected consecutive nodes (a single node can be left alone forming a supernode by itself). Notice that the new requirement does

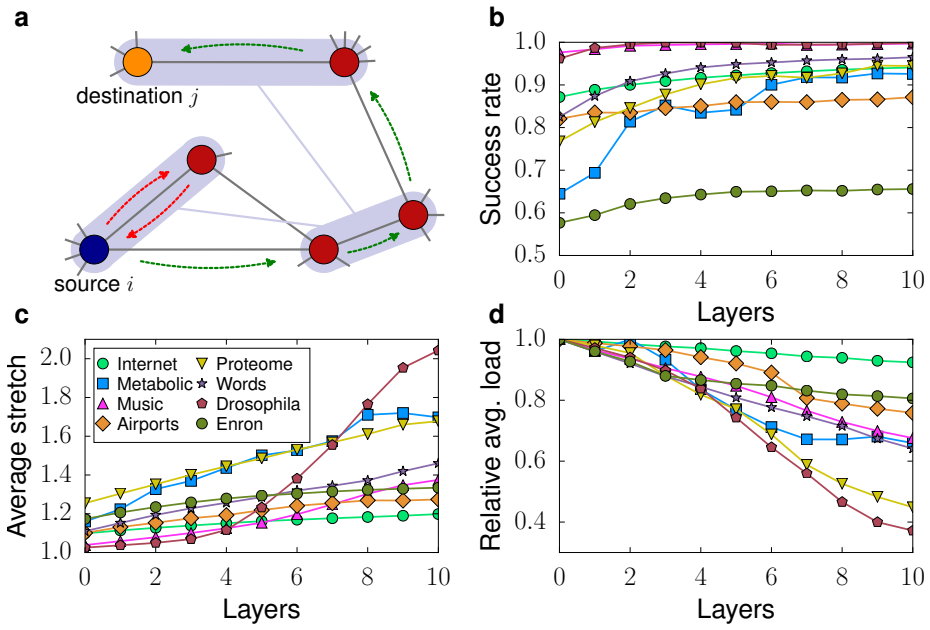


Figure 4.7: **Performance of the multiscale navigation protocol in real networks. a**, Illustration of the multiscale navigation protocol. Red arrows show the unsuccessful greedy path in the original layer of a message attempting to reach the target yellow node. Green arrows show the successful greedy path from the same source using both layers. **b**, Success rate as a function of the number of layers used in the process, computed for 10^5 randomly selected pairs. **c**, Average stretch $\langle l_g / l_s \rangle$, where l_g is the topological length of a path found by the algorithm and l_s is the actual shortest path length in the network. **d**, Relative average load $\langle \lambda_L^i \rangle_{\text{Hubs}} / \langle \lambda_0^i \rangle_{\text{Hubs}}$ of hubs, where λ_L^i is the fraction of successful greedy paths that pass through node i in the multiscale navigation protocol with L renormalized layers. The averages are computed over the 20 highest-degree nodes in the network.

not alter the self-similarity of the renormalized networks forming the multiscale shell nor the congruency with the hidden metric space [3]. Once L layers have been generated, the multiscale navigation protocol requires every node i to be provided with the following local information:

1. The coordinates $(r_i^{(l)}, \theta_i^{(l)})$ of node i in every layer l.

2. The list of (super)neighbours of node i in every layer as well as their coordinates.

3. Let SuperN(i,l) be the supernode to which i belongs in layer l. If SuperN(i,l) is connected to SuperN(k,l) in layer l, at least one of the (su-

per)nodes in layer $l-1$ belonging to SuperN(i,l) must be connected to at least one of the (super)nodes in layer $l-1$ belonging to SuperN(k,l); such node is called *gateway*. For every superneighbour of node SuperN(i,l) in layer l, node i knows which (super)node or (super)nodes in layer $l-1$ are gateways reaching it. Notice that both the gateways and SuperN$(i,l-1)$ belong to SuperN(i,l) in layer l so, in layer $l-1$, they must either be the same (super)node or different but connected (super)nodes.

4. If SuperN$(i,l-1)$ is a gateway reaching some supernode s, at least one of its (super)neighbours in layer $l-1$ belongs to s; node i knows which.

This information allows us to navigate the network as follows. Let j be the destination node to which i wants to forward a message, and let node i know j's coordinates in all L layers $(r_j^{(l)},\theta_j^{(l)})$. In order to decide which of its physical neighbours (i. e., in layer 0) should be next in the message-forwarding process, node i must first check if it is connected to j; in that case, the decision is clear. If it is not, it must:

1. Find the highest layer $l_{\max}$ in which SuperN$(i,l_{\max})$ and SuperN$(j,l_{\max})$ still have different coordinates. Set $l=l_{\max}$.

2. Perform a standard step of greedy routing in layer l: find the closest neighbour of SuperN(i,l) to SuperN(j,l). This is the current target SuperT(l).

3. While $l>0$, look into layer $l-1$:

   - Set $l=l-1$.
   - If SuperN(i,l) is a gateway connecting to some (super)node within SuperT$(l+1)$, node i sets as new current target SuperT(l) its (super)neighbour belonging to SuperT$(l+1)$ closest to SuperN(j,l).
   - Else node i sets as new target SuperT(l) the gateway in SuperN$(i,l+1)$ connecting to SuperT$(l+1)$ (its (super)neighbor belonging to SuperN$(i,l+1)$).

4. In layer $l=0$, SuperT(0) belongs to the real network and she is a neighbour of i, so node i forwards the message to SuperT(0).

We have tested this protocol on the eight different networks considered in this chapter and, as usual, we have considered the success rate and the stretch of successful path as performance metrics. Figure 4.7**b** shows the increase of the success rate as the number of renormalized layers L used in the navigation process is increased. Interestingly, the success rate always increases, even in systems with very high navigability in $l=0$ like the Internet, while this improvement

increases the stretch of successful paths only mildly, Fig. 4.7**c**. Counterintuitively, the slight increase of the stretch reduces the burden on highly connected nodes, see Fig. 4.7**d**. As the number of renormalized layers L increases, the average fraction of successful paths passing through the most connected hubs in the network decreases. The improvements come at the expense of adding information about the supenodes to the knowledge needed for standard greedy routing in single-layered networks. However, the trade-off between improvement and information overload is advantageous as, for many systems, the addition of just one or two renormalized layers already produces a notable effect.

4.6 The partition function in hyperbolic space

For the sake of completeness, in this section we show how the RGN presented in this chapter can be applied to the $\mathbb{H}^2$ model by exploiting its connection with statistical physics. Using the isomorphism between the two geometric models, the connection probability, Eq. (4.5), in the $\mathbb{H}^2$ model reads

$$p_{mn} = \frac{1}{1+e^{\frac{\beta}{2}(x_{mn}-\hat{R})}}. \tag{4.41}$$

Now, let $a_{mn}=1$ if the link between nodes m and n exists and $a_{mn}=0$ otherwise; Eq. (4.41) can be written as

$$P(a_{mn}) = \frac{e^{-\beta\frac{a_{mn}}{2}(x_{mn}-\hat{R})}}{1+e^{-\frac{\beta}{2}(x_{mn}-\hat{R})}}, \tag{4.42}$$

which means that, in the $\mathbb{H}^2$ model, every pair of nodes represents a fermionic state of energy $x_{mn}/2$ in the grand-canonical ensemble with $\hat{R}/2$ playing the role of the chemical potential. Indeed, since a network can be represented by the set $\{a_{mn}\}$, the likelihood of a given network is given by

$$P(\{a_{mn}\}) = \prod_{m<n} \frac{e^{-\beta\frac{a_{mn}}{2}(x_{mn}-\hat{R})}}{1+e^{-\frac{\beta}{2}(x_{mn}-\hat{R})}}, \tag{4.43}$$

that is, by the probability of the corresponding microstate of the gas of non-interacting fermions. The partition function of the system is

$$Z = \prod_{m<n}\sum_{a_{mn}=0}^{1} e^{-\beta\frac{a_{mn}}{2}(x_{mn}-\hat{R})} = \prod_{m<n}\left(1+e^{-\frac{\beta}{2}(x_{mn}-\hat{R})}\right). \tag{4.44}$$

When we apply the renormalization transformation, every node m (n) is mapped to a supernode i (j). We can rearrange the terms in the partition function according to such mapping as

$$Z = \prod_{i=1}^{\lfloor\frac{N}{r}\rfloor}\prod_{t=1}^{\frac{r(r-1)}{2}}\left(1+e^{-\frac{\beta}{2}(x_t-\hat{R})}\right)\prod_{i<j}\prod_{e=1}^{r^2}\left(1+e^{-\frac{\beta}{2}(x_e-\hat{R})}\right). \tag{4.45}$$

The first double product in the above expression corresponds to the partial sum over the links among the nodes within every supernode i (hence, there are $N(r-1)/2$ such terms), whereas the second double product represents the partial sum over the links among nodes in different supernodes i and j; thus, it contains $(N^2 - Nr)/2$ terms. According to the isomorphism with the $\mathbb{S}^1$ model,

$$e^{-\frac{\beta}{2}(x_{mn}-\hat{R})} = \left(\frac{\mu\kappa_m\kappa_n}{R\Delta\theta_{mn}}\right)^{\beta}, \tag{4.46}$$

so the rightmost term in Eq. (4.45) reads

$$\prod_{i<j}\prod_{e=1}^{r^2}\left(1+e^{-\frac{\beta}{2}(x_e-\hat{R})}\right) = \prod_{i<j}\prod_{e=1}^{r^2}\left(1+\left(\frac{\mu(\kappa_m\kappa_n)_e}{R\Delta\theta_e}\right)^{\beta}\right) = \prod_{i<j}\left(1+\Phi'_{ij}\right), \tag{4.47}$$

where Φ'_{ij} is given by Eq. (4.7). Using Eq. (4.9) and Eq. (4.11), which are fulfilled with the RG transformations derived in Section 4.4.1, yields

$$\begin{aligned}\prod_{i<j}\prod_{e=1}^{r^2}\left(1+e^{-\frac{\beta}{2}(x_e-\hat{R})}\right) &\approx \prod_{i<j}\left(1+\left(\frac{\mu}{R\Delta\theta}\right)^{\beta}\sum_{e=1}^{r^2}(\kappa_m\kappa_n)_e^{\beta}\right)\\ &= \prod_{i<j}\left(1+\left(\frac{\mu'\kappa'_i\kappa'_j}{R'\Delta\theta'_{ij}}\right)^{\beta}\right) = \prod_{i<j}\left(1+e^{-\frac{\beta}{2}(x'_{ij}-\hat{R}')}\right) = Z'.\end{aligned} \tag{4.48}$$

The leftmost term in Eq. (4.45) can be written as

$$\prod_{i=1}^{\lfloor\frac{N}{r}\rfloor}\prod_{t=1}^{\frac{r(r-1)}{2}}\left(1+e^{-\frac{\beta}{2}(x_t-\hat{R})}\right) = \prod_{i=1}^{\lfloor\frac{N}{r}\rfloor}\prod_{t=1}^{\frac{r(r-1)}{2}}\left(1+\left(\frac{\mu(\kappa_m\kappa_n)_t}{R\Delta\theta_t}\right)^{\beta}\right). \tag{4.49}$$

In the particular case of $r = 2$, we integrate consecutive nodes separated by a typical angular distance $\Delta\theta_t \approx 2\pi/N$. Hence, $R\Delta\theta_t \approx 1$, so

$$\prod_{i=1}^{\lfloor\frac{N}{r}\rfloor}\prod_{t=1}^{\frac{r(r-1)}{2}}\left(1+e^{-\frac{\beta}{2}(x_t-\hat{R})}\right) \approx e^{\sum_{i=1}^{\lfloor N/2\rfloor}\ln\left(1+(\mu(\kappa_m\kappa_n)_i)^{\beta}\right)} \approx e^{\frac{N}{2}\left\langle\ln\left(1+(\mu\kappa_m\kappa_n)^{\beta}\right)\right\rangle}. \tag{4.50}$$

Defining

$$\zeta \equiv e^{\left\langle\ln\left(1+(\mu\kappa_m\kappa_n)^{\beta}\right)\right\rangle} = e^{\int\ln\left(1+(\mu\kappa_m\kappa_n)^{\beta}\right)\rho_\kappa(\kappa_m)\rho_\kappa(\kappa_n)\mathrm{d}\kappa_m\mathrm{d}\kappa_n} \tag{4.51}$$

we can write Eq. (4.45) as

$$Z = \zeta^{N/2}Z', \tag{4.52}$$

where $Z' = \sum_{\{a_{ij}\}} e^{-\beta H'(\{a_{ij}\})}$.

4.7 Discussion

In the previous chapters, we have discussed the idea of geometric spaces underlying real complex networks by examining how unveiling them results in meaningful maps of real systems in terms of similarity between nodes. In this chapter, however, we focus on finding a valid source of scales for real networks using a hidden similarity space. As we show, our geometric-embedding-based approach can provide essential insights to our understanding of network topology. Indeed, the existence of a metric space underlying complex networks allows us to define a geometric renormalization group that reveals their multiscale nature. The geometric models of scale-free networks used in this thesis are shown to be self-similar under the RGN transformation, in accordance with the finding that self-similarity under geometric renormalization is a ubiquitous symmetry of real world scale-free networks, which provides a new evidence supporting the hypothesis that hidden metric spaces underlie real networks.

The renormalization group presented in this chapter is similar in spirit to the topological renormalization studied in [156, 80, 157, 91, 138, 143] and should be taken as complementary. Instead of using shortest paths as a source of length scales to explore the fractality of networks, we use a continuum geometric framework that allows us to unveil the role of degree heterogeneity and clustering in the self-similarity properties of networks. Indeed, a crucial point in the $\mathbb{S}^1$ is the explicit contribution of degrees to the probability of connection, giving the clue by which we can produce both short-range and long-range connections using a single mechanism captured in a universal connectivity law. The combination of similarity with degrees is a necessary condition to make the model predictive of the multiscale properties of real networks.

From a fundamental point of view, the geometric renormalization group introduced here has proven to be an exceptional tool to unravel the global organization of complex networks across scales and promises to become a standard methodology to analyze real complex networks. It can also help in areas like the study of metapopulation models, in which transportation fluxes or population movements happen both on a local and a global scale [52]. From a practical point of view, we envision many applications. In large-scale simulations, downscaled network replicas could serve as an alternative or guidance to sampling methods, or for fast track exploration of rough parameter spaces in the search of relevant regions. Downscaled versions of real networks could also be applied to perform finite size scaling, which would allow for the determination of critical exponents from single snapshots of their topology. Other possibilities include the development of a new multilevel community detection method [90, 5] that would exploit the mesoscopic information encoded in the different observation scales.

4.8 Additional information

4.8.1 Exact behaviour of the average degree for $r = 2$

To obtain an exact solution for the value of $\langle \kappa' \rangle$, we need to evaluate the $1/\beta$-th moment of the distribution of z',

$$\langle z'^{1/\beta} \rangle = \int_0^\infty z'^{1/\beta} \rho_{z'}(z') \mathrm{d}z', \tag{4.53}$$

from the Laplace transform of the distribution, which we obtained in Section 4.4.2,

$$\hat{\rho}'_z(s) = \left[C s^{\eta-1} \Gamma(1-\eta, s\kappa_0^\beta) \right]^r, \quad C = \frac{\gamma-1}{\beta \kappa_0^{1-\gamma}}. \tag{4.54}$$

Therefore, our first step is relating both expressions. Consider the integral

$$I = C' \int_0^\infty s^\alpha \hat{\rho}^{(n)}_{z'}(s) \mathrm{d}s = C' \int_0^\infty s^\alpha \int_0^\infty (-1)^n z'^n \rho_{z'}(z') e^{-sz'} \mathrm{d}z' \mathrm{d}s. \tag{4.55}$$

Taking into account that for $\alpha > -1$

$$\int_0^\infty s^\alpha e^{-sz'} \mathrm{d}s = z'^{-1-\alpha} \Gamma(1+\alpha), \tag{4.56}$$

we see that

$$I = C'(-1)^n \Gamma(1+\alpha) \int_0^\infty z'^{n-1-\alpha} \rho_{z'}(z') \mathrm{d}z'. \tag{4.57}$$

Now, setting $C' = (-1)^n \Gamma(1+\alpha)^{-1}$ and $n-1-\alpha = 1/\beta$, $I = \langle z'^{1/\beta} \rangle$. However, since $\alpha = n-1-1/\beta > -1$ and $n \in \mathbb{N}$, the smallest n we can choose is $n = 1$, so $\alpha = -1/\beta$. Finally, we can write

$$\langle \kappa' \rangle = -\frac{1}{\Gamma\left(1-\frac{1}{\beta}\right)} \int_0^\infty s^{-1/\beta} \hat{\rho}'_{z'}(s) \mathrm{d}s, \tag{4.58}$$

where $\hat{\rho}_{z'}(s)$ is given in Eq. (4.23).

Our problem has now been reduced to an integral depending on the first derivative of $\hat{\rho}_{z'}(s)$. To compute it, we first use the same change of variable as in Eq. (4.21) to obtain

$$\hat{\rho}'_z(s) = -\int_{\kappa_0^\beta}^{\kappa_c^\beta} z \rho_z(z) e^{-sz} \mathrm{d}z = -C s^{\eta-2} \Gamma(2-\eta, s\kappa_0^\beta). \tag{4.59}$$

Let us now evaluate $\hat{\rho}'_{z'}(s)$,

$$\hat{\rho}'_{z'}(s) = r\hat{\rho}_z^{r-1}(s)\hat{\rho}'_z(s) = 2\hat{\rho}_z(s)\hat{\rho}'_z(s) = -2C^2 s^{2\eta-3}\Gamma(1-\eta, s\kappa_0^\beta)\Gamma(2-\eta, s\kappa_0^\beta) \quad (4.60)$$

and introduce this result into Eq. (4.58),

$$\begin{aligned}\langle\kappa'\rangle &= \frac{2C^2}{\Gamma\left(1-\frac{1}{\beta}\right)}\int_0^\infty s^{2\eta-3-1/\beta}\Gamma(1-\eta, s\kappa_0^\beta)\Gamma(2-\eta, s\kappa_0^\beta)\mathrm{d}s \\ &= \frac{2C^2}{\Gamma\left(1-\frac{1}{\beta}\right)}\frac{1}{\kappa_0^{\beta(2\eta-2-1/\beta)}}\int_0^\infty \omega^{2\eta-3-1/\beta}\Gamma(1-\eta,\omega)\Gamma(2-\eta,\omega)\mathrm{d}\omega \\ &= \kappa_0\frac{2(\gamma-1)^2}{\beta^2\Gamma\left(1-\frac{1}{\beta}\right)}\int_0^\infty \omega^{2\eta-3-1/\beta}\Gamma(1-\eta,\omega)\Gamma(2-\eta,\omega)\mathrm{d}\omega.\end{aligned} \quad (4.61)$$

We thus see that the integral that we need to solve takes the form

$$I(\xi, s_1, s_2) = \int_0^\infty x^\xi \Gamma(s_1, x)\Gamma(s_2, x)\mathrm{d}x, \quad \xi > -1. \quad (4.62)$$

In our case, $\xi = 2\eta - 3 - 1/\beta = (2\gamma-3)/\beta - 1 > -1 \Leftrightarrow \gamma > 3/2$. Integrating by parts,

$$\begin{aligned}I(\xi, s_1, s_2) &= \frac{1}{\xi+1}x^{\xi+1}\Gamma(s_1, x)\Gamma(s_2, x)\Big|_0^\infty \\ &\quad + \frac{1}{\xi+1}\int_0^\infty x^{\xi+1}\left(\Gamma(s_1, x)x^{s_2-1}e^{-x} + \Gamma(s_2, x)x^{s_1-1}e^{-x}\right)\mathrm{d}x \\ &= \frac{1}{\xi+1}\int_0^\infty \left(\Gamma(s_1, x)x^{\xi+s_2}e^{-x} + \Gamma(s_2, x)x^{\xi+s_1}e^{-x}\right)\mathrm{d}x.\end{aligned} \quad (4.63)$$

We can find a recurrence relation for the integrals in the last expression,

$$\begin{aligned}I'(\alpha, s) &= \int_0^\infty \Gamma(s, x)x^\alpha e^{-x}\mathrm{d}x = \frac{1}{\alpha+1}x^{\alpha+1}\Gamma(s, x)\Big|_0^\infty \\ &\quad + \frac{1}{\alpha+1}\int_0^\infty x^{\alpha+1}\left(x^{s-1}e^{-2x} + \Gamma(s, x)e^{-x}\right)\mathrm{d}x \\ &= \frac{1}{\alpha+1}\frac{1}{2^{\alpha+s+1}}\Gamma(\alpha+s+1, 2x)\Big|_0^\infty + \frac{1}{\alpha+1}\int_0^\infty \Gamma(s, x)x^{\alpha+1}e^{-x}\mathrm{d}x \\ &= \frac{1}{\alpha+1}\frac{1}{2^{\alpha+s+1}}\Gamma(\alpha+s+1) + \frac{1}{\alpha+1}I'(\alpha+1, s).\end{aligned} \quad (4.64)$$

Iterating yields

$$I'(\alpha,s)=\sum_{n=1}^{\infty}\frac{1}{\prod_{n'=1}^{n}(\alpha+n')}\frac{1}{2^{\alpha+s+n}}\Gamma(\alpha+s+n)=\sum_{n=1}^{\infty}\frac{\Gamma(\alpha+1)\Gamma(\alpha+s+n)}{\Gamma(\alpha+n+1)2^{\alpha+s+n}}. \tag{4.65}$$

Introducing this result into Eq. (4.63),

$$\begin{aligned}I(\xi,s_1,s_2)&=\frac{1}{\xi+1}\left(I'(\xi+s_2,s_1)+I'(\xi+s_1,s_2)\right)\\&=\frac{1}{\xi+1}\sum_{n=1}^{\infty}\left(\frac{\Gamma(\xi+s_2+1)\Gamma(\xi+s_1+s_2+n)}{\Gamma(\xi+s_2+n+1)2^{\xi+s_1+s_2+n}}+\frac{\Gamma(\xi+s_1+1)\Gamma(\xi+s_1+s_2+n)}{\Gamma(\xi+s_1+n+1)2^{\xi+s_1+s_2+n}}\right)\\&=\frac{1}{(\xi+1)2^{\xi+s_1+s_2}}\sum_{n=1}^{\infty}\frac{\Gamma(\xi+s_1+s_2+n)}{2^n}\left(\frac{\Gamma(\xi+s_2+1)}{\Gamma(\xi+s_2+n+1)}+\frac{\Gamma(\xi+s_1+1)}{\Gamma(\xi+s_1+n+1)}\right).\end{aligned} \tag{4.66}$$

Finally, Eq. (4.61) becomes

$$\begin{aligned}\langle\kappa'\rangle&=\kappa_0\frac{2(\gamma-1)^2}{\beta^2\Gamma\left(1-\frac{1}{\beta}\right)}\int_0^{\infty}\omega^{2\eta-3-1/\beta}\Gamma(1-\eta,\omega)\Gamma(2-\eta,\omega)\mathrm{d}\omega\\&=\kappa_0\frac{2(\gamma-1)^2}{\beta^2\Gamma\left(1-\frac{1}{\beta}\right)}I(2\eta-3-1/\beta,1-\eta,2-\eta)\\&=\frac{2^{1+\frac{1}{\beta}}(\gamma-1)^2\kappa_0}{\beta\Gamma\left(1-\frac{1}{\beta}\right)(2\gamma-3)}\sum_{n=1}^{\infty}\frac{\Gamma\left(n-\frac{1}{\beta}\right)}{2^n}\left(\frac{\Gamma\left(\frac{\gamma-2}{\beta}+1\right)}{\Gamma\left(\frac{\gamma-2}{\beta}+n+1\right)}+\frac{\Gamma\left(\frac{\gamma-2}{\beta}\right)}{\Gamma\left(\frac{\gamma-2}{\beta}+n\right)}\right).\end{aligned} \tag{4.67}$$

Using this result, Eq. (4.18) and $\kappa_0=\langle\kappa\rangle(\gamma-2)/(\gamma-1)$ we can write an expression for the exponent ν (defined in Section 4.4.2 by the expression $\langle k'\rangle=r^{\nu}\langle k\rangle$):

$$\nu=\frac{2}{\ln 2}\ln\left[\frac{2^{1+\frac{1}{\beta}}(\gamma-1)(\gamma-2)}{\beta\Gamma\left(1-\frac{1}{\beta}\right)(2\gamma-3)}\sum_{n=1}^{\infty}\frac{\Gamma\left(n-\frac{1}{\beta}\right)}{2^n}\left(\frac{\Gamma\left(\frac{\gamma-2}{\beta}+1\right)}{\Gamma\left(\frac{\gamma-2}{\beta}+n+1\right)}+\frac{\Gamma\left(\frac{\gamma-2}{\beta}\right)}{\Gamma\left(\frac{\gamma-2}{\beta}+n\right)}\right)\right]-1. \tag{4.68}$$

The above result is shown in Fig. 4.8.

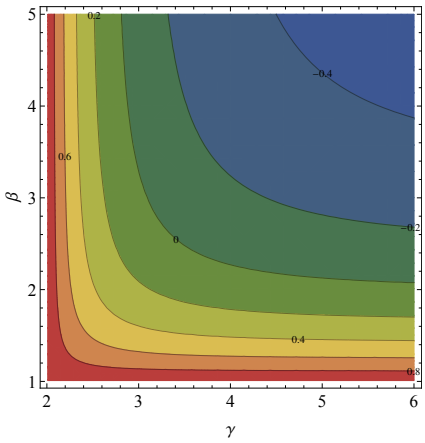


Figure 4.8: **Connectivity phase diagram.** Exact value of ν as a function of β and γ according to Eq. (4.68). The exact solution agrees with the solution in the power-law approximation used in Section 4.4.2 for large values of β or γ.

4.8.2 Renormalization in higher dimensions

In this subsection, we detail the RG transformations in the general case of a D-dimensional space, as well as the effect of the distortion of one-dimensional embeddings on the transformation.

As shown in Chapter 2, the connection probability Eq. (4.5) in D dimensions reads

$$p_{ij} = \frac{1}{1+\left(\frac{R\Delta\theta_{ij}}{(\mu\kappa_i\kappa_j)^{1/D}}\right)^{\beta}}, \tag{4.69}$$

where $\Delta\theta_{ij}$ is the angular separation between the two positioning vectors associated to the nodes embedded in the D-dimensional sphere of radius R. Consequently, Eq. (4.11) becomes

$$\left(\frac{R\Delta\theta}{\mu^{1/D}}\right)^{\beta} \frac{1}{\sum_{e=1}^{r^2}(\kappa_m\kappa_n)_e^{\beta/D}} = \left(\frac{R'\Delta\theta'_{ij}}{\left(\mu'\kappa'_i\kappa'_j\right)^{1/D}}\right)^{\beta'}. \tag{4.70}$$

Following the same reasoning as in the one-dimensional case, $\beta' = \beta$. Furthermore, we should also require the density of nodes to be preserved, which means that $R' = R/r^{1/D}$. Hence, setting $\mu' = \mu/r$ yields

$$\frac{R'\Delta\theta'}{\mu'^{1/D}} = \frac{R/r^{1/D}\Delta\theta}{(\mu/r)^{1/D}} = \frac{R\Delta\theta}{\mu^{1/D}}, \tag{4.71}$$

so the transformation further requires

$$\left(\kappa'_i\kappa'_j\right)^{\beta/D} = \sum_{e=1}^{r^2} (\kappa_m\kappa_n)_e^{\beta/D}, \tag{4.72}$$

which is fulfilled by Eqs. (4.13) and (4.15) with the change $\beta \to \beta/D$. Thus, the analysis of the RG flow of the average degree can be easily generalized by replacing β with β/D in all k'-related quantities.

In the γ-dominated region (now $\beta > D(\gamma-1)$), $\langle\kappa'\rangle$ does not depend on β, so Eq. (4.30) still holds and the average degree of the renormalized network is

$$\langle k'\rangle = C_0\frac{\mu}{r}r^{\frac{2}{\gamma-1}}\langle\kappa\rangle^2 = r^{\frac{2}{\gamma-1}-1}\langle k\rangle \to \begin{cases} \infty & \gamma<3 \\ cte. & \gamma=3 \\ 0 & \gamma>3 \end{cases} \tag{4.73}$$

In the β-dominated region, on the other hand, $\langle\kappa'\rangle$ becomes $\langle\kappa'\rangle = \langle\kappa\rangle r^{D/\beta}$, so

$$\langle k'\rangle = C_0\frac{\mu}{r}r^{2D/\beta}\langle\kappa\rangle^2 = r^{\frac{2D}{\beta}-1}\langle k\rangle \to \begin{cases} \infty & \beta<2D \\ cte. & \beta=2D \\ 0 & \beta>2D \end{cases} \tag{4.74}$$

Both exponents are equal at $\eta = 2$, that is, at $\beta = D(\gamma-1)$. Finally, the non-symmetric phase III, in which $\eta > 3$, corresponds to $\beta < \frac{1}{2}D(\gamma-1)$.

These results show that the dimension of similarity space does not alter the phase diagram qualitatively, but only quantitatively. However, it is interesting to mention that, as the dimension D increases, the small-world phase occupies a region with higher β networks. Moreover, the non-symmetric phase can be found for increasingly heterogeneous networks, that is, for smaller values of γ, at any given value of β.

In Fig. 4.9, we show the behaviour of the topological properties of networks of different dimensions where the RGN is performed using the embedding distances only. Interestingly, all four networks are almost indistinguishable, mainly due to the fact that the one-dimensional embedding is able to recover most of the original geometric information, as discussed in Chapter 2.

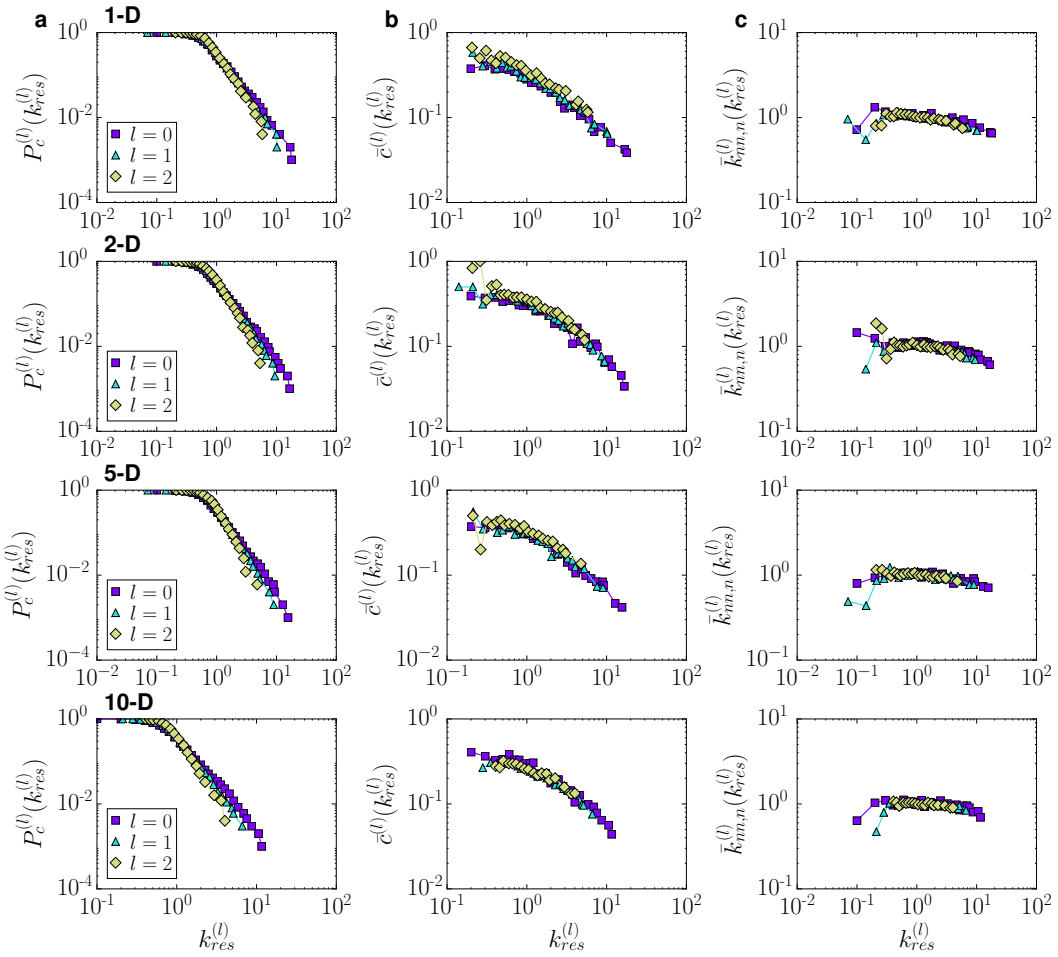


Figure 4.9: **RGN flow of topological properties.** Each row shows the RGN flows of the topological features of the networks of different dimensions from Fig. 2.5. For the transformations, only information from the one-dimensional embedding was used. **a**, Complementary cumulative distribution of rescaled degrees $k_{res}^{(l)} = k^{(l)}/\langle k^{(l)}\rangle$. **b**, Local clustering averaged over rescaled-degree classes. **c**, Normalized average nearest neighbour degree $\bar{k}_{nn,n}(k_{res}^{(l)}) = \bar{k}_{nn}(k_{res}^{(l)})\langle k^{(l)}\rangle/\langle (k^{(l)})^2\rangle$.

CHAPTER 5

The geometric nature of weights in real complex networks

This thesis has so far focused on the binary structure of networks, that is, in which all attributes of links are disregarded. As we have seen throughout, our results clearly support the existence of hidden metric spaces underlying real networks, so it is natural to question whether the weighted structure of real systems can also be explained by assuming a geometric origin of the intensities of links. In this chapter, which corresponds to Ref. [4], we present empirical evidence supporting such geometric origin of weights as well as a geometric model capable of reproducing real weighted networks accurately.

5.1 Beyond binary networks

Most of the complexity of networks is encoded into the intricate topology of the interactions among their components and into the layout of the intensities associated to such interactions (i.e., the weights). In real weighted networks, weights are coupled to the binary topology in a non-trivial way. This is manifested, for instance, in a non-linear relation between the strength of a node s (the sum of the total weight attached to it) and its degree k of the form $s \sim k^{\eta}$ [20, 136, 22]. Moreover, weights play a central role in their structural organization, function, and dynamics [20]. For instance, the quantification of the rich-club effect in real weighted networks, in sharp contrast to results in unweighted representations, unveils the formation of alliances in multipolarized environments or the lack of cohesion even in the presence of rich-club ordering [148]. Similarly, the propagation of emergent diseases in the international airports network is intimately linked to the number of passengers flying from one airport to the other [39]. A shift towards a paradigm of weighted networks is therefore in order to fully understand the behaviour and evolution of complex networks. However, advances in this area have been limited by the extreme heterogeneity and fluctuations that typically characterize the distribution of weights. In terms of the topic of this thesis, the relation between the layout of weights and the geometry underlying the network is not evident. The reason being that, even if the existence of a link depends on the metric distance between the nodes, there is no reason, a priori,

to expect that the same distance will affect its weight. For instance, in the airports network, the decision to set up a link between two cities depends on the airline companies operating at the two airports, a process affected by geopolitic and economic costs and by the expected flow of passengers that would eventually compensate such costs. However, once the connection is established, its weight is determined by the aggregation of the individual decisions of people using it, a process that may be affected by a different cost function.

In the remaining of the chapter, we present empirical evidence of the metric nature of weights in real biological, economic and transportation networks, which suggests that the hidden geometry paradigm can be extended to weighted complex networks. We then propose a general class of weighted networks embedded in hidden metric spaces that accurately reproduces many properties observed in real weighted networks. This model has the critical ability to fix the degree-strength distribution independently of the coupling of the topology and weighted organization with the metric space. It is therefore possible to isolate, and thus directly study, the effect of the coupling between the metric space and the weighted organization of real weighted networks. In fact, our results unveil that, in some systems, the coupling between weights and metric distances is weak, which in turn suggests that the formation of connections and the assignment of their magnitude might be ruled by different processes. Our empirical findings, combined with our new class of models, open the path towards the use of information encoded in the weights of the links to find more accurate embeddings of real networks, which in turn will improve the detection of communities, the prediction of missing links and provide estimates for the weights of such missing links.

5.2 Interplay between weights and triangles in real networks

In this section, we provide empirical evidence for the geometric origin of weights in real complex networks. The datasets that we have used are:

- The World Trade Web, describing significant trade exchanges between countries in 2013, as reconstructed in Chapter 3. The corresponding weights are trade volumes between pairs in US Million $ [2].

- The international network of global cargo ship movements, which consists of the number of shipping journeys between pairs of major commercial ports in the world in 2007 [89].

- The commodities network corresponding to the flows of the goods and services in US Million $ between industrial sectors in the United States in 2007 [81].
- The airports network, in which weights represent the number of passengers that flew between pairs of airports in the United States in 2013 [42].
- The commuting network, reflecting the daily flow of commuters between counties in the United States in 2000 [81].
- The metabolic network of the bacteria *Escherichia Coli (E. Coli)* K-12 MG1655, in which weights consist of the number of different metabolic reactions in which two metabolites participate [127, 155].
- The human brain network, representing the density of anatomical connections between subregions of the human brain as detected via diffusion tensor imaging [15].

Name	**Type**	**Nodes**	**Weights**	γ	η	$\langle k \rangle$	N
World Trade	Economic	Countries	US Million $	2.42	1.6	5.8	189
Cargo ships	Transportation	Ports	Shipping journeys	2.03	1.05	10.4	834
US Commodities	Economic	Economic sectors	US Million $	2.46	1.22	5.8	376
US Airports	Transportation	Airports	Passengers	1.76	1.7	8.6	884
US Commute	Transportation	Counties	People	4.31	2.02	4.3	3109
E. Coli	Biological	Metabolites	Common reactions	2.52	1.1	6.6	1100
Human brain	Biological	Brain regions	Connection density	7.14	0.86	24.1	501

Table 5.1: **Overview of the considered real-world network data.** The values of γ, $\langle k \rangle$ and N correspond to the familiar power-law degree exponent, average degree and system size, respectively, whereas η is the inferred exponent of the non-linear relation $\bar{s}(k) \sim k^{\eta}$ introduced in Chapter 1.

Their main topological and weighted features are summarized in Table 5.1. Except for the metabolic and human brain networks, all networks were filtered using the disparity filter explained in Chapter 3 [153] to preserve the most statistically significant connections. Many real weighted networks are generated from data by using a very broad definition of what constitutes a significant connection. This results in networks with huge average degrees and in which many links are noisy and weakly related to the overall functionality of the network. For instance, the US airports network contains links due to private flights (of the order of 10 passengers per year), which obviously follow different patterns of connection than the regular commercial airlines. Another interesting example is the World Trade Web, in which many trade interactions amount for less than one million dollars and are extremely volatile, appearing and disappearing from

year to year. Indeed, as we showed in Chapter 3, removing these noisy connections yields a significantly more congruent topology with real economic factors, such as the gross domestic product (GDP).

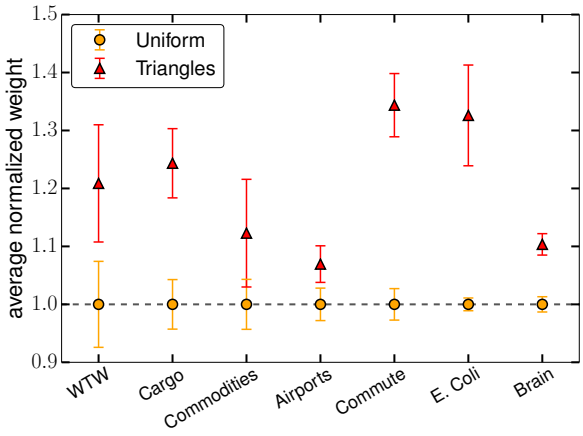


Figure 5.1: **Geometric nature of weights.** Comparison of the average normalized weights in the network (yellow circles) with the one measured by sampling links over triangles (red triangles) for the empirical datasets analyzed. The error bars correspond to an estimate of the standard deviation of the average value due to the finite size of the samples and are computed as $\sqrt{\mathrm{Var}(\omega^{\mathrm{norm}})/L}$, where $\mathrm{Var}(\omega^{\mathrm{norm}})$ is the variance of the normalized weights sampled uniformly or via the triangles, and L is the number of links.

As we discussed in Chapter 1, clustering, as a reflection of the triangle inequality, is the key topological property coupling the bare topology of a complex system and its effective underlying metric space [152]: if nodes A and B are close, and nodes B and C are also close, we expect nodes A and C to be close as well. Triangles are therefore more likely to exist between nodes that are nearby. Consequently, we expect that if the weights of connections depend on the distance between the connected nodes in the underlying metric space, they should be quantitatively different depending on the clustering properties of the connections. However, weights and clustering are known to be strongly influenced by the degrees of endpoint nodes [20, 136, 129], which prevents from a direct detection of the metric properties of weights due to the typical heterogeneity in the degrees of nodes in real networks. Thus, to compare links on equal footing, we define the normalized weight of an existing link connecting nodes i and j as $\omega_{ij}^{\mathrm{norm}} = \omega_{ij}/\bar{\omega}(k_i k_j)$, where $\bar{\omega}(k_i k_j)$ is the average weight of links as a function of the product of degrees of their endpoint nodes. By doing so, we decouple the weights and the topology, leaving the normalized weights seemingly randomly fluctuating around 1 (see Fig. 5.1).

Figure 5.1 also shows, however, that the fluctuations of the normalized weights are not uniform, as links involved in triangles tend to have larger normalized weights than the average link. Indeed, in some cases the difference can reach more than 30%. Sampling links over triangles is equivalent to sampling links proportionally to their multiplicity m (i.e., the number of triangles in which a link participates). Therefore, the results in Fig. 5.1 indicate that ω^{norm} and m are positively correlated variables, as corroborated by their Pearson correlation coefficient (see Secion 5.4.2). In Ref. [129], the authors also found local correlations between the multiplicity of links and the weights for different real networks. Note, however, that in that study weights were not normalized to discount the effects of the heterogeneity in the degrees of the nodes, so the detected weighted organization can not be taken as a signature of underlying metric properties.

Since triangles are a reflection of the triangle inequality in the underlying metric space, we expect nodes forming triangles to be close to one another. Thus, the higher average normalized weight observed on triangles strongly suggests a metric nature of weights, which is not a trivial consequence of the relation between weights and topology. This leads us to formulate the hypothesis that the same underlying metric space ruling the network topology—inducing the existence of strong clustering as a reflection of the triangle inequality in the underlying geometry—is also inducing the observed correlation between ω^{norm} and m. To prove this, we develop a realistic model of geometric weighted random networks, which allows us to estimate the coupling between weights and geometry in real networks.

5.3 A geometric model of weighted networks

As we stated in Chapter 1, seveal growing network models [27, 21, 174, 99, 13, 176, 165, 103] and the maximum-entropy class of models [108, 70, 71, 144, 145] have been proposed as generators of weighted networks. However, none of them is general enough to reproduce simultaneously the topology and weighted structure of real weighted complex networks. We introduce a new model based on a class of random networks with hidden variables embedded in a metric space [152, 33] that overcomes these limitations. In this model, the generator of the topology is the general $\mathbb{S}^D$ model, which we briefly review here for completeness; a detailed description was provided in Chapter 2. First, N nodes are uniformly distributed with constant density, which we fix to one, in a D-dimensional homogeneous and isotropic space, a D-sphere, and are assigned a hidden variable κ according to the probability density function $\rho_\kappa(\kappa)$. Two nodes i and j with hidden variables κ_i and κ_j separated by a metric distance d_{ij}

are connected with a probability

$$p\left(\kappa_i,\kappa_j,d_{ij}\right)=p\left(\chi\right),\,\chi=\frac{d_{ij}}{\left(\mu\kappa_i\kappa_j\right)^{1/D}}\,, \tag{5.1}$$

where $\mu>0$ is a free parameter fixing the average degree and $p\left(\chi\right)$ is an arbitrary positive function taking values within the interval $[0,1]$. The free parameter μ can be chosen such that $\bar{k}_i(\kappa_i)=\kappa_i$. Hence, κ correspond to the expected degrees of nodes, so the degree distribution can be specified through the probability distribution function $\rho_\kappa(\kappa)$ regardless of the specific form of $p\left(\chi\right)$. The freedom in the choice of this function allows us to tune the level of coupling between the topology of the networks and the metric space, which in turn allows us to control many properties such as the clustering coefficient and the navigability [152, 36].

To generate weighted networks, a second hidden variable σ is associated to each node. This new hidden variable can be correlated with κ so, hereafter, we assume that the pair of hidden variables (κ,σ) associated with the same node are drawn from the joint distribution $\rho(\kappa,\sigma)$. The weight of an existing link between two nodes with hidden variables κ_i, σ_i, κ_j and σ_j, respectively, and at a metric distance d_{ij} is assigned according to the distribution

$$\rho\left(\omega_{ij}|a_{ij}=1,\kappa_i,\sigma_i,\kappa_j,\sigma_j,d_{ij}\right)=\frac{1}{\bar{\omega}_{ij}}f\left(\frac{\omega_{ij}}{\bar{\omega}_{ij}}\right)\,, \tag{5.2}$$

where f is an arbitrary density function. The form of Eq. (5.2) allows us to control the expected weight between nodes i and j, as it is guaranteed to be proportional to $\bar{\omega}_{ij}$. Furthermore, since $\rho\left(\omega_{ij}|a_{ij}=1,\kappa_i,\sigma_i,\kappa_j,\sigma_j,d_{ij}\right)\mathrm{d}\omega_{ij}=f\left(\varepsilon_{ij}\right)\mathrm{d}\varepsilon_{ij}$ with $\varepsilon_{ij}\equiv\omega_{ij}/\bar{\omega}_{ij}$, we can write $\omega_{ij}=\varepsilon_{ij}\bar{\omega}_{ij}$, where ε_{ij} is now a positive random variable drawn from the probability density function $f\left(\varepsilon_{ij}\right)$. Now, the keystone of our model comes from the precise expected weight as a function of the hidden variables, that is, of $\bar{\omega}_{ij}$, according to which the weight of an existing link $\omega_{ij}=\varepsilon_{ij}\bar{\omega}_{ij}$ reads

$$\omega_{ij}=\varepsilon_{ij}\frac{\nu\sigma_i\sigma_j}{(\kappa_i\kappa_j)^{1-\alpha/D}d_{ij}^{\alpha}}, \tag{5.3}$$

with $\nu>0$ and $0\leq\alpha<D$. Notice that α dictates a trade-off between the contribution of degrees and geometry to weights. If $\alpha=0$ weights are independent of the underlying metric space and maximally dependent on degrees, while $\alpha=D$ implies that weights are maximally coupled to the underlying metric space with no direct contribution of the degrees. As we will show in the next subsection, the form of Eq. (5.3) ensures that $\bar{s}(\sigma)\propto\sigma$. The free parameter ν can then always be chosen such that $\bar{s}(\sigma)=\sigma$. The new hidden variable σ can therefore be

interpreted as the expected strength of a node, and the joint probability distribution function $\rho(\kappa,\sigma) = \rho_\kappa(\kappa)\rho(\sigma|\kappa)$ controls the correlation between degrees and strengths in the network. Indeed, the average strength of nodes with a given degree, $\bar{s}(k)$, relates to the first moment of the conditional distribution $\rho(\sigma|\kappa)$, $\bar{\sigma}(\kappa)$, so that the strength as a function of degree scales as $\bar{s}(k) \sim \bar{\sigma}(\kappa)$ for large k.

The relations $\bar{k}(\kappa) = \kappa$ and $\bar{s}(\sigma) = \sigma$—and consequently the relation between $\rho(\kappa,\sigma)$ and the degree-strength distribution—hold independently of the specific form of the connection probability $p(\chi)$ and of the noise distribution $f(\varepsilon)$. Besides conferring great versatility to our model, this conveys a degree of control over the weight distribution which is independent of the specification of degrees and strengths and, more importantly, opens the possibility to measure the metric properties of complex weighted networks.

5.3.1 General properties of the model

We now present an alternative derivation of the basic properties of the weighted structure of the model, different from the original one in Ref. [4]. We take a similar approach to that of Chapter 2, where we started our study of the $\mathbb{S}^D$ model by noting that, given a node i, every potential link to the $N-1$ other nodes can be regarded as and independent stochastic event. The assignment of weights in this model is not different in this respect. Given a node i with hidden degree and strength κ and σ, the weight between node i and a randomly chosen node will follow a distribution, which we will denote as $\Phi(\omega|\kappa,\sigma)$[1], independent of the specific position of i due to the symmetries of the space. Note that this distribution accounts for the fact that the weight is $\omega = 0$ if there is no link between them. The weighted structure of node i must be given by the sum of $N-1$ identically and independently random variables distributed according to $\Phi(\omega|\kappa,\sigma)$. This distribution is given by

$$\begin{aligned}\Phi(\omega|\kappa,\sigma) = \int \left[\delta(\omega)\left(1-p\left(\kappa,\kappa',d\right)\right) + \frac{1}{\bar{\omega}} f\left(\frac{\omega}{\bar{\omega}}\right) p\left(\kappa,\kappa',d\right)\right] \\ \times \rho\left(\kappa',\sigma'\right)\rho_d(d)\,\mathrm{d}\kappa'\mathrm{d}\sigma'\mathrm{d}d,\end{aligned} \tag{5.4}$$

where $\bar{\omega} = \frac{\nu\sigma\sigma'}{(\kappa\kappa')^{1-\alpha/D} d^\alpha}$, as given in Eq. (5.3). A thorough derivation of this result can be found in Section 5.6.

[1] We drop the indices i and j throughout this section for notation clarity, so that node-specific unprimed variables will refer to node i, whereas primed ones will correspond to the randomly chosen node.

5.3.1.1 Expected strength of a node

The expected strength of a node is given by the sum of $N-1$ weights, all of them independent and equally distributed in accordance with Eq. (5.4). Hence, we can write $\bar{s}(\kappa,\sigma) = (N-1)\,\bar{\omega}(\kappa,\sigma)$, with

$$\begin{aligned}\bar{\omega}(\kappa,\sigma) &= \int \omega\Phi(\omega|\kappa,\sigma)\,\mathrm{d}\omega \\ &= \int \frac{\omega}{\bar{\omega}} f\left(\frac{\omega}{\bar{\omega}}\right) p(\kappa,\kappa',d)\,\rho(\kappa',\sigma')\,\rho_d(d)\,\mathrm{d}\kappa'\mathrm{d}\sigma'\mathrm{d}d\mathrm{d}\omega.\end{aligned} \tag{5.5}$$

To proceed with the integration, we change variables to ε and χ, and we use the Euclidean approximation, as in Chapter 2, in which we assume that the only contribution to the integral comes from the region of small distances, so the curvature of the D-sphere can be neglected. Hence, we can write

$$\rho_d(d) \approx \frac{1}{A_D}\frac{2\pi^{D/2}}{\Gamma\left(\frac{D}{2}\right)} d^{D-1} \tag{5.6}$$

with $A_D = N$ due to unitary density, and extend the integration of distances to infinity. With these considerations, Eq. (5.5) becomes

$$\begin{aligned}\bar{\omega}(\kappa,\sigma) &= \frac{2\pi^{D/2}\nu\sigma\mu^{1-\alpha/D}}{A_D\Gamma\left(\frac{D}{2}\right)} \int_0^\infty \varepsilon f(\varepsilon)\,\mathrm{d}\varepsilon \int_0^\infty \chi^{D-1-\alpha} p(\chi)\,\mathrm{d}\chi \int \sigma'\rho(\kappa',\sigma')\,\mathrm{d}k'\mathrm{d}\sigma' \\ &= \frac{2\pi^{D/2}\nu\sigma\mu^{1-\alpha/D} I_1 I_2 \langle\sigma\rangle}{A_D\Gamma\left(\frac{D}{2}\right)},\end{aligned} \tag{5.7}$$

in which we have defined $I_1 = \int_0^\infty \varepsilon f(\varepsilon)\,\mathrm{d}\varepsilon$ and $I_2 = \int_0^\infty \chi^{D-1-\alpha} p(\chi)\,\mathrm{d}\chi$. Since $(N-1)/N \approx 1$ for large N, the strength of a node is given by

$$\bar{s}(\kappa,\sigma) = \frac{2\pi^{D/2}\nu\mu^{1-\alpha/D} I_1 I_2 \langle\sigma\rangle}{\Gamma\left(\frac{D}{2}\right)}\sigma. \tag{5.8}$$

This result justifies our particular choice of $\bar{\omega}$ in Eq. (5.3); for arbitrary f and p functions, the model yields a direct proportionality between the strength of a node and σ, which means that $\bar{s}(\kappa,\sigma) = \bar{s}(\sigma)$. Moreover, we can always choose ν such that $\bar{s}(\sigma) = \sigma$ and $\langle s\rangle = \langle\sigma\rangle$. This is achieved with

$$\nu = \frac{\Gamma\left(\frac{D}{2}\right)}{2\pi^{D/2}\mu^{1-\alpha/D} I_1 I_2 \langle\sigma\rangle}. \tag{5.9}$$

5.3.1.2 Fluctuations of the strength of a node

We can study the fluctuations of the strength using a similar procedure. Appealing to the independence between potential links, $\mathrm{Var}(s|\kappa,\sigma) =$

$(N-1)\,\mathrm{Var}(\omega|\kappa,\sigma)$, where $\mathrm{Var}(\omega|\kappa,\sigma)$ is the variance of $\Phi(\omega|\kappa,\sigma)$,

$$\mathrm{Var}(\omega|\kappa,\sigma) = \int \omega^2 \Phi(\omega|\kappa,\sigma)\,\mathrm{d}\omega - \left(\int \omega\Phi(\omega|\kappa,\sigma)\,\mathrm{d}\omega\right)^2. \tag{5.10}$$

The rightmost integral is given in Eq. (5.7), while the second moment is

$$\begin{aligned}\bar{\omega}^2(\kappa,\sigma) &= \int \frac{\omega^2}{\bar{\omega}} f\left(\frac{\omega}{\bar{\omega}}\right) p\left(\kappa,\kappa',d\right)\rho\left(\kappa',\sigma'\right)\rho_d(d)\,\mathrm{d}\kappa'\mathrm{d}\sigma'\mathrm{d}d\mathrm{d}\omega \\ &= \frac{2\pi^{D/2}\nu^2\sigma^2\mu^{1-2\alpha/D}}{A_D\Gamma\left(\frac{D}{2}\right)\kappa}\int_0^\infty \varepsilon^2 f(\varepsilon)\,\mathrm{d}\varepsilon \int_0^\infty \chi^{D-1-2\alpha} p\left(\chi\right)\mathrm{d}\chi \int \frac{\sigma'^2}{\kappa'}\rho\left(\kappa',\sigma'\right)\mathrm{d}k'\mathrm{d}\sigma' \\ &= \frac{2\pi^{D/2}\nu^2\mu^{1-2\alpha/D} I_3\left\langle\sigma^2/\kappa\right\rangle I_4}{A_D\Gamma\left(\frac{D}{2}\right)}\frac{\sigma^2}{\kappa},\end{aligned} \tag{5.11}$$

where we have denoted $I_3 = \int_0^\infty \varepsilon^2 f(\varepsilon)\,\mathrm{d}\varepsilon$ and $I_4 = \int_0^\infty \chi^{D-1-2\alpha} p\left(\chi\right)\mathrm{d}\chi$. Inserting this result into Eq. (5.10), the variance of the strength reads

$$\begin{aligned}\mathrm{Var}(s|\kappa,\sigma) &= \frac{2\pi^{D/2}\nu^2\mu^{1-2\alpha/D} I_3\left\langle\sigma^2/\kappa\right\rangle I_4}{\Gamma\left(\frac{D}{2}\right)}\frac{\sigma^2}{\kappa} - \frac{1}{A_D}\left(\frac{2\pi^{D/2}\nu\sigma\mu^{1-\alpha/D} I_1 I_2\langle\sigma\rangle}{\Gamma\left(\frac{D}{2}\right)}\right)^2 \\ &\xrightarrow{N\to\infty} \frac{2\pi^{D/2}\nu^2\mu^{1-2\alpha/D} I_3\left\langle\sigma^2/\kappa\right\rangle I_4}{\Gamma\left(\frac{D}{2}\right)}\frac{\sigma^2}{\kappa}.\end{aligned} \tag{5.12}$$

Thus, we immediately see that the relative fluctuations of the strength vanish as $\sqrt{\mathrm{Var}(s|\kappa,\sigma)}/\bar{s}(\kappa,\sigma) \sim 1/\sqrt{\kappa}$.

5.3.1.3 Strength of a node of a given degree

We can also address the behaviour of the strength of nodes of a given degree, $\bar{s}(k)$. For that purpose, however, we need the distribution of weights only over connections that have been established, which we will denote as $\tilde{\Phi}(\omega|\kappa,\sigma)$. This distribution, derived in Section 5.6, is given by

$$\tilde{\Phi}(\omega|\kappa,\sigma) = \frac{1}{\tilde{p}(\kappa)}\int \frac{1}{\bar{\omega}} f\left(\frac{\omega}{\bar{\omega}}\right) p\left(\kappa,\kappa',d\right)\rho\left(\kappa',\sigma'\right)\rho_d(d)\,\mathrm{d}\kappa'\mathrm{d}\sigma'\mathrm{d}d, \tag{5.13}$$

where

$$\tilde{p}(\kappa) = \int p\left(\kappa,\kappa',d\right)\rho_\kappa\left(\kappa'\right)\rho_d(d)\,\mathrm{d}\kappa'\mathrm{d}d \tag{5.14}$$

is the probability for node i to to connect to a randomly chosen node, which we calculated in Chapter 2. Its mean can be computed as

$$\begin{aligned}\tilde{\omega}(\kappa,\sigma) &= \int \omega\tilde{\Phi}(\omega|\kappa,\sigma)\,\mathrm{d}\omega \\ &= \frac{1}{\tilde{p}(\kappa)}\int \omega\Phi(\omega|\kappa,\sigma)\,\mathrm{d}\omega \\ &= \frac{\bar{\omega}(\kappa,\sigma)}{\tilde{p}(\kappa)}.\end{aligned} \tag{5.15}$$

This last result can be given a simple expression using Eqs. (5.7) and (5.9), and recalling that $\bar{k}(\kappa) = (N-1)\,\tilde{p}(\kappa) = \kappa$,

$$\tilde{\omega}(\kappa,\sigma) = \frac{\sigma}{A_D\tilde{p}(\kappa)} = \frac{\sigma}{\kappa}. \tag{5.16}$$

Finally, the expected strength of a node of degree k is given by

$$\begin{aligned}\bar{s}(k) &= \int s\rho(s|k)\,\mathrm{d}s = \frac{1}{P(k)}\int s\rho(s,k)\,\mathrm{d}s = \frac{1}{P(k)}\int s\rho(s,k,\sigma,\kappa)\,\mathrm{d}s\mathrm{d}\sigma\mathrm{d}\kappa \\ &= \frac{1}{P(k)}\int\left(\int s\rho(s|k,\sigma,\kappa)\,\mathrm{d}s\right)\rho(k|\sigma,\kappa)\,\rho(\sigma|\kappa)\,\rho_\kappa(\kappa)\,\mathrm{d}\sigma\mathrm{d}\kappa\end{aligned} \tag{5.17}$$

The integral inside the brackets is the expected strength of a node with degree k and hidden variables σ and κ. Since we know that each of the k links emerging from the node has expected weight $\tilde{\omega}(\kappa,\sigma)$ given in Eq. (5.16), the expected value of their sum is simply $k\tilde{\omega}(\kappa,\sigma)$. Furthermore, the term $\rho(k|\sigma,\kappa) = \rho(k|\kappa)$ was analyzed in Chapter 2, where we concluded that $\rho(k|\kappa) = e^{-\kappa}\frac{\kappa^k}{k!}$. Equation (5.17) hence becomes

$$\begin{aligned}\bar{s}(k) &= \frac{1}{P(k)}\int k\frac{\sigma}{\kappa}e^{-\kappa}\frac{\kappa^k}{k!}\rho(\sigma|\kappa)\,\rho_\kappa(\kappa)\,\mathrm{d}\sigma\mathrm{d}\kappa \\ &= \frac{1}{P(k)}\int e^{-\kappa}\frac{\kappa^{k-1}}{(k-1)!}\bar{\sigma}(\kappa)\,\rho_\kappa(\kappa)\,\mathrm{d}\kappa,\end{aligned} \tag{5.18}$$

thus further clarifying how the joint probability distribution function $\rho(\sigma,\kappa)$ controls the correlation between the degree and the strength of nodes. In fact, assuming no correlation, i.e., $\rho(\sigma,\kappa) = \rho_\kappa(\kappa)\,\rho_\sigma(\sigma)$, yields $\bar{s}(k) = \frac{P(k-1)}{P(k)}\langle\sigma\rangle$, which, for large k, is independent of the degree, as expected. If there are correlations, on the other hand, since the integrand becomes highly peaked around $k-1$, $\bar{s}(\kappa)$ has the same asymptotic behaviour as $\bar{\sigma}(\kappa)$.

5.3.2 The one-dimensional version

To use the model in the context of real weighted networks, we choose the circle as similarity space, so links are assigned in accordance with the $\mathbb{S}^1$ model with the familiar form of the connection probability

$$p(\chi) = \frac{1}{1+\chi^\beta} \quad \text{with} \quad \chi = \frac{d_{ij}}{\mu\kappa_i\kappa_j}, \tag{5.19}$$

where $d_{ij} = R\min\left(\left|\theta_i-\theta_j\right|, 2\pi-\left|\theta_i-\theta_j\right|\right)$, $R = N/(2\pi)$ and $\beta > 1$ tunes the clustering and quantifies the level of coupling between the network topology and the metric space. To obtain a scale-free degree distribution, hidden variables κ are distributed according to $\rho_\kappa(\kappa) \propto \kappa^{-\gamma}$ with $\kappa_0 < \kappa < \kappa_c$ and $\gamma > 1$.

Weights are assigned on top of the topology generated by the model. The noise distribution in Eq. (5.2) is chosen to be a gamma distribution of average $\langle\varepsilon\rangle = 1$, that is,

$$f(\varepsilon) = \frac{\lambda^\lambda}{\Gamma(\lambda)}\varepsilon^{\lambda-1}e^{-\lambda\varepsilon} \quad \text{with} \quad \left\langle\varepsilon^2\right\rangle = 1+\frac{1}{\lambda}. \tag{5.20}$$

This particular choice allows us to interpolate with a single parameter between a zero-noise limit when $\lambda \gg 1$, exponential noise when $\lambda = 1$, and strongly heterogeneous noise when $\lambda \ll 1$. Finally, to control the correlation between strength and degree and, therefore, to tune the strength distribution, we assume a deterministic relation between hidden variables σ and κ of the form $\sigma = a\kappa^\eta$, as observed in real complex networks [20]. From Eq. (5.18), we thus expect

$$\bar{s}(k) = \frac{ak\Gamma(k-\gamma+\eta,\kappa_0,\kappa_c)}{\Gamma(k-\gamma+1,\kappa_0,\kappa_c)} \sim ak^\eta\,. \tag{5.21}$$

Notice that the relation between average strength and degree in the previous expression is totally independent of the underlying metric space. Furthermore, given that the fluctuations of the degrees and the strengths vanish for large κ, we can assume the relation to become increasingly deterministic. Therefore, the strength distribution scales as $P(s) \sim s^{-\xi}$ for $s \gg 1$ with $\xi = (\gamma+\eta-1)/\eta$ since, if $s = ak^\eta$,

$$P(s)\mathrm{d}s = P(k)\mathrm{d}k \sim k^{-\gamma}\mathrm{d}k \propto s^{-\gamma/\eta+1/\eta-1}\mathrm{d}s = s^{-\frac{\gamma+\eta-1}{\eta}}\mathrm{d}s. \tag{5.22}$$

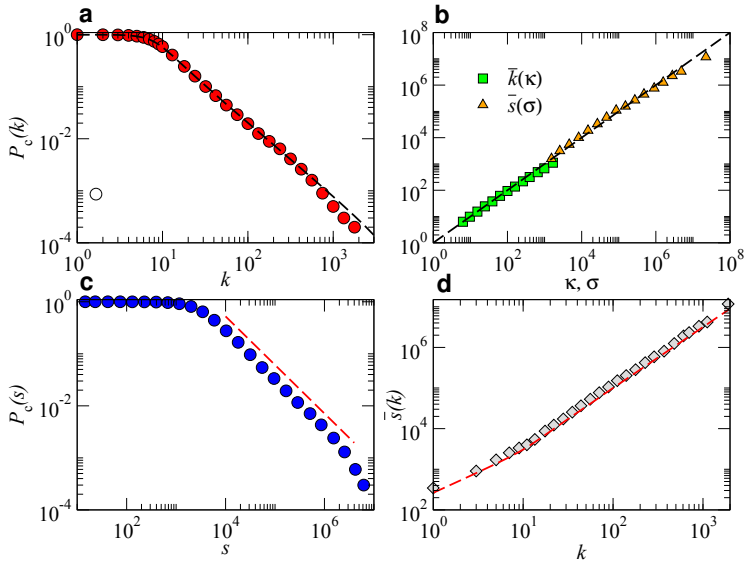


Figure 5.2: **Validation of the theoretical calculations.** The theoretical calculations, represented by the dashed lines, are obtained in Section 5.3.1, and the numerical simulations correspond to a single network of size $N = 10^4$ and parameters $\alpha = 0.4$, $\beta = 1.5$, $\gamma = 2.4$, $\eta = 1.5$, $\langle k \rangle = 20$, $\kappa_c = \kappa_0 N^{1/(\gamma-1)}$, $a = 100$, and noise $\langle \varepsilon^2 \rangle = 1.5$. **a**, Complementary cumulative degree distribution compared with the prediction from the $\mathbb{S}^D$ model. **b**, Average degree and average strength of nodes as a function of their hidden variables κ and σ. The deviation for high κ and σ is due to the finite size of the network and disappears as $N \to \infty$. **c**, Complementary cumulative strength distribution. The dashed line indicates a scaling $\propto s^{-\xi}$ with $\xi = (\gamma + \eta - 1)/\eta \simeq 1.93$, as expected. **d**, Average strength as a function of degree. The dashed line shows the prediction of Eq. (5.21).

Fig. 5.2 shows the basic topological and weighted properties of a network generated using Eqs. (5.19)–(5.21) and compares them to the theoretical predictions presented in this section. Apart from some expected fluctuations due to finite size, the agreement between the two is excellent.

5.3.3 The effect of the underlying geometry

Geometry has a strong effect on the strength and weight distributions, which depend on the coupling parameter α. In fact, as shown in Section 5.3.1, the second moment of the strength distribution $\langle s^2 \rangle$ is proportional to the integral $\int_0^\infty \chi^{D-2\alpha-1} p(\chi) \, \mathrm{d}\chi$, which increases with α. The origin of these fluctuations is rooted in the strong constraints that geometry imposes on the weights of individual links. In the absence of coupling (i.e., $\alpha = 0$) the metric distance between nodes does not influence the magnitudes of the weights. Consequently, the dis-

tribution of weights generated by the model is the original probability distribution function in Eq. (5.2) convoluted with the distribution of values of the ratio $\sigma\sigma'/\kappa\kappa'$. Conversely, in the case of strong coupling, short-range links are constrained to have larger weights, whereas long-range ones have small weights. This effect increases the heterogeneity in the weight distribution. Figure 5.3**a** shows this effect on synthetic networks generated with the model with identical parameters except for the value of α.

The same effect is visible in the local heterogeneity of the weights attached to a given node. To characterize such heterogeneity, we use the disparity measure defined as

$$Y_i = \sum_j \left(\frac{\omega_{ij}}{s_i}\right)^2 , \tag{5.23}$$

where ω_{ij} is the weight of the link between nodes i and j ($\omega_{ij} = 0$ if there is no link) and $s_i = \sum_j \omega_{ij}$ is the strength of the node [23]. Notice that this definition is similar to the one used in Chapter 3, except for a k_i factor. In Fig. 5.3**b**, we see that in the absence of coupling (i.e., $\alpha = 0$) the disparity scales as $Y_i \sim k_i^{-1}$ corresponding to the situation in which weights are roughly homogeneously distributed among the links [150]. On the other side of the spectrum, we see that under maximal coupling (i.e., $\alpha \lesssim 1$), the disparity decreases slower than k_i^{-1} meaning that weights are heterogeneous and that the large strength of nodes is due to a handful of links with large weights.

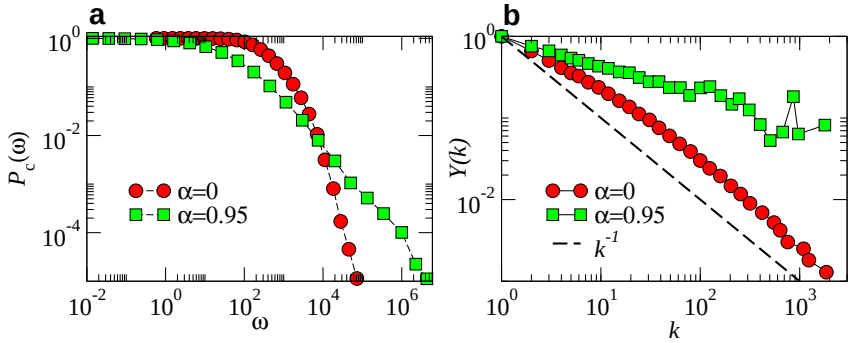


Figure 5.3: **Effect of the underlying geometry on the weights.** Simulations correspond to synthetic networks with the same parameters as in Fig. 5.2 but with two different values of the coupling parameter α. **a**, Complementary cumulative weight distribution for $\alpha = 0$ (no coupling with the hidden metric space) and $\alpha = 0.95$ (strong coupling). **b**, Average disparity of nodes as a function of their degree for the same values of the coupling α. The dashed line shows the scaling k^{-1}, corresponding to the a perfect equipartition of the strength of a node among its links.

5.4 Hidden metric spaces underlying real weighted networks

In Section 5.2, we showed that the normalized weights of links participating in triangles are higher, thus suggesting a coupling between the weighted organization of real weighted complex networks and an underlying metric space. We then presented a model that has the critical ability to fix the joint degree-strength distribution while, independently, varying the level of coupling between the weights and the metric space (parameter α). This opens the way to a definite proof of the geometric nature of weights in real complex networks, which inevitably must involve the triangle inequality, a fundamental property of any metric space.

For unweighted networks, a direct verification of the triangle inequality based on the topology without an embedding in a metric space is not possible. In contrast, if we assume weights to be influenced by distances in an underlying metric space, like in Eq. (5.3), then the numerical values in the weights of topological triangles must reflect the inequality. We now derive a method to test the triangle inequality in real weighted networks based on the observed weights only. To ensure that the metric properties of triples in the network are in correspondence to the metric properties of the corresponding triangles in the underlying space, only triples of nodes forming triangles in the network are taken into account to evaluate the triangle inequality. There are however two main challenges when one tries to apply this methodology. The first one is related to the fact that connections in the weighted $\mathbb{S}^1$ model depend not only on angular distances but also on hidden degrees, so we need a purely geometrical formulation of the weighted hidden metric space network model in which angular distances and degrees are combined into a single distance measure. Clearly, this problem can be easily solved using the isomorphism with the $\mathbb{H}^2$ model. The second issue, which is more involved, is related to the intrinsic noise present in the system due to the stochastic nature of the processes conforming it, which may blur the evaluation of the triangle inequality.

By using the mapping to hyperbolic space introduced in Chapter 2,

$$r_i = \hat{R} - 2\ln\left[\frac{\kappa_i}{\kappa_0}\right] \quad \text{with} \quad \hat{R} = 2\ln\left[\frac{N}{\mu\pi\kappa_0^2}\right], \tag{5.24}$$

which maps

$$\frac{d_{ij}}{\mu\kappa_i\kappa_j} \longrightarrow e^{\frac{1}{2}(x_{ij}-\hat{R})} \tag{5.25}$$

with $x_{ij} = r_i + r_j + 2\ln\frac{\Delta\theta_{ij}}{2}$ standing for the approximate hyperbolic distance between the two nodes, Eq. (5.3) becomes

$$\omega_{ij} = \varepsilon_{ij}\frac{\nu}{\mu^{\alpha}}\frac{\sigma_i\sigma_j}{\kappa_i\kappa_j}e^{-\frac{\alpha}{2}(x_{ij}-\hat{R})}, \tag{5.26}$$

from which we can isolate the hyperbolic distance, x_{ij}, between nodes i and j. The triangle inequality, $x_{ij} + x_{jk} \geq x_{ik}$, then becomes

$$\ln\left[\frac{\omega_{ij}\omega_{jk}}{\omega_{ik}}\left(\frac{\kappa_j}{\sigma_j}\right)^2\right] + \ln\left[\frac{\mu^{\alpha}}{\nu}\right] - \frac{1}{2}\alpha\hat{R} \leq \ln\left[\frac{\varepsilon_{ij}\varepsilon_{jk}}{\varepsilon_{ik}}\right]. \tag{5.27}$$

The first term in the left hand side of this inequality is a function of the actual weights and network topology and, thus, can be empirically estimated in any network. The next two terms on the left hand side have an explicit dependence on the parameter α (note that, according to Eq. (5.9), ν also depends on α). The term in the right hand side is a noise term whose mean value is close to zero.

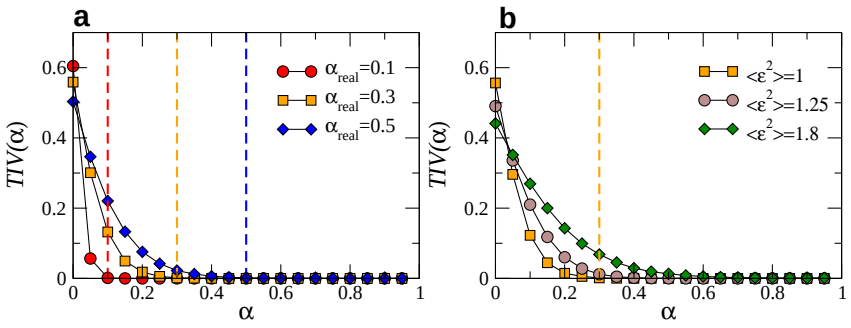


Figure 5.4: **Triangle inequality violation in synthetic networks.** Fraction of violations of the triangle inequality, $TIV(\alpha)$, (i.e. Eq. (5.27)) **a**, without noise and different values of α_{real} and **b**, with a fixed value of $\alpha_{\text{real}} = 0.3$ and different values of the noise $\langle\varepsilon^2\rangle$. In both cases, the topology is the same, with $\gamma = 2.5$, $\beta = 2$, $\eta = 1$, and $N = 10^4$. The vertical dashed lines indicate the values of α_{real} used to generate the different networks.

Let us first assume that this noise term is zero. In synthetic weighted networks generated under this condition, the inequality should hold approximately for any value of α in Eq. (5.27) equal to or larger than the value of α_{real} used to assign weights in the network. Note that it may not hold exactly even when α is greater than its real value due to the inherent noise in the estimation of the hidden variables κ and σ in Eq. (5.27), as well as the global parameters μ and ν. To minimize such uncertainty, we choose $\sigma = a\kappa^{\eta}$ and approximate κ by the degree of nodes. We propose to consider α in Eq. (5.27) as a free parameter and to measure the triangle inequality violation spectrum, $TIV(\alpha)$, defined as the fraction

of violations of the triangle inequality (i.e., triangles for which the left hand side of Eq. (5.27) is positive). In the absence of noise, $TIV(\alpha)$ should take a very small value when $\alpha \geq \alpha_{\text{real}}$ if the weighted structure of the network is congruent with the existence of an underlying metric space. In Fig. 5.4**a**, we show $TIV(\alpha)$ for synthetic networks generated with the model with different values of α_{real}. As expected, the curves fall rapidly precisely at $\alpha \gtrsim \alpha_{\text{real}}$, indicated by the dashed vertical lines.

In real situations, however, noise is typically present and has an impact on $TIV(\alpha)$. Indeed, Fig. 5.4**b** shows its behaviour for a fixed value of α_{real} and different values of the noise $\langle \varepsilon^2 \rangle$. This implies that two networks generated with different α_{real} can exhibit similar $TIV(\alpha)$ depending on their level of noise, so we need an independent measure of the noise to infer the value of α_{real} from the triangle inequality violation spectrum. For this purpose, we use the square of the coefficient of variation of the strength, which depends linearly on the noise $\langle \varepsilon^2 \rangle$. Combining these observations, we propose a procedure to infer the value of α_{real} for any real complex network based on the empirical $TIV(\alpha)$.

5.4.1 Inference via the triangle inequality

To test the triangle inequality in a given weighted complex network, we first find the parameters μ, β, and γ that best match the empirical topology. To achieve the optimal matching, we use the empirical sequence of degrees as input for the sequence of κ so that the fluctuations in the tail of the degree distribution of the input network are preserved. The sequence is then used to generate different weighted networks as follows. From the empirical relation strength-degree, we measure the proportionality factor a and the exponent η, as well as the first and second moments of the strength distribution $\langle s \rangle$ and $\langle s^2 \rangle$. For fixed values of α and of the fluctuations of the probability distribution function f, $I_3 = \langle \varepsilon^2 \rangle$, we generate a large number of synthetic weighted networks and measure the average value of $CV^2(s) = \text{Var}(s)/\langle s \rangle^2$ and its ensemble fluctuations. From our previous results, we expect the average value of $CV^2(s)$ to scale linearly with $\langle \varepsilon^2 \rangle$ as

$$CV^2(s) = \frac{\Gamma(D/2)\langle \sigma^2/\kappa \rangle^2 I_4}{2\pi^{D/2}\mu\langle \sigma \rangle^4 I_1^2 I_2^2} \langle \varepsilon^2 \rangle \, . \tag{5.28}$$

The left hand side in this equation can be directly measured from the network. The right hand side depends linearly on the noise $\langle \varepsilon^2 \rangle$ whereas its pre-factor depends both on the topology and on α_{real} through the integrals I_2 and I_4 (see Section 5.3.1 for technical details). This allows us to choose, among the set of values of α reproducing the observed triangle inequality violation spectrum, the one corresponding to the level of noise $\langle \varepsilon^2 \rangle$ that matches the observed fluctuations in the strength. The method is explained with the aid of a figure below.

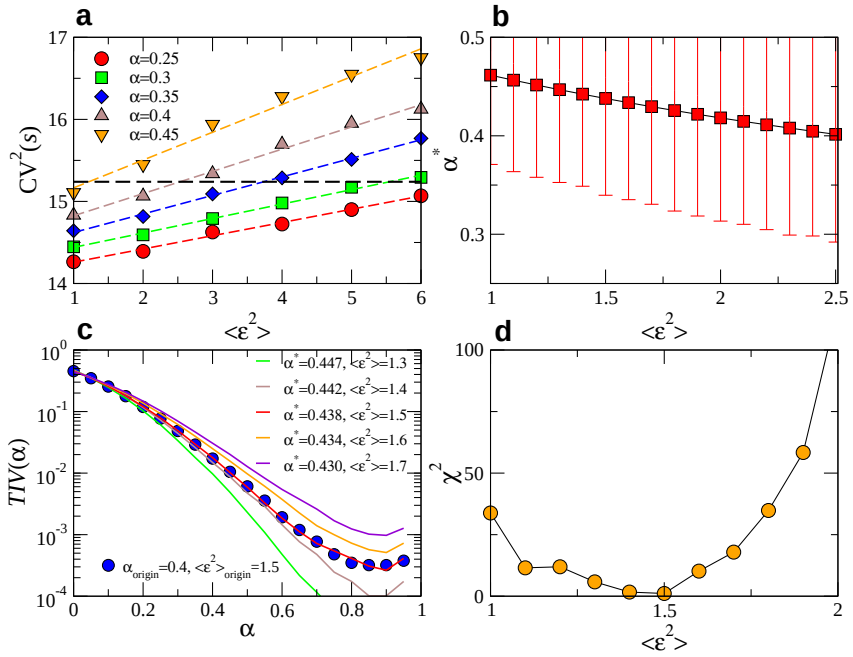


Figure 5.5: **Illustration of the test of the triangle inequality.** Test of the triangle inequality applied to a synthetic network generated with $\alpha_{\text{origin}} = 0.4$ and noise $\langle\varepsilon^2\rangle_{\text{origin}} = 1.5$. **a**, Square of the coefficient of variation of nodes' strength as a function of the noise $\langle\varepsilon^2\rangle$ in synthetic weighted networks with different values of α. The horizontal dashed line is the empirical value measured in the input network. **b**, Values of α^* as a function of the noise obtained from the intersection of the dashed line in **a** with the synthetic curves. **c**, $TIV(\alpha)$ curves for synthetic networks with the values of α^* and $\langle\varepsilon^2\rangle$ from **b** compared to the same function for the input network. **d**, χ^2 statistics obtained from the comparison of function $TIV(\alpha)$ between the model and input network.

In Fig. 5.5**a**, we show $CV^2(s)$ as a function of $\langle\varepsilon^2\rangle$ and different values of α as measured over synthetic networks used to infer the values of the parameters of a particular test network, also synthetic, generated with $\alpha_{\text{origin}} = 0.4$ and noise $\langle\varepsilon^2\rangle_{\text{origin}} = 1.5$. The intersection of these curves with the empirical value of $CV^2(s)$, that is, of the test network, defines a collection of α's as a function of the noise, $\alpha^*\left(\langle\varepsilon^2\rangle\right)$, which become the potential candidates to be the estimate of α_{real} (see Fig. 5.5**b**). In other words, instead of considering all possible values of coupling and noise, by using the coefficient of variation we reduce the space of parameters to be explored to a single dimension. Finally, for each pair $\left(\langle\varepsilon^2\rangle, \alpha^*\right)$ in Fig. 5.5**b** we measure the function $TIV(\alpha)$ over synthetic networks and compare it with the same function measured on the input network (see Fig. 5.5**c**). The comparison is performed by measuring the standard χ^2 statistics. The inferred value of α_{real} corresponds to the value of $\alpha^*(\langle\varepsilon^2\rangle)$ minimizing the value

of χ^2 (see Fig. 5.5**d**). The most remarkable aspect of the test is that it can be performed without any explicit embedding of the network and, thus, it can be readily applied to real networks for which an embedding is not available.

5.4.2 Reproducing real networks

Figures 5.6**a** and **b** show the $TIV(\alpha)$ curves for the real networks and the same curves for synthetic networks generated by our model using the inferred α_{real} and $\langle\varepsilon^2\rangle_{\text{real}}$ to be maximally congruent with the real data. In all cases, we find a very good agreement between theory and observations, which suggests a coupling with a hidden metric space as a highly plausible explanation of the observed weighted organization. Figure 5.6**c** shows the values of β (coupling topology and metric space) and α_{real} (coupling weights and metric space) inferred by our method. Notice that, except for the US airports network, α_{real} is always above 0.4, which indicates a clear and strong coupling between weights and the hidden underlying geometry. We also considered the networks for which an embedding of the binary structure was available and rescaled each weight by a factor $e^{-\alpha_{\text{real}} x_{ij}/2}$ to remove any dependence of weights on distances (see Eq. (5.26)), where x_{ij} is the hyperbolic distance between nodes in the embeddings [35, 155]. We then normalized and sampled the weights as in Fig. 5.1 and the results are shown in Fig. 5.6**d**. Strikingly, we see that the gap observed in Fig. 5.1 completely disappears in some of the networks or is significantly reduced in others. While the remaining gaps may be due to imprecisions in the embedding (i.e. the embedding procedure cannot take into account the information contained in the weights yet), these results nevertheless add their voice to the evidence pointing toward the geometric nature of the weights in real complex networks.

Name	$\rho_{m,\omega}$	$\rho_{m,\omega^{\text{norm}}}$	β	α_{real}	$\langle\varepsilon^2\rangle_{\text{real}}$	η	a
World Trade	0.68	0.10	2.5	0.41	1.3	1.63	3772
Cargo ships	0.19	0.10	1.85	0.65	1.7	1.05	83
US Commodities	0.24	0.05	1.3	0.65	1.2	1.22	3045
US Airports	0.72	0.03	1.4	0.15	1.4	1.72	10000
US Commute	0.51	0.17	2.2	0.59	1.4	2.02	719
E. Coli	0.73	0.45	2.2	0.45	1.3	1.09	1
Human brain	0.27	0.23	2.8	0.45	1.3	0.86	0.015

Table 5.2: **Parameters and information for the datasets.** Pearson correlation coefficients between the multiplicity, m, and the weight and normalized weight of links in the real networks considered in the chapter. Also, the parameters used to reproduce real networks with our model.

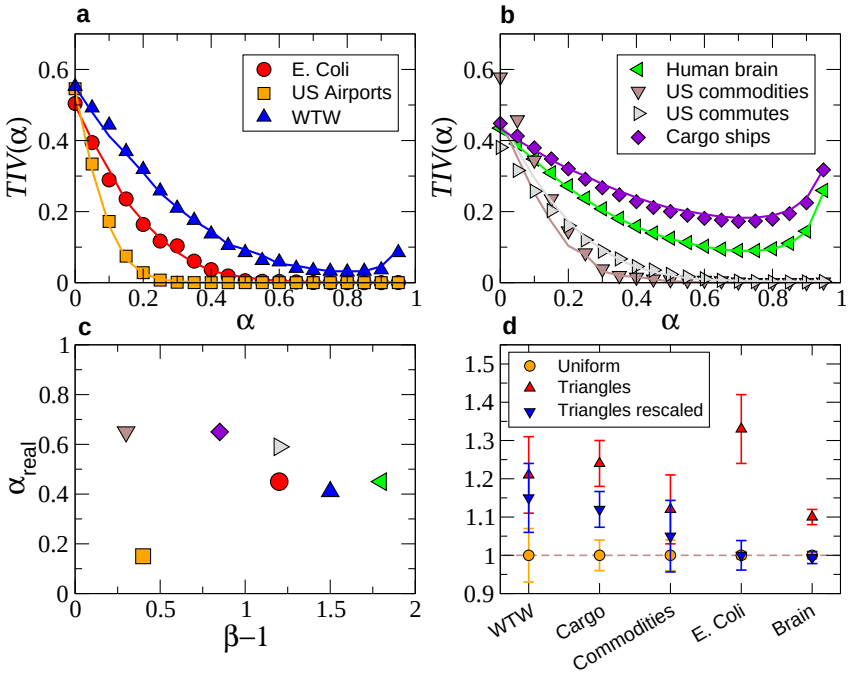


Figure 5.6: **Triangle inequality violation in real networks.** Triangle inequality violation curves (**a** and **b**) for all real networks considered in this study (symbols). Solid lines correspond to their model counterparts with the model parameters in Table 5.2. **c**, Inferred values of the coupling parameter α_{real} vs. $\beta-1$. The numerical values of these parameters can be found in Table 5.2. **d**, Average normalized weights in the network (yellow circles) with the one measured by sampling links over triangles (red triangles) for most of the empirical datasets analyzed. The error bars correspond to an estimate of the standard deviation of the average value (see the caption of Fig. 5.1 for details). The blue inverted triangles correspond to the red triangles but where the weights were rescaled by the factor $e^{-\alpha_{\text{real}} x_{ij}/2}$ to take into account the coupling between the weights and the hidden metric space. The airports and commuting networks could not be embedded into a metric space using current state-of-the-art methodology [35, 134] due to atypical topological features. These refer to a power-law degree distribution with an exponent below 2 in the case of the US airports network and a short-range repulsion effect in the connection probability for the commute network (i.e., people rarely commute from one suburb to another but rather commute to the major city in the area). This does not affect our general theory but rather prevents the state-of-the-art embedding algorithms to provide us with an embedding of these two networks.

Finally, we generated one network with our model for each dataset using these values for the parameters and then compared the resulting graphs with their real counterparts. The results are shown in Figs. 5.7–5.13. Since they all display the same features for different datasets, we explain their contents only here; each figure compares, for a given real system, the complementary cumulative degree distribution (**a**), the complementary cumulative strength distribution (**b**), the complementary cumulative weight distribution of links (**c**), the degree-dependent clustering coefficient (**d**), the average strength of nodes of degree k (**e**), and the disparity of nodes as a function of their degree (**f**). Solid lines represent the model networks while the symbols display the real properties. In all cases, the agreement between the model and the real networks is excellent. Remarkably, in the case of the weight distribution and disparity measure, such agreement is only achieved with the empirical value of α_{real} found via the test of the triangle inequality.

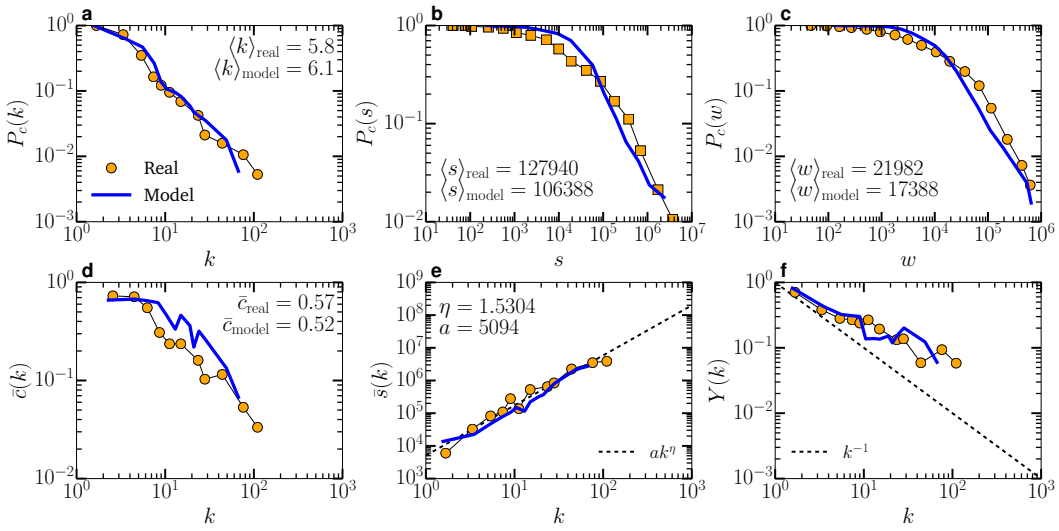


Figure 5.7: **Synthetic network and the WTW.**

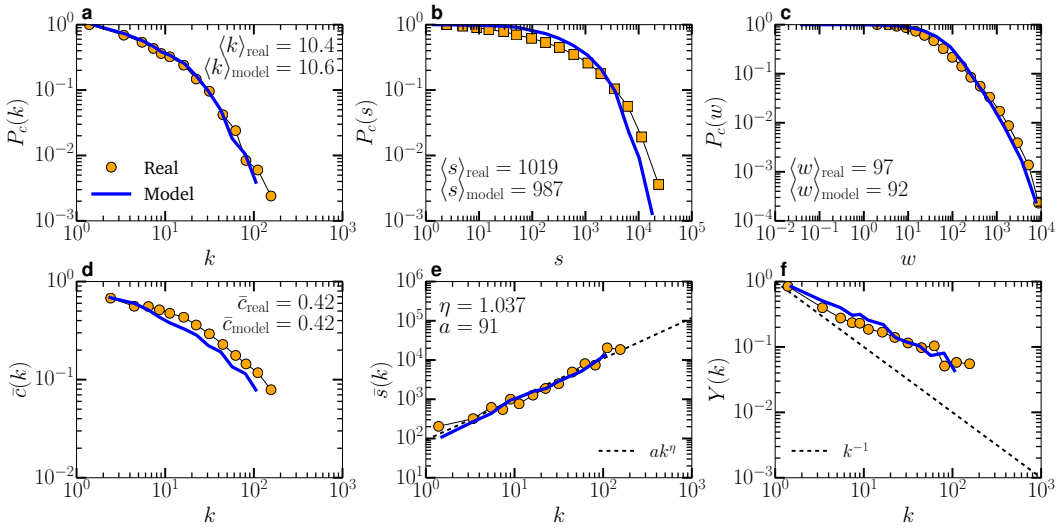


Figure 5.8: **Synthetic network and the Cargo ships network.**

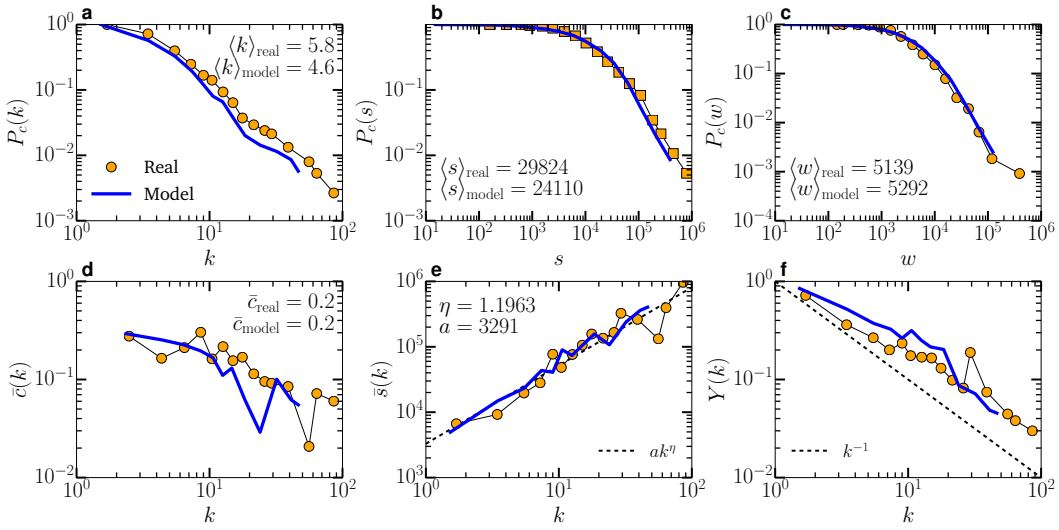


Figure 5.9: **Synthetic network and the US Commodities network.**

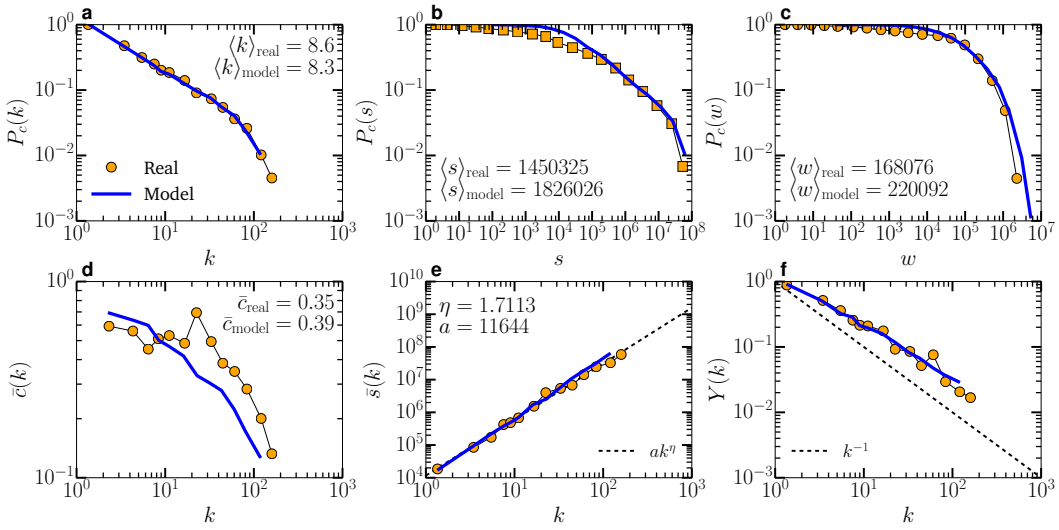


Figure 5.10: **Synthetic network and the US airports network.**

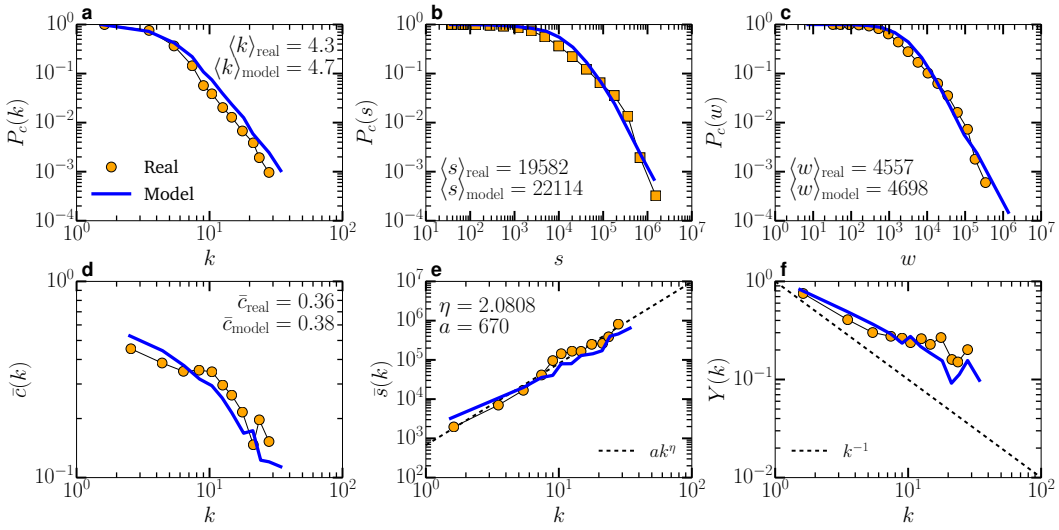


Figure 5.11: **Synthetic network and the US Commute network.**

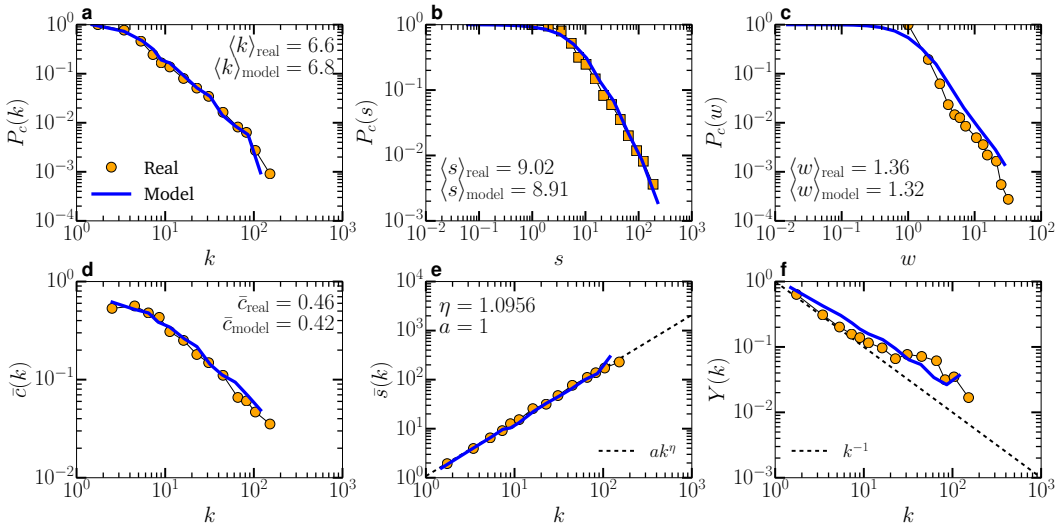


Figure 5.12: **Synthetic network and the iJO1366 *E. Coli* metabolic network.**

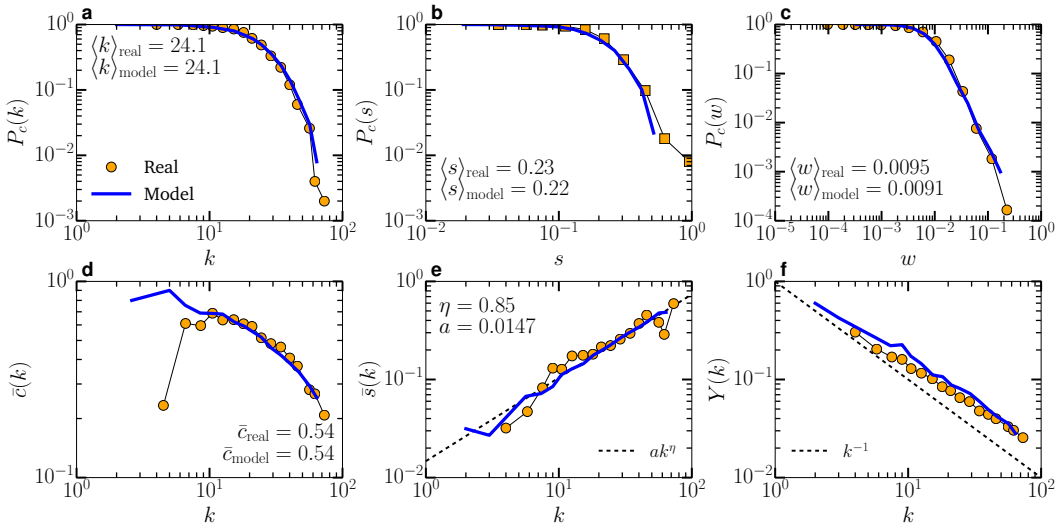


Figure 5.13: **Synthetic network and the Human brain network.**

5.5 Discussion

The metric character of many real complex networks—in which clustering is a direct consequence of the triangle inequality—has been extensively explored throughout this thesis. In this chapter, we addressed the metric nature of their weighted organization. We provided strong empirical evidence for the metric origin of the weighted architecture of real complex networks from very different domains. Our results suggest that the same underlying metric space ruling the network topology also shapes its weighted organization. It is important to notice that the distances between nodes implied by this metric space do not necessarily correspond to geographic distances (e.g., distances between ports on Earth) but are rather abstract and effective distances encoding several factors affecting the existence of connections and their intensity.

To account for these empirical findings, we proposed a very general model capable of reproducing the coupling with the metric space in a very simple and elegant way. This model allows us to fix the local properties of the nodes—their joint degree-strength distribution—while varying the coupling of the topology and, independently, of the weights with the hidden metric space. This critical property permits us to gauge quantitatively the effect of the metric space in real systems. In the case of the US airports network, we found quite remarkably that, while the coupling between the topology and the metric space is relatively strong, the coupling at the weighted level is quite weak. This strengthens the hypothesis that, in some systems, the formation of weights and topology obey different dynamics. Contrarily, we found strong coupling, both at the topological and weighted levels, even in networks that are not embedded in any obvious metric space like the metabolism of *E. Coli*, a system of metabolic reactions for which the hidden geometry is elucidated as a biochemical affinity space. This fact provides yet another empirical evidence towards the existence of hidden metric spaces shaping the architecture of these systems and, more generally, of real complex networks [152].

Our framework can be understood as a new generation of gravity models applicable to very different domains, including biology, information and communication technologies, and social systems. Indeed, Eq. (5.3) is a novel generalization of this concept to the case of weighted networks where

$$\frac{\sigma}{\kappa^{1-\alpha/D}} \tag{5.29}$$

plays the role of the "weighted mass" of nodes (as opposed to hidden degrees κ) and ensures that, once the network has been assembled, nodes have expected degree and strength κ and σ, respectively. Current gravity models predict the volume of flows between elements but cannot explain the observed topology of the interactions among them, as shown in works for the World Trade Web [59].

Our contribution overcomes this limitation and offers a gravity model that can reproduce both the existence and the intensity of interactions. This opens a new line of theoretic research on the coupling between topology, weighted structure, and geometry in complex networks.

Furthermore, our work opens the possibility to use information encoded in the weights of the links to find more accurate embeddings of real networks. Such improved embeddings are expected to allow the detection of communities or of missing links and to provide estimates of the weights of such missing links [104, 105, 175]. They can also be extremely helpful to implement navigation and searching protocols, such as greedy routing, which take into account not only the existence of connections but also their intensity.

In perspective, the hidden-metric-space weighted model and the maps of real complex systems that it will enable will lead to a deeper understanding of the interplay between the structure and function of real networks, and will offer insights on the impact they have on the dynamical processes they support and on their own evolutionary dynamics.

5.6 Calculation details

The distribution of the weight between a node with hidden variables κ and σ and a randomly chosen node in the system, $\Phi(\omega|\kappa,\sigma)$, can be obtained from Bayes' rule,

$$\Phi(\omega|\kappa,\sigma) = \frac{\rho(\omega,\kappa,\sigma)}{\rho(\kappa,\sigma)}, \tag{5.30}$$

where the joint probability distribution $\rho(\omega,\kappa,\sigma)$ results from the integration and summation of the distribution $\rho\left(\omega, a_{\kappa\kappa'},\kappa,\sigma,\kappa',\sigma',d\right)$ over the remaining random variables involved among two nodes in the model, that is,

$$\rho(\omega,\kappa,\sigma) = \sum_{a_{\kappa\kappa'}=0}^{1}\int \rho\left(\omega, a_{\kappa\kappa'},\kappa,\sigma,\kappa',\sigma',d\right)\mathrm{d}\kappa'\mathrm{d}\sigma'\mathrm{d}d, \tag{5.31}$$

where $a_{\kappa\kappa'}$ is the adjacency-matrix element encoding the existence of a connection between the two nodes. Now, $\rho\left(\omega, a_{\kappa\kappa'},\kappa,\sigma,\kappa',\sigma',d\right)$ can be easily expressed in terms of known conditional distributions:

$$\rho\left(\omega, a_{\kappa\kappa'},\kappa,\sigma,\kappa',\sigma',d\right) = \rho\left(\omega|a_{\kappa\kappa'},\kappa,\sigma,\kappa',\sigma',d\right)\rho\left(a_{\kappa\kappa'},\kappa,\sigma,\kappa',\sigma',d\right). \tag{5.32}$$

In the above expression,

$$\rho\left(\omega|a_{\kappa\kappa'},\kappa,\sigma,\kappa',\sigma',d\right) = \left(\frac{1}{\bar{\omega}}f\left(\frac{\omega}{\bar{\omega}}\right)\right)^{a_{\kappa\kappa'}}\delta(\omega)^{1-a_{\kappa\kappa'}}, \tag{5.33}$$

with $\bar{\omega} = \frac{\nu\sigma\sigma'}{(\kappa\kappa')^{1-\alpha/D}d^{\alpha}}$. The δ-function accounts for the null weight of a non-existing link. The joint probability distribution in Eq. (5.32) can be further expanded as

$$\begin{aligned}\rho\left(a_{\kappa\kappa'},\kappa,\sigma,\kappa',\sigma',d\right) &= \rho\left(a_{\kappa\kappa'}|\kappa,\sigma,\kappa',\sigma',d\right)\rho\left(\kappa,\sigma,\kappa',\sigma',d\right)\\ &= \rho\left(a_{\kappa\kappa'}|\kappa,\kappa',d\right)\rho(\kappa,\sigma)\,\rho\left(\kappa',\sigma'\right)\rho_d(d).\end{aligned} \tag{5.34}$$

In the last step, we have written the joint probability distribution in accordance with the independence between nodes and between hidden variables and spatial coordinates. We have also removed σ and σ' from the conditional distribution because links are established independently from these hidden variables; this distribution can be expressed as

$$\rho\left(a_{\kappa\kappa'}|\kappa,\kappa',d\right) = p\left(\kappa,\kappa',d\right)^{a_{\kappa\kappa'}}\left(1-p\left(\kappa,\kappa',d\right)\right)^{1-a_{\kappa\kappa'}}. \tag{5.35}$$

Introducing all these partial results into Eq. (5.30) yields Eq. (5.4),

$$\begin{aligned}\Phi(\omega|\kappa,\sigma) = \int &\left[\delta(\omega)\left(1-p\left(\kappa,\kappa',d\right)\right)+\frac{1}{\bar{\omega}}f\left(\frac{\omega}{\bar{\omega}}\right)p\left(\kappa,\kappa',d\right)\right]\\ &\times\rho\left(\kappa',\sigma'\right)\rho_d(d)\,\mathrm{d}\kappa'\mathrm{d}\sigma'\mathrm{d}d.\end{aligned} \tag{5.36}$$

The previous result represents the distribution of weight among node i and a randomly chosen node in the system, including the possibility for the two nodes not to be connected. The distribution conditioned to the fact that the connection exists, which we denoted as $\tilde{\Phi}(\omega|\kappa,\sigma)$ in Section 5.3.1, can be computed from its definition as

$$\tilde{\Phi}(\omega|\kappa,\sigma) = \rho(\omega|a_{\kappa\kappa'}=1,\kappa,\sigma) = \frac{\rho(\omega,a_{\kappa\kappa'}=1,\kappa,\sigma)}{\rho(a_{\kappa\kappa'}=1,\kappa,\sigma)}, \tag{5.37}$$

The joint probability distribution in the numerator can be obtained in a similar manner as we derived Eq. (5.31),

$$\rho(\omega,a_{\kappa\kappa'},\kappa,\sigma) = \int \rho\left(\omega,a_{\kappa\kappa'},\kappa,\sigma,\kappa',\sigma',d\right)\mathrm{d}\kappa'\mathrm{d}\sigma'\mathrm{d}d. \tag{5.38}$$

Using Eqs. (5.32)–(5.35), and evaluating the function at $a_{\kappa\kappa'}=1$, yields

$$\begin{aligned}\rho(\omega,a_{\kappa\kappa'}=1,\kappa,\sigma) = \int \frac{1}{\bar{\omega}} f\left(\frac{\omega}{\bar{\omega}}\right) p\left(\kappa,\kappa',d\right) \\ \times \rho(\kappa,\sigma)\,\rho\left(\kappa',\sigma'\right)\rho_d(d)\,\mathrm{d}\kappa'\mathrm{d}\sigma'\mathrm{d}d.\end{aligned} \tag{5.39}$$

The denominator in Eq. (5.37) can be computed considering that

$$\begin{aligned}\rho(a_{\kappa\kappa'}=1,\kappa,\sigma) &= \rho(a_{\kappa\kappa'}=1|\kappa,\sigma)\,\rho(\kappa,\sigma) \\ &= \rho(a_{\kappa\kappa'}=1|\kappa)\,\rho(\kappa,\sigma) = \tilde{p}(\kappa)\,\rho(\kappa,\sigma),\end{aligned} \tag{5.40}$$

where

$$\tilde{p}(\kappa) = \int p\left(\kappa,\kappa',d\right)\rho_\kappa\left(\kappa'\right)\rho_d(d)\,\mathrm{d}\kappa'\mathrm{d}d. \tag{5.41}$$

Notice that this result is fully consistent with the integration of Eq.(5.39) over ω. We can now write the distribution explicitly and obtain Eq. (5.13),

$$\tilde{\Phi}(\omega|\kappa,\sigma) = \frac{1}{\tilde{p}(\kappa)}\int \frac{1}{\bar{\omega}} f\left(\frac{\omega}{\bar{\omega}}\right) p\left(\kappa,\kappa',d\right)\rho\left(\kappa',\sigma'\right)\rho_d(d)\,\mathrm{d}\kappa'\mathrm{d}\sigma'\mathrm{d}d. \tag{5.42}$$

CHAPTER 6

Conclusions

During the last 20 years of network science, it has become a well-established fact that certain regularities permeate all complex networks, despite their topologies being non-trivial, in the sense that they are not completely regular nor purely random graphs. These universal features include scale-free degree distributions, sparseness, high clustering, small-worldness and community structure. In the case of weighted networks, a non-linear relation between strength and degree has been broadly observed as well.

Understanding the ubiquitousness of these features calls for the development of models including the minimal set of mathematical rules reproducing them. However, the origin of clustering remained elusive until the idea of latent similarity spaces underlying real networks was proposed. This hypothesis became the starting point of the new field of Network Geometry within the complex networks community, which also includes other approaches to the study of complex topologies. A notable example is simplicial complexes [171, 28, 29], which proposes models based on tetrahedra of different dimensions, representing interactions among more than two nodes, as fundamental building blocks of networks. The idea of latent similarity spaces, on which this thesis is based, resulted in a whole class of geometric models that have proven to have the potential not only to explain the origin of the high level of clustering coefficient observed in real networks, but even to unveil the hidden similarity space of real-world systems through what is called network embedding. This possibility confers the hidden metric spaces approach to complex networks a high degree of applicability beyond pure theoretical disquisitions regarding the structure of real systems. For instance, these maps of real similarity spaces reveal the community structure at a glance, and they can even be used to navigate networks efficiently using the so-called greedy-routing protocol. Needless to say, this makes the approach particularly appealing. Yet, it would have stood as a mere hypothesis were it not because a considerable amount of evidence supporting the existence of such hidden metric spaces has been accumulated ever since they were proposed.

Among the aforementioned geometric models, two are particularly important in this work: the $\mathbb{S}^1$ and the $\mathbb{H}^2$. The first one assumes a circle as a simple geometry for the latent similarity space occupied by nodes, while these are also equipped with another attribute encoding their proneness to acquire connec-

tions: a hidden degree. Nodes then establish connections according to a certain gravity-law-like connection probability. In the case of the $\mathbb{H}^2$ model, nodes do not have any attribute besides their positions in a space. They all lay on a finite disk in hyperbolic space, and their likelihood to connect depends only on the distance separating them. Strikingly, these two models were shown to be isomorphic, that is, equivalent descriptions of the same ensembles. As a result, the combination of these two models is a valuable tool in terms of the analysis of real networks, as it yields two notions of distance between the elements conforming the system: the $\mathbb{S}^1$ distance, standing as a measure of similarity between nodes, and the $\mathbb{H}^2$ distance, which encodes in a sole distance metric the likelihood for connections to be established, including both the popularity—or tendency to accumulate links—of nodes as well as their similarity simultaneously.

In the original versions of the $\mathbb{S}^1$ and $\mathbb{H}^2$ models, a uniform disposition of nodes on the angular dimension is customarily assumed. Yet, the embedding of real networks reveals that similarity-space coordinates are usually highly correlated, forming dense angular regions, or soft communities, as a consequence of the community structure of the network. This thesis contributes to the understanding of these models with a generalization for non-homogeneous angular distributions that proves that they are capable of accounting for these angular heterogeneities and match the target topological features. This is an important step in terms of extending the current state of these powerful models beyond several simplifying assumptions, which are usually adopted for convenience, but are not strictly necessary.

A crucial aspect of these two models is their congruency with real-world networks. Besides being able to generate realistic topologies, the geometric embeddings that they enable reveal meaningful similarity spaces of real systems. This thesis presents an example of this in terms of a thorough analysis of the network of international trade, or World Trade Web, covering a time period of 14 decades. By embedding the networks into hyperbolic space and using the mapping to the gravity-law representation, we assess in quantitative terms how the global economic system is being shaped, driven by forces that, although might seem contradictory, coexist: globalization, localization and hierarchization. Indeed, our geometric analysis of the trade system allows us to conclude that, although it is certainly true that the world is becoming more global in economic terms, with trade interactions becoming further-reaching over time, this is not the case for the significant trade channels; distance, understood in this context as an abstract, effective measure of similarity between countries, and henceforth not limited to geography, still matters. Furthermore, the congruency of the models and real systems can be exploited to devise new methodologies. For instance, our results show that similarity-distance-based community-detection methods seem to have higher resolution than other methods currently available

in the literature, in the sense that they can resolve smaller communities in real systems.

From a fundamental point of view, this thesis provides valuable insights to our understanding of the structure of complex networks. In the case of binary networks, our Geometric Renormalization Group unveils a previously unknown symmetry of real networks, namely, self-similarity with respect to similarity-scale transformations. Remarkably, this self-similarity is precisely explained by the congruency of real networks with the $\mathbb{S}^1$ model. The renormalization of the model reveals that any network admitting a good embedding must remain congruent with the model as it undergoes the scale transformation. This allows us to study the behaviour of the topology along the RG flow in analytical terms and to conclude that scale-free networks remain scale-free, with similar clustering and community structure. In addition, the behaviour of the average degree admits an interpretation in terms of the global connectivity of the network revealing a clear distinction between the small-world and non-small-world phases. This theoretical analysis further identifies two mechanisms as mainly responsible for small-worldness: degree heterogeneity and randomness.

As for our understanding of the weighted structure of real systems, this work presents empirical evidence of the geometric origin of link intensities. We propose a geometric model for weighted networks in which the weight of a link depends on the underlying distance between the endpoint nodes. This model, which has the crucial ability of controlling the binary and the weighted structures of the resulting networks simultaneously, is able to reproduce the weighted structure of real networks with high accuracy. We also show that, if weights have a geometric origin, they must obey an inequality derived from the triangle inequality fulfilled by the distances with which weights are correlated. Strikingly, this non-trivial relation between weights is confirmed on real systems, and the model reproduces the same behaviour almost perfectly, henceforth making it able to explain even very profound properties of real weighted networks. Moreover, we devise a method to measure the coupling of the weighted structure to the underlying metric space using the model, and we find real complex networks to be strongly coupled to the underlying geometry in most cases.

In addition to the fundamental insights that it provides, the hidden metric spaces approach to complex networks also brings us immediate applications. In technological scenarios, the greedy-routing navigation protocol could help to overcome scalability issues of currently used protocols, as it only requires nodes to be provided with local information. For this purpose, the multiscale navigation protocol could be especially valuable, as it increases the performance of the method and reduces the burden on the highest-degree nodes. Interestingly, although this new protocol requires nodes to be complemented with additional information with respect to the single-layer version of greedy routing, this ex-

tra information is local as well, which guarantees a suitable scalability of the method.

In this thesis we also propose an application that can be of particular relevance to any field in which simulating processes on networks is required: smaller-scale replicas of real systems. Since the replicas resemble their original versions to a high degree of accuracy, the dynamics running on them also behave in a similar fashion. Hence, they can be used, for example, to explore rough parameter spaces efficiently. This can be especially useful in the recent field of Quantum Complex Networks, which focuses on the structure and dynamics of networked quantum systems [123, 124, 122]. Due to the fast increase of the dimension of Hilbert space with the size of the network, computational studies of these systems could be limited by technological constraints. Hence, if quantum dynamics also prove to exhibit similar behaviours in smaller-scale replicas, these could pave the way to study the performance of quantum systems with real topologies.

It is worth emphasizing that the applications of the hidden metric spaces approach almost always require embedding networks according to some model. This is, in general, a difficult problem from the technical point of view, and more research in this direction is still needed in order to improve the current state-of-the-art embeddings on the $\mathbb{S}^1/\mathbb{H}^2$ models. Our work concerning heterogeneous angular distributions might have an impact in this respect. Typically, the likelihood maximization procedure only seeks the best angular coordinates, whereas hidden degrees are considered to be a function of degree only and known from the start. This hypothesis is a direct consequence of the simplifying assumptions usually contemplated in the models. Nevertheless, according to our results, an inhomogeneous angular distribution requires correcting hidden degrees in such a way that they depend on the hidden variables of all other nodes, which suggests that inferring hidden degrees via likelihood maximization as well as angles might noticeably improve the quality of embeddings of real-world networks with community structure. Along the same lines, it would be interesting to investigate whether our geometric model for weighted networks could be used to embed networks in which the intensity of the interactions is available. In principle, since in this model the weight of a link depends on the distance separating the two endpoint nodes in the hidden metric space, the weighted-model embedding could complement the geometric information unveiled by the purely topological one.

This thesis also seeks to be an invitation to pursue a generalization of current network-geometrical techniques to higher-dimensional similarity spaces. The behaviour of degrees in arbitrary dimensions shows that the $\mathbb{S}^D$ model, as proposed ten years ago, is equipped with the basic ingredients that confer it a high level of control over the resulting network topology. Furthermore, the model

proves that real topologies could emerge from higher-dimensional geometries as naturally as they do from one similarity dimension. Although we have not yet achieved a sufficient degree of understanding of the interplay between dimension and resulting network properties to infer the dimension underlying real networks, this certainly is an exciting line of research. Besides the theoretical challenge that it entails, identifying the real dimensionality of a system and embedding it accordingly will bring important insights in many respects, as it will allow, for instance, to identify the number and nature of the properties of nodes that influence their connectivity. Although there is a great amount of work still to be done, the first steps in this direction already reveal an important fact: real similarity spaces cannot have very high dimensions as a consequence of the limit imposed by high clustering. This result, along with the observation that the dimensionality reduction from low dimensionality to a one-dimensional representation yields a faithful description, also explains why so many real networks have been successfully embedded using the $\mathbb{S}^1/\mathbb{H}^2$ models.

Bringing all the pieces together, the general conclusion that we can draw from this thesis is that the existence of hidden metric spaces underlying real-world complex networks is not just a hypothesis. It provides an intuitive and reasonable explanation for a wealth of network properties like small-worldness, degree heterogeneity and, in particular, the high level of clustering observed in real networks; indeed, it has even been shown that the high density of triangles in sparse networks implies their geometric origin [95]. Furthermore, the possibilities of the approach go far beyond the realistic modelling of networks and, in fact, the evidence supporting the congruency of real systems and geometric models continues to accumulate. As an example, this thesis includes results related to utterly different aspects of the structure and function of complex networks and, yet, they are all unified in a common framework set by these geometric models. It is therefore worth pursuing this line of research and overcoming the technical challenges that it poses, as it is a promising approach in terms of fundamental understanding, but also with a great potential to provide us with new applications, possibly with implications for many other fields of knowledge.

APPENDIX A

List of publications

Publications related to the thesis

Articles

- *The hidden hyperbolic geometry of international trade: World Trade Atlas 1870-2013*
  G. García-Pérez, M. Boguñá, A. Allard and M. Á. Serrano
  Scientific Reports **6**, 33441 (2016)

- *The geometric nature of weights in real complex networks*
  A. Allard, M. Á. Serrano, G. García-Pérez and M. Boguñá
  Nature Communications **8**, 14103 (2017)

- *Multiscale unfolding of real networks by geometric renormalization*
  García-Pérez, M. Boguñá and M. Á. Serrano
  Nature Physics **14**, 583–589 (2018)

- *Soft communities in similarity space*
  G. García-Pérez, M. Á. Serrano and M. Boguñá
  Journal of Statistical Physics, in press (2018)

Book chapters

- *A new approach to international trade from Network Geometry. The World Trade Atlas 1870-2013*
  G. García-Pérez, M. Boguñá, A. Allard and M. Á. Serrano
  in Networks of International Trade and Investment: Understanding globalization through the lens of network analysis. Vernon Press (2018)
  ISBN: 978-1-62273-065-0.

Other publications

- *Complex architecture of primes and natural numbers*
  G. García-Pérez, M. Á. Serrano and M. Boguñá
  Physical Review E **90**, 022806 (2014)

- *Regulation of burstiness by network-driven activation*
  G. García-Pérez, M. Boguñá and M. Á. Serrano
  Scientific Reports **5**, 9714 (2015)

Code

- *World Trade Atlas interactive video tool*
  Guillermo García-Pérez
  http://morfeo.ffn.ub.edu/wta1870-2013/

APPENDIX B

Resum en català

Les xarxes complexes són representacions matemàtiques dels patrons d'interacció entre les diferents parts dels sistemes complexos. Les seves propietats topològiques juguen un paper crucial en els seus comportaments globals, incús donant lloc al que es coneix com a fenòmens emergents, de manera que és indispensable entendre no només com la topologia subjacent afecta el comportament del sistema, sinó també l'estructura de les xarxes en si mateixes. Notablement, durant els darrers 20 anys de recerca en xarxes complexes, ha quedat palès que hi ha certes regularitats comunes a pràcticament tots els sistemes, a pesar que les seves topologies són no trivials, en el sentit que no són ni completament regulars ni purament aleatòries. Aquestes propietats universals inclouen l'heterogeneïtat del nombre de veïns dels nodes, o degree, la baixa densitat de connexions, la presència de triangles, o clustering, la propietat de món petit, o small-world, i l'estructura de comunitats. En el cas de les xarxes pesades, també s'ha observat freqüentment una relació no lineal entre el pes total que intersecta un node, o strength, i el seu degree.

Entendre la universalitat d'aquestes propietats requereix el desenvolupament de models que incloguin un conjunt mínim de regles matemàtiques amb la capacitat de reproduir-les. Tanmateix, l'origen del clustering va romandre difícil d'explicar fins que es va proposar la idea d'espais de similitud subjacents a les xarxes reals; és a dir, que la similitud entre nodes, que es pot descriure matemàticament mitjançant la geometria, influencia la seva probabilitat de connexió. Aquesta hipòtesi, que es va convertir en el punt de partida d'un nou camp dins la comunitat de les xarxes complexes, ha donat com a resultat tot un conjunt de models geomètrics que han provat tenir el potencial no només per a explicar l'origen del clustering observat en sistemes reals, sinó inclús per desvelar els espais de similitud ocults rere les xarxes reals mitjançant un procediment anomenat embedding. Aquesta possibilitat atorga l'aproximació dels espais mètrics ocults a les xarxes complexes un alt nivell d'aplicabilitat que va més enllà de les pures disquisicions sobre l'estructura dels sistemes reals. Per exemple, aquests mapes dels espais de similitud reals revelen l'estructura de comunitats, i poden fins i tot fer-se servir per navegar les xarxes eficientment mitjançant un protocol anomenat greedy-routing. Evidentment, això fa que aquesta aproximació a l'estudi de les xarxes sigui particularment atractiva. Tanmateix, hauria sigut una mera hipòtesi de no ser per la considerable quantitat d'evidència a fa-

vor de l'existència d'aquests espais mètrics ocults que s'ha acumulat des que van ser proposats.

D'entre els models geomètrics, dos en són particularment importants en aquest treball: els anomenats $\mathbb{S}^1$ i $\mathbb{H}^2$. El primer assumeix un cercle com a simple geometria representant l'espai de similitud ocupat pels nodes, mentre que cada un d'ells també disposa d'un altre atribut—el degree ocult o hidden degree—que caracteritza la seva tendència a adquirir connexions. Els nodes estableixen connexions d'acord amb una certa funció de probabilitat amb la forma funcional d'una llei de gravetat. En el cas del model $\mathbb{H}^2$, els nodes no tenen cap atribut més que les seves posicions a un espai. Tots ells estan situats en un disc finit a l'espai hiperbòlic, i la seva probabilitat de connexió depèn de la distància que els separa en l'espai. Sorprenentment, es va trobar que aquests dos models són isomorfs, el qual vol dir que són descripcions equivalents dels mateixos conjunts de xarxes. Consegüentment, la combinació d'aquests dos models és una eina molt valuosa en termes d'anàlisi de xarxes reals, ja que dóna dues nocions de distància entre els elements que formen el sistema: la distància del model $\mathbb{S}^1$, representant la similitud entre nodes, i la del model $\mathbb{H}^2$, que codifica en una sola mètrica la probabilitat de què dos nodes estableixin una connexió, incloent-hi la popularitat, o tendència a establir connexions, dels nodes.

Al Capítol 2, després d'analitzar aquests dos models, presentem una generalització d'aquests per a distribucions angulars inhomogènies. A les versions originals dels models, s'acostuma a considerar una distribució angular uniforme. Els embeddings de xarxes reals, en canvi, mostren que les coordenades angulars en sistemes reals estan fortament correlacionades, formant regions angulars denses que anomenem soft communities, com a conseqüència de l'estructura de comunitats. Mitjançant l'algorisme inclòs al capítol, mostrem que els models són capaços d'incorporar aquestes inhomogeneïtats angulars i, al mateix temps, controlar les propietats topològiques resultants. Aquest és un pas important en termes de generalitzar l'estat actual d'aquests models més enllà de certes simplificacions que, tot i que s'acostumen a adoptar per conveniència, no són estrictament necessàries. A més, els nostres resultats mostren que els hidden degrees han de ser ajustats d'acord amb l'entorn local de cada node, el qual suggereix que els embeddings de xarxes reals es podrien millorar tenint en compte consideracions similars. Finalment, el capítol també inclou una descripció del model en dimensions superiors, $\mathbb{S}^D$, el qual mostra que els espais de similitud de dimensió més alta també poden generar xarxes realistes. Tot i que fer servir aquest model per a trobar embeddings de xarxes reals està més enllà de les nostres capacitats tècniques actuals, en podem extreure una conclusió molt important de la seva anàlisi. En augmentar la dimensió de l'espai subjacent, el nivell de clustering màxim que pot generar el model decreix, de manera que l'espai de similitud de xarxes amb clustering alt ha de ser baix. Aquest fet explica la bona

qualitat observada generalment en embeddings de xarxes reals en una sola dimensió, ja que la distorsió infligida per l'embedding d'una xarxa provinent d'un espai de dimensió baixa és limitada.

El poder combinat d'aquests dos models per a obrir noves maneres d'entendre les diferents dimensions que afecten l'estructura i evolució de les xarxes reals es demostra al Capítol 3, a on presentem una anàlisi detallada de la xarxa de comerç internacional, o World Trade Web, al llarg d'un període de 14 dècades. A partir dels embeddings d'aquestes xarxes a l'espai hiperbòlic i fent servir la identificació amb la representació en termes de la llei de gravetat del model $\mathbb{S}^1$, podem avaluar en termes quantitatius com el sistema econòmic global està sent modelat per forces que, tot i semblar contradictòries entre si, coexisteixen: la globalització, la localització i la jerarquització. Mentre que és cert que el món està esdevenint més global en termes econòmics, amb interaccions comercials cada vegada entre punts més remots del planeta, en aquest capítol mostrem que no només això no és cert en el cas de les interaccions comercials significatives, sinó que la distància, entesa en aquest context com a mesura abstracta de la similitud entre països i, per tant, no determinada per la geografia únicament, encara importa. De fet, la nostra anàlisi revela que les economies més petites s'estan allunyant gradualment, en termes comercials, de la resta del món, el qual implica que les seves oportunitats de comerciar amb altres economies petites disminueixen i que els seus espais comercials disponibles estan cada vegada més limitats a les grans potències comercials. Des del punt de vista metodològic en el camp de les xarxes complexes, els nostres resultats mostren que la detecció de comunitats basada en la distància de similitud sembla tenir una resolució més alta que altres mètodes disponibles a la literatura, en el sentit què poden resoldre comunitats més petites en sistemes reals.

A més d'aportar metodologies a l'estudi de sistemes reals particulars, l'aproximació basada en sistemes mètrics ocults pot descobrir noves propietats topològiques aparentment universals que podrien passar desapercebudes altrament. Al Capítol 4, partim de la hipòtesi que la dimensió de similitud defineix la noció d'escala en sistemes reals. Per a explorar aquesta idea, definim un Grup de Renormalització Geomètric, amb una transformació que canvia l'escala d'observació de xarxes de les quals es disposa d'un embedding al model $\mathbb{S}^1$. Aquest procediment revela una simetria de les xarxes reals anteriorment desconeguda: l'auto-semblança respecte a transformacions d'escala a l'espai de similitud subjacent. Notablement, aquesta auto-similitud troba una explicació precisa en la congruència entre les xarxes reals i el model $\mathbb{S}^1$. La renormalització del model mostra que qualsevol xarxa que admeti un bon embedding ha de romandre congruent amb el model a mesura que se sotmet a la transformació. Això permet estudiar el comportament de la topologia al llarg de la renormalització analíticament i concloure que les xarxes heterogènies roma-

nen heterogènies, amb clustering i estructura de comunitats similars. A més, el comportament del nombre mitjà de veïns admet una interpretació en termes de connectivitat global de la xarxa, el qual revela una clara distinció entre fases small-world i no small-world. Aquesta anàlisi teòrica també identifica dos mecanismes com a principals responsables de la propietat de small-world: l'heterogeneïtat de degree i l'aleatorietat. Més enllà de l'aportació a la nostra comprensió de les topologies complexes, el grup de renormalització també suposa aplicacions immediates. Al capítol, presentem un mètode per a construir rèpliques a petita escala de xarxes reals a les quals les dinàmiques es comporten de manera quasi idèntica a com ho fan a les xarxes originals. També incloem un algorisme de navegació que aprofita la informació del sistema a diferents escales per a millorar l'eficiència del protocol en una sola capa.

Tenint en compte la idoneïtat de l'aproximació dels espais mètrics ocults per a explicar la topologia complexa de les xarxes reals, sembla raonable qüestionar-se si la hipòtesi de l'espai de similitud subjacent només s'aplica a les xarxes binàries; certament, tots els resultats previs en aquest camp s'han limitat a aquesta classe de xarxes. Al Capítol 5, afrontem precisament aquest punt considerant xarxes pesades. En particular, partim de la conjectura de què, si les distàncies de similitud afecten els pesos dels links, ha d'existir una correlació positiva entre la multiplicitat d'un link (nombre de triangles als quals pertany) i el seu pes, ja que totes dues quantitats augmenten amb la proximitat entre els dos nodes a l'espai subjacent. Aquesta conjectura es confirma en tots els sistemes reals considerats al capítol, resultat que pot ser interpretat com a evidència de la natura geomètrica de l'estructura pesada de les xarxes reals. Consegüentment, proposem un model geomètric per a xarxes pesades al qual el pes d'un link depèn de la distància entre els dos nodes als extrems d'aquest. Tal com mostrem al capítol, aquest model té la capacitat crucial de controlar les estructures binària i pesada de les xarxes resultants simultàniament, de manera que pot reproduir les propietats d'aquesta classe de xarxes amb precisió. Tanmateix, mostrem que, si els pesos tenen un origen geomètric, han d'obeir una desigualtat derivada de la desigualtat triangular satisfeta per les distàncies amb les quals els pesos estan correlacionats. Sorprenentment, aquesta relació no trivial entre pesos es confirma en xarxes reals, i el model reprodueix el mateix comportament quasi exactament, de manera que aquest pot explicar fins i tot propietats molt subtils de xarxes pesades reals. Finalment, proposem un mètode per a mesurar l'acoblament de l'estructura pesada amb l'espai mètric subjacent, el qual és fort en xarxes reals, d'acord amb les nostres observacions.

Més enllà del coneixement fonamental que en podem extreure, aquestes troballes es poden considerar útils en altres camps que empren metodologies que provenen de les xarxes complexes. Per exemple, molts camps, incloent-hi l'economia, les ciències socials i la biologia, tracten amb sistemes interconnec-

tats, de manera que l'anàlisi d'aquests sistemes es podria beneficiar de la possibilitat de definir distàncies efectives entre nodes, el qual, a més, permet detectar comunitats amb alta resolució. Tanmateix, en algunes situacions tecnològiques, el protocol de navegació greedy-routing podria ajudar a resoldre els problemes d'adaptabilitat dels protocols actualment en ús. Per a aquest propòsit, el protocol de navegació multiescalar podria ser especialment valuós, ja que augmenta el rendiment del mètode a l'hora que redueix la càrrega de treball dels nodes de degree més alt. Finalment, les rèpliques a petita escala de xarxes reals són d'especial rellevància per a qualsevol camp que requereixi simular processos dinàmics en xarxes; donat que les rèpliques són extremadament semblants a les originals, les dinàmiques que corren sobre elles també exhibeixen comportaments similars. Així doncs, es poden fer servir, per exemple, per a explorar espais de paràmetres complicats eficientment. Això pot ser particularment útil en el recent camp de les Xarxes Complexes Quàntiques, el qual se centra en l'estructura i la dinàmica de sistemes quàntics interconnectats a on, a causa del ràpid creixement de la dimensió de l'espai de Hilbert amb la mida de la xarxa, els estudis computacionals d'aquests sistemes poden trobar limitacions tecnològiques. Per tant, si es demostra que les dinàmiques quàntiques també exhibeixen comportaments similars a les rèpliques a petita escala, aquestes podrien obrir el camí a avaluar l'eficiència de sistemes quàntics amb topologies reals.

Val la pena esmentar que les aplicacions de l'aproximació dels espais mètrics subjacents gairebé sempre requereixen trobar els embeddings de xarxes reals d'acord amb algun model. Aquest és, en general, un problema difícil des del punt de vista tècnic, i es requereix més recerca en aquesta direcció. Més enllà de millorar els actuals embeddings en els models $\mathbb{S}^1/\mathbb{H}^2$, és important trobar una manera d'identificar el nombre de dimensions subjacents als sistemes reals i obtenir els corresponents embeddings. Certament, això ens proporcionaria una gran quantitat de noves metodologies en termes d'anàlisi i detecció de comunitats. A més, també seria interessant explorar com el model per xarxes pesades proposat en aquesta tesi es podria aprofitar per a millorar la qualitat dels embeddings d'aquest tipus de xarxes.

Unint totes les peces, la conclusió general que podem extreure d'aquesta tesi és que l'existència d'espais mètrics subjacents a les xarxes complexes reals és més que una hipòtesi. A part de proporcionar-nos una explicació intuïtiva i raonable per l'elevada presència de triangles en xarxes reals, les possibilitats de l'aproximació van molt més enllà del modelatge de topologies realistes i, de fet, l'evidència a favor de la congruència dels sistemes reals i els models geomètrics es continua acumulant. Com a exemple, aquesta tesi inclou resultats relacionats amb aspectes completament diferents de l'estructura i la funcionalitat de les xarxes complexes, tots ells unificats dins un mateix marc teòric definit pels models geomètrics. Per tant, val la pena prosseguir amb aquesta línia de re-

cerca i resoldre les dificultats tècniques que suposa, ja que és una aproximació prometedora en termes de comprensió fonamental, però també amb un gran potencial per a proporcionar noves aplicacions.

References

In this thesis

[1] García-Pérez, G., Serrano, M. Á., and Boguñá, M. Soft communities in similarity space. *J. Stat. Phys.*, 2018. doi: 10.1007/s10955-018-2084-z. 24, 34

[2] García-Pérez, G., Boguñá, M., Allard, A., and Serrano, M. Á. The hidden hyperbolic geometry of international trade: World Trade Atlas 1870-2013. *Sci. Rep.*, 6:33441, 2016. doi: 10.1038/srep33441. 14, 15, 24, 25, 27, 31, 34, 46, 49, 50, 55, 62, 77, 112

[3] García-Pérez, G., Boguñá, M., and Serrano, M. Á. Multiscale unfolding of real networks by geometric renormalization. *Nat. Phys.*, 14(6):583–589, 2018. doi: 10.1038/s41567-018-0072-5. 15, 24, 44, 46, 79, 99

[4] Allard, A., Serrano, M. Á., García-Pérez, G., and Boguñá, M. The geometric nature of weights in real complex networks. *Nat. Commun.*, 8:14103, 2017. doi: 10.1038/ncomms14103. 15, 24, 25, 40, 111, 117

Other references

[5] Abou-Rjeili, A. and Karypis, G. Multilevel algorithms for partitioning power-law graphs. In *IEEE International Parallel & Distributed Processing Symposium (IPDPS)*, 124–124, 2006. 103

[6] Ahn, Y.-Y., Bagrow, J. P., and Lehmann, S. Link communities reveal multiscale complexity in networks. *Nature*, 466(7307):761–764, 08 2010. 83

[7] Alanis-Lobato, G. and Andrade-Navarro, M. A. Distance distribution between complex network nodes in hyperbolic space. *Complex Systems*, 25, 2016. 26

[8] Alanis-Lobato, G., Mier, P., and Andrade, M. Efficient embedding of complex networks to hyperbolic space via their laplacian. *Sci. Rep.*, 6, 2016. doi: doi:10.1038/srep30108. 14, 31

[9] Almog, A., Squartini, T., and Garlaschelli, D. A gdp-driven model for the binary and weighted structure of the international trade network. *New J. Phys.*, 17(1):013009, 2015. 50

[10] Alvarez-Hamelin, J. I., Dall'Asta, L., Barrat, A., and Vespignani, A. K-core decomposition of internet graphs: hierarchies, self-similarity and measurement biases. *Netw. Heterog. Media*, 3(2):371–393, 2008. 6

[11] Amaral, L. A. N., Scala, A., Barthelemy, M., and Stanley, H. E. Classes of small-world networks. *Proc. Natl. Acad. Sci. USA*, 97(21):11149–11152, 2000. doi: 10.1073/pnas.200327197. 3

[12] Anderson, J. E. and van Wincoop, E. Trade costs. *J. Econ. Lit.*, 42(3):691–751, 2004. 49

[13] Antal, T. and Krapivsy, P. L. Weight-driven growing networks. *Phys. Rev. E*, 71:026103, 2005. doi: 10.1103/PhysRevE.71.026103. 9, 115

[14] Arenas, A., Fernández, A., and Gómez, S. Analysis of the structure of complex networks at different resolution levels. *New J. Phys.*, 10(5):053039, 2008. 6, 34, 83

[15] Avena-Koenigsberger, A., Goni, J., Betzel, R. F., van den Heuvel, M. P., Griffa, A., Hagmann, P., Thiran, J.-P., and Sporns, O. Using Pareto optimality to explore the topology and dynamics of the human connectome. *Philos. Trans. R. Soc. B Biol. Sci.*, 369:20130530, 2014. doi: 10.1098/rstb.2013.0530. 113

[16] Baldwin, R. E. and Martin, P. *Two Waves of Globalisation: Superficial Similarities, Fundamental Differences*. NBER Working Paper 6904, National Bureau of Economic Research, Inc., 1999. 66

[17] Barabási, A. L. and Albert, R. Emergence of scaling in random networks. *Science*, 286:509–512, 1999. 8

[18] Barbieri, K. and Keshk, O., 2012. Correlates of War Project Trade Data Set Codebook, Version 3.0. [online] Available at `http://correlatesofwar.org`. 54

[19] Barbieri, K., Keshk, O. M., and Pollins, B. M. Trading data: Evaluating our assumptions and coding rules. *Conflict Management and Peace Science*, 26(5):471–491, 2009. doi: 10.1177/0738894209343887. 54, 55

[20] Barrat, A., Barthélemy, M., Pastor-Satorras, R., and Vespignani, A. The architecture of complex weighted networks. *Proc. Natl. Acad. Sci. USA*, 101 (11):3747–52, 2004. doi: 10.1073/pnas.0400087101. 6, 111, 114, 121

[21] Barrat, A., Barthélemy, M., and Vespignani, A. Weighted Evolving Networks: Coupling Topology and Weight Dynamics. *Phys. Rev. Lett.*, 92: 228701, 2004. doi: 10.1103/PhysRevLett.92.228701. 9, 115

[22] Barthélemy, M. Spatial networks. *Phys. Rep.*, 499(1-3):1–101, 2011. doi: 10.1016/j.physrep.2010.11.002. 6, 111

[23] Barthélemy, M., Barrat, A., Pastor-Satorras, R., and Vespignani, A. Characterization and modeling of weighted networks. *Physica A*, 346:34–43, 2005. doi: 10.1016/j.physa.2004.08.047. 123

[24] Bender, E. a. and Canfield, E. The asymptotic number of labeled graphs with given degree sequences. *J. Comb. Theory A*, 24(3):296–307, 1978. doi: 10.1016/0097-3165(78)90059-6. 8

[25] Bergstrand, J. H. and Egger, P. Gravity equations and economic frictions in the world economy. In Bernhofen, D., Falvey, R., Greenaway, D., and Krieckemeier, U., editors, *Palgrave Handbook Of International Trade*. Palgrave Macmillan Press, 2011. 49

[26] Bergstrand, J. H. The gravity equation in international trade: some microeconomic foundations and empirical evidence. *Rev. Econ. Stat.*, 67(3): 474–481, 1985. 14, 49

[27] Bianconi, G. Emergence of weight-topology correlations in complex scale-free networks. *Europhys. Lett.*, 71:1029–1035, 2004. doi: 10.1209/epl/i2005-10167-2. 9, 115

[28] Bianconi, G. and Rahmede, C. Complex quantum network manifolds in dimension $d > 2$ are scale-free. *Sci. Rep.*, 5:13979, 2015. 139

[29] Bianconi, G. and Rahmede, C. Network geometry with flavor: From complexity to quantum geometry. *Phys. Rev. E*, 93:032315, 2016. doi: 10.1103/PhysRevE.93.032315. 139

[30] Blasius, T., Friedrich, T., Krohmer, A., Laue, S. Efficient embedding of scale-free graphs in the hyperbolic plane. *IEEE/ACM Trans. Netw.*, 26(2): 920–933, 2018. doi: 10.1109/TNET.2018.2810186. 14, 31

[31] Blondel, V. D., Guillaume, J.-L., Lambiotte, R., and Lefebvre, É. Fast unfolding of communities in large networks. *J. Stat. Mech.*, 2008(10):P10008, 2008. 71, 84

[32] Boettcher, S. and Brunson, C. Renormalization group for critical phenomena in complex networks. *Front. Physiol.*, 2:102, 2011. doi: 10.3389/fphys.2011.00102. 79

[33] Boguñá, M. and Krioukov, D. Navigating ultrasmall worlds in ultrashort time. *Phys. Rev. Lett.*, 102(5):058701, 2009. doi: 10.1103/PhysRevLett.102.058701. 16, 40, 115

[34] Boguñá, M. and Pastor-Satorras, R. Class of correlated random networks with hidden variables. *Phys. Rev. E*, 68(3):036112, 2003. doi: 10.1103/PhysRevE.68.036112. 22, 42

[35] Boguñá, M., Papadopoulos, F., and Krioukov, D. Sustaining the Internet with Hyperbolic Mapping. *Nat. Commun.*, 1:62, 2010. doi: 10.1038/ncomms1063. 14, 15, 16, 17, 24, 25, 27, 28, 29, 31, 37, 46, 63, 64, 88, 98, 128, 129

[36] Boguñá, M., Krioukov, D., and Claffy, K. C. Navigability of complex networks. *Nat. Phys.*, 5:74–80, 2009. 16, 40, 116

[37] Bollobás, B. A Probabilistic Proof of an Asymptotic Formula for the Number of Labelled Regular Graphs. *Eur. J. Combin.*, 1(4):311–316, 1980. doi: 10.1016/S0195-6698(80)80030-8. 8

[38] Bolt, J. and van Zanden, J. L., 2014. The Maddison Project: collaborative research on historical national accounts. [online] Available at `http://www.ggdc.net/maddison/maddison-project/home.htm`. 56

[39] Brockmann, D. and Helbing, D. The Hidden Geometry of Complex, Network-Driven Contagion Phenomena. *Science*, 342(6164):1337–1342, 2013. doi: 10.1126/science.1245200. 111

[40] Broder, A., Kumar, R., Maghoul, F., Raghavan, P., Rajagopalan, S., Stata, R., Tomkins, A., and Wiener, J. Graph structure in the Web. *Comput. Netw.*, 33:309–320, 2000. doi: 10.1016/S1389-1286(00)00083-9. 3

[41] Brun, J.-F., Carrère, C., Guillaumont, P., and de Melo, J. Has distance died? evidence from a panel gravity model. *World Bank Econ. Rev.*, 19:99–120, 2005. 77

[42] BTS, 2015. U.S. Bureau of Transportation Statistics. [online] Available at `http://transtats.bts.gov`. 113

[43] Cagé, J. and Gadenne, L. Tax revenues, development, and the fiscal cost of trade liberalization, 1792-2006. *Working paper, The Warwick Economics Research Paper Series (TWERPS) 1132, University of Warwick*, 2016. 77

[44] Candellero, E. and Fountoulakis, N. Clustering and the hyperbolic geometry of complex networks. In Bonato, A., Graham, F. C., and Pralat, P., editors, *Algorithms and Models for the Web Graph*, volume 8882 of *Lecture Notes in Computer Science*, 1–12. Springer International Publishing, 2014. ISBN 978-3-319-13122-1. doi: 10.1007/978-3-319-13123-8_1. 24

[45] Cannon, J., Floyd, W., Kenyon, R., and Parry, W. *Flavors of Geometry*, chapter Hyperbolic Geometry. MSRI, Berkeley, 1997. 12

[46] Chase-Dunn, C., Kawano, Y., and Brewer, B. D. Trade globalization since 1795: Waves of integration in the world-system. *Am. Sociol. Rev.*, 65:77–95, 2000. 66

[47] CIA, 2015. The World Factbook. [online] Available at https://www.cia.gov/library/publications/the-world-factbook/. 56

[48] Claffy, K., Hyun, Y., Keys, K., Fomenkov, M., and Krioukov, D. Internet mapping: from art to science. In *CATCH*, 2009. 82

[49] Cohen, R. and Havlin, S. Scale-free networks are ultrasmall. *Phys. Rev. Lett.*, 90:058701, 2003. 5, 79

[50] Cohen, R. and Havlin, S. *Complex Networks: Structure, Robustness and Function*. Cambridge University Press, 2010. ISBN 9780521841566. 2

[51] Colizza, V., Flammini, A., Serrano, M. Á., and Vespignani, A. Detecting rich-club ordering in complex networks. *Nat. Phys.*, 2:110–115, 2006. 54

[52] Colizza, V., Pastor-Satorras, R., and Vespignani, A. Reaction-diffusion processes and metapopulation models in heterogeneous networks. *Nat. Phys.*, 3:276–282, 2007. 103

[53] Costa, L. D. F., Rodrigues, F. A., Travieso, G., and Boas, P. R. V. Characterization of complex networks: A survey of measurements. *Adv. Phys.*, 56:78, 2005. doi: 10.1080/00018730601170527. 5

[54] Cover, T. M. and Thomas, J. A. *Elements of information theory*. Wiley-Interscience, Hoboken, New Jersey, second edition, 2006. ISBN 978-0-471-24195-9. 71

[55] COW, 2012. The Correlates of War Project. [online] Available at http://www.correlatesofwar.org/. 54

[56] de Solla Price, D. Networks of scientific papers. *Science*, 149:510–515, 1965. 3

[57] Dorogovtsev, S. N. and Mendes, J. F. F. *Handbook of Graphs and Networks*, chapter Accelerated Growth of Networks. Wiley-VCH, Berlin, 2003. 2

[58] Dorogovtsev, S. N., Goltsev, A. V., and Mendes, J. F. F. Ising model on networks with an arbitrary distribution of connections. *Phys. Rev. E*, 66: 016104, 2002. doi: 10.1103/PhysRevE.66.016104. 95

[59] Dueñas, M. and Fagiolo, G. Modeling the International-Trade Network: a gravity approach. *J. Econ. Interact. Coord.*, 8:155–178, 2013. doi: 10.1007/s11403-013-0108-y. 49, 60, 78, 134

[60] Erdős, P. and Rényi, A. On the evolution of random graphs. *Publications Mathematics Institute Hungarian Academy of Science*, 5:17, 1960. 7

[61] Erdős, P. and Rényi, A. On random graphs. *Publ. Math.*, 6:290–297, 1959. 7

[62] Euler, L. Solutio problematis ad geometrian situs pertinentis. *Comentarii academiae scientarum Petropolitanae*, 8:128–140, 1736. 2

[63] Faloutsos, M., Faloutsos, P., and Faloutsos, C. On Power-law relationship of the internet topology. In *SIGCOMM'99*, volume 53, 1689–1699, 1999. ISBN 9788578110796. doi: 10.1017/CBO9781107415324.004. 3

[64] Fortunato, S. Community detection in graphs. *Phys. Rep.*, 486(3):75–174, 2010. doi: 10.1016/j.physrep.2009.11.002. 34

[65] Frankel, J. *Regional Trading Blocs in the World Economic System*. Institute for International Economics, Washington DC, 1997. 71

[66] Frankel, J. and Rose, A. An estimate of the effect of common currencies on trade and income. *Q. J. Econ.*, 117(2):437–466, 2002. doi: 10.1162/003355302753650292. 49

[67] Friedman, T. L. *The world is flat: A brief history of the twenty-first century*. Farrar, Straus and Giroux, 2005. 66, 77

[68] Fujita, M., Krugman, P., and Venables, A. *The Spatial Economy: Cities, Regions and International Trade*. MIT Press, Cambridge, MA, 1999. 66, 77

[69] García-Pérez, G., Boguñá, M., Allard, A., and Serrano, M. Á., 2016. The World Trade Atlas 1870-2013 interactive tool. [online] Available at `http://morfeo.ffn.ub.edu/wta1870-2013`. 18, 51, 58, 62, 63, 66, 77

[70] Garlaschelli, D. The weighted random graph model. *New J. Phys.*, 11:073005, 2009. doi: 10.1088/1367-2630/11/7/073005. 9, 115

[71] Garlaschelli, D. and Loffredo, M. Generalized Bose-Fermi Statistics and Structural Correlations in Weighted Networks. *Phys. Rev. Lett.*, 102:038701, 2009. doi: 10.1103/PhysRevLett.102.038701. 9, 115

[72] Garlaschelli, D. and Loffredo, M. I. Fitness-dependent topological properties of the world trade web. *Phys. Rev. Lett.*, 93(18):188701, 2004. 49, 50

[73] Gfeller, D. and De Los Rios, P. Spectral coarse graining of complex networks. *Phys. Rev. Lett.*, 99:038701, 2007. doi: 10.1103/PhysRevLett.99.038701. 79

[74] Ghemawat, P. Why the world isn't flat. *Foreing Policy*, 159:54–60, 2007. 66, 77

[75] Ghemawat, P. and Altman, S. Depth index of globalization 2013: And the big shift to emerging economies. Technical report, IESE, ST-310-E, 11/2013, 2013. 77

[76] Gilbert, E. N. Random graphs. *Ann. Math. Stat.*, 30(4):1141–1144, 12 1959. 8

[77] Gleditsch, K. S. Expanded trade and gdp data. *J. Conflict Resolut.*, 46:712–724, 2002. 54

[78] Gleditsch, K. S., 2015. Kristian Skrede Gleditsch's data. [online] Available at `http://ksgleditsch.com/exptradegdp.html`. 56

[79] Gnedenko, B. and Kolmogorov, A. *Limit distributions for sums of independent random variables*. Addison-Wesley series in statistics. Addison-Wesley, 1968. 91

[80] Goh, K. I., Salvi, G., Kahng, B., and Kim, D. Skeleton and fractal scaling in complex networks. *Phys. Rev. Lett.*, 96:018701, 2006. 79, 103

[81] Grady, D., Thiemann, C., and Brockmann, D. Robust classification of salient links in complex networks. *Nat. Commun.*, 3:864, 2012. doi: 10.1038/ncomms1847. 113

[82] Gulyás, A., Bíró, J. J., Kőrösi, A., Rétvári, G., and Krioukov, D. Navigable networks as Nash equilibria of navigation games. *Nat. Commun.*, 6:7651, 2015. doi: 10.1038/ncomms8651. 15, 46

[83] Handelsman, R. A. and Lew, J. S. Asymptotic expansion of laplace convolutions for large argument and tail densities for certain sums of random variables. *SIAM J. Math. Anal.*, 5(3):425–451, 1974. doi: 10.1137/0505045. 91

[84] Herfindahl, O. C. *Copper Costs and Prices: 1870-1957*. John Hopkins University Press, Baltimore, MD, USA, 1959. 57

[85] Hirschman, A. O. The paternity of an index. *Am. Econ. Rev.*, 54:761–762, 1964. 57

[86] IMF, 2015. International Monetary Fund, Direction of Trade Statistics (DOTS). [online] Available at http://www.imf.org/en/Data. 54

[87] Jeong, H., Tombor, B., Albert, R., Oltvai, Z. N., and Barabási, a. L. The large-scale organization of metabolic networks. *Nature*, 407(6804):651–4, 2000. doi: 10.1038/35036627. 3

[88] Johnston, L. and Williamson, S. H., 2015. 'What Was the U.S. GDP Then?', MeasuringWorth. [online] Available at http://www.measuringworth.org/usgdp/. 56

[89] Kaluza, P., Kölzsch, A., Gastner, M. T., and Blasius, B. The complex network of global cargo ship movements. *J. R. Soc. Interface*, 7(48):1093–1103, 2010. doi: 10.1098/rsif.2009.0495. 112

[90] Karypis, G. and Kumar, V. A fast and high quality multilevel scheme for partitioning irregular graphs. *SIAM J. Sci. Comput.*, 20:359–392, 1999. 103

[91] Kim, J. S., Goh, K. I., Hahng, B., and Kim, D. Fractality and self-similarity in scale-free networks. *New J. Phys.*, 9:177, 2007. 79, 103

[92] Klimt, B. and Yang, Y. Introducing the enron corpus. In *CEAS*, 2004. URL http://dblp.uni-trier.de/db/conf/ceas/ceas2004.html#KlimtY04. 82

[93] Krioukov, D., Papadopoulos, F., Vahdat, A., and Boguñá, M. Curvature and temperature of complex networks. *Phys. Rev. E*, 80:035101, 2009. 12, 25, 27, 98

[94] Krioukov, D., Papadopoulos, F., Kitsak, M., Vahdat, A., and Boguñá, M. Hyperbolic geometry of complex networks. *Phys. Rev. E*, 82:036106, 2010. 9, 12, 24, 25, 26, 27, 35, 37, 98

[95] Krioukov, D. Clustering implies geometry in networks. *Phys. Rev. Lett.*, 116:208302, 2016. doi: 10.1103/PhysRevLett.116.208302. 46, 143

[96] Krugman, P. *Geography and Trade*. MIT Press, Cambridge, MA, 1991. 66, 77

[97] Krugman, P. Regionalism versus multilateralism: analytical notes. In J. de Melo, J. and Panagariya, A., editors, *New Dimensions in Regional Integration*, 58–89. Cambridge University Press, Cambridge, 1993. 71, 77

[98] Krugman, P. R. Growing world trade: Causes and consequences. *Brookings Pap. Eco. Ac.*, 1995(1):327–377, 1995. 66

[99] Kumpula, J. M., Onnela, J.-P., Saramäki, J., Kaski, K., and Kertész, J. Emergence of Communities in Weighted Networks. *Phys. Rev. Lett.*, 99:228701, 2007. doi: 10.1103/PhysRevLett.99.228701. 9, 115

[100] Kunegis, J. KONECT – The Koblenz Network Collection. In *Proc. Int. Conf. on World Wide Web Companion*, 1343–1350, 2013. 82

[101] Kadanoff L. P. *Statistical Physics: Statics, Dynamics and Renormalization*. World Scientific, Singapore, 2000. 79

[102] Leskovec, J., Lang, K. J., Dasgupta, A., and Mahoney, M. W. Community structure in large networks: Natural cluster sizes and the absence of large well-defined clusters. *Internet Mathematics*, 6(1):29–123, 2009. doi: 10.1080/15427951.2009.10129177. 82

[103] Li, M., Wang, D., Fan, Y., Di, Z., and Wu, J. Modelling weighted networks using connection count. *New J. Phys.*, 8:72, 2006. doi: 10.1088/1367-2630/8/5/072. 9, 115

[104] Liben-Nowell, D. and Kleinberg, J. The link-prediction problem for social networks. *J. Am. Soc. Inf. Sci. Technol.*, 58(7):1019–1031, 2007. doi: 10.1002/asi.v58:7. 135

[105] Lü, L. and Zhou, T. Link prediction in weighted networks: The role of weak ties. *Europhys. Lett.*, 89(1):18001, 2010. 135

[106] Mandelbrot, B. B. *The Fractal Geometry of Nature*. Freeman, San Francisco, 1982. 79

[107] Maslov, S., Sneppen, K., and Zaliznyak, A. Detection of topological patterns in complex networks: Correlation profile of the internet. *Physica A*, 333(1-4):529–540, 2004. doi: 10.1016/j.physa.2003.06.002. 5

[108] Mastrandrea, R., Squartini, T., Fagiolo, G., and Garlaschelli, D. Enhanced reconstruction of weighted networks from strengths and degrees. *New J. Phys.*, 16:043022, 2014. doi: 10.1088/1367-2630/16/4/043022. 9, 115

[109] McFarland, D. D. and Brown, D. J. Social distance as a metric: a systematic introduction to smallest space analysis. *EO Laumann. Bonds of Pluralism: The Form and Substance of Urban Social Networks. New York: John Wiley*, 213–252, 1973. 14

[110] McPherson, M., Smith-Lovin, L., and Cook, J. M. Birds of a feather: Homophily in social networks. *Annu. Rev. Sociol.*, 27(1):415–444, 2001. doi: 10.1146/annurev.soc.27.1.415. 14

[111] Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N., Chklovskii, D., and Alon, U. Network motifs: simple building blocks of complex networks. *Science*, 298(5594):824–827, 2002. doi: 10.1126/science.298.5594.824. 5, 6

[112] Milo, R., Itzkovitz, S., Kashtan, N., Levitt, R., Shen-Orr, S., Ayzenshtat, I., Sheffer, M., and Alon, U. Superfamilies of evolved and designed networks. *Science*, 303(5663):1538–1542, 2004. 82

[113] Moreno, Y. and Pacheco, A. F. Synchronization of kuramoto oscillators in scale-free networks. *Europhys. Lett.*, 68(4):603, 2004. 98

[114] Muscoloni, A. and Cannistraci, C. V. A nonuniform popularity-similarity optimization (npso) model to efficiently generate realistic complex networks with communities. *New J. Phys.*, 20:052002, 2018. 35

[115] Muscoloni, A., Thomas, J. M., Ciucci, S., Bianconi, G., and Cannistraci, C. V. Machine learning meets complex networks via coalescent embedding in the hyperbolic space. *Nat. Commun.*, 8:1615, 2017. doi: 10.1038/s41467-017-01825-5. 14, 21, 31

[116] Newman, M. E. J. Modularity and community structure in networks. *Proc. Natl. Acad. Sci. USA*, 103:8577–8582, 2006. 70

[117] Newman, M. E. J. and Girvan, M. Finding and evaluating community structure in networks. *Phys. Rev. E*, 69:026113, 2004. doi: 10.1103/PhysRevE.69.026113. 6, 34

[118] Newman, M. *Networks: An Introduction*. Oxford University Press, Inc., New York, NY, USA, 2010. ISBN 0199206651. 2

[119] Newman, M. and Watts, D. Renormalization group analysis of the small-world network model. *Phys. Lett. A*, 263(4–6):341 – 346, 1999. doi: 10.1016/S0375-9601(99)00757-4. 79

[120] Newton, I. *Philosophiae naturalis principia mathematica*. J. Societatis Regiae ac Typis J. Streater, 1687. 1

[121] Newton, I. *A Treatise of the System of the World*. Fayram, 1731. 1

[122] Nokkala, J., Arzani, F., Galve, F., Zambrini, R., Maniscalco, S., Piilo, J., Treps, N., and Parigi, V. Reconfigurable optical implementation of quantum complex networks. *New J. Phys.*, 20(5):053024, 2018. 142

[123] Nokkala, J., Galve, F., Zambrini, R., Maniscalco, S., and Piilo, J. Complex quantum networks as structured environments: engineering and probing. *Sci. Rep.*, 6:26861, 2016. 142

[124] Nokkala, J., Maniscalco, S., and Piilo, J. Non-markovianity over ensemble averages in quantum complex networks. *Open Syst. Inf. Dyn.*, 24(04): 1740018, 2017. doi: 10.1142/S1230161217400182. 142

[125] Openflights network dataset – KONECT, September 2016. URL `http://konect.uni-koblenz.de/networks/openflights`. 82

[126] Orsini, C., Dankulov, M. M., Colomer-de Simón, P., Jamakovic, A., Mahadevan, P., Vahdat, A., Bassler, K. E., Toroczkai, Z., Boguñá, M., Caldarelli, G., Fortunato, S., and Krioukov, D. Quantifying randomness in real networks. *Nat. Commun.*, 6:8627, 2015. 6

[127] Orth, J. D. et al. A comprehensive genome-scale reconstruction of *Escherichia coli* metabolism. *Mol. Syst. Biol.*, 7:535, 2011. 113

[128] Ortiz, E., Starnini, M., and Serrano, M. Á. Navigability of temporal networks in hyperbolic space. *Sci. Rep.*, 7(1):15054, 2017. doi: 10.1038/s41598-017-15041-0. 17

[129] Pajevic, S. and Plenz, D. The organization of strong links in complex networks. *Nat. Phys.*, 8:429–436, 2012. doi: 10.1038/nphys2257. 114, 115

[130] Papadopoulos, F., Psounis, K., and Govindan, R. Performance preserving topological downscaling of internet-like networks. *IEEE J. Sel. Area. Comm.*, 24(12):2313–2326, 2006. doi: 10.1109/JSAC.2006.884029. 94

[131] Papadopoulos, F., Krioukov, D., Boguñá, M., and Vahdat, A. Greedy forwarding in dynamic scale-free networks embedded in hyperbolic metric spaces. In *INFOCOM*, 2010. 16

[132] Papadopoulos, F. and Psounis, K. Efficient identification of uncongested internet links for topology downscaling. *SIGCOMM Comput. Commun. Rev.*, 37(5):39–52, 2007. doi: 10.1145/1290168.1290173. 94

[133] Papadopoulos, F., Kitsak, M., Serrano, M. A., Boguñá, M., and Krioukov, D. Popularity versus similarity in growing networks. *Nature*, 489(7417): 537–540, 2012. 13, 35

[134] Papadopoulos, F., Aldecoa, R., and Krioukov, D. Network geometry inference using common neighbors. *Phys. Rev. E*, 92:022807, 2015. doi: 10.1103/PhysRevE.92.022807. 14, 31, 129

[135] Pastor-Satorras, R. and Vespignani, A. Epidemic spreading in scale-free networks. *Phys. Rev. Lett.*, 86(14):3200–3203, 2001. 95

[136] Popović, M., Štefančić, H., and Zlatić, V. Geometric Origin of Scaling in Large Traffic Networks. *Phys. Rev. Lett.*, 109:208701, 2012. doi: 10.1103/PhysRevLett.109.208701. 6, 111, 114

[137] Radicchi, F., Castellano, C., Cecconi, F., Loreto, V., and Parisi, D. Defining and identifying communities in networks. *Proc. Natl. Acad. Sci. USA*, 101: 2658–2663, 2004. 6, 34

[138] Radicchi, F., Ramasco, J. J., Barrat, A., and Fortunato, S. Complex networks renormalization: Flows and fixed points. *Phys. Rev. Lett.*, 101:148701, 2008. doi: 10.1103/PhysRevLett.101.148701. 79, 103

[139] ROC, 2015. ROC's Bureau of Foreign Trade. [online] Available at `http://www.trade.gov.tw/English/`. 55

[140] Rolland, T., *et al.* A proteome-scale map of the human interactome network. *Cell*, 159(5):1212 – 1226, 2014. doi: 10.1016/j.cell.2014.10.050. 82

[141] Ronhovde, P. and Nussinov, Z. Multiresolution community detection for megascale networks by information-based replica correlations. *Phys. Rev. E*, 80:016109, 2009. doi: 10.1103/PhysRevE.80.016109. 83

[142] Rosson III, C. P., Runge, C. F., and Hathaway, D. E. International trade agreements. In Hallberg, M., Sptize, R., and Ray, D., editors, *Food, Agriculture, and Rural Policy into the Twenty-First Century*. Westview Press, 1994. 77

[143] Rozenfeld, H. D., Song, C., and Makse, H. A. Small-world to fractal transition in complex networks: A renormalization group approach. *Phys. Rev. Lett.*, 104:025701, 2010. doi: 10.1103/PhysRevLett.104.025701. 79, 103

[144] Sagarra, O., Pérez Vicente, C., and Díaz-Guilera, A. Statistical mechanics of multiedge networks. *Phys. Rev. E*, 88:062806, 2013. doi: 10.1103/PhysRevE.88.062806. 9, 115

[145] Sagarra, O., Pérez Vicente, C. J., and Díaz-Guilera, A. Role of adjacency-matrix degeneracy in maximum-entropy-weighted network models. *Phys. Rev. E*, 92:052816, 2015. doi: 10.1103/PhysRevE.92.052816. 9, 115

[146] Serrà, J., Corral, A., Boguñá, M., Haro, M., and Arcos, J. L. Measuring the evolution of contemporary western popular music. *Sci. Rep.*, 2, 2012. doi: 10.1038/srep00521. 82

[147] Serrano, M. Á. Phase transition in the globalization of trade. *J. Stat. Mech.*, L01002, 2007. 49, 50, 51

[148] Serrano, M. Á. Rich-club vs rich-multipolarization phenomena in weighted networks. *Phys. Rev. E*, 78:026101, 2008. 54, 111

[149] Serrano, M. Á. and Boguñá, M. Topology of the world trade web. *Phys. Rev. E*, 68:015101, 2003. 49, 50

[150] Serrano, M. Á., Boguñá, M., and Pastor-Satorras, R. Correlations in weighted networks. *Phys. Rev. E*, 74:055101, 2006. doi: 10.1103/PhysRevE.74.055101. 123

[151] Serrano, M. Á., Boguñá, M., and Vespignani, A. Patterns of dominant flows in the world trade web. *J. Econ. Interact. Coord.*, 2(2):111–124, 2007. 49, 50, 51, 57

[152] Serrano, M. Á., Krioukov, D., and Boguñá, M. Self-similarity of complex networks and hidden metric spaces. *Phys. Rev. Lett.*, 100:078701, 2008. 9, 11, 22, 24, 26, 35, 40, 60, 114, 115, 116, 134

[153] Serrano, M. Á., Boguñá, M., and Vespignani, A. Extracting the mutiscale backbone of complex weighted networks. *Proc. Natl. Acad. Sci. USA*, 106:6483–6488, 2009. 52, 57, 82, 113

[154] Serrano, M. Á., Garlaschelli, D., Boguñá, M., and Loffredo, M. The world trade web: Structure, evolution and modeling. In Caldarelli, G., editor, *Complex Networks*, Encyclopedia of Life Support Systems (EOLSS). Eolss Publishers, Oxford, UK, 2010. 49, 50, 51

[155] Serrano, M. Á., Boguñá, M., and Sagués, F. Uncovering the hidden geometry behind metabolic networks. *Mol. Biosyst.*, 8:843–850, 2012. doi: 10.1039/c2mb05306c. 12, 14, 15, 16, 24, 25, 31, 34, 46, 82, 88, 113, 128

[156] Song, C., Havlin, S., and Makse, H. A. Self-similarity of complex networks. *Nature*, 433:392–395, 2005. 79, 103

[157] Song, C., Havlin, S., and Makse, H. A. Origins of fractality in the growth of complex networks. *Nat. Phys.*, 2:275–281, 2006. 79, 103

[158] Squartini, T. and Garlaschelli, D. Jan tinbergen's legacy for economic networks: From the gravity model to quantum statistics. In Abergel, F., Aoyama, H., Chakrabarti, K. B., Chakraborti, A., and Ghosh, A., editors, *Econophysics of Agent-Based Models*, 161–186. Springer International Publishing, Cham, 2014. ISBN 978-3-319-00023-7. doi: 10.1007/978-3-319-00023-7_9. 14, 49

[159] Stanley, H. E. *Introduction to Phase Transitions and Critical Phenomena*. Oxford Univ. Press, Oxford, 1971. 79

[160] Stiglitz, J. E. *Making globalization work*. W. W. Norton & Company, New York, London, 2006. 66, 77

[161] Takemura, S.-y., *et al.* A visual motion detection circuit suggested by drosophila connectomics. *Nature*, 500(7461):175–181, 2013. 82

[162] Tinbergen, J. *Shaping The World Economy: Suggestions For An International Economic Policy*. The Twentieth Century Fund, New York, 1962. 14, 49

[163] Travers, J. and Milgram, S. An Experimental Study of the Small World Problem. *Sociometry*, 32(4):425–443, 1969. doi: 10.2307/2786545. 5

[164] UN, 2015. United Nations Commodity Trade Statistics Database (UN COMTRADE). [online] Available at https://comtrade.un.org/db/. 54

[165] Wang, W.-X., Hu, B., Zhou, T., Wang, B.-H., and Xie, Y.-B. Mutual selection model for weighted networks. *Phys. Rev. E*, 72:046140, 2005. doi: 10.1103/PhysRevE.72.046140. 9, 115

[166] Watts, D. J. and Strogatz, S. H. Collective dynamics of "small-world" networks. *Nature*, 393:440–442, 1998. 3, 8, 79

[167] WB, 2015. The World Bank open data. [online] Available at http://data.worldbank.org/. 55, 56

[168] Wilson, K. G. The renormalization group: Critical phenomena and the kondo problem. *Rev. Mod. Phys.*, 47(4):773–840, 1975. 79

[169] Wilson, K. G. The renormalization group and critical phenomena. *Rev. Mod. Phys.*, 55:583–600, 1983. doi: 10.1103/RevModPhys.55.583. 79

[170] WTO, 2015. World Trade Organization, Regional Trade Agreements Information System (RTA-IS), [online] Available at http://rtais.wto.org/UI/PublicMaintainRTAHome.aspx. 56, 71

[171] Wu, Z., Menichetti, G., Rahmede, C., and Bianconi, G. Emergent complex network geometry. *Sci. Rep.*, 5:10073, 2015. 139

[172] Yao, W. M. and Fahmy, S. Downscaling network scenarios with denial of service (dos) attacks. In *2008 IEEE Sarnoff Symposium*, 1–6, 2008. doi: 10.1109/SARNOF.2008.4520099. 94

[173] Yao, W. M. and Fahmy, S. Partitioning network testbed experiments. In *2011 31st International Conference on Distributed Computing Systems*, 299–309, 2011. doi: 10.1109/ICDCS.2011.22. 94

[174] Yook, S., Jeong, H., Barabási, A.-L., and Tu, Y. Weighted Evolving Networks. *Phys. Rev. Lett.*, 86:5835–5838, 2001. doi: 10.1103/PhysRevLett.86.5835. 9, 115

[175] Zhao, J., Miao, L., Yang, J., Fang, H., Zhang, Q.-M., Nie, M., Holme, P., and Zhou, T. Prediction of links and weights in networks by reliable routes. *Sci. Rep.*, 5:12261, 2015. 135

[176] Zheng, D., Trimper, S., Zheng, B., and Hui, P. M. Weighted Scale-Free Networks with Stochastic Weight Assignments. *Phys. Rev. E*, 67:040102(R), 2003. doi: 10.1103/PhysRevE.67.040102. 9, 115

[177] Zuev, K., Boguñá, M., Bianconi, G., and Krioukov, D. Emergence of soft communities from geometric preferential attachment. *Sci. Rep.*, 5:9421, 2015. 13, 35, 36, 37, 39, 88