UNIVERSITA' DEGLI STUDI DI PARMA

Dipartimento di Fisica e Scienze della Terra
Macedonio Melloni

Dottorato di ricerca in Fisica
Ciclo XXVIII

Asymptotic theory of time varying networks with memory and heterogeneous activation pattern

Coordinatore:
Chiar.mo Prof. Cristiano Viappiani

Tutor:
Chiar.ma Prof. Raffaella Burioni

Dottorando:
Enrico Ubaldi

Anno Accademico 2015-2016

Asymptotic theory of time varying networks with memory and heterogeneous activation pattern

Candidate: Enrico Ubaldi
Coordinator: Prof. Cristiano Viappiani
Tutor: Prof. Raffaella burioni

January 2016

Abstract

The recent availability of large-scale, time-resolved and high quality digital datasets allowed for a deeper understanding of the structure and properties of many real-world networks. The underlying complexity found in such systems called for an appropriate framework to systematically analyze and characterize their properties. To this extent the recently founded field of network-science provided the necessary theoretical grounding to tackle the study and characterization of the relevant properties of both complex networks and dynamical processes undergoing on top of them. Particular attention has been devoted to comprehend the structure and the dynamics of social interactions in real-world complex networks. Moreover, the empirical evidence of a temporal dimension in networks prompted the switch of paradigm from a static representation of networks to a time varying one.

In this work we focus on the framework of Time-Varying-Networks in real world social systems. By means of the Activity-Driven paradigm (a modeling tool belonging to the family of Time-Varying-Networks) we define and develop a dynamical model mimicking the evolution of real-world social networks.

The developed framework allows for the encoding of generative mechanisms shaping the large scale properties of real-world networks. Among all of these mechanisms, three of them seem to play a central role in the social networks' evolution: the individual's propensity to engage in social interactions, its social capital allocation and the burstiness of interactions amongst social actors.

The former mechanism accounts for the temporal dependence of interactions between individuals. This is a fundamental ingredient to consider so as to describe and characterize both the topological and the diffusion properties of social networks, as human interactions are continuously rearranged in time. Among the possible ways to implement a temporal dimension on the topology of the network, the activity-driven-Network framework encodes the propensity for an individual i to engage a social interaction in a constant activity potential a_i. The latter measures the probability per unit time for the node to interact with another agent on the network.

The second ingredient (social capital allocation) accounts for the correlations between subsequent events of a single node. In other words, the social capital allocation of each agent is not random and people tend to distribute their social events and attention toward already contacted nodes (strong ties), reserving a smaller amount of their social activity towards less frequently activated edges (weak ties). To empirically measure and analytically model a mechanism able to reproduce the real-world behavior of people is a relevant problem in social science. The importance of a solution of such a task is not just of academical interest but it also allows for a better understanding on diffusion processes on social networks (e.g. epidemic and rumor spreading processes) and on how to control them (e.g. avoid a pandemic outbreak, infer the diffusion path of a piece of information etc.).

We base our study and characterization of social capital allocation extending a simple mechanism aiming at the description of edge creation that has been recently proposed. The latter captures the mechanism with a simple reinforcement rule, that inhibits the creation of new ties once a node gets active, thus favoring the reinforcement of the already established connections. To this extent we present a thorough analysis of social capital allocation mechanism in 7 real-world datasets. Then, given the results of the empirical analysis we develop a data-driven model and analytically solve it. We then present a robust check of the results against both numerical simulations and real-world datasets.

The final ingredient (burstiness) accounts for the non-poissonian temporal activation pattern of social ties in real-world networks. Indeed, social interactions displays an heterogeneous, heavy-tailed distribution of the waiting time w (or inter-event time) that can be reasonably described by a power law. This property is the fingerprint of the bursty nature of human dynamics. When developing a model or analysis of the long-time evolution and diffusion properties of social networks, one has to consider this mechanism, as it introduces strong heterogeneities in the networks' temporal dynamics. We then further develop the activity driven model so as to account for both the reinforcement process and the bursty activation of ties. The latter is introduced by its simplest formulation, i.e. by imposing an inter event time distribution following a power-law. We analytically solve the system dynamics in the asymptotic limit, finding a non trivial phase diagram of the network evolution due the interplay of the two mechanisms: there are situations in which the burstiness governs the evolution of the network and others where the dynamics is completely determined by the social capital allocation process. Also in this case, we test our analytical results against numerical simulations and empirical datasets, finding a good agreement between the two.

In the last chapter we then present an additional attempt to develop a complete dynamical model of Time-Varying-Networks that is based on the Kauffman's *adjacent-possible* theory. This model overcomes some limitations of the previously proposed models. For instance, it accounts for the growing size of social network and naturally introduces some of the real-world networks' topological properties. Furthermore, the model naturally introduces heterogeneous behaviors in the system, and it is also able to reproduce the empirical functional form of the edges reinforcement process. Even in this case we present an analytical solution of the model and we check our results with numerical simulations comparing them with the real-world networks under investigation.

Contents

Preface

The work presented in this dissertation was carried between the Physics and Earth Sciences Department at the University of Parma and the MOBS laboratory at the Northeastern University in Boston in the period between January 2013 and December 2015. Part of this work has been also done at Institute for Scientific Interchange (ISI) in Turin. I thank these institutions for their kind hospitality.

Most of the work presented in this thesis has been done within different scientific collaborations, whose members I kindly acknowledge for giving me the possibility and the honor to work with them.

Chapters 4 and 5 are the result of different collaborations with Nicola Perra, Màrton Karsai, Alessandro Vezzani, Kayuan Sun and Alessandro Vespignani. They are based on the following papers:

Asymptotic theory for the dynamic of networks with heterogeneous social capital allocation, arXiv:1509.04563, 2015.

The interplay between burstiness and social capital allocation in social networks (to be submitted).

Epidemic spreading in modular time-varying networks (to be submitted).

Chapter 6 is the result of a collaboration with Francesca Tria and Vittorio Loreto and it is based on a paper in preparation.

Introduction and motivations

The research activity on complex systems is an important and still growing field of statistical mechanics. The methods belonging to this research area are being implemented and applied to many research fields spanning from economy to medicine. Moreover, in the recent years a lot of attention has been devoted to the continuously growing field of network-science, as many real-world systems can be modeled as graphs. Also in this case, many of these equivalent networks displayed complex features, thus giving the birth to the complex networks term. The interdisciplinary nature of such a filed calls for a universally accepted definition of what we are referring to when we talk about complex networks or *complex system* in general.

Firstly, it is important to stress that complex does not mean something merely complicated. We are surrounded by complicated systems and object every day: the computer we use at work, the car that drives us home and even our phones are composed by many elements designed to interact and perform different tasks. Nevertheless they are really merely complicated as they are engineered accordingly to a precise, supervised blueprint. This is where they differ from a *complex system* [1]. A complex system is an emergent phenomena, as it is a spontaneous outcome of the interactions amongst the many units composing it: there is no supervision, no blueprint and no engineering. In other words, complex systems consist of a large number of elements capable of interacting with each other and the surrounding environment in order to organize in specific emergent structure, so that their behavior cannot be described from simple extrapolations of the properties of their constitutive units. One usually resume this feature with the well-known phrase by P.W. Anderson *"More is different"* [2]: the same elements assembled in large number can cause different macroscopic and dynamical behaviors, and holistic approaches have to be used.

Another peculiar property of complex systems are their scale-free structure and fluctuations. The former is usually visualized with the complex structure of fractals object in which the same level of complication is found at all the possible scales at which we inspect the element [3]. The non-linearity of diverging fluctuations is instead summarized with the renowned *"Butterfly effect"*: a small change (the flapping of a butterfly's wings) can lead to a big difference (a storm) later. This resumes in a nutshell the mechanism of critical phenomena where macroscopic rearrangements across the entire system are generated by the combined action of infinitesimal localized interactions [1, 4].

Now we are ready to give a possible definition of what a complex network is. As already pointed out, a great variety of systems in nature, economy, society and

technology can be modeled as networks. In the latter, the vertices represent the single agents belonging to the system, while the edges connecting them represent a relation or interaction among those elements. If, after this abstraction has been made, the complex features are still present we will refer to these complex systems as complex networks [5]. This results applies to a huge number of systems belonging to the most disparate fields, from the net of sexual contacts to the nervous system, from the social interactions among individuals to the biological interactions between proteins in the cells. The mathematical framework is based on Graph Theory [6, 7], one of the fundamental parts of discrete mathematics founded by Euler in 1736 with the famous solution of the Königsberg bridges problem.

To support the interest that mathematicians and physicists put in the research on complex networks, the unprecedented availability of large, high-quality digital datasets allowed the network-science field to deepen our understanding on the structure and dynamical properties of these complex networks. In particular, the apparently ubiquitous heavy-tailed distributions characterizing the structural properties revealed the all-scale complications typical of complex systems. These heterogeneities appear to be common features of a large number of real-world networks, along with other complex topological properties. It is now acknowledged that the complex topology and dynamics on networks is the natural outcome of the evolution of such networks. This evidence is far from being just accidental and raises the question of the existence of some general organizing principles determining the emergence of common properties in such very different contexts. This consideration led the network-science community to switch the focus to the microscopic processing governing the arrangement, appearance and disappearance of vertices and links.

To this extent, a first round of studies focused on a static representation of networks. In the latter, the nodes and the edges connecting them are fixed and immutable in time. In this framework the focus is on the connectivity properties of nodes and dynamical processes are naturally implemented on top of the underlying network fabric. The significant results found in real-world datasets taught us that many networks show a small but significant number of nodes, called hubs, whose degree (number of contacts) is larger than the average value. Indeed, we expect social networks to contain a few popular individuals with many acquaintances, the WWW to feature a few very popular websites cumulating a huge number of incoming links. Moreover in biological systems there are just a few metabolites that take part in most of the metabolic processes, there is a minute fraction of airports collecting a huge number of connecting flights etc.. In these networks there is no characteristic scale: this is why they are often called *scale-free* networks.

Besides their heterogeneous nature, real-world networks feature the so called *small-world* effect: the average distance between pairs of nodes is very small and typically grows as the logarithm (or slower) of the network size (i.e. its number of vertices), in contrast with what happens in regular lattices. This feature is well known in the sociological context as the six degrees of separation after the pioneering works of Milgram in 1967 [8].

As a last remark on static networks, we recall the tendency of nodes to form

groups of densely interconnected elements, namely communities. This properties is usually revealed by a high clustering coefficient.

Given all these non-trivial features, many efforts were done in order to find a model that could generate all these peculiar properties at once. The (not-yet completed) trail to this goal started with the simplest model proposed by Erdös and Rènyi in the 60s [9–11]. In their work they proposed the simplest implementation of a *random graph*: each pair of nodes is connected with a given probability p. Though far from reality, the model generates a random networks with a pronounced small-world effect. However, a very small clustering coefficient is found and an heterogeneous (Poissonian) connectivity distribution is found.

To overcome the low-clustering coefficient, in 1998 Watts and Strogatz introduced a new mechanism able to generate both a small-world effect and a high clustering coefficient [12]. Despite the progress made, the skewed, heavy tailed connectivity distribution was still missing.

In 1999 A. L. Barabási and R. Albert proposed a novel model published in one of the most cited papers of the last decades [13]. In the paper they proposed a generative model for networks that provides an heavy-tailed connectivity distribution, high clustering and small-world effect. The intuition was that, in order to obtain a skewed degree distribution, there must be a multiplicative process governing the nodes attachment and edges cannot be connected randomly. To this end, the model states that nodes tend to connect to vertices which already have a large number of connections: preferential attachment. After the introduction of this model, many others have been proposed as variations of this mechanism and helped to understand the processes shaping real-world networks. Besides, a great endeavor has been put on the characterization of how the heterogeneous topology and the complex nature of these networks affect the dynamical processes undergoing on top of them.

On the other hand, the effort put in order to understand the evolution of complex networks (and social networks in particular) generated a lot of attention to the mechanism governing the creation, arrangement and deletion of links. A second round of works in network-science revealed the presence of a temporal dimension in social networks, giving rise to the Time-Varying-Networks (TVN) paradigm. This advance moved the focus of the research work from the structure of the network to the growth of the network itself. Indeed, in the first network-science works the object of the study was the network topology (the vertices and edges) and a dynamical process undergoing on the network was modeled and implemented separately in a top-down approach. Compared to this picture, TVN approach moves information about when things happen from the dynamical system to the network, the structure on which the dynamics actually takes place, a more bottom-up way of thinking.

Systems that can be modeled as temporal networks are everywhere, as we expect contact in social and natural environments not to be always active and available to the agents of the underlying web: we are not (luckily) always at the phone with our friends, the spread of biological viruses and diseases (again, luckily) is not always possible as contacts between infected and susceptible nodes are not always active. This new approach poses some fundamental question: how does the temporal di-

mension affect the dynamical processes on the network (e.g. reaction diffusion ones) and how can we encode the dynamics on the network itself? In this work we focus on a model belonging to a family of the TVN models class, i.e. an *activity-driven* (AD) model of network-generation. Though there are many ways to implement a temporal dimension in networks, the AD approach allows for a rigorous analytical approach whose results gives accurate prediction on the network evolution.

In the AD approach each node is given an activity potential, i.e. the propensity for the node to engage a social interaction with another, randomly chosen actor in the network. The interaction is instantaneous, meaning that the edge is created and deleted before the next interaction on the network takes place. This simple prescription strongly modifies both the topological and the dynamical properties of the network. Indeed, the instantaneous network features only one edge between the two active nodes at that time. It is then customary to integrate these interactions over a time window Δt to get an *integrated-network* and then let the dynamical process evolve on this aggregated network (also called *snapshot*). Given the temporal dependence of the edges, the sparser nature of the network and the different connectivity patterns of subsequent snapshots, the behavior of the dynamical processes undergoing on top of the network fabrics is deeply modified with respect to the same processes evolving on static networks: the first passage time of a random walker on a node varies, as well as the point at which an epidemic process does not die out anymore and results in an outbreak.

The AD framework introduces a dynamics on the edge creation but it does not tell anything on how, in a social network, the social acts of an agents are arranged amongst its neighbors. A second question then arises: how do people invest their social capital? In other words, is it possible to measure and define a mechanism able to reproduce the growth of real-world ego-nets? As in the previous points, the recent and unprecedented availability of large and high quality datasets helps us to answer these questions. In this work we tackle this problems by applying a *data-driven* approach. We measure and characterize the social capital allocation in real-world datasets and then propose and implement a reinforcement process on top of an AD model. As reasonably expected, we find that social interactions are not randomly arranged in a social network but they are rather concentrated towards the closest neighbors of the node. This behavior can be resumed in the fact that we are more prone to interact with already known pals, through already established ties, rather than always connect or contact a new node in the network. As for the switch from static to TVN, also the introduction of a reinforcement process affects both the topology of the network and the dynamical properties of the reaction-diffusion processes unfolding on top of it. In particular, the critical properties of different epidemics processes (e.g. SIR and SIS) change in a different way when reinforcement is introduced. On the topology side, the reinforcement process inhibits the creation of new edges thus slowing down the growth in time of the average degree of a node. If this growth was linear in the random allocation model, now we find this growth to be sub-linear.

An additional level of complexity found in real-world systems is the supplementary heterogeneity of the time scales of the inter-event time, i.e the time between two consecutive interactions of a single node. Again, thanks to the constantly increasing possibility to access the time-resolved activity of most components of selected complex systems, such as time-resolved mobile phone calls, Twitter citations, and scientific collaborations, we find that the temporal activity patterns is bursty in many systems. Burstiness then appears to be an ubiquitous property of complex, real-world systems. In few words, burstiness can be described in term of bursts of activity, i.e. a significantly enhanced activity levels over short periods of time followed by long periods of inactivity. In particular, in human dynamics burstiness has been reduced to the fat-tailed nature of the response time distribution or the inter-event time between two successive interactions of the same user. Once present, burstiness can affect the spreading of diseases or resource allocation. In the AD model we can consider the system events as an activity pattern that can be thought of as a discrete signal. In this formalism the activity pattern follows a Poisson process, as the probability of an event is time-independent. In this case the inter-event time between two consecutive events follows an exponential distribution. To account for the real-world heterogeneous inter-event time distributions, we implement such a waiting time distribution on top of the AD model with reinforcement finding a non-trivial interplay between the two mechanisms. Indeed, the competition between "memory" and burstiness breeds a rich phase diagram composed by the parameters characterizing the two mechanisms. In this work we present the topological characterization of the networks generated accordingly to this improved model, leaving the dynamical characterization for future work. It is clear, however, that the introduction of burstiness will further modify the behavior of the reaction-diffusion processes' dynamics. Indeed, the heavy tails of the inter-event time distribution are further slowing down the contact rate between nodes.

Despite the intrinsic importance of the modeling tools and efforts made to catch the real-world networks evolution, all of the proposed and analyzed models do share a limitation: the heterogeneity (either on the activity of a node or on the inter event time distribution) has to be put by hand or by enforcing a preferential attachment rule. For instance, in the AD model we have to externally enforce the activity distribution, when we add the reinforcement process we have to define its functional form and implement it in the numerical simulations, and when adding burstiness to the system we enforced a skewed a waiting time distribution. It is then a comprehensible objection to ask whether or not is it possible to develop a fundamental, *ab-initio*, model that could generate by itself some (if not all) of the observed heterogeneity and social actors' behavior found in real-world systems. To this extent, we investigate how the social exploration of an individual determines his next step in the process of social capital allocation. In other words, we look at how a novelty (i.e. a new contact) introduced in the ego-net of a social actor (i.e. the set of the nodes that have been in contact with the node) regulates the target of the next social activities of the node itself.

It is reasonable to expect that novelties may lead our social life, as our daily lives are spiced with little novelties: to hear a new song, taste a new food, meet a new

friend. Sometimes, our contact with a specific novelty may correlate with a later novelty, as one of these first-time experiences may sparks another. For example, meeting a person we enjoy may prompt us to spend more time with him/her and eventually meet him/her friends as well. The notion that one new thing sometimes triggers another is, of course, commonsensical. But a recent work, Tria et. al [14] quantitatively show that, indeed, everyday novelties truly do pave the way for other future novelties.

In this approach, novelties are then seen as an event able to extend the so called *adjacent possible*. This mechanism falls in the framework of Kauffman's *adjacent-possible* theory and has successfully been applied to the individuals' exploration of ideas and information, generating the so called *dynamics of correlated novelties*. In their work, Tria et al. [14] modeled the dynamics of novelties encountered by a single agent by means of a single Polya's urn.

In the last chapter of this thesis, we generalize this approach to define a dynamical model of the social networks' evolution . The adjacent possible is now the social circle of a person plus all the neighbors of its friends. To get in contact with an agent never contacted before results in the expansion of the space available for the social exploration of the agent. Surprisingly the introduced model accomplish to reproduce many of the desired properties of real-world social networks: the growth in size of a network with time, the measured reinforcement mechanism, the heterogeneous activity and degree distributions and the high value of the clustering coefficient.

To sum up, we showed many challenging, difficult and intriguing questions that motivate this work. Evaluating the mechanisms at the base of the complex-networks formation and to analytically predict their behavior are two deeply related problems that can be attacked within one general approach well grounded in Physics, namely the multi-agent approximation and the Master Equation formalism. We present a walk, not random, into the complexity and in the holistic Physics perspective. In particular Chapter 1 is dedicated to the introduction of basic concepts of graph theory, while in Chapter 2 we present real-world networks and their basic generative models. In Chapter 3 the general framework of Time-Varying-Networks is introduced, with particular focus on the possible modeling approaches. In Chapter 4 we introduce our Activity-Driven model and deeply investigate its properties together with the effect of the heterogeneous activity and the reinforcement process on the network evolution and dynamics. We then show the numerical results validating our analytical approach and we test the model against real-world datasets, finding a very good agreement between the analytical predictions and the empirical data. We also present a preliminary study on how the time-varying nature of the system affect an epidemics process, focusing on a realization of he network where the reinforcement process is introduced by means of a community structure of the network.

Then, in Chapter 5 we further develop the model introduced in the previous Chapter to include burstiness in our analysis. Again, a thorough analytical characterization of the system is carried on and we test our predictions with both numerical simulations and empirical data. We also analyze the phase space created by the

interplay of the reinforcement mechanism and the bursty behavior.

Finally, in Chapter 6 we introduce the basic model based on the theory of the adjacent possible, by means of a generalized Polya's urn. We solve the model in the asymptotic limit and we discuss the properties of the generated network. We then give a preliminary recipe to map this model to the corresponding activity-driven version.

Chapter 1

Introduction to Network Theory

Networks are the basic ingredient of this work. In this Chapter we introduce the principal concepts of networks and graph theory. We define different type of networks and the basic measures and quantities under examination: degree, shortest path length, clustering coefficient, betweenness and all their statistical properties.

1.1 Network definition and basic concepts

A network can be thought as any systems that allows an abstract representation as a graph whose nodes (vertices) correspond to the elements composing the system and whose edges (links) identify the relations occurring between the elements. As a matter of fact, many real systems can be modeled as networks. The theoretical framework of networks thus provides a convenient and powerful representation of the interactions in complex systems where the mapping of the interactions among a huge number of individuals determines the properties of the systems themselves.

Research on networks has a long tradition that ideally started with the work of Leonhard Euler in solving the Königsberg bridges problem in 1736. Since then a lot of attention has been put on the study of networks as this subject is found in graph theory, sociology, physics, communication research, biology and a large bibliography of textbooks is found in literature [6, 7, 15–18].

1.1.1 Vertices and edges

As a first insight, let us recall that the simplest definition of a graph G is the undirected one. In the latter we define an undirected graph G as the pair of sets $G = (\mathcal{V}, \mathcal{E})$, where $\mathcal{N}$ is a set of vertices (the elements or the nodes), and $\mathcal{E}$ is a set of *unordered* pairs of edges (the links or edges). An edge (i, j) then connects the two vertices i and j which are said to be *adjacent* or *connected* (or, in a more common way, i and j are dubbed neighbors). The size (cardinality) of the $\mathcal{N}$ set is the number of vertices N, denoting the order of the graph G. This defines the physical size of many real systems as N denotes the number of agents found in the network under study. The size E of the $\mathcal{E}$ set is the number of edges and, formally

speaking, defines the size of the graph G. The maximum number of edges E is $\binom{N}{2}$. A graph containing exactly $\binom{N}{2}$ edges is called a *complete N-graph*.

On the other hand, a directed graph D has the same definition as the undirected one exception made for the pairs representing the edges that are now *ordered*. In this case an edge (i, j) points from the source node i to the target node j. The connection between the two nodes is possible only from i to j and not the other way around (unless also the (j, i) edge is present in $\mathcal{E}$). In the undirected case however, an edge (i, j) connects the two nodes in both the directions.

To graphically depict a graph one usually print a set of dots (the vertices) and join them by lines between the pairs of connected vertices (the links). From a mathematical point of view it is convenient to define a graph by means of the adjacency matrix $\mathbf{A} = a_{ij}$, i.e. a $N \times N$ matrix whose entries are:

$$a_{ij} = \begin{cases} 1 & \text{if } (i,j) \in \mathcal{E} \\ 0 & \text{if } (i,j) \notin \mathcal{E}. \end{cases} \tag{1.1.1}$$

In an undirected graph $\mathbf{A}$ is symmetric (i.e. $a_{ij} = a_{ji}$) while in the directed case the first and second indexes correspond to the source and the target, respectively.

We then define the sparseness of a graph to quantify the density of connections on the network itself. If the number of edges E scales with the number of nodes N as $E \sim N^{\alpha}$ with $\alpha < 2$ we call the graph *sparse*, otherwise the graph is said to be *dense*. In other words, we can define the connectance or density of a graph $\mathcal{D}$ as the ratio between the number of edges E divided by the maximum possible number of edges $\mathcal{D} = E/[N(N-1)/2]$. A graph is then sparse if $\mathcal{D} \ll 1$. As the adjacency matrix's $\mathbf{A}$ entries of a sparse graph are mostly 0, the most convenient way to mathematically (and numerically) represent a sparse graph is by the adjacency list $l(i, \nu \in \mathcal{N}(i))$. The latter is a list including all the first neighbors of node i and it is usually named *neighborhood* or, in social networks, the *ego-net* of node i.

An important characterization usually done on large graphs is the one concerning local structures as *cliques* or communities. These are defined as highly interconnected subgraphs that are poorly connected to the nodes outside the community itself. This characterization leads to the study of network modularity and clustering techniques aimed at the individuation of major cluster in a given network [19–25].

1.1.2 Paths and distance

Another important feature of graphs and complex networks in general is the reachability of vertices, i.e. to determine whether or not it exists a path connecting one vertex to another within a given graph. One can easily tackle this problem by means of the adjacent matrix $\mathbf{A}$. A path $\mathcal{P}_{i_0,i_n}$ connecting i_0 to i_n and of length n in a graph $\mathcal{G} = (\mathcal{N}, \mathcal{E})$ is defined as an ordered collections of $n+1$ vertices $\mathcal{N}_{\mathcal{P}} = \{i_0, i_1, \ldots, i_n\}$ and n edges $\mathcal{E}_{\mathcal{P}} = \{(i_0, i_1), (i_1, i_2), \ldots, (i_{n-1}, i_n)\}$, so that $i_\alpha \in \mathcal{E}$ and $(i_{\alpha-i}, i_\alpha) \in \mathcal{N}$ for every α. The number $\mathcal{N}_{ij}$ of paths of length n connecting the two nodes i_0 and i_n is given by the a_{ij} element of the n-th power of the adjacency matrix, i.e. $\mathcal{N}_{ij} = (\mathbf{A}^n)_{ij}$. Based on these definitions, a graph is said to be connected if it exists at least one path connecting any two vertices on the web. There could be groups of

nodes that are connected among themselves (components) and two components $\mathcal{C}_1$ and $\mathcal{C}_2$ are disconnected if there is no path from a node i in the first component to a node j belonging to the second one. In a random graph, a component that scales with the size of the network N (thus diverging in the $N \to \infty$ limit) is called the *giant-component* $\mathcal{G}$.

The notion of path naturally introduces the concept of distance between nodes. For instance, the distance between two nodes i and j is defined as the number of edges l_{ij} composing the shortest path connecting i and j. To give a complete definition of distance, two vertices belonging to two disconnected components have infinite distance $l_{ij} = \infty$. Moreover, distance is symmetric in directed graphs ($l_{ij} = l_{ji}$), while this is not true in general in undirected graphs. An additional application of the notion of distance is the definition of a graph's $\mathcal{G}$ *diameter*, d_G. The latter is defined as the maximum distance between two nodes, i.e.

$$d_G = \max_{i,j \in \mathcal{N}} l_{ij}. \tag{1.1.2}$$

By means of Eq. (1.1.2) we can also define the *average shortest path length* $\langle l \rangle$ as the average value of l_{ij} over all the pairs of nodes in the network. It is evident that $\langle l \rangle \leq d_G$ by definition.

Trivial examples of characteristic distances in graphs include the complete graph ($\langle l \rangle = 1$), the hypercubic lattice in D dimensions ($\langle l \rangle \sim N^{1/D}$) and the majority of random graphs where the average shortest path length scales as the logarithm of the network size N, i.e. $\langle l \rangle \sim \ln N$. The last results enlighten the slower growth of the average distance in random graphs with respect to regular lattices. This phenomenon steams from the fact that two randomly chosen vertices are connected by a small shortest path even at a large network size and constitutes the so-called *small-world effect* [8, 26, 27].

Though important, distance measures give us global information about the network structure but does not tell us a lot about the importance or *centrality* that nodes have in the network structure. The definition of centrality usually depends on the relevant quantities related to the process or task we want our network to fulfill, so that many definitions of centrality are found in literature. However, the most commonly used ones are the degree, closeness or betweenness centralities of a given vertex. For what concerns edges, their centrality is usually measured in terms of their betweenness centrality.

The simplest definition of centrality is the *degree centrality* as it depends on the degree k_i of the node i alone. The degree k_i is defined as the number of vertices incident on the vertex i. This definition of degree centrality simply measures how well a node is connected to other vertices in the network.

Another possible definition of centrality is the *closeness centrality*. The latter expresses the average distance of a vertex i to all others as

$$g_i = \frac{1}{\sum_{j \neq i} l_{ij}}. \tag{1.1.3}$$

As g_i is inversely proportional to the average distance of a node from the other vertices, large closeness centrality is given to nodes well connected through short paths to the other nodes.

The two measures of centrality proposed so far are somewhat limited as they completely overlook the importance that a node may have in connecting different regions or components of a network. To overcome this restraint, a more powerful definition of centrality is the renowned *betweenness centrality* b_i of a node i [28,29]. This centrality is defined as the number of shortest paths between all the pairs of vertices that pass through a given vertex. To formally define b_i, let us introduce σ_{hj} as the total number of shortest paths from h to j and $\sigma_{hj}(i)$ as the number of these paths passing through i. The betweenness centrality b_i for node i is then defined as:

$$b_i = \sum_{h \neq j \neq i} \frac{\sigma_{hj}(i)}{\sigma_{hj}}. \tag{1.1.4}$$

This definition moves the focus from the most connected nodes or the most "connected ones" to the better connected ones or rather to the vertices that play a central role in keeping different components of the network connected. This definition can be easily extended to edges. Indeed, it is possible to count the number of shortest paths passing through a given edge (i, j) for all the pairs of vertices of the network. The vertices or edges with a high value of betweenness centrality are often called bridges and their importance lies in their central role in connecting different regions: if these nodes are removed from the network, the graph quickly becomes disconnected and the average path between nodes increases. It is then evident that networks featuring many bridges or highly central nodes are more fragile and prone to failure in case of an external damage.

1.1.3 Clustering

An additional analysis of interest in network analysis is the quantification of the clustering, i.e. a quantification of how the structure of a local neighborhood is "dense" or interconnected. A typical application of this measure is to quantify the overlap of the neighborhoods (i.e. the set of nodes connected to each vertex) of two or mode nodes connected with each other in social networks. Indeed, in many natural networks we observe a tendency to create dense subnets (or even *cliques*) in the neighborhood of a selected vertex. This behavior can be explained as that if node i is connected to node j and j is connected to l then, with a very high probability, also i and l are connected. The clustering of an undirected network can be measured with the *clustering coefficient* C [12]. The latter measures the cohesion of a local group and it is defined for a single node i of degree k_i as:

$$C(i) = \frac{e_i}{k_i(k_i - 1)/2}, \tag{1.1.5}$$

where e_i is the number of edges connecting the neighbors of i over the maximum number of such edges (given that i has degree $k_i > 1$). If $k_i \leq 1$ the clustering coefficient is set to $C(i) = 0$. Given the clustering coefficient of a single node we

can define the average clustering coefficient as the simple average over all the nodes $i = 1, \ldots, N$ of $C(i)$:

$$\langle C \rangle = \frac{1}{N} \sum_{i=1}^{N} C(i), \tag{1.1.6}$$

so that the clustering coefficient is normalized, i.e. $0 \leq C(i) \leq 1$.

1.2 Statistical properties of networks

So far, we introduced local measures and quantities that cannot be transferred painlessly in the task of analyzing large networks. As we already pointed out, the measure of betweenness centrality takes a diverging computational time, the local measure of clustering coefficient gives us only a local and node-dependent quantification of clustering and so on.

The recent and unprecedented availability of large scale datasets and the constantly increasing computational power of modern cluster allow for the gathering and manipulation of very large real-world networks. In these systems the tools provided by statistical mechanics and the ones specifically developed by network-science come to hand. The attention is thus shifted from local elements to statistical measures taking into account the global behavior of the quantities under investigation.

1.2.1 Degree and betweenness

First in line comes the degree distribution $\rho(k)$ of a given undirected network that is defined as the probability for a node to have degree k. It is evaluated by the construction of the normalized histogram of the degrees of all the N nodes in the network. For directed networks one can define both the in-degree distribution $\rho(k_{in})$ and the corresponding out-degree one $\rho(k_{out})$. Given the degree distribution $\rho(k)$ one can easily compute the average degree of a network that reads

$$\langle k \rangle = \frac{1}{N} \sum_{i} k_i = \sum_{k} k \rho(k) \equiv \frac{2E}{N}, \tag{1.2.1}$$

as each edge connects two vertices. For the same reason, in a directed network the average values of both the in-degree and out-degree distribution must coincide, so that

$$\langle k_{in} \rangle = \sum_{k_{in}} k_{in} \rho(k_{in}) = \langle k_{out} \rangle = \sum_{k_{out}} k_{out} \rho(k_{out}) \equiv \frac{\langle k \rangle}{2}. \tag{1.2.2}$$

We can then generalize the computed moment of the degree distribution and define the n-th moment $\langle k^n \rangle$,

$$\langle k^n \rangle = \sum_{k} k^n \rho(k). \tag{1.2.3}$$

As we will show in Section 2.5 the first and second moment of the distribution play a central role in determining the dynamical properties of a reaction-diffusion process

undergoing in complex networks. As a last remark, let us note that a sparse graph has an average degree that is much smaller than the size of the network itself, i.e. $\langle k \rangle \ll N$. In Chapters 2-3 we will show how such properties of the degree distribution will help us identifying different families of networks.

As with the degree, we can also define the $P(b)$ probability distribution of betweenness centrality, i.e. the probability for a node to have betweenness b. By repeating the same reasoning we can define the average betweenness $\langle b \rangle$. The value of the latter is fixed by the average shortest path length $\langle l \rangle$. Indeed, by simply re-arranging the sums in the betweenness definition we get

$$\sum_i b_i = \sum_i \sum_{h,j \neq i} \frac{\sigma_{hj}(i)}{\sigma_{hj}} = \sum_{h \neq j} \frac{1}{\sigma_{hj}} \sum_{i \neq h,j} \sigma_{hj}(i). \tag{1.2.4}$$

The last sum may be rewritten as $\sum_{i \neq h,j} \sigma_{hj}(i) = (l_{hj} - 1)\sigma_{hj}$, so that $\sum_i b_i = N(N-1)(\langle l \rangle - 1)$. By plugging the last result in Eq. (1.2.4) we get

$$\langle b \rangle = (N-1)(\langle l \rangle - 1). \tag{1.2.5}$$

The betweenness then usually scales as $\mathcal{O}(N)$, so that it is customary to define a normalized version of it as $\tilde{b} \equiv b/N$.

Another property of betweenness is that it usually correlates with the degree of the node [30, 31], i.e.

$$b_i \sim k_i^{\eta}, \tag{1.2.6}$$

where η is a positive exponent whose value depends on the network under investigation. The nodes featuring both a large betweenness b and degree k are called *hubs* and they play a central role in processes of network exploration as they will be easily visited and discovered. Moreover, as a lot of edges are pointing at them and they fall in most of the shortest paths between nodes in the network they will carry a large load of the network's traffic, usually ending up in a congestion of the node. On the other hand, fluctuation of the network topology may generate nodes with a large betweenness but low degree k. As we already pointed out, these are bridges connecting different populations or components of the network and they usually appear in network where spatial constraints prevent the hubs to establish long-range links [32, 33].

1.2.2 Nodes mixing

As one can easily imagine, the degree distribution itself does not tell much about the topological properties of a network. Indeed, the distribution itself does not tell anything about how nodes connect to each other with respect to their intrinsic properties: it is reasonable to expect nodes not to connect among themselves irrespective of their features (degree, clustering of their neighborhood etc.). For instance, in many networks specific mixing patterns has been observed. A typical case of such a mixing is the "assortative mixing" that refers to the propensity of nodes in social or ecological systems to connect to other nodes featuring similar

properties (e.g. in a social network people will prefer to be connected with other nodes sharing a similar interest).

On the other hand, there are situations in which a "disassortative mixing" is found, i.e. a node tends to interact with nodes sharing different properties or opposite attributes.

The mixing pattern can be defined with respect to the most different type or property of a node [34]. Nevertheless, the research work in network-science and sociology usually focus on a particular mixing, i.e. the mixing by vertex degree. This mixing refers to the probability for a node of degree k to connect to a node of degree k' and it is studied by means of a multipoint degree correlations function. In many real-world social networks this mixing is found to be assortative (i.e. hubs tend to interact with other hubs and low-degree nodes are more likely to connect to other low-degree nodes). In many technological and biological networks, a disassortative mixing is found. The typical example is the worldwide network of airports: though hubs are connected to each other, the majority of their neighbors are small airports often connected to the hub alone. Thus, most of the hub's neighbors are small degree nodes, while it is very likely for a loosely connected airport to be connected to a hub. This two mechanism are then inducing a disassortative mixing in the network.

The correlations of nodes mixing are usually characterized by means of the conditional probabilities $P(k', k'', \ldots, k'^{(n)}|k)$ that a vertex of degree k is connected to n other vertices whose degrees read $k', k'', \ldots, k'^{(n)}$. Given this definition, a network is said to be uncorrelated when it displays no structure in the conditional probability, so that the only relevant measure is the degree distribution $\rho(k)$.

A more direct quantity is defined as the Pearson assortativity coefficient r introduced in the 2002 by Newman [35]. The quantity r is defined as

$$r = \frac{\sum_e j_e k_e/E - [\sum_e (j_e + k_e)/(2E)]^2}{[\sum_e (j_e^2 + k_e^2)/(2E)] - [\sum_e (j_e + k_e)/(2E)]^2}, \tag{1.2.7}$$

where j_e and k_e denote the degree of the extremities of a given link e. This quantity varies from -1 (disassortative mixing) to $+1$ (assortative mixing), despite the measure could be misleading in the correlation function has a non-monotonous behavior as the coefficient r weights more the more abundant degree classes [35].

A more accurate insight on the mixing of topological properties of a network is given by the average nearest neighbors degree of a vertex i

$$k_{nn,i} = \frac{1}{k_i} \sum_{j \in \mathcal{N}(i)} k_j, \tag{1.2.8}$$

where the sum spans the neighbors j of node i.

Starting from Eq. (1.2.8) we can get a deeper insight on degree correlations by averaging the degree of nearest neighbors $k_{nn}(k)$ over nodes of degree k, getting [36, 37]

$$k_{nn}(k) = \frac{1}{N_k} \sum_{i|k_i=k} k_{nn,i}, \tag{1.2.9}$$

where N_k is the number of nodes of degree k. We can now write a condition on the degree correlations of connected vertices as, on average, we find

$$k_{nn}(k) = \sum_{k'} k' P(k'|k). \tag{1.2.10}$$

If the degree k' of the first neighbors is uncorrelated to k, then $P(k'|k)$ is a function of k' alone and $k_{nn}(k)$ is a constant in k. On the other hand, if correlations are present, we can distinguish two kind of networks. If $k_{nn}(k)$ is an increasing function of k, hubs will more likely interact with hubs (assortative mixing). In the other case $k_{nn}(k)$ is a descending function of k, and high degree nodes will be linked with nodes with a low average degree and vice-versa (disassortative mixing).

As a last remark, let us note that for some functional form of the degree distribution a network with a completely uncorrelated degree-degree distribution is not always feasible due to some structural constraints [21, 38–40].

In the case of random, uncorrelated networks we can derive an explicit form for the conditional probability $P(k'|k)$. In this case the $P(k'|k)$ does not depend on k but only on the k' degree of the node pointed by the link. The probability for an edge to point to a node of degree k' then equals the number of edges departing from nodes of degree k' divided by the total number of edges, i.e.

$$P_{\text{uncorr}}(k'|k) = \frac{k' N_{k'}}{\sum_{k''} k'' N_{k''}} = \frac{1}{\langle k \rangle} k' \rho(k'), \tag{1.2.11}$$

where we used the fact that the degree probability distribution $\rho(k) = N_k/N$. This result tells us that even in truly random network the probability for an edge to point to a node of a given degree k' is not uniform but proportional to its degree. In other words, when following a randomly chosen edge of the network it is more likely to end up in a vertex featuring a large degree (a hub).

As for the conditional probability, other quantity that is worth to re-define in a degree-dependent way is the clustering coefficient $C(i)$ defined in Eq. (1.1.5). Indeed, the average clustering coefficient $\langle C \rangle$ defined in Eq. (1.1.6) just measures a global tendency on the network for edges to close triangles and gives us only a first indication of the network's cohesiveness. It is then customary to re-define the clustering coefficient $C(i) \to C(k)$ as a function depending on the degree k alone, averaging the clustering coefficient C_i over nodes of same degree k

$$C(k) = \frac{1}{N_k} \sum_{i|k_i=k} C(i), \tag{1.2.12}$$

where N_k is the number of vertices featuring degree k. The presence of a $C(k)$ dependence on k may be due to the hierarchical structure of the network that can generate complex structure in the three-vertex correlation pattern.

1.2.3 Link weight

So far we focused only on the connectivity or degree of a node and we treated all the edges of a given network as they all are equal and indistinguishable. In real-world systems however, the heterogeneity are found not only in the nodes' degree k but also on the capacity or intensity of their edge: the weight of the edges. Indeed, certain edges of a technological network may be able to carry out more traffic than others, whilst in a social network some links could be activated more frequently than other etc.. To account for this additional complication one usually generalizes the entries a_{ij} of the adjacency matrix $\mathbf{A}$ to be the weight of the (i, j) edge, w_{ij} ($w_{ij} = 0$ if nodes i, j are not connected): the weight w_{ij} may assume any value as it measures different physical properties such as capacity, bandwidth, traffic or intensity.

Given the weight of an edge one can define a central quantity of networks' properties, i.e. the *strength* s_i defined as

$$s_i = \sum_{j \in \mathcal{N}(i)} w_{ij} \tag{1.2.13}$$

where the sum covers all the nodes in the neighborhood of i. The strength s_i can be regarded as a natural generalization of the measure of degree. Usually, in the approximation in which the edges weights do not depend on the topology, the weight is usually proportional to the node degree, i.e. $s_i \sim \langle w \rangle k_i$. On the other hand, when such correlations are present we find the strength to grow as $s_i \simeq A k^{\eta}$, with $\beta \geq 1$. As for the degree, also in this case it is possible to define and then measure a probability distribution of both the edges weights $P(w)$ and nodes strengths $P(s)$.

Despite its importance, the simple measure of the node's strength may give misleading results as it does not tell us how the strength of a node is distributed among its neighbors. In particular, we could have a network with an high clustering coefficient but with the edges forming the triangles with low weight so that the highly coherent structure is, in fact, overestimated. On the other hand we could have a hub with a large total strength that could be concentrated in few connections toward other hubs with the rest of the edges featuring a low weight. In this case a simple topological measure would reveal a strongly disassortative mixing, while, taking into account the edges weight, one would find an assortative behavior, as the majority of the hubs' strength is put in the connection toward other hubs. To overcome this limitation, has been defined the *weighted clustering coefficient* [41–44]

$$C^w(i) = \frac{1}{s_i(k_i - 1)} \sum_{j,h} \frac{(w_{ij} + w_{ih})}{2} a_{ij} a_{ih} a_{jh}, \tag{1.2.14}$$

where a_{ij} is the entry of the adjacency matrix $\mathbf{A}$. The $C^w(i)$ quantity measures the clustering of the node i but weighing it with respect to the actual strength of node i. The pre-factor is a normalization constant that ensures that $0 \leq C^w(i) \leq 1$. As for the unweighted clustering coefficient $C(i)$ and average clustering coefficient C defined in Eq. (1.1.5) and Eq. (1.1.6), it is possible to define an average $C^w = \langle C^w(i) \rangle_i$ and degree dependent $C^w(k) = \langle C^w(i) \rangle_{k_i = k}$. This weighted version of the clustering coefficient can be compared to the unweighted version C. If one finds

$C > C^w$ then the topological clustering is due to edges with low weight. If, on the other hand, $C < C^w$ then the triples in the system are more likely to be closed by "heavy" edges.

As for the clustering coefficient, it is also possible to generalize the concept of average nearest neighbor degree $k_{nn,i}$ (see Eq. (1.2.8) and following) to its weighted form $k^w_{nn,i}$ [41]

$$k^w_{nn,i} = \frac{1}{s_i} \sum_{j=1}^{N} a_{ij} w_{ij} k_j, \tag{1.2.15}$$

together with the $k_{nn,i}(k)$ weighted average nearest neighbors degree for nodes of degree k.

The weighted quantities so far introduced are simply re-defining topological properties taking into account the edges weight. However, they do not tell us anything on how the strength s_i of a node i is allocated among its k_i edges: the strength, though large, could be homogeneously distributed on the k_i nodes of weight s_i/k_i or concentrated in one (or few) edges. To overcome this limitation, a more powerful measure of the heterogeneity of a node's i strength allocation over its k_i edges can be measured with the $Y_2(i)$ quantity or the local entropy, respectively.

The first is defined as [45]

$$Y_2(i) = \sum_{j \in \mathcal{N}(i)} \left[\frac{w_{ij}}{s_i} \right]^2, \tag{1.2.16}$$

that is $\sim 1/k_i$ for homogeneous distributions of the node's strength while $Y_2 \sim 1/m$ (with $m = \mathcal{O}(1)$) for heterogeneous strength allocation.

The local entropy $f(i)$ is instead a practical application of the Shannon Entropy, measuring the weight distribution on the k_i edges, i.e.

$$f(i) = -\frac{1}{\ln k_i} \sum_{j \in \mathcal{N}(i)} \frac{w_{ij}}{s_i} \ln \left[\frac{w_{ij}}{s_i} \right]. \tag{1.2.17}$$

As for the Shannon entropy, $f(i)$ is maximum (equals 1) when the weight are homogeneously distributed, while $f(i) = 0$ if we reach the maximum heterogeneity, i.e. all the strength allocated on a single weight.

This completes the tools here introduced. Let us stress that the weighted quantities here introduced play a central role in the study, characterization and prediction of dynamical processes on top of real-world networks where a great heterogeneity of the bandwidth or traffic capacity of edges is found. There are, however, other quantities of practical interest that take into account the weight of edges and the strength of nodes, e.g. the weighted topological distance and the minimum spanning tree. More tools and details can be found in the (vast) literature of the field [46–48].

Chapter 2

Complex networks

Many real-world networks feature strong heterogeneities that, if visualized in a graphical representation of the network, transmit the idea of a random, complicated and somehow arbitrary arrangement of nodes and vertices. Given this complex scenario, it is important to measure and characterize how a network is heterogeneous and when this heterogeneity results in a truly complex system.

In this Chapter we introduce the idea of what is a complex networks and show which real-world networks belong to this category.

We then use the relevant features of these networks to define different families of networks. We also present some generative models able to reproduce the self-organized evolution as well as the scale-free and small-world properties. We then introduce the basic concepts, mechanisms and results of reaction-diffusion processes undergoing on top of complex networks.

2.1 Networks and complexity

A graphical visualization of a large scale network as shown in Fig. 2.1 may give the impression of a complicated, somehow random system generated by a generative process difficult to describe and shaped in an apparently aimless way.

Despite the undeniable difficulties to catch at a first eye-sight the underlying structure, features and organization of such real-world networks, their apparently complicated nature is misleading as these system are just *complex* [1].

These systems are the main objects under investigation in the intense research activity of the last years in complex systems. As already discussed, a complex system does not imply that the system is complicated. On the contrary complex systems are emergent phenomena: they are the spontaneous outcome of the interactions among the system constitutive units. Moreover, their behavior cannot be described from simple extrapolations of the properties of their constitutive units. The study of each subpart of the system in isolation does not allow an understanding of the whole system or of its dynamics. This is mainly due to the presence of scale-free structure and diverging fluctuations [1, 3, 5].

Within this framework, a possible definition of complex network is as follows: many real-world complex systems admit an abstract mathematical representation as a graph in which nodes or vertices are the units/elements of the system and links

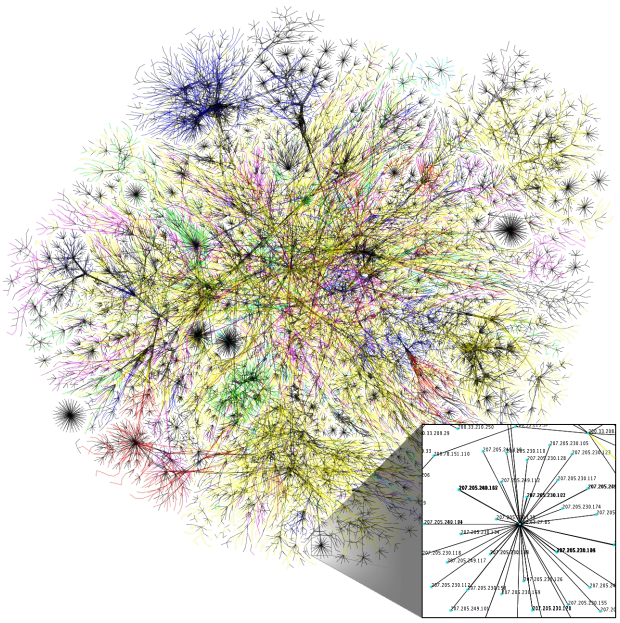


Figure 2.1: Partial map of the Internet based on the January 15, 2005 data found on opte.org. Each line is drawn between two nodes, representing two IP addresses. The length of the lines are indicative of the delay between those two nodes. This graph represents less than 30of the Class C networks reachable by the data collection program in early 2005. Lines are color-coded according to their corresponding RFC 1918 allocation. Credits: The Opte Project.

or edges represent a relation or interaction among those elements. If, after this abstraction has been made, the complex features are still present [5] we will refer to these complex systems as complex networks.

All of the large real-world networks feature the characteristic properties of complex networks: the WWW, the internet and many social and biological systems are all structures that grow in time by following complicated dynamical rules without a general blueprint. They all self-organize to adapt and resist to damages, removals and they evolve to optimize their performance to accomplish the network's main task (resource allocation, information spreading etc.). At the end of such evolution or rearrangement processes all these systems feature an emergent structure with non trivial properties and regularities. Moreover, their complex nature allows for a better plasticity and resilience from external attack or damage.

As we will show in the following, these heterogeneous, scale free properties of complex networks are usually indicated by their heavy tailed, skewed distributions of degree and strength, together with the small world effect and other complex topological properties such as the presence of hierarchical and community-like structures.

Moreover, most of these networks are dynamically shaped in time according to local processes that define the topological and dynamical properties of the system. The fact that networks evolution outcome is an ubiquitous complex topology may hardly be described as a coincidence. Indeed, it prompts for a deeper understanding and research work to measure and characterize the general organizing principles

that might shape and determine the emergence of this architecture in very different contexts.

As we will show in Section 2.4 and Chapter 3, the modeling effort has been shifted from the static representation of networks to their generating processes that govern the appearance and disappearance of vertices and links. Prompted by the continuously growing availability of digital, large and high quality network datasets, this approach has triggered the development of new classes of models which predict both the large-scale properties and temporal evolution of the networks from basic principles regulating the interactions among its nodes.

A detailed review of the concepts here introduced can be found in literature [18, 49–55].

2.2 Real-world networks

As already pointed out, the growing availability of large, high quality and longitudinal datasets paved the way for a detailed structural and dynamical analysis of real-world complex networks.

So far, we can distinguish two main families of systems, i.e. the infrastructure systems and the natural or living systems. We can further divide networks within these two families: the structural systems may be divided in technological virtual or cyber networks such as the WWW, or physical, distributing networks like power grids, the road web etc. For what concerns the natural systems we can model as networks the social interactions (social networks), food web, the proteins in a living cell etc..

Social and biological networks are actually the ones at the center of the research field in the last years, as many of the network science research topics belong to this area and as the majority of sociologists refer to our society as a networked one.

In these two main categories the dynamical processes shaping the network generate an intricate and structured topology. In some cases, like airport or power distribution networks, these mechanism are a mix of both social and economical mechanism cooperating to the creation of the network topology.

Such an ubiquitous presence of networks calls for a characterization of the properties and elements of the prototypes usually found in literature, as these networks feature very different nature of their element.

The paradigm of networks and one of the fundamental application of network theory is the field of social networks. Since the seminal work of Moreno [56] and the definition of a sociogram, the dynamics and arrangement of social interaction has been under a constant and busy analysis and study. In social networks the nodes are people and the edges linking them may be the most diverse kind of interactions: phone calls, sentimental relations, sex intercourse, scientific collaboration or, recently, interactions on on-line social networks [12, 13, 22, 26, 29, 57–68].

The recently available datasets coming from this last kind of interactions, together with the digitalization of many human activities, allow us to deepen our understanding on human dynamics and on the social interactions' guiding principles. These research topics are not only interesting per se, but they also help us to

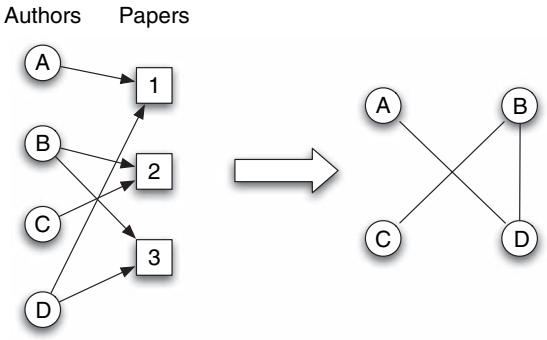


Figure 2.2: A typical example of a bipartite graph: nodes are divided into two separate sets such that every edge links a node of one set to a node of the other set. For a co-authorship network the authors and the published papers form two distinct classes. Each edge represents the authorship of an author in a given paper. On the right we show the corresponding one mode projection, i.e. a unipartite graph. Here the nodes of the author class share an edge if they are connected to a common node of the papers class, thus defining the co-authorship network in which two authors share an edge if they co-authored at least one paper. Image credits: Barrat, Barthèlemy, Vespignani [15].

better understand the spreading of diseases, emergence of consensus and retrieval of information together with knowledge diffusion.

These network are usually bipartite (the presence of actors in a movies or the authors of a scientific paper), but the one mode projection is usually used. In the latter, we put an edge connecting two nodes (actors or researcher) if they both have played or wrote the same movie or paper. Given this definition, we can reintroduce the notion of degree and weight of an edge. In this multi agent system it is worth defining the weight of he (i, j) link as

$$w_{ij} = \sum_p \frac{\delta_i^p \delta_j^p}{n_p - 1}, \tag{2.2.1}$$

where the sum runs over all the possible social interactions p (all the recorded movies or all the written scientific papers), n_p is the number of actors of authors of the p-th interaction and δ_i^p is 1 if the i-th author took place in the p-th paper, and 0 otherwise.

This weight introduces the proximity between two authors or characters that have a stronger tie connecting them if they collaborated together in many events.

Other fields in which network theory is frequently applied are the transportation networks and technological networks. The former field, as the others, has gained from the large datasets recently made available. In this networks nodes are for instance airports or cities and edges are connecting flight or highways going from one place to the other. The importance of this topic is found in its application to the formulation of epidemic forecasts as they take into account the commuting and travel of people around the globe [32, 41, 69–77]. As for social networks and as already said in Section 1.2.2 also these networks feature heterogeneity in the nodes

degree (there are hubs and leaves) and strength.

As for the technological networks, the two most renowned examples belonging to the common knowledge and everyday experience, are the World Wide Web (WWW) and the Internet.

The latter is the ensemble of router, computers, cables connecting devices worldwide. Its exponential growth makes it difficult to extrapolate the key features of the network. Indeed, the size of the network has grown by almost six order of magnitude since the first realization of the net and the components and interconnections found in the Internet are as heterogeneous as possible (link with different traffic capacity, server with thousands of incoming/outgoing connections, local area networks, etc..). Moreover, the absence of a blueprint and the diverse economical, physical and technical constraints led to a very complicate structure and topology [50, 78].

On the other hand, the WWW is probably the most famous virtual network developed on top of the Internet infrastructure. Due to the completely unregulated growth and the characteristic freedom, the Web is larger both in size and complexity than the underlying Internet network, to the point that is difficult (if not even impossible) to estimate the basic properties of the graph (number of pages, links between them etc.). Even in this case the network exploration and characterization must be carried out by special software packages (crawlers) that, starting from a web page, iteratively follows all the links therein to the pages reached from the first one. The process is then repeated to generate a graph representation of the local WWW [79–81].

As a last remark, let us recall that, besides the Internet, WWW, and social networks, the large interest on network science is given by the possibility to apply it to a wide range of biological networks such as the Protein Interaction Network (PIN), food webs and web of sexual relations. The application of network science in these field allowed for the individuation of the protein responsible for gene knockout lethality, the interactions in a given ecosystem, and the prevention of sexually transmitted diseases, respectively [82–87].

The possibility to apply the complex networks approach to such a large set of real-world systems by looking at digital datasets calls for a rigorous treatment and statistical analysis of the data. Indeed, in many cases the complete network topology cannot be retrieved due to technical, time or resources constraints. To this end, an increasing part of the network science works are aimed at a better sampling and analysis of network and how to take into account the missing part of them [88–92].

2.3 Network families

Despite the differences in the real-world networks presented so far, thanks to the toold developed by network science it has been possible to discover some common and shared features of these networks that allowed for a cataloging of them. In particular we distinguish two fundamental features that help us in defining different

families of networks: *small-world*, *clustering* effects and the network *heterogeneities*.

The small world property refers to the fact that, as already shown in Section 1.1.1, two randomly chosen nodes of a network are separated by an average distance $\langle l \rangle$ much smaller than the system scale N. The distance usually scales as $\langle l \rangle \sim \ln N$ so that we recover the famous "six degrees of separation" of Milgram [8].

In many networks this property is crucially important as it allows for a fast and direct connection of nodes in the network. For example, if the diameter of Internet were to scale as $N^{1/2}$, each dns or http request would have to cross ~ 1000 routers and servers before reaching the actual target, slowing down the functioning of the infrastructure.

To quantify the small world effect one usually measures the $M(l)$ quantity, i.e. the number of nodes within a distance l from any given vertex. While in a regular lattice $M(l) \sim l^{\eta}$, in a small world network this quantity grows exponentially (or faster). From a modeling point of view, this property can be replicated by imposing that each possible edge between pairs of nodes in the system is created with the same probability p (see the Erdos-Rény model in Section 2.4.1). It is then possible to show that the average distance grows exactly with the logarithm of the network size N. This result may reasonably explain the ubiquitous presence of the small world effect in real-world systems, as we expect the presence of randomness and stochastic behavior in the dynamical processes shaping these networks.

An important result later found (see Section 2.4.1) is that the small world effect can easily coexists with an high clustering coefficient. An high cohesiveness is usually found in real-world networks and especially in social networks. However, a truly random model of edge creation cannot account for this cohesion, as the clustering coefficient usually gets very small in random networks. This means that the high clustering coefficient together with the community and components ordering are a memory of a grid-like ordering. The breakthrough of Watts and Strogatz was to develop a model able to reproduce these two features at once [12].

The main properties of real-world networks that was still missing in the networks generated by the model was the heterogeneity (or the appearance of heavy tails) in the network distributions of either degree k or strength s.

Though some real-world networks display an homogeneous distribution of degree, betweenness and weighted quantities (so that they are well described by light tailed distribution such as Gaussian or Poissonian ones), other networks feature statistically heterogeneous distributions of the same quantities that are now found to be heavy-tailed (skewed distributions like power-laws one).

To illustrate the difference between an homogeneous and an heterogeneous network let us focus on the degree distribution $\rho(k)$. In many real-world networks we find few nodes with very large degree (hubs) and many others with low degree (leaves or loosely connected nodes). The presence of hubs and heterogeneities on the connectivity pattern is of larger impact than one can trivially imagine, yielding in many cases degree distribution with heavy tails [13] as shown in Fig. 2.3.

In the large scale limit, heavy tailed distributions do not feature a characteris-

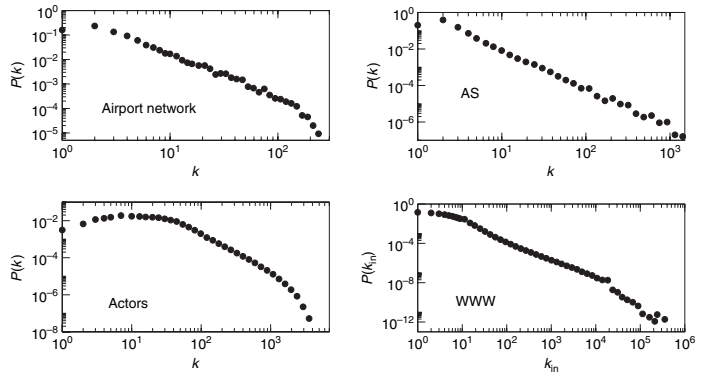


Figure 2.3: Degree distribution $\rho(k)$ of real-world networks. Worldwide airport network (IATA data); actors' collaboration network (Internet Movie database at http://www.imdb.com/); map of the Autonomous Systems of the Internet obtained by the DIMES mapping project (http://www.netdimes.org); map of the WWW collected in 2003 by the WebBase project. Image from [15].

tic scale or an average behavior. Indeed, as these distributions usually span three or more orders of magnitude, there is a non negligible probability to find in heterogeneous systems a node whose degree or strength strongly deviates from the average value. Usually, all of the real-world distributions can be approximated by a power law (or Pareto distribution) dependence on the measured parameter value, e.g. $\rho(k) \sim k^{-}\mu$.

The heavy tailed nature of these distributions translates to a very large (if not diverging) level of degree fluctuations. Mathematically speaking, if we let the degree distribution to be $\rho(k) \propto k^{-\gamma}$, with $2 < \gamma \leq 3$, the average degree is well defined as

$$\langle k \rangle = \int_{k_{\min}}^{\infty} k\rho(k)dk, \tag{2.3.1}$$

where $k_{\min}$ is the minimum value of k found in the sytem. However, if we now compute the variance $\sigma^2 = \langle k^2 \rangle - \langle k \rangle^2$, this is dominated by the second moment of the distribution. The latter is even diverging with the system scale as

$$\langle k^2 \rangle \propto \int_{k_{\min}}^{k_c} k\rho(k)dk = \int_{k_{\min}}^{k_c} k^{2-\gamma}dk \sim k_c^{3-\gamma}, \tag{2.3.2}$$

where k_c is the upper cutoff of the distribution for a finite size system. Thus, $\langle k^2 \rangle$ and then σ^2 diverges as $k_c \to \infty$. The presence of such a large fluctuations implies that the system has no characteristic scale, being *scale-free*. This also means that if we change the scale of the system by letting $k \to \lambda k$, the shape and behavior of the distribution does not change.

To quantify the heterogeneity of a network it is possible to define the parameter

$$\mathcal{K} = \frac{\langle k^2 \rangle}{\langle k \rangle}, \tag{2.3.3}$$

so that the normalized variance of the distribution reads $\sigma^2 = \mathcal{K}/\langle k \rangle - 1$: scale free network then have $\mathcal{K} \to \infty$, while homogeneous ones have $\mathcal{K} \sim \langle k \rangle$. In other words,

a network is said to be scale free if $\mathcal{K} \gg \langle k \rangle$. As we will see in Section 2.5 this parameter will affect all the relevant properties of dynamical processes evolving on top of complex networks as well as their critical features.

Of course, these heterogeneity are not only found in the degree distribution but they affect all the relevant measures of the networks under investigation: strength, betweenness, and weight. The measure and the frequent approximation of these distributions with a power-law fit raise some issues about the correct way to measure and characterize the power law behavior in real-world data [93]. As a matter of fact, real data observables intrinsically feature a cut-off in their value as the real-world systems cannot be of infinite scale. Due to this complication, alternative functional forms of these distributions are taken into account: Log-Normal distribution, truncated power-law (a power law with an exponential cutoff) and stretched exponential to name only a few [26, 94]. Moreover, several rigorous procedures have been developed in order to tell if a given series of data follows or not a power-law behavior together with the statistical methods to find the best functional form to fit a dataset [94]. We will take advantage of these procedures in our analysis of real-world datasets as we will show in Chapter 4.

2.4 Generative models

We now quickly present some generative models for static networks, i.e. networks without a temporal dimension. In other words, when a link is created it stays in the network forever and there is no dynamics of edge removal/rearrangement. We will introduce the temporal dimension in the next chapter.

2.4.1 Random graphs

The first and more basic examples of generative models of a static random graph are the ones by Erdos-Rény, and Molloy and Reed, respectively [9–11, 95]. These are the most basic models in the sense that there is no information on or knowledge of the underlying mechanism generating the graph. Given this lack of insights, the simplest possible assumption is that each of the $N(N-1)/2$ edges is created with a given *connection probability* p. In an alternative definition, given N nodes, we randomly chose E pairs of nodes as vertices of the E edges of the network (this is actually the original formulation of the Erdos-Rény model). The outcome of this generating procedure is the $G_{N,p}$ (or the $G_{N,E}$ for the second approach) graph. In particular, there are $2^{N(N-1)/2}$ possible representation of a $G_{N,p}$ graph, given that each edge can be either present or not. Many other average quantities are easily derived. For instance, the average number of edges of a $G_{N,p}$ graph is simply $\langle E \rangle = \frac{1}{2}N(N-1)p$, so that the average degree reads

$$\langle k \rangle = \frac{2\langle E \rangle}{N} \simeq Np, \tag{2.4.1}$$

for $N \gg 1$. Eq. (2.4.1) tells us that the average degree is diverging with the system size N so it is natural to impose a size-dependent connection probability scaling as

$p(N) = \langle k \rangle / N$ with the system size N. Moreover, when $\langle k \rangle < 1$ we have a network composed by many subgraphs not interconnected among them. When $\langle k \rangle$ trespasses the value of 1, we observe a phase transition equivalent to the percolation one. For $\langle k \rangle > 1$ a giant component (whose size is proportional to the network size) emerges.

In the same way one can evaluate the degree distribution $\rho(k)$, given that the probability to have a node of degree k is to create a node connected to k nodes and not connected to the remaining $N - k - 1$, i.e.

$$\rho(k) = \binom{N-1}{k} p^k (1-p)^{N-k-1}. \tag{2.4.2}$$

If we let $N \to \infty$ and $p = \langle k \rangle / N$, we can approximate the binomial distribution with the Poisson one:

$$\rho(k) = e^{-\langle k \rangle} \frac{\langle k \rangle^k}{k!}. \tag{2.4.3}$$

The Erdős-Rény model then generates a homogeneous network as the degree distribution exponentially falls and features small fluctuations. Indeed, the approximation in which we consider the average node with the average degree $\langle k \rangle$ as representative of the network works well in the large scale characterization of the network. One can also show that the Erdős-Rény model generates a small clustering coefficient. The latter, given that connections among a nodes' neighbors are independently created with probability p, reads

$$\langle C \rangle = p = \frac{\langle k \rangle}{N}, \tag{2.4.4}$$

so that $\langle C \rangle \to 0$ when $N \to \infty$. Finally, it is possible to evaluate the average shortest path length $\langle l \rangle$ that scales logarithmically with the system size N as

$$\langle l \rangle = \frac{\log N}{\log \langle k \rangle}, \tag{2.4.5}$$

signalizing the presence of the small-world effect in the Erdős-Rény graphs.

Starting from the E-R generative model, many other generative procedures have been developed, so as to generate networks with an arbitrary degree distribution [95–97]. Even in these cases we recover the small world signature and the selected degree distribution. Despite the lack of heterogeneities in the network structure, the simplicity and power of these models allowed them to be the basic paradigm in network modeling for almost forty years.

A possible way to improve the completely random Erdős-Rény (ER) model is by using models featuring *fitness* or *hidden-variable* mechanism. Indeed, in the ER model an edge between a pair of nodes is created with the same probability p irrespective of the nodes under examination. It is reasonable to expect that in some systems a node will prefer to interact or connect to other nodes in the system featuring some properties. For instance, in a social layer, we tend to connect with people sharing our same interests or to the ones that hold the information that we need or we are looking for.

To model this mechanism, a possible way is to assign to each node i of the network a fitness described by a continuous or discrete variable x_i distributed accordingly to a probability distribution $\rho(x)$. Two nodes i and j are then connected with a probability $p_{ij} = f(x_i, x_j)$, where f is a given function [98, 99]. If $f(x_i, x_j)$ is constant we fall back to the E-R model. Otherwise, a node of fitness x has, on average, a degree $k(x)$ given by

$$k(x) = N \int_0^\infty f(x, y)\rho(y)dy \equiv NF(x), \tag{2.4.6}$$

so that the degree distribution reads

$$\rho(k) = \int dx\rho(x)\delta[k - k(x)] = \rho\left[F^{-1}\left(\frac{k}{N}\right)\right] \frac{d}{dk}F^{-1}\left(\frac{k}{N}\right), \tag{2.4.7}$$

where we supposed $F(x)$ to be monotonic in x. To get a scale free network we can then impose a fitness distributed as a power-law and a connectivity function $f(x, y)$ linear in x, y. Despite the resulting heavy tailed degree distribution, the result is somewhat artificial as we are introducing the heterogeneity with the fitness distribution.

A more interesting result is found if we let $\rho(x) \sim e^{-x}$ and

$$f(x, y) = \theta[x + y - z(N)], \tag{2.4.8}$$

where $\theta(x)$ is the Heaviside function and $z(N)$ a threshold depending on the system size N. In this case the degree distribution is found to be $\rho(k) \sim k^{-2}$, so that even from a peaked fitness distribution, by correctly tuning the connectivity function $f(x, y)$ we can recover a scale free network [99]. Besides the degree distribution, even the assortativity $k_{nn}(k)$ and the clustering spectrum $C(k)$ behave as power laws. This framework has been further developed in the following years and is now known as the hidden-variable framework [100]. Using this set of tools, it is possible to determine the topological structure of the resulting network by the defined fitness distribution and connection probability $f(x, y)$.

A feature lacking in the large scale limit of random networks is the high clustering coefficient usually found in real-world systems [8, 101].

As we already discussed, given the undeniable noisy and stochastic behavior of real-world systems, the high clustering must be, in some way, the reminiscence of an ordered lattice-like edge distribution. The encoding of a tunable clustering coefficient C in a simple generative model was a challenging task until the work of Watts and Strogatz [12].

The groundbreaking idea is to start from a regular circular lattice where each node is connected to the $2m$ nearest neighbors and then add some noise (or randomness) by rewiring each edge with probability p. This last step is applied to each node of initial circle proceeding either in clock or anticlockwise.

In this way the ordered structure is perturbed and shortcuts between distant nodes are introduced. The density of these shortcuts is tuned with the parameter

p, being the average degree fixed at $\langle k \rangle = 2m$. The degree distribution can be evaluated and reads

$$\rho(k) = \sum_{n=0}^{\min(k-m,m)} \binom{m}{n} (1-p)^n p^{m-n} \frac{(pm)^{k-m-n}}{(k-m-n)!} e^{-pm}, \text{ for } k \geq m. \tag{2.4.9}$$

In the $p \to 1$ limit the above distribution reduces to a Poisson distribution of the $k' = k - m$ variable with average value $\langle k' \rangle = m$ as in the E-R model.

The real difference with respect to a random network is the clustering coefficient $\langle C \rangle$ that can be tuned with p from its maximum value ($\langle C \rangle = 3(m-1)/2(2m-1)$ for $p = 0$) to lower ones by increasing p (until $\langle C \rangle$ vanishes for $p \to 1$). Together with the clustering coefficient, even the average shortest $\langle l \rangle$ path length changes with p: $\langle l \rangle$ scales as N for $p = 0$, while it logarithmically grows even for very small values of p.

The smallest value of p at which the small-world effect can be observed scales with the size of the network as [27]

$$\langle l \rangle = N^* F\left(\frac{N}{N^*}\right), \tag{2.4.10}$$

where N^* scales as $1/p$.

The advances introduced by the Watts-Strogatz model was the possibility to fine-tune the clustering coefficient $\langle C \rangle$ and to explain the structural cohesiveness of real-world systems as a reminiscence of an initially ordered structure reshaped by some environmental noise [102]. The feature still not reproduced is the scale-free nature of the generated network. Several models have been introduced to generate this feature [103–105], as we will show later in Section 2.4.2, a more clean and elegant way to get a scale free network was introduced just two years later by Albert and Barabási.

As a last remark let us spend some words on the *exponential random graph* formalism. The latter stems from the observation that the set of the $\binom{N(N-1)/2}{E}$ networks generated by the $E - R$ model can be seen as a microcanonical ensemble of all the possible adjacency matrix $\mathbf{A}$ representation (where the degrees of freedom are now the a_{ij} entries of $\mathbf{A}$).

It is then possible to define a statistical weight $P(\mathbf{A})$ of each realization of $\mathbf{A}$ as

$$P(\mathbf{A}) = \frac{\exp\left[\sum_i \theta_i z_i(\mathbf{A})\right]}{\mathcal{K}(\{\theta_i\})}, \tag{2.4.11}$$

being θ_i a set of parameters, $z_i(\mathbf{A})$ a set of the network's statistical observables and $\mathcal{K}(\{\theta_i\})$ a normalization factor.

The observable z_i may vary from the simple average degree $\langle k \rangle$ of the graph to a distribution of a certain observable in the network. The set of parameters θ_i has to be estimated by comparison with the real data via techniques using various likelihood-based estimators [106–108].

The connections with the field of equilibrium statistical physics are evident, as it can be shown that exponential random graph models are equivalent to the statistical

mechanics of Boltzmann and Gibbs for networks [52, 109, 110]. In other words, the probability for the system (our network) to be in the configuration $\mathbf{A}$ is given by the distribution $P(\mathbf{A})$ that minimizes the Gibbs entropy

$$S[P] = -\sum_{\mathbf{A}} P(\mathbf{A}) \ln P(\mathbf{A}). \tag{2.4.12}$$

We can then apply the same machineries of the statistical mechanics in order to impose or measure relevant quantities in the ensemble of the exponential random graphs.

2.4.2 Evolving networks

The main limitation of the models just introduced is that they deal with a static representation of the network: the edges are created at once when the network gets created and there is not any kind of rearrangement on the network connectivity pattern. However, it is commonsensical to expect a dynamics on the links creation and deletion, as, for instance, in a social network friendships, relationships etc. get continuously created and interrupted. As these systems are in a constant evolution we are dealing with non-equilibrium systems: they are driven by microscopical mechanisms driving the system's configuration to explore certain regions of the phase space, due to either a dynamical drift or to the initial conditions of the system.

Dynamical networks Due to the continuous appearance of nodes and edges in the system, it is customary not to study the system taking into account all of its constraints but rather focus on the dynamical evolution of the network.

A temporal dependence of the system's status has to be introduced and we switch the paradigm from the statical probability $P(\mathbf{A})$ for a network to be described by the realization $\mathbf{A}$ to the dynamical probability $P(\mathbf{A}, t)$ for the network realization $\mathbf{A}$ at time t. The evolution of this probability is then described in terms of a Master Equation (ME), a linear differential equation accounting for all the possible transition of the network representation $\mathbf{A} \to \mathbf{B}$ per unit time. These transition amplitudes are proportional to a certain rate $r_{\mathbf{A}\to\mathbf{B}}$ (or transition probability in the case of a discrete time evolution), measuring the rate of transition between the two representations $\mathbf{A}$ and $\mathbf{B}$. If each evolution step does not depend on the history of the system (i.e. if we have no "memory" in the evolution process) we are then dealing with a Markovian process and the evolution of $P(\mathbf{A}, t)$ reads:

$$\partial_t P(\mathbf{A}, t) = \sum_{\mathbf{B}\neq\mathbf{A}} [P(\mathbf{B}, t) r_{\mathbf{B}\to\mathbf{A}} - P(\mathbf{A}, t) r_{\mathbf{A}\to\mathbf{B}}]. \tag{2.4.13}$$

This approach introduces a trade off between the easiness of the mathematical and numerical handling of the ME and the fact that one has to provide the correct formulation of the dynamical laws governing the evolution of the system. Indeed, it is not always trivial to measure and characterize the underlying processes driving the evolution of the real-world networks [111, 112].

As a reference example, let us quickly introduce the simple model of a continuously growing network in which new nodes appear and edge are created, removed

and rearranged [113]. When a node enters the network it creates m reciprocal links toward nodes already present in the network. The simplest observable of the system is the degree distribution that can be measured as the number of nodes N_k that have degree k. This quantity evolves accordingly to

$$\partial_t N_k = r_{k-1\to k} N_{k-1} - r_{k\to k+1} N_k + \delta_{k,m}, \tag{2.4.14}$$

where the last term accounts for the entrance of the node in the group of nodes of degree m. To solve the ME and thus the evolution of the network one should then define the two rates $r_{k-1\to k}$ and $r_{k\to k+1}$.

If the defined rates allows for an analytical solution of the ME, one can solve the system evolution in time. For networks where a new node is added at each evolution step, the natural time scale is the network's size N itself. One can then describe the entire system by looking at the probability $p(k, s, t)$ that a node entered in the system at time s has degree k at time $t \geq s$. To recover the degree probability distribution $P(k, t)$ at time t it is then sufficient to integrate (or sum in the discrete case) over all the possible entrance time

$$P(k,t) = \frac{1}{N} \sum_{s=0}^{t} p(k,s,t), \tag{2.4.15}$$

together with the expression for the average degree value of the node entered at time s at time t:

$$k_s(t) = \sum_{k=0}^{\infty} k p(k,s,t). \tag{2.4.16}$$

If we now assume that the only feature of a node determining its future evolution is its degree alone, we can write the evolution of the dynamical rate equation governing the $k_s(t)$ in terms of the probability $\Pi[k_s(t)]$ that an edge is attached to the s-th node, i.e.

$$\frac{\partial k_s(t)}{\partial t} = m\Pi\left[k_s(t)\right]. \tag{2.4.17}$$

All of the dynamical features of each model defined are then encoded in the probability or growth kernel $\Pi\left[k_s(t)\right]$.

Of course, our analysis could have focused on an alternative quantity of the system, e.g. the number of vertices $N(k|l)$ of degree k sharing a vertex with a node of degree l.

Moreover, despite the analytical approach given by this formalism, it is clear that one has to precisely measure and characterize the dynamical processes shaping the system. Indeed, a wrong definition of the growth kernel Π can easily lead to non-accurate analytical predictions. However, the switch of the focus from the topology of the network to the dynamics of its evolution is the correct way to catch the emergence of complex behavior and critical features of the network.

Preferential attachment The most famous implementation of such a mechanism is the Barabási-Albert class of models [13]. In this class of generative models the wiring process at the base of the network evolution is governed by a *preferential-attachment* rule: new nodes appearing in the network do not connect at random but

will pair more likely with node with a larger degree. The hand-waving explanation of this mechanism is that in many real-world systems nodes with an high degree (i.e., web pages featuring a large number of incoming links or people maintaining a large number of social interactions) will have more probability to get connected by new node appearing in the system. This feature is dubbed in many way: rich-get-richer phenomenon, Matthew effect [114], the Gibrat principle [115], or cumulative advantage [116].

The Barabási Albert model (AB model) combines the preferential attachment mechanism with the growth of the network by defining two simple rules:

- **Growth:** the network starts with m_0 nodes connected vertices. At each evolution step we let a new node enter the system and wire m connections to the nodes already present in the system.
- **Preferential attachment:** each new edge is connected with the s-th node with a probability proportional to its degree k_s.

The generated network has then average degree $\langle k \rangle = 2m$ and features a degree distribution falling as a power law $\rho(k) \sim k^{-3}$ [6, 13, 117].

The model also allows for an analytical approach within the evolving networks formalism as the growth rete $\Pi[k_s(t)]$ is simply related to the probability for a node of degree k_s to acquire a new link, i.e.

$$\Pi[k_s(t)] = \frac{k_s(t)}{\sum_j k_j(t)}. \tag{2.4.18}$$

If we now work in the large k limit (i.e. we let k to be a continuous variable), and by recalling that at time t we added tm edges (and each of them contributes as 2 in the total degree sum) we find

$$\frac{\partial k_s(t)}{\partial t} = \frac{m k_s(t)}{2mt + m_0 \langle k \rangle_0}, \tag{2.4.19}$$

being $\langle k \rangle_0$ the average connectivity of the initial m_0 nodes.

The solution of Eq. (2.4.19) is

$$k_s(t) \simeq m \left(\frac{t}{s} \right)^{1/2}, \tag{2.4.20}$$

where we applied the $t, s \gg m\langle k \rangle_0$ limit. The degree distribution can then be obtained by integrating over all the entrance time and, for an AB model, we find the $t \to \infty$ degree distribution $\rho(k)$ to be:

$$\rho(k) = 2m^3 k^{-3}. \tag{2.4.21}$$

Other relevant quantities that can be estimated in the AB model are the average clustering coefficient

$$\langle C \rangle_N = \frac{m}{8N} (\ln N)^2, \tag{2.4.22}$$

and the average shortest path

$$\langle l \rangle \sim \frac{\log(N)}{\log\log(N)}. \tag{2.4.23}$$

It is also possible to extend the model and define the preferential attachment to relate with a superimposed, dynamical fitness x_s that is weighted by the node's degree k_s. In this way, as the fitness is always drawn from an arbitrary distribution $\rho(x)$, very fit nodes can be attractive even with a small degree and vice-versa.

The overwhelming success of this model dwells in its simple formulation that allows for a straightforward numerical approach and with the ability to create a small world effect and a scale-free network with a simple attachment rule. Even more interest toward this model has flourished when it has been discovered that the preferential attachment rule is based on the relative values of the fitness, rather than on its absolute value [118]. In other words, there is no need to correctly measure the absolute value of the fitness from real-world data but it suffices to rank the nodes accordingly to the selected attribute.

The AB model is of course a first order approximation of a possible mechanism of network growth and it is not meant to be a realistic model of real-world networks. However, a wealth of works flourished from this seminal paper, trying to introduce and replicate all the most diverse features (clustering coefficient, power law degree distribution with exponent $2 < \mu < 3$ etc.). We refer to the literature for a deeper insight [34, 49, 52].

There have also been different approaches proposed in order to explain the skewed degree distribution of real-world networks. Above all let us mention the *copy and duplication* models [119, 120]. In the latter, when a new node i is introduced in the network it chooses at random a node j already present and generate k_j links pointing to the neighbors of j (thus copying the j-th neighborhood). Then, with probability α, each of these k_j links are rewired to a randomly chosen node, otherwise with probability $1 - \alpha$ they are left as is. This copy mechanism makes sense on the WWW, where a new page (the node i) of a given topic (the same topic of the node j in the example) will be more likely to point to the pages already connected to a page of the same topic. The rewiring probability α accounts for the "curiosity" of the writer that may find some unrelated pages interesting.

This mechanism models the preferential attachment in a subtle way, and has been implemented in the growth modeling of the WWW, biological networks, citation networks, gene-protein interactions and social networks.

Another possible way to tackle the network growth modeling problem is via the trade-off and optimization models, where the system arranges its elements and connections so as to optimize (or minimize) a global constraint (or cost function). The most famous families of these models are the Heuristically-Optimized-Trade-off (HOT) [121] and the global optimizations applied to road traffic [122], food webs [123], transportation networks [124] and air travel network [125].

2.4.3 Toward a Data-Driven approach

The models presented so far focus on the reproduction of a skewed degree distribution and, at most, at a small-world effect in their generated networks. It is clear, however, that higher order statistical properties of the networks are found in real-world complex networks. The clustering coefficient is usually higher than in the AB models, the community structure is intricate and the assortativity of these networks has a non trivial functional form. Moreover, in many cases the technological or spatial constraints shape the network topology in synergy with other complex mechanisms.

The need to measure, model and reproduce these additional properties led to a change in the approach typically used on these systems. Instead of a a-priori modeling effort, the attention is now posed on the so-called data-driven approach. The salient features of a system can be measured and some processes characterized thanks to the growing availability of large datasets. These models sometimes accounts for so many specific properties and mechanism to the point they do not allow for an analytical approach or predictions on much quantities of the system.

It is then a common guideline to develop models that catch very general properties of the network and use them as a general reference to characterize dynamical processes undergoing in a wide array of system prototypes. On the other hand, one must keep in mind that a complete and detailed description of such systems may actually be out of reach if one sticks with a general framework and wants to develop an "universal" generative model of the network dynamics. In other words, many complex systems (especially the technological and social ones) are shaped by an intricate mix of different mechanism leading the topological evolution. The interplay of these mechanisms may play a central role in determining the outcome of dynamical processes and cannot be neglected [41, 73, 125–137].

It is clear that each approach presented so far has some pros as well as some cons. The exponential random models are based on the solid foundations of statistical physics but it is cumbersome to develop and often present technical difficulties that limit their application spectra. For example, this approach is not suitable when dealing with out of equilibrium networks and system whose size, properties and interaction evolves in time. As this is the case for the large majority of real-world systems, other approaches have to be pursued.

That is why, in the era in which *Big-Data*, informatics revolution and numerical methods are the common tool to investigate real-world complex networks, the dynamical approach is preferred in general. The latter allows for a finer modeling of the microscopic mechanism and the evolution rates that govern the growth of the network. Moreover, they allow for both a simpler numerical approach at even very large system scales and a more comprehensive analytical approach. This framework has been applied to the most diverse field of research and human activity as it also simply encodes the dynamical nature of these system in a simple and intuitive way.

The availability of these large scale datasets raises some conceptual gap that are still under discussion.

The most relevant one is the issue of "universality". Indeed, many of the ana-

lyzed datasets describing real-world systems apparently very different and diverse, such as food webs or the internet, feature common properties and behavior in their large scale representations. This is due to the fact that they are governed by the same set of statistical laws and that they behave in a similar way in the common phase-space. Universality however, does not mean equivalence, as we always have to remember that each system has some peculiarity and different mechanism on a local scale. Universality must then be intended at a global, coarse and large scale level. The similarities found are used to group the real-world networks in different classes that still share most of their macroscopic properties.

The counterpart of the dynamical approach (and of all the data-driven approaches in general) is the issue of models validation and reliability of the statistical models implemented. While the statistical modeling based on the maximization of entropy takes into account all of the statistical observables, the dynamical approach can be more handy for large datasets and their evolutionary properties.

The same role that data cover in the two cases is also different. In the statistical approach the data are used to estimate the optimal value of the parameters that are then used to forecast distributions of other relevant quantities of the network under study. These results can finally be compared with the behavior found in data.

In the dynamical approach instead the data are used to validate the outcome of the evolving process. In some cases (the data-driven models) the parameters of these microscopic dynamical rules are inferred directly from the data when their measurement is possible.

To cut the chase, each approach and model has some shortcomings. From a modeling point of view is beyond the scope of the physical approach to take into account all of the mechanism playing a role in a system evolution as one is usually interested in the behavior of some quantities, while others are neglected. Moreover, there are models whose aim is to realistically reproduce a process and some others that are just aiming at a general understanding of the fundamental mechanisms governing a certain system. This dichotomy is also found in the modeling efforts of the dynamical processes as we will show in the next section.

2.5 Dynamical processes in heterogeneous networks

We will now introduce different dynamical processes on complex networks. In particular we will be speaking about a specific class of these: Reaction-Diffusion processes.

These are used to model a huge variety of phenomena in which local quantities obey physical reaction diffusion equations. Within the same general framework here introduced, we will describe the diffusion and localization of the the spreading of infectious diseases and the evolution of opinion and consensus in social networks.

2.5.1 Formalism

As already seen in Section 2.4.2 the master equation (ME) formalism allows an analytical approach for dynamical systems in which many agents interact with each

other. This formalism can also be applied to reaction-diffusion processes taking place on a network composed by N nodes. The procedure is as follows: for each node i let us introduce a variable σ_i characterizing its dynamical state, i.e., the evolution of particular attribute of that node. The complete set of these state variables for all the vertices in the network defines the microstate of the whole system, and we will address it as $\sigma(t) = \{\sigma_1(t), \sigma_2(t), \ldots, \sigma_N(t)\}$. This system will evolve in time accordingly to the dynamical rules that regulate the system dynamics. The configuration will then move in the phase space of all the possible system's configurations, passing in one evolving time step from a configuration σ_a to the next one σ_b with a given rate $W(\sigma_a \to \sigma_b)$. As the phase space of real-world complex systems is huge, one usually focuses on the probability $P(\sigma, t)$ of finding the system at time t in a given configuration σ and then studies the evolution of $P(\sigma, t)$ by means of the master equation. In the continuous time approximation we can write the ME of the system in terms of the transition rates $W(\sigma_a \to \sigma_b)$:

$$\partial_t P(\sigma, t) = \sum_{\sigma'} \left[P(\sigma', t) W(\sigma' \to \sigma) - P(\sigma, t) W(\sigma \to \sigma') \right]. \tag{2.5.1}$$

The rates $W(\sigma' \to \sigma)$ are in general function of all configurations. It is quite common to consider cases in which the change of state of a node i is determined only by local interactions with its nearest neighbors. In this case we can write:

$$W(\sigma' \to \sigma) = \prod_i \omega(\sigma_i' \to \sigma_i | \sigma_j), \tag{2.5.2}$$

with j belonging to the neighborhood of i, i.e. $j \in \mathcal{N}(i)$. This is where the network topology enters in the definition of the dynamics, i.e. by defining the transition rates for each node that will "feel" it neighbors when updating its state. In the same way of Section 2.4.2 we can the evaluate the expectation value of a certain measurable quantity A by the *phase space average*, i.e. a weighted average over all the possible configurations at time t giving

$$\langle A(t) \rangle = \sum_{\sigma} A(\sigma) P(\sigma, t). \tag{2.5.3}$$

An isolated system maximizes its entropy and reaches a uniform stationary equilibrium distribution $P_{eq}(\sigma)$ with the same probability of being in any of the fixed energy accessible configurations. However, real system are never isolated. They are always coupled with the external environment, that we can consider as a heat bath that fixes the equilibrium temperature T of the system. In this case the stationary distribution is no longer uniform but characterized by the Boltzmann-Gibss distribution:

$$P_{eq}(\sigma) = \frac{1}{Z} e^{-H(\sigma)/k_b T}, \tag{2.5.4}$$

being H the Hamiltonian of the system and Z is the partition function

$$Z = \sum_{\sigma} e^{-H(\sigma)/k_b T}, \tag{2.5.5}$$

that correctly normalizes the $P_{eq}(\sigma)$ probability.

In this way the equilibrium state is characterized by the detailed balance condition that sets the net probability current between pairs of configurations to zero when $P = P_{eq}$:

$$P_{eq}(\sigma)W(\sigma \to \sigma') = P_{eq}(\sigma')W(\sigma' \to \sigma). \tag{2.5.6}$$

This condition holds only for equilibrium systems, whereas in systems out of the equilibrium currents between microstates do not balance.

Moreover, the complete solution of the ME can be found in a relatively small set of systems and models. That is why some approximations are needed. The most powerful and famous approximation scheme is the homogeneous assumption or mean-field theory. In this approach the system is considered homogeneous and the correlations between microstate variables are neglected. The interactions (field) felt by any element in the system are the same (mean). In other words, each agent feels a mean field given by the average over the full system. In other words the probability p_x for a given agent i to be in the state x is independent on i. So, neglecting the correlations we can write the probability of any configuration as a factorization of single node probabilities:

$$P(\sigma) = \prod_i p_{\sigma_i}. \tag{2.5.7}$$

This approach can be applied to both diffusion processes and $A + B \to 2B$ reaction-diffusion processes. We will present here only the latter referring to the literature for the first case [15, 138–145].

Let us consider as example a fundamental process in which each node can be only in two states $\sigma_i = \{A, B\}$. Let us fix the dynamical rule of the process as $A + B \to 2B$, i.e. the only way for a node in state A to switch to state B is to get in contact with a node in state B. This transition occurs with rate β so that the transition rates for a node i in contact with j read:

$$\begin{aligned} \omega(A \to A|\sigma_j = A) = \omega(B \to B|\sigma_j = A) = \omega(B \to B|\sigma_j = B) = 1, \\ \omega(A \to B|\sigma_j = B) = \beta. \end{aligned} \tag{2.5.8}$$

We can now write the number of nodes $N_A(t)$ and $N_B(t)$ that at time t are in state A and B, respectively as

$$N_A(t) = \sum_{\sigma,i} \delta_{\sigma_i,A} P(\sigma, t), N_B(t) = \sum_{\sigma,i} \delta_{\sigma_i,B} P(\sigma, t), \tag{2.5.9}$$

so that we can write the ME for $N_B(t)$ as

$$\begin{aligned} \partial_t N_B(t) = \sum_{\sigma,i} \delta_{\sigma_i,B} \partial_t P(\sigma, t) = \\ \sum_{\sigma',i} \left[\omega(\sigma'_i \to \sigma_i = B|\sigma'_j) P(\sigma', t)\right] - N_B(t). \end{aligned} \tag{2.5.10}$$

Now we can introduce the mean-field approximation stating that the probability for each node to be in the state A or B is $p_A = N_a/N$ and $p_B = N_B/N$, respectively.

We can then write:

$$\sum_{\sigma'}\omega(\sigma_i' \to \sigma_i = B|\sigma_j')P(\sigma',t) = \\ \sum_{\sigma_j'}[\omega(\sigma_i' = A \to \sigma_i = B|\sigma_j')P(\sigma',t)p_A \prod_{j\in\mathcal{N}(i)} p_{\sigma_j'} + \\ \omega(\sigma_i' = B \to \sigma_i = B|\sigma_j')P(\sigma',t)p_B \prod_{j\in\mathcal{N}(i)} p_{\sigma_j'}]. \tag{2.5.11}$$

By considering that $\omega(\sigma_i' = B \to \sigma_i = B|\sigma_j') = 1$ irrespective of σ_j' and that $\omega(\sigma_i' = A \to \sigma_i = B|\sigma_j') = \beta$ if at least one of the k neighbor of i is in state B. This happens with probability $1-(1-p_B)^k$ (and assuming that all the nodes of the network have degree k) we get:

$$\partial_t N_B(t) = \beta N_A \left[1 - \left(1 - \frac{N_B}{N}\right)^k\right], \tag{2.5.12}$$

that in the $N_B/N \ll 1$ yields:

$$\partial_t N_B(t) = \beta k \frac{N_A N_B}{N}. \tag{2.5.13}$$

The equation for A is given by the conservation rule $N_A = N - N_B$. The expression Eq. (2.5.13) is the mean-field solution for the basic reaction process $A + B \to 2B$ that is part of a wide range of epidemic spreading phenomena that we will discuss in detail in the following sections.

2.5.2 Epidemics models

Using the general framework of dynamical processes on complex networks just introduced we will study another application of those concepts: basic epidemic models. In the simplest formulation, the nodes (or the population) are divided into different compartments depending on the stage of the disease [146, 147] such as susceptible S, those who can contract the infection, infectious I, those who have already contracted the infection, and recovered R, those who have recovered from the disease. Additional stages of the disease can be introduced depending on the type of the disease. Let us consider a population of N individuals and let us define the number of individuals in the class $[m]$ at the time t as $X^{[m]}(t)$. By assuming a constant number of nodes N in time we get $N = \sum_m X^{[m]}(t)$.

The transitions between different compartments depends on the modeling details of the specific disease, as well as the rate at which susceptible contract the infection (force of infection) that has generally two possible form [148]:

- frequency dependent or mass action transmission (equivalent to a homogeneous or mean-field approximation);
- density dependent or pseudo mass action.

The first kind applies when the number of contacts is independent of the population size but rather on the fraction of infectious individuals. The second kind instead assumes that as the population size increase so does the contact rate. These kind of processes are anyway binary interactions among individuals, so that they can be modeled as reaction-diffusion processes of the $A + B \to 2B$ family.

Sticking with a mass action transmission the variation of the m compartment's population $X^{[m]}$due to the process is given by

$$\sum_{h,g} \nu^m_{h,g} a_{h,g} X^{[h]} X^{[g]} \frac{1}{N}, \tag{2.5.14}$$

where $\nu^m_{g,h} = [-1, 0, 1]$ and $a_{h,g}$ is the transition rate of the process.

Another kind of transition is typically considered, i.e. a spontaneous transition from a compartment $[m]$ to another $[n]$. The latter commonly accounts for the spontaneous recovery of infected individuals from a certain disease and is described as

$$B \to A, \tag{2.5.15}$$

so that the variation of the number of individuals $X^{[m]}$ is given by $\sum_h \nu^m_h a_h X^{[h]}$, where again $\nu^m_h = [-1, 0, 1]$ and a_h is the transition rate. By accounting also for this term we get the general deterministic reaction rate equations for the quantity $X^{[m]}$:

$$\partial_t X^{[m]} = \sum_{h,g} \nu^m_{h,g} a_{h,g} X^{[h]} X^{[g]} \frac{1}{N} + \sum_h \nu^m_h a_h X^{[h]}. \tag{2.5.16}$$

By using this result we can easily derive the dynamical equations for the three basic models here introduced, i.e. SI, SIS, SIR.

2.5.3 Epidemics models in homogeneous networks

We will now quickly recall the main properties of such processes undergoing on homogeneous networks and we will later show the differences found when considering the same processes on top of heterogeneous networks.

We will start considering the SI model in which every node can only exist in two discrete states, susceptible or infected. The probability that a susceptible acquires the infection from any given infected neighbor in a time interval dt is proportional to the pathogen spreading rate β. In this model individuals that enter in the I state remain permanently infectious. The $I(t)$ or $i(t) = I(t)/N$ can just increase over time. Every infected node attempts to infect a connected susceptible vertex with probability βdt. The probability of getting infected having n infected neighbors is $1 - (1 - \beta dt)^n$. Considering k contacts per node and assuming $\beta dt \ll 1$ we find

$$1 - (1 - \beta dt)^{ki} \simeq \beta k i dt, \tag{2.5.17}$$

that, plugged in the Eq. (2.5.16) gives us

$$\partial_t i(t) = \beta \langle k \rangle i(t)[1 - i(t)], \tag{2.5.18}$$

as we already found in Eq. (2.5.13), and note that $1 - i(t) = s(t)$.

Also in the SIS model individuals exist in two class only. The disease transmission is described as in the SI model but infected individuals may recover and come back in the susceptible class again due to a spontaneous transition with probability μdt, where μ is the recovery rate. Using the general Eq. (2.5.16) for this processes we have:

$$d_t i(t) = -\mu i(t) + \beta\langle k\rangle i(t)[1 - i(t)]. \tag{2.5.19}$$

the $S \to I \to S$ cycle can lead to a stationary endemic state with a constant number of infected individuals and thus a time independent fraction $i(t) = i_{eq}$.

In the same way we can derive the evolution equations for the SIR model. In this model the infected individuals recover with rate μ and enter a new compartment R of removed individuals [149]. Using the general Eq. (2.5.16) for our three compartments model we get:

$$d_t s(t) = -\beta\langle k\rangle i(t)[1 - r(t) - i(t)], \tag{2.5.20}$$

for S,

$$d_t i(t) = -\mu i(t) + \beta\langle k\rangle i(t)[1 - r(t) - i(t)], \tag{2.5.21}$$

for I, and finally

$$d_t r(t) = -\mu i(t), \tag{2.5.22}$$

for R. In this model as in the SIS model we have a characteristic time scale μ^{-1} for the recovery of individuals. The processes of infection and recovery are competing and one may rule out the other depending on the infection and recovery rate relative values (e.g. for $\mu^{-1} \ll \beta^{-1}$ the recovery is faster and there is no time for the infection to outbreak and viceversa).

All the models defined so far can be easily solved at the early stage of the epidemics when we can assume that the number of infected individuals is a very small fraction of the whole population. We can solve the differential equations in the limit $i(t) \ll 1$ with a linear approximation neglecting all the terms order $\mathcal{O}i(t)^2$. The solution of the simplified version of Eq. (2.5.18) for the $i(t)$ population in the SI model reads:

$$i(t) \simeq i_0 e^{\beta\langle k\rangle t}, \tag{2.5.23}$$

where i_0 is the density of infected individuals at the beginning of the evolution. It is clear, since in the exponential all the factors are positive, the epidemic always propagates in the population infecting all the individuals.

By repeating the above procedure we can find a solution for $i(t)$ for the SIS and SIR models starting from Eq. (2.5.19) and Eq. (2.5.21), respectively. The common dynamical equation reads:

$$d_t i(t) \simeq \mu i(t) + \beta\langle k\rangle i(t), \tag{2.5.24}$$

whose solution is

$$i(t) \simeq i_0 e^{t/\tau}, \tag{2.5.25}$$

where $\tau^{-1} = \beta\langle k\rangle - \mu$. In this case the argument of the exponential is not defined positive and, if

$$\beta < \frac{\mu}{\langle k\rangle}, \tag{2.5.26}$$

we get a negative term at the exponential so that the infection dies out exponentially fast. The epidemic outbreak will not affect a finite portion of the population and will die out in a finite time. We have then an epidemic threshold:

$$\tau^{-1} = \mu(R_0 - 1) > 0, \tag{2.5.27}$$

where $R_0 = \beta\langle k\rangle/\mu$ is the basic reproductive rate in the SIS and SIR models. The spreading will occur provided that $R_0 > 1$. However, the stochastic fluctuations may lead to the extinction of the epidemics even if the system is above the threshold. As shown in [150] the extinction probability of a epidemic starting with i_0 infected individuals is:

$$P^{\text{ext}} = \frac{1}{R_0^{i_0}}. \tag{2.5.28}$$

2.5.4 Epidemics processes in heterogeneous networks

The results found so far applies for homogeneous networks. However, as we already showed in Chapters 1 and 2 many real social and technological networks (mobility networks, the web of sexual contacts and internet) are far from being homogeneous. The hypothesis that each individual in the system has the same number of connections $k \simeq \langle k\rangle$ (that we used in the previous sections) is not a good approximation. The fluctuations play a main role in determining the epidemic properties and the spreading may be favored in heterogeneous networks [50, 71, 128, 146].

Even in this case we will consider a degree block approximation: all nodes with the same degree are statistically equivalent. The quantity we will study are:

$$i_k = \frac{I_K}{N_K}, \quad s_k = \frac{S_K}{N_K}, \tag{2.5.29}$$

and the corresponding global averages read

$$i = \sum_k \rho(k) i_k, \quad s = \sum_k \rho(k) s_k. \tag{2.5.30}$$

As for the SI model we already know that the system will be totally infected independently of the spreading rate.

However, the topological fluctuations introduce an interesting effect on the spreading velocity. Considering the class of degree k and defining $\theta k(t)$ the density of infected neighbors of vertices of degree k the evolution equations read:

$$d_t i_k(t) = \beta[1 - i_k(t)] k \theta_k(t). \tag{2.5.31}$$

While in a homogeneous network the last term is equal to the density of infected nodes, in a heterogeneous network it is in general a very complicated term that has to account for the different degree classes and their connections. The simplest case to analyze is a network with no degree correlations:

$$\theta_k(t) = \theta(t) = \frac{\sum_{k'}(k'-1)P(k')i_{k'}(t)}{\langle k\rangle}. \tag{2.5.32}$$

By substituting Eq. (2.5.32) in Eq. (2.5.31) we get

$$d_t i_k(t) = \beta k \theta(t). \tag{2.5.33}$$

Then by multiplying both sides of this expression by $\sum_k (k-1)P(k)$ and summing over k we get:

$$d_t i_k(t) = \beta\theta(t)\left(\frac{\langle k^2\rangle}{\langle k\rangle} - 1\right). \tag{2.5.34}$$

By fixing $i_k(t=0) = i_0$ we get:

$$i_k(t) = i_0\left[1 + \frac{k(\langle k\rangle - 1)}{\langle k^2\rangle - \langle k\rangle}(e^{t/\tau} - 1)\right], \tag{2.5.35}$$

where

$$\tau = \frac{\langle k\rangle}{\beta(\langle k^2\rangle - \langle k\rangle)}. \tag{2.5.36}$$

It is clear that the fraction of infected individuals increases exponentially, with a faster growth for higher degree nodes. The growth time scale is measured by the heterogeneity ratio $\langle k^2\rangle/\langle k\rangle$, that, for scale free networks with exponent $2 < \alpha \leq 3$, is diverging as $N \to \infty$. Then, in an uncorrelated scale-free network we would have a virtually instantaneous rise of the epidemic size. The reason for that is quite intuitive. Once the disease has reached the hubs it can spread rapidly among the network. Multiplying both sides for $\rho(k)$ and summing over all the k values we can find $i(t)$:

$$i(t) = \left[1 + \frac{\langle k\rangle^2 - \langle k\rangle}{\langle k^2\rangle - \langle k\rangle}\left(e^{t/\tau} - 1\right)\right]. \tag{2.5.37}$$

If we let some correlations on the nodes degree into the system we can write [128]

$$\theta_k = \sum_{k'} i_{k'} \frac{k'-1}{k'} P(k'|k). \tag{2.5.38}$$

Neglecting terms of $\mathcal{O}(i^2)$ order we get:

$$d_t i_k(t) = \sum_{k'} \beta k \frac{k'-1}{k'} P(k'|k) i_{k'}(t) \equiv \sum_{k'} C_{k,k'} i_{k'}(t), \tag{2.5.39}$$

i.e. a linear system of differential equations given by the matrix $C = C_{k,k'}$. The solution will be a linear combination of exponential functions of the form $e^{\Lambda_i t}$, where Λ_i are the eigenvalues of the matrix C. We can then approximate

$$i(t) \sim e^{\Lambda_m t}, \tag{2.5.40}$$

using the largest eigenvalue Λ_m that can be express by a lower bound coming from the Frobenius theorem [100]

$$\lambda_m^2 \geq \min_k \sum_l (l-1)P(l|k)(k_{nn}(l)-1), \tag{2.5.41}$$

that diverges for scale-free networks with $2 < \alpha \leq 3$ as k_{nn} is diverging to, thus confirming the exponentially fast outbreak of the epidemic process in the network.

By following the same path we can solve the SIS and SIS models. In these cases we get the starting equation that reads

$$d_t i_k(t) = \beta k s_k(t)\theta_k(t) - \mu i_k(t), \tag{2.5.42}$$

where $s_k(t) = 1 - i_k(t)$ for the SIS and $s_k(t) = 1 - i_k(t) - r_k(t)$ for the SIR model. By re-applying the linear approximation and the uncorrelated network we find the characteristic time scale of the dynamical process to be

$$\tau = \frac{\langle k \rangle}{\beta \langle k^2 \rangle - (\mu + \beta)\langle k \rangle}, \tag{2.5.43}$$

that can be either positive or negative. Indeed, $\tau > 0$ if

$$\frac{\beta}{\mu} \geq \frac{\langle k \rangle}{\langle k^2 \rangle - \langle k \rangle}. \tag{2.5.44}$$

For scale-free networks with exponent $2 < \alpha \leq 3$ in the limit of infinite size the second moment diverges, so we have a null epidemic threshold. In real cases the threshold is not zero, but really small. This is an important result that confirm how heterogeneous networks behave in a completely different way from homogeneous networks. Scale-free networks are then an ideal topology for the spreading of infectious diseases. Fortunately the prevalence for small spreading rates is very small and this plays as an advantage in vaccination strategies of great effectiveness.

Besides the early stages of the dynamics evolution also the $t \to \infty$ is very interesting. For the SI model we trivially find that $i(t \to \infty) = 1$.

In the SIS model instead we can set Eq. (2.5.42) to equal 0 so that we find the stationary regime of $i_k(t \to \infty) = i_k$. By substituting i_k in Eq. (2.5.32) we find

$$\theta = \frac{1}{\langle k \rangle} \sum_k (k-1)P(k)\beta k \theta \mu + k\beta\theta. \tag{2.5.45}$$

We can then explicitly calculate the epidemic threshold from this equation as shown in [151] just considering that the condition is given by the value of β and μ for which it is possible to obtain a non-zero solution θ^*. Eventually we find:

$$\frac{\beta}{\mu} = \frac{\langle k \rangle}{\langle k^2 \rangle}. \tag{2.5.46}$$

If the degree distribution $\rho(k) \propto k^{-\alpha}$ we find

$$\frac{\beta}{\mu} = \begin{cases} \frac{\alpha-3}{m(\alpha-2)} & \text{if } \alpha > 3 \\ 0 & \text{if } \alpha \leq 3. \end{cases} \tag{2.5.47}$$

To find the density of infected individuals in the stationary state we then have to solve the self-consistent equation for θ in the limit β/μ approaching the epidemic threshold. Depending on the value of α we find:

$$i_\infty = \begin{cases} e^{\frac{-mu}{m\beta}} & \text{if } \alpha = 3 \\ \left(\frac{\beta}{\mu}\right)^{\frac{1}{3-\alpha}} & \text{if } 2 < \alpha < 3 \\ \left(\frac{\beta}{\mu} - \frac{\alpha-3}{m(\alpha-2)}\right)^{\frac{1}{\alpha-3}} & \text{if } 3 < \alpha < 4. \end{cases} \tag{2.5.48}$$

Finally in the SIR model $i_\infty = 0$. The epidemics dies due to the depletion of the susceptible individuals that after recovering move into the removed compartment. The interesting quantity is then the total number of individuals affected by the infection: $r_\infty = \lim_{t\to\infty} r(t)$. By starting from the system of differential equation for the SIR model in the assumption that $i_k(0) = i_0 \simeq 0$ and $s_k(0) \simeq 1$, we can integrate directly the equation getting:

$$s_k(t) = e^{-\beta k \Phi(t)},\ r_k(t) = \mu \int_0^t d\tau i_k(\tau), \tag{2.5.49}$$

where

$$\phi(t) = \int_0^t d\tau \theta(\tau) = \frac{1}{\langle k \rangle \mu} \sum_k (k-1)\rho(k) r_k(t). \tag{2.5.50}$$

By taking the derivative of both sides of Eq. (2.5.49) and by imposing the stationary solution by forcing $d_t \Phi(t) = 0$ we find

$$\mu \Phi_\infty = 1 - \frac{1}{\langle k \rangle} - \frac{1}{\langle k \rangle} \sum_k (k-1)\rho(k) e^{-\beta k \Phi_\infty}, \tag{2.5.51}$$

whose non-trivial solution is bound to the $r_\infty = \sum_k \rho(k)(1 - e^{\beta k \Phi_\infty})$, defining the well known epidemic threshold

$$\frac{\beta}{\mu} > \frac{\langle k \rangle}{\langle k^2 \rangle - \langle k \rangle}. \tag{2.5.52}$$

As a last remark let us note that these results show how the heterogeneity found in real networks lowers the epidemic threshold. This is a worrying scenario. However it is possible to take advantage of the heterogeneity developing new defensive strategies extremely effective. Just as a remark, one of the most common method leverage on the scale-free network property to be strongly affected by targeted damage. If a few of the most connected nodes are removed the network suffers a huge reduction of its ability to carry informations [15]. This can be extremely helpful in case of the spreading of infectious diseases. A targeted immunizations in which we progressively make immune the hubs will be very effective since the principal actors in the spreading will be blocked.

Chapter 3

Time Varying Networks

In this chapter we present in Sections 3.1 and 3.2 a quick review of the systems and the situations in which a temporal dimension is needed so as to correctly model both the network evolution and the dynamical processes taking place on top of it. Then in Section 3.3 we introduce the theoretical and modeling framework of the *Activity-Driven* networks, a particular class of temporal network to which our model belongs. In Section 3.4 we present the first measures of the social capital allocation, together with some known results. Finally, in Section 3.5 we will introduce the concept of burstiness and its main properties and known features.

3.1 The need for a temporal dimension

The classical, static approach of network science has recently switched to a time dependent representation of the network structure and edges dynamics. Indeed, the static networks representation happens to be unrealistic and wrong in many social and technological systems where the dynamics of the topology and edges rearrangement is comparable to (if not faster than) the time scales of the dynamical processes undergoing on the network fabrics. In these systems, interactions may last for an infinitesimal time interval or rather for a non-negligible period of time. Think for example at a phone call lasting several minutes and the quick interaction between two web servers exchanging a small text email. It is also reasonable to expect that, like we observed with the static network topology, the temporal structure and dynamics of edges activation may affect the dynamical processes and their outcomes through the network.

The introduction of a temporal dimension challenges our mathematical and computational modelling and completely changes the focus of the research work: in the traditional network theory one was mainly interested about the how and when something was happening in the dynamical system on the network. Now the dynamics and the systems are the network itself.

Moreover, the variation of the networks' connectivity patterns and the ongoing dynamical processes are usually coupled, as their time scales are comparable. This additionally complicates the analytical treatment of these systems and calls for a robust and thorough study of the tools to use and approximations that one can apply to study these systems.

Given the complication introduced it is worth asking when the framework of temporal networks is a suitable choice for the analysis and modeling of complex networks? It is clear that the temporal structure should share similar properties with respect to the static complex networks framework: the system should consist of agents that interact pairwise, the interactions should have both some degree (but not too much) of randomness and some (but again not too much) regularity.

On one hand, the historically adopted procedure to project a temporal network to a static graph will always result in a loss of information (see Fig. 3.1 for details). In other cases, however, this loss of information is a reasonable trade-off to make for the more complicated analysis and modeling needed for the temporal graph approach.

Another important ingredient is the dynamical system of interest as different dynamical processes may respond differently to a specific temporal structure. For instance, a bursty (i.e. as we will see in Section 3.5, a temporal contact pattern with sudden increase of activity followed by long pauses of the edges activation) edge dynamics could spread more easily a disease or an information through the network [59, 152–162]. In a situation with more uncorrelated or evenly distributed times of contact, the disease's concentration (or the individuals' interest about a rumor) would have time to fall below the dangerous level between the contacts. Thus for such a dynamical system, bursty edge activity would play a far more important role than for a system where the dynamics can be modeled as a Poissonian process, as is the case for many network-based models of disease spreading.

As already pointed out, a sufficient condition for a system to suit a temporal-network framework relates to the time scales of the dynamical system on the network: if the former is too rapid compared to the dynamics of the contacts, then there is no need to model the system as a temporal network. The prototypical example is the Internet where the data packets travel much faster than the topology changes. To sum it up, if the system's temporal and topological connections evolution is coupled with the dynamics of interest, then temporal networks is the ideal theoretical framework to stick with.

As the temporal networks framework can be applied to a variety of real-world systems the field of research is quite interdisciplinary. This resulted in a variety of names and terminology given to the object of study: temporal graphs, evolving graphs, time-varying networks, time-aggregated graphs, time-stamped graphs, dynamic networks, dynamic graphs, dynamical graphs, and so on [163–166].

3.2 Types of temporal networks

The most commonsensical and trivial example of a temporal interaction is the person-to-person communication. The recently availability of records of electronic one-to-one communication are particularly suitable for the temporal network approach. These datasets have been used to model spreading dynamics of information or electronic viruses and, in accordance with the results, several immunization and

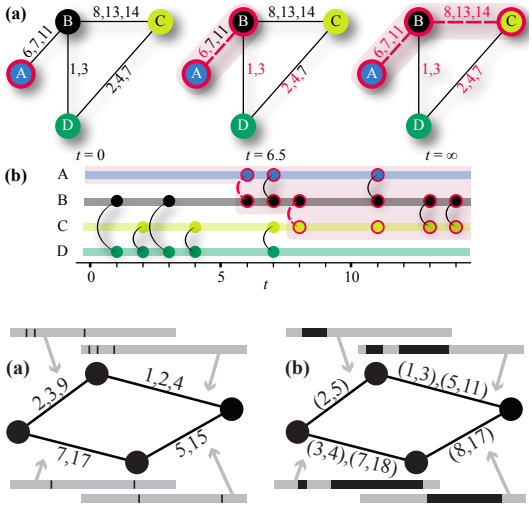


Figure 3.1: (Top) Illustration of a simple temporal network and its main differences with respect to a static one, i.e. the reachability issue and the intransitivity of temporal networks. In (a), the times of the contacts between vertices A–D are indicated on the edges. Assume that, for example, a disease starts spreading at vertex A and spreads further as soon as a contact occurs. The dashed lines and vertices show this spreading process for four different times. The spreading will not continue further than what is indicated in the $t = \infty$ picture, i.e. D cannot get infected. However, if the spreading started at vertex D, the entire set of vertices would eventually be infected. Aggregating the edges into one static graph cannot capture this effect that arises from the time ordering of contacts. Panel (b) visualizes the same situation by showing the temporal dimension explicitly. The colors of the lines in (b) matches the vertex colors in (a).
(Bottom) Contact sequences and interval graphs. This figure illustrates the two fundamental temporal network representations, i.e. contact sequences (a) and interval graphs (b). The times of the contacts are shown next to the edges. The contacts time-lines (grey bars) are shown for clearness. In these the contacts are marked by black bars or fields and the time lines range from $t = 0$ to $t = 20$. In the former, contacts occur at points in time while the contacts are extended in time in the second case. Image credits from reference [165].

control procedure has been developed for containing the spread of malware and electronic viruses in mobile devices. These are based on the temporal centrality measures that we will quickly introduce later.

Another class of real-world systems allowing for a temporal description are the one-to many information dissemination processes. These are, for example, the broadcast of information through microblogs (like Twitter) [167, 168], e-mails and post on on-line social media [169] and posting of information on websites. In the latter case, the time dimension is studied to analyze the circadian patterns of Wikipedia editorial activity, allowing to estimate the geographical distribution of editors [170].

An additional example of temporal contacts are the proximity patterns of humans, i.e. the contact pattern on who is close to whom at what time. These kind of data have historically been gathered in small scale and with questionnaires to

confined space gatherings of people like fraternities or offices. The recent availability of cheap Radio-Frequency-Identification-Devices (RFID) and the Bluetooth technology installed on the majority of mobile devices allowed to collect large scale and long time interactions of several people in open space. The pioneering studies was carried out by the Reality Mining [171] and the SocioPatterns project [163]. The RFID methodology has been implemented in networks of patients [172], school children [173], and conference attendees [72, 72, 163, 172, 174, 175]. Advanced data-mining procedures, spectral analysis methods and procedures borrowed from the tensor-theory have then been applied in order to develop, also in this case, optimal procedures to contain and isolate epidemics processes in such critical social systems as hospitals and schools [72, 76, 163, 172–176].

The temporal network approach can also be of great interest in many cell and microbiology systems [177]. For instance, the interactome, i.e. the set of molecular interactions in a cell whose vertices are proteins that can connect one to the other to perform biological functions. Despite these interactions are usually represented as a static graph, the biological functionality stems from the chain of control and temporal activation of the connections between these agents [178]. In the same way, one can also represent gene expression and regulatory networks as temporal networks [179], where the vertices are gene about to be transcribed, while an edge represents a generic functional relationships by which one gene affect the transcription of another (via feedback from RNAs or proteins).

Other biological systems allowing for a network formalism are the web of neural connections, which displays several levels of structural and temporal connectivity, from the spiking patterns of individual neurons to the functional connections put in place between brain areas working to perform similar tasks/response to the same set of external stimuli [180,181]. Another example are ecological networks capturing the interactions between species, either trophic, i.e. a sort of food-chain, or mutualistic, i.e. representing how two species engage a cooperation that benefits both [182].

As for the technological classes of networks, we list the systems that originally caused the interest and the early theoretical developments on temporal networks, i.e. the distributed computation units developed by computer scientists. These are composed by independent computational units spread out over some network [183].

For what concerns the infrastructural family (e.g. the air-transport network), in the majority of these networks the topology changes so slowly that there is no point in modeling them as temporal networks [164].

3.2.1 Temporal topological structure

Most of the so far introduced topological quantities, structural properties and observables do not trivially translate to the framework of temporal networks. Thus, many of these methods need to be updated from their definition dealing with static networks.

To this extent, let us first note that the temporal networks can be divided into two main families: instantaneous and interval graphs, the main difference being the temporal extension of the modeled interaction with respect to the dynamical process. We show the two types of representation in Fig. 3.1.

In the first representation (Fig. 3.1(B)(a)) the durations of the interactions are negligible and the system can be represented by a *contact sequence*, i.e. a set of C *contacts*, triples (i, j, t) where $i, j \in \mathcal{N}$ and t denotes the time at which the contact takes place. These representation applies to communication data (e-mails, phone calls, text messages, etc.), and physical proximity data where the duration of the contact is less important (e.g. sexual networks). The commonly adopted representation of these graphs is to group all the contacts taking place in a given time window (a discrete *time-step*) into one graph (or "graphlet"). The temporal network is then thought as a time sequence of such graphlets. Though simple to remember and explain, this representation is somewhat misleading as it makes it tempting to think about the temporal-network structure as an evolving static network structure, instead of the more correct definition of contact sequences. Furthermore, in many real datasets the density of active nodes at a given time is negligible, as there usually are only a few edges present among tens of thousands of vertices.

In the second class of temporal networks one is dealing with *interval graphs*. The activation of one edge occur in a given interval $T_e = \{(t_1, t_1'), \ldots (t_n, t_n')\}$, where the parentheses indicate the periods of activity and the unprimed times mark the beginning of the interval and the primed quantities mark the end. The static graph in which an edge between i and j is present if and only if there is a contact between i and j is called the *(time) aggregated graph*. The interval graphs framework applies to proximity networks, seasonal food webs where a certain specie is prey of another during some time of the year, and infrastructural systems like the Internet.

Like for static graphs, it can be useful to define the adjacency index, stating whether or not two nodes i and j are connected at time t

$$a(i, j, t) = \begin{cases} 1 & \text{if } i \text{ and } j \text{ are connected at time } t \\ 0 & \text{otherwise} \end{cases} \tag{3.2.1}$$

Another assumption usually made is that a triple of a contact sequence never occurs twice, i.e. the contacts can be uniquely ordered. The topological structure of static networks can be characterized by an abundance of measures (see, e.g., [184]).

The measures on the topological structure of static networks are either local (i.e. based on connections between neighbors, as the degree or clustering coefficient) or global (i.e. between larger sets of nodes, such as path lengths, network diameter and centrality). With the addition of a temporal dimension these quantities need to be re-defined.

For instance, the degree k is usually measured as the number of edges in the graph aggregated over a certain time window Δt. On the other hand, some other properties are directly influenced by the order of link activations. As an example, *paths* that transmit anything through the network need to follow time-ordered sequences of contacts that, just like the temporal networks, evolve with time. These paths are the hallways inside which a dynamical process is constrained to move. While in a static network a path is simply defined as a sequence of always active edges, in temporal networks paths must necessarily consist of sequences of link activations that follow one another in time. That is why we can define paths are as sequences of contacts with non-decreasing times that connect sets of vertices, i.e.

"time-respecting" paths [185]. As an example, in Fig. 3.1, we find time-respecting paths from A to D (for example $(\mathrm{A}, \mathrm{B}, 7), (\mathrm{B}, \mathrm{C}, 8), (\mathrm{C}, \mathrm{D}, 11)$) but none from A to E.

Let us also stress that the time-respecting connection is not transitive, in that the existence of time-respecting paths from i to j and j to k does not imply that there is a path from i to k. Indeed, the path from i to k via j exists only if the path connecting i to j completes before the beginning of the $j-k$ path.

The set of nodes that can be reached from node i in a time $t \in [t_0, T]$ via time-respecting paths is called the *set of influence* of i. The average fraction of vertices in the sets of influence of all vertices is the *reachability ratio* of the network. These two quantities are important in the analysis of disease spreading, as they quantify the set of vertices that can eventually be infected if i is the source of infection.

Even in the case of a well time-respecting connected graph, dynamical processes could be slowed down from too long waiting times spent on the vertices of the underlying network. Specifically, limitations on the times that a path is allowed to spend at vertices (i.e. the times between two consecutive contacts on a path) may be from below (i.e. the process must be allowed to wait for some time before the next contact on the path) or from above, as for spreading dynamics, where the transmission has to happen before infectious nodes recover, or before a piece if information looses its novelty so that nodes are not forwarding it anymore [164].

Let us note that connectivity, as in directed graphs, is not a symmetric relation for directed or temporal graphs. We can transpose the static definition of connectivity to temporal networks by means of the time-respective paths: two vertices i and j of a temporal network are *strongly connected* if there is a directed, time-respecting path connecting i to j and vice versa. Otherwise, they are *weakly connected* if there are undirected time-respecting paths from i to j and the other way around.

As for static networks we can also define a distance on a temporal network. While in static networks the *geodesic distance* between two vertices is defined as the length of the *shortest path* joining them, when the dimension of time is added one defines similar quantities characterizing how *quickly* vertices can reach each other through time-respecting paths. A possible way is to define distance as the *duration* of a time-respecting path, i.e. the time difference between the last and first contacts on the path (this is also dubbed *temporal path length*) [164]. In the same way, it is possible to define the *fastest* time-respecting path between two nodes, i.e. the shortest time within which a walker can move from i to j. This quantity is better known as the *latency*.

Latency can be used to "benchmark" the *velocity* of a network, i.e. measuring how quickly (on average) vertices can spread a piece of information or something else to a large fraction of the network along the contact sequences. Though simple to describe, the average definition is cumbersome as it strongly depends on the time window finiteness [164, 186].

An important set of tool to transpose in the temporal networks formalism is the one regarding centrality measures: these measures aim at the identification of

the important vertices beyond the degree. Depending on the task the network has to perform, a rank of the node can be done accordingly to the average distance to other vertices or by the importance for shortest paths connecting other vertices (see Section 1.1.2 for details).

As one could have expected, even in this case we do not have a unique candidate for the temporal version of a centrality measure. The most straightforward approach is to simply replace the role of paths in static networks by time-respecting paths. Following this prescription, the *closeness centrality* C_C [16] for static networks is defined as

$$C_C(i) = \frac{N-1}{\sum_{j \neq i} d(i,j)}, \tag{3.2.2}$$

where $d(i,j)$ is the geodesic distance between i and j. For temporal networks the focus is on how quickly a vertex may on average reach other vertices, thus defining the temporal closeness centrality as [187]

$$C_C(i,t) = \frac{N-1}{\sum_{j \neq i} \lambda_{i,t}(j)}, \tag{3.2.3}$$

where $\lambda_{i,t}(j)$ is the latency between i and j. This quantity, just like the static closeness centrality, is not working unless there is a time-respecting path from j to i ending within time t. To get rid of this strict constraint (especially if t is late in the observation period) it is possible to average out the time dependence [164].

The downside of this approach is that there it completely drops information belonging to the intervals with infinite latencies. To overcome this limitation an augmented closeness centrality measure based on reciprocal latencies (in the same way as the above-mentioned "efficiency") can be defined [164, 188].

Finally, *betweenness centrality* C_B is another important centrality quantity that measures the fraction of shortest paths passing through the focal vertex (or edge) [189, 190]. Let us recall that, on static networks, betweenness centrality is formally defined as in Eq. (1.1.4). This definition can be easily generalized to temporal networks: we add a dependence on time t and count the fraction of either shortest or fastest time-respecting paths that pass through the focal vertex [191].

Another way to tackle the problem of centrality measures is to take advantage of a diffusion process that randomly travels through the network. The rank of a node can then be measured as the probability for that vertex to be occupied by that walker [192]. While in static graphs this approach yields matrix-based centrality measures like the *eigenvector centrality* and *PageRank* [16, 190], in temporal networks this measure deals three-dimensional tensors representing the network. To avoid the need to work with tensor algebra, a possible solution is the following one:

1. Start with a centrality value 1 at each vertex.

2. At every contact between vertices i and j, let the C_E values after the contact at time-step t be

$$C_E^{t+1}(i) = \zeta C_E^t(i) + (1-\zeta) C_E^t(j) \tag{3.2.4}$$

and

$$C_E^{t+1}(j) = \zeta C_E^t(j) + (1-\zeta) C_E^t(i). \tag{3.2.5}$$

By cycling over the whole connection sequences, we obtain a centrality distribution that can be used as a generalization of the eigenvector centrality. The parameter ζ sets the rate of centrality transmitted at a contact: a larger value of ζ puts a bigger emphasis on recent contacts.

3.2.2 Models of temporal networks

As already shown in Chapter 2, the development of a model generally has two purposes: explain the mechanisms generating peculiar properties of certain systems (e.g. the preferential attachment to explain the scale-free degree distribution in complex networks) or to provide a synthetic "playground" with tunable characteristic where to test the outcomes of dynamical processes via numerical simulations. Another application of real-world datasets is the randomized reference networks, a framework aiming at a deeper understanding of the role of correlations in real-world spreading processes: the recorded contact sequences are rearranged by selected randomization procedures. The outcome of a dynamical process on this random version of the network is then compared to the corresponding real-world version. The number of time varying networks found in the literature is booming and we will here introduce the historically relevant ones, together with some randomized reference models. Then, in Section 3.3 we will introduce the main framework used in this work, i.e. the activity driven networks family.

Temporal exponential random graphs Based on their static definition (see Section 2.4 for details), this group of models has a connection to the Ising model [193, 194]. Indeed, one can find also in this case a partition function that, within this framework, is a time-varying one. The parameters regulating the temporal dependence of the network contacts are inferred from empirical data and they usually tune the importance and frequency of chosen topological elements and subgraphs, such as triangles and stars. As for the centrality measures, the basic problem in temporal exponential random graphs is the measure of such parameters. The problem is that given an observed time evolution, one should use the same set to determine the probability of a contact to happen between a pair of vertices at a given time.

Despite this complication, the models belonging to the exponential random graphs family are commonly used in social science, both for quantifying the topological and temporal structures and as a reference generative model for dynamical processes.

Models of social group dynamics Basic mechanisms shaping the evolution of a group of individuals has been encoded in a framework for modeling social networks [156, 195, 196]. In this modeling scheme edges are transitory social ties, such as being in a conversation with a person. The evolution of these contacts is then modeled through a master equation that governs the entrance and exit rates of individuals from a group of a certain size. The main idea beneath this mechanism is a sort of reinforcement of interactions, as the longer an agent interacts with a group, the more it is likely to stick with the given group. The interactions modeled by this framework belong to the interval graphs, as contacts last for a long period of time.

Several models of this family also encode in their evolution rules the mechanisms belonging both to earlier social network models (e.g. focal and cyclic closure, tie strength reinforcement [20], and triad interactions) and to more modern processes such as the social interaction task execution from a priority queue, thus introducing a bursty activation pattern of the links [159].

Contact network models In this model one tries to extend the static graph framework to allow for a turnover of neighbors. The prescription is to select two edges with some probability every time step and swapping them. Since the edges are simply rewired, the topology and the contacts dislocation have to be determined using some other generative models [197].

The real-world situation that this kind of contact sequences try to mimic is the change of partnerships to generate contact structures for disease spreading simulations. The model rules are as follows [198]:

1. (a) A new partnership is formed with probability ρ. The individuals participating in the pair are chosen according to the mixing function ϕ:
   i. draw two random individuals i and j;
   ii. decide whether they form a pair according to $\phi(i, j)$;
   iii. if yes, done; else go to i.
   (b) Repeat step 1a $N/2 - P$ times.
2. In every pair consisting of a susceptible and an infected, the disease is transmitted with probability λ.
3. Every pair splits up with probability σ.

The mixing function may tune the nodes' mixing by their degree to be assortative [16],:

$$\phi(i, j) = 1 - \xi + \xi \frac{k_i k_j}{k_{\max}^2} \tag{3.2.6}$$

or disassortative:

$$\phi(i, j) = 1 - \xi + \xi \frac{(k_i - k_j)^2}{k_{\max}^2}. \tag{3.2.7}$$

The parameter ξ sets the strength of the mixing pattern and $k_{\max}$ is an upper limit of the degree.

A wealth of following works explored possible development of the stochastic pair-formation model, as the "dynamic random graph models with memory" [199] and the evolving interval graphs with structure [200, 201].

3.2.3 Randomized reference models

When analyzing the networks' structural properties, in order to claim significance of quantitative topological characteristics of the empirical graph one needs to compare it with a reference model. In static networks the configuration model is usually adopted as a benchmark, as it yields a maximally random network by randomly rewiring pairwise the links of the original network. The differences between the empirical and the random network can then be made by a direct comparison to

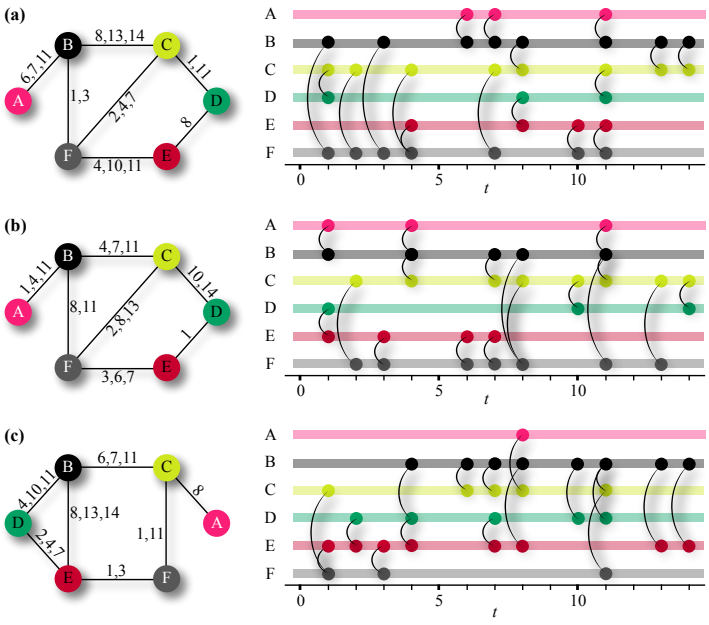


Figure 3.2: Illustration of two types of randomization null-models for contact sequences. (a) shows a contact sequence (the same as in Fig. 3.1). In (b) we apply the Randomly Permuted times procedure that preserves the network's topology and so that contacts happen the same number of time per edge. In (c) the contact we show the Randomized edges (RE) procedure instead. Here the conserved pattern is the time sequence of the contacts along an edge, as well as the degree sequence of the original network. However, all other structures of the topology are destroyed.

averages in the randomized reference ensemble or by measuring to which extent the dynamics of a certain process differs when run on the empirical networks and on the reference ensemble.

A similar approach can be applied to the contact sequences of temporal graphs. In the latter, one randomly reshuffles the event sequences accordingly to a given scheme. This is done so as to remove all the possible time-domain structures and activity correlations on nodes and edges. However, given the (many) different structures and time-scales of temporal correlations, a unique and general-purpose procedure procedure cannot be developed. Each null model rather deletes a selected type of correlations so that one can isolate their contribution to either the topological characteristic or dynamical processes in a given time-domain of the real-world network. Given the possibility to sharply switch-off a given type of correlations, it is customary to apply in sequence all of these procedure and monitor how the dynamics of the process depends on the reference models, to isolate the role of different temporal and topological correlations on the process.

We quickly review here some of these temporal null models that can be applied to contact sequence [186, 202].

Randomized edges (RE) This method can be interpreted as the temporal analogous of the configuration model. From a practical point of view, the procedure is

defined as follows:

1. Go over all edges sequentially.
2. For every edge (i, j), pick another edge (i', j').
3. With a probability 1/2 replace (i, j) and (i', j') by (i, j') and (i', j), otherwise replace them by (i, i') and (j, j').
4. If the move in step 3 created a self-edge or multiple edge, then undo it and start over from step 2.

The contact between nodes lasts a for a constant, small time lapse. The two alternatives in step 3 are to avoid possible correlations due to the datasets structure, that may return the vertices of an edge in a specific order.

By applying this procedure we are switching-off the effect of the network topology, that is, the wiring diagram of the original network. An assumption implicitly made is that the dynamics is governed by the characteristic times of the edge activation, rather than the nodes activity. Indeed, this procedures changes both the number of activation and the time of each contact for each vertex, while both the degree k and the temporal correlations are preserved (see Fig. 3.2 for details).

Randomly permuted times (RP) The complementary approach of the RE procedure is to keep the topological structure fixed and reshuffle the contact times. This is a much simpler approach with respect to the edge rewiring scheme discussed above and it can be accomplished in mainly two ways: randomly exchange the time stamps of all contacts, or just randomly reshuffle the order of time stamps in an array or vector. As this procedure switches-off the temporal correlations on edges activation and completely destroys the time ordering of the event, it is frequently applied to study the effects of burstiness and inter-contact time distributions on dynamical processes. On the other hand, this procedure retains the topological structure and link weight, together with the overall rate of events in the network at every point in time (see Fig. 3.2 for details).

Randomized edges with randomly permuted times (RE + RP) This null model combines the RE and the RP (in this exact strict order) so that it firstly randomize the network structure and then destroys the temporal correlations of all the contacts. The outcome is then a temporal graph, where all the structural and temporal correlations are removed. The only quantity retained is the overall rate of contacts (such as daily/weekly pattern), as the total number of event in a given time window does not change.

Random times (RT) To account for how temporal rates of contacts affect a dynamical process, one needs to go over the RP approach. Indeed, although it destroys burstiness of events on individual vertices and edges as well as correlations between events, the aggregated rate of events will still follow the typical circadian and weekly patterns of human activity [203–208].

Even though there are evidences that these activity patterns do not affect simple epidemics processes such as SI spreading [186, 207], there are other situations in

which such patterns play a role more important than the mere burstiness of contacts: SIS or SIR spreading models.

To account for circadian and seasonal patterns, the random times (RT) null model destroys such regularities: each contact on each edge is assigned a random time within the observation time window of the original data. This procedure still preserves the network structure and the total number of activations of each edge as for the RP null model. Any difference of the outcomes of a dynamical process run on a RP and RT realizations should then be due to the network-level temporal patterns of the empirical data. Besides the flat, homogeneous extraction of the activation times for each contact event it is also possible to re-arrange them accordingly to any chosen distribution or process, such as the Poisson process [186].

Randomized contacts (RC) This procedure preserves the graph topology while it randomly re-distributes the contacts among the E edges. Given an initial skewed, heavy-tailed edges' weight distribution we end up with a graph whose edges' weights are homogeneously distributed following a binomial distribution. In this way, the effect of the heterogeneity on number of contacts per edge is ruled out, together with the order of events.

3.3 The Activity-Driven approach

All the models presented so far display (sometimes not at once) two main limitations: they all are connectivity driven and they assume the time scales of the topological rearrangement to be either larger or shorter than the dynamical process undergoing on top of the network.

Even though these models, mainly represented by the preferential mechanism, has been under a thorough investigations in the last 15 years [13,50,79,100,111,117], the possibility for the interactions among the elements to happen on a very short scale is not taken into consideration. However, in many real-world systems the rearrangements of contacts and the evolution of the topological structure happens at time scales challenging the dynamical processes undergoing on top of the system's network [165].

The modeling of such a limit (where topology evolution and dynamics are strictly coupled at least in their characteristic time scales) has been recently introduced in network science [209], while before that it was a prerogative of adaptive systems [210, 211].

Moreover, the characterization of how the network topology evolution and the connectivity dynamics affect dynamical processes has been studied in the two opposite limits: quenched network dynamics (the time scales of the edges dynamics are supposed to be much larger than the process itself) or of an annealed edges dynamics (the dynamical process evolves on a sample extracted from a reservoir of the network representations).

While this approach is extremely convenient from an analytical and numerical point of view, we already largely discussed that a precise modeling of both the network topology and the contact sequence features of a real-world network is ex-

tremely important. Indeed, different topological or temporal structures may results in (very) different outcomes of a diffusion or spreading process. The same, of course, holds for a correct characterization and modeling of the edges activation pattern and the time scales of both the inter-event time and the duration of the interactions [5, 175, 212, 213].

To this end, the activity driven approach aims at the modeling of the evolution of rapidly growing networks, capturing the process of accumulating connections over time. All the measurable quantities such as the cumulative degree, the corresponding degree distribution and other topological properties have to be thought in a time-integrated perspective of the system.

In their work, Perra *et. al.* [209], focused on an accurate modeling of the dynamics of activity-driven networks, by measuring the activity of the agents forming the system from empirical, time-resolved and large-scale network datasets.

The new concept they introduced is to move the dynamics and the link activity (i.e. its propensity to engage an interaction) from the links and topological measures (such as the degree of a node in the preferential attachment model) to the node itself. This is done by assigning to (or measuring for) each node i its activity potential a_i. The latter measures the fraction of interactions that node i performs with respect to the total number of events measured in a certain temporal windows. The activity then sets a clock (or activation rate) that determines the temporal interaction pattern of node i within the network.

The first analysis has been performed on three large-scale network datasets, by means of which a first measure of the activity potential was carried out. The three systems under investigation were: Collaborations in the journal "Physical Review Letters" (PRL) published by the American Physical Society, messages exchanged over the Twitter microblogging network, and the activity of actors in movies and TV series as recorded in the Internet Movie Database (IMDb). In the PRL dataset the network is composed by the authors of scientific papers published on PRL connected by an edge if the two authors collaborated in writing one article. In the Twitter case nodes correspond to the users and the (i, j) edge is drawn if i and j ever exchanged a message. Finally, the actor network is obtained by drawing an undirected link between any two actors who have participated in the same movie or TV series.

As anticipated, the *activity potential* a_i of node i is defined as the number of interactions that i_i performed, in a given time window Δt, by i divided by the total number of interactions I_{TOT} made by all the nodes in the same time window. As for many properties measure in real-world systems, the distribution of such activity potential $F(a)$ is found to be highly heterogeneous and heavy tailed as we show in Fig. 3.4 (these tails are often approximated as a power-law). This comes as no surprise, as real-world social networks and human activity are known to feature complex properties usually found in complex systems [159,207] (here we present this first approach while in Section 4.3 we will show that the results here introduced hold for other datasets).

Remarkably, on the contrary of the degree distribution and other structural

characteristics of the networks, the distribution $F(a)$ weakly depends on the time scale over which the activity potential is measured. Additionally, we find that the distribution $F_c(x)$ is skewed and fairly broadly distributed. In other words, the activity potential a_i estimates the probability that the agent i is involved in any given interaction in the system. This can be thought as the propensity for node i to engage in a social interaction. The dynamics of the system is then fixed by the activity distribution $F(a)$, i.e. the probability for a node i to have an activity potential a.

The role of agents' activity can be well understood by looking at Fig.3.3: as we broaden the time window to aggregate the interaction of nodes (we show the time-aggregated co-authorships over 1, 10, and 30 years), we observe a non-stationary and growing connectivity pattern. Moreover, both the cumulative degree distribution and the network structure are affected by the quickly evolving contact pattern among nodes.

3.3.1 Modeling framework

Given the definition of the activity potential a_i, we are led to a natural definition of the model. We start with a network $\mathcal{G}$ composed by N nodes (agents) and assign to each of them an activity potential a_i, defined as the probability per unit time δt to create new contacts or interactions with other individuals. The average number of active nodes per unit time in the system is $\langle a \rangle N$. Without loosing generality we can assume a unit time interval $\delta t = 1$, so that we can define the activity potential values to fall in the $\epsilon \leq a_i \leq 1$ range. Here, ϵ is a lower cut-off on a_i in order to avoid possible divergences of $F(a)$ for $a \to 0$.

The evolution process then reads as follows (see Fig.3.4-D):

- At each discrete time step t the network G_t starts with N disconnected vertices;
- With probability $a_i \delta t = a_i$ each vertex i becomes active and generates m links that are connected to m other randomly selected vertices. Non-active nodes can still receive connections from other active vertices;
- At the next time step $t + \delta t$, all the edges in the network G_t are deleted. From this definition it follows that all interactions have a constant duration $\tau_i = \delta t$.

The just outlined model is random and Markovian. Indeed, it has no memory (nodes choose to activate or not and who to contact regardless of their story) and all the interactions are established among randomly chosen nodes. The topological and evolution properties of such a network are then completely encoded in the activity potential distribution $F(a)$.

In particular, the interaction rates among nodes and thus the activity potential distribution define the network structure of both the cumulated and the instantaneous connectivity patterns (i.e. the snapshot of the network given by the active edges between time t and $t + \delta t$). Specifically, one can trace the origin of the degree distribution heterogeneity (i.e. the appearance of hubs) to the heterogeneous activity of the network elements.

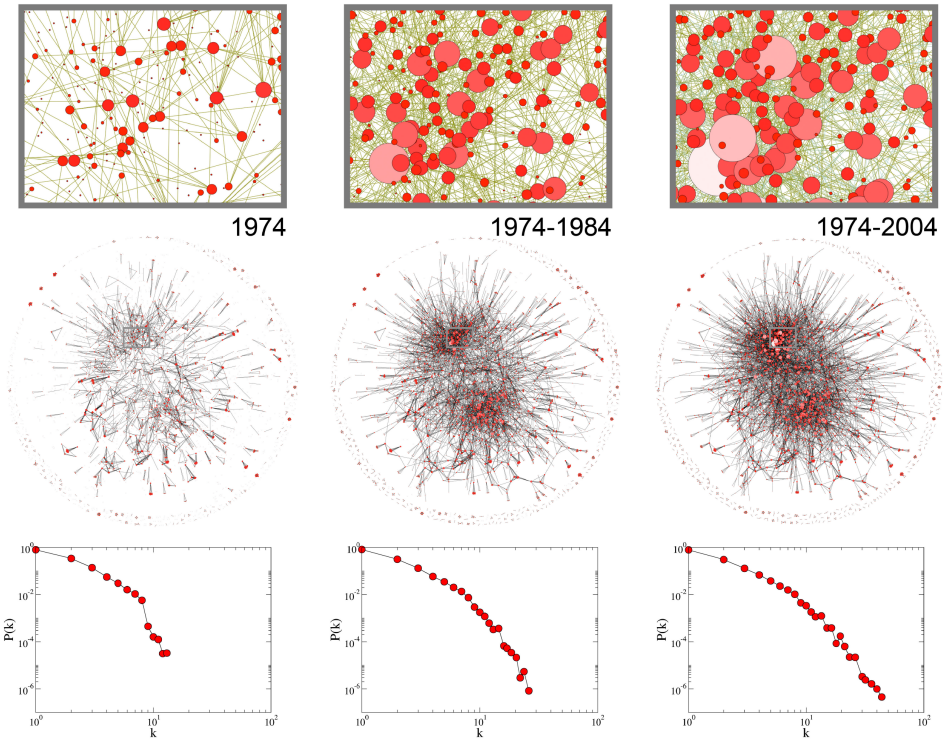


Figure 3.3: Network visualization and degree distribution of the PRL dataset considering three different aggregation time window. The first two rows show the set of authors who wrote at least one paper in the period between 1960 and 1974 (first column), from 1974 to 1984 (second column) and 1974 − 2004 (third column). The upper row represents a blown up perspective of a particular network region of the central row. In each visualization the size and color of the nodes is proportional to their degree. In the third row we show the degree distributions for the different time windows, that consider the complete set of authors.

Moreover, the model can be easily coupled with dynamical processes without the need of a time scale separation.

Extensive numerical simulations show that model, by plugging in a suitable activity distribution resembling the empirical one, recovers the same qualitative behavior observed in Fig. 3.3.

Taken aside, each temporal snapshot of the network evolution is a simple random graph with low average connectivity. By accumulating the connections activated in a given time window of length T we observe a skewed $P_T(k)$ degree distribution. Nodes with a larger activity value are then more likely to become hubs and vice-versa. The emergence of this highly connected nodes is no more due to a positional advantage in degree space (as in preferential attachment). In that model the nodes with higher degree benefit from a rich-get-richer mechanism caused by the passive attraction of more and more connections. In the activity driven model instead, the appearance of hubs stems from the very heterogeneity of social propensity measured

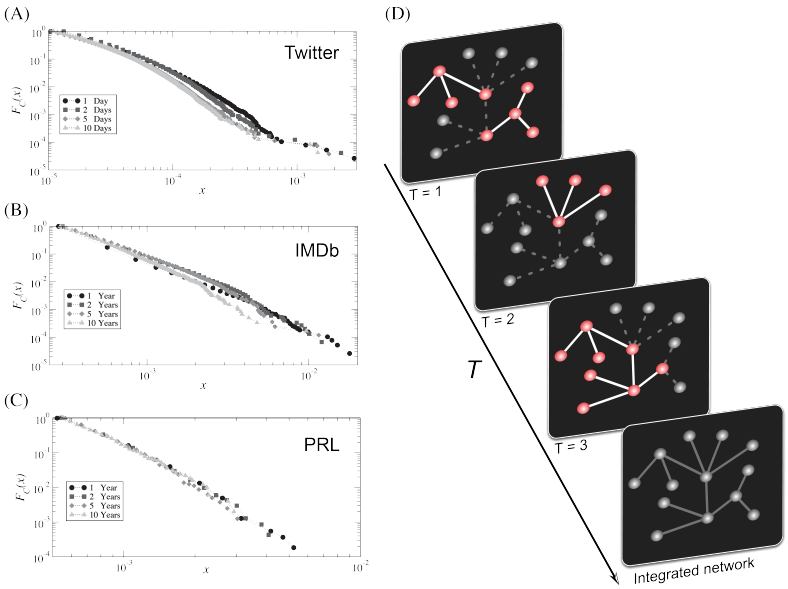


Figure 3.4: Cumulative distribution of the activity $F_C(a)$ as measured in the empirical datasets in four different time windows: (A) Twitter, (B) IMDb, and (C) for PRL. In panel (D) we show a schematic representation of the model ($N = 13$ nodes, $m = 3$ and we show three different time steps). The red nodes represent the firing/active nodes. The final visualization represents the network after integration over all time steps.

in the real-world system: hubs (or the most popular individual) are the nodes willing to repeatedly engage in social interactions.

These hand-waving considerations can be shown also from the analytical side. The integrated network $G_T = \bigcup_{t=0}^{t=T} G_t$ is defined as the union of all the networks obtained all the $T+1$ time steps. Each active node creates m links so that the average number of active edges per unit time is $E_t = mN\langle a \rangle$, yielding the average degree per unit time

$$\langle k \rangle_t = \frac{2E_t}{N} = 2m\langle a \rangle. \tag{3.3.1}$$

The instantaneous network is sparse and mainly composed by a set of stars with degree $k \geq m$ plus some vertices with low degree (the nodes that passively got called by the active nodes). The corresponding integrated network, on the other hand, will generally be dense, being the union of many random graphs at previous times. In fact, for large time T and network size N, we can approximate the nodes degrees as continuous variables, one finds that agent i will have at time T a degree in the integrated network given by (see Section 4.2.1 for details)

$$k_i(T) = N\left(1 - e^{-Tma_i/N}\right). \tag{3.3.2}$$

It can then easily be shown that the degree distribution $P_T(k)$ of the integrated network at time T takes the form:

$$P_T(k) \sim F\left[\frac{k}{Tm}\right], \tag{3.3.3}$$

Remarkably, the degree distribution is found to follow the same functional form of the individual activity distribution. While this results is recovered in numerical simulations, empirical data show a different scaling form of the two distributions. As we will discuss in Section 3.4, this is due to features that the random model does not capture: links already explored are more likely to get activated again, social relations have a lifetime distribution (persistence) and other structural constraints (communities or weighted interactions) may be relevant.

3.3.2 Dynamical processes

We now present some recent results regarding dynamical processes undergoing on top of a temporal activity driven network.

First of all we analyze a simple SIS epidemic process and provide the explicit analytical expression for the epidemic threshold.

In an homogeneous population the behavior of the epidemics is controlled by the reproductive number $R_0 = \beta/\mu$, where $\beta = \lambda\langle k \rangle$ is the per capita spreading rate that takes into account the rate of contacts of each individual (see Section 2.5.3 for details). As the reproductive number identifies the average number of secondary cases generated by a primary case in an entirely susceptible population, the epidemic threshold is set at $R_0 = 1$: if $R_0 > 1$ the epidemics can reach an endemic state and spread into a closed population. In complex networks we have shown that the epidemic threshold depends on the topological properties of the networks. In

particular, for annealed network the heterogeneous mean-field approaches predict an epidemic threshold that is inversely proportional to the second moment of the network's degree distribution [15, 214]:

$$\beta/\mu > \langle k\rangle^2/\langle k^2\rangle. \tag{3.3.4}$$

These results do not hold if we introduce a time dependence of the connectivity pattern, especially if this process occurs on the same time scale of the dynamical process (see Fig. 3.5 for details). In particular, the two predictions applied to the time-aggregated network lead to misleading results in both the threshold and the epidemic magnitude as a function of β/μ. The problem resides in the fact that the static approach assumes that all the edges are always available to carry the contagion process. So, even though the epidemic threshold discounts the different average degree of the networks in the factor $\beta = \lambda\langle k\rangle$, the temporal nature of the network is completely neglected.

To account for the activity pattern of the agents we have to work with no time aggregation of the network connectivity and with a mean-field approximation. By imposing a functional form of the activity distribution $F(a)$, we write the evolving equation for the number of infected individuals in the class of activity rate a, at time t, i.e. I_a^t. The number of infected individuals of class a at time $t + \Delta t$ given by:

$$I_a^{t+\Delta t} = -\mu\Delta t I_a^t + I_a^t + \lambda m(N_a^t - I_a^t)a\Delta t \int da' \frac{I_{a'}^t}{N} + \lambda m(N_a^t - I_a^t)\int da' \frac{I_{a'}^t a'\Delta t}{N}, \tag{3.3.5}$$

where N_a is the total number of individuals with activity a. In Eq. (3.3.5), the third term on the right side takes into account the probability that a susceptible of class a is active and contacts an infected node acquiring the infection with probability λ. The last term takes into account the possibility for a susceptible node to get contacted from any infected active individual. The solution of Eq. (3.3.5) yields the following epidemic threshold for the activity driven model:

$$\frac{\beta}{\mu} > \frac{2\langle a\rangle}{\langle a\rangle + \sqrt{\langle a^2\rangle}}, \tag{3.3.6}$$

where the activity rate of each actor is considered, thus accounting for the actual dynamics of interactions. Moreover, the result does not depend in any way on the time-aggregated network and sets the epidemic threshold as a function of the interaction rate of the nodes. The outbreak condition can then be measured and predicted on the characteristic time scale of both the network and spreading process dynamics [209].

Other dynamical and searching processes that have been investigated on activity driven networks are the random walk, the mean first passage time and the percolation threshold. Their features have been found to strongly differ with respect to the static (or aggregated) counterparts.

In particular, letting W random walkers spread on an activity driven network we find that the stationary probability W_a for a random walker to be found on a

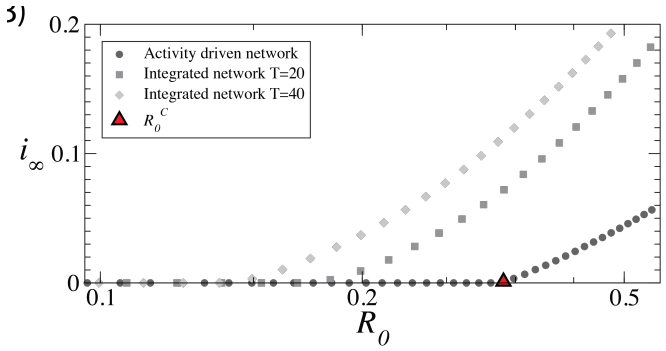


Figure 3.5: The density of infected nodes, I_∞, in the stationary state, obtained from numerical simulations of the SIS model on a network generated according to the proposed model and two other networks resulting from an integration of the model over 20 and 40 time steps, respectively. The parameters of the simulations are $N = 10^6$, $m = 5$, $F(a) \propto a^{-\gamma}$ with $\gamma = 2.1$ and $\epsilon \leq a \leq 1$ with $\epsilon = 10^{-3}$. Results are averaged over 10^2 system representations. The critical reproductive number R_0^C as found in Eq. (3.3.6) is marked by a red triangle.

node of activity a reads [215]

$$W_a = \frac{amw + \Phi}{a + m\langle a \rangle}, \tag{3.3.7}$$

where $w = W/N$ is the walkers density and $\Phi = \int aF(a)W_a da$. This results depends on the activity of the node and quickly saturates to a constant value as a grows. This is at variance with the static case [15] where the number of walkers in each node of degree k, is a linear function of the degree: $W_k \sim k$. The difference is due to the limited connectivity that higher activity nodes have in the instantaneous snapshot of the network. Indeed, they have on average $k \sim m$ connections at each time step, and therefore they cannot collect many walkers. This is at variance with the hubs present in the integrated network.

In the same way, the Mean-First-Passage-Time T_i (MFPT) (i.e. the average time needed for a walker to arrive at node i starting from any other node in the network) is found to be [215]

$$T_i = \frac{NW}{ma_i W + \sum_j a_j W_j}. \tag{3.3.8}$$

Thus, the MFPT of a node i is inversely proportional to its activity a_i plus a constant contribution from the "mean field". This is, again, at variance with the results of quenched and annealed networks where T_i is the reciprocal of the stationary state of the random walk, $T_i = W/W_i$. In this case the difference is due to low activity nodes that can trap for a very long time walkers spreading on the network.

As for the other dynamical processes studied on the activity driven networks, we recall the optimal immunization procedure developed in reference [216], the study of the percolation threshold by means of the hidden variable formalism done in [217] and the study of the integrated network's topological properties in [218].

3.4 Social Capital allocation

As already pointed out, the activity driven model just outlined is a random model that completely misses correlations and second order structure of the human interaction pattern. Indeed, it is reasonable to expect that different forces govern the evolution and shaping of social relationships making them far from random [219–226]. In addition, we expect individuals to reinforce and re-activate more likely the edges and connections that they already explored in a previous time [227, 228]. The availability of large digital datasets prompted an intense research work aimed at the understanding of what mechanisms control the dynamics of activating or deactivating social ties. The mechanisms so far presented dealt with connectivity ranking in the social network (e.g. the degree in the preferential attachment), to homophily or assortativity. Nevertheless, they completely miss the transposing of correlation of social events in the modeling framework. The importance of a correct measure and characterization of both temporal activation patterns and correlations mechanisms on human dynamics are key elements in order to give a correct description of social networks' properties [22, 186, 229], dynamical features [63, 72, 156, 163, 163, 165, 166, 172, 173, 175, 186, 209, 230–233], and the behavior of processes unfolding in social systems [22, 165, 186, 215, 218, 229, 234–239].

Recent studies tried to measure and characterize how individuals invest their social capital, i.e. how they allocate their limited energy, time, attention and emotional closeness in their ego-net [240]. The latter is the set of all the social neighbors that an individual contacts or interacts with, and is usually the microscopical quantity under investigation and the basic element to model. The most famous results on social interactions limitation is probably the Dunbar's number [228, 241]: this is a cognitive limit on the number of stable social relationships due to the neocortical volumes [242–245]. This limit imposed by neocortical processing capacity seems to define the number of individuals with whom it is possible to maintain stable interpersonal relationships. Quite remarkably, this results been recently validated by means of Twitter data [246].

An additional complication, when using electronic communication sources, is that they typically lack information on the nature of the social relationships, i.e. the strength and the nature of a given interaction. On the other hand, data collections based on survey alone cannot scale to large sample sizes and they usually feature a bias due to the limited memory and accuracy of the people recalling detailed communication events [63].

It is now customary, for detailed and precise sociological works to combine digital, automatically retrieved datasets with periodical surveys and human-compiled reports. As an example we report the work done in Reference [63], where three main results are found:

- There is a common and robust pattern in the way people allocate their social events across the members of their ego-net. Indeed, there are a few individuals (emotionally close to the ego) that receive a large fraction of calls: these links are called *strong-ties* [219]. On the other hand, the rest of the contacted alters sums up to a small fraction of the social activity: these are the so-called

weak-ties [219];

- the single individual still retains a particular social signature corresponding to his peculiar way of communication allocation;
- the latter fingerprint remains stable in its shape even during period of intense social turnover (e.g. the authors focus on the switch from high-school to college).

Thus, the heterogeneity usually found in the degree and in the activity distribution are also found in the behavior that individuals follow in they time and social capital allocation to their alters (see Fig. 3.6 for details).

A recent work further investigated the implication of the second point of the above list [242]. In particular, a measure and characterization of different "social strategies" and different patterns of links activation-deactivation has been observed in a large datasets of mobile phone calls. A method to detect links activation-deactivation in finite-size datasets is developed (see Fig. 3.6 for details) and two main social strategies (or social network explorations) are found: social keepers and social explorer. While the former class of individuals tends to interact with a small and fixed set of alters (even for high activity values), the latter is more likely to explore and connect to new nodes in the network at every interaction. Moreover, a detailed characterization of correlations between an edge activation/deactivation and a subsequent deactivation/activation is performed, together with a temporal characterization of the inter-event pattern (see reference [242] for details and Section 3.5 for more information).

All of these findings suggest that there should be an underlying mechanism shaping and driving the social exploration that could be common (or at least very similar) in diverse social layers. Other characterizations and possible explanation of these mechanism have been proposed and we refer to the literature for a deeper insight [41, 62, 138, 199, 204, 225, 226, 247–251].

Despite these first approaches, it is still lacking a general dynamical system framework able to relate the emerging connectivity pattern of social networks to the combined action of social actors activity and their heterogeneity in distributing resources in social capital allocation. To fill this gap, in the next section we will introduce our chosen mechanism to account for correlations on the individuals' activity.

3.4.1 The $p(k)$ reinforcement rule

Among all the possible way to measure and define correlations and to implement a inhibition of new edges creation, we rely on the reinforcement process recently proposed by Karsai *et. al.* [229].

In this work, the activity driven framework is expanded so as to account for the heterogeneous nature of individuals' social interactions. This is achieved by means of a simple memory effects encoded in a non-Markovian reinforcement process. The introduction of this mechanism generates two fundamental results: (i) it inhibits the

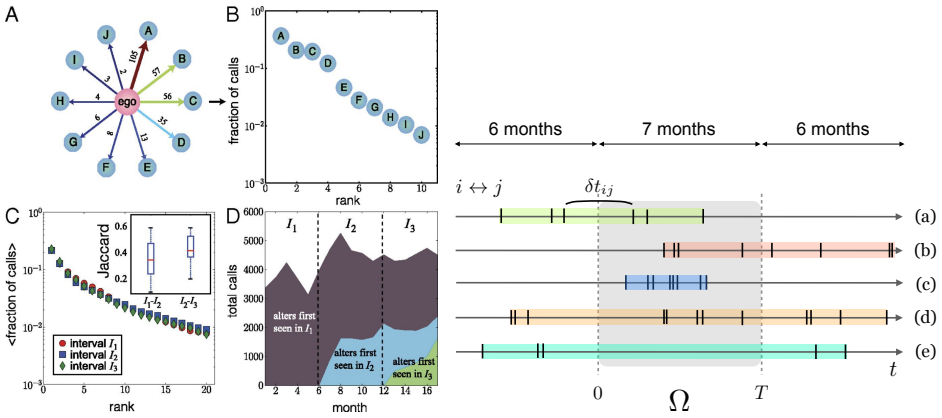
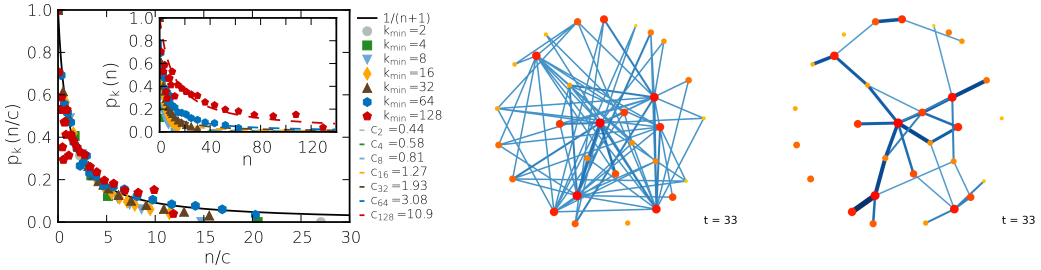


Figure 3.6: (Left panel) (A): The ego-net of a given individual and its alters clockwise ordered by link weight. (B) The rank of such weights for the ego-net of plot (A). As one can see, a large fraction of a node's activity points toward few nodes of its alters (the y-scale is logarithmic). (C) This allocation of time and attention is preserved in different periods within the observation windows: three curves are plotted, showing the rank plot of the links weight. The same pattern is observed in all the three cases. (D) The network turnover shown by means of the total numbers of calls by the participants to their alters, divided between alters that have for the first time appeared in their networks in each of the intervals. Image from [63].
(Right panel) The typical pattern of links activation/deactivation. Each line refers to a different tie leaving the ego node. Each social event toward the ego's alter is marked by a vertical segment. δt_{ij} is the inter-event time for link (i, j). An edge is activated when it is observed for the first time and it gets deactivated when the last contact between two nodes is recorded. Image from [242].



(a) (b) (c)

Figure 3.7: (a) The $p_b(k)$ function for different degree classes b as measured from empirical data rescaled by sending $k \to k/c_b$. In the inset we show the original curves. (b) A network with distributed activity as resulting from a memory-less evolution after 33 time steps. (c) The same network after the same number of steps with the reinforcement process turned on and set to $p(k) = 1/(1+k)$. In both the plots the color is proportional to the node activity (the redder the more active) and the edge width is proportional to the link weight w_{ij}.

creation of new edges by a node that gets active and (ii) it generates heterogeneity on the edges' weight. Indeed, the edges appearing first are the ones that are more likely to get activated more times in the following network's evolution.

The reinforcement process is defined through the probability $p(k)$, i.e. the probability that the next communication event of a node that already contacted k different alters in the network results in the establishment of a new, $k+1$-th link toward a node never contacted before. This probability then sets the rate, for each value of the cumulative degree k, of the increment of degree from $k \to k+1$ for an active node.

The activity driven model is then redefined as follows: once active, the node i of degree k_i will call a node j already belonging to its ego-net $\mathcal{N}(i)$ with probability $1 - p(k_i)$. The node j is chosen in a uniform way among the k_i neighbors of i. Otherwise, with probability $p(k)$ node i will call a new, randomly chosen and never contacted before node j not belonging to the neighborhood of i. Note that a node i that is not active at a given time t has still the chance to get contacted by another node j. This can be achieved in two complementary ways: either from a node j that is already in the ego-net of i and, once active, calls i with probability $(1-p(k_j))/k_j$ (where the $1/k_j$ accounts for the probability for j to exactly choose i among its k_j neighbors), or by a node l that never got in contact with i that, once active, calls i with probability $p(k_l)/(N-1-k_l)$. Here the $1/(N-1-k_l)$ term accounts for the probability for l to exactly choose i out of the nodes outside the ego-net of l.

The measure of the $p(k)$ has initially been carried out on the Mobile-Phone-Call datasets (MPC), a huge datasets comprehending more than 6 million users of an European mobile company (see Section 4.3.2 for details). As strong heterogeneities are found in the real-world network, nodes have been grouped by their cumulative degree k so as to average the reinforcement process on nodes sharing a common topological property. Then, for each class b, we measure the probability that the

next event of a node of cumulative degree k results in the connection toward a new node, thus incrementing the degree $k \to k+1$. The results are shown in Fig. 3.7. The $p_b(k)$ function of each nodes class b can be approximated by the expression

$$p_b(k) = \frac{c_b}{k + c_b} = \frac{1}{1 + \frac{k}{c_b}}, \tag{3.4.1}$$

that is well defined as $p_b(k) < 1$ for $k > 0$ and $p_b(k = 0) = 1$. The value of c_b sets the decay of the $p_b(k)$ function, as it weights the cumulative degree k. The larger c_b the slower the $p_b(k)$ will drop to zero. As a last remark, the curves of different classes con be rescaled on a single reference curve by sending $k \to k/c_b$ as we show in Fig. 3.7.

The introduction of this reinforcement mechanism deeply affects the structure and topology of both the instantaneous and the aggregated network. As we can see from Fig. 3.7 (b-c), the creation of new edges is inhibited with respect to the memory-less version of the activity driven model, so that the resulting integrated network is sparser. Moreover, we observe the appearance of strong and weak ties, i.e. we see an heterogeneity on the links' weight distribution. Indeed, links that are appearing first in time are more likely to be re-activated later, as the vertices of the edge will be pushed by the reinforcement rule $p(k)$ to re-activate already established ties instead of new ones.

As one can imagine this modification of the network evolution will strongly affect the dynamical processes as we will show in the next section.

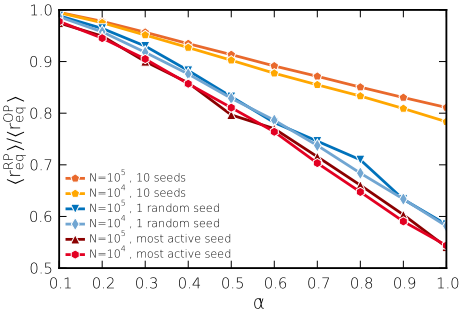
Moreover, in Section 4.1.2 we will present a more robust and thorough characterization of the reinforcement process and develop an extended version of the link creation mechanism encoded in Eq. (3.4.1).

3.4.2 Dynamical processes

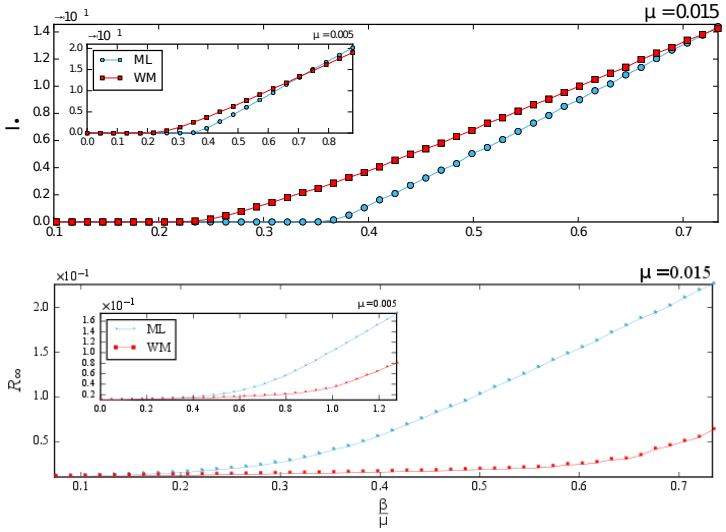
As a first insight, let us compare the outcomes of a rumor spreading process in a memory-less activity driven network and in the corresponding version featuring the reinforcement process $p(k)$ [229]. Let us recall that in a rumor spreading process people are either ignorant (compartment I), spreader (S) or stifler (R). The transition scheme is $I + S \to 2S$ with probability λ and $S + R \to 2R$ or $S + S \to 2R$ with probability λ.

As one can see in Fig. 3.8 (a), the outcome of the same process run on the two different underlying evolving networks is remarkably different. In particular, the density of the stifler r_{eq}^{RP} is smaller for every value of α, meaning that the rumor diffusion is slowed down by the correlated links activation dynamics. In particular, the repetition of contacts along the same edges has two consequences: (i) to slow down the diffusion of a rumor that is localized in a restrict number of individuals interacting between them and (ii) once a node gets stifler, all the nodes interacting with him will likely become stifler as well, since the repeated interactions may lead more quickly to a $R + S \to 2R$ transition.

The same reasoning and results can be found when running epidemics processes on top of these networks as found in references [76,176]. As we show in Fig. 3.8 (b),



(a)



(b)

Figure 3.8: (a) The ratio $\langle r_{eq}^{RP} \rangle / \langle r_{eq}^{ML} \rangle$ between the average number of stifler in an activity driven network with a reinforcement process (RP) and a memory less one (ML). The ratio is evaluated in the stationary limit of the process. (b) (Top) The I^{∞} density of infectious individuals in a ML model (blue curve) and in a RP model (red curve) as a function of β/μ. (Bottom) The R^{∞} density for ML (blue curve) and RP evolution (red curve) as a function of β/μ.

the reinforcement process has an opposite influence on the SIS and SIR epidemics processes. Indeed, while the epidemic threshold lowers with respect to the memory-less (ML) case for what concerns the SIS process, in the SIR model we observe the opposite behavior.

While the latter result can be explained in the same terms of the rumor spreading process, the former can be explained as follows: the repeated interactions between a set of individuals bound by strong ties allow for the maintenance of local hotbeds for the infectious disease. Indeed, once an alter of a tightly connected group gets infected, all the other nodes can infect as well. Then, as this group of nodes will remain infected, they will slowly spread the infection around the network through new connections established in the later part of the network's evolution. This in-

crements both the attack rate and the outbreak possibility of the epidemic process, resulting in a lower threshold value of the β/μ ratio.

This result is quite important as it highlights the weakness of strong ties: not only the repeated connections slow down the rumor diffusion on a network, but they also allow for SIS type epidemics processes to spread more efficiently on the net [76, 229, 252]. This his a counterpart of the historical result made by Granovetter, i.e. the "strength of weak ties" [219]. The latter was focusing on the role played by weak, occasionally-activated ties that are acting as bridges and shortcuts between closed communities of the network thus enhancing its spreading and diffusion performances.

As we will show in the next section, another mechanism profoundly affecting the real-world networks' dynamics is the burstiness of the human interactions.

3.5 Burstiness and edges activation time scales

Another feature that in Chapter 5 we will encode in our model is the burstiness of human activity.

So far, all the modeling tools presented feature a inter-event time distribution either peaked around one single value or following a Poisson distribution. Indeed, all the generative models of Chapter 2 feature an evolution where each new edges are added or activated precisely every δt time, i.e. for every evolution step being δt the measure of a time interval. On the other hand, the activity driven model just introduced features a Poisson distribution of inter event time: the average inter-event τ_i value for a node of activity a_i is $\tau_i \propto 1/a_i$. Moreover, the probability distribution $P(\delta t)$ of the inter-event times δt exponentially decays, thus showing finite fluctuations [155–157, 159, 209, 229].

Poisson processes have been (and still are) widely used and adopted to quantify the dynamics of human activity and modelling traffic, congestions and incidents in technological networks, mainly because of their easy and natural analytical tractability. However, the access to large scale and time-resolved digital datasets revealed that the timing of many human activities feature a skewed distribution of the inter-event times. These are usually well approximated by heavy tailed distributions (power-law, Log-Normal, Pareto distribution etc..).

The differences introduced by this heterogeneous temporal activation pattern are outstanding: while in a Poisson process the consecutive events are uniformly spaced in time so that very long waiting times intervals are forbidden, in the bursty (heavy-tailed) process one observes localized bursts (peaks) of very intense activity followed by long periods of inactivity.

The roots of this highly heterogeneous behavior have been addressed to the decision-based queuing process that humans apply when tackling tasks and allocating their time and energy in their everyday activities [159]. In other words, whenever an individual is presented with multiple tasks and chooses which one to execute first by some perceived priority parameter, the waiting time of the various tasks will follow a Pareto distribution [157, 160]. This is in contrast with the first-

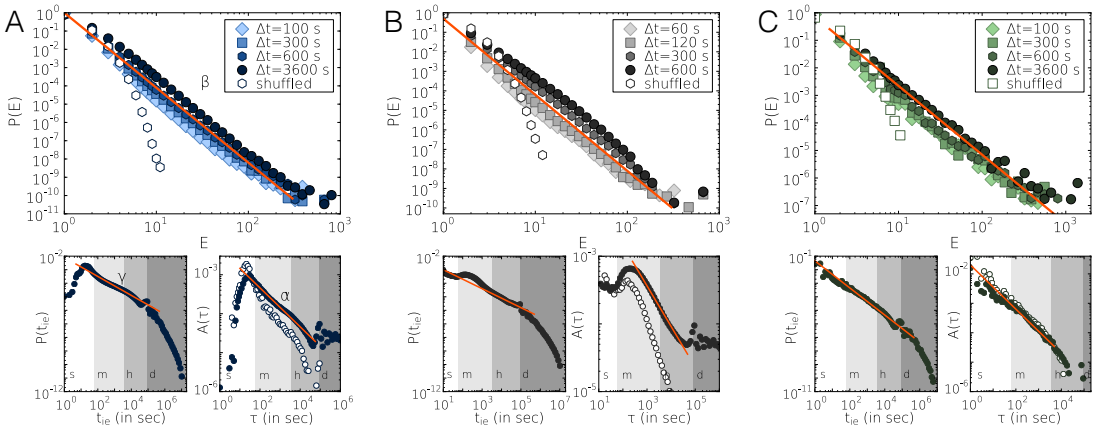


Figure 3.9: (Top row) The probability distribution $P(E)$ to observe E events in a burst of maximum duration of Δt (see legends). Data are compared to the randomly reshuffled temporal sequence (white symbols). The data refer to (A) mobile phone calls, (B) chat messages and (C) email sequence. (Bottom row) (Left panels) The probability distribution $P(t_{ie})$ of observing an inter-event time $\delta ti, i+1 = t_{ie}$. (Right panels) The autocorrelation functions $A(\tau)$. Figure from [160].

come-first-serve and random task execution that leads to a Poisson-like dynamics instead.

3.5.1 Circadian rhythms and human activity patterns

As just pointed out, heavy tailed inter-event time distributions are the typical signature of a bursty nature of social contacts and human dynamics.

Besides the queuing process, another mechanisms has been suggested as a cause of this behavior: inhomogeneities due to the circadian and weekly activity patterns of individuals. It is reasonable to expect that the circadian and other long-time cycles of our lives (seasons, school cycle, etc.) may cause inhomogeneity of natural and societal factors. Malmgrem [205, 253] proposed a possible explanation by means of a combinations of two Poisson processes featuring different time scales. In this way, the long inter-event inactivity are attributed to night-time and weekend inactivity.

However, a more recent study revealed that the bursty behavior remains even after a de-seasoning procedure. This procedure completely removes the circadian and weekly patterns from the time series of mobile phone communication events of individuals [160]. However, heavy tails in the inter-event time distributions are still found after the procedure is applied, thus promoting the human task execution-based mechanism as the possible cause of the burstiness in social and communication layers [155].

As an additional complications, there is evidence that the bursty behavior is caused and/or goes together with both memory processes and activity correlations in human dynamics [155, 159, 160, 254].

Regarding the correlations, they cannot be catch by only looking at the inter-event time distribution or the autocorrelation function. Moreover, the lengths of the bursts themselves follow a scale-free probability distribution (see Fig. 3.9 for details). These bursty trains can be explained in terms of memory effects. Specifically, also humans (as earthquakes and firing neurons) seem to feature a threshold mechanism governing their activation in the social layer [160].

As for the memory part, it has been recently shown that different kind of long-term memory effects are found in the links creation and deactivation [254]. Moreover, by combining these mechanisms it is possible to naturally generate the heterogeneous network dynamics found in real-world systems.

3.5.2 Fingerprints of bursty behavior

As already said, the most easily found and measured fingerprint of a bursty behavior is a scale free, or at least heavy tailed, inter-event time t_{ie} distribution $P(t_{ie})$. As we show in the bottom plots of Fig. 3.9 this feature is found in many real-world systems, from email exchange, to mobile phone calls passing through online chats and messaging platforms.

To replicate such feature by imposing a broad inter event time distribution is the first and most naive way to implement burstiness in a time-varying network model. This is how we will implement burstiness in our model as we will show in Chapter 5.

Of course, there are more mechanisms interplaying and determining the temporal activation pattern of human dynamics. Among them, we recall the short time correlations and long-term memory mechanisms that interact to shape the edges' activation. In particular, we recall that even the bursts lengths and sizes (measured as the number of events E belonging to a single burst of activity) follows a scale-free distribution. This is shown in the top panels of Fig. 3.9 for different upper cut-offs of the single burst length Δt (i.e. by measuring $P(E)$ in bursts whose temporal length is at most Δt).

Other works aiming at the modeling and characterization of burstiness features have been proposed in the last years [163, 172–174, 254].

Specifically, the results found in reference [254] associate the emergence of several aspects of the temporal heterogeneities of real-world complex networks as the outcome of long-term memory mechanisms governing the activation and deactivation of edges. In particular, four main mechanisms are proposed:

- *contact self-reinforcement* (CSR): the more a given edge has been active, the less likely for it is to be terminated. This is encoded in a $f_l(\tau)$ function that may depend on the activation's length $\tau = t - t_{(i,j)}$, i.e. the time passed from the (i, j) link activation at time $t_{(i,j)}$. By now, it should come as no surprise that empirical measures of this quantity reveal a scale free form of $f_l(\tau) \sim \tau^{-\gamma}$.
- *activity self-reinforcement* (ASR): the same mechanism of the above point is applied to a node's activation pattern. The more recently a node has been

active, the more likely it is for it to engage a new social event. This is encoded in the $f_a(\tau) \sim \tau^{-\gamma}$ function, measuring the probability for a node whose last activation was at time $t = t_{a_i}$ to get active again at time $t = t_{a_i} + \tau$.

- *agent-centric preferential attachment* (APA) This mechanism makes nodes that have been recently active a good fit for node getting active and looking for an alter to interact with. Indeed, the probability for a node to get chosen by another node establishing a connection is proportional to a function $\Pi_a(\tau) \sim \tau^{-\gamma}$.

- *link-centric preferential attachment* (LPA) This is the last mechanism proposed in this modeling framework and it accounts for the reinforcement of already established ties. Indeed, the probability for an active node to choose a given alter is found to decay with the time passed by the last interactions between the two. In other words, an agent initiating a contact will chose with probability $\Pi_l(\tau) \sim \tau^{-\gamma}$ an agent it has recently been in contact with, being τ the time passed since their last interaction.

After an empirical measurement and a fine tuning of these mechanisms' parameters, a remarkable agreement between synthetic and empirical data is found, both for the inter-event and dynamical processes outcome.

In particular broad distributions of contact and inter-event times and of the number of contacts per link are reproduced, as well as for what concerns the agents' activities.

The further possibility to switch one or more mechanisms also allows the determination of the impact of each type of heterogeneity on dynamical processes on temporal networks.

Amongst all, the results highlight the crucial roles of the inter-event time distribution $p(\Delta\tau_i)$ and the edges weight distribution $F(w_{i,j})$ [174, 254]. It is this result that prompted us to model burstiness, as a first level of approximation, by means of a heavy tailed inter-event distribution. As we already shown, the heterogeneous distribution of weight is naturally introduced by the reinforcement process introduced in Section 3.4.1.

The same approach has also been recently adopted in reference [154]. In the latter, particular attention has been devoted to the study and characterization of the aging behavior, that is known to be found in systems whose inter-event time distribution between the addition of connections is power-law distributed [255, 256].

In particular, this effect is evident when aggregating the network's connections over an arbitrary time-window $[t_i, t_f = t_i + t]$, thus setting the starting integration time $t_i \neq t_0$, where t_0 is the starting time of observation of the system. Though a preliminary measure on the effects of a varying time-window on the integrated network has been already carried out [236, 257], one is here interested in the breaking of the time-translation invariance of the network's topological properties. Indeed, the degree distribution $\rho(k)$ depends both on the integration time window's length $t = t_f - t_i$ and the aging time t_i at which the aggregation of connections starts.

Specifically, the work studies the integrated properties of a Non-Poissonian-Activity-Driven model (NoPAD model). Burstiness is introduced by means of a renewal process [258], in particular with a time-dependent activation rate $a_i(t)$ for node i. The inter-event time distribution $\Psi_i(w)$ for node i is given by

$$\Psi_i(w) = a_i(w) \exp\left[-\int_0^w a_i'(w')dw'\right]. \tag{3.5.1}$$

If $a_i(t)$ is set to be constant we fall back to the Poisson process of the regular activity driven model. The inter-vent time distribution can be further generalized so as to account for heterogeneous activation pattern (e.g. average inter-event time), thus including an additional parameter ξ, distributed among the nodes accordingly to a given distribution $\eta(\xi)$. The waiting-time distribution then reads $\Psi(w, \xi_i)$.

The evaluation of the degree distribution $\rho(k)$ is made in terms of the hidden variable and generative functions formalisms [67,100,218,259] , where the key point is to characterize the kernel probability $\Pi_t(i,j)$ that two nodes i and j are linked within time t. The latter lets then define the degree distribution

$$\rho(k) = \sum_{\vec{h}} F(\vec{h}) g(k|\vec{h}), \tag{3.5.2}$$

where $\vec{h}$ is the hidden variable, $F(\vec{h})$ its distribution and $g(k|\vec{h})$ is the conditional probability that a vertex featuring hidden variable $\vec{h}$ has degree k. The generating function for $g(k|\vec{h})$ reads

$$\ln \hat{g}(z|\vec{h}) = N \sum_{\vec{h}'} F(\vec{h}') \ln\Big[1 - (1-z)\Pi(\vec{h}, \vec{h}')\Big]. \tag{3.5.3}$$

The $\Pi_t(i,j)$ probability can be written [218] as $\Pi_t(i,j) = 1 - [1-(1/N)]^{r_i}[1-(1/N)]^{r_j}$, where r_i counts the number of activations of node i. The model can then be solved by noting that the hidden variables are the vector $\vec{h} = (r, \xi)$, following the probability distribution $F_t(r,\xi) = \eta(\xi)\chi_t(r|\xi)$, being $\chi_t(r|\xi)$ the distribution of number of activation of a node with time t and heterogeneity ξ. The degree distribution at time t, starting the integration at the initial time $t_i = t_0 = 0$, then reads

$$\rho_t(k) \simeq \sum_{\xi} \eta(\xi)\chi_t(k - \langle r\rangle_t|\xi), \tag{3.5.4}$$

where $\langle r\rangle_t = \sum_\xi \eta(\xi) \sum_r r\chi_t(r|\xi)$. A power-law distributed inter-event time can then be obtained by fixing $\chi_t(r|\xi) \sim 1/(\xi t)^a \cdot e^{\omega(a,\xi t) r^{1/(1-a)}}$, where $\omega(\alpha, u) = -[1-\alpha][(\alpha/u)^\alpha \Gamma(1-\alpha)]^{1/(1-\alpha)}$. This leads to an inter-event time distribution $\Psi(w,\xi) = \alpha\xi(\xi t+1)^{-(\alpha+1)}$ (being the exponent $0 < \alpha < 1$). If one also sets a broad distribution $\eta(\xi) \propto (\xi/\xi_0)^{-(b+1)}$ of the heterogeneity parameter ξ (with $b > \alpha$), in the $k \ll (\xi_0 t)^\alpha$ we get

$$\rho_t(k) \sim (\xi_0 t)^b (k - \langle r\rangle_t)^{-1-b/\alpha}, \tag{3.5.5}$$

then fixing the relation between the exponent μ driving the tail of the degree distribution $\rho(k) \sim k^{-\mu}$ and the exponent of the inter-event distribution $\Psi(w) \sim w^{-(1+\alpha)}$, i.e.:

$$\mu = 1 + b/\alpha. \tag{3.5.6}$$

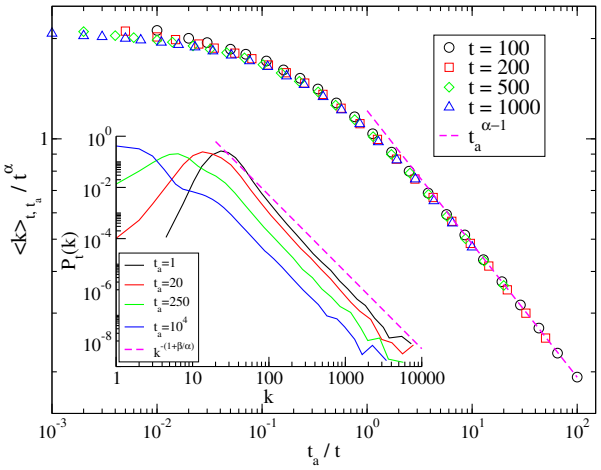


Figure 3.10: The rescaled average degree $\langle k\rangle_{t_i,t}/t^\alpha$ (symbols)as a function of the aging time t_i (data refers to numerical simulations featuring $N = 10^6$ nodes, $\alpha = 0.6$, $\xi_0 = 1$, and $b = 1.2$). The analytical prediction of Eq. (3.5.7) is shown (red dashed line). In the inset it is shown the $\rho_t(k)$ degree distribution for different aging time t_i and an integration time of $t = 100$ (solid lines). The asymptotic behavior predicted by Eq. (3.5.6). Figure from reference [154].

Given that the distribution of renewal events $\chi(r|\xi)$ defines the topological properties of the systems, in the presence of a power-law distributed inter-event time (with an exponent $\alpha < 1$) the NoPAD model will be affected by aging effects.

This can be easily shown by looking at the degree distribution $\rho_{t_i,t}(k)$ found by integrating the network's contact between time t_i and $t_i + t$. As shown in Fig. 3.10, the tail of the distribution does not change as one increases t_i. However, the peak of the distribution moves to smaller values of degree k. We can easily understand this behavior in terms of the average number of activations $\langle r\rangle_t$ in the network up to time t for $\alpha < 1$:

$$\langle r\rangle_t \simeq (\xi_0 t)^\alpha. \tag{3.5.7}$$

As the average degree $\langle k\rangle_{t,t_a}$ of the network equals the number of activation events multiplied by two, we find it to depend on the integration time as

$$\langle k\rangle_{t,t_i} \sim [(t_i + t)^\alpha - t_i^\alpha]\,. \tag{3.5.8}$$

Thus, if t_i is comparable with t we find a non-trivial dependence of the average degree on t_i and α [154]. This phenomenon is due to a sort of "thermalization" of the system when we let it age for a long aging time t_i before starting to aggregate its connections. Indeed, when the inter-event time distribution features a diverging average value (i.e. when $\alpha < 1$), the larger t_i the more likely it is to find many nodes to be in a large gaps of activity within the observation window of length t (this is due to the fact that all the nodes are synchronized at time $t = 0$).

The same aging effect has also been measure in empirical temporal networks [154]. As already noted, empirical evidence of such aging and heavy-tailed inter-event time

distributions may be mingled with other mechanisms (population fluctuations, correlations on the activity, clustering etc.), so that a more comprehensive analysis of the different contribution to the aging process of temporal networks is needed.

Chapter 4

Asymptotic theory of Poissonian TVN

In this chapter, as the model is data-driven, in Section 4.1 we first introduce the seven datasets used to derive the model and the fundamental, preliminary measures we carried over on them. In particular, we conduct a rigorous analysis of the activity distribution $F(a)$ and introduce a precise measure of the edges' reinforcement mechanism $p(k)$.

We then present in Section 4.2 the developed model and the adopted analytical framework, together with the corresponding analytical predictions. Then, in Section 4.3 we present the numerical check of our analytical results and we later test of our model against real-world datasets.

Finally, in Section 4.4 we will present a preliminary characterization of an epidemic process on top of a modular activity-driven network. In the latter, we apply the memory-less version of the activity driven model and we insert correlations on edges activation by means of an enforced community structure.

4.1 First insight on the data

4.1.1 Datasets

We analyze seven datasets containing time-stamped information about three different type of social interactions: scientific collaborations, Twitter mentions, and mobile phone calls. All of these datasets are represented as time-varying networks. Each node describes an individual and each time-resolved event within the temporal sequence describes a social act linking two (or more in the case of scientific collaborations) alters. The nature of connections is different according to the specific dataset. Links might represent a collaboration resulting in a publication in a scientific journal, a Twitter mention, or a mobile phone call. We considered five scientific collaborations networks obtained from five different journals (PRA, PRB, PRD, PRE, and PRL) of the American Physical Society (APS), one Twitter mentions network (TMN), and one mobile phone network (MPN).

American Physical Society

The *APS* dataset contains the five co-authorship networks of five journals of the *A*merican *P*hysical *S*ociety, i.e., *Physical Review A*, *B*, *D*, *E* and Letters (*L*).

The various datasets contains the data referring to all the issues of the single journals from their first issue up to a certain edition, specifically:

- *PRA* from January 1970 to December 2006;

- *PRB* and *PRD* from January 1970 to December 2007;

- *PRE* from January 1993 to December 2006;

- *PRL* from February 1960 to December 2006.

Each dataset is composed by several files (one per month). Each file has as many lines as the number of papers published in that month. Finally, each line contains the IDs of the authors of the specific paper. For instance, the typical head of a file is:

```
Author_000    Author_001    Author_002  #First Paper with 3 authors
Author_003    Author_004                #Second Paper with 2 authors
 . . .          . . .          . . .          . . .
```

The data are cleaned so as to not take into account the papers with a single author.

When analyzing this dataset we define the user's activity a_i as the number of engaged collaborations (e.g. an author i that publish two papers, the first with 3 co-authors and the second with a single co-author, has activity $a_i = 4$).

Twitter Mention Network

The dataset of *Twitter* is composed by 273 daily files covering the period between January the 1st to September the 30th2008. The dataset contains the so called *fire-hose*, i.e., all the $16,329,466$ citations done by all the $536,210$ users in the given period. The nodes in the network are connected via $2,620,764$ edges.

Each file contains the daily events with the structure:

```
Citer_ID_00    Cited_ID_00    # Event 0
Citer_ID_01    Cited_ID_01    # Event 1
Citer_ID_02    Cited_ID_02    # Event 2
   . . .          . . .          . . .
```

This dataset is not cleaned, as we have all the events that happened on the platform in the selected period.

When analyzing this dataset we define the user's activity a_i as the number of citation made by i, i.e. the number of events actually engaged by the node i.

Mobile Phone Network

The dataset of the *M*obile *P*hone *N*etwork (*MPN*) is composed by a single file containing the $1,949,624,446$ time ordered events with 1 second resolution covering the period between January and July of 2008 for $6,779,063$ users of a single operator with 20% market share in an undisclosed European country.

The dataset contains all the events from and toward users of the company (so that even the calls from non-company users to company users and vice-versa are taken into account). As a result, we have $33,160,589$ nodes (of which $6,779,063$ are users of the selected company) that are connected via $92,784,825$ edges.

We split the huge list of events in 98 files (each of them containing more or less the same number of events) for computing convenience. Each file contains events with the structure:

```
Caller_ID    Called_ID    Company_Caller   Company_Called   # Event 0
Caller_ID    Called_ID    Company_Caller   Company_Called   # Event 1
Caller_ID    Called_ID    Company_Caller   Company_Called   # Event 2
  . . .        . . .          . . .            . . .          . . .
```

where `Company_Caller` and `Company_Called` are the value of the provider company of the called and caller nodes, respectively (e.g. the value is set to 1 if the node is a customer of our company, 0 otherwise).

When analyzing this dataset we define the user's activity a_i as the number of calls done by the node, i.e. the number of calls actually engaged by the node i.

4.1.2 Activity Distribution

In order to characterize the time-varying and reinforcement properties of such networks we first measure the activity a_i. Formally, a_i is defined as the fraction of interactions in which node i is engaged per unit of time with respect to all the interactions per unit time occurring in the network. This quantity describes the propensity of nodes i to be involved in social interactions. Empirical measurements in a wide set of social networks show broad distributions of activity [209, 215, 229, 236, 251].

For the datasets presented in Section 4.3.2 we first evaluate, for each node i, the total number u_i of events engaged by the node itself. For instance u_i is the number of calls made by the node i in the *MPC* dataset or the number of citations done by i in the *Twitter* dataset.

We then define the node activity a_i as the ratio between the i-th node's number of events and the total number of events observed in the dataset, i.e. $a_i = u_i/u_{\text{tot}}$ where $u_{\text{tot}} = \sum_j u_j$. Thus, a_i falls in the range $a_i \in [\epsilon, 1.0)$ with $\epsilon = \min_i(u_i)/u_{\text{tot}}$. We then introduce and compute the activity distribution $F(a)$, i.e. the probability for the activity to have value a. In Fig. 4.2 we show the resulting activity distribution for each analyzed dataset, while in Table (4.1) we show the best candidate functional form for the $F(a)$ distribution of each dataset. The latter is estimated using the methods found in [94].

In particular, we compare the goodness of fit on the $F(a)$ distribution of the functional forms found in Table 4.1, i.e. power-law, truncated power-law, stretched exponential and log-normal distribution. The procedure for each dataset and each functional form reads as follows:

PDF	$F(a)$
Power Law	$a^{-\nu}$
Stret. Exp.	$a^{\nu-1}\exp\left[-\lambda a^{\nu}\right]$
Trunc. PL	$a^{-\nu}\exp\left[-\lambda a\right]$
Log-Normal	$\frac{1}{a}\exp\left[-\frac{(\ln(a)-\mu)^2}{2\sigma_a^2}\right]$

Figure 4.1: The functional form of the activity PDF $F(a)$ took into account. The selected PDF are, from top to bottom: power law, stretched exponential (Stret. Exp.), power law with cutoff (Trunc. PL) and the Log-Normal distribution.

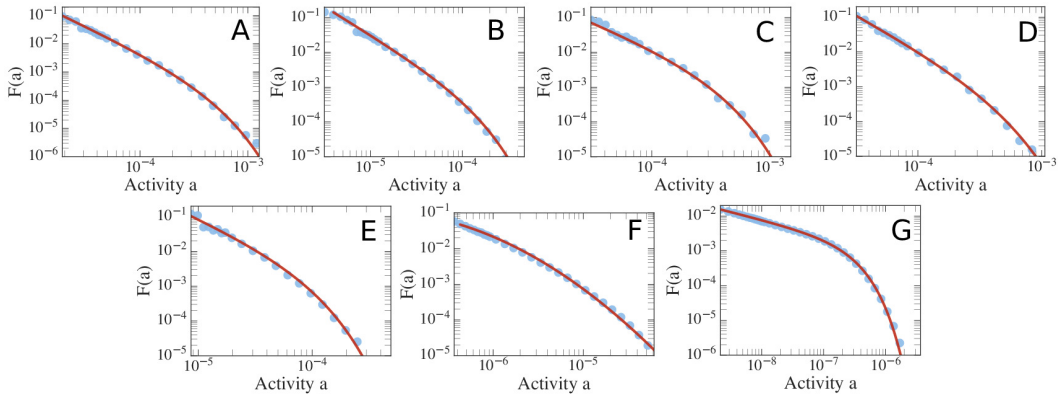


Figure 4.2: The experimental activity distribution $F(a)$ for (A) PRA, (B) PRB, (C) PRD, (D) PRE, (E) PRL, (F) TMN and (G) MPN (blue points). We also show the best candidate fit of the $F(a)$ distribution (red solid lines) featuring the functional form and parameters found in Table (4.1). In all the plots we show the data and fit ranging from the lower bound x_{min} to the 99.9% of the measured data, thus excluding from the visible area the top 0.1% of the activity values (see Table (4.1) for the lower bound details).

Dataset	Dist.	Parameters	KS_d	%	$\mathcal{L}$
TMN	LogN	$a_{\min} = 4.28e-7$, $\mu = -14.02$, $\sigma = 1.71$	$1.5e-2$	52	$-2.26e+3$
PRA	Trunc	$a_{\min} = 1.90e-5$, $\lambda = 3.14e+3$, $\alpha = 1.789$	$1.5e-2$	32	-301
PRB	Trunc	$a_{\min} = 6.31e-6$, $\lambda = 7.96e+3$, $\alpha = 1.638$	$1.4e-2$	41	-744
PRD	Trunc	$a_{\min} = 4.54e-5$, $\lambda = 4.02e+3$, $\alpha = 1.37$	$1.6e-2$	27	-286
PRE	Trunc	$a_{\min} = 4.24e-5$, $\lambda = 3.32e+3$, $\alpha = 1.92$	$1.5e-2$	23	-264
PRL	Trunc	$a_{\min} = 1.10e-5$, $\lambda = 1.55e+4$, $\alpha = 1.47$	$1.7e-2$	31	-577
MPN	Trunc	$a_{\min} = 2.17e-9$, $\lambda = 3.82e+6$, $\alpha = 0.448$	$9.5e-3$	94	$-1.6e+4$

Table 4.1: The candidate functional form of the activity distribution for each analyzed dataset, the evaluated parameters (see Table (4.1) for the analytical expressions), the Kolmgorov-Smirnov distance KS_d, the percent % of nodes in the dataset that have activity $a_i \geq a_{\min}$ and the normalized log-likelihood $\mathcal{L}$. In the parameters we include $a_{\min}$ that is the value of the activity that minimizes the KS distance. This is the lower bound for the functional form behavior, i.e. the point at which data behave as the functional form.

- we fit the $F(a)$ taking into account all the nodes featuring $a_i \geq x_{\min}$, where $x_{\min}$ is the lower bound of the distribution. The fit is performed using the maximum likelihood estimators (MLE) that return the optimal values of the parameters;

- once the optimal parameters are found we compute the Kolmogorov-Smirnov distance ($KD_d(x_{\min})$) between the analytical and experimental complementary cumulative distribution function (CDF);

- we then apply this procedure for different $x_{\min}$ and set the $a_{\min} = \min_{x_{\min}} KS_d(x_{\min})$ lower bound value as the one that minimizes the KS_d.

We then repeat this procedure for all the functional forms of the $F(a)$ and we then compare them with the *likelihood ratio test* $\mathcal{R}$ combined with the p-value that gives the statistical significance of $\mathcal{R}$ [94, 260]. The result of this procedure gives us the best candidate for the $F(a)$ for each dataset as shown in Table (4.1). We find that a truncated power law is the best candidate for all the APS datasets together with the MPN one. On the other hand, in the TMN we find a log-normal distribution as the best candidate for the dataset.

Our datasets provide evidence that nodes within the same activity class (i.e. node with similar values of activity a_i) can feature very different memory behavior. In particular agents with large activity may connect to very few different nodes (strong reinforcement) or establish new links at almost every step (weak reinforcement). For this reasons each node i of the network is naturally classified according to her activity a_i and her final degree k_i, i.e. the total number of different agents that have been connected to i in the considered time window.

We then define a binning procedure that let us group together the similar nodes, i.e. nodes with similar activity and final degree. We divide the nodes in N_{act} activity classes so that within each activity class the most active node performs at most 1.5 times the events of the least active node. We also tested that, by changing this 1.5 values by both lowering and raising it the reinforcement process measure does not significantly change.

Then, with the same procedure, we further group the nodes within each activity class a according to their final degree, thus defining $N_{\text{deg}}(a)$ final degree classes. The nodes are therefore divided in $N_b = \sum_{a=1}^{N_{\text{act}}} N_{\text{deg}}(a)$ activity-degree classes. From now on, unless differently stated, whenever we mention the nodes' class or bin b we will be referring to one of these N_b classes.

Reinforcement process

The activity a_i sets the clock for the activation of each node, however it does not provide any information on how each node invests its social capital in exploring new ties or reinforcing already established ties [242].

We then measure the probability $p_b(k)$ that the next social act for the nodes in the class b that have already contacted k nodes will result in the establishment of a new, $k+1$-th, tie.

To measure the reinforcement process of each system, we count all the communication events $e_b(k)$ engaged by every node i of the b-th class when it has degree $k_i = k$. In other words, $e_b(k)$ is the total number of events engaged by the nodes of the b-th class at degree k.

Each time an event engaged by a node i of the b-th class results in a degree increase $k_i = k \to k_i = k + 1$, we increment the counter $n_b(k)$ by 1. In other words, $n_b(k)$ is the total number of events that the nodes belonging to the b-th and featuring degree k perform toward a new node. Of course, if a node i of the b-th class with degree $k_i = k$ increases its degree to $k_i = k + 1$ because it gets called by a new node, the $n_b(k)$ counter is not incremented.

The best estimate of the probability for a new node to get establish a new connection at degree k then reads:

$$f_b(k) = \frac{n_b(k)}{e_b(k)}, \tag{4.1.1}$$

where $n_b(k)$ and $e_b(k)$ are the event counters as defined above. We can give an estimate of the uncertainty on $f_b(k)$, by assuming that at a given degree k the events are independent (i.e. there are no correlations between users) and by checking that $1 \ll n_b(k) \ll e_b(k)$ so that the STD $\sigma(f_b(k))$ of $f_b(k)$ reads:

$$\sigma(f_b(k)) = \sigma_b(k) = \sqrt{\frac{f_b(k)(1 - f_b(k))}{e_b(k)}}. \tag{4.1.2}$$

As shown in Figure 4.3 $p_b(k)$ is in general a decreasing function of k. This observation resonates with previous research and empirical findings suggesting that our social interactions are bounded by cognitive and temporal constraints [228, 242–244]. Indeed, the larger the number of alters in our social circle, the smaller the probability that the next social act will be towards a new tie.

Given the behavior of the $f_b(k)$ shown in Fig. 4.3, we then fit $f_b(k)$ with the proposed reinforcement function $p_b(k, \beta)$ (see Eq. (4.1.7) below for details):

$$p_b(k, \beta) = \left(1 + \frac{k}{c(b)}\right)^{-\beta}, \tag{4.1.3}$$

where $c(b)$ is the social propensity of the b-th bin, k is the cumulative degree and β is the reinforcement strength, that will be kept fixed for all the nodes in the system. In particular, for each class b and with a fixed β, we optimize the parameter $c(b)$, by minimizing the function $\chi^2_b(\beta)$:

$$\chi^2_b(\beta) = \sum_{k=1}^{K_b} \frac{[f_b(k) - p_b(k, \beta)]^2}{\sigma_b(k)^2}, \tag{4.1.4}$$

where the index k runs over the K_b points of the b-th bin's curve and $\sigma_b(k)$ is as defined in Eq. (4.1.2). By repeating this procedure for each value of $\beta \in [0, 5.0]$ we find, for each class b, a $\chi^2_b(\beta)$ curve.

In Fig. 4.4 we show the behavior of $\chi^2_b(\beta)$. For each class b we find a minimum of $\chi^2_b(\beta)$ at a certain $\beta_{opt}(b)$ (see the horizontal lines in the heat-map-like panels of Fig. 4.4).

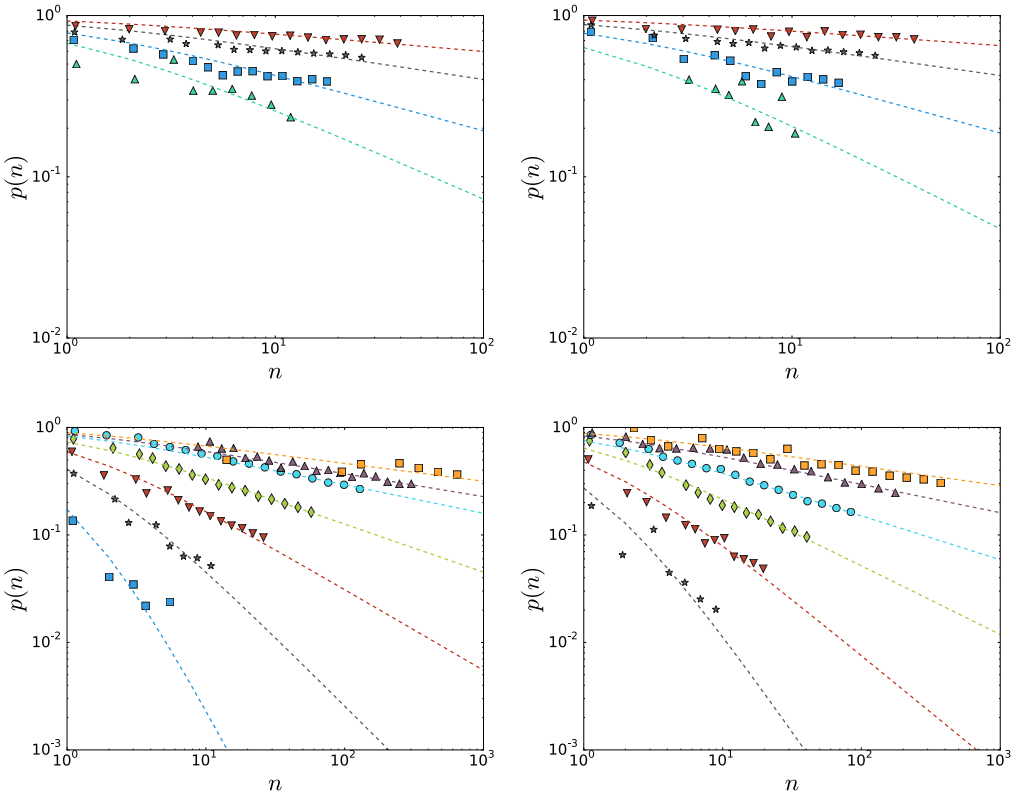
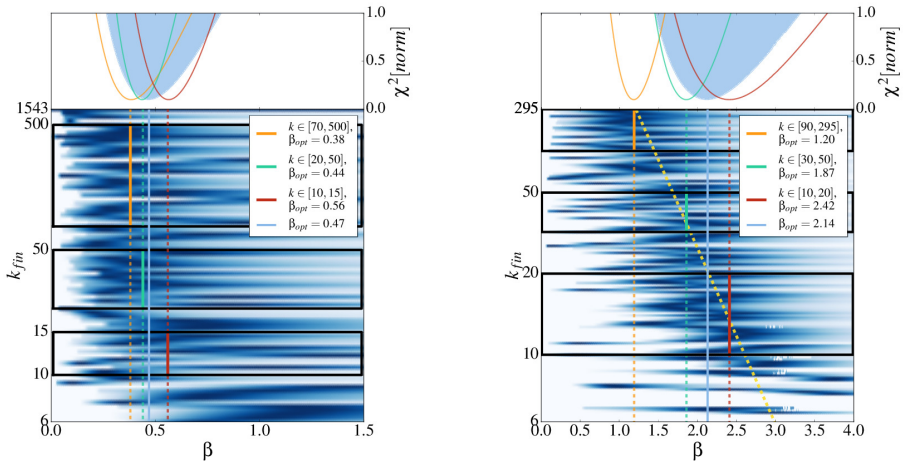


Figure 4.3: The $f_b(k)$ functioning as empirically measured (symbols). We show the results concerning (a) PRE, (b) PRL, (c) TMN, and (d) MPN. Each curve corresponds to a different node bin.



(a) (b)

Figure 4.4: The heat-map-like value of $-\ln[\chi^2_b(\beta)]$ (bottom plots). We plot the exponent β on the x-axes and the different bins b sorted by their final degree on the y-axes. The color-map is proportional to $-\ln[\chi^2_b(\beta)]$ representing the goodness of fit: the darker, the higher. The cyan vertical line is the value of β_{opt} defined in Eq. (4.1.6), while the other vertical lines represent the same quantity evaluated in the three black boxes corresponding to different final degree intervals. (Top plots) The curve $\chi^2(\beta)$ as defined in Eq. (4.1.5) (up-filled curve) and the same quantity for the three final degree intervals. For Twitter (a): a single value of $\beta_{opt} = 0.47$ fits most of the curves and only some bins b deviate from the average behavior. (b) MPC: in this case we observe different behaviors depending on the final degree. Thus, a single $\beta_{opt} = 2.14$ does not fit all the curves. We also show a "guide-to-the-eye" to highlight this feature (yellow dashed line).

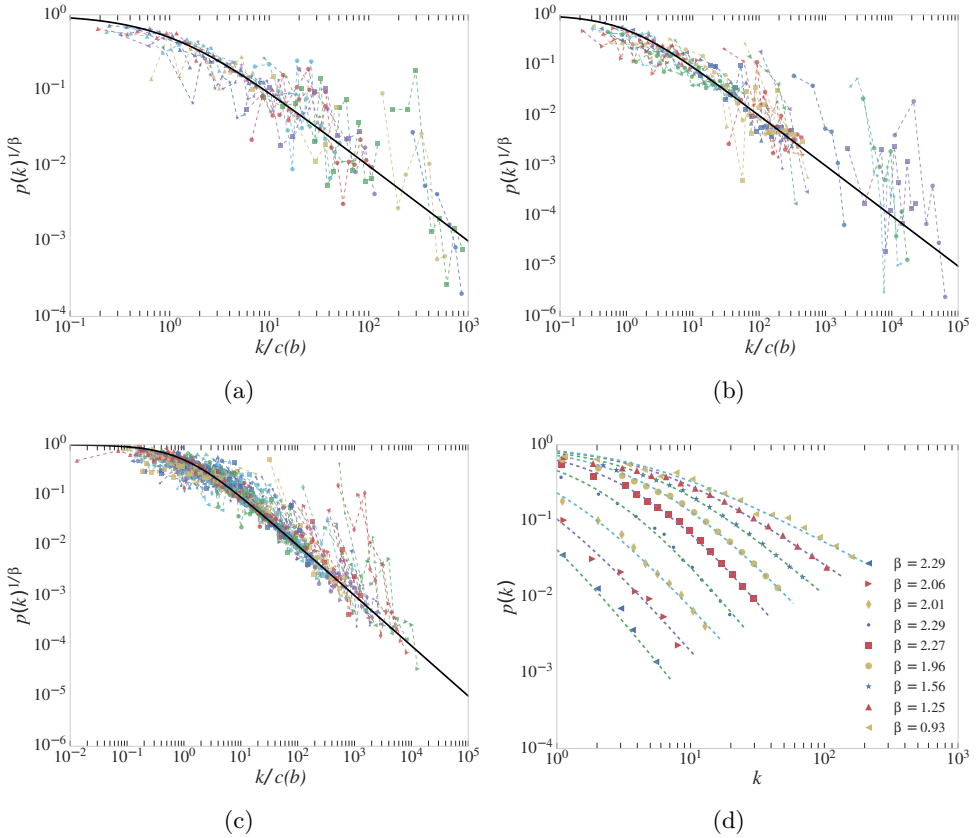


Figure 4.5: Plot of the experimental $p_b(k)$ curves for the (a) PRE, (b) PRL, (c) TMN and (d) MPC datasets. In the (a-c) cases the k is rescaled as $k \to k/c(b)$, where $c(b)$ is the constant for the b-th class curve at $\beta = \beta_{\text{opt}}$. The $p_b(k/c(b))$ points are then rescaled sending $p_b(k/c(b))^{1/\beta_{\text{opt}}}$. In the (d) panel for MPC we simply plot $p_b(k)$ as a function of k with no rescaling, given that each curve features its own β optimal value β_{opt} as shown in the legend.

Moreover, Fig. 4.4 shows that there are two different behaviors. Specifically, in the TMN case (see Fig. 4.4 (a)), one value of $\beta_{opt} = 0.47$ fits most of the curves, exception made for some outsiders: the value of $\beta_{\text{opt}}(b)$ that maximizes the $1/\chi^2_b(\beta)$ is practically the same for all the bins. On the contrary, in the MPC case the maximum of the $1/\chi^2_b(\beta)$ function follows a diagonal path ranging from a larger $\beta_{\text{opt}}(b)$ for bins with lower final degree to a smaller $\beta_{\text{opt}}(b)$ for larger degree bins. In this case a single β_{opt} cannot fit all the curves and we have to consider a multi-β model where each class b features a different optimal value of β, $\beta_{\text{opt}}(b)$.

In Fig. 4.5 we present the rescaled $p_b(k)$ curves for the PRA, PRD, PRE, PRL, TMN and MPC datasets. In the first five cases we show the rescaled curves obtained by substituting $k \to k/c_b$ and then plotting $p_b(k) \to p_b(k)^{1/\beta_{\text{opt}}}$. As one can see, the curves nicely collapse on the reference curve $(1+k)^{-1}$. In the MPC case we show instead the original curves, each one fitted with its own $\beta_{\text{opt}}(b)$. The latter parameter falls in the $1.2 \lesssim \beta_{\text{opt}}(b) \lesssim 3.0$ interval for most of the curves as we also show in Fig. 4.4.

To quantitatively define the β_{opt} parameter, let us define the total mean square deviation $\chi^2(\beta)$ as

$$\chi^2(\beta) = \sum_{b=1}^{N_b} \left[\chi_b^2(\beta)\right], \tag{4.1.5}$$

where N_b is the total number of curves, i.e. the number of activity-degree bins b. Then, for the single exponent case, the function $\chi^2(\beta)$ allows to define β_{opt} as:

$$\beta_{\text{opt}} = \min_\beta(\chi^2(\beta)). \tag{4.1.6}$$

In the multi-β case instead, we compute the different values of the exponent $\beta_{\text{opt}}(b)$ found in the system by grouping the memory classes b accordingly to their final degree as shown in Fig. 4.4. The optimal value of $\beta_{\text{opt}}(b)$ is found to be minimum for the bins featuring a large final degree, i.e. $\beta_{\text{min}} \equiv \beta_{\text{opt}} \sim 1.2$, which, as we will show in Section 4.2.3, is the exponent driving the evolution of the network.

To corroborate the results just outlined, we show in Fig. 4.6 the box plot of the $\beta_{\text{opt}}(b)$ distribution for different groups of nodes classes b grouped by their final degree. We note that the APS and TWT datasets are well approximated by a single β_{opt} as the distribution of $\beta_{\text{opt}}(b)$ within each sub-group of nodes is compatible with the global optimal value β_{opt}. On the other hand, in the MPN case we see that the large final-degree classes have their $\beta_{\text{opt}}(b)$ distribution centered around a smaller value of $\beta_{\text{min}} \sim 1.2$. As already anticipated, this value will lead the asymptotic growth of the system as we will show in Section 4.2.3.

As a last remark we present in Fig. 4.7 (a, b) the measured distribution of the constant $c(b)$ for the MPN and TMN datasets. We show the distribution for all the nodes in the network and for each activity class a, i.e. the group of nodes featuring similar activity. The values of this constants are distributed but peaked around an average value. Moreover, the distribution of the $c(b)$ parameter within each activity closely follows the global one. The distribution of the social attitude $c(b)$ then appears to be a global, activity independent feature of the nodes in the system. Finally, in Fig. 4.7 (c) we show how the average value of the $c(b)$ constants, $\langle c\rangle = \langle c(b)\rangle_b$, differs from one dataset to the other varying from $\langle c\rangle = 0.8$ in PRB to $\langle c\rangle = 1.7$ in TMN and $\langle c\rangle = 4.6$ in the MPN case, respectively.

The above empirical findings suggest that the mechanism governing the allocation of social capital follow a general form that in its simplest analytical form can be written as:

$$p_b(k) = \left(1 + \frac{k}{c_b}\right)^{-\beta_b}. \tag{4.1.7}$$

In this expression, β_b modulate the tendency to explore new connections, while c_b define the intrinsic characteristic limit of the individual to maintain multiple ties. Although one could imagine more complicate analytical forms, we use this parsimonious approach to characterize the different data sets. Interestingly, we find that in the five co-authorship networks and Twitter, the exponent β is the same regardless of the class b.

Furthermore, the values of c_b are typically peaked around a well defined value. More in detail, we can rescale the proposed functional form in each class b by

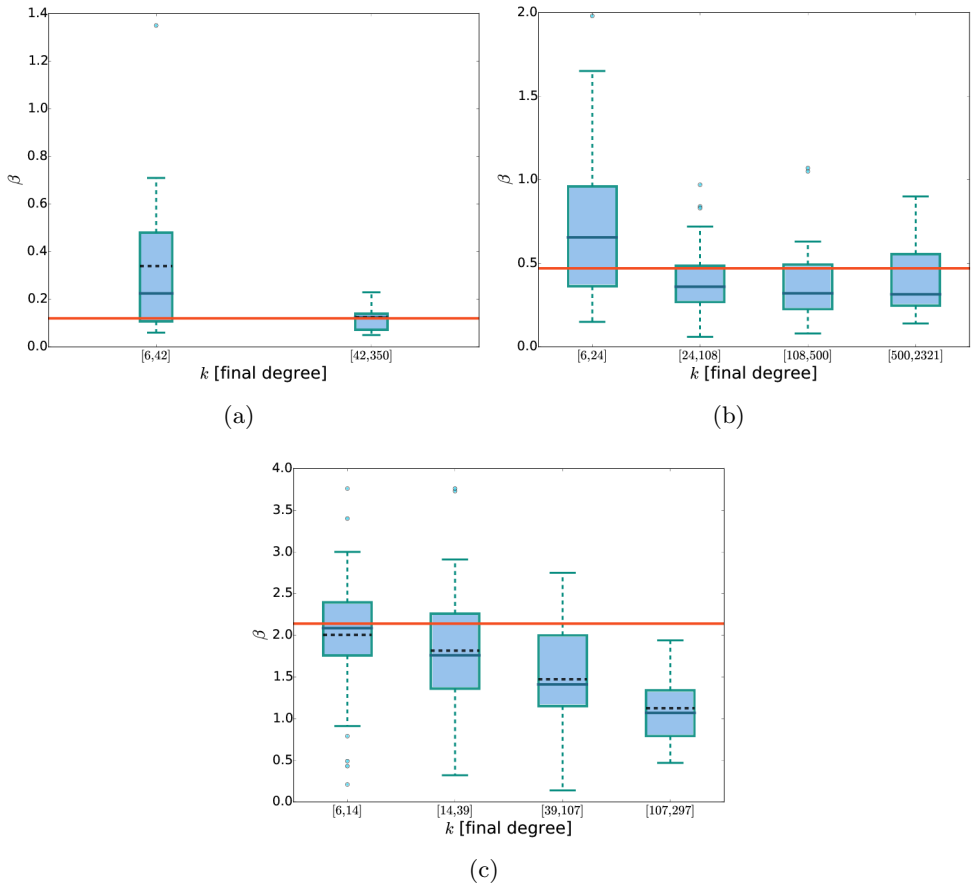
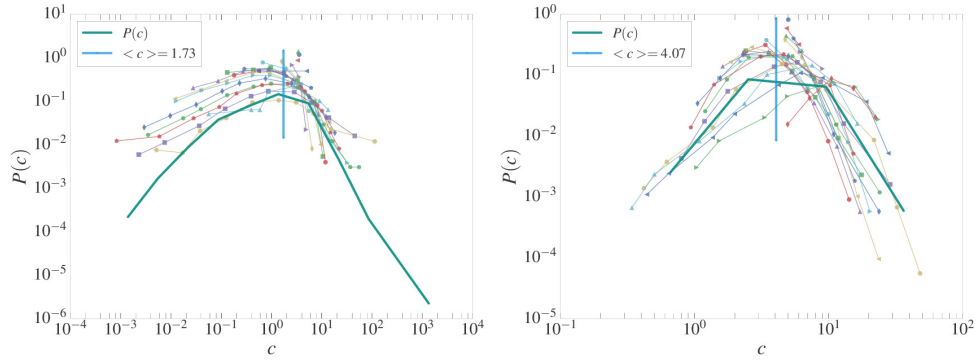
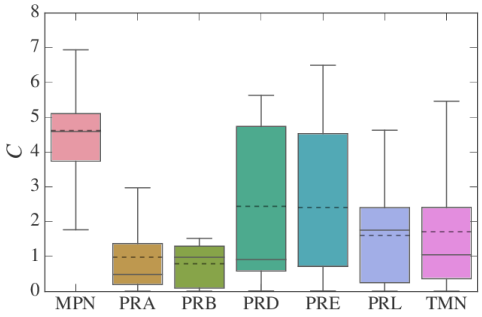


Figure 4.6: The box plot representing the distribution for different range of nodes classes b of the $\beta_{\rm opt}(b)$ for (a) PRB, (b) TMN and (c) MPN. We also show the global optimal value $\beta_{\rm opt}$ (horizontal red line) as found in Eq. (4.1.6). The height of the box corresponds to the lower and upper quartile values of the distribution and the horizontal solid line corresponds to the distribution's median, while the dashed lines indicates the average value for each range of final degree. The whiskers extend from the box to values that are within $1.5x$ the quartile range. As one can see, in both the PRB and TMN datasets the optimal values $\beta_{\rm opt}$ is compatible with the distribution found in all the nodes class ranges (we find the same result for all the other APS datasets analyzed). On the other hand, in the MPN the distribution of $\beta_{\rm opt}(b)$ lowers as the final degree of the class increases. The last group of nodes classes is no more compatible with the overall optimal $\beta_{\rm opt}$, being the distribution centered around $\beta_{\rm opt} \sim 1.1$, in agreement with our estimation of $\beta_{\rm min} = 1.2$.





Dataset	β_{opt}	$\langle c(b) \rangle$
PRA	0.20	1.5
PRB	0.13	0.8
PRD	0.28	2.3
PRE	0.25	2.4
PRE	0.15	1.6
TMN	0.47	1.7
MPN	1.20	4.6

(d)

Figure 4.7: The $P(c)$ distribution of the constant $c(b)$ for the (a) TMN, and (b) MPN case (solid green line). We also compare the global $P(c)$ distribution with the distribution of the $c(b)$ values found within each activity class (solid lines and points): we find that the distribution of the $c(b)$ parameter is more or less activity independent as most of the distribution of the single activity classes follows the same functional form of the total distribution $P(c)$. We then report the average value $\langle c \rangle$ of the $c(b)$ constant for each dataset (vertical cyan line). The latter reads 1.71 for TMN and 4.62 for MPN. Note that in the single β case we evaluate $c(b)$ as the values of $c(b)$ that best fits the b-th $p_b(k, \beta)$ curve fixing β at its optimal value ($\beta = \beta_{opt}$). On the other hand, in the multi-β case we evaluate $c(b)$ as the ones that best fits the $p_b(k, \beta(b))$ curve, where the exponent is now foxed to the $\beta(b)$ value for the memory class b, i.e. to the optimal value for the class b. (c) The box plot showing the global distribution of the constant $c(b)$ in all the datasets analyzed. The height of the box corresponds to the lower and upper quartile values of the distribution and the horizontal solid line corresponds to the distribution's median, while the dashed lines indicates the average value for each range of final degree. The whiskers extend from the box to values that are within $1.5x$ the quartile range. (d) In this table we report all the values of the reinforcement exponent β_{opt} and the average reinforcement constant $\langle c(b) \rangle$. For the MPN case we report the $\beta_{\text{opt}} = \beta_{\text{min}}$ and the constant values are evaluated for each nodes class b using its optimal value of β, $\beta_{\text{opt}}(b)$.

defining the variable $x_b = k/c_b$, yielding

$$p_b(x_b)^{\frac{1}{\beta}} = (1 + x_b)^{-1}. \tag{4.1.8}$$

In the presence of a single exponent β characterizing the system, as shown in Figure 4.5 (a-c), all empirical curves do collapse on the reference function $(1 + x)^{-1}$. The data collapse however is not occurring in the case of the MPN dataset. In the latter we find a more heterogeneous scenario in which different nodes' classes are characterized by different values of β_b and c_b (see Figures 4.6 and 4.7).

4.2 The Model

Given the preliminary measures performed we now introduce the modeling framework we use in this work. This is based on a Master Equation formalism.

4.2.1 Memory-Less approach

To show the consistence of our analytical approach with previous results found in references [209, 218] let us develop the memory-less version of the model.

In the latter, (i) we start with a graph $\mathcal{G}$ of N nodes initially disconnected. (ii) Each node i is initially assigned with an activity potential a_i determining its activation ratio, i.e. the probability to get active in a given time interval δt reads $a_i \delta t$. (iii) When active, a node connects to a randomly chosen node of the network. For each evolution step, each node is given the possibility to get active. (iv) After all the connections have been established, we remove them and pass to the sequent evolution time step resuming from step (iii).

The $P(k,t)$ distribution

The quantity we investigate in this work is the $P_i(k,t)$ probability distribution measuring the probability for a node of activity a_i to have degree k at time t.

The discrete time equation for $P_i(k,t)$ reads:

$$P_i(k,t+1) = \tag{4.2.1}$$

$$a_i \frac{N-k}{N} P_i(k-1,t) + a_i \frac{k}{N} P_i(k,t) + P_i(k-1,t) \sum_{j \nsim i} a_j \sum_h \frac{P_j(h,t)}{N} +$$

$$P_i(k,t) \sum_{j \nsim i} a_j \sum_h P_j(h,t) \frac{N-1}{N} + P_i(k,t) \sum_{j \sim i} a_j + P_i(k,t)(1 - \sum_j a_j). \tag{4.2.2}$$

The equation is obtained in the approximation where $a_i \ll 1$, so that between time t and $t+1$ only one site can be active. We will also consider the approximation $1 \ll k \ll N$ i.e. the integrated number of neighbors of a site is much larger than 1 but much smaller than the total number of sites. The first term of the sum represents the probability that the site i is active and a new link is added to the system. The second term is the probability that the site i is active but this site

connects to a site that has been already linked. In the third and fourth terms, the symbol $\sum'$ denotes the sum over the sites that are not yet connected to i. In particular, the third term represents the probability that one of these sites is active and that it connects to i. The fourth term is the probability that one of these sites is active but no link between j and i is established. The fifth term is the probability that one of the sites already connected to i is active; in this case no new link is added to i. Finally, the last term represents the probability that at time t all the sites are not active. For $k \ll N$, the second term can be neglected. After some algebra we obtain the equation:

$$P_i(k,t+1) - P_i(k,t) = -\left(P_i(k,t) - P_i(k-1,t)\right)\left(a_i + \frac{1}{N}\sum_{j\nsim i} a_j\right) \tag{4.2.3}$$

where we have taken into account that $\sum_h P_j(h,t) = 1$. For $k \ll N$, we assume that $\frac{1}{N}\sum_j{}' a_j = \langle a \rangle$ i.e. the average value of the activity. In the limit of large time and large k we can write a continuous equation in t and k, and neglect the index i since all the sites with the same activity follow the same equation. We finally get:

$$\frac{\partial P_i(k,t)}{\partial t} = (a_i + \langle a \rangle)\left(-\frac{\partial P_i(k,t)}{\partial k} + \frac{\partial^2 P_i(k,t)}{\partial k^2}\right). \tag{4.2.4}$$

The solution of Eq. (4.2.4) is straightforward:

$$P_i(k,t) = (2\pi(a_i + \langle a \rangle)t)^{-\frac{1}{2}} \exp(-\frac{(k - (a_i + \langle a \rangle)t)^2}{2t(a_i + \langle a \rangle)}). \tag{4.2.5}$$

The average degree $\langle k(a,t) \rangle$

In the large time limit the solution presented in Eq. (4.2.5) reduces to a delta function:

$$P_i(k,t) = \delta(k - (a_i + \langle a \rangle)t). \tag{4.2.6}$$

This relation provides the simple average degree growth

$$\langle k(a,t) \rangle = (a + \langle a \rangle)t \tag{4.2.7}$$

that was already obtained in previous works [209, 218] and presented in Section 3.3.1.

The degree distribution $\rho(k)$

Besides the $P(k,t)$ functioning and the average degree growth for nodes of activity a, Eqs. (4.2.5) and Eq. (4.2.7) also set the relation between the activity distribution $F(a)$ and the asymptotic degree distribution $\rho(k)$ of a network. Indeed, one can write the predicted $\rho(k)$ starting from the $F(a)$ distribution and substituting $k \sim at$ as found in Eq. (4.2.7). By assuming an activity distribution $F(a) \propto a^{-\nu}$ the predicted degree distribution reads:

$$\rho(k) \propto k^{-\nu}. \tag{4.2.8}$$

As one can see, the heterogeneous degree distribution is naturally introduced by the broad distribution of the activity potential a.

4.2.2 Plugging in the reinforcement process

The model presented in Section 4.2.1 is a basic model as it contains no correlations on an agent's story at all. In particular, the probability for a node i to re-call an already contacted node is independent of the node degree. While simple to describe and solve analytically, this model is not realistic, as there are no correlations in the each agent's history. Moreover, the probability to call an already contacted node is always small as $k/N \ll 1$ (and thus the probability to call a new node remains ~ 1 even at large degree k). However, as shown by the above empirical measures and previously in Section 3.4, real-world systems features a strong reinforcement process on the established links, as the probability $p_i(k)$ to call a new node at degree k decreases as the degree k increases.

For this reason we introduce an extended version of the model described in [229] which includes a reinforcement function $p_i(k)$ that measures the probability for an active node i, that has already contacted k different nodes in the network, to call a new node instead of an already contacted one.

The $P(k,t)$ distribution

We then introduce the reinforcement process $p_i(k)$ using the functional form presented in Eq. (4.1.7), i.e. the probability of adding a new link for the node i of degree k, reads:

$$p_i(k) = (1 + k/c_i)^{-\beta}. \tag{4.2.9}$$

By plugging Eq. (4.2.9) into Eq. (4.2.1) for node i, we get:

$$\begin{aligned} P_i(k,t+1) = P_i(k-1,t)\left[a_i p_i(k-1) + \sum_{j \nsim i} a_j \sum_h \frac{p_j(h)}{(N-h)} P_j(h,t)\right] + \\ P_i(k,t)\left[a_i[1-p_i(k)] + \sum_{j \nsim i} a_j \sum_h \left(1 - \frac{p_j(h)}{N-h} P_j(h,t)\right)\right] + \\ P_i(k,t)\left[1 - \sum_j a_j\right], \end{aligned} \tag{4.2.10}$$

where N is the number of nodes in the network, $\sum_{i \nsim j}$ is the sum over the nodes not yet connected to i and $\sum_j$ is the sum over all the N nodes of the network. Each term of Eq. (4.2.10) corresponds to a particular event that may take place in the system, as already presented in the paper. For instance, the first term of the l.h.s. of Eq. (4.2.10) takes into account the increment of the node i's degree from $k-1$ to k. This may happen whether because node i gets active and contacts a new node in the system with probability $a_i p_i(k-1)$ or because a node j never contacted before gets active and calls exactly node i with probability $a_j p_j(h)/(N-h)$, being h the degree of j. In the same way, the second line takes into account that node i does not change degree k whether because it calls an already contacted node or because the non contacted nodes call other nodes in the network. The last line of Eq. (4.2.10) considers the possibility that no node in the network gets active.

If we now substitute Eq. (4.2.9) in Eq. (4.2.10), after some algebra we get:

$$P_i(k,t+1)-P_i(k,t)=\frac{a_i c_i^\beta}{(k-1+c_i)^\beta}P_i(k-1,t)-\frac{a_i c_i^\beta}{(k+c_i)^\beta}P_i(k,t)$$
$$-(P_i(k,t)-P_i(k-1,t))\sum_{j\nsim i}a_j\sum_h\frac{P_j(h,t)c_j^\beta}{(N-h)(h+c_j)^\beta}. \quad (4.2.11)$$

Then, by applying the same approximations of large degree k and time t we obtain the continuous equation:

$$\frac{\partial P_i(k,t)}{\partial t}=-a\frac{c_i^\beta}{k^\beta}\frac{\partial P_i(k,t)}{\partial k}+\frac{a_i c_i^\beta}{2k^\beta}\frac{\partial^2 P_i(k,t)}{\partial k^2}+\frac{a_i\beta c_i^\beta}{k^{\beta+1}}P_i(k,t)+ \quad (4.2.12)$$
$$+\left(\frac{1}{2}\frac{\partial^2 P_i(k,t)}{\partial k^2}-\frac{\partial P_i(k,t)}{\partial k}\right)\int da_j F(a_j)a_j\int dc_j\rho(c_j|a_j)\int dh\frac{c_j^\beta}{h^\beta}P_j(h,t),$$

where $\rho(c_j|a_j)$ is the probability for a node j of activity a_j to have reinforcement constant c_j.

The long time asymptotic solution of Eq. (4.2.12) is of the form:

$$P_i(k,t)\propto\exp\left[-A\frac{(k-C(a_i,c_i)t^{\frac{1}{1+\beta}})^2}{t^{1/(1+\beta)}}\right]. \quad (4.2.13)$$

Moreover, $C(a,c)$ is a constant depending on the activity a and the reinforcement constant c that follows the:

$$\frac{C(a,c)}{1+\beta}=\frac{ac^\beta}{C(a,c)^\beta}+\int da' F(a')\int dc'\rho(c',a')\frac{a'c'^\beta}{C(a',c')^\beta}. \quad (4.2.14)$$

We do not have an exact solution for $C(a,c)$, however $C(a,c)\simeq(ac^\beta)^{1/(1+\beta)}$ for large a.

Let us note that Eq. (4.2.13) can be obtained setting the variable $x=k-C(a)t^{\frac{1}{1+\beta}}$ and substituting it in Eq. (4.2.12) and imposing that $|x|\ll t^{\frac{1}{1+\beta}}$ from Eq. (4.2.12):

$$\frac{\partial P_i(x,t)}{\partial t}=\frac{a_i\beta c_i^\beta}{C(a_i,c_i)^{1+\beta}t}\left(x\frac{\partial P_i(x,t)}{\partial x}+P_i(x,t)\right)+\frac{C(a_i,c_i)}{2(1+\beta)t^{\frac{\beta}{1+\beta}}}\frac{\partial^2 P_i(x,t)}{\partial x^2} \quad (4.2.15)$$
$$-\frac{\partial P_i(x,t)}{\partial x}\int da_j F(a_j)\int dc_j\rho(c_j|a_j)\int dy\frac{a_j\beta c_j^\beta}{C(a_j,c_j)^{1+\beta}t}P_j(y,t)y.$$

The solution of the latter equation is of the form

$$P_i(x,t)\approx t^{-\frac{1}{2(1+\beta)}}\exp\left(-\frac{Ax^2}{t^{1/(1+\beta)}}\right) \quad (4.2.16)$$

thus confirming that x can be considered much smaller than $t^{\frac{1}{1+\beta}}$.

The average degree $\langle k(a,t)\rangle$

An important consequence of equations (4.2.13) and (4.2.14) is that, for a system featuring a reinforcement strength β, the average degree of the nodes belonging to a class b of activity a and constant c grows as:

$$\langle k(a,c,t)\rangle \propto C(a,c)\cdot t^{\frac{1}{1+\beta}}. \tag{4.2.17}$$

In particular, $\langle k(a,c,t)\rangle \propto (at)^{\frac{1}{1+\beta}}$ for large values of the activity a.

As expected, the average degree grows slower than in the memoryless case ($\beta = 0$) where the average degree grows linearly in time, as found in Eq. (4.2.7). Specifically, the exponent β weighting the reinforcement process's strength affects the growth exponent: the stronger the reinforcement attitude (larger β), the slower the growth of the average degree $\langle k(a,t)\rangle$ in time.

The degree distribution $\rho(k)$

The presence of a reinforcement process also affects the asymptotic behavior of the degree distribution $\rho(k)$. Indeed, as already shown in Eq. (4.2.17) gives us the relation between the degree k and the activity a at a given time t, as $k \propto a^{\frac{1}{1+\beta}}$. Thus, given an activity distribution $F(a)$, we can infer the functional form of the degree distribution $\rho(k)$ by substituting $a \to k^{\frac{1}{1+\beta}}$, finding:

$$\rho(k)dk \propto F(k^{(1+\beta)})k^{\beta}dk. \tag{4.2.18}$$

Specifically, by supposing a power-law activity distribution $F(a) \propto a^{-\nu}$ and considering that the degree distribution for a class b is described by Eq. (4.2.13), we obtain

$$\rho(k) \propto k^{-[(1+\beta)\nu-\beta]}. \tag{4.2.19}$$

where we integrated over time t and reinforcement constant $c(b)$ and we considered the asymptotic regime of large time and activity.

In other words, the connectivity patterns emerging from social interactions can be inferred knowing the propensity of individuals to be involved in social acts, the activity, and the strength of the reinforcement towards previously establish ties, β. Finally it is worth remarking that Eqs. (4.2.17, 4.2.19) are not affected by the distribution of c_i. This is an important result as it reduces the number of relevant parameters necessary to define the temporal evolution of the system.

4.2.3 The multi-β case

As we already mentioned, in the MPN dataset we find the evolution of social ties described by a distribution of β rather than a single value of it. This observation points to a more heterogeneous distribution of social attitudes with respect to the other six analyzed datasets. Arguably, such tendency might be driven by the different functions phone calls serve enabling us to communicate with relatives, friends or rather to companies, clients etc.. The need to introduce different values of β in the system complicates the model beyond analytical tractability. Nevertheless, we

can show that the leading term of the evolving average degree can be described by introducing a simplified model, in which the nodes of the system feature different values of β, and undergo a simplified dynamics that neglects, for every node, the effects of links established by others agents in the network.

Simplified version of the model

Thus, the model focuses on a single agent that can only call other nodes in the network (i.e. we neglect the contributions coming from the incoming calls) and whose parameters are a_i, β_i and c_i. In this approximation we have to solve a modified version of Eq. (4.2.10), obtained by discarding all the terms containing the activity a_j of the nodes $j \neq i$:

$$P_i(k, t+1) = a_i p(k-1) P_i(k-1, t) + P_i(k, t) \left[a_i (1 - p(k)) + (1 - a_i) \right]. \quad (4.2.20)$$

The continuum limit for large degree k and time t of Eq. (4.2.20) is:

$$\frac{\partial P}{\partial t} = -a \left(\frac{c}{k} \right)^{\beta} \left[\frac{\partial P}{\partial k} - \frac{1}{2} \frac{\partial^2 P}{\partial k^2} \right]. \quad (4.2.21)$$

The solution for $P_i(k, t)$ is:

$$P_i(k, t) \propto \exp \left[-A \frac{\left(k - C_i t^{\frac{1}{1+\beta_i}} \right)^2}{t^{1/(1+\beta_i)}} \right], \quad (4.2.22)$$

where the C_i now reads:

$$C_i = \left[(1 + \beta_i) c_i^{\beta} a_i \right]^{\frac{1}{1+\beta_i}}. \quad (4.2.23)$$

The asymptotic growth of the $\langle k(a,t) \rangle$

Thanks to Eqq. (4.2.22) and (4.2.23) we can write the average degree $\langle k_i(t) \rangle$ growth as:

$$\langle k_i(t) \rangle \propto C_i t^{\frac{1}{1+\beta_i}}. \quad (4.2.24)$$

The result found in Eq. (4.2.24) holds for a single class of nodes with a given set of activity a_i and reinforcement constant c_i and strength β_i.

The average degree $\langle k(a,t) \rangle$ for the activity class a can be computed by integrating over the different values of β_i and c_i:

$$\langle k(a,t) \rangle = \int dc' \int d\beta' \rho(\beta', c'|a) C(a, c', \beta') (t)^{\frac{1}{1+\beta'}} \quad (4.2.25)$$

where $\rho(\beta, c|a)$ is the probability for a node of activity a to have a reinforcement exponent and constant equal to β and c. By assuming that the distribution of the exponent β is independent from a and c we can factor out the time-depend term obtaining for the activity class a:

$$\langle k(a,t) \rangle \propto \int d\beta' \rho(\beta') t^{\frac{1}{1+\beta'}}, \quad (4.2.26)$$

where $\rho(\beta)$ is the probability distribution of the β parameter.

Let us assume that $\rho(\beta)$ can be written as a sum of Kroenecker δ-functions, i.e.:

$$\rho(\beta) = \frac{1}{\sum_i C_i} \sum_{i=1}^{N_\beta} C_i \delta(\beta - \beta_i). \tag{4.2.27}$$

By plugging Eq. (4.2.27) in Eq. (4.2.26) we find that:

$$\langle k(a,t) \rangle \propto \sum_{i=1}^{N_\beta} C_i t^{\frac{1}{1+\beta_i}} \xrightarrow{t\to\infty} t^{\frac{1}{1+\beta_{\min}}}, \tag{4.2.28}$$

so that the minimum value of β_i, $\beta_{\min}$ leads the asymptotic behavior of the $\langle k(a,t) \rangle$ function.

In other words, we find that even in this case $\langle k(a,t) \rangle$ evolves as in Eq. (4.2.17) but with β substituted by $\beta_{\min}$. It is interesting to notice that the nodes characterized by $\beta_{\min}$ are those with the weak tendency to reinforce already established social ties. They are social explorers [242]. Notably, our results, indicate that they lead the growth of average connectivity of the network.

4.3 Results

We now test our results against both synthetic and empirical data.

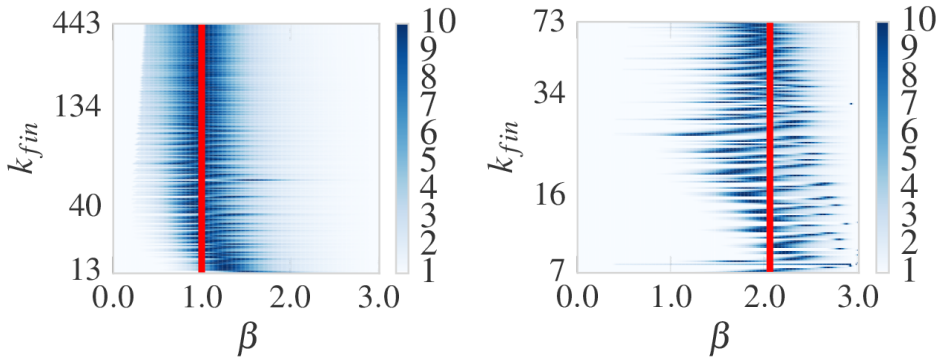
4.3.1 Numerical results

Single β

To check the results of Section 4.2.2 we run numerical simulations featuring the following parameters:

- $N = 10^6$ nodes;
- activity $a \in [\epsilon, 1.0]$ with $\epsilon = 10^{-3}$, power -law distributed so that $F(a) \propto a^{-\nu}$ with $\nu = 2.1$;
- single value of the reinforcement exponent $\beta = \{0.5, 1.0, 1.5, 2.0\}$ and a fixed $c = 1$ for all the nodes;
- $T = 10^5$ evolution steps.

We start with no edge in the system and we draw for each node the activity a_i from the distribution $F(a)$. At each step a randomly chosen node gets active with probability a_i. An active node then connects with probability $p_i(k)$ with a randomly chosen node which have not been yet connected to i or, with probability $1 - p_i(k)$, the node calls an already contacted node and no new connection is added to the system. An evolution step corresponds to N of these elementary steps, i.e. for each evolution step we give, on average, the possibility to make a call to every node in the network.



(a) (b)

Figure 4.8: The heat-map of $-\ln(\chi_b^2(\beta))$ obtained from the simulation in the same way as in Figure 4.4 from real data. As one can see the recovered β_{opt} (red vertical line) is in excellent agreement with the value used in the simulation: 1.0 in the (a) panel and 2.06 in the in the (b) panel.

The results are in excellent agreement with the analytical predictions. First, in Fig. 4.8 we show that the analysis presented in Section 4.1.2 correctly recovers the reinforcement exponent β_{opt}. Indeed the minimum of the $\chi_b^2(\beta)$ are vertically aligned with the value of β fixed in the simulations.

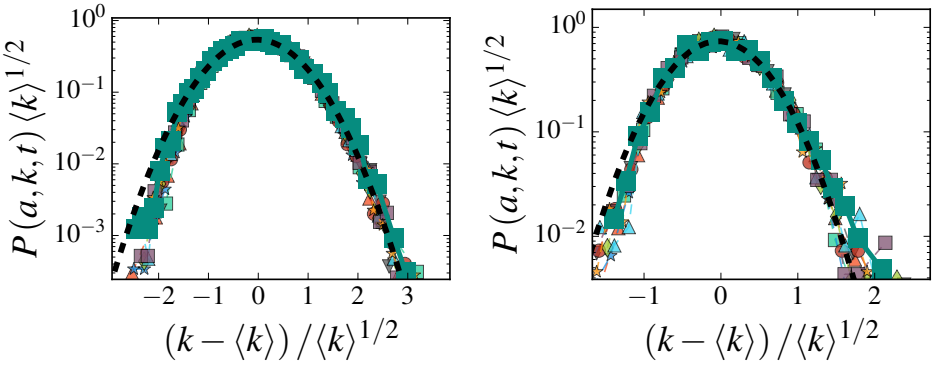
Then, in Fig. 4.10 we present the asymptotic growth of the average degree for an activity class (i.e. a collection of nodes bins b featuring similar activity values) and we compare it with the analytical prediction $\langle k(a,t)\rangle \propto (at)^{\frac{1}{1+\beta}}$. In Fig. 4.9 we show that the shape and the evolution of the $P_i(k,t)$ distribution follows the predicted form of Eq. (4.2.13).

The last check regards the overall degree distribution $\rho(k)$ that should follow Eq. (4.2.19). In Fig. 4.11 we compare the activity distribution $F(a) \propto a^{-\nu}$ and the degree distribution $\rho(k)$. The exponent μ leading the $\rho(k) \propto k^{-\mu}$ is in good agreement with the analytically predicted value $\mu = [(1+\beta)\nu - \beta]$.

Mutli-β

To investigate the multi-β case we performed further numerical simulations considering networks with the following parameters:

- $N = 10^6$ nodes;
- activity $a \in [\epsilon, 1.0]$ with $\epsilon = 10^{-3}$, power -law distributed so that $F(a) \propto a^{-\nu}$ with $\nu = 2.1$;
- (a) reinforcement exponent $\beta = [0.5, 1.5, 2.5]$ with probability $[1/6, 1/3, 1/2]$ (i.e. one sixth of the nodes has $\beta = 0.5$, one third $\beta = 1.5$ and a half of them $\beta = 2.5$ regardless of their activity) and (b) $\beta = [1.0, 1.5, 2.0]$ with equal probability $1/3$.
- fixed $c = 1$ for all the nodes;



(a) (b)

Figure 4.9: The probability distribution $P_a(k,t)$ for a selected activity class a in the simulations with exponent $\beta = 1.0$ (a) and $\beta = 2.0$ (b). We compare different evolution times (see legend) by rescaling the degree $k \to \tilde{k} = (k - \langle k(a,t)\rangle)/\langle k(a,t)\rangle^{1/2}$ on the x-axis and the distribution itself $P_a(k,t) \to \langle k(a,t)\rangle^{1/2} P(a,\tilde{k},t)$ on the y-axis, where $\langle k(a,t)\rangle$ is the average degree at time t for the nodes belonging to the activity class a. We also show the fit of the large time $P(a,k,t)$ with a Gaussian curve (black dashed line) as predicted in Eq. (4.2.13).

- $T = 2 \cdot 10^5$ evolution steps.

The numerical procedure is similar to the one described in the previous section, the difference being that we compute the attachment probability $p_i(k)$ taking into account the reinforcement exponent β_i of the node itself.

In Fig. 4.12 we show that, in both the cases, we can recover the behavior described in Section 4.1.2 for real data. In particular Fig. 4.12(a) (related to the $\beta \in [1,2]$ case) we observe a clear diagonal pattern of the optimal values of the exponent $\beta(b)$ for the b bins that minimize the $\chi^2_b(\beta)$. In particular $\beta(b)$ varies from $\beta \sim 2.0$ values for lower degree nodes bins up to $\beta \sim 1.0$ values for the larger final degree nodes bins. The figure recalls the situation of the MPC dataset presented in Fig. 4.4 (b).

In Fig. 4.13 we show the asymptotic growth of the average degree $\langle k(a,t)\rangle$ together with the predicted asymptotic behavior proportional to $t^{\frac{1}{1+\beta_{\min}}}$. As one can see, numerical results and the suggested analytical solution are in very good agreement in both the cases.

4.3.2 Test on empirical datasets

First, in Fig. 4.14 we present the measured $P(k,t)$ distribution for different datasets. We observe two different results. In all the APS datasets the $P(k,t)$ is found to correctly scales as a Gaussian and the predicted distribution nicely catches the empirical data. In the TMN and MPN instead we find a skewed distribution with a fat tail in the large degree interval.

For what concerns the MPN datasets, the simplest explanation for this behavior can be the presence in the system of different values of reinforcement strength β.

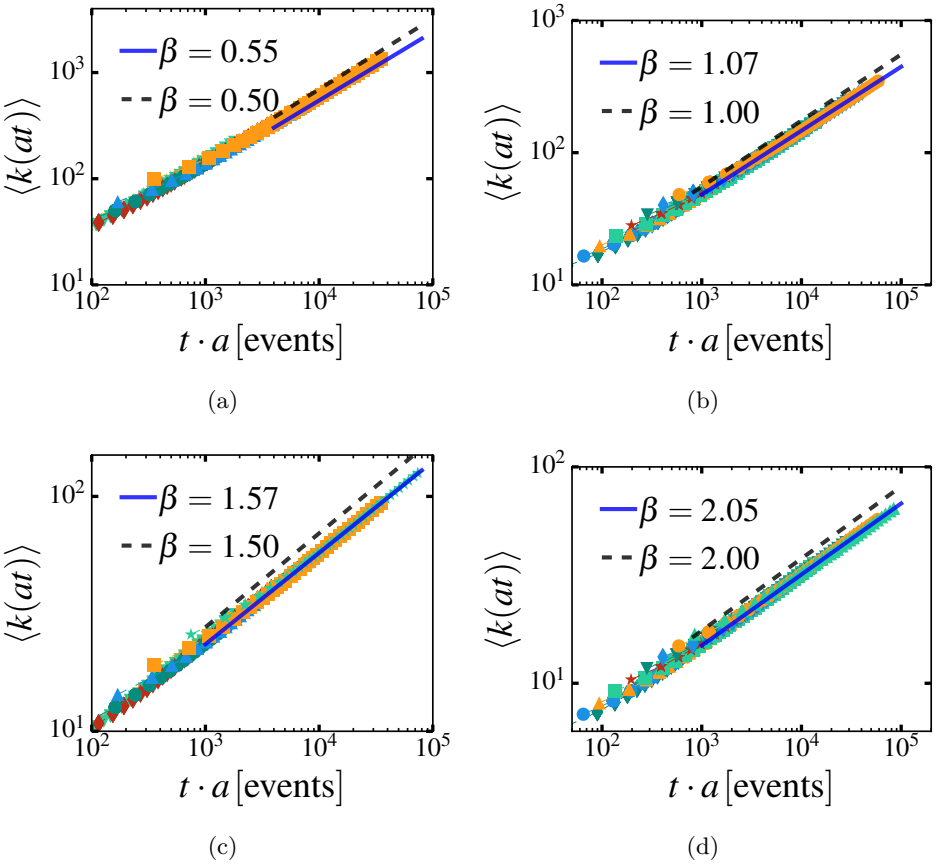


Figure 4.10: The average degree $\langle k(at)\rangle$ for different activity classes in the $\beta = 0.5$ (a), $\beta = 1.0$ (b), $\beta = 1.5$ (c) and $\beta = 2.0$ (d) case. The time is rescaled with activity $t \to at$, so that all the curves collapse on a single behavior. We also fit $\langle k(at)\rangle \propto (t/A)^{\frac{1}{1+\beta^*}}$ (blue solid line) and compare the simulation with the analytical result $\langle k(at)\rangle = A \cdot t^{\frac{1}{1+\beta}}$ (dashed line).

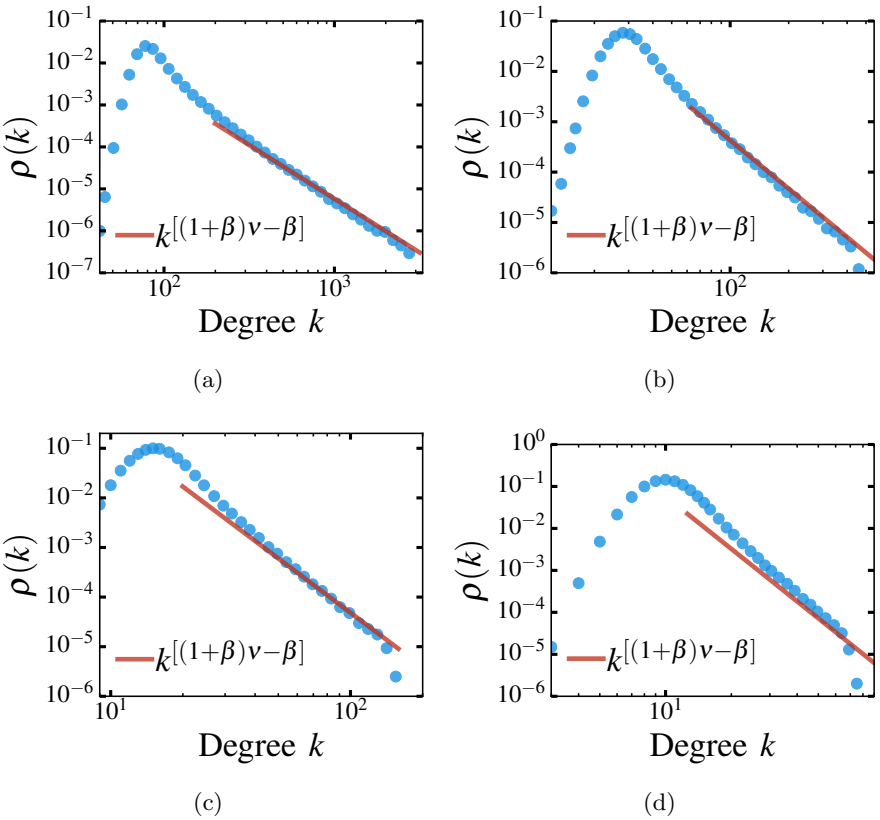


Figure 4.11: The resulting degree distribution of simulations (blue circles) featuring $\beta = 0.5$ (a), $\beta = 1.0$ (b), $\beta = 1.5$ (c) and $\beta = 1.0$ (d). The analytical predictions (given $F(a) \propto a^{-\nu}$, with $\nu = 2.1$) for the scaling exponent are sown (red solid lines).

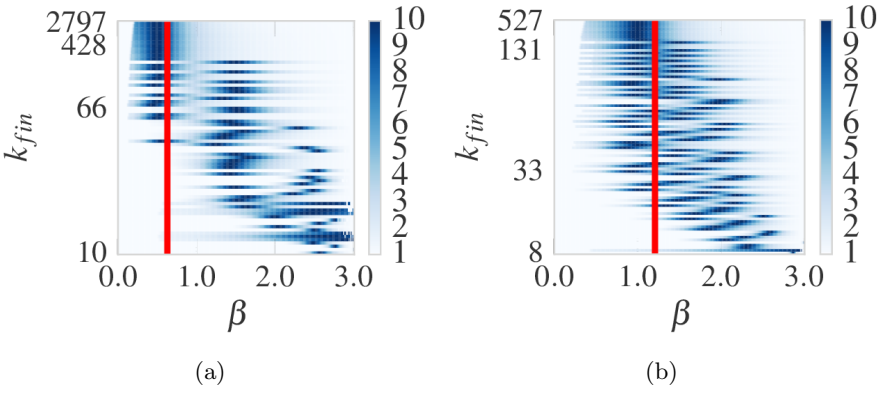
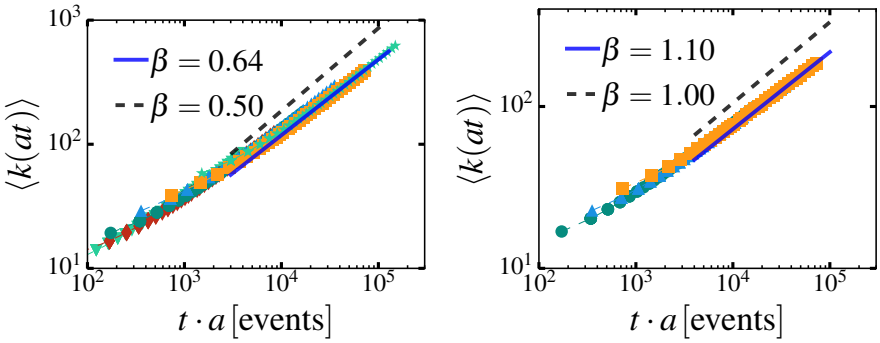
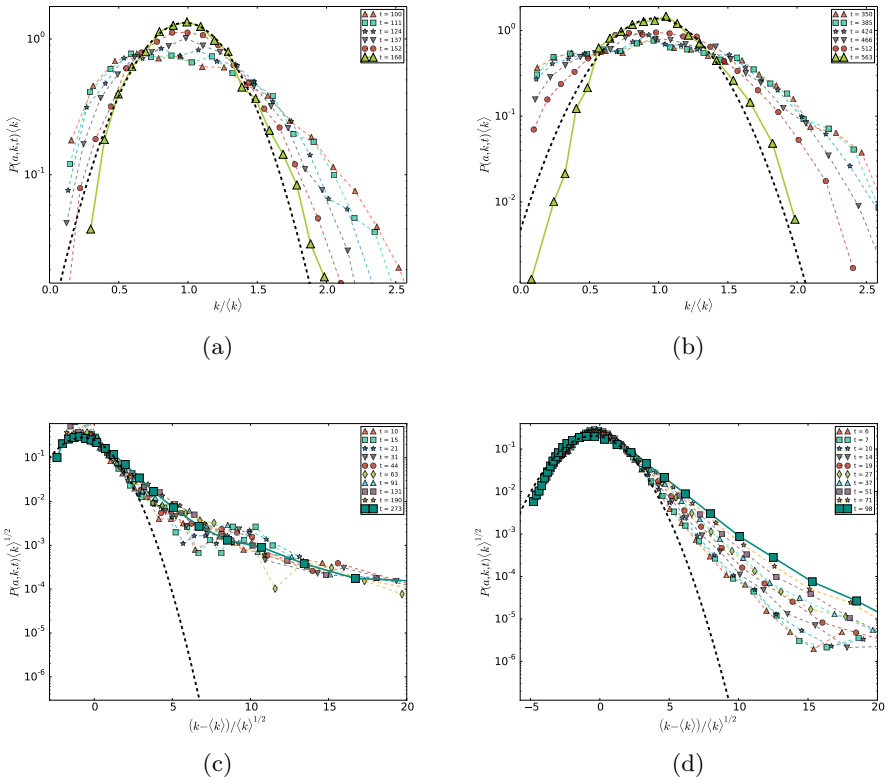


Figure 4.12: The heat-map like matrix of $-\ln(\chi^2(\beta))$ for the simulation with $\beta \in [0.5, 1.5, 2.5]$ (a) and $\beta \in [1.0, 1.5, 2.0]$ (b); the two panels are qualitatively similar to Figure 4.4 that accounts for real data.



(a) (b)

Figure 4.13: The average degree $\langle k(at) \rangle$ for different activity classes in the (a) $\beta \in [0.5, 1.5.2.5]$ and (b) $\beta \in [1.0, 1.5, 2.0]$ case. The time is rescaled with activity $t \to at$, so that all the curves collapse. We also plot the fit $\langle k(at) \rangle \propto (t/A)^{\frac{1}{1+\beta^*}}$ in the long time limit (cyan solid line) and the predicted asymptotic growth $\langle k(at) \rangle = A \cdot t^{\frac{1}{1+\beta_{\min}}}$ (dashed line).



(a) (b)

(c) (d)

Figure 4.14: The $P(k,t)$ distributions (symbols) at different times (see legends for details) for a given activity class of nodes. The data refers to (a) PRE, (b) PRL, (c) TMN, and (d) MPN. We also show the reference normal distribution predicted for the long time asymptotic limit (black dashed lines).

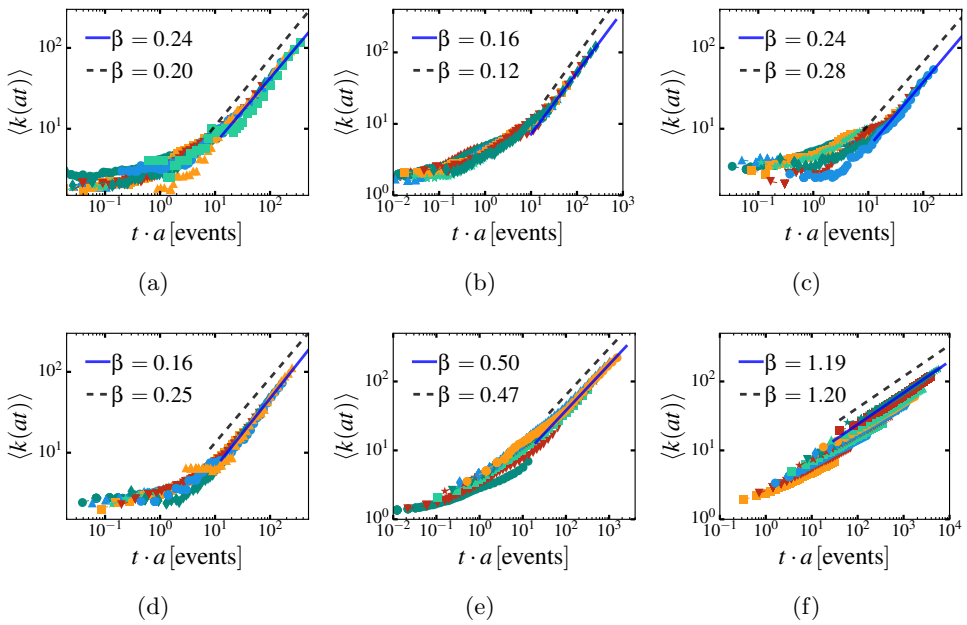


Figure 4.15: The average degree $\langle k(at)\rangle$ (each data series corresponds to a different activity class) for: (a) PRA, (b) PRB, (c) PRD, (d) PRE, (e) TWT and (f) MPC. We compare the data for $\langle k(a,t)\rangle$ with the expected behavior (dashed lines) $(at)^{1/(1+\beta_{\rm opt})}$: in panels (a-e) $\beta_{\rm opt}$ has been evaluated according Eq. (4.1.6), while in the (f) case we use $\beta_{\rm opt} = \beta_{\rm min} = 1.2$. We also plot the power-law fit $\langle k(a,t)\rangle \propto (at)^{1/(1+\beta^*)}$ (solid lines) for comparison.

This can lead to an heterogeneous growth of the degree within a single activity class. So, even though the average degree time-dependence can be forecast by means of the minimum value of β (as we show in Fig. 4.15f), the $P(k,t)$ distribution may be multi-modal. Indeed, we correctly catch the shape of the distribution's peak while we miss the tail where the outliers are found.

For what concerns the TMN dataset instead, we will show in the next chapter that the tails in the $P(k,t)$ distribution may stem from the bursty nature of the datasets.

In Fig. 4.15 we present the comparison between prediction and real data for the PRA, PRB, PRD, PRE, TMN and MPN datasets. The $\langle k(a,t)\rangle$ curve of each activity class is shown with the time rescaled with the activity of each activity class, i.e. $t \to at$. In the MPC case we use as $\beta_{\rm opt} = \beta_{\rm min} = 1.2$ (the β value found in the largest degree bins of Fig. 4.4 (b)). In all the other cases, as the $\beta_{\rm opt}$ fits correctly most of the curves, we use the $\beta_{\rm opt}$ value returned by our analysis.

Finally, in Fig. 4.16 we present the degree distributions, together with the predicted functional form of degree distribution as found in Table (4.2). In the latter we show which parameters of the degree distribution $\rho(k)$ can be inferred knowing the functional form of the activity distribution $F(a)$ and the reinforcement strength β. The change of variable resulting from the relation $\langle k(a,t)\rangle \propto (at)^{1/(1+\beta)}$ then fixes the scale-free parameters of the degree distribution $\rho(k)$. As an example,

PDF	$F(a)$	$\rho(k)$
Power Law	$a^{-\nu}$	$k^{-[(1+\beta)\nu-\beta]}$
Stret. Exp.	$a^{\nu-1}\exp\left[-\lambda a^{\nu}\right]$	$k^{[(1+\beta)(\nu-1)+\beta]}\exp\left[-\tau k^{(1+\beta)\nu}\right]$
Trunc. PL	$a^{-\nu}\exp\left[-\lambda a\right]$	$k^{-[(1+\beta)\nu-\beta]}\exp\left[-\tau k^{(1+\beta)}\right]$
Log-Normal	$\frac{1}{a}\exp\left[-\frac{(\ln(a)-\mu)^2}{2\sigma_a^2}\right]$	$\frac{1}{k}\exp\left[-\frac{(\ln(k)-\gamma)^2}{2\left(\frac{\sigma_a}{1+\beta}\right)^2}\right]$

Table 4.2: The functional form of the activity PDF $F(a)$ and the predicted functional form of the $\rho(k)$ degree distribution as found in Eq. (4.2.18), i.e. by replacing $a \to k^{1+\beta}$. This substitution fixes the scale free parameters of the resulting distribution, i.e. the exponent of the power-law and of the k terms in the first three cases, and the STD $\sigma_k = \frac{\sigma_a}{1+\beta}$ in the Log-Normal case. The free parameters over which we fit the degree distribution are: (i) the cut-off τ in the stretched exponential and power-law with cut-off and (ii) the γ mean value in the Log-Normal case. The selected PDF are, from top to bottom: power law, stretched exponential (Stret. Exp.), power law with cutoff (Trunc. PL) and the Log-Normal distribution.

let us present the detailed procedure in the case of a power-law distributed activity.

Given the activity probability density function $F(a)da$ we know from the solution of the master equation that the degree and the activity are linked by the

$$\langle k(a,t)\rangle \propto (at)^{\frac{1}{1+\beta}}, \tag{4.3.1}$$

so that

$$a \propto k^{1+\beta}, \tag{4.3.2}$$

and

$$da = (1+\beta)k^{\beta}dk. \tag{4.3.3}$$

By substituting Eq. 4.3.3 in Eq. 4.3.1 we get

$$F(a)da \to \rho(k) = F(k^{1+\beta})k^{\beta}dk. \tag{4.3.4}$$

If $F(a) \propto a^{-\nu}$ we get:

$$\begin{aligned} F(a)da &\propto Ca^{-\nu}da \\ \rho(k)dk &\propto F(k^{1+\beta})k^{\beta}dk = Ck^{-[(1+\beta)\nu-\beta]}dk. \end{aligned} \tag{4.3.5}$$

The same procedure can then be repeated for other functional forms of the $F(a)$ (we present the results in Table (4.2)).

4.4 Dynamical processes

As a first insight on how the temporal dimension and the correlations of human activity affect dynamical processes, we investigated an epidemic spreading process

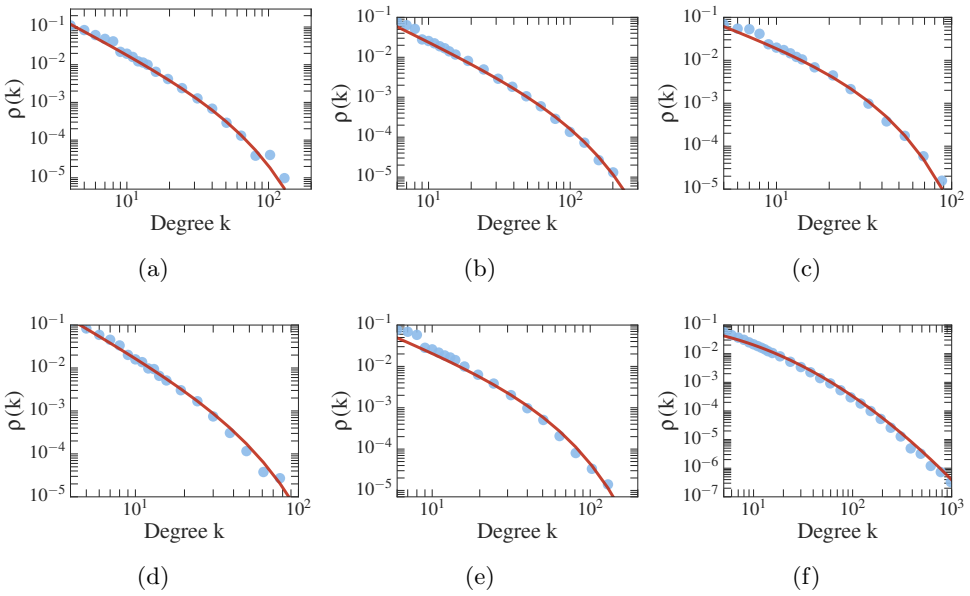


Figure 4.16: The degree distribution $\rho(k)$ for: (a) PRA, (b) PRB, (c) PRD, (d) PRE, (e) PRL and (f) TMN (blue circles). We compare the results with the predicted behavior of Table (4.2) given the parameters of Table (4.1) (red solid lines). We use the single value of $\beta_{\rm opt}$ defined by Eq. (4.1.6) in all the cases. As in Fig. 4.2 we show the data and fit ranging from the lower bound to the 99.9% of the measured data, thus excluding from the visible area the top 0.1% of the degrees values.

on top of an activity driven network. Indeed, the effects introduced by communities and time-varying connectivity patterns on dynamical processes have been scrutinized separately. However, the two attributes are deeply connected. The presence of groups, think for example the interactions network of students in a school, introduces specific dynamical constraints that deeply affect the outcome of spreading processes [72, 163, 172–174].

4.4.1 Modular structure

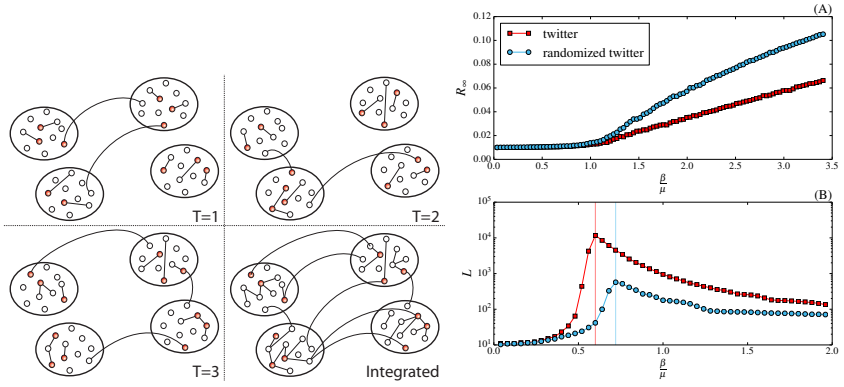
Here we implement a memory-less version of the activity driven model in which we plug links reinforcement and correlations on edges activation by means of an imposed community structure of the network. In other words, nodes are grouped in communities of different sizes and are more likely to call nodes belonging to their same community rather than alters outside their group. Here we try to study the interplay of these two features. To this extent, we introduce a model of time-varying networks with tunable modularity, able to capture several features of real temporal graphs. We derive a complete analytical characterization of the model, and we study the behavior of the Susceptible-Infected-Recovered (SIR) and the Susceptible-Infected-Susceptible (SIS) epidemic processes unfolding on its fabrics.

Remarkably, while the presence of tightly connected clusters inhibits SIR processes, it favorites the spreading of SIS-like diseases. Indeed, using simple physical and statistical arguments we show that the interplay between time-varying and modular properties lower the epidemic threshold of SIS dynamics.

In practical terms, we consider a heterogeneous population of N nodes each of them characterized by an activity rate a_i extracted from a distribution $F(a) \propto a^{-\nu}$, with $a \in [\epsilon, 1]$. Furthermore, each node is assigned to a group/community. To consider empirical evidences the size of each community is extracted from a heavy-tailed distribution, i.e. $P(s) \propto s^{-\omega}$ with $s \in [s_{min}, \sqrt{N}]$ [261, 262]. Given these settings a dynamical model of the network's evolution is defined by the following steps (see Fig. 4.17(a)). At each time t, the network, G_t, starts with N disconnected nodes. With probability $a_i \Delta t$ each vertex i is active and willing to create m connections. With probability μ each link is generated within the node's community, and with probability $1 - \mu$ with nodes in any other groups. In both cases nodes are selected randomly. At the next time step $t + \Delta t$ all the edges in G_t are deleted. All the interactions have a constant duration Δt. In the following, without loss of generality, we fix $\Delta t = m = 1$. We present a detailed analytical analysis of the evolving network in Appendix A.1.

4.4.2 Results

Let us turn our attention on the dynamical properties of SIR and SIS processes unfolding on the proposed model (we refer to Section 3.4.2 for the processes definition and evolving rules). Specifically, we focus on a central concept of contagion phenomena: the epidemic threshold. This quantity defines the conditions necessary for the spreading of the illness. In annealed networks the threshold is determined by



(a) (b)

Figure 4.17: (a) Schematic representation of the model. In red we show active nodes. Straight lines and arcs describe links connecting nodes in the same or in different communities respectively. In the bottom right panel we show the integrated network obtained as the union of G_1, G_2, G_3.
(b) Panel A: R_∞ as a function of β/γ for SIR processes diffusing on Twitter (red squares) and on the randomized Twitter dataset (blue circles). Panel B: L as a function of β/γ for a SIS models evolving on the same two networks. Each point is the average of 10^2 independent simulations started from 1% of random seeds. We fix $\gamma = 0.5$.

the moments of the degree distribution $P(k)$, while in static graphs the expression is given by the principle eigenvalue of the adjacency matrix $\mathbf{A}$. In time-varying networks instead, the threshold is determined by interplay between the timescales of the contagion and network evolution processes [76, 176, 217, 259, 263, 264].

Indeed, the contagion is described by a timescale, τ_P, defined by the recovery time, i.e. γ^{-1}. Instead, the network by a timescale, τ_G, defined by the convolution of the activity rates, a^{-1}, of each node. Starting from these observations, we can derive the mean-field level dynamical equations describing the contagion process in modular activity driven networks. Let us define the activity block variable S_a^t, I_a^t, and R_a^t as the number of susceptible, infected and recovered individuals, respectively, in the class of activity a at time t. This allows us to write the mean-field evolution of the number of infected individuals, for a SIR process, in each group of nodes with activity a

$$d_t I = -\gamma I + \lambda \langle a \rangle I + \lambda \Theta + \lambda \mu \sum_s (\langle a \rangle^s - \langle a \rangle) I^s, \tag{4.4.1}$$

$$\begin{aligned} d_t \Theta &= -\gamma \Theta + \lambda \langle a^2 \rangle I + \lambda \langle a \rangle \Theta + \\ &+ \lambda \mu \sum_s \left[(\langle a^2 \rangle^s - \langle a^2 \rangle) I^s + (\langle a \rangle^s - \langle a \rangle) \Theta^s \right], \end{aligned} \tag{4.4.2}$$

where I^s and I are the number of infected in communities of size s and in the whole network, respectively. Moreover, we defined $\Theta = \sum_a a I_a$, and $\Theta^s = \sum_a a I_a^s$. $\langle a^x \rangle^s = \sum_a N_a^s a^x / s$ describes the moments of the activity distribution in any community of size s. The epidemic threshold can be derived evaluating the principal

eigenvalue of the Jacobian matrix of the system of differential equations in I and Θ [176].

As expected, the spreading condition is determined by the interplay between the timescale of the contagion process and the timescales of the network. In the opposite limit $\mu \to 1$ networks are extremely modular. Fluctuations become important and the symmetry between SIR and SIS breaks, as we already shown in Section 3.4.2.

In order to numerically determine the threshold of SIR models we study the final fraction of recovered nodes, R_∞, as a function of β/γ. Indeed, this quantity acts as order parameter of a second order phase transition [15]. For SIS processes instead, the order parameter is the final fraction of infected individuals, I_∞. The numerical estimation of this quantity is challenging as requires the precise determination of endemic configurations. For this reasons we follow Ref. [217] measuring the life time of the disease, L, that acts as the susceptibility. This quantity is defined as the average time the disease needs to either die out or reach a macroscopic fraction, Y, of the populations. Without loss of generality, we start our simulations setting 1% of randomly selected nodes as initial infected seeds, $\gamma = 0.01$, and $Y = 0.5$.

Although, the modeling framework presented captures realistic activity and community size distributions it completely neglects others important features of real networks such as burstiness, and more complex temporal/structural correlations. It is then crucial grounding the picture emerging from synthetic models with a real world system. To this extent, we consider the online temporal, and modular network of Twitter Mention Network. We focus on tree months of data coarse-grained at a time resolution of one day. To single out the effects introduced by communities on contagion processes, we consider also a randomized version of the dataset. Here the interactions at each time are shuffled, modules are destroyed, but the sequence of activation times for each node, the final time integrated degree distribution, and the degree distribution at each time step are preserved. Using these two networks we study the dynamical properties of SIR and SIS processes unfolding on their structure. In Figure 4.17(b)A-B we present the results. Interestingly, the modular properties of the real network seems to slightly influence the threshold of SIR models. Nevertheless, the presence of communities reduces the impact of the disease, i.e. final the number of recovered nodes. In the case of SIS processes instead, communities have a larger effect shifting the threshold to smaller values. These results qualitatively confirm what observed in synthetic systems.

Chapter 5

Letting burstiness in

So far we measured, characterized and modeled two of the three main mechanisms playing a central role in social networks evolution: the individual's activity and its social capital allocation. The missing ingredient is burstiness. Here we present a preliminary analysis and modeling of bursty behaviors within the activity driven framework.

In Section 5.1 we first present the preliminary measure aimed at the burstiness characterisation. Given these results, we introduce the analytical framework used to develop and inspect the dynamical model. We solve the model and show the relevant results concerning the asymptotic behavior and the scaling properties of both the $P(k,t)$ distribution and average degree $\langle k(a,t)\rangle$. We then test our results against numerical simulations and check the analytical predictions on real-world datasets.

Finally, in Section 5.2 we point out some of the shortcomings of the models presented up to this point together with some considerations on activity driven models. This discussion will then lead us to the next chapter, in which we will introduce a novel approach to the development of temporal and dynamical models of real-world social networks's evolution.

5.1 The model

In the following we analyze the seven real world datasets presented in the previous chapter to characterize and measure their burstiness. Based on these measures, we further improve our activity-driven-like model so as to account both for the reinforcement process of ties and the bursty inter-event time distribution. Especially, burstiness is encoded by enforcing a skewed inter-event time distribution while the reinforcement mechanism is still defined as before.

We then build the analytical framework and solve the system dynamics in the long time limit, finding a non trivial phase diagram for the interplay of the two introduced mechanism. Indeed, we find situations in which the burstiness governs the evolution of the network and others where the dynamics is completely determined by the social capital allocation process. Interestingly, if reinforcement is sufficiently strong, burstiness can be sub-leading even in the presence of large inter-event time fluctuations.

Once we analytically derived the predictions on the relevant network's quantities, we test them against numerical simulations and empirical the datasets featuring burstiness.

As a matter of fact, we find the burstiness to be strong enough to change the asymptotic predictions with respect to the previous chapter in only one dataset, i.e. the twitter mentions network. We then repeat the comparison between analytical prediction and empirical data finding a good agreement between the two.

5.1.1 Burstiness characterization

The first empirical measure we carry out regards the inter-event time distribution for the various datasets. As we showed in Section 3.5 this is the first fingerprint of a bursty behavior as well as the easiest way to implement it in a model.

As we show in Fig. 5.1, the waiting time distribution $\Psi(w_i)$ is found to approximately fall as a power-law in the right tail in all the three layers of human activity we examine, i.e. $\Psi(w_i) \sim w_i^{-(1+\alpha)}$.

Since in the TMN (APS) datasets the data are daily (monthly) aggregated we infer the inter-event time distribution for $w_i \lesssim 24$h ($w_i \lesssim 1$month) by assuming the events done by a node within a single day (month) to be homogeneously distributed during the 24 hours of the day (30 days of the month). As we are measuring the α exponent leading the $\Psi(w_i) \propto w_i^{-(1+\alpha)}$ in the right tail of the distribution this assumption does not change the resulting α. The MPN datasets is instead time-resolved to the second, so that we can accurately measure the inter-event time distribution with no approximations at all.

In Fig. 5.1 we show that the values of the exponent α differ from one dataset to the other. In particular we found the MPN and all the APS datasets to feature values of α and reinforcement strength β that, as we will show later, introduce no corrections to the analytical prediction of the model presented in the previous chapter.

On the other hand, the TMN dataset features an $\alpha \lesssim 1$. With this exponent value, the results here presented introduce some corrections in both the functional form of the $P(k,t)$ and the average degree growth.

For what concerns the reinforcement process, we use the parameters and functional form as shown in the previous chapter as this measure is time-independent. Specifically $\beta = 0.47$ exponent in the TMN.

We also checked that sorting the nodes accordingly to their average inter-event time τ_i and final degree k_i instead of by their activity does not change the measured value of β that reads $\beta = 0.50$ instead of $\beta = 0.47$.

5.1.2 Analytical results

The activity driven time-varying network formalism described so far belongs to the family of the Poisson processes, as for each node we have a characteristic time length $\tau_i \sim 1/a_i$ of the inter event time distribution $P(w_i)$ (i.e. the time w_i between two consecutive social interactions of node i). Moreover, the distribution features a

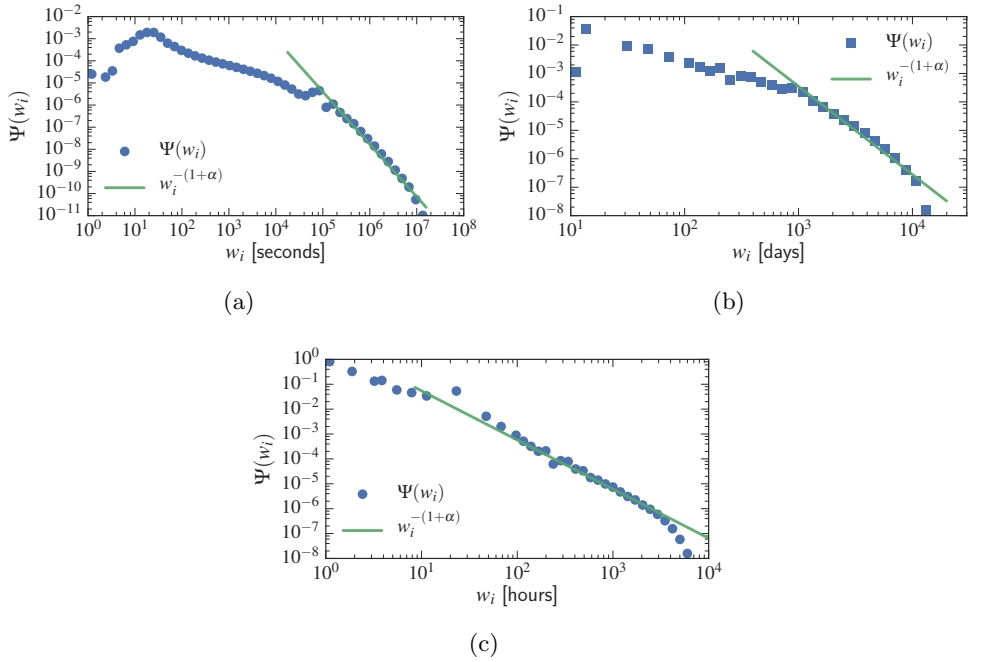


Figure 5.1: The waiting-time distribution $\Psi(w_i)$ for (a) Mobile phone network (circles), (b) the co-authorship network of the Physical Review B journal (squares), and (c) the TMN. We also show the fitting curve of the right tail $\Psi(w_i) \propto w_i^{-(1+\alpha)}$ giving $\alpha \sim 1.45$ for the mobile calls dataset, $\alpha \sim 2.1$ for PRB and $\alpha = 0.95$ in the TMN. Given these results, and provided that the (minimum) value of β found in the mobile phone call dataset is $\beta_{\min} = 1.2$ we conclude that the first two system are above the $\alpha = (2\beta+2)/(2\beta+1)$ curve, thus falling in the SupBR regime. The TMN falls in the StrBR region instead.

finite variance as the fluctuations of the inter event time are of the same time scale of τ_i itself.

This is at variance with the preliminary measures of Section 5.1.1, where we found a strongly heterogeneous waiting time distribution $P(w_i)$that can be well approximated by a power-law decay in the long time limit.

Given these insights, we implement the bursty nature of human dynamics by replacing the activity driven network with an imposed inter event waiting time w_i for node i drawn from a power-law distribution $\Psi(w_i)$:

$$\Psi(w_i) = \frac{\alpha}{\tau_i^{-\alpha}} w_i^{-(1+\alpha)}, \quad w_i \in [\tau_i, +\infty), \tag{5.1.1}$$

where α is the exponent characterizing the distribution and τ_i is the lower cut-off of the distribution. The latter represents the characteristic timescale for the node i, i.e. $\tau_i \sim 1/a_i$, as the γ-th moment of the distribution $\Psi(w_i)$ reads $\langle w_i^\gamma \rangle \sim \tau_i^\gamma$. Therefore, if we let also the τ_i to be heterogeneously distributed as

$$\Phi(\tau_i) \propto \tau_i^{\nu-1}, \tag{5.1.2}$$

for small τ_i, we obtain a network in which the corresponding activity potential a_i is broadly distributed. In particular the activity distribution behaves as $F(a_i) \propto a_i^{-(\nu+1)}$ for large a_i.

By adding the reinforcement process in the picture we can now formulate our developed model. We start defining the network $\mathcal{G}$ containing N nodes. Initially, we assign to each node i a lower cut-off τ_i of the inter event time distribution accordingly to Eq. (5.1.2) and we then set the integrated degree $k_i = 0$ for all nodes. We start the evolution of the system by extracting, for each node i, the time w_i at which the node will get active for the first time. We then activate one node at a time by their next activation time. When active, an agent i calls a new, randomly chosen agent in the network with probability $p_i(k_i) = (1 + k_i/c_i)^{-\beta_i}$; in this case the value of k_i is increased by one both for the connecting and the connected nodes. Otherwise, with probability $1 - p_i(k_i)$, the agent i interacts with a randomly chosen node which have been already connected to i. The interaction is then removed and we fix the time of the next activation for node i by drawing a waiting time w_i from the probability distribution of Eq. (5.1.1).

Given this definition of the model, we can write the master equation (ME) in the single agent approximation, i.e. agents can only attach to other nodes and never get contacted by them. In this approximation, the problem can be analytically solved. In the single agent approach τ_i is fixed to the value $\tau_i = \tau_0$ of the considered agent. Using these assumptions, let us call $Q(k,t)$ the probability that an agent makes a call at time t and after the call its degree is k. The ME governing the evolution of $Q(k,t)$ then reads

$$\begin{aligned} Q(k,t) = \tilde{\mathcal{C}} \int_{\tau_0}^{+\infty} \frac{Q(k-1, t-t')}{t'^{\alpha+1}} \frac{c^\beta}{(c+k-1)^\beta} dt' + \\ \tilde{\mathcal{C}} \int_{\tau_0}^{+\infty} \frac{Q(k, t-t')}{t'^{\alpha+1}} \left(1 - \frac{c^\beta}{(c+k)^\beta}\right) dt' + \delta(k,0)\delta(t,0), \end{aligned} \tag{5.1.3}$$

where α is the exponent driving the inter-event time distribution $P(t) = \tilde{\mathcal{C}} t^{-(\alpha+1)}$ and $\tilde{\mathcal{C}}$ is the normalization constant $\tilde{\mathcal{C}} = \alpha \tau_0^{\alpha}$. The first term accounts for the probability that the nodes gets active and calls, with probability $p_i(k)$, a never contacted node, while the second term accounts for the probability for the active node to contact an already contacted neighbor.

The $P(k,t)$ distribution

To obtain the probability distribution $P(k,t)$ that the agent has degree k at time t we must integrate Eq. (5.1.3) so that:

$$P(k,t) = \int_{\tau_0}^{t} dt' Q(k, t-t') \int_{t'}^{+\infty} d\tau \frac{1}{\tau^{\alpha+1}}. \tag{5.1.4}$$

Let us perform the Fourier Transform of Eq. (5.1.3) in time, by sending the integration variable $t \to (t-t')$ we get:

$$\begin{aligned} \tilde{Q}(k,\omega) = & \\ & \tilde{\mathcal{C}} \Bigg[\frac{c^{\beta}}{(c+k-1)^{\beta}} \tilde{Q}(k-1,\omega) \int_{\tau_0}^{\infty} \frac{e^{i\omega t'}}{t'^{\alpha+1}} dt' + \\ & \left(1 - \frac{c^{\beta}}{(c+k)^{\beta}}\right) \tilde{Q}(k,\omega) \int_{\tau_0}^{\infty} \frac{e^{i\omega t'}}{t'^{\alpha+1}} dt' \Bigg] + \\ & \delta(k,0). \end{aligned} \tag{5.1.5}$$

By taking the limit $k \to \infty$ of Eq. (5.1.6) we end up with

$$\begin{aligned} \tilde{Q}(k,\omega) = & \\ & \tilde{\mathcal{C}} \left[\left(\frac{c}{k}\right)^{\beta} \left[\tilde{Q}(k-1,\omega) - \tilde{Q}(k,\omega) \right] \int_{\tau_0}^{\infty} \frac{e^{i\omega t'}}{t'^{\alpha+1}} dt' + \tilde{Q}(k,\omega) \int_{\tau_0}^{\infty} \frac{e^{i\omega t'}}{t'^{\alpha+1}} dt' \right] + \\ & \delta(k,0). \end{aligned} \tag{5.1.6}$$

The issue is now to compute the integral appearing in Eq. (5.1.6). There are three intervals of the exponent α leading to three different results. We present the detailed derivation in Appendix A.2, while we resume here the main results.

In particular we find the $P(k,t)$ distribution to be:

$$P(k,t) \simeq \begin{cases} \frac{1}{(t/\tau_0)^{\frac{\alpha}{1+\beta}}} f_{\alpha\beta}\left(A'_{\alpha,\beta} \frac{k}{(t/\tau_0)^{\frac{\alpha}{1+\beta}}} \right) & \text{if } \alpha < 1, \\ \frac{1}{(t/\tau_0)^{\frac{1}{\alpha} - \frac{1}{(1+\beta)}}} f_{\alpha\beta}\left(A'_{\alpha,\beta} \frac{k - v(t/\tau_0)^{1/(1+\beta)}}{(t/\tau_0)^{\frac{1}{\alpha} - \frac{1}{(1+\beta)}}} \right) & \text{if } 1 < \alpha < \frac{2\beta+2}{2\beta+1}, \\ \frac{1}{(t/\tau_0)^{\frac{1}{2(1+\beta)}}} \exp\left[-A_{\beta} \frac{\left(k - C_{\beta}(t/\tau_0)^{\frac{1}{1+\beta}}\right)^2}{(t/\tau_0)^{1/(1+\beta)}} \right] & \text{if } \alpha > \frac{2\beta+2}{2\beta+1}, \end{cases} \tag{5.1.7}$$

where $f_{\alpha\beta}(x)$ is a non-Gaussian scaling function (see [265]), v is the drift velocity of the peak of the distribution and $A_{\alpha,\beta}$, A_{β} and C_{β} are constants depending on α and β.

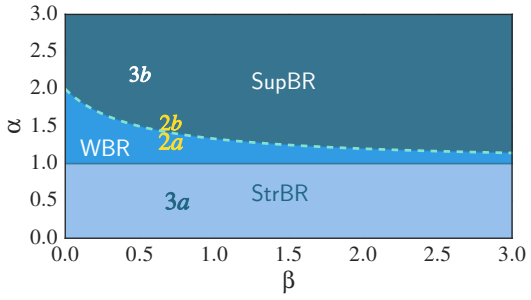


Figure 5.2: The phase diagram for the scaling behavior of the system. We report the delimiting lines of the different scaling regions as found in Eq. (5.1.7). Specifically for $\alpha < 1$ we are in the **Str**ong **B**urstiness **R**egime (StrBR). On the other hand in the $1 < \alpha < 2$ region we find two different behavior and we report the delimiting curve $\alpha = (2\beta + 2)/(2\beta + 1)$ as found in the second case of Eq. (5.1.7) (dashed line). For α below the delimiting line we find a **W**eak **B**urstiness **R**egime (WBR), while above the line we fall in the **Sup**pressed **B**urstiness **R**egime (SupBR). We also show the position on the phase diagram of the simulations presented in this work: we test the transition between the WBR and SupBR regimes in Fig. 5.3 (yellow tags), while in Fig. 5.4 we present the results corresponding to the StrBR and SupBR regions (white and blue tags).

The average degree $\langle k(t)\rangle$

As a consequence of Eq. (5.1.7), we also show the expected growth of the average degree $\langle k(t)\rangle$:

$$\langle k(t)\rangle \propto \begin{cases} t^{\alpha/(1+\beta)} & \text{if } \alpha < 1, \\ \\ t^{1/(1+\beta)} & \text{if } \alpha > 1. \end{cases} \tag{5.1.8}$$

The dynamical phase diagram is summarized in in Fig. 5.2. For $\alpha < 1$, burstiness strongly affects the behavior of the system: the exponents governing the scaling of the $P(k,t)$ distribution and the growth of the average degree $\langle k(t)\rangle$ depend on the value of α and β. In such **Str**ong **B**urstiness **R**egime (StrBR) the scaling function $f_{\alpha\beta}(x)$ is not Gaussian and the exponent leading the growth of $\langle k(t)\rangle$ depends on both the burstiness exponent α and the reinforcement strength β. On the other hand, for $\alpha > (2\beta + 2)/(2\beta + 1)$ we have a **Sup**pressed **B**urstiness **R**egime (SupBR), where the dynamics is independent of α and the reinforcement driven behavior, is fully recovered with a Gaussian scaling function and a connectivity growing as $t^{1/(1+\beta)}$. Finally, for $1 < \alpha < (2\beta+2)/(2\beta+1)$ the average connectivity grows as as $t^{1/(1+\beta)}$ as in the systems without burstiness, while the scaling function is not Gaussian and its scaling length depends on the burstiness exponent α; we will name this behavior **W**eak **B**urstiness **R**egime (WBR).

Degree distribution $\rho(k)$

The degree distribution can be evaluated recalling the scaling form of $P(k,t)$ in the different phases of the system. In particular, starting from Eqq. (5.1.7) and (5.1.8)

we can evaluate, at fixed time t, the $\rho(k)$ distribution finding:

$$\rho(k) = \int_{\tau_{\min}}^{\tau_{\max}} F(\tau_0) P(k,t) d\tau_0, \tag{5.1.9}$$

where $F(\tau)$ is the distribution of the lower-cut off τ, i.e. the lower cut-off of the inter-event time distribution for each agent. If $F(\tau) = \delta(\tau - \tau_0)$ the $\rho(k)$ is trivially the $P(k,t)$. If we instead let the τ of the system be distributed as $F(\tau) \propto \tau^{\nu-1}$ we can evaluate the degree distribution $\rho(k)$. Let us recall that such a distribution of the lower cut-off τ corresponds to an activity distribution going as:

$$F(a) \propto a^{-(\nu+1)}, \tag{5.1.10}$$

where we used the fact that $a \propto \tau^{-1}$.

For the $\alpha < 1$ case we get:

$$\begin{aligned} \rho(k) &\propto \int \tau^{\nu-1} \tau^{-\alpha/(1+\beta)} f_{\alpha\beta}\left(\frac{k}{(t/\tau)^{\alpha/(1+\beta)}}\right) d\tau = \\ &= \int \tau^{\nu-1} \tau^{-\alpha/(1+\beta)} f'_{\alpha\beta}\left(\tau k^{(1+\beta)/\alpha}\right) d\tau = [\tau' = \tau k^{(1+\beta)/\alpha}] = \\ &= \int k^{-1} \left(\frac{\tau'}{k^{(1+\beta)/\alpha}}\right)^{\nu-1} f'_{\alpha\beta}\left(\tau'^{\alpha/(1+\beta)}\right) \frac{d\tau'}{k^{(1+\beta)/\alpha}} = \mathcal{C} k^{-\left[\frac{1+\beta}{\alpha}\nu+1\right]}, \end{aligned} \tag{5.1.11}$$

where $\mathcal{C}$ is a constant with respect to k.

For $\alpha > 1$ we can use the scaling variable $x = k - t^{1/(1+\beta)}$ as a $\delta(k - t^{1/(1+\beta)})$ so that we recover the result of the previous chapter:

$$\rho(k) \propto k^{-[(1+\beta)\nu+1]}, \tag{5.1.12}$$

the only difference being that we previously considered $F(a) \propto a^{-\nu}$.

5.1.3 Numerical results

The non trivial dependence on β and α of the transition line between WBR and SupBR highlights the complex interplay between burstiness and social capital allocation occurring for $1 < \alpha < 2$; Fig. 5.3 shows that the curve $\alpha = (2\beta+2)/(2\beta+1)$ marks a transition from a Gaussian to a non Gaussian scaling function, providing a numerical support to the analytical asymptotic results of Equations (5.1.7), (5.1.8). In particular we observe that in the Fig. 5.3 (a) panel the left tail of the long time curve is slowly increasing and the peak of the distribution cannot be fitted with a normal distribution as it is asymmetric with respect to the average value. On the other hand, in Fig. 5.3 (b) we observe the opposite behavior: the long time curve is slowly converging to the Gaussian scaling and the peak of the distribution is symmetric and well fitted by a normal PDF.

Finally, we find that the single-agent results provide a qualitatively correct description also of the multi-agent case (where multi agents with different activity values are present) as we show in Fig. 5.4. The main difference is, of course, that it takes a larger evolution time for the system to converge to its asymptotic behavior.

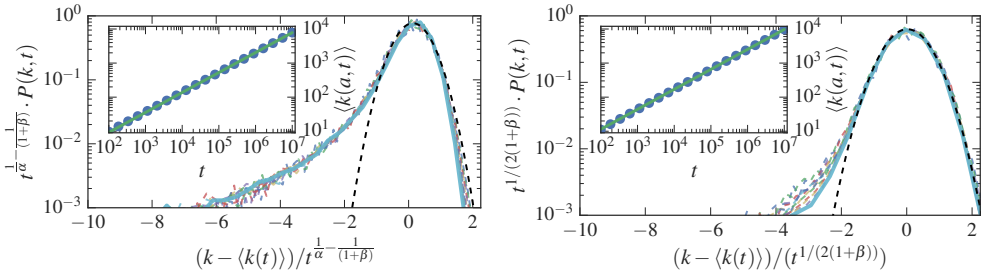


Figure 5.3: (color online) The scaling of the $P(k,t)$ function for (top) $\beta = 0.7$, $\alpha = 1.35$ (WBR region) and (bottom) $\beta = 0.7$, $\alpha = 1.6$ (SupBR regime). In each plot we show the $P(k,t)$ curves of the logarithmically spaced times between $t = 10^4$ and $t = 10^7$ averaged over 10^5 representation of the single agent evolution. The curves referring to the longest time $t = 10^7$ are shown in solid thick line, while shorter times are shown in dashed lines. We show for comparison the fit of the peak of the $P(k,t)$ distribution with a normal distribution (black dashed lines) that correctly matches the SupBR data (bottom) while it completely misses the WBR case (top) as expected by looking at Eq. (5.1.7). In the insets we plot the $\langle k(t)\rangle$ curve for the two cases (circles) and the corresponding analytical prediction of Eq. (5.1.8) (green solid lines).

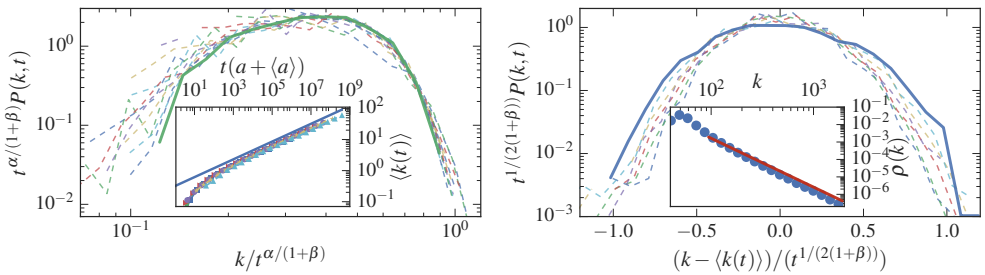
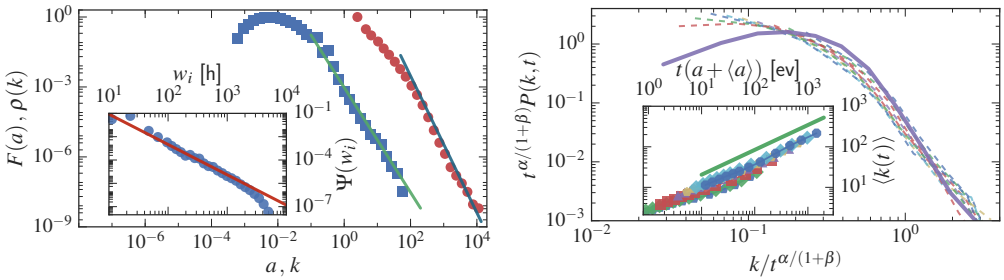


Figure 5.4: (color online). Top: the rescaled $P(k,t)$ distribution for τ_i distributed with $\nu = 1.4$, $\alpha = 0.5$, $\beta = 0.75$. Curves refer to ten logarithmically spaced times between $t = 10^5$ and $t = 10^8$ (dashed lines, the longest time is shown in solid line). The data correspond to the StrBR regime. In the inset we show the average degree growth with time for different activity classes (symbols) rescaled sending $t \to t(a + \langle a\rangle)$, where a is the activity value of the class and $\langle a\rangle$ is the average value of the activity in the system. The analytical prediction of Eq. (5.1.8) is shown for comparison (blue solid line). (Bottom): the $P(k,t)$ distribution for $\nu = 1.2$, $\beta = 0.5$ and $\alpha = 2.2$, at seven logarithmically spaced times in the $10^4 \leq t \leq 10^6$ range. As the τ_i are distributed following a power-law, we show in the inset the experimental degree distribution at the final time (blue circles) and the analytical prediction (red solid line) of Eq. (5.1.13). In these numerical simulation we let the system evolve with the full dynamics, i.e. the agents can both contact and get contacted by other actors in the network.



(a) (b)

Figure 5.5: (a) The activity distribution $F(a \propto 1/\tau_i)$ (blue squares) and the relative fit $F(a) \propto a^{-(1+\nu)}$ (green solid line) giving $\nu = 1.25$ as measured from the TMN dataset. We also show the integrated degree distribution $\rho(k)$ (red circles) and the predicted behavior (blue solid line) found in Eq. (5.1.13). In the inset we plot the waiting-time distribution $\Psi(w_i)$ (blue circles) and the $\Psi(w_i) \propto w_i^{-(1+\alpha)}$ fit (red solid line) giving $\alpha = 0.95$. (b) The $P(a, k, t)$ distribution of Twitter for a selected activity class and at different times, with the degree k rescaled accordingly to the $\beta = 0.47$ found in the previous chapter and the just computed values of α. In the inset we show the average degree $\langle k(t)\rangle$ growth for different activity classes a (symbols) rescaling time as $t \to t(a + \langle a\rangle)$ as in the top panel of Fig. 5.4. The predicted behavior of Eq. (5.1.8) is shown for comparison (green solid line).

In the many agents model we can also consider the prediction of the degree distribution at a given time among the different agents. In particular if the time scales τ_i are distributed according to Eq. (5.1.2), from Eq. (5.1.7) we get that at a given time t the degrees are distributed for large k as

$$\rho(k) \propto \begin{cases} k^{-[\nu(1+\beta)/\alpha+1]} & \text{if } \alpha < 1, \\ \\ k^{-[\nu(1+\beta)+1]} & \text{if } \alpha > 1. \end{cases} \tag{5.1.13}$$

In the insets of Fig. 5.4 we show that the numerical data are well described by the asymptotic behavior in (5.1.13).

5.1.4 Comparison with empirical data

To make some contact with experimental results, we check how the rescaled $P(k, t)$ compare to the proposed scaling picture. We analyze the Twitter Mentions Network (TMN) datasets where we found an inter event time distribution approximately following a power-law (see Fig. 5.5 for details).

Notice that, given the measured value of the exponent $\alpha \sim 0.95$, we expect the TMN system to fall in the StrBR region and indeed the data correctly scale as expected, as one can see from Fig. 5.5.

Moreover, let us note that the measured value of $\alpha = 0.95$ slightly smaller than one introduces a small correction on the average degree growth exponent with respect to the $1/(1 + \beta)$ of the previous chapter. This correction let the analytical

prediction to get even closer to the empirical measure of the average degree growth exponent.

As a last remark, the degree distribution is nicely predicted by starting from the inter-event time distribution and the measured α and β exponents as we show in Fig. 5.5 (a).

5.2 Perspectives

In the previous and current chapters we firstly introduced the activity driven modeling, and then improved it with an heterogeneous, enforced waiting time distribution. Despite its simplicity, the model catches several large-scale quantities of real-world networks: the asymptotic degree growth, the shape of the $P(k,t)$ distribution and the connection between the activity (or inter-event time) distribution and the degree one.

As already stated, more complex and "second order" mechanisms are present in human dynamics. For instance, the reinforcement of links do not happen by randomly choosing an alter that a node already contacted. We showed that there are temporal correlations making it more likely for recently activated edges to get active again [254].

On the other hand, one cannot consider the characterization of burstiness as complete by simply imposing a skewed waiting-time distribution. As we discussed, there are more complex structures and cyclical temporal patterns that has to be taken in consideration [155, 157, 159, 160, 196, 208, 254].

Nevertheless, our approach provides a powerful framework that can be easily extended and improved by means of additional or different mechanisms shaping the network evolution.

As a second consideration, we only did a preliminary analysis and characterization of dynamical processes on the Time-Varying-Networks that our model generates. Indeed, our work calls for a deeper insight on how burstiness and memory affect not only the asymptotic properties of the network but also the outcome of epidemics processes undergoing on top of it.

Though we characterized the response of SIS and SIR processes to the insertion of correlations by means of a community structure, one should also test these dynamical processes on top of networks generated by the model just introduced.

As one could expect, both the spreading of diseases and diffusion of rumors will slow down because of the introduction of both reinforcement process and burstiness. Indeed, we showed that edges reinforcement slows down SIR processes and rumor spreading. In the same way it is easy to expect that highly fluctuating inter-event times will act as a "quarantine" for such spreading processes. This concept can be summarized in the recently coined "small but slow world" [186].

Chapter 6

A novel approach to temporal networks modeling

6.1 The need to overcome Activity-Driven approach

In the previous two chapters we introduced and developed an activity driven model featuring burstiness and links reinforcement. The model precisely reproduces many asymptotic properties of the real-world networks and then serves as a valid testbed for the analysis and characterization of dynamical processes. This model represents a relevant advance in the state of the art of activity driven networks modeling. Indeed, it provides analytical predictions of systems featuring different mechanism that are usually considered separately: the heterogeneous activity potential of nodes, their social capital allocation and the bursty nature of human interactions.

The model, however, presents some shortcomings. There are some mechanisms and properties usually found in real-world systems that can be hardly instilled in this modeling framework, while some others are put by hand in the formulation of the model. For instance, think about the heterogeneous activity and degree distributions. These are enforced in the model by an arbitrary choice of the activity distribution. Another delicate point regards the non-equilibrium nature of real-world social networks as they are always found to change in size (typically growing in time). The addition of a supplementary mechanism regulating the entrance of nodes in the network would make the analytical approach unfeasible and increase the modeling complications beyond the scope.

As an additional point, the reinforcement process we encoded in the model is data-driven (i.e. empirically measured from data) and its microscopical interpretation is somewhat mysterious: though we can give a physical meaning to the reinforcement constant c and strength β that each node features, we cannot say a lot about their origin and, most of all, give an explanation on the underlying processes that change their values from one dataset to the other.

It is then worth asking whether is it possible or not to find a more basic modeling framework able to reproduce some of this real-world features from first principles. To this end, in this chapter we propose a novel approach to the modeling of social networks based on the Kauffman's theory of the *adjacent possible* [266]. This model not only replicates all the so far analyzed asymptotic properties of empirical net-

works, but it also naturally reproduces some of the features missed by the activity driven model.

6.2 Correlated novelties

The Kauffman's *adjacent possible* theory [266] has been initially studied in the field of molecular and biological evolution, and it has later been applied to the study of innovation and technological evolution. In short, the adjacent possible consists of all those things (that could be ideas, songs, genomes, technological products, scientific concepts etc. depending on the given context) that are just one step away from what we are aware of or from what actually exists. The path leading to these new things can be formed either by incremental modifications or recombinations of existing ideas/concepts. When we accomplish to enter in contact with one of these new things, part of our adjacent possible becomes actual and gets immediately surrounded by a newly introduced adjacent possible, i.e. the things we can reach always expands and advance forward in front of us just like a moving horizon. Stated in another way, this framework predicts that once we get aware of a one new thing we are implicitly led to encounter another.

This way of thinking applies to our daily lives in which a novelty is often encountered: a new song, a new friend or a new TV-series we watch. Sometimes, one of these novelties may pave the way to a successive one: if we like a certain director we will be more likely to watch all of his movies. In the same way, a WWW page capturing our attention may lead us to explore some of its link to other concepts and pages.

However, the lack of digital records on human activities in the pre-internet era made it hard to measure if and how these commonsensical considerations actually apply to everyday life. Now, however, the availability of large, longitudinal records of human activity, allows for a precise measure of how everyday novelties appear in our life and how they correlate with each other [14].

6.2.1 The single-urn model

The evolution and expansion of the adjacent possible has been recently modeled in terms of a simplified mathematical model based on Polya's urn [267, 268]. The classical version of this model is generalized to allow for novelties to occur and to trigger (facilitate) further novelties. Specifically, the model considers an urn $\mathcal{U}$ containing N_0 distinct elements, represented by balls of different colors (see Fig. 6.1 for details). Each color corresponds to an unique ID representing a word used in a conversation or songs we have listened to. A given sequence $\mathcal{S}$ of these IDs then corresponds to the series of words in a book or to the list of songs listened by an individual. When we withdraw an element from the urn we annotate it in the sequence $\mathcal{S}$ and put it back in the urn together with ρ additional copies of it, thereby reinforcing that ID's likelihood of being drawn again in later extractions. This introduces a "rich-get-richer" dynamics leading to skewed distributions of the number of balls representing the different IDs.

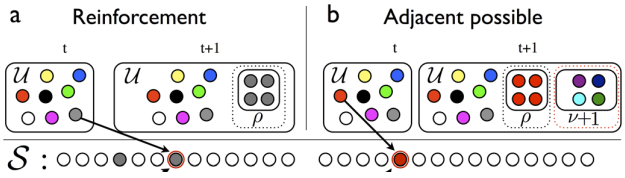


Figure 6.1: Single urn model with novelty triggering (a), (b). (a) The reinforcement step of the evolution. An element (the gray ball) that had previously been drawn from the urn $\mathcal{U}$ is drawn again, annotated in the sequence $\mathcal{S}$ and put back into the urn with ρ additional gray balls. (b) Generic adjacent possible step of the evolution: new ball (red) is withdrawn from $\mathcal{U}$, noted in the sequence and put back into $\mathcal{U}$ with $\nu + 1$ brand new balls and the ρ balls of the reinforcement step. Figure from [14].

To account for the adjacent possible expansion, when something novel occurs the model lets the contents of the urn itself to enlarge. This is done whenever a novel (never extracted before) element appears in the sequence $\mathcal{S}$: the novel ID is put back in the urn together with its ρ reinforcement copies and $\nu + 1$ new distinct IDs that are not present neither in the urn $\mathcal{U}$ nor in the sequence $\mathcal{S}$.

Specifically, the evolution proceeds as follows:

- an element is randomly withdrawn from $\mathcal{U}$ with uniform probability and added to $\mathcal{S}$;
- we put back in $\mathcal{U}$ the extracted element together with ρ copies of it;
- if the extracted element has never appeared before in the sequence $\mathcal{S}$ we add $\nu + 1$ brand new items in the urn $\mathcal{U}$.

In the last step, the $\nu + 1$ new elements represent the set of new possibilities triggered by the novelty, that actually expands the adjacent possible. The number of elements N of the sequence $\mathcal{S}$, i.e. the length of the sequence, equals the number of evolution steps t we repeated the above procedure.

6.2.2 Results

The time dependence by which novelties are introduced in the system can be measured by determining the growth of the number $D(N)$ of distinct elements found in the sequence $\mathcal{S}$ after N evolution steps. Empirical measures reveal a sublinear power-law growth of $D(N) \simeq N^{\eta}$ in time, thus following the Heaps' law [269]. As a consequence, the entrance rate of novelties decreases over time as $N^{\eta-1}$.

To analytically predict the exponent η, let us note that the total number of elements in the urn $\mathcal{U}$ after t evolution steps is $|\mathcal{U}|_t = N_0 + (\nu + 1)D + \rho t$. The dynamics equation for $D(t)$ then reads

$$\frac{dD(t)}{dt} = \frac{U_D(t)}{U(t)} = \frac{N_0 + \nu D(t)}{N_0 + (\nu + 1)D(t) + \rho t}, \tag{6.2.1}$$

where $U_D(t)$ is the number of elements in the urn that at time t have not yet appeared in $\mathcal{S}$, and $U(t)$ is the total number of elements within the urn at time

t. In the numerator we count $\nu D(t)$ unseen balls, as each time a new element is introduced in the sequence $\nu + 1$ brand new elements are added to U while the chosen element is no longer new (-1). In the long time and large $D(t)$ limit, Eq. (6.2.1) can be approximated as

$$\frac{dD(t)}{dt} = \frac{\nu D(t)}{(\nu+1)D(t) + \rho t}, \tag{6.2.2}$$

whose solution reads:

$$D(t) \sim \begin{cases} [(\rho-\nu)t]^{\nu/\rho}, \text{ if } \nu < \rho; \\ \\ \frac{\nu-\rho}{\nu+1} t, \text{ if } \nu > \rho; \\ \\ \frac{\nu}{\nu+1} \frac{t}{\log t}, \text{ if } \nu = \rho. \end{cases} \tag{6.2.3}$$

Another relevant quantity measured in real-world dataset is the frequency of occurrence of a given element within a temporal sequence. In particular, the frequency-rank distribution is found to approximately fall as a power-law with exponent γ (i.e. a Zipf's law since this quantity is discrete) [14, 93, 270]. The frequency-rank distribution is the distribution of number of occurrences by rank, in decreasing order of magnitude. For example, if a sequence contains 10, 20 and 15 appearances of items A,B and C, the frequency-rank distribution is 20,15,10.

In real datasets, the exponent γ fitting the rightmost part of the empirical Zipf's law is found in good agreement with the prediction $\eta = 1/\gamma$, being η the exponent of the Heap's law leading the growth of $D(t)$ [158, 271].

In the modeling scheme, the analytical prediction for the frequency-rank distribution can be estimated by evaluating the number of occurrences n_i of an element i in the sequence $\mathcal{S}$:

$$\frac{dn_i}{dt} = \frac{n_i \rho + 1}{N_0 + aD + \rho t}. \tag{6.2.4}$$

Then, by substituting Eq. (6.2.3) we find two limits:

$$\frac{dn_i}{dt} \simeq \begin{cases} \frac{n_i}{t}, \text{ if } \nu \leq \rho; \\ \\ \frac{\rho n_i}{\nu t}, \text{ if } \nu > \rho, \end{cases} \tag{6.2.5}$$

where t_i denotes the time at which element i appeared for the first time in the sequence. By taking into account the initial condition $n_i(t_i) = 1$ the solution of Eq. (6.2.5) is

$$n_i(t) = \begin{cases} \frac{t}{t_i}, \text{ if } \nu \leq \rho; \\ \\ \left(\frac{t}{t_i}\right)^{\rho/\nu}, \text{ if } \nu > \rho. \end{cases} \tag{6.2.6}$$

To get the Zipf's law describing the frequency-rank distribution of IDs in the sequence $\mathcal{S}$, we evaluate the cumulative distribution $P(n_i \leq n) = P(t_i \geq t/n) = 1 - P(t_i < t/n)$, obtaining:

$$P(t_i < \frac{t}{n}) \simeq n^{-\nu/\rho}, \tag{6.2.7}$$

irrespective of ν and ρ. The probability density function then reads $P(n) = \partial P(n_i < n)/\partial n \sim n^{-(1+\nu/\rho)}$.

6.2.3 Generalization to N-urns

The single urn model can account for the exploration of the adjacent possible for a single individual. It is then reasonable to make an attempt to extend the model so that it can replicate the formation and evolution of social ties among many individuals interacting in a social system.

To this end we propose a generalization of the model in which many urns are interacting amongst them, thus mimicking social interactions between people. In the model each urn plays the role of an individual: it is identified by a unique identifier (ID) and it features a social circle represented by the IDs that are present within the urn itself. The people (IDs) that a given urn will get in contact with are then chosen from these IDs. These IDs correspond to some of the other urns present in the system, and determine the urn (person) with whom the extracting urn will get in contact with. Moreover, each urn contains a special set of $\nu + 1$ IDs (that we dubbed the urn's *sons*) that are exchanged by the urn with the other urns it gets in contact with. Thus, the generic urn i also features a *father* f_i, whose ID is defined to be the ID that created i as a son.

When two urns i and j get in contact we map the reinforcement mechanism of the single urn model in the following way: i adds to its urn ρ copies of j and vice-versa. To account for the addition of novelties of each urn we proceed as follows: whenever two urns i and j get in contact for the first time, i passes a copy of its $\nu + 1$ sons to j that in turn adds them into its urn and vice-versa. In this way, each time an urn i connects with an urn j never contacted before (i.e. the i-th urn experiences a new contact) we facilitate the exploration of new ties and connections for i by adding the sons of j to its urn (i.e. we let novelties to correlate with some others).

In this way, both the size and the number of urns found in the system are changing at each evolution step. We then define the probability for an urn i to get active at a given evolution step t proportional to its size n_i, i.e. the number of balls within the i-th urn. Once the urn i is active we select the "called" urn by a weighted extraction among the D_i IDs contained in the i-th urn. The weight of each ID is proportional to the number of elements with that ID present in the i-th urn. In other words, we uniformly extract an element from the i-th urn, thus accounting for the diverse number of replicas that each ID has in the urn itself. The i-th ID together with the extracted j-th ID (the called actor) then appear in the temporal sequence of events $\mathcal{S}$ as a single contact (i, j).

In a more formal way the model is defined as follows:

1. we start with 2 urns, u_0 and u_1 that have the other agent in the urn (u_0 has 1 copy of u_1 and vice versa); the urns also contains the $\nu + 1$ distinct identities (IDs) of their respective sons. The urns of these $2(\nu + 1)$ sons are initially present in the system as empty urns; the events sequence $\mathcal{S}$ is initially empty;

2. for each time step we extract a "calling" urn i with probability proportional to the urns size and then a "called" urn j amongst the ID-s in the calling urn;

3. we then add ρ copies of i in the j's urn and vice-versa;

4. if it is the first time that j gets called by another node it creates $\nu + 1$ new agents (empty urns) into the system and a copy of each of them into the j's urn; these $\nu + 1$ IDs are the so called j's sons;

5. if it is the first time that i and j get in contact we add a copy of each son of j into the urn of i and vice-versa; it may happen that i is a son of j (and vice-versa). In that case we omit the copy of the i (j) son of j (i) into the i's (j's) urn thus copying ν sons instead of $\nu + 1$; this procedure avoid the possibility for a node to interact with itself;

6. another peculiar event is that a son of i (j) may already be contained in the j's (i's) urn. In this case we simply put an additional copy of the son in the respective urn as if the two ID are simply exchanging their sons. We also checked that the inhibition of this "weak" reinforcement does not affect the exponent characterizing the evolution of the network (i.e. the γ exponent does not change from one case to the other).

7. after all these exchange and reinforcement are done, we note the interaction between i and j in the main sequence $\mathcal{S}$.

Each evolution step is then defined as a repetition of the $2 \to 7$ steps of the just outlined procedure.

6.2.4 Analytical results

Given the definition of the model we can analytically tackle the analysis of the asymptotic behavior of the system evolution. To this end, we focus on two observables of the system, i.e. the number of distinct IDs $D(t)$ in the events sequence $\mathcal{S}$ and the number of edges (multiplied by two) $A(t) = \sum_{i=1}^{D(t)} k_i(t)$, where k_i is the number of distinct IDs that contacted node i up to the evolution step t (i.e. the degree of node i at time t).

To write the time dependence of the two observables we must introduce some more quantities as to correctly take into account the different contribution to the system evolution. In particular we define $F(t) = \sum_{i=1}^{D(t)} k_{f_i}(t)$, i.e. the sum for each node i of the degree of its father f_i at time t. As suggested from numerical simulations, this sum grows with the same time-dependence as $A(t)$ but with a different multiplying constant so that $F(t) = fA(t)$. We also define $\tilde{p}$ to be the probability for an urn to be connected to its "father". Prompted by numerical simulations we will assume this probability to be constant in time and node-independent. Finally, we define $N(t)$ as the total number of balls in the system, i.e. $N(t) = \sum_{i=1}^{D(t)} n_i(t)$, where $n_i(t)$ is the number of balls in the i-th urn. Given the network definition we have:

$$N(t) = N_0 + 2\rho t + (\nu + 1)D(t) + (\nu + 1)A(t), \tag{6.2.8}$$

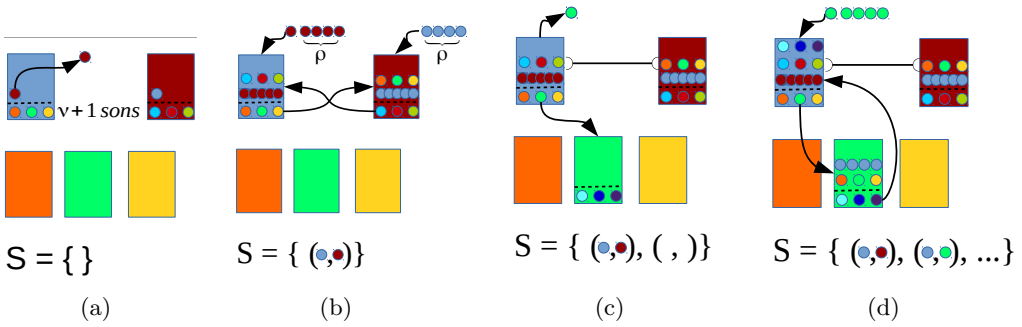


Figure 6.2: Two step of the evolution of the Polya's urn model generalized to N urns. (a) We start with an empty sequence $\mathcal{S}$ and two full urns (cyan and red) containing a copy of their respective ID and the $\nu + 1$ balls of their sons. The sons of the cyan urn appear in the second row and are initially empty urns. The extraction selects the cyan urn as the calling one. From this urn we withdraw the red ID. The first contact is then *(cyan, red)*. (b) We annotate this event in the contacts sequence $\mathcal{S}$ and put back ρ copies of the red ID into the cyan urn and vice-versa. As this is the first contact between these two urns we also exchange their $\nu + 1$ sons (for instance the cyan urns passes the orange, green, and yellow IDs to the red urn). (c) We then draw an edge connecting the cyan urn to the red one. Future contacts between these two urns will only result in their reinforcement. We then proceed with the second extraction. The cyan urn turns out to be the caller urn again, but this time it withdraws the ID of its green son. As the son's urn is empty, it creates its $\nu + 1$ sons elements together with their empty urns (that are not shown for simplicity). (d) The *(cyan, green)* event is recorded in the sequence and the two urns repeat the reinforcement and sons exchange steps.

where t is the number of evolution steps and $N_0 = 2 + 2(\nu + 1)$ is the initial number of balls in the system.

Given these definitions, we can write the master equation (ME) governing the evolution of the $D(t)$. In the long time limit the differential form of this ME reads:

$$\frac{dD(t)}{dt} = \frac{(\nu - \tilde{p})D(t) + (\nu + 1 - f)A(t)}{N(t)}, \tag{6.2.9}$$

where we took into account that the number of balls present in the system that are still unseen in the sequence $\mathcal{S}$ is the number of created sons $(\nu + 1)D(t)$ plus the number of exchanged (and thus duplicated) sons $(\nu + 1)A(t)$ minus the copies of the IDs already present in the sequence $\mathcal{S}$. These are the $D(t)$ copies of the IDs extracted in $\mathcal{S}$ and the missing balls exchanged between a son and a father $\tilde{p}D(t)$ plus the $F(t)$ copies of each ID in $\mathcal{S}$ around the system spread by the respective father.

In the same way we can evaluate the master equation governing the evolution of the degree of the single node $k_i(t, t_i)$, i.e. the degree at time t of an urn whose appearance time is t_i (i.e. the time of appearance of urn i in the system). To write the equation governing the $k_i(t, t_i)$evolution we have to account for the different contributions to the possibility for the urn i to contact (or get contacted by) a new ID in the network. In particular this probability is proportional to the number of "new" balls in the i-th urn $(\nu + 1)(1 + k_i)$, i.e. the sons created at the appearance time and the copies of the sons got from the $k_i(t, t_i)$ established ties. In addition we have to sum the $k_{f_i}(t, t_{f_i})$ copies of the i-th ID that its father spread around the network within the urns of its k_{f_i} neighbors. We then have to subtract the over-counted balls, in particular the $k_i(t, t_i)$ copies of i and of its sons that are no more "new" as for each established tie we burn an ID contained in the i-th urn. We also have to account for the possibility for the node to have contacted its father $\tilde{p}$ that results in the loss of one ball plus the $\nu/2$ balls of the brothers that are counted twice as once an ID i contacts one of its brothers j we burn two balls instead of one (the j copy in the i-th urn and vice-versa). Gathering all the outlined terms we get:

$$\frac{dk_i(t, t_i)}{dt} = \frac{(\nu/2 + 1 - \tilde{p}) + (\nu + f)k_i(t, t_i)}{N(t)}, \tag{6.2.10}$$

being t_i the entrance time of the i-th urn in the system. Considering the boundary condition $k_i(t = t_i, t_i) = 1$ and approximating $N(t) \simeq 2\rho t$ the solution of Eq. (6.2.10) reads:

$$k_i(t, t_i) = -\frac{(\nu/2 + 1 - \tilde{p})}{(\nu + f)} + \mathcal{C}t^{(\nu+f)/(2\rho)}, \tag{6.2.11}$$

where $\mathcal{C} = [1 + (\nu + 1 - \tilde{p})/(\nu + f)]/t_i^{(\nu+f)/(2\rho)}$, so that:

$$\begin{aligned} k_i(t, t_i) = -\frac{\nu + 1 - \tilde{p}}{\nu + f} + \left(1 + \frac{\nu + 1 - \tilde{p}}{\nu + f}\right)\left(\frac{t}{t_i}\right)^{(\nu+f)/(2\rho)} = \\ = -\mathcal{Q} + (1 + \mathcal{Q})\left(\frac{t}{t_i}\right)^{(\nu+f)/(2\rho)}, \end{aligned} \tag{6.2.12}$$

where we set $\mathcal{Q} = (\nu + 1 - \tilde{p})/(\nu + f)$. Now we can evaluate the $A(t)$ sum by substituting the just found functioning of $k_i(t, t_i)$:

$$A(t) = \sum_{i=1}^{D(T)} k_i(t, t_i) \simeq \int_0^t k(t, t') \frac{\partial D(t')}{\partial t'} dt', \tag{6.2.13}$$

where we dropped the i index from $k(t, t')$ and where we introduced the number of urns that entered the system at t' as the differential of the number of urns $D(t)$. Prompted by numerical simulation we set $D(t) = qt^\gamma$, where q is an unknown constant and γ the exponent leading the time evolution of the number of urns $D(t)$. By substituting Eq. (6.2.12) and the $D(t)$ form in Eq. (6.2.13) we find:

$$\begin{aligned} A(t) \simeq \gamma q \int_0^t k(t, t') t^{\gamma-1} dt' = \gamma q \int_0^t \left[-\mathcal{Q} + (1 + \mathcal{Q})\, t^{(\nu+f)/(2\rho)} \right] t^{\gamma-1} dt' = \\ = \left[-\mathcal{Q} + (1 + \mathcal{Q}) \frac{\gamma}{\gamma - \mathcal{M}} \right] D(t) = r(\gamma) D(t), \end{aligned} \tag{6.2.14}$$

where $\mathcal{M} = (\nu + f)/(2\rho)$. Eq. (6.2.14) tells us that $A(t)$ evolves in time with the same exponent of $D(t)$ and the two are bound by a proportionality constant $r(\gamma)$ so that $A(t) = r(\gamma)D(t)$. We can now solve the system by substituting Eq. (6.2.14) in Eq. (6.2.9) getting:

$$\frac{dD(t)}{dt} = \frac{(\nu - \tilde{p}) + (\nu + 1 - f) r(\gamma)}{2\rho t} D(t), \tag{6.2.15}$$

where, again, we approximated $N(t) \simeq 2\rho t$ as we expect $\gamma < 1$. By substituting $D(t) = qt^\gamma$ in Eq. (6.2.15) we get a second-order equation whose positive solution gives us the predicted value of γ that reads:

$$\gamma = \frac{\mathcal{B} + \sqrt{\mathcal{B}^2 + 4[2\nu(1 - f) + 2\tilde{p}f - \tilde{p} - f + 1]}}{4\rho} \xrightarrow{\rho,\nu\to\infty} \frac{3}{2}\frac{\nu}{\rho} = \frac{3}{2}\mathcal{R}^{-1}, \tag{6.2.16}$$

where $\mathcal{B} = 3\nu - \tilde{p} + \frac{\nu+1-f}{\nu+f}(f - \tilde{p})$ and where we introduced the ratio $\mathcal{R} = \rho/\nu$. Note that the solution of Eq. (6.2.16) holds in the $\mathcal{R} > 3/2$ region. For $\mathcal{R} \leq 3/2$ we cannot approximate $N(T) \simeq 2\rho t$ as the $D(t)$ and the $A(t)$ terms are now comparable to the linear term $2\rho t$. In this case one should solve the following set of coupled equations:

$$\begin{cases} \frac{dk(t,t')}{dt} = \frac{\nu+1-\tilde{p}+(\nu+f)k_i(t,t')}{2\rho t+(\nu+1)[A(t)+D(t)]} \\ A(t) = \int_0^t k(t, t') \frac{dD(t')}{dt'} dt' \\ \frac{dD(t)}{dt} = \frac{(\nu-\tilde{p})D(t)+(\nu+1-f)A(t)}{2\rho t+(\nu+1)[A(t)+D(t)]}. \end{cases} \tag{6.2.17}$$

We leave this task for future work.

We can now go back to Eq. (6.2.14) and substitute Eq. (6.2.16) therein to get the proportionality constant $r(\gamma)$ between $D(t)$ and $A(t)$ in the $\nu, \rho \to \infty$ limit:

$$r(\gamma) = \frac{\frac{3}{2}\mathcal{R}^{-1} + \frac{1}{2}\mathcal{R}^{-1}}{\frac{3}{2}\mathcal{R}^{-1} - \frac{1}{2}\mathcal{R}^{-1}} = 2. \tag{6.2.18}$$

Eq. (6.2.18) tells us that the proportionality constant does not depend (in the $\rho \gg 1$ limit) on the ratio $\mathcal{R}$.

As a last remark, let us note that this definition of the model generates a network with a strong hierarchical structure. Indeed, the urns created as sons can only be called for the first time from nodes belonging to a previous generation of the network. Moreover, nodes pass to newly contacted alters their fixed $\nu + 1$ sons, irrespectively of their urn composition. Such a structure and evolution rules could be relevant to model some kind of systems. Think for example to the academia and research world: nodes appearing first in the network are the professors or the establishment of the research in one field. Then, newly introduced nodes (the sons) corresponds to newly enrolled students. These students can get in contact with other nodes of the network only if they're put in contact by their supervisors (their fathers). Once they have been introduced outside of their circle, they can initiate to develop and accumulate a personal set of contacts (the collaborators). However, the initial set of $\nu+1$ sons is the only "legacy" that a node spreads around the network, i.e. regardless of its history and contact, the node's sons are the only novelties that the node will spread to others.

There are however systems in which we expect a more flexible and history-dependent spread of neighbors and recommendations throughout the network. Indeed, we can think of the sons passage to a newly contacted alter as a recommendation that a node is giving to the others based on its experience. Think for example of a network of friendship. Based on the individuals' history, one can stick with the first friends he made in time (the sons). There could be cases where a friend found later in time can literally emerge and overcome the already established friendships by repeated contacts and interactions with a node. This reflects in the fact that when we enter in contact with new pals we transfer to them a sample of our current social circle and not a snapshot of our initial ego-net. In other words, a new friend we meet is more likely to get in contact with our most frequented friends rather than our historically first ones.

That is why the model has to be generalized so as to let nodes spread a sample of their urns instead of their sons. This mechanisms allows both to break the rigid hierarchical structure of the network and to mimic in a more realistic way the social interactions. However, we leave this task for future work and here we will only discuss the previous version of the model.

6.2.5 Numerical results

The first quantity we measure is the γ exponent leading the growth of both the $D(t)$ and $A(t)$ quantities. As we show in Fig. 6.3, the exponents measured in numerical simulations are found in very good agreement with the analytical prediction for various values of $\mathcal{R}$.

Let us note that, as expected, the analytical predictions applies to numerical data for $\mathcal{R} > 3/2$. Moreover, we observe that the probability $\tilde{p}$ for a node to be connected to its father, the ratio $f = F(t)/A(t)$ and the proportionality constant

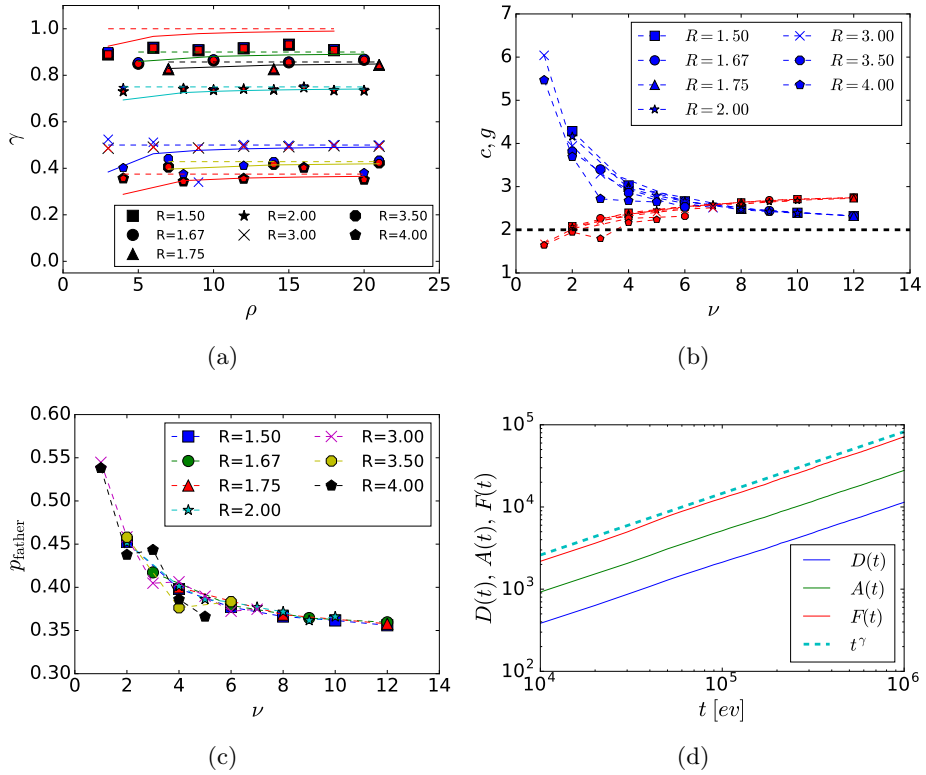


Figure 6.3: (a) The exponents γ as a function of ρ for different values of the ratio $\mathcal{R}$ (see legend). We show the exponent γ from the fit of $D(t)$ (black symbols), $A(t)$ (blue symbols) and $F(t)$ (red symbols). For every ratio analyzed we plot the prediction of Eq. (6.2.16) (solid lines) together with the asymptotic values of γ (dashed lines). (b) The constants of proportionality $r(\gamma)$ (blue symbols) and f (linking $F(t) = f \cdot A(t)$) (red symbols) for different values of ratio $\mathcal{R}$ as a function of ν. We also show the $r(\gamma) = 2$ asymptotic limit (black dashed line). (c) The $\tilde{p}$ probability for an urn to be attached to its father as a function of ν for different ratios $\mathcal{R}$. As in the previous case the curves seems to collapse in a single behavior (or they weakly depends on $\mathcal{R}$). (d) The numerical $D(t)$, $A(t)$ and $F(t)$ curves (solid lines) and the predicted time dependence t^{γ} (cyan dashed line). The simulations refers to a system with $\rho = 16$, $\nu = 8$ (hence $\mathcal{R} = 2$) and an evolution time of 10^6 steps.

$r(\gamma)$ all seem to weakly depend on the ratio $\mathcal{R}$, as they collapse on a single behavior when plotted against ν instead of ρ as we show in Fig. 6.3.

6.2.6 Making contact with the TVN framework

So far, we focused on the asymptotic growth of the size $D(t)$ of the sequence $\mathcal{S}$ and on the (double) number of links $A(t)$ found in the network.

Though the results are interesting and resembles some of the asymptotic properties of real-world time-varying networks, we are just scratching the surface. Remarkably, we can go deeper and show that the contact sequences generated by such a model can be put in closer correspondence with empirical datasets. Indeed, the model here introduced basically generates a sequence of contacts amongst nodes in a network. So what if we pretend that these data are an empirical dataset and we analyse it with the tools and procedures outlined in Chapter 4? Remarkably many of the previously measured real-world networks properties are found together with other features that were completely missed by the previous activity driven model.

As a first insight, we show in Fig. 6.4 (a,b) the analysis of the correlation between the measured activity a_i of a node and its entrance time t_i (i.e. the time, in events, at which the node appears for the first time in the temporal sequence).

As one can see, the activity of a node correlates with the inverse of its entrance time in the numerical simulations. The situation is of course more heterogeneous in the TMN datasets even though we can observe a weak diagonal pattern even here. Moreover, as shown in Fig. 6.4 (d), the size of the empirical sequence $D(t)$ grows following a Heaps' law (i.e. $D(t) \simeq t^{\gamma}$ with $\gamma < 1$), as we already found in Fig. 6.3 (d).

Another relevant results is that we recover an heterogeneous activity distribution as we show in Fig. 6.5.

The activity distribution is strikingly broad, and the power-law behavior is found for many orders of magnitude.

The more surprising results is maybe the one found measuring the attachment probability $p(k)$ from numerical data. As we show in Figg. 6.6 and 6.7, both the empirical behavior and the functional form of the $p(k)$ reinforcement function are precisely reproduced by the synthetic data.

We show that a single value of $\beta = \beta_{\text{opt}}$ catches the functional form of the $p(k)$ function for most of the nodes' bins. To corroborate this results, in Fig. 6.7 we show the rescaled $p(k \to k/c_b)^{1/\beta}$ curves for different values of ρ and ratio $\mathcal{R}$. As we found in the empirical case, all the curves collapse on the reference curve $p(x) = 1/(1+x)$, so that the empirical behavior is reproduced with high accuracy.

As an additional check, we measure the average degree growth $\langle k(a,t) \rangle$ with time and compare the growth with the predicted functioning $t^{1/(1+\beta)}$.

We present the results in Fig. 6.8, where we show a nice agreement between synthetic data and analytical predictions. Moreover, we also recover a Gaussian-like distribution of the probability distribution $P(k,t)$ as we show in Fig. 6.9.

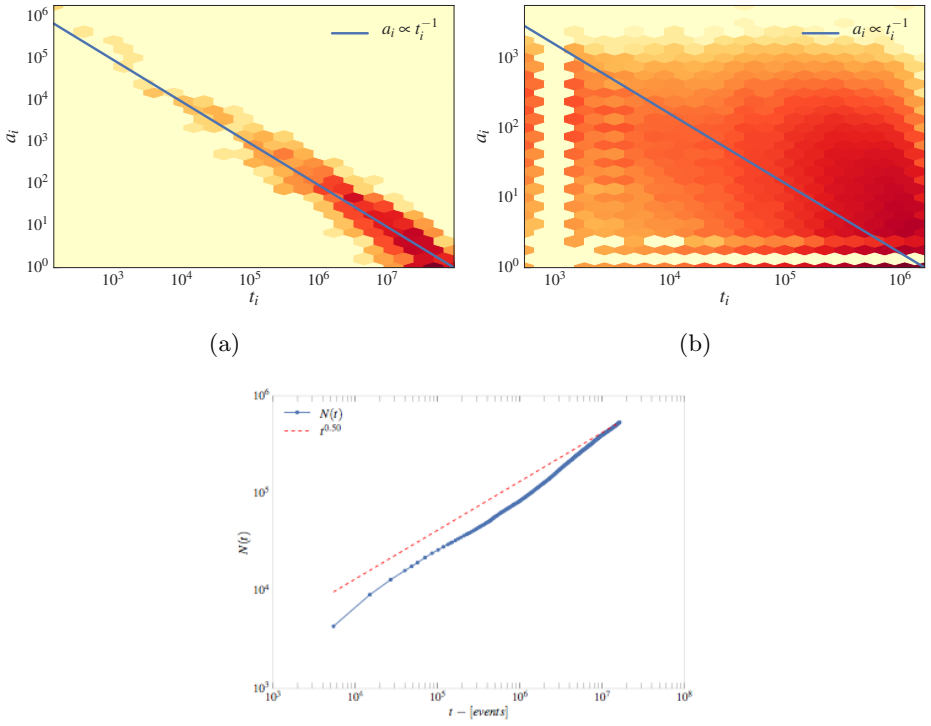


Figure 6.4: The activity a_i vs the entrance time t_i for (a) numerical simulation featuring $\rho = 3$, $\nu = 1$ and (b) TMN. The color code is logarithmically scaled (the redder the higher count for a particular bin). We also show the reference line $a_i \simeq 1/t_i$ (blue solid line). (c) The number of distinct nodes in the TMN sequence as a function of events number (blue symbols). We also show a power-law approximation of the temporal growth $D(t) \sim t^{0.5}$ (red dashed line).

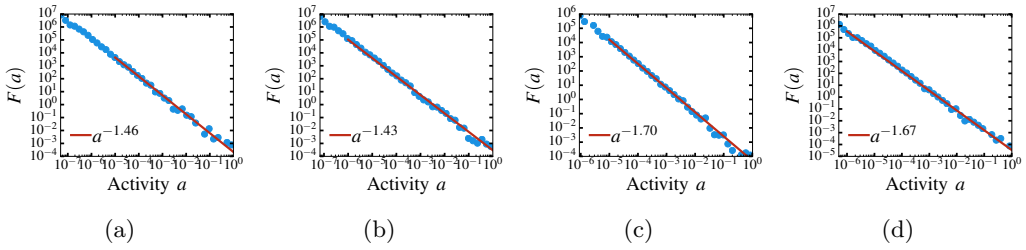
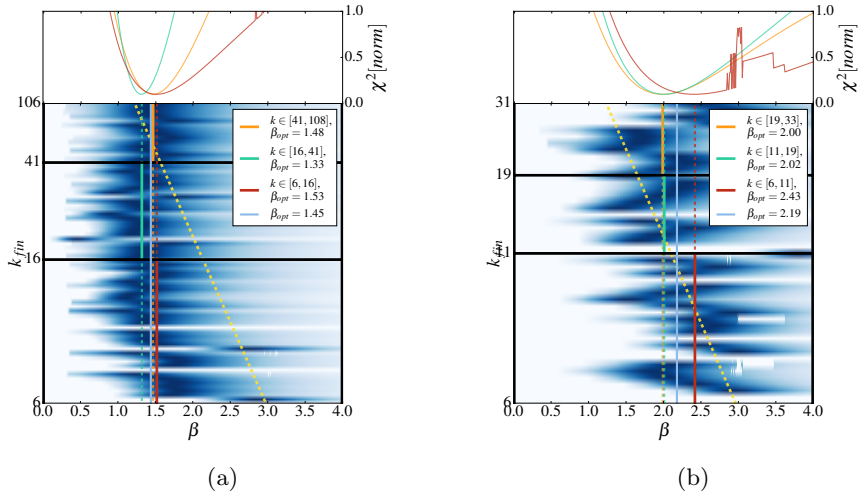
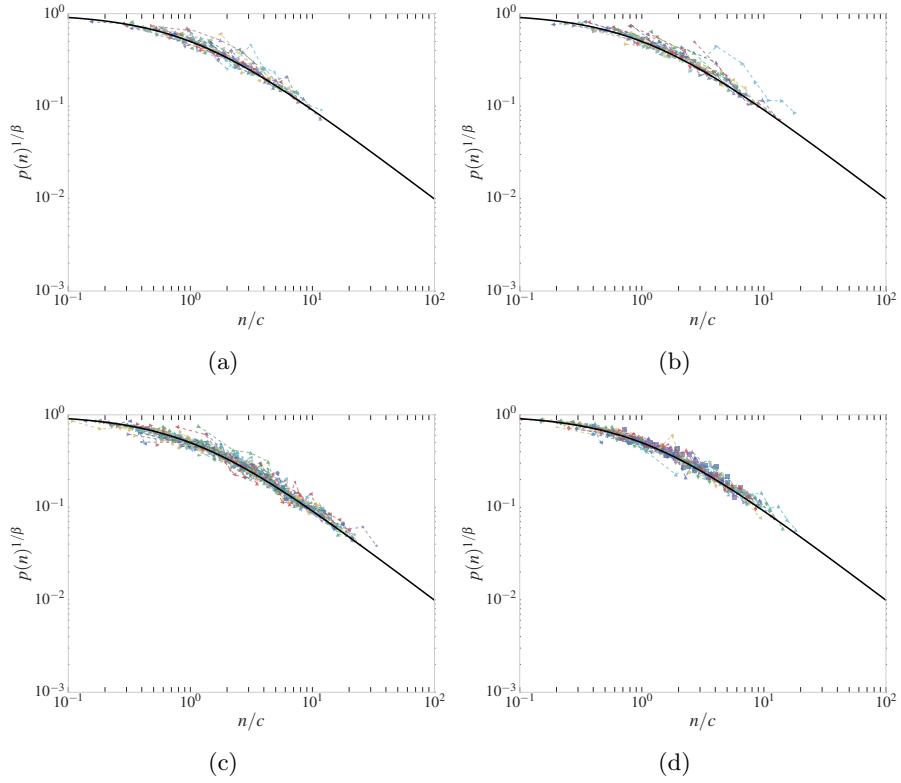


Figure 6.5: The activity distribution $F(a)$ (blue circles) as measured in numerical simulations featuring (a) $\rho = 9$, $\nu = 3$, (b) $\rho = 18$,$\nu = 6$, (c) $\rho = 10$, $\nu = 5$, and (d) $\rho = 20$, $\nu = 10$. We also show the power-law fit (Zipf's law) $F(a) \propto a^{-\nu}$ reporting the resulting exponent in the legends.



(a) (b)

Figure 6.6: The heat-map like plot of the $\chi^2(\beta)$ for numerical simulations featuring (a) $\rho = 12$, $\nu = 8$ and (b) $\rho = 20$, $\nu = 10$. The optimal β is shown (vertical cyan line) together with the optimal value for different sets of nodes bins. The β_{opt} values read (a) 1.4 and (b) 2.2.



(a) (b)

(c) (d)

Figure 6.7: The rescaled attachment probability $p(x = k/c)^{1/\beta}$ (symbols) as measured for different activity classes in numerical simulations featuring (a) $\rho = 9$, $\nu = 3$, (b) $\rho = 18$,$\nu = 6$, (c) $\rho = 10$, $\nu = 5$, and (d) $\rho = 20$, $\nu = 10$. We also show the reference curve $p(x) = (1 + x)^{-1}$ (black solid line).

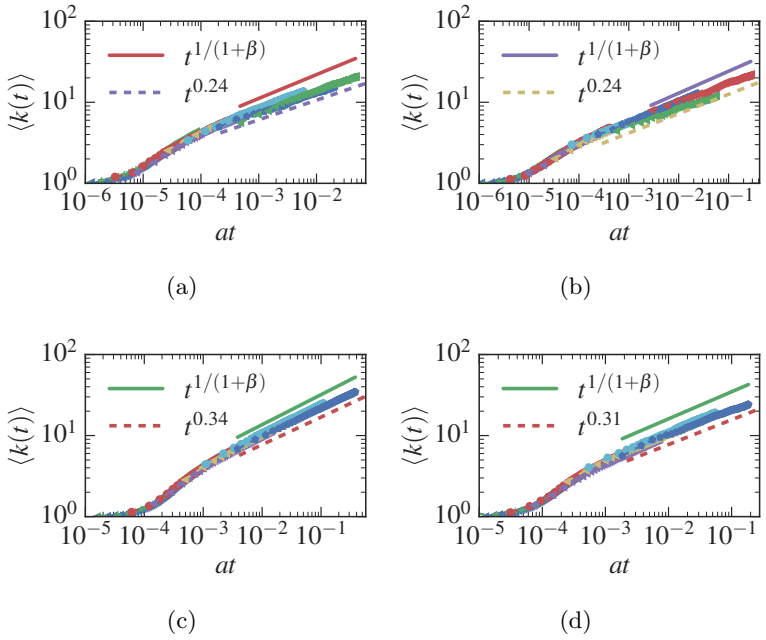


Figure 6.8: The average degree growth $\langle k(a,t)\rangle$ (symbols) as measured for different activity classes in numerical simulations featuring (a) $\rho = 9$, $\nu = 3$, $\beta = 2.4$, (b) $\rho = 18$,$\nu = 6$, $\beta = 2.6$, (c) $\rho = 10$, $\nu = 5$, $\beta = 1.7$, and (d) $\rho = 20$, $\nu = 10$, $\beta = 2.1$. We also show the analytical prediction $\langle k(a,t)\rangle \propto t^{1/(1+\beta)}$ (red solid line) together with the fit t^{η} (dashed line, see the legends for the actual values). In all the case we find a good agreement between the two results.

In Fig. 6.10 we also show that the measured activity distribution exponent ν and reinforcement strength β correctly predict the μ exponent of the degree distribution $\rho(k) \propto k^{-\mu}$.

Remarkably, even in this case we find a very good agreement between synthetic data and analytical predictions. As for the activity distribution, also the degree $\rho(k)$ one is found to be heterogeneous and to span with a power-law like functioning different orders of magnitude.

As a final measure, we checked the topological differences between the networks generated with the urns model, the activity driven ones and the empirical ones. Specifically, we look at the average clustering coefficient $\langle C\rangle$ and the Pearson degree assortativity r as defined in Eqq. (1.1.6) and (1.2.7), respectively.

As we show in Table (6.1), the urns model generates networks with an higher average clustering coefficient $\langle C\rangle \simeq 0.25$ with respect to the activity driven model (in which we get networks with $\langle C\rangle \simeq 10^{-5}$). Moreover, we also find these networks to feature a positive degree assortativity r, thus making more contact with some of the real-world networks under considerations, specifically the ones belonging to the APS datasets.

In this final chapter we introduced, developed and characterized a network dy-

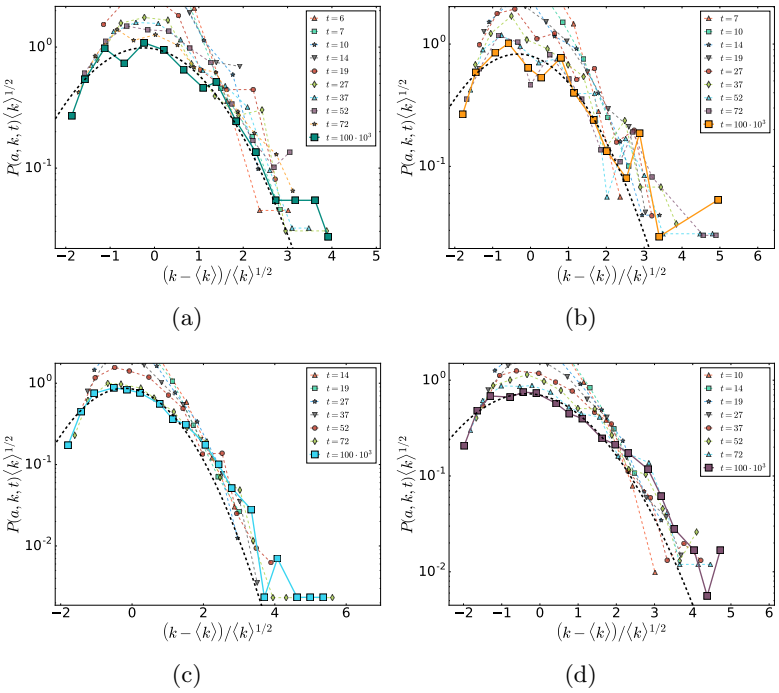


Figure 6.9: The rescaled probability distribution $P(k,t)$ (symbols) as measured at different times for a single activity class in numerical simulations featuring (a) $\rho = 9$, $\nu = 3$, (b) $\rho = 18$,$\nu = 6$, (c) $\rho = 10$, $\nu = 5$, and (d) $\rho = 20$, $\nu = 10$. We also show the reference Gaussian form for the long-time asymptotic limit (black dashed line).

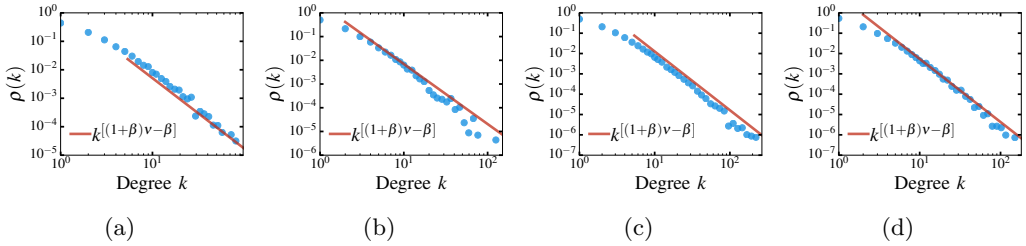


Figure 6.10: The degree distribution $\rho(k)$ (blue circles) as measured in numerical simulations featuring (a) $\rho = 9$, $\nu = 3$, (b) $\rho = 18$,$\nu = 6$, (c) $\rho = 10$, $\nu = 5$, and (d) $\rho = 20$, $\nu = 10$. We also show the analytical prediction $\rho(k) \propto k^{-[(1+\beta)\nu-\beta]}$ fitting the right tail of the distribution (red solid line).

Dataset	$\langle C \rangle$	r
AD $\beta = 1$, $\nu = 2.1$	$3 \cdot 10^{-5}$	-0.06
Urns $\rho = 3$, $\nu = 1$	0.27	0.19
TMN	0.16	-0.03
PRB	0.63	0.07
PRL	0.66	0.08

Table 6.1: The average clustering coefficient $\langle C \rangle$ and the Pearson degree assortativity coefficient r as found in: (i) numerical simulations on an activity driven model with reinforcement featuring $\beta = 1.0$ and $\nu = 2.1$, (ii) urns model featuring $\rho = 3$ and $\nu = 1$, (iii) the Twitter mention network, (iv) the Phys. Rev. B., and (v) The Phys. Rev. Letters.

namical model based on the adjacent possible theory. Though we only presented preliminary measures of the most relevant topological and dynamical properties of the network, results are encouraging and calls for a deeper analysis of the generated networks.

In particular, one of the issue is to map the time of the synthetic model with the real-world time. Indeed, the network is changing in size and the first urns appearing are the one more likely to become the hubs and thus to dominate the overall dynamics. How can we map this transient to a real-world model?

Moreover, the hierarchical structure of the network generates a positive degree assortativity that is found only in some of the analyzed datasets. As already stated, modifications of the model so as to allow for a more "flexible" network topology must be taken into account. To this end, the development of a model in which the novelties spread around the network are sampled by the urns in a dynamical way, rather than in a static one with the sons of an urn being the only exchanged "novelties", has to be considered.

Lastly, a rigorous temporal characterization of the model is also needed to measure the generated burstiness. Indeed, even though we did not discussed about it, there are indications that the waiting time and inter-event time distribution are heavy-tailed as in real-world systems. However, this feature could be due to the serial nature of the model alone (that evolves one event at a time), so that further investigations are necessary in this direction.

We will tackle these problems in future works. By now, let us say that this approach seems really promising as it naturally introduces several complex features without the need to insert them by hand in the model. Moreover, the simplified single-version of the model has already been proved to correctly model the adjacent-possible exploration of single-agents so that it is reasonable to expect it to apply even to social interactions.

Conclusion

In this work we presented the basic ideas and concepts of complex networks as well as the modeling framework of time-varying-networks.

Within this general approach, one can analyze and characterize different processes and problems: the heterogeneous human activation patterns, the correlations and strategies of humans' social exploration and the asymptotic properties of the social networks under consideration.

In the first part of this thesis we introduced the relevant measures and quantities found in network science. We presented some generative models focused on the static representation of social networks and how the main features of real-world networks can be replicated from a numerical and modeling point of view.

In the second part we then presented evidences calling for the introduction of a temporal dimension in networks, thus introducing the modeling framework of time-varying-network, focusing on the activity driven modeling scheme. We showed how the diverse temporal activation and contact pattern affects the dynamical processes and how the instantaneous and cumulative topologies differ.

We then focused on the modeling of correlations found in human dynamics and social activity. Specifically, we presented a thorough measure and characterization of the reinforcement process of social links, pointing out how people tend to interact through already established ties rather than continuously attach to new individuals in the network. Specifically, we find this mechanism to feature the same functional form in very diverse layer of human interactions (mobile phone calls, Twitter mentions and scientific collaborations) and to be characterized by two parameters alone: the reinforcement constant c and the reinforcement strength β.

We then framed these empirical findings in a simple stochastic model of network evolution, and we derive a general asymptotic theory of the network dynamic, together with the general scaling laws for the time-dependence of the node's degree distribution and average connectivity value. We compared the analytical predictions against both numerical simulations and seven real world datasets finding a striking agreement between the two. Moreover, we also studied how correlations on social capital investment affect dynamical processes.

Then, we let into the model the bursty activation pattern typically found in social networks. To this end we encoded burstiness on edges activation by its simplest

formulation, i.e. by imposing a power-law distributed inter-event time distribution. We analytically solved the asymptotic evolution of this model, finding a non trivial phase diagram of the network evolution due to the interplay of burstiness and reinforcement process. Interestingly, if reinforcement is sufficiently strong, burstiness can be sub-leading even in the presence of large inter-event time fluctuations. These results have then been tested against numerical simulations and we also checked their consistence with an empirical dataset featuring burstiness (the twitter mentions network) finding a good agreement between the two.

The introduced model represents a relevant step forward in the modeling framework of activity driven networks. In particular, the model reproduces with striking accuracy the asymptotic properties of all the analyzed real-world networks. Moreover it also provides analytical predictions accounting for both the reinforcement process and burstiness that are found to be in very good agreement with empirical data.

The importance of such a characterization goes beyond the analytical point of view. Indeed, the generated contact sequences may be applied as a synthetic testbeds to forecast the outcome of dynamical processes such as epidemic processes or information spreading. Furthermore, the developed analytical framework allows for a simple implementation of different functional forms of the activity distribution, the inter-event time and the reinforcement process. It is also open to the insertion of additional mechanism in the system dynamical evolution. This is an additional strength of the adopted modeling framework. Indeed, we expect the continuously growing data availability to allow for a precise measurement of new and more complex mechanisms shaping the network evolution. These mechanisms will eventually have to be taken into account in a modeling framework, and our one, thanks to its ME formalism, provides an handy way to accomplish this task.

Though its consistent analytical predictions, the activity driven model comes with some shortcomings. For instance, it does not capture the modular structure of social networks and requires the system to be constant in size. Furthermore, the heterogeneous behavior of individuals has to be put by hand within the modeling framework (usually setting a power-law distributed activity potential) and the reinforcement mechanism's functional form cannot be derived from first principles. The latter is encoded in the model with an imposed functional form as measured from real-world data. In the last part of this thesis we then asked ourselves whether or not it is possible to develop a more basic model, able to reproduce all the so far analyzed properties of empirical networks.

We tried to answer to this last question by taking advantage of the adjacent possible theory. This theoretical framework has already been successfully applied to the modeling of how individuals explore knowledge, ideas or an artistic field. In all these layers of knowledge explorations, our curiosity brings us to experience novelties and, once we encounter one novelty, we are prompted to others as our adjacent possible (i.e. our horizon) expands. One can then reasonably expect that a similar mechanism also applies to social exploration. This consideration let us

tackle the modeling task from a novel point of view: social interactions are seen as novelties occurring in our lives and thus correlates in time as one novelty paves the way for the following one.

This mechanism is encoded in a model based on a generalization of the Polya's urn with triggering effect. In particular, each social interaction triggers two different mechanisms, depending on the novelty introduced with such contact: a reinforcement process (a rich-get-richer mechanism) and an insertion of novelties (expansion of the adjacent possible).

The model generates a sequence of contacts between the nodes of the network and naturally introduces some of the missing features that the activity driven approach could not provide: the reinforcement mechanism $p(k)$ is naturally reproduced (both in its functional form and in the variety of social strategies), the network size grows in time following a Heaps' law and the heterogeneous activity and degree distribution are automatically generated by the reinforcement mechanism. Moreover, both the hierarchical structure and an higher clustering coefficient are reproduced without the need of additional mechanisms or evolving rules.

The modeling framework introduced in this work paves the way to a deeper understanding of the emergence and evolution of social ties. The agreement between the analytical predictions and observed behaviors in seven real datasets, describing different types of social interactions, are encouraging steps in this direction. Moreover, our results are a starting point for the development of predictive tools able to forecast the growth and evolution of social systems, based not just on regression models, but on a more rigorous analysis of the ego-network's dynamics.

Finally, the novel approach based on the adjacent-possible theory opens the fascinating possibility to quantify the human propensity to social interactions and network explorations from first principles. We also think that leveraging this modeling perspective may give us new and appealing insight on how individuals collaborate so as to improve the overall network's performance or to develop innovative ideas by being continuously pollinated by new contacts and, thus, novelties.

Appendix A

Appendix

A.1 Analytical solution of the activity driven model with communities structure

We now present in a more detailed and comprehensive way the definition of the model, the key properties of the latter and the results (both analytical and numerical) found.

A.1.1 The Model

The network is defined by means of the following parameters:

- the total number N of nodes in the network;
- the activity distribution parameters, i.e. the lower cut-off ϵ and the leading exponent ν, so that $F(a) = a^{-\nu}$ for $a \in [\epsilon, 1]$;
- the lower cut-off $s_{\min}$, the upper limit $s_{\max}$ and the exponent ω describing the community size s distribution, i.e. $P(s) = s^{-\omega}$ for $s \in [s_{\min}, s_{\max}]$;
- μ, i.e. the probability that, once active, a node calls inside the community, so that $(1 - \mu)$ is the probability to call outside the community;

We initialize the network extracting N activity values from a given activity distribution $F(a)$ and then grouping the nodes in communities of size s drawn from a size-distribution $C(s)$. The latter spans the $s \in [100, \sqrt{N}]$ interval. Once we initialized the network we let it evolve following the time-varying activity driven framework. At each time step t we start with N disconnected nodes. Each node gets active with probability $a_i dt$ at each time step dt and call a randomly chosen node inside (outside) its own community with probability μ $(1 - \mu)$. At time $t + 1$ we delete all the edges and repeat the above procedure. Each node i will then have a set of neighbors that have been contacted or have contacted the node during the network growth. The size of such a set is the integrated degree $k_T(i)$ of the node i. Of these k_T neighbors, k_c will be inside the i's community (community degree) and $k_o = k_T - k_i$ will be external to the community.

A.1.2 The Network growth

In order to easily describe the network growth it is worth noting that, since the model is memoryless, the in-community degree k_c and the out-community degree k_o are decoupled and can, in fact, be treated separately recalling that $k = k_c + k_o$.

Even the activity potential a_i of each node can be "split" in two components: the in-community activity μa_i and the complementary out-community $(1-\mu)a_i$. Indeed, each node will point a fraction μ of its own events toward the community, while the remaining $1-\mu$ will be directed outside the community itself. Each node will then experience a mean field of activity $\mu\langle a\rangle$ coming from the community (provided that the community is large enough) and a supplementary external field $(1-\mu)\langle a\rangle$ coming from the rest of the network.

The network time scales

As a first insight, let us note that the in-degree time dependence can be easily approximated with a probabilistic consideration. Each node i of activity a_i within a community of size s has $s-1$ available edges. Now, for each time step of the dynamics, the edge e_{ij} will be created with probability $\mu(a_i/(s-1) + a_j/(s-1))$; then, on average, each edge emanating from i will be activated with probability $c(a_i,\mu)/(s-1) = \mu(a_i + \langle a\rangle)/(s-1)$. The probability $P(a_i,s,t)$ for an edge pointing to i not to be activated after t time steps then reads:

$$P(\mu a_i, s, t) = \left(1 - \frac{c(\mu a_i)}{(s-1)}\right)^t . \tag{A.1.1}$$

Since the in-degree $k_c \in [0, s-1]$, we can write the $k_c(\mu a_i, s, t)$ as:

$$k_c(\mu a_i, s, t) = (s-1)\left(1 - P(\mu a_i, s, t)\right). \tag{A.1.2}$$

We also note that Eq. (A.1.1) gives us an estimation of the characteristic time $\tau(\mu a_i, s)$ that takes for a node of activity a_i to saturate the in-degree $k_c \to (s-1)$. Indeed, we can rewrite Eq. (A.1.1) as:

$$\begin{aligned} P(\mu a_i, s, t) = \exp\left[t \ln\left(1 - \frac{c(\mu a_i)}{s-1}\right)\right] = \exp\left(-\frac{t}{\tau(\mu a_i, s)}\right) \quad \Rightarrow \\ \tau(\mu a_i, s) = -\left[\ln\left(1 - \frac{c(\mu a_i)}{s-1}\right)\right]^{-1} . \end{aligned} \tag{A.1.3}$$

So, as expected, the saturation time (i.e. the typical time for k_c to be of the same order of s) decreases as the activity μa_i and/or the community size s grows.

Generalizing the above reasoning, the characteristic time for a community to have the majority of the nodes saturated is obtained by evaluating the probability $P_e(\mu, s)$ for an edge e_{ij} in a community of size s to be created (on average) in a single evolution step, that is the number of edges activated in one step divided by the total number of edges:

$$P_e(\mu, s) = \frac{2s\mu\langle a\rangle}{s(s-1)}. \tag{A.1.4}$$

The probability for one edge not to be created after t time steps is then:

$$\begin{aligned} P_e(\mu, s, t) &= \left(1 - \frac{2\mu\langle a\rangle}{s-1}\right)^t \quad \Rightarrow \\ P_e(\mu, s, t) &= \exp\left[-\frac{t}{\tau_c(\mu, s)}\right] \quad \Rightarrow \\ \tau_c(\mu, s) &= -\left(\ln\left(1 - \frac{2\mu\langle a\rangle}{s-1}\right)\right)^{-1}, \end{aligned} \tag{A.1.5}$$

where $\tau_c(\mu, s)$ represents the typical time by which the majority of the nodes of a community has a degree $k_c \sim s$.

Note that, in the evaluation of both $\tau(\mu a_i, s)$ and $\tau_c(\mu, s)$, we did not take into account the difference between edges pointing to an active node and the ones pointing to a less active one. Nevertheless, this is a simple estimation that, as we will show later, correctly catches the general behavior of the in-degree k_c for any value of μ, a_i and s. Besides, when computing the key features of the evolving network, we are now able to distinguish two different regimes: the short time range $t \ll \tau(\mu a_i, s)$ at which $k_c(t) \ll s$ and, on the contrary, the long time limit $t \gtrsim \tau_c(\mu, s)$ where $k_c(t) \sim s$ for any activity value a_i.

The Master Equation and the $P(a, k, t)$

We can now write down the Master Equation (ME) for the quantities $P_c(a_i, s, k_c, t)$ and $P_o(a_i, s, k_o, t)$, that is, the probability for a node of activity a_i belonging to a community of size s to have degree in (out) degree k_c (k_o) at time t.

For the in-degree k_c distribution we get the following equation:

$$\begin{aligned} P_c(a_i, s, k_c, t+1) = P_c(a_i, s, k_c, t)\left[\mu a_i \frac{k_c}{s} + \sum_{j\sim i}\mu a_j + \frac{1}{s}\sum_{j\nsim i}\mu a_j\left(\frac{1-s}{s}\right)\right] \\ P_c(a_i, s, k_c - 1, t)\left[\mu a_i\left(1 - \frac{k_c}{s}\right) + \frac{1}{s}\sum_{j\nsim i}\mu a_j\right] + P_c(a_i, s, k_c, t)\left[1 - \sum_j \mu a_j\right] \end{aligned}$$

where $\sum_{j\nsim i}$ and $\sum_{j\sim i}$ are the sum over the nodes not connected with the i-th node and the sum over the neighbors of i, respectively.

After some algebra we obtain the equation:

$$\begin{aligned} P_c(a_i, s, k_c, t+1) - P_c(a_i, s, k_c, t) = \\ -\left[P_c(a_i, s, k_c, t) - P_c(a_i, s, k_c - 1, t)\right]\left(\mu a_i - \mu a_i \frac{k_c}{s} + \frac{1}{s}\sum_{j\nsim i}\mu a_j\right). \end{aligned} \tag{A.1.6}$$

By passing to the continuum limit we eventually get the differential equation:

$$\begin{aligned} \frac{\partial P_c(a, s, k_c, t)}{\partial t} = \\ -\left(\mu a_i - \mu a_i \frac{k_c}{s} + \frac{1}{s}\sum_{j\nsim i}\mu a_j\right)\left[\frac{\partial P_c(a, s, k_c, t)}{\partial k} - \frac{1}{2}\frac{\partial^2 P_c(a, s, k_c, t)}{\partial k^2}\right], \end{aligned} \tag{A.1.7}$$

where we dropped the a_i index since we expect all the nodes of a given activity to behave in the same way. The first term in the *r.h.s.* of Eq. (A.1.7) has two different values based on the in-degree k_c of the i-th node: for $t \ll \tau(\mu a_i, s)$ we have $k_c \ll s$ and the sum $1/s \sum_{j \nsim i}$ can be approximated with $\mu\langle a \rangle$, while we can neglect the k_c/s term. On the other hand, as time passes, we converge to the $k_c \to s$ limit ($t \gtrsim \tau(\mu a_i, s)$): we can then neglect the term proportional to $1/s \sum_{j \nsim i}$ and consider only the $\mu a_i k_c/s$ one. As expected, in the $t \gtrsim \tau \mu a_i, s$ limit, the $P_c(a, s, k_c, t)$ time derivative goes to zero, as it will tend to $\delta(k_c - (s-1))$.

Let us focus on the $t \ll \tau(\mu a_i, s)$ limit. When discarding the k_c/s term, a solution of the above equation is:

$$P_c(a, s, k_c, t) = C \exp\left[-\frac{(k_c - \mu(a + \langle a \rangle)t)^2}{2t\mu(a + \langle a \rangle)}\right], \tag{A.1.8}$$

where C is a normalization constant.

For the out-degree k_o, by following the same procedure, we recover the same results with μ substituted by $(1 - \mu)$:

$$P_o(a, s, k_o, t) = C \exp\left[-\frac{(k_o - (1-\mu)(a + \langle a \rangle)t)^2}{2t(1-\mu)(a + \langle a \rangle)}\right]. \tag{A.1.9}$$

Since $N \gg s_{\max}$, the out-degree $k_o \ll N$ for any time t of the process, thus we assume that Eq. (A.1.9) is valid for all the time scales analyzed. Also note that, as expected, the net effect of the parameter μ is just a time rescaling of the in-community and out-community activity, respectively.

Then, discarding the $t \sim \tau_c(\mu, s)$ time range, in the $t \gtrsim \tau_c(\mu, s)$ time limit the $P_c(a, s, k_c, t)$ converges to the $\delta(k_c - (s-1))$ distribution. In fact, all the nodes will have all their edges activated.

Let us now resume the results found in this section:

$$P_c(a, s, k_c, t) \propto \begin{cases} \exp\left[-\dfrac{(k_c - \mu(a + \langle a \rangle)t)^2}{2\mu(a + \langle a \rangle)t}\right] & \text{for } t \ll \tau(\mu a, s) \text{ or } (k_c \ll s) \\ \\ \delta(k_c - (s-1)) & \text{for } t \gg \tau(\mu a, s), \end{cases} \tag{A.1.10a}$$

(A.1.10b)

$$P_o(a, s, k_o, t) \propto \exp\left[-\frac{(k_o - (1-\mu)(a + \langle a \rangle)t)^2}{2(1-\mu)(a + \langle a \rangle)t}\right] \qquad \text{for } \forall t, \tag{A.1.10c}$$

where we now distinguish between the in-community degree distribution $P_c(a, s, k_c, t)$ and the out degree distribution $P_o(a, s, k_o, t)$. The latter however, is independent on the community size and we can then define $P_o(a, k_o, t) = P_o(a, s, k_o, t)$.

So far we treated the two probability functions separately when, in fact, k_c and k_o are bound by the relation $k = k_c + k_o$. The total degree distribution $P(a, s, k, t)$ will

then be determined by the convolution of both the $P_c(a,s,k_c,t)$ and $P_o(a,k-k_c,t)$ so that:

$$P(a,s,k,t) = \int_0^k dk_I P_c(a,s,k_I,t) P_o(a,k-k_I,t), \tag{A.1.11}$$

where we integrate over all the possible arrangements of the k edges.

In the $t \ll \tau(\mu a_i, s)$ limit, by substituting Eq. (A.1.10a) and (A.1.10d) in Eq. (A.1.11) we sum the two exponents getting:

$$P(a,s,k,t) = C \int_0^k dk_I \exp\left[-\frac{(k_I - \mu(a+\langle a\rangle)t)^2}{2\mu(a+\langle a\rangle)t} - \frac{((k-k_I) - (1-\mu)(a+\langle a\rangle)t)^2}{2(1-\mu)(a+\langle a\rangle)t} \right], \tag{A.1.12}$$

where C is, again, a normalization constant.

By combining the two terms and after some algebra we get:

$$P(a,s,k,t) = C \int_0^k dk_I \exp\left[-\frac{(k_I - \mu k)^2}{2\mu(1-\mu)(a+\langle a\rangle)t} + \frac{k^2 - 2k(a+\langle a\rangle)t + (a+\langle a\rangle)^2 t^t}{2(a+\langle a\rangle)t} \right]. \tag{A.1.13}$$

The integration over k_I gives:

$$\begin{aligned} P(a,s,k,t) \;=\; & C\left[\mathrm{Erf}\left(\frac{(1-\mu)k}{\sqrt{2\mu(1-\mu)(a+\langle a\rangle)}} \right) - \mathrm{Erf}\left(-\frac{\mu k}{\sqrt{2\mu(1-\mu)(a+\langle a\rangle)}} \right) \right] \\ & \exp\left[-\frac{k^2 - 2k(a+\langle a\rangle)t + (a+\langle a\rangle)^2 t^t}{2(a+\langle a\rangle)t} \right], \end{aligned} \tag{A.1.14}$$

where $\mathrm{Erf}\, x$ is the error function evaluated at x. Now, let us recall that we are working in the $t \ll \tau(\mu a_i, s)$ limit. If we want to evaluate the $P(a,k,t) = \int_{s_{\min}}^{s_{\max}} ds P(s) P(a,s,k,t)$, we then have to work in the $t \lesssim min_s(\tau_c(\mu,s))$ so that in every community we can use Eq. (A.1.14) as the true value of the $P(a,s,k,t)$. The integration over the different community size s is then straightforward since the terms are independent on s giving $P(a,k,t) = P(a,s,k,t)$ as found in Eq. (A.1.14). Note that this result holds for any value of μ, N and a.

The computation of $P(a,k,t)$ in the $t \gtrsim \langle \tau_c(\mu,s)\rangle_s$ limit is more complicated and we have to assume that $k_c = s-1$ for each node in a community of size s. The $P_o(a,s,k_o,t)$ will still be approximated by Eq. (A.1.10d). The starting point is:

$$\begin{aligned} P(a,k,t) = \int_{s_{\min}}^{s_{\max}} ds P(s) \int_0^k dk_I \delta(k_I - (s-1)) P_o(a,k-k_I,t) = \\ \int_{s_{\min}}^{s_{\max}} ds P(s) \exp\left[-\frac{[k-(s-1)-(1-\mu)(a+\langle a\rangle)t]^2}{2(1-\mu)(a+\langle a\rangle)t} \right]. \end{aligned} \tag{A.1.15}$$

By working out the exponential term, we eventually get:

$$\exp\left[-\frac{(k-(1-\mu)(a+\langle a\rangle)t)^2 + (s-1)\left[(s-1) - 2(k-(1-\mu)(a+\langle a\rangle)t)\right]}{2(1-\mu)(a+\langle a\rangle)t} \right]. \tag{A.1.16}$$

The rise of new terms proportional to s^2 and ks makes it difficult to perform the integral in Eq. (A.1.15).

We can, however, give a solution by looking at the $P(a, s, k, t)$. When setting $P_c(a, s, k_c, t) = \delta(k_c - (\bar{s} - 1))$, Eq. (A.1.15) reduces to:

$$P(a, s, k, t) = \begin{cases} P_o(a, k - (s-1), t) & \text{for } k \geq s - 1, \quad \text{(A.1.17a)} \\ 0 & \text{for } k < s - 1. \quad \text{(A.1.17b)} \end{cases}$$

In the simple case $P(s) = \delta(s - \bar{s})$, and $P(s) = \delta(s - \bar{s})$, Eq. (A.1.17) reduces to:

$$P(a, k, t) = \begin{cases} P_o(a, k - (\bar{s} - 1), t) & \text{for } k \geq \bar{s} - 1, \quad \text{(A.1.18a)} \\ 0 & \text{for } k < \bar{s} - 1. \quad \text{(A.1.18b)} \end{cases}$$

The average degree $\langle k(a,t) \rangle$

Now that we computed the $P(a, s, k, t)$ for both the $t \lesssim \langle \tau(\mu a, s) \rangle_s$ and for $t \gtrsim \langle \tau(\mu a, s) \rangle_s$ limit, we can provide a simple expression for the average degree. As we already showed $\langle k_c(a, s, t) \rangle$ grows as:

$$\langle (\rangle k_c(a, s, t)) = (s - 1) \left(1 - exp \left(- \frac{t}{\tau(\mu a_i, s)} \right) \right) \quad \text{(A.1.19)}$$

and the $k_o(a, s, t)$ grows as the mean value of the distribution found in Eq. (A.1.10d), that is:

$$k_o(a, t) = (1 - \mu)(a + \langle a \rangle)t, \quad \text{(A.1.20)}$$
$$\text{(A.1.21)}$$

so that the average total degree $k_T(a, t)$ for nodes of activity a belonging to communities of size s reads, for $t \lesssim \tau(\mu a_i, s)$:

$$\langle k_T(a, s, t) \rangle = k_c(a, s, t) + k_o(a, t) = (a + \langle a \rangle)t. \quad \text{(A.1.22)}$$

As time grows toward the $t \gtrsim \langle \tau(\mu a, s) \rangle_s$ regime, we cannot approximate anymore the $k_c(\mu a, s, t)$ but we insert Eq. (A.1.2), so that

$$\langle k_T(a, s, t) \rangle = k_o(a, t) + (s - 1) \left(1 - P(\mu a_i, s, t) \right). \quad \text{(A.1.23)}$$

for $t \gtrsim \langle \tau(\mu a, s) \rangle_s$.

The above equations then predict that $k_T(a, t)$ has a linear growth proportional to $a + \langle a \rangle$ for short time limit (Eq. (A.1.22)) and then a transition for $t \sim \langle \tau(\mu a, s) \rangle_s$ followed by a second linear growth proportional to $(1 - \mu)(a + \langle a \rangle)$. These regimes correspond to the initial growth where both the in-community and the put-community degree are growing linearly in time, followed by the slowing down of the in-community degree that is saturating to $s - 1$ and then another linear regime driven by the $(1 - \mu)(a + \langle a \rangle)$ coefficient, meaning that only the out-community degree is growing.

The degree distribution $\rho(k)$

Now that we have the expression of the average degree, it is straightforward to write the degree distribution. At all the time scales we found $\langle k_T(a,t)\rangle \propto Ct$, where C is a time-independent coefficient. Then, $k \propto at$, so that $da = dk$. If now $F(a)da = a^{-\nu}da$ we have that:

$$F(a)da = a^{-\nu}da \;\Rightarrow\; \rho(k)dk \propto k^{-\nu}dk. \tag{A.1.24}$$

Eq. (A.1.24) tells us that the degree distribution $\rho(k)$ has the same exponent of the activity distribution.

A.1.3 Comparison with numerical simulations

To check the analytical predictions of Section A.1.2 we performed numerical simulation. In particular we realized 20 representations of a network featuring:

- $N = 10^5$ nodes with modularity $\mu = 0.9$ evolving for 10^5 evolution steps;
- activity potential distributed following the $F(a) \propto a^{-\nu}$ with $\nu = 2.1$ and $a \in [10^{-3,1}]$ interval;
- power-law distributed community sizes $P(s) \propto s^{-\beta}$ with $\beta = 2.1$ and $s \in [100, \sqrt{N}]$.

In order to analyze the collective behavior of the nodes we group them by their activity and community size, thus defining b classes of nodes. We average over the representations of the network and for each class of nodes b we evaluate:

- $P_c(a,s,t)$, $P_o(a,s,t)$, $P(a,s,t)$ for $t \ll \tau(\mu a_i, s, t)$ and $t \gtrsim \tau(\mu a_i, s, t)$;
- the average degree $\langle k_c(a,s,t)\rangle$, $\langle k_o(a,t)\rangle$ and $\langle k_T(a,s,t)\rangle$;
- the degree distribution $\rho(k)$.

In Figures A.1 and A.2 we present the results and the comparison with analytical predictions concerning the in-, out-community and total degree, respectively.

As one can see, we found an excellent agreement between analytical and numerical results, and the rescaling of all the observables with the characteristic time $\tau(\mu a_i, s)$, community size s and activity potential a lets all the curves collapse on an unique, general behavior.

As a last check we present in Fig. A.3 the comparison of the degree distribution $\rho(k_c)$, $\rho(k_o)$ and $\rho(k_T)$ with the predicted behavior of Eq. (A.1.24) $\rho(k) \propto k^{-\nu}$, being ν the exponent driving the activity distribution $F(a) \propto a^{-\nu}$.

Again, exception made for the $\rho(k_c)$ that saturates to the distribution of communities sizes $P(s)$, we find a good agreement between the predicted behavior and the numerical results.

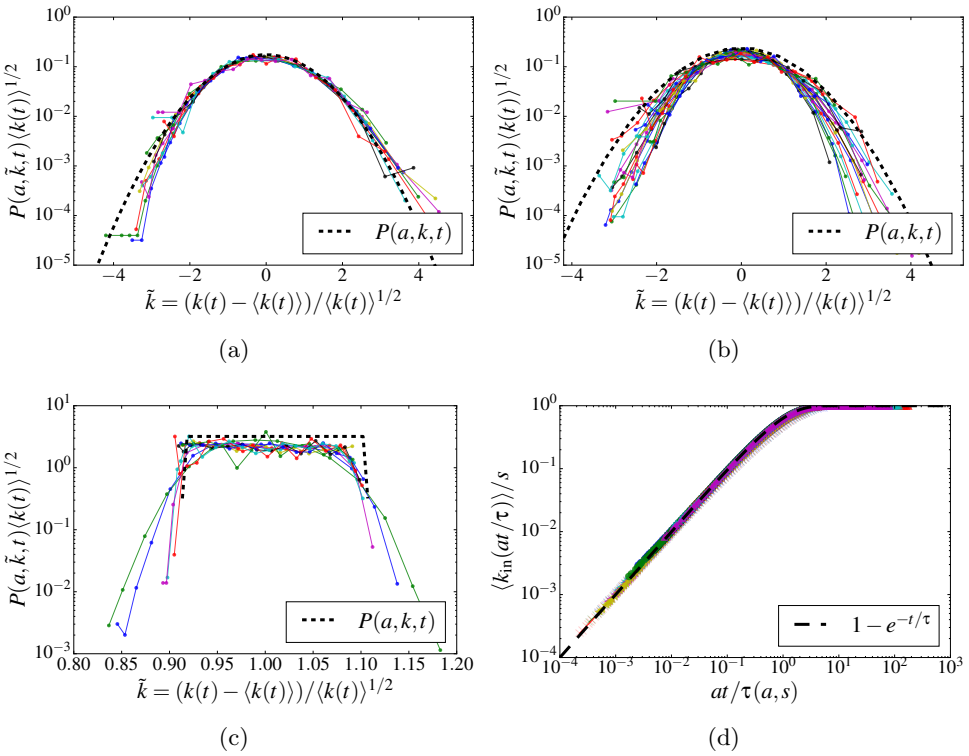


Figure A.1: The rescaled $P_c(a, s, k, t)$ probability distribution as found for a selected node class at different times (symbols): (a) $t \lesssim \tau(\mu a_i, s)$, (b) $t \sim \tau(\mu a_i, s)$ and (c) $t \gg \tau(\mu a_i, s)$. The function is rescaled accordingly to Eq. (A.1.8), i.e., by sending $k \to \tilde{k} = (k - \langle k(t)\rangle)/\langle k(t)\rangle^{1/2}$ and plotting $P(a, \tilde{k}, t) \cdot \langle k(t)\rangle^{1/2}$. In all the figures we also plot the analytical predictions of Eq. (A.1.10) (dashed black lines). In (d) we plot $\langle k_c(t)\rangle$ for different nodes classes, i.e., for nodes featuring different activity potential a and belonging to communities of different size s (symbols). The data are rescaled sending the time $t \to \tilde{t} = \frac{at}{\tau(a,s)}$ and then plotting $\langle k_c(\tilde{t})\rangle/s$. We also show the analytical prediction of Eq. (A.1.19) computed for a node of activity $a = \langle a\rangle$ and belonging to a community of size $s = \langle s\rangle$.

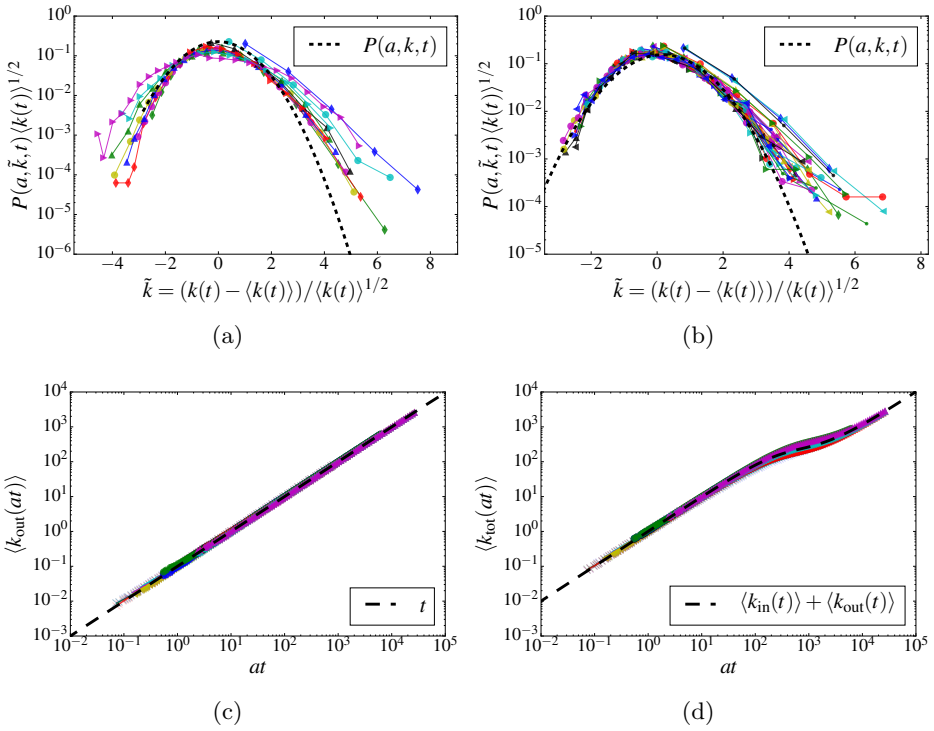


Figure A.2: (a) The rescaled $P(a,k,t)$ probability distribution for the out-community degree as found for a selected node class at different times (symbols) and the analytical prediction of Eq. (A.1.10) (black dashed line). (b) The rescaled $P(a,s,k,t)$ probability distribution of the total degree $K_T(a,s,t)$ as found for a selected node class at different times (symbols) and the analytical prediction of Eq. (A.1.14) (black dashed line). The functions are rescaled accordingly to Eq. (A.1.8), i.e., by sending $k \to \tilde{k} = (k - \langle k(t)\rangle)/\langle k(t)\rangle^{1/2}$ and plotting $P(a,\tilde{k},t)\cdot\langle k(t)\rangle^{1/2}$. In (c) we plot $\langle k_o(t)\rangle$ for different nodes classes, i.e., for nodes featuring different activity potential a (symbols). The data are rescaled sending the time $t \to \tilde{t} = at$ and then plotting $\langle k_o(\tilde{t})\rangle$. We also show the analytical prediction of Eq. (A.1.21) computed for a node of activity $a = \langle a\rangle$ (black dashed line). Finally, in (d) we show $\langle k_T(a,s,t)\rangle$ for different nodes classes (symbols) with the time t rescaled by the activity a. We also show the analytical prediction of Eq. (A.1.23) for a node of activity $a = \langle a\rangle$ belonging to a community of size $s = \langle s\rangle$ (black dashed line).

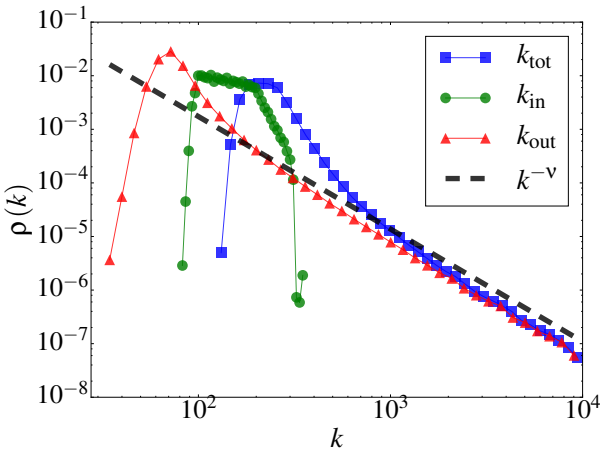


Figure A.3: The $\rho(k_T)$ (blue squares), $\rho(k_c)$ (green circles) and $\rho(k_o)$ (red triangles) compared with the $\rho(k) \propto k^{\nu}$ (black dashed line).

A.2 Leading terms evaluation for the model with burstiness

A.2.1 Integral Contributions

Here we present the computation of the integral appearing in Eq. (5.1.6). Depending on the value of the exponent α, we find three different results. In all the three cases we will perform the integral by taking the $\omega \to 0$ limit, i.e. take into account only the long-time, asymptotic region of the solution.

- $0 < \alpha < 1$: in this regime we can take the $\omega \to 0$ limit and expand the exponential term $e^{i\omega t} \sim 1 + i\omega t + \mathcal{O}(t^2)$. However, the first term proportional to t diverges as $\alpha < 1$. We can use the following trick to estimate the first diverging contribution of the integral:

$$\tilde{\mathcal{C}} \frac{\partial}{\partial \omega} \int_{\tau_0}^{\infty} \frac{e^{i\omega t'}}{t'^{\alpha+1}} dt' = i\tilde{\mathcal{C}} \int_{\tau_0}^{\infty} \frac{e^{i\omega t'}}{t'^{\alpha}} dt', \tag{A.2.1}$$

and by defining $|\omega| t' = x$ we get

$$i\tilde{\mathcal{C}} \int_{|\omega|\tau_0}^{\infty} \frac{e^{ix\,\mathrm{sign}(\omega)}}{\left(\frac{x}{|\omega|}\right)^{\alpha}} \frac{dt'}{|\omega|} = i\tilde{\mathcal{C}}|\omega|^{\alpha-1} \int_{|\omega|\tau_0}^{\infty} \frac{e^{ix\,\mathrm{sign}(w)}}{x^{\alpha}} dt' =$$

$$i\tilde{\mathcal{C}}|\omega|^{\alpha-1} \int_{|\omega|\tau_0}^{\infty} \frac{\cos(x\,\mathrm{sign}(\omega)) + i\sin(x\,\mathrm{sign}(\omega))}{x^{\alpha}} dt' \sim$$

$$i\tilde{\mathcal{C}}|\omega|^{\alpha-1} \left[\frac{(\tau_0|\omega|)^{1-\alpha}}{\alpha - 1} + \Gamma(1-\alpha) \left[\cos\left(-\frac{\pi\alpha}{2}\right) + i\,\mathrm{sign}(\omega) \sin\left(-\frac{\pi\alpha}{2}\right) \right] \right]. \tag{A.2.2}$$

Now we can integrate the last term of Eq. (A.2.2) in $d\omega$ to get the leading

term of the integral in the $\omega \to 0$ limit

$$\tilde{\mathcal{C}} \int_{\tau_0}^{\infty} \frac{e^{i\omega t'}}{t'^{\alpha+1}} dt' \sim$$
$$\sim 1 - (\tau_0\omega)^\alpha \frac{\alpha}{\alpha} \Gamma(1-\alpha) \left[\cos\left(-\frac{\pi\alpha}{2}\right) + i\,\mathrm{sign}(\omega) \sin\left(-\frac{\pi\alpha}{2}\right)\right] =$$
$$= 1 - (\tau_o\omega)^\alpha A_\alpha, \tag{A.2.3}$$

where the 1 comes from the first term of the exponential expansion and where we dropped the terms $\propto \omega$ as they die out faster than the ones $\propto \omega^\alpha$. We also encoded the complex constant multiplying the $(\tau_0\omega)^\alpha$ term in the A_α symbol.

- $1 < \alpha < 2$: in this case we can apply the same procedure of the previous case but expanding the series to the second order, thus deriving twice with respect to ω. The leading terms of the integral of Eq. (5.1.6) are:

$$\tilde{\mathcal{C}} \int_{\tau_0}^{\infty} \frac{e^{i\omega t'}}{t'^{\alpha+1}} dt' \sim$$
$$1 + i(\omega\tau_0)\frac{\alpha}{(\alpha-1)} + \tag{A.2.4}$$
$$\frac{(\tau_0\omega)^\alpha}{(\alpha-1)} \Gamma(2-\alpha) \left[\cos\left(\frac{\pi\alpha}{2}\right) - i\,\mathrm{sign}(\omega) \sin\left(\frac{\pi\alpha}{2}\right)\right],$$

and we define $C_\alpha = \frac{1}{(\alpha-1)}\Gamma(2-\alpha)\left[\cos\left(\frac{\pi\alpha}{2}\right) - i\,\mathrm{sign}(\omega)\sin\left(\frac{\pi\alpha}{2}\right)\right]$.

- $\alpha > 2$: in this case it is sufficient to consider the first three orders of the integral expansion so that:

$$\tilde{\mathcal{C}} \int_{\tau_0}^{\infty} \frac{e^{i\omega t'}}{t'^{\alpha+1}} dt' \sim 1 + i(\omega\tau_0)\frac{\alpha}{(\alpha-1)} - (\omega\tau_0)^2 \frac{\alpha}{(\alpha-2)}. \tag{A.2.5}$$

A.2.2 Asymptotic solution $\alpha < 1$

Using Eq. (A.2.3) we can rewrite Eq. (5.1.6) as

$$A_\alpha(\tau_0\omega)^\alpha Q(k,w) + \left(\frac{c}{k}\right)^\beta \frac{\partial Q(k,\omega)}{\partial k} = \delta(k,0). \tag{A.2.6}$$

To solve the equation we introduce the variable $h = k^{1+\beta}$ so that:

$$A_\alpha(\tau_0|\omega|)^\alpha Q(h,\omega) + c^\beta(1+\beta)\frac{\partial Q(h,\omega)}{\partial h} = \delta(h,0). \tag{A.2.7}$$

We Fourier transform this equation in space sending $h \to q$ getting

$$A_\alpha(\tau_0|\omega|)^\alpha Q(q,\omega) + ic^\beta(1+\beta)qQ(q,\omega) = 1, \tag{A.2.8}$$

so that:

$$Q(q,\omega) = \frac{1}{A_\alpha(\tau_0|\omega|)^\alpha + iqc^\beta(1+\beta)}. \tag{A.2.9}$$

Now let us introduce the variable h also in Eq. (5.1.4) calling $R(h,\omega) = P(h^{1/(1+\beta)},\omega)$. We then perform the FT of Eq. (5.1.4) with respect to h calling q the transformed variable and getting:

$$R(q,\omega) = \tilde{Q}(q,\omega) \int e^{i\omega t'} \int_{t'}^{\infty} \frac{\tilde{\mathcal{C}}}{\tau^{\alpha+1}} dt' d\tau. \tag{A.2.10}$$

The integral on the r.h.s. of the last equation is the same as the one in Eq. (A.2.3) with $\alpha \to \alpha - 1$ so that:

$$R(q,\omega) = \frac{B_\alpha |\omega|^{\alpha-1}}{A_\alpha (\tau_0|\omega|)^\alpha + ic^\beta(1+\beta)q}, \tag{A.2.11}$$

where $B_\alpha = A_{\alpha-1}\tau_0^\alpha$. This equation is the same of Eqq. (8) and subsequent in reference [265] therefore we have

$$P(h,t) = \frac{1}{c^\beta(\beta+1)(t/\tau_0)^\alpha} f_\alpha \left(\frac{h}{c^\beta(\beta+1)(t/\tau_0)^\alpha} \right) \tag{A.2.12}$$

where f_α is the Lévy function. Reintroducing the degree variable $k = h^{1/(1+\beta)}$ we find:

$$P(k,t) = \frac{k^\beta}{c^\beta(t/\tau_0)^\alpha} f_\alpha \left(\frac{k^{\beta+1}}{c^\beta(\beta+1)(t/\tau_0)^\alpha} \right), \tag{A.2.13}$$

that can be rewritten as:

$$P(k,t) = \frac{1}{t^{\alpha/(1+\beta)}} \tilde{f}_{\alpha\beta} \left(\frac{k}{(t/\tau_0)^{\alpha/(1+\beta)}} \right) = \frac{1}{t^{\alpha/(1+\beta)}} \tilde{f}_{\alpha\beta} \left(\tilde{k} \right), \tag{A.2.14}$$

where $\tilde{k} = k/(c^\beta(\beta+1)(t/\tau_0)^\alpha)^{1/(1+\beta)}$ and $\tilde{f}_{\alpha\beta}(\tilde{k})$ is an unknown function of $\tilde{k}$.

Equations (A.2.14) and (A.2.13) states that the peak of the $P(k,t)$ distribution (i.e. the average degree) grows as $t^{\alpha/(1+\beta)}$.

We can show these last findings in an alternative way. Prompted by the results of the numerical simulations we suppose a scaling form of the $P(k,t)$ that reads:

$$P(k,t) \simeq \frac{1}{t^\gamma} P\left(\frac{k^{1+\beta}}{t^\gamma} \right) \xrightarrow{h=k^{1+\beta}} P(h,t) \simeq \frac{1}{t^\gamma} P\left(\frac{h}{t^\gamma} \right). \tag{A.2.15}$$

If we now compute the space and time Fourier transform of Eq. (A.2.15) calling q the transformed variable of h we find

$$\begin{aligned} P(q,\omega) &= \iint e^{i\omega t + iqh} \frac{1}{t^\gamma} P\left(\frac{h}{t^\gamma} \right) dt dh = \\ &= \iint e^{it' + ih'} \frac{1}{\omega} G\left(\frac{h'/q}{(t'/\omega)^\gamma} \right) dt' dh' = \frac{1}{\omega} \tilde{G}\left(\frac{q}{\omega^\gamma} \right), \end{aligned} \tag{A.2.16}$$

where we renamed $\omega t \to t'$ and $qh \to h'$.

By comparing the last term of the last line of Eq. (A.2.16) with Eq. (A.2.11) we find that $\gamma = \alpha$ so that we recover the results shown in Eq. (A.2.14).

The resulting scaling regime reads:

$$P(h,t) = \frac{1}{c^\beta(\beta+1)(t/\tau_0)^\alpha} f_{\alpha\beta} \left(\frac{h}{c^\beta(\beta+1)(t/\tau_0)^\alpha} \right) \tag{A.2.17}$$

where $f_{\alpha\beta}$ is a non Gaussian scaling function. Reintroducing the connectivity k we get:

$$P(k,t) = \frac{k^\beta}{c^\beta(t/\tau_0)^\alpha} f_\alpha\left(\frac{k^{\beta+1}}{c^\beta(\beta+1)(t/\tau_0)^\alpha}\right) = \left(\frac{t}{\tau_0}\right)^{\frac{-\alpha}{\beta+1}} \tilde{f}_{\alpha\beta}\left(\frac{k}{\left(\frac{t}{\tau_0}\right)^{\alpha/(\beta+1)}}\right), \tag{A.2.18}$$

where we grouped the $k/(t/\tau_0)^{\alpha/(\beta+1)}$ terms and where $\tilde{f}_{\beta\alpha}(\tilde{k})$ is an unknown, non-Gaussian scaling function of $k/\left(\frac{t}{\tau_0}\right)^{\alpha/(\beta+1)}$.

Moreover the overall time scaling of the average degree of an activity class reads:

$$k \sim c^{\beta/(1+\beta)}(at)^{\alpha/(1+\beta)} \tag{A.2.19}$$

where we used the fact that τ_0 is the inverse of the activity a.

A.2.3 Asymptotic solution $1 < \alpha < 2$

First of all let us compute the functional form of the Fourier transformed $P(k,t)$. As in the previous case, we find that the solution follows the:

$$\frac{1}{t^\gamma} P\left(\frac{k^{1+\beta}-t}{t^\gamma}\right) = \frac{1}{t^\gamma} P\left(\frac{h-t}{t^\gamma}\right), \tag{A.2.20}$$

where $h = k^{1+\beta}$. If we now follow the same procedure as in Equations (A.2.15) and following, we find:

$$\begin{aligned}
&\int e^{i\omega t+iqh}\frac{1}{t^\gamma}P\left(\frac{h-t}{t^\gamma}\right)dtdh = \ [\epsilon = h-t] \ = \int e^{i\omega t+iq(\epsilon+t)}\frac{1}{t^\gamma}P\left(\frac{\epsilon}{t^\gamma}\right)dtdh = \\
&[(\omega+q)t = t';\, q\epsilon = q'] \ = \int e^{it'+i\epsilon'}\left(\frac{\omega+q}{t'}\right)^\gamma P\left(\frac{\epsilon'}{q}\left(\frac{\omega+q}{t}\right)^\gamma\right)\frac{dt'}{\omega+q}\frac{d\epsilon'}{q} = \\
&\qquad\qquad \frac{1}{\omega+q}g\left(\frac{q}{(\omega+q)^\gamma}\right),
\end{aligned} \tag{A.2.21}$$

so that, calling $\tilde{\omega} = \omega + q$ we have $P(q,\tilde{\omega}) = \frac{1}{\tilde{\omega}}g\left(\frac{q}{\tilde{\omega}^\gamma}\right)$.

Using Eq. (A.2.4) we can rewrite Eq. (5.1.6) as

$$\begin{aligned}
Q(k,\omega) = &-\left(\frac{c}{k}\right)^\beta\frac{\partial Q}{\partial k} + \frac{1}{2}\left(\frac{c}{k}\right)^\beta\frac{\partial^2 Q}{\partial k^2} + i\omega\langle\tau\rangle Q(k,\omega) + \\
&+ Q(k,\omega) + \omega^\alpha A_\alpha Q(k,\omega) + \frac{\beta c^\beta}{k^{\beta+1}}Q(k,\omega) + \delta(k,0),
\end{aligned} \tag{A.2.22}$$

where $\langle\tau\rangle$ is the first momentum of the inter event time distribution (i.e. its mean value) and the second-last term comes from the expansion of the $(c/(c+k-1))^\beta$ term. We then introduce as in the previous case the $h = k^{\beta+1}$ variable, so that $1/k^\beta \cdot \partial/\partial k \to \beta\partial/\partial h$ and $1/k^\beta \cdot \partial^2/\partial k^2 \sim (1+\beta)^2 h^{\beta/(\beta+1)} \cdot \partial^2/\partial h^2$. We then Fourier transform with respect to space sending $h \to q$ getting:

$$\begin{aligned}
&-i\omega\langle\tau\rangle P(q,\omega) - A_\alpha\omega^\alpha P(q,\omega) = -i\beta c^\beta q P(q,\omega) + \\
&\frac{1}{2}(1+\beta)^2\int e^{iqh}h^{\beta/(\beta+1)}\frac{\partial^2 P(h,\omega)}{\partial h^2}dh + \beta c^\beta\int h^{-1}e^{iqh}P(h,\omega)dh + 1,
\end{aligned} \tag{A.2.23}$$

where $P(h,\omega) = Q(h^{1/(1+\beta)},\omega) = Q(k,\omega)$. Now let us focus on the integral containing the second-derivative term. We can rewrite it as:

$$\begin{aligned} \int e^{iqh} h^{\beta/(1+\beta)} \frac{\partial^2}{\partial h^2} \int e^{-i\omega t} P(h,t) dh dt = \\ \iint e^{iqh'} e^{-i\omega' t} (h' + \beta c^\beta t)^{\beta/(1+\beta)} \frac{\partial^2}{\partial h'^2} P((h' + \beta c^\beta t), t) dh' dt, \end{aligned} \tag{A.2.24}$$

where we introduced $\omega' = \omega - \beta c^\beta q$ and $h' = h - \beta c^\beta t$. Note that we can approximate $(h' + \beta c^\beta t) \simeq \beta c^\beta t$, as we expect $h' \ll t$. If we now integrate by parts Eq. (A.2.24) and re-write $P(h' + \beta c^\beta t, t)$ as $\int e^{i\tilde{\omega}t + i\tilde{q}h'} P(\tilde{q},\tilde{\omega}) d\tilde{q} d\tilde{\omega}$ we get:

$$\begin{aligned} -q^2 \iint dh' dt e^{iqh'} e^{-i\omega' t} (\beta c^\beta t)^{\beta/(1+\beta)} \int e^{i\tilde{\omega}t + i\tilde{q}h'} P(\tilde{q},\tilde{\omega}) d\tilde{q} d\tilde{\omega} = \\ = -q^2 \iint dt d\tilde{\omega} e^{-i\omega' t + i\tilde{\omega}t} (\beta c^\beta t)^{\beta/(1+\beta)} P(q,\tilde{\omega}), \end{aligned} \tag{A.2.25}$$

where we used the fact that the integrals in dq and $d\tilde{q}$ give us a $\delta(q,\tilde{q})$. Now we insert our prediction on the $P(q,\tilde{\omega})$ of Eq. (A.2.21). In addiction we call $\omega' t = t'$ and, once the substitution is done, $\tilde{\omega}/\omega' = y$ getting:

$$-q^2 \iint dy dt' e^{-it' + iyt'} (\beta c^\beta t'/\omega')^{\beta/(1+\beta)} \tilde{\omega} f\left(\frac{q}{\tilde{\omega}^\gamma}\right) = -\frac{q^2}{\omega'^{\beta/(1+\beta)}} H\left(\frac{q}{\tilde{\omega}^\gamma}\right), \tag{A.2.26}$$

where $H(x)$ is an unknown scaling function.

Now, making the $\omega \to \omega'$ substitution and putting the result of Eq. (A.2.26) in Eq. (A.2.23) we get:

$$\begin{aligned} -i\omega' \langle\tau\rangle f\left(\frac{q}{\omega^\gamma}\right) - A_\alpha (\omega' + \beta c^\beta q)^\alpha f\left(\frac{q}{\omega^\gamma}\right) + \\ + \frac{1}{2}(1+\beta)^2 \left(\frac{q}{\omega'^{\frac{1}{2}\beta/(1+\beta)}}\right)^2 H\left(\frac{q}{\tilde{\omega}^\gamma}\right) = 1. \end{aligned} \tag{A.2.27}$$

Now we collect an ω' term from all the members to isolate the leading order finding

$$\begin{aligned} -i\langle\tau\rangle f\left(\frac{q}{\omega^\gamma}\right) - A_\alpha \left(\omega'^{1-1/\alpha} + \beta c^\beta \frac{q}{\omega'^{1/\alpha}}\right)^\alpha f\left(\frac{q}{\omega^\gamma}\right) + \\ + \frac{1}{2}(1+\beta)^2 \left(\frac{q}{\omega'^{\frac{1}{2}\left(\frac{\beta}{(1+\beta)}+1\right)}}\right)^2 H\left(\frac{q}{\tilde{\omega}^\gamma}\right). \end{aligned} \tag{A.2.28}$$

Now if the term proportional to A_α is the leading one we find that the exponent $\gamma = 1/\alpha$, otherwise, if the term proportional to $H(q/\omega'^\gamma)$ leads the evolution we find $\gamma = (2\beta+1)/(2\beta+2)$. The separation between the two cases is set at the point where the two exponents are equal, e.g. the A_α term wins if

$$\alpha < \frac{2\beta+2}{2\beta+1}. \tag{A.2.29}$$

In the other case the term proportional to $H(q/\tilde{\omega}^\gamma)$ wins and we recover the calculation of the previous work (see also Section A.2.4).

Given these results, to show the scaling form of the $P(k,t)$ distribution let us recall the assumed scaling form of the $P(k,t)$ of Eq. (A.2.20). As already said the drift of the peak of the distribution goes as $\langle k(t)\rangle \propto t^{1/(1+\beta)}$, so that the variable inside the distribution function reads:

$$\frac{k^{1+\beta} - vt}{t^{\gamma}}. \tag{A.2.30}$$

If we rewrite $k^{\beta+1} = (k - t^{1/(1+\beta)} + t^{1/(1+\beta)})^{1+\beta}$, we can introduce the variable $x = k - t^{1/(1+\beta)} \ll 1$ and expand Eq. (A.2.30) in x:

$$k^{1+\beta} \simeq t + (1+\beta)xt^{\beta/(1+\beta)} + \mathcal{O}(x^2). \tag{A.2.31}$$

By substituting Eq. (A.2.31) in Eq. (A.2.30) we find the distribution $P(k,t)$ to scale as:

$$\frac{1}{t^{\gamma - \frac{\beta}{1+\beta}}} P\left(\frac{x}{t^{\gamma - \frac{\beta}{1+\beta}}}\right), \tag{A.2.32}$$

where $x = k - t^{1/(1+\beta)}$. Then in the Gaussian scaling $\gamma = \frac{2\beta+1}{2(\beta+1)}$ so that the scaling form reads:

$$\frac{1}{t^{\frac{1}{2(1+\beta)}}} P\left(\frac{x}{t^{\frac{1}{2(1+\beta)}}}\right), \tag{A.2.33}$$

while in the other case $\gamma = \frac{1}{\alpha}$ so that the scaling form reads:

$$\frac{1}{t^{\frac{1}{\alpha} - \frac{\beta}{1+\beta}}} P\left(\frac{x}{t^{\frac{1}{\alpha} - \frac{\beta}{1+\beta}}}\right). \tag{A.2.34}$$

As a last remark, we stress that, regardless of the α and β exponent, the peak of the distribution, i.e. the average value $\langle k(t)\rangle$ grows as

$$\langle k(t)\rangle \propto k^{\frac{1}{1+\beta}}. \tag{A.2.35}$$

This is in good agreement with numerical simulations and also allows for the prediction of the degree distribution $\rho(k)$.

A.2.4 Asymptotic solution $\alpha > 2$

Using Eq. (A.2.5) we can rewrite Eq. (5.1.6) as

$$\begin{aligned} Q(k,\omega) = -\left(\frac{c}{k}\right)^{\beta} \frac{\partial Q}{\partial k} + \frac{1}{2}\left(\frac{c}{k}\right)^{\beta} \frac{\partial^2 Q}{\partial k^2} + i\omega\langle\tau\rangle Q(k,\omega) + Q(k,\omega) + \\ + \omega^2\langle\tau^2\rangle Q(k,\omega) + \frac{\beta c^{\beta}}{k^{\beta+1}} Q(k,\omega) + \delta(k,0), \end{aligned} \tag{A.2.36}$$

where $\langle\tau^2\rangle$ is the second moment of the $P(t)$ inter-event time distribution. The calculations in this case can be done in the direct space and we recover the results of the previous work, in particular we find the $P(k,t)$ to scale as a Gaussian with mean value growing as $(t/\tau_0)^{1/(1+\beta)}$ and variance growing as $t^{1/[2(1+\beta)]}$:

$$P(k,t) \simeq \frac{1}{(t/\tau_0)^{\frac{1}{2(1+\beta)}}} \exp\left[-A_{\beta} \frac{\left(k - C_{\beta}(t/\tau_0)^{\frac{1}{1+\beta}}\right)^2}{(t/\tau_0)^{1/(1+\beta)}}\right]. \tag{A.2.37}$$

This results has the same functional form of the one found in Section A.2.3 for the $\alpha > (2\beta + 2)/(2\beta + 1)$ case.

Bibliography

[1] J. J. Binney, N. J. Dowrick, A. J. Fisher, and M. Newman. *The Theory of Critical Phenomena: An Introduction to the Renormalization Group*. Oxford University Press, Inc., New York, NY, USA, 1992.

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

[3] B. B. Mandelbrot. *The fractal geometry of nature /Revised and enlarged edition/*. 1983.

[4] Per Bak (auth.). *How Nature Works: the science of self-organized criticality*. Copernicus, 1 edition, 1996.

[5] Carter T. Butts. Revisiting the foundations of network analysis. *Science*, 325(5939):414–416, 2009.

[6] B. Bollobas. *Random Graphs*. Cambridge University Press, 2001.

[7] Bela Bollobas. *Modern Graph Theory*. Springer, 1998.

[8] Stanley Milgram. The small world problem. *Psychology Today*, 67(1):61–67, 1967.

[9] Paul Erdős and Alfréd Rényi. On random graphs i. *Publicationes Mathematicae (Debrecen)*, 6:290–297, 1959 1959.

[10] Paul Erdos and Alfred Renyi. On the evolution of random graphs. *Publ. Math. Inst. Hungary. Acad. Sci.*, 5:17–61, 1960.

[11] P. Erdős and A. Rényi. On the strength of connectedness of a random graph. *Acta Mathematica Hungarica*, 12(1-2):261–267, 1961.

[12] Duncan J. Watts and Steven H. Strogatz. Collective dynamics of /'small-world/' networks. *Nature*, 393(6684):440–442, 06 1998.

[13] Albert-László Barabási and Réka Albert. Emergence of scaling in random networks. *Science*, 286(5439):509–512, 1999.

[14] F. Tria, V. Loreto, V. D. P. Servedio, and S. H. Strogatz. The dynamics of correlated novelties. *Sci. Rep.*, 4, 07 2014.

[15] Alessandro Vespignani Alain Barrat, Marc Barthelemy. *Dynamical Processes on Complex Networks*. 1 edition, 2008.

[16] Mark Newman. *Networks: An Introduction*. Oxford University Press, 2010.

[17] *Random graphs as models of networks*. Wiley-VCH Verlag GmbH and Co. KGaA, 2005.

[18] Guido Caldarelli. *Scale-Free Networks: Complex Webs in Nature and Technology (Oxford Finance Series)*. Oxford University Press, 2013.

[19] E. A. Leicht and M. E. J. Newman. Community structure in directed networks. *Phys. Rev. Lett.*, 100:118703, Mar 2008.

[20] Jussi M. Kumpula, Jukka-Pekka Onnela, Jari Saramäki, Kimmo Kaski, and János Kertész. Emergence of communities in weighted networks. *Phys. Rev. Lett.*, 99:228701, Nov 2007.

[21] Ginestra Bianconi and Matteo Marsili. Emergence of large cliques in random scale-free networks. *EPL (Europhysics Letters)*, 74(4):740, 2006.

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

[23] Matteo Marsili, Fernando Vega-Redondo, and František Slanina. The rise and fall of a networked society: A formal model. *Proceedings of the National Academy of Sciences of the United States of America*, 101(6):1439–1442, 2004.

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

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

[26] L. A. N. Amaral, A. Scala, M. Barthélémy, and H. E. Stanley. Classes of small-world networks. *Proceedings of the National Academy of Sciences*, 97(21):11149–11152, 2000.

[27] Marc Barthélémy and Luís A. Nunes Amaral. Small-world networks: Evidence for a crossover picture. *Phys. Rev. Lett.*, 82:3180–3183, Apr 1999.

[28] Linton C Freeman. A set of measures of centrality based on betweenness. *Sociometry*, pages 35–41, 1977.

[29] M. E. J. Newman. Scientific collaboration networks. i. network construction and fundamental results. *Phys. Rev. E*, 64:016131, Jun 2001.

[30] K.-I. Goh, B. Kahng, and D. Kim. Universal behavior of load distribution in scale-free networks. *Phys. Rev. Lett.*, 87:278701, Dec 2001.

[31] M. Barthélemy. Betweenness centrality in large complex networks. *The European Physical Journal B - Condensed Matter and Complex Systems*, 38(2):163–168, 2004.

[32] R. Guimerá and L.A.N. Amaral. Modeling the world-wide airport network. *The European Physical Journal B - Condensed Matter and Complex Systems*, 38(2):381–385, 2004.

[33] Alain Barrat, Marc Barthélemy, and Alessandro Vespignani. The effects of spatial constraints on the evolution of weighted complex networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2005(05):P05003, 2005.

[34] M. E. J. Newman. Mixing patterns in networks. *Phys. Rev. E*, 67:026126, Feb 2003.

[35] M. E. J. Newman. Assortative mixing in networks. *Phys. Rev. Lett.*, 89:208701, Oct 2002.

[36] Romualdo Pastor-Satorras, Alexei Vázquez, and Alessandro Vespignani. Dynamical and correlation properties of the internet. *Phys. Rev. Lett.*, 87:258701, Nov 2001.

[37] Alexei Vázquez, Romualdo Pastor-Satorras, and Alessandro Vespignani. Large-scale topological and dynamical properties of the internet. *Phys. Rev. E*, 65:066130, Jun 2002.

[38] André Auto Moreira, José S. Andrade, and Luís A. Nunes Amaral. Extremum statistics in scale-free network models. *Phys. Rev. Lett.*, 89:268703, Dec 2002.

[39] M. Bogu, R. Pastor-Satorras, and A. Vespignani. Cut-offs and finite size effects in scale-free networks. *The European Physical Journal B*, 38(2):205–209, 2004.

[40] Hamed Seyed-allaei, Ginestra Bianconi, and Matteo Marsili. Scale-free networks with an exponent less than two. *Phys. Rev. E*, 73:046113, Apr 2006.

[41] A. Barrat, M. Barthélemy, R. Pastor-Satorras, and A. Vespignani. The architecture of complex weighted networks. *Proceedings of the National Academy of Sciences of the United States of America*, 101(11):3747–3752, 2004.

[42] Jukka-Pekka Onnela, Jari Saramäki, János Kertész, and Kimmo Kaski. Intensity and coherence of motifs in weighted complex networks. *Phys. Rev. E*, 71:065103, Jun 2005.

[43] M. Ángeles Serrano, Marián Boguñá, and Romualdo Pastor-Satorras. Correlations in weighted networks. *Phys. Rev. E*, 74:055101, Nov 2006.

[44] Jari Saramäki, Mikko Kivelä, Jukka-Pekka Onnela, Kimmo Kaski, and János Kertész. Generalizations of the clustering coefficient to weighted complex networks. *Phys. Rev. E*, 75:027105, Feb 2007.

[45] Marc Barthelemy, Bernard Gondran, and Eric Guichard. Spatial structure of the internet traffic. *Physica A: Statistical Mechanics and its Applications*, 319:633 – 642, 2003.

[46] J. Clark and D.A. Holton. *A First Look at Graph Theory.* World Scientific, 1991.

[47] T.H. Cormen, C.E. Leiserson, R.L. Rivest, and C. Stein. *Introduction To Algorithms*. MIT Press, 2001.

[48] D. Jungnickel. *Graphs, Networks and Algorithms*. Algorithms and Computation in Mathematics. Springer, 2007.

[49] S.N. Dorogovtsev, J.F.F. Mendes, and A.N. Samukhin. Principles of statistical mechanics of uncorrelated random networks. *Nuclear Physics B*, 666(3):396 – 416, 2003.

[50] Alessandro Vespignani Romualdo Pastor-Satorras. *Evolution and Structure of the Internet: A Statistical Physics Approach*. Cambridge University Press, 1 edition, 2007.

[51] Guido Caldarelli Alessandro Vespignani Guido Caldarelli, Alessandro Vespignani. *Large scale structure and dynamics of complex networks: from information technology to finance and natural science*. Complex Systems and Interdisciplinary Science. World Scientific Publishing Company, 2007.

[52] Réka Albert and Albert-László Barabási. Statistical mechanics of complex networks. *Rev. Mod. Phys.*, 74:47–97, Jan 2002.

[53] S. N. Dorogovtsev and J. F. F. Mendes. Evolution of networks. *Advances in Physics*, 51(4):1079–1187, 2002.

[54] M. E. J. Newman. The structure and function of complex networks. *SIAM Review*, 45(2):167–256, 2003.

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

[56] J.L. Moreno and H.H. Jennings. *Who shall survive?: A new approach to the problem of human interrelations*. Nervous and mental disease monograph series. Nervous and mental disease publishing co., 1934.

[57] M. E. J. Newman. The structure of scientific collaboration networks. *Proceedings of the National Academy of Sciences*, 98(2):404–409, 2001.

[58] M. E. J. Newman. Scientific collaboration networks. ii. shortest paths, weighted networks, and centrality. *Phys. Rev. E*, 64:016132, Jun 2001.

[59] Márton Karsai, Kimmo Kaski, and János Kertész. Correlated dynamics in egocentric communication networks. *PLoS ONE*, 7(7):e40612, 07 2012.

[60] Vasyl Palchykov, Kimmo Kaski, Janos Kertesz, Albert-Laszlo Barabasi, and Robin I. M. Dunbar. Sex differences in intimate relationships. *Sci. Rep.*, 2, 04 2012.

[61] D. Zinoviev and S. Llewelyn. Co-Evolution of Friendship and Publishing in Online Blogging Social Networks. *ArXiv e-prints*, January 2014.

[62] Giovanna Miritello, Esteban Moro, and Rubén Lara. Dynamical strength of social ties in information spreading. *Phys. Rev. E*, 83:045102, Apr 2011.

[63] Jari Saramäki, E. A. Leicht, Eduardo López, Sam G. B. Roberts, Felix Reed-Tsochas, and Robin I. M. Dunbar. Persistence of social signatures in human communication. *Proceedings of the National Academy of Sciences*, 111(3):942–947, 2014.

[64] Scott D. McClurg and David Lazer. Political networks. *Social Networks*, 36(0):1 – 4, 2014. ¡ce:title¿Special Issue on Political Networks¡/ce:title¿.

[65] M. E. J. Newman, Stephanie Forrest, and Justin Balthrop. Email networks and the spread of computer viruses. *Phys. Rev. E*, 66:035101, Sep 2002.

[66] José J. Ramasco, S. N. Dorogovtsev, and Romualdo Pastor-Satorras. Self-organization of collaboration networks. *Phys. Rev. E*, 70:036106, Sep 2004.

[67] M. E. J. Newman, S. H. Strogatz, and D. J. Watts. Random graphs with arbitrary degree distributions and their applications. *Phys. Rev. E*, 64:026118, Jul 2001.

[68] Gerald F. Davis, Mina Yoo, and Wayne E. Baker. The small world of the american corporate elite, 1982-2001. *Strategic Organization*, 1(3):301–326, 2003.

[69] M. Karsai, G. Iñiguez, K. Kaski, and J. Kertész. Complex contagion process in spreading of online innovation. *ArXiv e-prints*, May 2014.

[70] D. Balcan, V. Colizza, B. Goncomplex calves, H. Hu, J. J. Ramasco, and A. Vespignani. Multiscale mobility networks and the spatial spreading of infectious diseases. *Proc Natl Acad Sci USA*, 106, 2009.

[71] Duygu Balcan, Vittoria Colizza, Bruno Goncomplex calves, Hao Hu, José J. Ramasco, and Alessandro Vespignani. Multiscale mobility networks and the spatial spreading of infectious diseases. *Proceedings of the National Academy of Sciences*, 106(51):21484–21489, 2009.

[72] Juliette Stehlé, Nicolas Voirin, Alain Barrat, Ciro Cattuto, Vittoria Colizza, Lorenzo Isella, Corinne Régis, Jean-Francomplex cois Pinton, Nagham Khanafer, Wouter Van den Broeck, and Philippe Vanhems. Simulation of an seir infectious disease model on the dynamic contact network of conference attendees. *BMC Medicine*, 9(1):1–15, 2011.

[73] V. Colizza, A. Barrat, M. Barthélemy, and A. Vespignani. The role of the airline transportation network in the prediction and predictability of global epidemics. *Proc Natl Acad Sci USA*, 103, 2006.

[74] Michael T. Gastner and M. E. J. Newman. Diffusion-based method for producing density-equalizing maps. *Proceedings of the National Academy of Sciences of the United States of America*, 101(20):7499–7504, 2004.

[75] Emanuele Massaro and Franco Bagnoli. Epidemic spreading and risk perception in multiplex networks: A self-organized percolation method. *Phys. Rev. E*, 90:052817, Nov 2014.

[76] K. Sun, A. Baronchelli, and N. Perra. Epidemic Spreading in Non-Markovian Time-Varying Networks. *ArXiv e-prints*, April 2014.

[77] Wouter Broeck, Corrado Gioannini, Bruno Goncalves, Marco Quaggiotto, Vittoria Colizza, and Alessandro Vespignani. The gleamviz computational tool, a publicly available software to explore realistic epidemic spreading scenarios at the global scale. *BMC Infectious Diseases*, 11(1):37, 2011.

[78] Mark Crovella and Balachander Krishnamurthy. *Internet Measurement: Infrastructure, Traffic and Applications*. John Wiley & Sons, Inc., New York, NY, USA, 2006.

[79] Reka Albert, Hawoong Jeong, and Albert-Laszlo Barabasi. Internet: Diameter of the world-wide web. *Nature*, 401(6749):130–131, 09 1999.

[80] Andrei Broder, Ravi Kumar, Farzin Maghoul, Prabhakar Raghavan, Sridhar Rajagopalan, Raymie Stata, Andrew Tomkins, and Janet Wiener. Graph structure in the web. *Comput. Netw.*, 33(1-6):309–320, June 2000.

[81] Lada A. Adamic and Bernardo A. Huberman. The web's hidden order. *Commun. ACM*, 44(9):55–60, September 2001.

[82] Albert-Laszlo Barabasi and Zoltan N. Oltvai. Network biology: understanding the cell's functional organization. *Nat Rev Genet*, 5(2):101–113, 02 2004.

[83] Fredrik Liljeros, Christofer R. Edling, Luis A. Nunes Amaral, H. Eugene Stanley, and Yvonne Aberg. The web of human sexual contacts. *Nature*, 411(6840):907–908, 06 2001.

[84] Anne Schneeberger, Catherine H Mercer, Simon A J Gregson, Neil M Ferguson, Constance A Nyamukapa, Roy M Anderson, Anne M Johnson, and Geoff P Garnett. Scale-free networks and sexually transmitted diseases: a description of observed patterns of sexual contacts in britain and zimbabwe. *Sex Transm Dis*, 31(6):380–387, Jun 2004.

[85] Jose M Montoya, Stuart L Pimm, and Ricard V Sole. Ecological networks and their fragility. *Nature*, 442(7100):259–264, Jul 2006.

[86] Jennifer A. Dunne, Richard J. Williams, and Neo D. Martinez. Food-web structure and network theory: The role of connectance and size. *Proceedings of the National Academy of Sciences*, 99(20):12917–12922, 2002.

[87] Jennifer A. Dunne, Richard J. Williams, and Neo D. Martinez. Network structure and biodiversity loss in food webs: robustness increases with connectance. *Ecology Letters*, 5(4):558–567, 2002.

[88] Willinger Walter, Govindan Ramesh, Jamin Sugih, Paxson Vern, and Shenker Scott. Scaling phenomena in the internet: Critically examining criticality. *Proceedings of the National Academy of Sciences of the United States of America*, 99(3):2573–2580, 2002.

[89] P.J. Carrington, J. Scott, and S. Wasserman. *Models and Methods in Social Network Analysis*. Structural Analysis in the Social Sciences. Cambridge University Press, 2005.

[90] Aaron Clauset and Cristopher Moore. Accuracy and scaling phenomena in internet mapping. *Phys. Rev. Lett.*, 94:018701, Jan 2005.

[91] A. Barrat, I. Alvarez-Hamelin, L. Dall'Asta, A. Vázquez, and A. Vespignani. Sampling of networks with traceroute-like probes. *Complexus*, 3(1-3):83–96, 2006.

[92] Fabien Viger, Alain Barrat, Luca Dall'Asta, Cun-Hui Zhang, and Eric D. Kolaczyk. What is the real size of a sampled network? the case of the internet. *Phys. Rev. E*, 75:056111, May 2007.

[93] MEJ Newman. Power laws, pareto distributions and zipf's law. *Contemporary Physics*, 46(5):323–351, 2005.

[94] Aaron Clauset, Cosma Rohilla Shalizi, and M. E. J. Newman. Power-law distributions in empirical data. *SIAM Review*, 51(4):661–703, 2009.

[95] Michael Molloy and Bruce Reed. A critical point for random graphs with a given degree sequence. *Random Structures and Algorithms*, 6(2-3):161–180, 1995.

[96] MICHAEL MOLLOY and BRUCE REED. The size of the giant component of a random graph with a given degree sequence. *Combinatorics, Probability and Computing*, 7:295–305, 9 1998.

[97] William Aiello, Fan Chung, and Linyuan Lu. A random graph model for power law graphs. *Experiment. Math.*, 10(1):53–66, 2001.

[98] Bo Söderberg. General formalism for inhomogeneous random graphs. *Phys. Rev. E*, 66:066121, Dec 2002.

[99] G. Caldarelli, A. Capocci, P. De Los Rios, and M. A. Muñoz. Scale-free networks from varying vertex intrinsic fitness. *Phys. Rev. Lett.*, 89:258702, Dec 2002.

[100] Marián Boguñá and Romualdo Pastor-Satorras. Class of correlated random networks with hidden variables. *Phys. Rev. E*, 68:036112, Sep 2003.

[101] S. Wasserman and K. Faust. *Social Network Analysis: Methods and Applications*. Structural Analysis in the Social Sciences. Cambridge University Press, 1994.

[102] Steven H. Strogatz. Exploring complex networks. *Nature*, 410(6825):268–276, 03 2001.

[103] Petter Holme and Beom Jun Kim. Growing scale-free networks with tunable clustering. *Phys. Rev. E*, 65:026107, Jan 2002.

[104] S. N. Dorogovtsev, J. F. F. Mendes, and A. N. Samukhin. Size-dependent degree distribution of a scale-free growing network. *Phys. Rev. E*, 63:062101, May 2001.

[105] M. Ángeles Serrano and Marián Boguñá. Tuning clustering in random networks with arbitrary degree distributions. *Phys. Rev. E*, 72:036133, Sep 2005.

[106] Stanley Wasserman and Philippa Pattison. Logit models and logistic regressions for social networks: I. an introduction to markov graphs andp. *Psychometrika*, 61(3):401–425, 1996.

[107] Ashish Sanil, David Banks, and Kathleen Carley. Models for evolving fixed node networks: Model fitting and model testing. *Social Networks*, 17(1):65–81, 1995.

[108] Tom A. B. Snijders. The statistical evaluation of social network dynamics. *Sociological Methodology*, 31(1):361–395, 2001.

[109] S. N. Dorogovtsev, A. V. Goltsev, J. F. F. Mendes, and A. N. Samukhin. Spectra of complex networks. *Phys. Rev. E*, 68:046109, Oct 2003.

[110] Juyong Park and M. E. J. Newman. Statistical mechanics of networks. *Phys. Rev. E*, 70:066117, Dec 2004.

[111] S. N. Dorogovtsev, J. F. F. Mendes, and A. N. Samukhin. Structure of growing networks with preferential linking. *Phys. Rev. Lett.*, 85:4633–4636, Nov 2000.

[112] P. L. Krapivsky and S. Redner. Organization of growing random networks. *Phys. Rev. E*, 63:066123, May 2001.

[113] P.L. Krapivsky and S. Redner. Rate equation approach for growing networks. In Romualdo Pastor-Satorras, Miguel Rubi, and Albert Diaz-Guilera, editors, *Statistical Mechanics of Complex Networks*, volume 625 of *Lecture Notes in Physics*, pages 3–22. Springer Berlin Heidelberg, 2003.

[114] R.K. Merton. *Social Theory and Social Structure*. Enlarged ed. Free Press, 1968.

[115] Herbert A. Simon. A behavioral model of rational choice. *The Quarterly Journal of Economics*, 69(1):99–118, 1955.

[116] Derek De Solla Price. A general theory of bibliometric and other cumulative advantage processes. *Journal of the American Society for Information Science*, 27(5):292–306, 1976.

[117] Albert-László Barabási, Réka Albert, and Hawoong Jeong. Mean-field theory for scale-free random networks. *Physica A: Statistical Mechanics and its Applications*, 272(1–2):173–187, 10 1999.

[118] Santo Fortunato, Alessandro Flammini, and Filippo Menczer. Scale-free network growth by ranking. *Phys. Rev. Lett.*, 96:218701, May 2006.

[119] Jon M. Kleinberg, Ravi Kumar, Prabhakar Raghavan, Sridhar Rajagopalan, and Andrew S. Tomkins. *Computing and Combinatorics: 5th Annual International Conference, COCOON'99 Tokyo, Japan, July 26–28, 1999 Proceedings*, chapter The Web as a Graph: Measurements, Models, and Methods, pages 1–17. Springer Berlin Heidelberg, Berlin, Heidelberg, 1999.

[120] R. Kumar, P. Raghavan, S. Rajagopalan, D. Sivakumar, A. Tomkins, and E. Upfal. Stochastic models for the web graph. In *Proceedings of the 41st Annual Symposium on Foundations of Computer Science*, FOCS '00, pages 57–, Washington, DC, USA, 2000. IEEE Computer Society.

[121] Alex Fabrikant, Elias Koutsoupias, and ChristosH. Papadimitriou. Heuristically optimized trade-offs: A new paradigm for power laws in the internet. In Peter Widmayer, Stephan Eidenbenz, Francisco Triguero, Rafael Morales, Ricardo Conejo, and Matthew Hennessy, editors, *Automata, Languages and Programming*, volume 2380 of *Lecture Notes in Computer Science*, pages 110–122. Springer Berlin Heidelberg, 2002.

[122] J G WARDROP. Road paper. some theoretical aspects of road traffic research. *Proceedings of the Institution of Civil Engineers*, 1(3):325–362, 1952.

[123] Diego Garlaschelli, Guido Caldarelli, and Luciano Pietronero. Universal scaling relations in food webs. *Nature*, 423(6936):165–168, 05 2003.

[124] Jayanth R. Banavar, Amos Maritan, and Andrea Rinaldo. Size and form in efficient transportation networks. *Nature*, 399(6732):130–132, 05 1999.

[125] Marc Barthélemy and Alessandro Flammini. Optimal traffic networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2006(07):L07002, 2006.

[126] R. Guimerà, A. Díaz-Guilera, F. Vega-Redondo, A. Cabrales, and A. Arenas. Optimal network topologies for local search with congestion. *Phys. Rev. Lett.*, 89:248701, Nov 2002.

[127] Vittoria Colizza, Jayanth R. Banavar, Amos Maritan, and Andrea Rinaldo. Network structures from selection principles. *Phys. Rev. Lett.*, 92:198701, May 2004.

[128] Marián Boguá, Romualdo Pastor-Satorras, and Alessandro Vespignani. Epidemic spreading in complex networks with degree correlations. In Romualdo Pastor-Satorras, Miguel Rubi, and Albert Diaz-Guilera, editors, *Statistical Mechanics of Complex Networks*, volume 625 of *Lecture Notes in Physics*, pages 127–147. Springer Berlin Heidelberg, 2003.

[129] Marián Boguñá and Romualdo Pastor-Satorras. Epidemic spreading in correlated complex networks. *Phys. Rev. E*, 66:047104, Oct 2002.

[130] Michele Catanzaro, Marián Boguñá, and Romualdo Pastor-Satorras. Generation of uncorrelated random scale-free networks. *Phys. Rev. E*, 71:027103, Feb 2005.

[131] Priya Mahadevan, Dmitri Krioukov, Kevin Fall, and Amin Vahdat. Systematic topology analysis and generation using degree correlations. *SIGCOMM Comput. Commun. Rev.*, 36(4):135–146, August 2006.

[132] M. Barthélemy. Crossover from scale-free to spatial networks. *EPL (Europhysics Letters)*, 63(6):915, 2003.

[133] A. Barrat, M. Barthélemy, and A. Vespignani. *Dynamical processes on complex networks*. 2008.

[134] S. H. Yook, H. Jeong, A.-L. Barabási, and Y. Tu. Weighted evolving networks. *Phys. Rev. Lett.*, 86:5835–5838, Jun 2001.

[135] Alain Barrat, Marc Barthélemy, and Alessandro Vespignani. *Traffic-Driven Model of the World Wide Web Graph*, chapter Traffic-Driven Model of the World Wide Web Graph, pages 56–67. Springer Berlin Heidelberg, Berlin, Heidelberg, 2004.

[136] Alain Barrat, Marc Barthélemy, and Alessandro Vespignani. Weighted evolving networks: Coupling topology and weight dynamics. *Phys. Rev. Lett.*, 92:228701, Jun 2004.

[137] G. Bianconi. Emergence of weight-topology correlations in complex scale-free networks. *EPL (Europhysics Letters)*, 71(6):1029, 2005.

[138] R. Lambiotte, V. Salnikov, and M. Rosvall. Effect of Memory on the Dynamics of Random Walks on Networks. *ArXiv e-prints*, January 2014.

[139] E. Agliari and R. Burioni. Random walks on deterministic scale-free networks: Exact results. *Phys. Rev. E*, 80(3):031125, September 2009.

[140] Ralf Metzler and Joseph Klafter. The random walk's guide to anomalous diffusion: a fractional dynamics approach. *Physics Reports*, 339(1):1 – 77, 2000.

[141] C. Gardiner. *Stochastic Methods: A Handbook for the Natural and Social Sciences*. Springer Series in Synergetics. Springer Berlin Heidelberg, 2009.

[142] S.K. Ma. *Statistical Mechanics*. World Scientific, 1985.

[143] D. Chandler. *Introduction to Modern Statistical Mechanics*. Oxford University Press, 1987.

[144] R. Balescu. *Statistical Dynamics: Matter Out of Equilibrium*. Imperial College Press, 1997.

[145] K. Huang. *Statistical mechanics*. Wiley, 1987.

[146] R.M. Anderson and R.M. May. *Infectious Diseases of Humans: Dynamics and Control*. Dynamics and Control. OUP Oxford, 1992.

[147] O. Diekmann, J. A. P. Heesterbeek, and J. A. J. Metz. On the definition and the computation of the basic reproduction ratio r 0 in models for infectious diseases in heterogeneous populations. 28(4):365–382, 1990.

[148] Matt J. Keeling and Pejman Rohani. *Modeling Infectious Diseases in Humans and Animals*. Princeton University Press, 2008.

[149] W. O. Kermack and A. G. McKendrick. A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 115(772):700–721, 1927.

[150] N.T.J. Bailey. *The Mathematical Theory of Infectious Diseases and Its Applications*. Mathematics in Medicine Series. Griffin, 1975.

[151] Romualdo Pastor-Satorras and Alessandro Vespignani. Epidemic spreading in scale-free networks. *Phys. Rev. Lett.*, 86:3200–3203, Apr 2001.

[152] Jon Kleinberg. Bursty and hierarchical structure in streams. In *Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '02, pages 91–101, New York, NY, USA, 2002. ACM.

[153] Ravi Kumar, Jasmine Novak, Prabhakar Raghavan, and Andrew Tomkins. On the bursty evolution of blogspace. *World Wide Web*, 8(2):159–178.

[154] Antoine Moinet, Michele Starnini, and Romualdo Pastor-Satorras. Burstiness and aging in social temporal networks. *Phys. Rev. Lett.*, 114:108701, Mar 2015.

[155] K.-I. Goh and A.-L. Barabási. Burstiness and memory in complex systems. *EPL (Europhysics Letters)*, 81(4):48002, 2008.

[156] Juliette Stehlé, Alain Barrat, and Ginestra Bianconi. Dynamical and bursty interactions in social networks. *Phys. Rev. E*, 81:035101, Mar 2010.

[157] Alexei Vázquez, João Gama Oliveira, Zoltán Dezsö, Kwang-Il Goh, Imre Kondor, and Albert-László Barabási. Modeling bursts and heavy tails in human dynamics. *Phys. Rev. E*, 73:036127, Mar 2006.

[158] M. Ángeles Serrano, Alessandro Flammini, and Filippo Menczer. Modeling statistical properties of written text. *PLoS ONE*, 4(4):e5372, 04 2009.

[159] Albert-Laszlo Barabasi. The origin of bursts and heavy tails in human dynamics. *Nature*, 435(7039):207–211, 05 2005.

[160] Márton Karsai, Kimmo Kaski, Albert-László Barabási, and János Kertész. Universal features of correlated bursty behaviour. *Sci. Rep.*, 2, 05 2012.

[161] Jean-Pierre Eckmann, Elisha Moses, and Danilo Sergi. Entropy of dialogues creates coherent structures in e-mail traffic. *Proceedings of the National Academy of Sciences of the United States of America*, 101(40):14333–14337, 2004.

[162] Zimo Yang, Ai-Xiang Cui, and Tao Zhou. Impact of heterogeneous human activities on epidemic spreading. *Physica A: Statistical Mechanics and its Applications*, 390(23–24):4543 – 4548, 2011.

[163] Ciro Cattuto, Wouter Van den Broeck, Alain Barrat, Vittoria Colizza, Jean-Francomplex cois Pinton, and Alessandro Vespignani. Dynamics of person-to-person interactions from distributed rfid sensor networks. *PLoS ONE*, 5(7):e11596, 07 2010.

[164] Raj Kumar Pan and Jari Saramäki. Path lengths, correlations, and centrality in temporal networks. *Phys. Rev. E*, 84:016105, Jul 2011.

[165] P. Holme and J. Saramäki. Temporal networks. *Physics Reports*, 519:97–125, October 2012.

[166] P. Holme and J. Saramäki. *Temporal Networks*. Understanding Complex Systems. Springer Berlin Heidelberg, 2013.

[167] Akshay Java, Xiaodan Song, Tim Finin, and Belle Tseng. Why We Twitter: Understanding Microblogging Usage and Communities. In *Procedings of the Joint 9th WEBKDD and 1st SNA-KDD Workshop 2007*, pages 56–65. Springer, August 2007.

[168] Haewoon Kwak, Changhyun Lee, Hosung Park, and Sue Moon. What is twitter, a social network or a news media? In *Proceedings of the 19th International Conference on World Wide Web*, WWW '10, pages 591–600, New York, NY, USA, 2010. ACM.

[169] David Liben-Nowell and Jon Kleinberg. Tracing information flow on a global scale using internet chain-letter data. *Proceedings of the National Academy of Sciences*, 105(12):4633–4638, 2008.

[170] Taha Yasseri, Robert Sumi, and János Kertész. Circadian patterns of wikipedia editorial activity: A demographic analysis. *PLoS ONE*, 7(1):e30091, 01 2012.

[171] Nathan Eagle and Alex (Sandy) Pentland. Reality mining: Sensing complex social systems. *Personal Ubiquitous Comput.*, 10(4):255–268, March 2006.

[172] Lorenzo Isella, Mariateresa Romano, Alain Barrat, Ciro Cattuto, Vittoria Colizza, Wouter Van den Broeck, Francesco Gesualdo, Elisabetta Pandolfi, Lucilla Ravà, Caterina Rizzo, and Alberto Eugenio Tozzi. Close encounters in a pediatric ward: Measuring face-to-face proximity and mixing patterns with wearable sensors. *PLoS ONE*, 6(2):e17144, 02 2011.

[173] Juliette Stehlé, Nicolas Voirin, Alain Barrat, Ciro Cattuto, Lorenzo Isella, Jean-Francomplex cois Pinton, Marco Quaggiotto, Wouter Van den Broeck, Corinne Régis, Bruno Lina, and Philippe Vanhems. High-resolution measurements of face-to-face contact patterns in a primary school. *PLoS ONE*, 6(8):e23176, 08 2011.

[174] Laetitia Gauvin, André Panisson, Ciro Cattuto, and Alain Barrat. Activity clocks: spreading dynamics on temporal networks of human contact. *Sci. Rep.*, 3, 10 2013.

[175] André Panisson, Alain Barrat, Ciro Cattuto, Wouter Van den Broeck, Giancarlo Ruffo, and Rossano Schifanella. On the dynamics of human proximity for data diffusion in ad-hoc networks. *Ad Hoc Networks*, 10(8):1532 – 1543, 2012. Special Issue on Social-Based Routing in Mobile and Delay-Tolerant Networks.

[176] Su-Yu Liu, Andrea Baronchelli, and Nicola Perra. Contagion dynamics in time-varying metapopulation networks. *Phys. Rev. E*, 87:032805, Mar 2013.

[177] B.Ø. Palsson. *Systems Biology: Properties of Reconstructed Networks*. Cambridge University Press, 2006.

[178] Teresa M Przytycka, Mona Singh, and Donna K Slonim. Toward the dynamic interactome: it's about time. *Briefings in Bioinformatics*, 11(1):15–29, 01 2010.

[179] Sophie Lèbre, Jennifer Becq, Frédéric Devaux, Michael PH Stumpf, and Gaëlle Lelandais. Statistical inference of the time-varying structure of gene-regulation networks. *BMC Systems Biology*, 4(1):1–16, 2010.

[180] Olaf Sporns, Dante R. Chialvo, Marcus Kaiser, and Claus C. Hilgetag. Organization, development and function of complex brain networks. *Trends in Cognitive Sciences*, 8(9):418–425, 2016/01/14.

[181] Ed Bullmore and Olaf Sporns. Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat Rev Neurosci*, 10(3):186–198, 03 2009.

[182] C. Pahl-Wostl. *The Dynamic Nature of Ecosystems: Chaos and Order Entwined*. Wiley, 1995.

[183] S. Ghosh. *Distributed Systems: An Algorithmic Approach*. Chapman & Hall/CRC Computer and Information Science Series. CRC Press, 2006.

[184] L. da F. Costa, F. A. Rodrigues, G. Travieso, and P. R. Villas Boas. Characterization of complex networks: A survey of measurements. *Advances in Physics*, 56(1):167–242, 01 2007.

[185] Petter Holme, Christofer R. Edling, and Fredrik Liljeros. Structure and time evolution of an internet dating community. *Social Networks*, 26(2):155–174, 5 2004.

[186] M. Karsai, M. Kivelä, R. K. Pan, K. Kaski, J. Kertész, A.-L. Barabási, and J. Saramäki. Small but slow world: How network topology and burstiness slow down spreading. *Phys. Rev. E*, 83:025102, Feb 2011.

[187] J. Tang, S. Scellato, M. Musolesi, C. Mascolo, and V. Latora. Small-world behavior in time-varying graphs. *Phys. Rev. E*, 81:055101, May 2010.

[188] Nicola Santoro, Walter Quattrociocchi, Paola Flocchini, Arnaud Casteigts, and Frédéric Amblard. Time-Varying Graphs and Social Network Analysis: Temporal Indicators and Metrics. In *3rd AISB Social Networks and Multiagent Systems Symposium (SNAMAS*, pages 32–38, United Kingdom, May 2011.

[189] Matthew O. Jackson. *Social and Economic Networks*. Princeton University Press, Princeton, NJ, USA, 2008.

[190] Easley David and Kleinberg Jon. *Networks, Crowds, and Markets: Reasoning About a Highly Connected World*. Cambridge University Press, New York, NY, USA, 2010.

[191] John Tang, Mirco Musolesi, Cecilia Mascolo, Vito Latora, and Vincenzo Nicosia. Analysing information flows and key mediators through temporal centrality metrics. In *Proceedings of the 3rd Workshop on Social Network Systems*, SNS '10, pages 3:1–3:6, New York, NY, USA, 2010. ACM.

[192] Scott A. Hill and Dan Braha. Dynamic model of time-dependent complex networks. *Phys. Rev. E*, 82:046105, Oct 2010.

[193] Garry Robins, Pip Pattison, Yuval Kalish, and Dean Lusher. An introduction to exponential random graph (p*) models for social networks. *Social Networks*, 29(2):173–191, 5 2007.

[194] Mladen Kolar, Le Song, Amr Ahmed, and Eric P. Xing. Estimating time-varying networks. *Ann. Appl. Stat.*, 4(1):94–123, 03 2010.

[195] Kun Zhao, Juliette Stehlé, Ginestra Bianconi, and Alain Barrat. Social network dynamics of face-to-face interactions. *Phys. Rev. E*, 83:056109, May 2011.

[196] Hang-Hyun Jo, Raj Kumar Pan, and Kimmo Kaski. Emergence of bursts and communities in evolving weighted networks. *PLoS ONE*, 6(8):e22687, 08 2011.

[197] Erik Volz and Lauren Ancel Meyers. Susceptible–infected–recovered epidemics in dynamic contact networks–infected–recovered epidemics in dynamic contact networks. *Proceedings of the Royal Society of London B: Biological Sciences*, 274(1628):2925–2934, 2007.

[198] Mirjam Kretzschmar and Martina Morris. Measures of concurrency in networks and the spread of infectious disease. *Mathematical Biosciences*, 133(2):165–195, 4 1996.

[199] Tatyana S. Turova. Dynamical random graphs with memory. *Phys. Rev. E*, 65:066102, Jun 2002.

[200] Tom A. B. Snijders, Johan Koskinen, and Michael Schweinberger. Maximum likelihood estimation for social network dynamics. *Ann. Appl. Stat.*, 4(2):567–588, 06 2010.

[201] Tom A. B. Snijders, Gerhard G. van de Bunt, and Christian E. G. Steglich. Introduction to stochastic actor-based models for network dynamics. *Social Networks*, 32(1):44–60, 1 2010.

[202] Petter Holme. Network reachability of real-world contact sequences. *Phys. Rev. E*, 71:046119, Apr 2005.

[203] W.E. Leland and D.V. Wilson. High time-resolution measurement and analysis of lan traffic: Implications for lan interconnection. In *INFOCOM '91. Proceedings. Tenth Annual Joint Conference of the IEEE Computer and Communications Societies. Networking in the 90s., IEEE*, pages 1360–1366 vol.3, Apr 1991.

[204] P. Holme. Network dynamics of ongoing social relationships. *EPL (Europhysics Letters)*, 64(3):427, 2003.

[205] R. Dean Malmgren, Daniel B. Stouffer, Adilson E. Motter, and Luís A. N. Amaral. A poissonian explanation for heavy tails in e-mail communication. *Proceedings of the National Academy of Sciences*, 105(47):18153–18158, 2008.

[206] R. Dean Malmgren, Daniel B. Stouffer, Andriana S. L. O. Campanharo, and Luís A. Nunes Amaral. On universality in human correspondence activity. *Science*, 325(5948):1696–1700, 2009.

[207] Hang-Hyun Jo, Márton Karsai, János Kertész, and Kimmo Kaski. Circadian pattern and burstiness in mobile phone communication. *New Journal of Physics*, 14(1):013055, 2012.

[208] Tao Zhou, Zhi-Dan Zhao, Zimo Yang, and Changsong Zhou. Relative clock verifies endogenous bursts of human dynamics. *EPL (Europhysics Letters)*, 97(1):18006, 2012.

[209] N. Perra, B. Goncomplex calves, R. Pastor-Satorras, and A. Vespignani. Activity driven modeling of time varying networks. *Sci. Rep.*, 2, 06 2012.

[210] Erik Volz and Lauren Ancel Meyers. Epidemic thresholds in dynamic contact networks. *Journal of The Royal Society Interface*, 6(32):233–241, 2009.

[211] Damon Centola, Juan Carlos González-Avella, Víctor M. Eguíluz, and Maxi San Miguel. Homophily, cultural drift, and the co-evolution of cultural groups. *Journal of Conflict Resolution*, 51(6):905–929, 2007.

[212] Carter T. Butts. A relationa event framework for social action. *Sociological Methodology*, 38(1):155–200, 2008.

[213] Martina Morris and Mirjam Kretzschmar. Concurrent partnerships and the spread of hiv, 1997.

[214] Alessandro Vespignani. Modelling dynamical processes in complex socio-technical systems. *Nat Phys*, 8(1):32–39, 01 2012.

[215] N. Perra, A. Baronchelli, D. Mocanu, B. Goncomplex calves, R. Pastor-Satorras, and A. Vespignani. Random Walks and Search in Time-Varying Networks. *Physical Review Letters*, 109(23):238701, December 2012.

[216] Suyu Liu, Nicola Perra, Márton Karsai, and Alessandro Vespignani. Controlling contagion processes in activity driven networks. *Phys. Rev. Lett.*, 112:118702, Mar 2014.

[217] Michele Starnini and Romualdo Pastor-Satorras. Temporal percolation in activity-driven networks. *Phys. Rev. E*, 89:032807, Mar 2014.

[218] Michele Starnini and Romualdo Pastor-Satorras. Topological properties of a time-integrated activity-driven network. *Phys. Rev. E*, 87:062807, Jun 2013.

[219] Mark S. Granovetter. The strength of weak ties. *American Journal of Sociology*, 78(6):1360–1380, 1973.

[220] Noah Friedkin. A test of structural features of granovetter's strength of weak ties theory. *Social Networks*, 2(4):411–422, 1980.

[221] Nan Lin, Walter M. Ensel, and John C. Vaughn. Social resources and strength of ties: Structural factors in occupational status attainment. *American Sociological Review*, 46(4):393–405, 1981.

[222] Mark Granovetter. The strength of weak ties: A network theory revisited. *Sociological Theory*, 1(1983):201–233, 1983.

[223] Jacqueline Johnson Brown and Peter H. Reingen. Social ties and word-of-mouth referral behavior. *Journal of Consumer Research*, 14(3):350–362, 1987.

[224] Reed E. Nelson. The strength of strong ties: Social networks and intergroup conflict in organizations. *The Academy of Management Journal*, 32(2):377–401, 1989.

[225] Daniel Z. Levin and Rob Cross. The strength of weak ties you can trust: The mediating role of trust in effective knowledge transfer. *Management Science*, 50(11):1477–1490, 2004.

[226] Pasquale De Meo, Emilio Ferrara, Giacomo Fiumara, and Alessandro Provetti. On facebook, most ties are weak. *Commun. ACM*, 57(11):78–84, October 2014.

[227] Julianne Holt-Lunstad, Timothy B. Smith, and J. Bradley Layton. Social relationships and mortality risk: A meta-analytic review. *PLoS Med*, 7(7):e1000316, 07 2010.

[228] R. I. M. Dunbar. The social brain hypothesis and its implications for social evolution. *Annals of Human Biology*, 36(5):562–572, 01 2009.

[229] Marton Karsai, Nicola Perra, and Alessandro Vespignani. Time varying networks and the weakness of strong ties. *Sci. Rep.*, 4:4001, 02 2014.

[230] José Luis Iribarren and Esteban Moro. Impact of human activity patterns on the dynamics of information diffusion. *Phys. Rev. Lett.*, 103:038702, Jul 2009.

[231] A. Clauset and N. Eagle. Persistence and periodicity in a dynamic proximity network. In *DIMACS Workshop on Computational Methods for Dynamic Interaction Networks*, pages 1–5, 2007.

[232] Petter Holme and Fredrik Liljeros. Mechanistic models in computational social science. *Frontiers in Physics*, 3(78), 2015.

[233] Konstantin Klemm, Víctor M. Eguíluz, Raúl Toral, and Maxi San Miguel. Nonequilibrium transitions in complex networks: A model of social interaction. *Phys. Rev. E*, 67:026120, Feb 2003.

[234] M. Morris. Telling tails explain the discrepancy in sexual partner reports. *Nature*, 365:437, 1993.

[235] Luis EC Rocha, Fredrik Liljeros, and Petter Holme. Information dynamics shape the sexual networks of internet-mediated prostitution. *Proceedings of the National Academy of Sciences*, 107(13):5706, 2010.

[236] Bruno Ribeiro, Nicola Perra, and Andrea Baronchelli. Quantifying the effect of temporal resolution on time-varying networks. *Sci. Rep.*, 3, 10 2013.

[237] R. Pfitzner, I. Scholtes, A. Garas, C.J Tessone, and F. Schweitzer. Betweenness preference: Quantifying correlations in the topological dynamics of temporal networks. *Phys. Rev. Lett.*, 110:19, 2013.

[238] Eytan Bakshy, Itamar Rosenn, Cameron Marlow, and Lada Adamic. The role of social networks in information diffusion. In *Proc. ACM Intl. World Wide Web Conf. (WWW)*, pages 519–528, 2012.

[239] James H. Fowler and Nicholas A. Christakis. Cooperative behavior cascades in human social networks. *Proceedings of the National Academy of Sciences*, 107(12):5334–5338, 2010.

[240] David Lazer, Alex Pentland, Lada Adamic, Sinan Aral, Albert-László Barabási, Devon Brewer, Nicholas Christakis, Noshir Contractor, James Fowler, Myron Gutmann, Tony Jebara, Gary King, Michael Macy, Deb Roy, and Marshall Van Alstyne. Computational social science. *Science*, 323(5915):721–723, 2009.

[241] R. I. M. Dunbar. Neocortex size as a constraint on group size in primates. *Journal of Human Evolution*, 22(6):469–493, 1992.

[242] Giovanna Miritello, Rubén Lara, Manuel Cebrian, and Esteban Moro. Limited communication capacity unveils strategies for human interaction. *Sci. Rep.*, 3, 06 2013.

[243] James Stiller and R. I. M. Dunbar. Perspective-taking and memory capacity predict social network size. *Social Networks*, 29(1):93–104, 1 2007.

[244] Joanne Powell, Penelope A. Lewis, Neil Roberts, Marta García-Fiñana, and R. I. M. Dunbar. Orbital prefrontal cortex volume predicts social network size: an imaging study of individual differences in humans. *Proceedings of the Royal Society of London B: Biological Sciences*, 2012.

[245] Zoltan Toroczkai and Kevin E. Bassler. Network dynamics: Jamming is limited in scale-free systems. *Nature*, 428(6984):716–716, 04 2004.

[246] Bruno Goncomplex calves, Nicola Perra, and Alessandro Vespignani. Modeling users' activity on twitter networks: Validation of dunbar's number. *PLoS ONE*, 6(8):e22656, 08 2011.

[247] F. Caravelli, A. Hamma, and M. Di Ventra. Scale-free networks as an epiphenomenon of memory. *ArXiv e-prints*, December 2013.

[248] J. M. Read, K. T. Eames, and W. J. Edmunds. Dynamic social networks and the implications for the spread of infectious disease. *J R Soc Interface*, 5, 2008.

[249] Gueorgi Kossinets and Duncan J. Watts. Empirical analysis of an evolving social network. *Science*, 311(5757):88–90, 2006.

[250] M. Formentin, A. Lovison, A. Maritan, and G. Zanzotto. Universal activity pattern in human interactive dynamics. *ArXiv e-prints*, May 2014.

[251] Mario V. Tomasello, Nicola Perra, Claudio J. Tessone, Márton Karsai, and Frank Schweitzer. The role of endogenous and exogenous mechanisms in the formation of r&d networks. *Scientific Reports*, 4:5679 EP –, 07 2014.

[252] I. Scholtes, N. Wider, R. Pfitzner, A. Garas, C. Juan Tessone, and F. Schweitzer. Slow-Down vs. Speed-Up of Diffusion in Non-Markovian Temporal Networks. *ArXiv e-prints*, July 2013.

[253] R. Dean Malmgren, Daniel B. Stouffer, Andriana S. L. O. Campanharo, and Luís A. Nunes Amaral. On universality in human correspondence activity. *Science*, 325(5948):1696–1700, 2009.

[254] Christian L. Vestergaard, Mathieu Génois, and Alain Barrat. How memory generates heterogeneous dynamics in temporal networks. *Phys. Rev. E*, 90:042805, Oct 2014.

[255] M. Henkel, H. Hinrichsen, M. Pleimling, and S. Lübeck. *Non-Equilibrium Phase Transitions: Volume 2: Ageing and Dynamical Scaling Far from Equilibrium*. Theoretical and Mathematical Physics. Springer, 2011.

[256] J. Klafter and I.M. Sokolov. *First Steps in Random Walks: From Tools to Applications*. OUP Oxford, 2011.

[257] Gautier Krings, Marton Karsai, Sebastian Bernhardsson, Vincent Blondel, and Jari Saramaki. Effects of time window size and placement on the structure of an aggregated communication network. *EPJ Data Science*, 1(1):4, 2012.

[258] D.R. Cox. *Renewal Theory*. Methuen science paperbacks. Methuen, 1970.

[259] M. E. J. Newman, I. Jensen, and R. M. Ziff. Percolation and epidemics in a two-dimensional small world. *Phys. Rev. E*, 65:021904, Jan 2002.

[260] Jeff Alstott, Ed Bullmore, and Dietmar Plenz. powerlaw: A python package for analysis of heavy-tailed distributions. *PLoS ONE*, 9(1):e85777, 01 2014.

[261] Santo Fortunato. Community detection in graphs. *Physics Reports*, 486(3–5):75–174, 2 2010.

[262] A Lancichinetti, S Fortunato, and F Radicchi. New benchmark in community detection. *Phys. Rev. E*, 78:046110, 2008.

[263] Guangming Ren and Xingyuan Wang. Epidemic spreading in time-varying community networks. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 24(2):–, 2014.

[264] Kihong Chung, Yongjoo Baek, Daniel Kim, Meesoon Ha, and Hawoong Jeong. Generalized epidemic process on modular networks. *Phys. Rev. E*, 89:052811, May 2014.

[265] R Burioni, G Gradenigo, A Sarracino, A Vezzani, and A Vulpiani. Rare events and scaling properties in field-induced anomalous dynamics. *Journal of Statistical Mechanics: Theory and Experiment*, 2013(09):P09022, 2013.

[266] S.A. Kauffman and N.M.) Santa Fe Institute (Santa Fe. *Investigations: The Nature of Autonomous Agents and the Worlds They Mutually Create*. SFI working papers. Santa Fe Institute, 1996.

[267] G. Pólya. Sur quelques points de la théorie des probabilités. *Annales de l'institut Henri Poincaré*, (2):117–161, 1930.

[268] H. Mahmoud. *Polya Urn Models*. Chapman & Hall/CRC Texts in Statistical Science. CRC Press, 2008.

[269] H.S. Heaps. *Information retrieval, computational and theoretical aspects*. Library and information science. Academic Press, 1978.

[270] G.K. Zipf. *Human behavior and the principle of least effort: an introduction to human ecology*. Addison-Wesley Press, 1949.

[271] Linyuan Lü, Zi-Ke Zhang, and Tao Zhou. Zipf's law leads to heaps' law: Analyzing their relation in finite-size systems. *PLoS ONE*, 5(12):e14139, 12 2010.