UNIVERSITÀ
DI PARMA

UNIVERSITÀ DEGLI STUDI DI PARMA

DOTTORATO DI RICERCA IN FISICA

CICLO XXXI

Dynamical processes and memory effects on temporal networks.

Coordinatore:
Chiar.mo Prof Cristiano VIAPPIANI

Tutor:
Chiar.ma Prof. Raffaella BURIONI

Dottorando:
Michele TIZZANI

Anni 2015/2018

Abstract

Understanding the mechanism behind the diffusion of a disease has always been a crucial problem for health and society. In particular the evaluation of the threshold above which there is an emergence of an epidemic state is one of the fundamental problems for disease control. From the mathematical point of view, many models have been formalized over the years. In particular the introduction of networks science has significantly improved both the analysis and the prediction capability of these phenomena giving more compelling results. Many of the studies in epidemics have been carried on static networks, but in the last few years the interest toward time-varying networks has rapidly grown. Despite the increased complexity from the introduction of the time variable, some models allow an analytical understanding of the spreading processes. In particular in our work we will consider the activity-driven model in which the time variable is embedded in the propensity of each individual to interact at a certain time. Social ties are also driven by the memory people have of each other, preferring old acquaintances interactions to new ones. In this work we explore the effects that the memory mechanism in a time-varying network has on a spreading process focusing on the epidemic dynamics.
We focus our attention on two standard epidemic models: the susceptible-infected-susceptible (SIS) and the susceptible-infected-recovered (SIR), describing respectively diseases that don't or do confer immunity after the infection. We formulate an activity-based mean-field approach obtaining analytically the epidemic threshold as a function of the parameters describing the distribution of activities and the strength of the memory effects. In particular we consider the asymptotic regime in which the infection starts only when the people have had a sufficiently large number of connections in their social circle. In this limit the dynamical process can be seen as an activity-driven process evolving on an effective static graph. Our results show that memory amplifies the activity fluctuations reducing the threshold and enhancing the epidemic spreading in both the SIS and SIR models. To numerically prove our findings we simulate the epidemic process on both the time-evolving and the effective static networks, varying the memory parameter and the

starting time of the infection. Comparing the theoretical model with the numerical simulations we confirm our predictions in the asymptotic limit. We also show that in the preasymptotic regime there are strong aging effects making the epidemic threshold deeply affected by the starting time of the outbreak. In particular, for short starting times of the infection, the correlations induced by the memory produce strong backtracking effect in both the SIS and SIR processes, lowering or increasing the epidemic threshold respectively. We discuss in detail the origin of the model-dependent preasymptotic corrections, setting the bases for potential epidemics control methods on correlated temporal networks.

Contents

Abstract 2
Preface 5

Introduction 6

Chapter 1. Static networks 10
1.1. Basic definitions of static networks 10
1.2. Properties of complex networks 16
1.3. Models of Complex Networks 19
1.4. Random walk on static networks 24

Chapter 2. Temporal networks 27
2.1. Representing temporal network 27
2.2. Statistical properties of TVN 31
2.3. Models of temporal networks 32
2.4. Activity-driven network 33

Chapter 3. Epidemic Models 45
3.1. Introduction 45
3.2. Traditional models 45
3.3. Epidemics on static networks 50
3.4. Epidemics on time evolving networks 58

Chapter 4. Epidemic Spreading and Aging in Temporal Networks with Memory 64
4.1. Introduction 64
4.2. The model 65
4.3. Analytical results 66
4.4. Linear Stability Analysis 72
4.5. Numerical simulations 72

Chapter 5. Conclusions 79

Bibliography 81

Preface

The work presented in this dissertation as been carried on at the Department of Mathematics, physics and Computer Science of the University of Parma. The first three chapters introduce the main subjects of this work which that are static and time-varying networks, and epidemic processes.

Chapter 4 is the result of the collaboration with Claudio Castellano, Stefano Lenti, Enrico Ubaldi, Alessandro Vezzani and Raffaella Burioni, and it is based on the paper: *Epidemic Spreading and Aging in Temporal Networks with Memory Ref. [118].*

Introduction

We live in a complex and interconnected world. Complex systems, from microscopic to macroscopic scale, are formed of many interacting elements. From atoms to chemical compounds, from cells to organisms, from people to society, the interactions among the elements of each system defines a hierarchy of complexity that spans through different fields of science. From these examples we can see that the common characteristics of a complex system are that they are composed of a large number of interacting *agents* exhibiting *emergence*, i.e. a self-organized collective behavior not discernible from the single action of the agents.

In the last few years, the need for a new language to describe complexity has lad to the science of complex networks. There are many examples of complex networks around us. We could define two main classes of real networks, infrastructures an natural systems [12]. In the first category we found virtual structures, like the World Wide Web, or physical structures like power greed, and transportation networks. On the other hand we can refer to natural networks as to the structures forming or form by living entities, like biological and social systems. As we can see networks science touches different subjects, and for this reason most of the definitions describing the network are borrowed for different scientific fields.

The network paradigm can be very useful to study dynamical processes, such as information diffusion or epidemic spreading, which can be seen as additional ingredients evolving on top of the network structure. This approach allows studying the interplay that exists between the dynamical process and the structure from both sides.

The first approach to study dynamical system on complex networks is to consider a static approximation of the graph, where the time-scales of the evolution of the network are either too slow or too fast respect to the dynamics of the process on top of it.

On the other hand in most social and information systems, time scales of networks dynamics are often comparable to the time scales of the dynamical processes taking place on top of them. The diffusion of online information and the spreading of transmitted diseases in a population are typical examples of such processes. In these cases the static representation of the network is not able to grasp all the features of the rapidly changing topology [10, 38, 2, 9, 124]. Modern technologies are able to measure and monitor the evolution of interactions with a high time

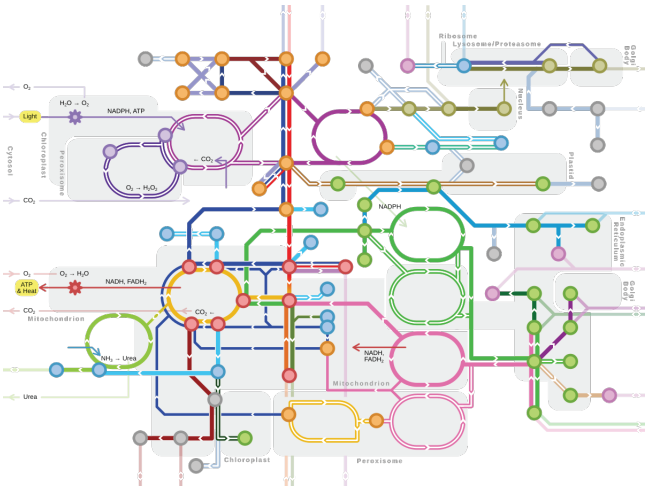

FIGURE 0.0.1. Metabolic Network. Orange nodes: carbohydrate metabolism. Violet nodes: photosynthesis. Metabolic metro Red nodes: cellular respiration. Pink nodes: cell signaling. Blue nodes: amino acid metabolism. Grey nodes: vitamin and cofactor metabolism. Brown nodes: nucleotide and protein metabolism. Green nodes: lipid metabolism. Source `https://en.wikipedia.org/wiki/Metabolic_network`

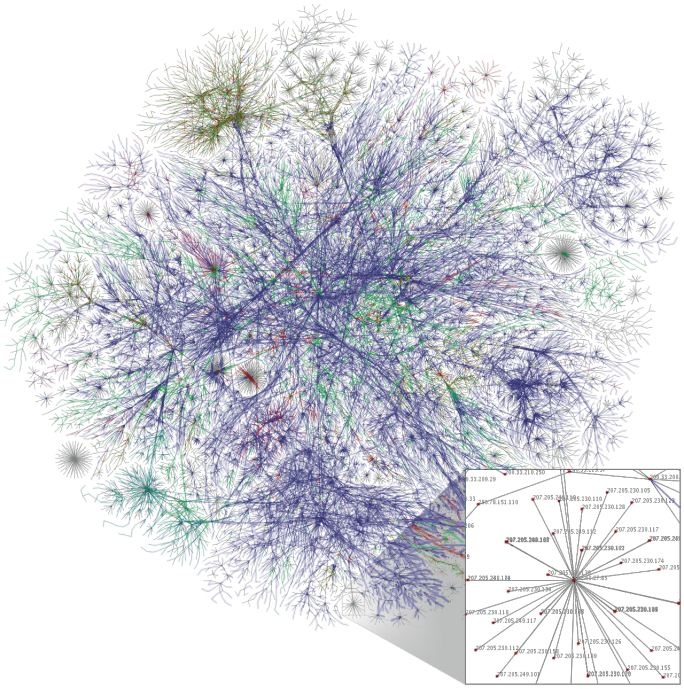

FIGURE 0.0.2. World Wide Web map. visualization of routing paths through a portion of the Internet. Source: `https://en.wikipedia.org/wiki/Internet`

resolution [29], calling for new theories to understand the effect of time-varying topologies on dynamical processes.

Especially in social systems the time evolution of the contacts is the result of the human activity, a quantity that can be easily measured from the available large scale and time-resolved datasets [101]. This analysis shows that human activities are typically highly heterogeneously distributed, and this has strong effects on network evolution. To explicitly include the effect of activity distributions on the network dynamics it has been introduced the activity-driven networks [94]. In this framework, each agent is endowed with a degree of freedom that encodes the propensity of the individual to engage in a social event, establishing a link with another agent in the system.

When links are randomly established among agents, activity-driven models have been studied in detail [95, 94, 111, 99], uncovering the effects of heterogeneous activity distributions on network topology and on dynamical processes, such as random walks and epidemic processes.

However, in general agents do not connect randomly to their peers [45, 72, 102]. During their activity, individuals remember their social circles and they are more inclined to interact with their history of contacts, establishing strong and weak ties with their peers [36, 116]. Recently this problem has been tackled by applying a data-driven approach and measuring the tie allocation mechanism in real systems, introducing a memory process to activity-driven models[55, 60]. As reasonably expected, social interactions are not randomly established but they are rather concentrated towards already contacted nodes, with a reinforcement process encoded in a single measurable memory parameter. The memory process tunes the network evolution, that can be predicted at large times [121, 23, 59], and it is also expected to influence dynamical processes. Indeed it has been shown that it changes the spreading rate in a diffusion process, slowing it down in some cases and speeding it up in others [100, 103, 64, 62, 53, 54, 98]. Similarly in epidemic spreading on activity-driven networks it can be shown that memory can lower or increase the epidemic threshold in SIS or SIR model, respectively [117]. This happens when the epidemic process and the network evolution start at the same time. However in presence of a memory process, as observed also in other fields[47], the network evolution could introduce aging in the process [73], and this could further influence the spreading dynamics. In our work, we analyze these phenomena giving a full understanding of their effects on the epidemic dynamics.

We formulate the activity-based mean-field model and analytically derive the epidemic threshold as a function of the memory parameter and the activity distribution for both the SIS and the SIR models. In particular we consider the asymptotic limit in which the epidemic process starts when the individuals have already reached a certain average number of contacts in their social circle. In this regime we can consider the epidemic evolution as a dynamical process evolving on an effective static

network. The analytic results show that the memory amplifies the activity fluctuations lowering the the epidemic threshold respect to the memoryless case for both the SIS and the SIR processes. We compare the analytical model with numerical simulation on both the time-evolving and the effective static networks, confirming our predictions.

The aging effects are recovered in the preasymptotic regime, when, for short starting time of the infection, the memory induces correlations among the infection probabilities of the nodes already contacted. Because of these correlations both the SIS and SIR present backtracking effects which lower or increase the epidemic threshold respect to the mean-field result. In this work we will discuss the reasons of this deviation opening new horizons for controlling and understanding disease and information spreading in networks with high correlations.

This work is organized as follows. In Chapter 1 we will introduce the basic concepts and models of static networks. In Chapter 2 we will introduce time-varying networks, and in particular the activity-driven framework. In Chapter 3 we will explore the main models of epidemic spreading on both static and time-varying networks. Chapter 4 is the results of original research on the epidemic spreading in time-varying networks with memory.

CHAPTER 1

Static networks

In this chapter we will introduce the basic concepts and models of static networks [83, 129, 4] that will be useful to understand some of the results on epidemic models described in the third chapter. In this framework, the structure of the system doesn't evolve in time, and we will see, later in this work, how this property affects the dynamics of a process evolving on the network.

In the first section we will introduce the basic definition of the graph theory, while in the second section we will consider the statistical properties of complex networks. In section 1.3 we will describe some of the principal models of complex static networks, and finally in the last section of we will introduce the random walk formulation on static networks.

1.1. Basic definitions of static networks

The natural theoretical framework to study complex networks is *graph theory* [18, 21, 4]. A graph, or a network, $G = (V, L)$ is a structure consisting in a set of vertices, or nodes V and a collection of links, or edges, L, such that V is non null and L is formed by pairs of elements of V. A subset $G'(V', L')$ of a graph G is called *subgraph* if $V' \subset V$ and $L' \subset L$.

The number of elements N in V represents the *order* of the network, while the number of element in L, i.e. the total number of links, is denoted by K and represents the *size* of the network, so that it is possible to define a graph also by its order and size $G(N, K)$.

In a network $G(V, L)$ two nodes i and j are said to be *adjacent* or *neighboring* if there is a link between them, this can be expressed in the matrix representation by the *adjacency matrix* A which elements A_{ij} are defined as follows

$$A_{ij} = \begin{cases} 1 & (i,j) \in L \\ 0 & (i,j) \notin L \end{cases}$$

We can introduce the main definitions of a network according to the properties of A, and of the sets V and L.

- If we associate a real number w_{ij} to an edge between two nodes i and j, the graph is *weighted* (Figure 1.1.1(d)). An example is given by the different strengths of social ties [13], in biological systems like food webs,

where weights represents different carbon flows between spices [66], or in transportation networks, where they represent the traffic flow [87].

- If a node i can be connected to itself, $A_{ii} \neq 0$, we have a *loop* or *self-edge* (figure 1.1.1(f)*)*. In this case an example is given by the network of transcription interactions in the E.Coli bacteria, where the self-edge is the transcription factors that regulate the transcription of their own genes [130].
- If two nodes i and j are connected with more than one link, $A_{ij} > 1$, we have *multiedges* and the graph is a *multigraph* (figure 1.1.1(c)).
- If the sets of nodes and links are unordered, the the graph is *undirected* (figure 1.1.1(a)), in this case two adjacent nodes i and j are mutually connected $A_{undirected}$ Eq.1.1.1. Most of the graphs we will consider from now on belong to this category. The simplest example to picture is friendship, which is a mutual interaction.
- If the set of nodes and links is ordered, the graph is *directed* (figure 1.1.1(b)), which mean that the edge between two nodes has a defined direction and generally nodes are not mutually connected $A_{directed}$ Eq.1.1.2. An example of this type of connection is given by the citation networks [109], where, because of the temporal order of the publications, a cited article can't cite back the source of the citation.

$$A_{undirected} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \tag{1.1.1}$$

$$A_{directed} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \tag{1.1.2}$$

- The number of links attached to a node i defines the *degree* k_i of the node (figure 1.1.2).
  In an undirected graph the degree can be expressed in term of the adjacency matrix as
  $$k_i = \sum_{j=1}^{N} A_{ij}$$
  in this case every edge has two end so that the total number of edges is
  $$K = \frac{1}{2} \sum_{i=1}^{N} k_i$$

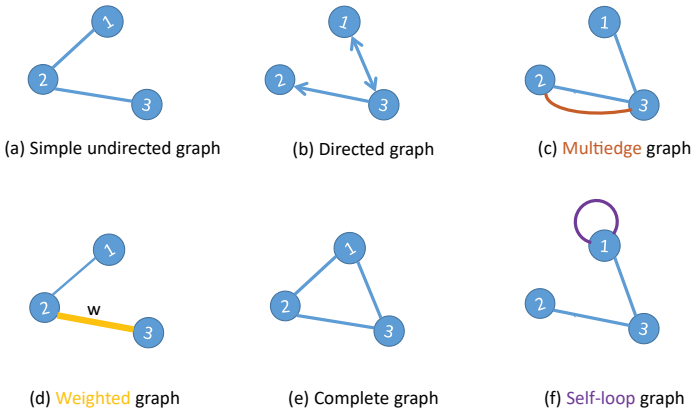

FIGURE 1.1.1. (a) Simple undirected graph; (b) directed graph; (c) multiedge graph; (d) weighted graph; (e) complete graph; (f) self-loop graph.

For a directed graph we need to distinguish between the outgoing and the incoming edges of the node i, defining respectively the *in-degree* (figure 1.1.2(b)) k_i^{in} and the *out-degree* (figure 1.1.2(c)) k_i^{out}

$$k_i^{in} = \sum_{j=1}^{N} A_{ji}$$

$$k_i^{out} = \sum_{j=1}^{N} A_{ij}$$

the total in-going number of edges is equal to the total out-going edges

$$K = \sum_{i=1}^{N} k_i^{in} = \sum_{j=1}^{N} k_i^{out}$$

hence the mean in-degree is equal to the mean out-degree.

For weighted networks we can define a *weighted degree* k_i^w for a node i given by

$$k_i^w = \sum_{j=1}^{N} A_{ij}^w$$

We can also define the strength of as the generalization of the degree for weighted networks as

$$s_i = \sum_{j} w_{ij}$$

where the sum is over all the neighbors of i.

We will not go in further details with weighted networks, and from now on we will refer only to undirected graphs, except when specified.

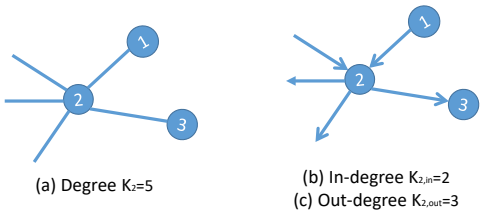

FIGURE 1.1.2. Degree for undirected (a) and directed (b) (c) networks

- The *connectance* or *density* κ of the graph is the ratio between the total number of links K and the maximum number of possible links K_{max}

$$\kappa = \frac{K}{K_{max}} = \frac{K}{\begin{pmatrix} N \\ 2 \end{pmatrix}}$$

  A graph G is *sparse* if $K \ll N^2$, i.e. $\kappa \to 0$ for $N \to \infty$, *dense* if $K = \mathcal{O}(N^2)$ and κ is constant in the limit of $N \to \infty$, or *complete* if $K = K_{max} = \binom{N}{2} = N(N-1)/2$, i.e. the nodes are all connected together. When analyzing or simulating sparse networks it is computational convenient to define the *adjacency list* which, given a node i, is the set $l = (i, s \in L(i))$ of all its first neighbors.
- The *k-core* of a graph G is the biggest subgraph in which all the nodes have at least degree k.

To understand how to move across a network we need to introduce further definitions that characterize the metric of the system.

- A series of consecutive edges connecting i_0 to i_n through n edges is a *path* $\mathcal{P}_{i_0,i_n}$ of *length* n , which mathematically speaking is subgraph $G'(V', L')$ of an ordered collection of $n+1$ vertexes V' and n edges such that $i_s \in V$ and $(i_{s-1}, i_s) \in L$ for all s. When a path passes once through all the nodes, not necessary using every edges, it is an *Hamiltonian path*. On the other hand if a path passes through all the edges, but not necessary through all the nodes, it is an *Eulerian path* [83].
- A closed path forms a *circuit* when $i_0 = i_n$ or a *cycle* if all nodes of the circuit are distinct, circuits from Hmiltonian or Eulerian paths are called Hamiltonian or Eulerian circle respectively. A set of k connected nodes without a cycle forms a *tree* of order k, and a set of disconnected trees form a *forest*. A tree of order k with maximum diameter 2 forms a *star*.
- If there is a path between every couple of nodes the graph is sad to be *connected,* and the property of being connected is the *connectivity*.
- A connected subgraph forms a *component*, while a complete subgraph forms a *clique*.
- A component that scales as the size of the network N diverging in in the infinite size limit, is called *giant component [20].*

- If it is possible to divide a graph in n classes such that all the vertexes in the same class are not adjacent, the graph is called *n-partite* graph. In the special case of $n = 2$ we have a bipartite graph [7]. An example of bipartite graph is given by the affiliation network, in which a two sports clubs share the same player during two season of a championship [51].
- A *tree* is a connected graph without cycles, while a *forest* is a not connected acyclic graph, i.e. composed by multiple trees. The natural social example of a tree is the genealogy graph, which is also directed if we consider the relation of being son to the next node.
- The *distance* d_{ij} between two nodes i and j , is the shortest path length to travel from i to j and is given by

$$d_{ij} = \min \left\{ \sum_{k,l \in \mathcal{P}_{i,j}} A_{kl} \right\}$$

  Another definition of distance can be introduced substituting the adjacency matrix A_{kl} with $A_{kl}^{-1} \neq 0$, which in case of simple graphs with entries either 1 or 0 make equal sense, but for weighted graph

$$d_{ij}^{w} = \min \left\{ \sum_{k,l \in \mathcal{P}_{i,j}} A_{kl}^{w} \right\}$$

  and

$$d_{ij}^{w} = \min \left\{ \sum_{k,l \in \mathcal{P}_{i,j}} [A_{kl}^{w}]^{-1} \right\}$$

  have different meaning.
- The *diameter* of a graph, is the maximum distance between two nodes

$$D = \max_{i,j} d_{ij}$$

  for example we could ask what is the diameter of the World Wide Web [5].
- The *average shortest path length* or *characteristic path length* is the average geodesic distance over all couple of nodes

$$\langle d \rangle = \frac{1}{N(N-1)} \sum_{i,j} d_{ij}$$

- Some time is more convenient to use the harmonic mean of the distance introducing the *efficiency* [63], which defines how efficiently a network exchange information.

$$\langle e \rangle = \frac{1}{N(N-1)} \sum_{i,j} [d_{ij}]^{-1}$$

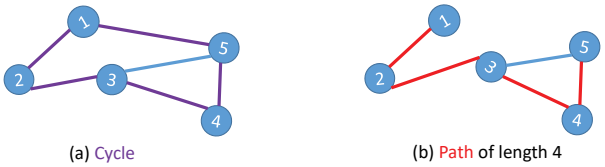

FIGURE 1.1.3. Cycle (a) and path (b) for an undirected graph.

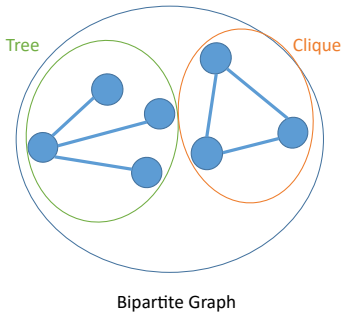

FIGURE 1.1.4. Bipartite graph, clique and tree.

To study the importance of a node i in a network we introduce some of the main *centrality measures*. These indicate for example how influential is an individual in a social network, or help to identify super-spreaders in epidemic processes.

- The simplest centrality measure is the *degree centrality* defined by the degree of the node i.
- The *closeness centrality* is the average shortest path from the node i to all the others. The more a node is close to the others the more is central.

$$g_i = \frac{1}{\sum_{i \neq j} d_{ij}}$$

- The *betweenness centrality* quantifies the capability for a node to be a bridge between the others. The more edges componing shortest paths pass through i the more the node is central

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

where σ_{hj} is the total number of shortest path from h to j and $\sigma_{hj}(i)$ are the ones that pass through i.

We can measure the tendency for the nodes of a graph to be connected between each other, and characterize the local structure of the neighbors of the node i studying the clustering of the network.

- The *clustering coefficient* is number of links around a node i

$$C_i = \frac{2}{k_i(k_i - 1)} \sum_{j,k} A_{ij} A_{jk} A_{ki}$$

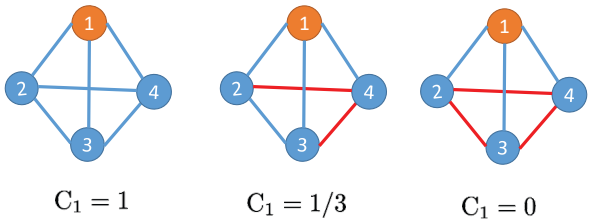

FIGURE 1.1.5. Cluster coefficient for the node 1.

for a vertex with $k_i > 1$, which is the rate between the number of pairs of connected neighbors of i and number pairs of neighbors of i. It measure the local group cohesiveness

The definition can be extended to directed networks considering the possible direction of the edges.

$$C_i^{in} = \frac{2}{k_i^{in}(k_i^{in}-1)} \sum_{j,k} A_{ij} A_{jk} \frac{(A_{ki} + A_{ki})}{2}$$

$$C_i^{out} = \frac{2}{k_i^{out}(k_i^{out}-1)} \sum_{j,k} A_{ij} A_{jk} \frac{(A_{ki} + A_{ki})}{2}$$

In some occasion can be also useful to evaluate the average clustering coefficient given by

$$\langle C \rangle = \frac{1}{N} \sum_{i=1}^{N} C_i$$

1.2. Properties of complex networks

1.2.1. Degree distribution. A fundamental statistical quantity to characterize a network is the degree distribution $P(k)$. For undirected graphs it is defined as the probability that a random chosen node has degree k, while for directed graphs we have to distinguish between in-degree $P(k^{in})$ and out-degree $P(k^{out})$ where the same meaning applies to k^{in} and k^{out}.

The nth moment of the distribution is given

$$\langle k^n \rangle = \sum_k k^n P(k)$$

or in the continuous limit

$$\langle k^n \rangle = \int dk k^n P(k)$$

while for a directed graph we have

$$\langle k_{in}^n \rangle = \sum_k k_{in}^n P(k_{in}) = \langle k_{out}^n \rangle = \sum_k k_{out}^n P(k_{out})$$

and

$$\langle k_{in}^n \rangle = \int dk k_{in}^n P(k_{in}) = \langle k_{out}^n \rangle = \int dk k_{out}^n P(k_{out})$$

Correlated and uncorrelated networks. The degree distribution completely defines the statistical properties of an uncorrelated network. However in most real networks connectivity patterns present significant correlations that affect both the topological properties of the network and the dynamical processes evolving on it.

In correlated networks [104, 90] the probability that a node with degree k is simultaneously connected to n other nodes of degree $k', ..., k'^{(n)}$ depends on k, and it is represented by $P(k', ..., k'^{(n)}|k)$. In fact in general, nodes interacts among each other respect their intrinsic properties defining specific mixing patterns. Let's consider the simplest case of a node with degree k connected to a node with degree k', the probability $P(k'|k)$ must satisfy the normalization condition

$$\sum_{k'} P(k'|k) = 1 \tag{1.2.1}$$

and the detailed balance condition

$$kP(k'|k)P(k) = k'P(k|k')P(k') \tag{1.2.2}$$

which means that the total number of links from vertexes of degree k to vertexes of degree k' must be the same pointing from vertexes of degree k' to vertexes of degree k for an undirected graph.

Introducing the joint degree distribution $P(k, k')$ representing the probability that two connected nodes have degree k and k' respectively, and using the conditions 1.2.1 and 1.2.2, it is possible to obtain the degree distribution as

$$P(k) = \frac{\langle k \rangle}{k} \sum_{k'} P(k, k')$$

In this case the network is completely characterized by the degree distribution $P(k)$ and the first conditional probability

$$P(k'|k) = \frac{\langle k \rangle \, P(k, k')}{kP(k)} \,,$$

and in particular for uncorrelated networks

$$P(k'|k) = \frac{k'P(k')}{\langle k \rangle} \,.$$

The evaluation of $P(k'|k)$ for a network of finite size N, is not easy, and sometimes to better understand the mixing topological properties of the network it is convenient to introduce the *average nearest neighbors degree* of a node i as

$$k_{nn,i} = \frac{1}{k_i} \sum_{j=1}^{N} A_{ij} k_j$$

and the average degree of the nearest neighbors with degree k as

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

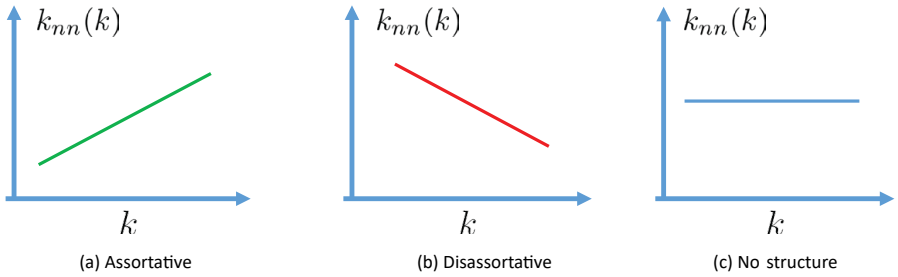

FIGURE 1.2.1. Average degree of neighbors in the assortative (a), disassortative (b) and non-structured (c) cases.

In absence of degree correlations $k_{nn}(k)$ is a constant

$$k_{nn}(k) = \frac{\langle k^2 \rangle}{\langle k \rangle}$$

and it is independent from k.

If the system presents degree correlations, then when $k_{nn}(k)$ is an increasing function of k, the graph is called *assortative* (figure 1.2.1(a)), while if it is a decreasing function of k it is *disassortative* (figure 1.2.1(b))[79]. In assortative networks the nodes tend to connect to their connectivity peers, while in disassortative networks nodes with low degree are more likely connected with highly connected ones.

To study the clustering of the network including the correlations among vertexes can be useful to define the clustering spectrum, i.e. the average clustering coefficient restricted to the nodes of degree class k [127]

$$C(k) = \frac{1}{P(k)N} \sum_{i|k_i=k} C_i \,.$$

Homogeneous and heterogeneous networks. Another distinction between networks according to their degree distribution is between *homogeneous* or *heterogeneous*. In the first case the functional form of $P(k)$ is light tailed, like a Gaussian or a Poissonian, while in the second case $P(k)$ is heavy tailed. The peculiarity of heavy-tailed distributions is that the average degree does not represent any special value for the distribution, because, even if a random choice will typically pick a node with low degree, the probability to extract a node with large degree is still significant. This property, in the absence of an intrinsic scale for the degree's fluctuations, defines the *scale-free* networks [25, 3]. As a matter of fact if the degrees distribution is power-law

$$P(k) = Bk^{-\gamma}$$

and, considering $2 < \gamma \leq 3$, the average degree is well defined and bounded as

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

while the second moment

$$\langle k^2 \rangle = \int_{k_{min}}^{\infty} k^2 P(k) dk$$

diverges, and the fluctuation of the degree, that in this case depends on the size of the system, is unbounded. The heterogeneity properties translates in a high level of degree fluctuations and in the absence a characteristic scale for the degree. A parameter to identify the scale-free behavior can be defined as

$$\kappa = \frac{\langle k^2 \rangle}{\langle k \rangle}$$

so that if $\kappa \gg \langle k \rangle$ the network is considered scale-free.

Scale-free networks are particularly suited to describe several real-world networks [31]. For example the presence of hubs, nodes with degree highly exceeding the average, in many real systems, is a clear manifestation of this property [22, 1].

As we will see in this dissertation, the difference between heterogeneous and homogeneous networks play a fundamental role in the studying of dynamical processes evolving on the network.

1.2.2. Small world. Travers and Millgram in the 1960s [119], in their experiment, studied how many people are needed to handout a letter passed from person to person to reach a given target. From the results we have the famous *six degrees of separation* theory, asserting that most people in the world are connected by short paths of length six.

The average shortest path length introduced before is an indicator of how far from each others nodes are in a network. In particular when $\langle d \rangle \sim \log N$ or shorter, the network has the *small-world* propriety.

Despite being a characteristic of random graphs, as we will see later, this property has been observed in many real networks and, it is associated with a certain level of clustering, as we will see in the Watts and Strograts model [132].

1.3. Models of Complex Networks

In this section we will present some of the generative models of static networks. All the possibility to create a static graph described here, relies on the definition of the particular choice for probability to form a connection between two nodes.

1.3.1. Random networks.

1.3.1.1. *Erdös and Rényi (ER) graph.* The most simple model for complex networks is the Erdös and Rényi (ER) random graph model [37]. We can define two generative processes for this model, forming two kind of ensemble $G_{N,K}$ and $G_{N,p}$ that are directly related to the canonical and gran canonical ensemble in statistical mechanics [18].

In the first, and original formulation, the pairs of nodes forming a number K of edges are randomly chosen among the total number of nodes N. In the second scenario the rule to build the network is that each of the $N(N-1)/2$ total number of edges is created with probability p. In particular the number of possible representation of $G_{N,p}$ is $2^{N(N-1)/1}$ considering that each edge either exists or not. This defines an ensemble $G_{N,p}$ of all the graphs for which the probability of having k edges is $p^k(1-p)^{K-k}$, consequently we can easily see that the probability that a random chosen node has degree k is given by the binomial distribution

$$P(k) = \begin{pmatrix} N-1 \\ L \end{pmatrix} p^k (1-p)^{N-1-k}$$

where L is the total number of links. For this model of random graphs many analytical results can be obtained in the thermodynamic limit $N \to \infty$, which can also be extended to the first model. In particular the average degree is constant,$\langle k \rangle = (N-1)p$, and the degree distribution becomes a Poissionian

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

One of the most important properties of an ER graph is the presence of a phase transition from low-density with few links and mostly small components, to high-density states in which a large fraction of vertexes are included in a single giant component. The threshold is determined by the critical probability $p_c = 1/N$ corresponding to the critical average degree $\langle k \rangle_c = 1$. For $p < p_c$, and large N the graph has no component of size grater then $\mathcal{O}(\ln N)$ with more than one cycle. For $p = p_c$ the largest component has size of order $N^{2/3}$, while for $p > p_c$ a component of size N exists. Considering as the order parameter the size of the largest component, this transition is of the second order, and it is in the same universality class of the mean field percolation phase transition.

In the ER graphs the average clustering coefficient is $\langle C \rangle = p = \langle k \rangle / N$, which means that it decreases with the graph size for fixed $\langle k \rangle$. Moreover the average shortest path length is given by

$$\langle d \rangle \sim \frac{\ln N}{\ln \langle k \rangle}$$

giving the ER graphs the small-world property.

1.3.1.2. *Configuration model.* The natural extension of the ER model is to consider instead of the Poisson distribution, an arbitrary distribution $P(k)$ for the degree. To this end we introduce the *configuration model* defined as the ensemble of graphs formed by configurations generated with the following recipe [80]. First we assign to each node i a degree k_i, representing the number of half-links, called *stubs*, choosen from a fixed degree sequence $D = \{k_i\}$, for $i = 1, ..., N$ such that

$$\sum_{i=1}^{N} k_i = 2L ,$$

where L is the number of links, and the fraction of nodes with degree k tends to $P(k)$ for large N. Second we join together pairs of stubs randomly accordingly to their degree.

This process generate with equal probability every possible graphs compatible with the fixed degree sequence, of which each configuration can be obtained in $\prod_i k_i!$ ways, considering that the number of permutation of the stubs of a node i is k_i. From the generating process two main limitation of the configuration model stand out: the sum of all the k_i must add up to an even number, and it is not possible to avoid the formation of self-edges and multiedges.

In this framework the probability p_{ij} that two nodes i and j are connected given by

$$p_{ij} = \frac{k_i k_j}{2L-1},$$

i.e. the probability for the node i to connect to a node j with degree k_j, $k_j/(2L-1)$, times the number of stubs k_i attached to i. Notice that in the limit of large m the probability becomes

$$p_{ij} = \frac{k_i k_j}{2L}.$$

Hence in this model a node has lower degree than its typical neighbor, describing the criterion of "I have less friends then my friends".

The configuration model allows analytical calculation for different quantity for instance the clustering coefficient is given by

$$C = \frac{\left[\langle k^2\rangle - \langle k\rangle\right]^2}{N\langle k\rangle^3}$$

which, as expected in a random graphs, is zero in the limit of $N \to \infty$. On the other hand considering a highly skewed degree distribution, the factor $\langle k^2\rangle / \langle k\rangle^2$ can be very large, and C is not completely negligible for large finite graphs.

Also the presence of a giant component can be evaluated using the Molloy-Reed criterion [75], i.e. consider a given degree distribution $P(k)$, when

$$G = \sum_k k(k-2)P(k) > 0$$

and the maximum degree is not too large, a giant component is most likely to appear. On the other hand if $G < 0$ and the maximum degree is not too large, the size of the largest component is $\mathcal{O}(k_{max}^2 \ln N)$.

Using the generating function formalism [82] Newman et al have also characterized the average shortest path length given by

$$\langle l\rangle = \frac{\ln(N/z_1)}{\ln(z_2/z_1)} + 1$$

in the limit of $N \ll z_1$ and $z_2 \ll z_1$, where z_s is the average number of neighbor at distance s.

Another property of the configuration model is that the probability of finding more than one path between any pair of nodes is $\mathcal{O}(N^{-1})$ in case of well-behaved distributions. If this property is fundamental for the solution of the model, on the other hand it also not not true for most real networks.

1.3.1.3. *Scale-free random graphs.* A very significant subclass of random graphs closer to real networks consists in the scale-free random graphs. There are many generative processes that lead to power-law degree distributed random graphs, we will introduce some of them here and talk more extensively later.

The Newman's approach is to use the generating functions method introducing a degree distribution $P(k) \sim k^{-\gamma}$ and finding the expression $C \sim N^{(3\gamma-7)/(\gamma-1)}$ for the clustering coefficient. In this case for $\gamma < 7/3$, we can say that there can be more than one edge between two nodes sharing a common neighbor, and hence C increases with the size of the system N. While for $\gamma > 7/3$ C tends to zero for large graphs.

It also possible to generate the graph assigning a uniform probability to all random graphs with a number of nodes k given by $N(k) = e^{\alpha}k^{-\gamma}$. In this case a giant component exists when $\gamma < \gamma_c \sim 3.47875..$ [18].

Another possible approach is the *fitness model* introduced by Caldarelli et al [24]. In this case we start with N isolated nodes, and consider each node i to have a fitness η_i, which is a real number drawn from the fitness distribution $\rho(\eta)$, then for every couple of nodes i and j a link is extracted with probability $p_{i,j} = f(\eta_i, \eta_j)$, where f is a symmetric function [108], and if it is constant we obtain the ER model. This model generate a power-law $P(k)$ for many fitness distribution, indeed if we consider a node with fitness η, its average degree is

$$k(\eta) = N \int_0^{\infty} f(\eta, \varphi) P(\varphi) d\varphi = NF(\eta)$$

and the degree distribution is given by

$$P(k) = \int d\eta P(\eta) \delta\left[k - k(\eta)\right] = P\left(\left[F^{-1}\left(\frac{k}{N}\right) \partial_k F^{-1}\left(\frac{k}{N}\right)\right]\right)$$

considering $F(\eta)$ to be monotonic in η.

Let's give an example considering $P(\eta) = e^{-\eta}$, and

$$f(\eta, \varphi) = \theta\left[\eta + \varphi - \kappa(N)\right]$$

where κ is a predetermined threshold and θ is the Heaviside function. This process generates a scale free network with degree distribution $P(k) \sim k^{-2}$ from a picked fitness distribution. In this case both the assortativity $k_{nn}(k)$ and the clustering spectrum $C(k)$ are power-laws.

1.3.1.4. *Watts and Strogatz.* Many real systems have both the small-world property and the high clustering coefficient, one of the simplest method to have them both was formulated by Watts and Strogatz (WS) [132]. The basic feature is a

rewiring procedure of the edges with a probability p. The generating process considers a ring of N nodes each symmetrically connected to its $2m$ nearest neighbors, considering that the total number of links is $L = mN$, each edge is then rewired with probability p or preserved with probability $1-p$, moving its endpoint to a new node randomly chosen from the graph, excluding multiedges or loops. Notice that if $p = 0$ we have a regular lattice with clustering coefficient $C = (3k-3)/(4k-2)$, while for $p = 1$ we reproduce a random graph with $k_{min} = m$, with distance of order $\log N/\log k$, and very low clustering coefficient $C \sim 2k/N$. Hence we can see p as tuning parameter for the randomness of the graph, keeping the number of edges constant, and for $0 < p < 1$ we generate graphs that have the two wanted characteristics. The clustering coefficient in this case is

$$C = \frac{3(k-1)}{2(2k-1)}(1-p)^3 .$$

Also the degree distribution is exactly computed and has the form

$$P(k) = \sum_{n=0}^{\min(k-m,n)} \left(\begin{array}{c} m \\ n \end{array} \right) (1-p)^n p^{m-n} \frac{(pm)^{k-m-n}}{(k-m-n)!} e^{-pm} \qquad \text{for } k \geq m$$

which in the limit of $p \to 1$ gives us the Poisson distribution as expected.

Regarding the average path length we can estimate a number of partial exact results, and some scaling results. In the limit of $p \to 0$ the typical path length is $l = N/4k$, while in the large p limit $l \sim \log N$ which is the small-world property. Barthélémy and Amaral [14] formulate the following scaling relation for l

$$l = \xi g(N/\xi)$$

where ξ is a correlation length that depends on p and $g(x)$ is an unknown scaling function that depends only on the system dimension and on the geometry of the graph, and in the two limit situation takes the values

$$g(x) = \begin{cases} x & x \gg 1 \\ \log x & x \ll 1 \end{cases}$$

they also showed that for small p, $\xi \sim p^{-\tau}$ where τ is a constant.

1.3.1.5. *Preferential attachment.* The preferential attachment model, formulated by Barabasi and Albert (BA) [4], was inspired by the formation of the World Wide Web and it is characterized by two main ingredients:

- *growth*: which means that starting with m_0 isolated nodes, at each time step $\Delta t = 1, 2, ..., N - m_0$ a new node j with $m \leq m_0$ number of links is added to the network and it will connect to a preexisting node i.
- *preferential attachment*: the probability Π for a new node to connect to a preexisting node i, depends on the degree k_i.

$$\Pi(k_i) = \frac{k_i}{\sum_j k_j} .$$

After t time steps the size of the network will be $N = m_0 + t$ and the number of edges will be mt.

The idea is that, as for the WWW, nodes with high degree form new link with higher rate than the ones with low-degree. In the long time limit this model produces a power-law degree distribution $P(k) \sim k^{-\gamma}$ with $\gamma = 3$. This result can be easily obtained considering that each time a new link attaches to a preexisting node i, the degree increases as follows

$$\partial_t k_i = m\Pi(k_i)\,,$$

where we have considered k_i as a continuous variable. The last expression can be written explicating the probability $\Pi(k_i)$, and considering the k limit becoming the following

$$\partial_t k_i(t) = \frac{m k_i(t)}{2mt + m_0 \langle k \rangle_0},$$

where $\langle k \rangle_0$ is the average degree of the initial m_0 nodes. Solving for k_i with the initial condition that each node starts at t_i with m edges we obtain that

$$k_i(t) \simeq m \left(\frac{t}{t_i} \right)^{1/2}.$$

In the large time limit the degree distribution is given by

$$P(k) = 2m^3 k^{-3}$$

while the clustering coefficient is given by

$$C = \frac{m}{8N} (\ln N)^2$$

and the average shortest path is

$$\langle l \rangle = \frac{\log N}{\log \log N}\,.$$

Notice that in the BA model the growth of the network is only in the generative process, and the result is a static network.

1.4. Random walk on static networks

We want to study the evolution of the number of elements passing through a node using the random walk formalism [68, 86].

In the simplest Markovian random walk formalism we study a diffusive process on an undirected, unbiased network. In a network with N nodes and adjacency matrix A_{ij}, a walker i at times t hops to one of its $k_i = \sum_j A_{ij}$ neighbors chosen with equal probability, at time $t+1$, hence we define the transition probability π_{ij} to go from i to j as

$$\pi_{ij} = \frac{A_{ij}}{k_i} = \frac{A_{ij}}{\sum_{j=1}^{N} A_{ij}}\,.$$

We want to study the evolution equation of the occupation probability distribution $p_i(t)$ for a node i to be visited at a certain time t, which in the discrete time model

is given by

$$p_j(t+1) = \sum_{i=1}^{N} \pi_{ji} p_i(t)\,. \tag{1.4.1}$$

When Eq. 1.4.1 holds also for $t = 0$, and considering $\Pi = \{\pi_{ij}\}$ as the transition matrix, we can write in the vector form as

$$\boldsymbol{p}(t+1) = \Pi \boldsymbol{p}(t)$$

The stationary distribution $\boldsymbol{p}^*$ is given by the fixed point solution of the equation such that $\boldsymbol{p}^* = \Pi \boldsymbol{p}^*$. For connected graphs containing at least one odd cycle, the Perron-Frobenius theorem guarantees the existence of $\boldsymbol{p}^*$such that

$$\lim_{t\to\infty} \Pi^t \boldsymbol{p}(0) = \boldsymbol{p}^*\,.$$

In this case all the occupation probability distributions converge to the stationary distribution. For undirected networks the stationary distribution is given by

$$p_i^* = \frac{k_i}{\sum_j k_j}$$

while if the network is also unweighted $p_i^* = k_i/K$, where K is the total number of links. In this case the stationary distribution must also follow the detail balance condition

$$p_i^* \pi_{ij} = p_j^* \pi_{ji}$$

which means that the flow of probability in each direction of the edge must be equal at the equilibrium.

We now want to focus on a method particularly useful in case of heterogeneous networks, assuming that all the nodes with the same degree are statistically equivalent. As we will discuss in chapter 3, in this approximation, nodes are characterized only by their degree. Let's define the occupation number W_i representing the number of walkers for the node i, and the total number of walkers $W = \sum_i W_i$. The degree block variable

$$W_k = \frac{1}{NP(k)} \sum_{i\in k} W_i$$

where $P(k)$ is the degree distribution, and the sum is over all the nodes with degree equal to k. Introducing the transition rate r for a node with degree k to go to a node of degree k', we can write the mean-field equation for the variation in time of the walkers $W_k(t)$ in each degree class given by

$$\partial_t W_k(t) = -rW_k(t) + k \sum_{k'} P(k'|k) \frac{r}{k'} W_{k'}(t)$$

where the first term account for the walkers moving out the node with rate r, and the second is the diffusive term describing the walkers that moving from the neighbors visit the node with degree k.

Because we are considering uncorrelated networks the conditional probability $P(k'|k) = k'P(k')/\langle k\rangle$, and studying the stationary condition $\partial_t W_k(t) = 0$ we obtain the solution

$$W_k = \frac{k}{\langle k\rangle}\frac{W}{N}\,.$$

The probability to find a diffusive walker in a node with degree k is consequently given by

$$p_k = \frac{W_k}{W} = \frac{k}{N\langle k\rangle}$$

hence it is more probable for a walker to diffuse into a highly connected node.

We will use this method to describe epidemic processes on the network in Chapter 3.

CHAPTER 2

Temporal networks

In many situations, the introduction of the topology is enough to give an insight on the mechanisms characterizing the system. However in most real situations connections between individuals evolve in time. When studying a dynamical process evolving on a graph, we can consider two opposite time-scale limits in which the static network framework is still a good approximation. The first one is when the network evolution is much slower than the dynamical process evolving on it. The second one is the opposite limit, in which the evolution of the network is much faster than the time-scales of the dynamical process evolution. In this last case, the dynamic unfolds on the annealed static network [114, 41], which is described by a mean-field version of the adjacency matrix, giving a good approximation the process.

However in most real systems, such as social systems, both the structure of the graph and the dynamical processes on it evolve on comparable time scales, and the previous two approximations don't apply. In this case it is necessary to introduce the time variable in the network definition introducing the *time-varying networks (TVN)* [85, 50, 49, 67].

Most of the properties valid for static networks can't be easily translated to the TVN. This chapter is organized as follows: in section 2.1 we will introduce the basic definitions of time-evolving networks, while in section 2.2 we will consider some of their statistical properties. In section 2.3 we will give some examples of TVN models, and, in particular, in section 2.4 we will focus on the activity-driven model which is at the center of this work.

2.1. Representing temporal network

We can describe static networks with either the adjacency matrix or the adjacency list representation. Also for temporal networks, there are many possible representations of the system. We will focus mainly on the *event-based,* and *snapshot* representations.

Let's consider a TVN formed by a set of nodes N during an *observation interval* $[0, T]$. In the event-based, we consider the TVN as an ordered set $\mathcal{C}$ of time-stamped links called *events* or *contacts* recorded in the observation interval. The

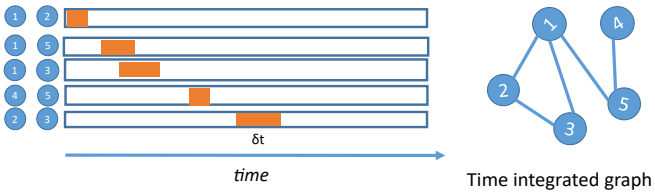

FIGURE 2.1.1. Event-based representation.

general definition of a contact between two nodes is given by the quadruplet

$$c_{ij}(t, \delta t) = (i, j, t, \delta t)$$

where $0 \leq t \leq T$ is the time of the interaction and δt is its duration, and if $c_{ij}(t, \delta t) = c_{ji}(t, \delta t)$ then the contact is *symmetric*. The duration of an event can be very long, and sometimes, as we will see later, can be broadly distributed, but in many situations is much smaller than the inter-event time and δt can be neglected, obtaining $c_{ij}(t) = (i, j, t)$. This approximation yields to a useful simplification for both numerical and analytical analysis of TVN.

The other representation consists in describing the TVN as a discrete time series of the network. In this case, we consider a *snapshot* G_t representing the instant configuration of the graph at time t. The system is then defined by the ordered sequence of snapshots $G_{[0,T]} = \{G(t_0), G(t_1), ..., G(T)\}$ called *time aggregated graph*, where T is the total number of time intervals considered. Each snapshot of the system at time $t \in [0, T]$ can be fully described by the *adjacency index* or *adjacency matrix* $A(t)$ at time t , whose elements are

$$a_{ij}(t) = \begin{cases} 1 & i \text{ and } j \text{ are connected at time } t \\ 0 & \text{otherwise} \end{cases}$$

The sequence of adjacency indexes $A = \{A(t_0), A(t_1), ..., A(T)\}$ forms the *time aggregated adjacency matrix* at time T. The snapshot model is a discrete time representation useful to develop a matrix-based model of temporal networks, and allows incorporating the time variable in the mathematical formulation through the definition of an *adjacency tensor*. With this representation, we can study the macroscopic properties of the TVN seen as a time-evolving structure.

Event-based representation at discrete time can turn in the snapshot representation without loss of information, on the other hand, the transformation of continuous time event in snapshots leads to a loss of temporal information due to the discretization process.

2.1.1. Measures. Walking in a static network from a node i to a node j, depends only on the existence of a set of links between the two nodes. On the other hand, in TVN, a walker must wait for the times of the intermediary nodes to be connected to move around the network. Moreover, the arrow of time establishes

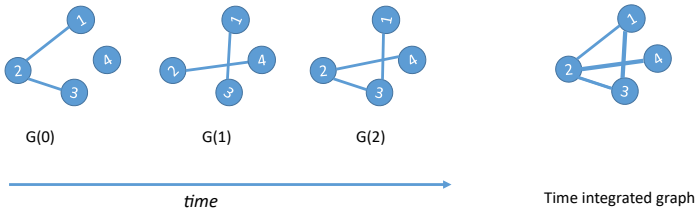

FIGURE 2.1.2. Snapshot representation.

a causality constraint, not allowing the use of past events. These properties make the definition of temporal walk very different from the static network framework.

In the contact-based representation a *temporal walk* from a node n_i to a node n_j is a contact sequence $[(n_i, n_1, t_1), ..., (n_{j-1}, n_j, t_j)]$ ordered in time, $t_i < t_{n_1} < ... < t_j$. Similarly, in the snapshot representation, the contact (n_l, n_{l+1}, t_{l+1}) is interpreted as the link of the snapshot $G(t_l)$ such that $a_{n_l, n_{l+1}}(t_{l+1}) \neq 0$.

If a walker visits each node between n_i and n_j, we have a *temporal path* or *time-respecting path* from n_i to n_j. Notice that in the aggregated network, temporal walks and paths are always overestimated, because of the presence of all the links created during a time interval.

A node n_j is *reachable* from a node n_i if there is a temporal walk between them. The set of nodes that can be reached starting from a node n_i is called *set of influence* of n_i. We can also define the *reachability ratio* as the average fraction of nodes in the set of influence of all nodes and the set of vertexes that reach n_i through temporal paths in a certain time window, as the *source set of* n_i. The ordered nature of time-respecting paths set a fundamental difference with static networks, hence in TVN the reachability is not a transitive property.

2.1.2. Metric. Most TVN contacts have neither the symmetry nor the transitivity properties, hence we can not define a proper metric, nonetheless we can introduce the concept of distance either from the topological or from the temporal perspective, obtaining three main definitions valid for both the event-based and the snapshot representations. Let's consider a temporal path $[(n_i, n_1, t_1), ..., (n_{j-1}, n_j, t_j)]$ from n_i to n_j

The *topological shortest distance* is given by the minimum number of hops h necessary to go from n_i to n_j along the temporal path

$$d_{short}(n_i, n_j, t) = \min\{h : t_1 \geq t\}$$

The *temporal shortest path* or the *foremost distance* is defined as the minimum amount of time to travel between two nodes.

$$d_{fore}(n_i, n_j, t) = \min\{t_j - t : t_1 \geq t\}$$

Another distance in the time domain, favors the temporal paths with shortest duration independently from the starting time

$$d_{travel}(n_i, n_j, t) = \min\{t_j - t_1 : t_1 \geq t\}$$

For each distance $d_{def}(n_i, n_j, t)$, where *def* stands for the three definitions, we can introduce the average distance over all the pairs of nodes given by

$$L_{def} = \frac{1}{N(N-1)} \sum_{i \neq j} d_{def}(n_i, n_j, t)$$

and the *diameter* of the TVN

$$D_{def} = \max_{n_i n_j} d_{def}(n_i, n_j, t)$$

Notice that if a point is not reachable then $d_{def}(n_i, n_j, t) = \infty$, and to avoid the divergence of L we can also define the *temporal global efficiency* as

$$E = \frac{1}{N(N-1)} \sum_{i \neq j} \frac{1}{d_{def}(n_i, n_j, t)}$$

.

We can also measure the recency of an information exchange between two nodes, or the *information latency* of n_i respect to n_j at time t.

The temporal view $\phi(n_i, n_j, t)$ that a node n_i has about a node n_j at time t is the latest time $t' < t$ such that a temporal path starting from n_j at t' reaches n_i before t. The information latency is given by the difference $t - \phi(n_i, n_j, t)$, and represents the time passed since the most updated information about n_j was obtained by n_i. Considering the information passed to n_i from all the network, we can define the *vector clock* $\phi(n_i, t) = [\phi(n_i, n_j, t)]_{j=1,..,N}$ as the list of the most up to date information that n_i has about all the network [61].

2.1.3. Components. The definition of temporal paths affects also the concepts of *connectivity* and *components* in TVN. For static networks, components are defined as the sets of nodes for which a path always exists, moreover nodes are either connected or not. In particular, in undirected graphs, the connectivity is a reflexive, symmetric and transitive property. On the other hand in temporal graphs the symmetry property doesn't hold anymore. In analogy with directed static graphs, two nodes n_i and n_j in a TVN are *strongly connected* if the temporal paths connecting n_i to n_j and vice versa are directed, while they are *weakly connected* if the temporal paths connecting them are undirected [84]. By using the reachability and the definitions of connectivity it is possible to introduce different type of *temporal components*: $\mathrm{OUT}_T(n_i)$ ($\mathrm{IN}_T(n_i)$) is the temporal out-component (in-component) of the node n_i, i.e. the set of nodes that can be reached from i (from which i can be reached) in the TVN, while the strongly connected $\mathrm{SCC}_T(n_i)$ (weakly connected $\mathrm{WCC}_T(n_i)$) component of the node n_i is the set of nodes from

which i can be reached and which can be reached (undirectelly) from i in the TVN [84].

2.1.4. Centrality. Also for centrality measures, we can consider the time-dependent and the time-independent definition.

Two examples of centrality measures based on time-dependent distance are:

- The *temporal closeness centrality*
$$C_C(n_i, t) = \frac{1}{N-1} \sum_{i \neq j} \frac{1}{d_{def}(n_i, n_j, t)}$$
measure how quickly an edge reach other edges on average [88].
- The *temporal betweenness centrality* takes in to account the fraction of shortest time-respecting paths that pass through a focal vertex
$$C_{n_i}^B(t) = \frac{1}{(N-1)(N-2)} \sum_{n_i \neq n_j} \sum_{\substack{k \neq j \\ k \neq i}} \frac{U(n_i, t, n_j, n_k)}{\sigma_{jk}}$$
where σ_{jk} is the number of shortest path from n_j to n_k, and $U(n_i, t, n_j, n_k)$ is the number of temporal shortest paths from n_i to n_j in which node i is traversed from the path in the snapshot t [33].

2.2. Statistical properties of TVN

Let's consider the contact-based representation of the network considering null the duration of the events. The event sequence associated either to each vertex or each link, is given by $\{t_1, ..., t_e\}$, where now the time t_i and the number of evens e are properties of a single node or link and not of the whole network. In this section, we will see some of the statistical features of this set, which will help us to define some models of temporal networks.

2.2.1. Burstiness. Many events, especially in human behavioral networks, occur in a short period of time generating a burst, and tend to be separated by long time intervals. An event sequence generated from a Poisson process lacks of this property. In fact, if we define the inter-event of a node i, $\tau_i = t_{i+1} - t_i$, as the time between two consecutive events of an edge, then a Poisson generating process give the inter-event distribution

$$\psi(\tau) = \sigma e^{-\sigma\tau}$$

whit σ a parameter. On the other hand real data analysis shows that most human processes have power-law distributed inter-events, i.e.

$$\psi(\tau) \sim \tau^{-\alpha}$$

The latest case defines one of the fingerprints of the bursty behavior of real social systems [42, 128, 11].

To measure the burstiness of a process we can introduce the *coefficient of variation* defined as the standard deviation of $\{\tau_i\}$ divided by its average

$$CV = \frac{\sqrt{\frac{1}{e-1}\sum_{i=1}^{e-1}(\tau_i - \langle\tau\rangle)^2}}{\langle\tau\rangle},$$

where $\langle\tau\rangle = (e-1)^{-1}\sum_{i=1}^{e-1}\tau_i$. For a Poisson distribution of inter-event times $CV = 1$, while for periodic sequence of events $CV = 0$. The normalized form of the CV is called *burstiness* B

$$B = \frac{CV - 1}{CV + 1}$$

which varies between -1 and 1. For a Poisson process $B = 0$, while for a periodic sequence $B = -1$. Extremely burtsy behaviors have $B = 1$ ($CV \to \infty$).

Another statistical measure is the *local variation* LV defined as

$$LV = \frac{3}{e-1}\sum_{i=1}^{e-1}\left(\frac{\tau_i - \tau_{i+1}}{\tau_i + \tau_{i+1}}\right)^2$$

While the CV is greatly affected by extreme large inter-event time, the LV remains confined in the interval $[0, 3)$ [106].

2.3. Models of temporal networks

In the last few years the number of studies of dynamical processes on temporal networks has significantly grown. A variety of dynamical processes can be well described using this framework. In this section, we will mention some of the most relevant further looking into the activity-driven framework which will be extensively used in the description of epidemic processes on TVN.

Social group dynamics models. In this model links represent social ties, as facing another individual, and it is possible to write a master equation representing the changing of the number of people in a group of a given size [115, 42]. This approach describes the observation "the more isolated is an individual from a group the less it is probable that it will interact with the group, and the longer it is interacting with the group the more it is likely to stay in the group". In this case the TVN representation more suited to model the process is the interval graph's formalism.

Randomized reference models. To assess the importance of a topological feature in empirical networks analysis, it is generally useful to have a reference model to compare the data with. In this type of models event sequences of the original network are randomly shuffled in a fashion that removes correlations in the time domain. Considering the variety of possible temporal correlations and time scales for different systems, there are many way of achieving this purpose,

for example switching off selected types of correlation can help to understand their contribution to the empirical network. Because of their purpose of annihilate time correlation these models are also called temporal *null models*. Considering a contact sequence TVN, we can define some of these models depending on the type of randomization process. If we go over all the edges sequentially and randomly substitute them with another pair, following some restrains, then we have the randomized edges model (RE). Another option would be to randomly permute the contact times while keeping the graph structure and the number of contacts intact, in this case we have the randomly permuted times model (RP). We can also do both the randomization described before, destroying all correlations except for patterns in the contact rate. Instead of keeping the set of times of the original contact sequence just permuting them as in the RP, one could assign a random time in the observation time window of the original empirical data, still conserving the graph structure and the total number of events for each link. It is also possible to randomize the contacts between the edges (RC).

For more details on the topic it is possible to consult the Ref. [40].

2.4. Activity-driven network

Especially in social systems, interactions among individuals rapidly change in time, and the time scale of the evolution of the network is comparable to the time scale of the dynamical process evolving on it. In this case we have seen that the snapshot representation is well suited to describe the system, and the topological properties can be captured by the time-integrated view of the network. The activity-driven model [94] belongs to this category of TVN, and it is the result of empirical studies of human activity from large data sets. The main concept is to embed the dynamical feature of the system in the node itself, instead of studying the links and the topology of the network.

In this model to each vertex v_i is associated an activity potential a_i which is a random variable, extracted from a given distribution $F(a)$, representing the tendency of the node to form a certain arbitrary chosen number m of links at time Δt. Hence in each snapshot the node v_i is active with probability $a_i \Delta t$, and the activity can be considered as a clock determining the temporal interaction patterns of each node. In the limit of $\Delta t \to 0$ activation of a node follows a Poisson process.

The activity-driven model is simply enough to allow analytical investigation of different properties of the structure and the dynamic of the network, and it is also able to reproduce the heterogeneity of many real systems.

Generating process and properties. The system starts with a network with N nodes, each characterized by the activity a_i distributed according to $F(a)$. The generative process, considering that all interactions have a constant duration, is the following.

- At each time step t the snapshot G_t starts with all the nodes disconnected.
- With probability $a_i \Delta t$ each vertex activates and connects m edges to m randomly selected different nodes. The nodes not activated can still receive connections from active nodes.
- At the subsequent time step $t + \Delta t$ all the edges in the network G_t are deleted.

At this level the model is random and Markovian, hence it has no memory of the previous time steps, and the interaction between the nodes are randomly chosen. The activity distribution $F(a)$ completely determines the topological and evolutionary properties of the network.

At first we can consider the properties of each snapshot. The number of active links per unit time is $L_t = mN \langle a \rangle$, so that the average degree per unit time is $\langle k \rangle = 2m \langle a \rangle$, where the two is the consequence of the undirected nature of each link and $\langle a \rangle = \int aF(a)da$ is the average activity. Snapshots are generally sparse networks formed by stars with degree $k \geq m$.

Time integrated network. We are also interested in the properties of the time integrated network $G = \bigcup_{t=0}^{T} G_t$ defined by the union of all the snapshots generated in $T+1$ time steps. This network is generally dense and, if T is long enough, we obtain a complete graph. If we consider the integrated network normalized by the number of snapshot, i.e. the total time, then its adjacency matrix is given by

$$A_{ij} = \frac{m(a_i + a_j)}{N}$$

neglecting the probability for an active node to contact another active node. From this relation we can find the average degree of the integrated network for the node i given by

$$\overline{k_i} = \sum_{j=1}^{N} A_{ij} = m(a_i + \langle a \rangle) \tag{2.4.1}$$

From the equation 2.4.1, noticing the monotonic relation between the degree and the activity, the following expression for the degree distribution holds $\rho(k)dk = F(a)da$. Hence, taking $m = 1$, in the limit $N \gg T \gg 1$ and $T^2 \gg k \gg 1$

$$\rho(k) \sim \frac{1}{T} F\left(\frac{k}{T} - \langle a \rangle\right).$$

The degree distribution of the time integrated network has the same form of the activity distribution $F(a)$.

From the analysis of big data sets [44, 94] it is possible to show that most human activities are heterogeneous, and characterized by a power-law activity distribution with exponent ν

$$F(a) = \frac{1 - \nu}{1 - \epsilon^{1-\nu}} a^{-\nu}$$

where $a \in [\epsilon, 1]$, $0 < \epsilon \ll 1$ is a cutoff introduced to avoid divergences of the distribution in the limit of zero activity. In this case the degree distribution is

$$\rho(k) \sim k^{-\nu}$$

On the other hand if $F(a) = \delta_{a,a_0}$ then the asymptotic form of the degree distribution is $\rho(k) \sim \delta_{k,k_0}/T$, while the exact form is a Poisson distribution centered at $2Ta_0$.

Starnini and Pastor-Satorras [112] noticed also that, unless $F(a)$ is a delta function, the average degree correlation for integrated network is given in the limit of large k by

$$\frac{k_{nn}^T(k) - 1}{T} \simeq 2\langle a \rangle + \sigma_a^2 \left(\frac{k}{T}\right)^{-1}$$

where $\sigma_a^2 = \langle a^2 \rangle - \langle a \rangle^2$ is the variance of the activity. For delta distributed activity, on the other hand

$$k_{nn}^T(k) \simeq 1 + 2Ta_0$$

Hence, for non constant activity distribution the time integrated network shows a disassortative mixing behavior, at odds with real social networks, which are considered assortative. Notice, however, that in case of power-law distributed activities with small variance σ_a (of order $\epsilon^{\nu-1}$ for $\nu < 3$ and order ϵ^2 for $\nu > 3$), the network can be considered approximately uncorrelated.

In the limit of large k it is also possible to compute the clustering coefficient of the integrated network which in its asymptotic form is given by

$$\frac{\bar{c}(k)}{T} \simeq \frac{2\langle a \rangle}{T} + \frac{2\sigma_a^2}{N}\left(\frac{k}{T}\right)^{-1}$$

which is small compared to random networks.

We can also study the eigenspectrum of the time integrated adjacency matrix [110]

$$A\boldsymbol{u} = \Lambda \boldsymbol{u}$$

and assuming that the eigenvector has the property $\sum_{i=1}^N u_i = 1$ we can obtain the eigenvalues

$$\Lambda_\pm = m\left(\frac{\sum_{i=1}^N a_i}{N} \pm \sqrt{\frac{\sum_{i=1}^N a_i^2}{N}}\right)$$

where the $N-2$ other eigenvalues are zero. As we will see in the next chapter the largest eigenvalue Λ_+ explains also the analytical result obtained for the epidemic threshold in the activity driven model.

.

2.4.1. Master equation. To study the characteristics of the time integrated network, we can also use the master equation formalism, considering the evolution of the probability $P_i(k,t)$ that a node with activity a_i has degree k at time t [120].

We will extensively use this framework to add different levels of complexity to the activity-driven model.

2.4.1.1. *Simple activity-driven model.* In the simple activity-driven model, considering the approximations where $a_i \ll 1$ so that only one site can be active between two consecutive times, and that $1 \ll k_i \ll N$, the discrete-time equation for $P_i(k,t)$ is

$$\begin{aligned} P_i(k,t+1) = \\ a_i \frac{N-k}{N} P_i(k-1,t) + a_i \frac{k}{N} P_i(k,t) - P_i(k-1,t) \sum_{j \nsim i} a_j \sum_h \frac{P_j(h,t)}{N} + \\ P_i(k,t) \sum_{j \nsim i} a_j \sum_h \frac{(N-1)P_j(h,t)}{N} + P_i(k,t) \sum_{j \sim i} a_j + P_i(k,t) \left(1 - \sum_j a_j\right) \end{aligned} \tag{2.4.2}$$

The first term is the probability that the site i is active and a new edge is added to the system. The second term is the probability that the site i is active but connects to an already linked site. The symbol $\sum_{j \nsim i}$ represent the sum over the nodes j that are still not connected to i. Hence the third and the fourth terms are the probabilities that one of these nodes j is active and either contact or not i respectively. The fifth is the probability that one of the nodes already connected to i ($j \sim i$) is active and no new link is added to i. The last term is the probability that all the nodes are inactive at time t.

We can neglect the second term for $k \ll N$ and obtain the equation

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

In this approximation, considering that $\frac{1}{N} \sum_{j \nsim i} a_j = \langle a \rangle$, the equation in the continuous time limit for $P_i(k,t)$ is

$$\partial_t P_i(k,t) = (a_i + \langle a \rangle) \left(-\partial_k P_i(k,t) + \partial_k^2 P_i(k,t)\right)$$

In this case the solution is given by

$$P_i(k,t) = [2\pi(a_i + \langle a \rangle)t]^{-1/2} \exp\left(-\frac{(k - a_i + \langle a \rangle t)^2}{2t(a_i + \langle a \rangle)}\right)$$

In the long time regime, this equation reduces to a delta function

$$P_i(k,t) = \delta(k - a_i + \langle a \rangle t)$$

and the average degree of the time integrated network for a node of activity a at time t is given by

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

Noticing the monotonic relation between the degree and the activity, also the degree distribution is a power law with the same exponent, $\rho(k) \sim k^{-\nu}$.

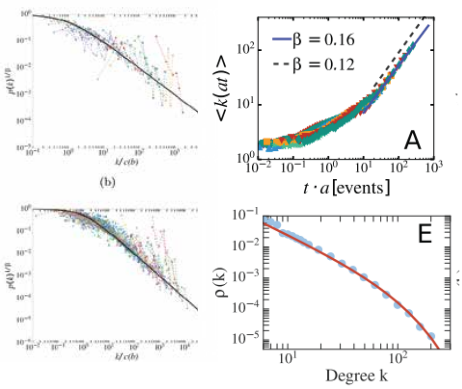

FIGURE 2.4.1. On the left the experimental curves of the reinforcement probability $p(k)$, from the PRE and the MPC datasets. On the right, from the top the measure of the average degree $\langle k(at)\rangle = \overline{k(at)}$ for different activity classes from the PRB dataset, below the degree distribution for the PRA dataset [120].

2.4.1.2. *Memory process.* In most real systems concerning human interaction we have memory processes, representing the tendency for an individual to reconnect with a previously contacted node. In TVN this phenomenon is particularly relevant. In fact because of the time evolution of the contacts, the reinforcement process strongly affects the structure of the network. In their empirical study of the datasets from American Physical Society, Twitter Mention Network and Mobile Phone Networks, Ubaldi et al [122] measured the probability $p(k_i)$ for a node i with a history of k_i number of contacts to establish a new link, finding the expression

$$p_i(k_i) = \left(1 + \frac{k_i}{c_i}\right)^{-\beta} \tag{2.4.3}$$

where c is the characteristic limit for an individual to maintain multiple contacts, β is the memory parameter and k_i is the degree of the node i in the time integrated graph (2.4.1). In this framework when a node is active, it contacts a new randomly chosen node with probability $p(k)$ or randomly connects to a node in its history of contacts with probability $1 - p(k)$.

Introducing the memory process in the equation 2.4.2 the equation for the $P_i(k,t)$ is given by

$$
\begin{aligned}
P_i(k,t+1) = \\
P_i(k-1,t)\left[a_i p_i(k-1) + \sum_{j\nsim i} a_j \sum_h \frac{p_j(h)P_j(h,t)}{N-h}\right] + \\
P_i(k,t)\left[a_i\left[1-p_i(k)\right] + \sum_{j\nsim i} a_j \sum_h \left(1 - \frac{p_j(h)P_j(h,t)}{N-h}\right)\right] + \\
P_i(k,t)\left(1 - \sum_j a_j\right)
\end{aligned}
\tag{2.4.4}
$$

The node i can go from $k-1$ to k contacts in two ways described by the second line of the equation. The first is for i to be active and contact a new node with probability $a_i p_i(k-1)$, and the second is for i to be contacted by an active node j never contacted before. Similarly the third line considers that the node i does not change degree k either because it calls an already contacted node or because the non contacted nodes call other nodes in the network. The last line describe the situation in which no nodes in the network activate. Using the expression 2.4.3 we can write the equation as

$$
\begin{aligned}
P_i(k,t+1) - P_i(k,t) = \frac{a_i c_i^\beta}{(k-1-c_i)^\beta} P_i(k-1,t) - \frac{a_i c_i^\beta}{(k+c_i)^\beta} P_i(k,t) \\
- \left(P_i(k,t) - P_i(k-1,t)\right) \sum_{j\nsim i} a_j \sum_h \frac{c_j^\beta P_j(h,t)}{(N-h)\,(h+c_j)^\beta}
\end{aligned}
$$

and applying the same approximation of the memory-less case we obtain the equation for the continuous time limit given by

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

where $\rho(c_j|a_j)$ is the probability for a node j of activity a_j to have memory constant c_j. In this case the solution for the master equation for the probability $P_i(k,t)$ for a node i with activity a_i to have a degree k at time t, in the long time limit and for $k \gg N$, is given by

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

Hence the average degree of the time integrated network for the nodes of activity a at time t is

$$
\overline{k(a,c,t)} \propto C(a,c)t^{\frac{1}{\beta+1}}
$$

where $C(a,c)$ is a constant depending on the activity which follows the recurrence relation

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

As we will see in the fourth chapter it is possible to consider $c=1$ without loosing generality. In the presence of the memory process the relation between the activity and the degree is given by $k \propto a^{1/(\beta+1)}$, and to find the degree distribution, we can use the relation

$$\rho(k) \sim k^\beta F\left(k^{1+\beta}\right) dk$$

In the special case of a power law activity distribution $F(a) \sim a^{-\nu}$, we have that

$$\rho(k) \sim k^{-[(1+\beta)\nu-\beta]}$$

We will extensively use these results in the last chapter where we will use this model to study two epidemic processes on the activity-driven network with memory. Moreover this model can be extended to multiple β as showed in [120].

2.4.1.3. *Burstiness.* The activity-driven model considered up to now is Poissonian. Now we want to study a non-Poissonian AD process (NoPAD) introducing the burstiness in the system [74, 32]. In this framework to each individual i is associated a time dependent activity $a_i(t)$ that the node is active for the first time at time t. The activation of each node is regulated by an inter-event or waiting time distribution

$$\Psi_i(w) = a_i(t) \exp\left[-\int_0^w a_i(w')dw'\right]$$

assuming that receiving a call from another individual doesn't affect the time between two activation events. In the limit of constant activity the model coincides with the simple AD.

Because the activities depend on the nodes, so must the waiting times, and we need to introduce a node dependent parameter ξ_i such that $\Psi_i(w) = \Psi(w,\xi_i)$.

The degree distribution of the time integrated network can be found using the hidden variable formalism [112] defining

$$\rho(k) = \sum_{\vec{h}} H(\vec{h}) g(k|\vec{h})$$

where $H(\vec{h})$ is the distribution of the hidden variable $\vec{h}$, and $g(k|\vec{k})$ is the conditional probability that a node with hidden variable $\vec{h}$ has degree k. Introducing the connection probability $\Pi(\vec{h}_i, \vec{h}_j)$ to create a link between the nodes i and j it is possible to write the equation for the generating function $\hat{g}(k|\vec{h})$ of the conditional probability

$$\ln \hat{g}(z|\vec{h}) = N \sum_{\vec{h}'} H(\vec{h}') \ln\left[1-(1-z)\Pi(\vec{h},\vec{h}')\right]$$

The hidden variable is vectors $\vec{h} = (r, \xi)$, where r is the number of activation that for a node with time t and heterogeneity ξ, is distributed according to $\chi_t(r|\xi)$. Restricting the analysis to the time interval $[0, t]$, Moinet et al found the approximated solution for the degree distribution at time t given by

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

where $\eta(\xi)$ is the distribution of the heterogeneity ξ of a node and $\langle r \rangle_t$ is the average number of activation of the node at time t.

Empirical findings suggest to consider the special case in which the distribution for the waiting times is

$$\Psi(w, \xi) = \alpha\xi(\xi w + 1)^{-(1+\alpha)}$$

where $0 < \alpha < 1$ is the exponent of the distribution. When the heterogeneity parameter is broadly distributed, $\eta(\xi) \sim (\xi/\xi_0)^{-b-1}$ $(b > \alpha)$, and in the limit of $k \gg (\xi_0 t)^{\alpha}$ the degree distribution is

$$\rho_t(k) = (\xi_0 t)^b (k - \langle r \rangle_t)^{-1-(b/\alpha)}.$$

In this framework the distribution of number of activation events $\chi_t(r|\xi)$ affects the topological properties of the graph, for heavy-tailed waiting times distributions, suggesting aging effects affecting the model. In fact, as showed in Ref. [74, 32], introducing the aged degree distribution $\rho_{t_a,t}(k)$, where t_a is the aging time, the average degree of the network integrated from time t_a depends on t_a and is given by

$$\langle k \rangle_{t,t_a} \sim (t_a + t)^{\alpha} - t_a^{\alpha}\,.$$

Notice that in the limit $t \gg t_a$ the average degree is

$$\langle k \rangle \sim t^{\alpha}$$

and the aging effects can be neglected. On the other hand for $t \ll t_a$ the average degree depends only on the aging time t_a

$$\langle k \rangle_{t_a,t} \sim t^{\alpha-1}\,.$$

The prediction of NoPAD model are compatible with empirical data gathered from the scientific collaboration network in PRL Society [107]

2.4.1.4. *Burstiness and memory.* Considering both the burstiness and the memory [23] processes the generative model of the network is the following:

(1) The first activation time τ_i for each node is extracted from the $\Psi(\tau, \xi_i)$ before starting the network evolution
(2) The time t is set on the node j with the smallest activation time $t = \tau_j$.
(3) The site j then contacts a new node with probability $p_j(k_j)$, or a site in its history of contacts with probability $1 - p_j(k_j)$. In the last case the integrated degree k_i of all the nodes i remains the same.
(4) A waiting time w_j is drawn from $\Psi(w, \xi_i)$ and τ_j is updated to $w_j + \tau_j$.

(5) Return to step 2

To obtain analytical results, Ubaldi et al, considered the approximation in which individuals can only contact other nodes and never be contacted. In this case they consider the evolution of a single agent 0 with its waiting time distribution $\Psi(w, \xi_0)$ and memory function $p(k)$, and study the master equation for the probability $Q(k,t)$ that the individual makes a connection at time t and after that has degree k. The $P(k,t)$ then is obtained integrating over the time and the waiting time

$$P(k,t) = \int_0^t dt' Q(k, t-t') \int_{t'}^{\infty} dw \Psi(w, \xi_0)$$

In general the results for the $P(k,t)$ depends on the average inter-event time of $\Psi(w, \xi_0)$, and on the asymptotic behavior of $\Psi(w, \xi_0)$ for large w. In particular there are three intervals of interest for the exponent α that leads to different $P(k,t)$ results [23].

$$P(k,t) \simeq \begin{cases} \frac{1}{(t/w_0)^{\frac{\alpha}{1+\beta}}} f_{\alpha\beta}\left(A'_{\alpha\beta} \frac{k}{(t/w_0)^{\frac{\alpha}{1+\beta}}}\right) & \alpha < 1 \\ \frac{1}{(t/w_0)^{\frac{1}{\alpha}-\frac{\alpha}{1+\beta}}} f_{\alpha\beta}\left(A'_{\alpha\beta} \frac{k - v(t/w_0)^{\frac{1}{(1+\beta)}}}{(t/w_0)^{\frac{1}{\alpha}-\frac{\alpha}{1+\beta}}}\right) & 1 < \alpha < \frac{2\beta+2}{\beta+1} \\ \frac{1}{(t/w_0)^{\frac{1}{2(1+\beta)}}} \exp\left(-A'_{\beta} \frac{\left(k - C_\beta (t/w_0)^{\frac{1}{(1+\beta)}}\right)^2}{(t/w_0)^{\frac{1}{1+\beta}}}\right) & \alpha > \frac{2\beta+2}{\beta+1} \end{cases}$$

where $f_{\alpha\beta}$ is a non-Gaussian scaling function, v is the drift velocity of the peak of the distribution. $A_{\alpha\beta}$, A_β and C_β are constant depending on the parameters β and α.

The average degree then can be written as

$$\overline{k(t)} = \begin{cases} t^{\frac{\alpha}{1+\beta}} & \alpha < 1 \\ t^{\frac{1}{1+\beta}} & \alpha > 1 \end{cases}$$

The equation for the degree distribution can be evaluated at fixed time considering

$$\rho(k) = \int F(a_i) P(a_i, k, t) da_i$$

When the activity distribution has a power-law decay, the degree distribution is given by

$$\rho(k) \sim \begin{cases} k^{-\left(\frac{1+\beta}{\alpha}(\nu-1)+1\right)} & \alpha < 1 \\ k^{-((1+\beta)\nu-\beta)} & \alpha > 1 \end{cases}$$

2.4.1.5. *Attractivness ADA.* A further extension of the simple activity-driven networks was introduced by Pozzana et al [98]. In their model they include the characteristic of social systems to distinguish between more or less popular individual. The main idea is that a node i might be more popular then the others introducing the concept of *attractivness* b_i. In this framework, when a node i is active it will target a node j with a probability depending on the j's attractiveness b_j.

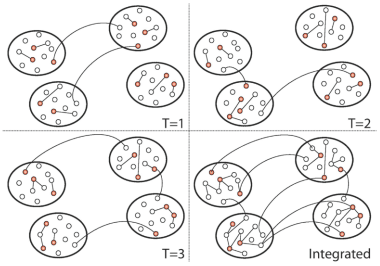

FIGURE 2.4.2. Schematic representation of the model. Straight lines represent the contact in the same community, arch represent the connections between communities. The active nodes are colored in red. [77]

The distributions of the activity $F(a)$ and of the acttractiveness $G(b)$ can be either uncorrelated, or correlated affecting differently the dynamical processes running on the network.

We will see in the next chapter how this model can be used to study epidemic processes.

Modular activity driven. Datasets analysis stressed out the organization of real networks in communities or modules, where the density of connection is much larger than the density of links between communities. To include this feature in the activity driven model Nadini et al [77] considered a network with N nodes and tunable *modularity*, where the size s of the communities is drawn from a given distribution $P(s)$. The heterogeneity of the modules size grasped from real data, suggested a heavy-tailed form for $P(s) \sim s^{-\omega}$, with $s \in [s_{min}, \sqrt{N}]$.

In this framework each node is progressively assigned to a module of size s extracted from $P(s)$ and the generative process of the ADM network is the following:

- At each time the graph starts with N disconnected nodes.
- Each node is activated with probability $a_i \Delta t$ and creates m links (m can be set to one).
- Each link connects randomly within the community with probability μ or outside the community with probability $\mu' = 1 - \mu$.
- At $t + \Delta t$ all links are deleted.

Where Δt is the constant duration of the interactions which can be set to one.

They considered the master equation for the probabilities $P_c(s, k_c)$ and $P_o(s, k_o)$, which are respectively the probability for a node of activity a_i to belong to a community of size s and have in-degree k_c or out-degree k_o respect to the community at time t. In the limit of large time $t \gg 1$ and large degree $k \gg 1$, they found the

analytic solution for the master equation for both the probability distributions.

$$P_c(s,k_c) \propto \begin{cases} \exp\left[-\frac{(k_c-\mu(a-\langle a\rangle)t)^2}{2\mu(a+\langle a\rangle)t}\right] & t \ll \tau_c(s) \\ \delta(k_c-(s-1)) & t \gg \tau_c(s) \end{cases}$$

$$P_o(s,k_o) \propto \exp\left[-\frac{(k_o-\mu'(a-\langle a\rangle)t)^2}{2\mu'(a+\langle a\rangle)t}\right] \quad \forall t$$

Notice that, while the in-community probability P_c depends on the size of the community, the out-community probability P_o doesn't.

Considering that $k_c + k_o = k$ they determined the total probability distribution as

$$P(s,k) = \int_0^k P_c(s,k_c)P_0(k-k_c)dk_c$$

In this framework they study the evolution of the average in-community degree of each node given by

$$\overline{k_c(a,s,t)} = (s-1)\left[1-\exp\left(-\frac{t}{\tau(a,s)}\right)\right]$$

where $\tau(a,s)$ is the characteristic time that it takes for the degree $k_c(a,s,t)$ to become maximal, i.e. $k_c(a,s,t) \sim s-1$. On the other hand the out-community average degree is given by

$$\overline{k_o(a,t)} = \mu'(a+\langle a\rangle)$$

The total average degree is then given by

$$\overline{k(a,s,t)} = \begin{cases} (a+\langle a\rangle)t & t \ll \tau(a,s) \\ \mu'(a+\langle a\rangle)t + s - 1 & t \sim \tau(a,s) \\ \mu'(a+\langle a\rangle)t & t \gg \tau(a,s) \end{cases}$$

The long time evolution of the degree is linear in time, hence for power-law activity distribution $F(a) = a^{-\nu}$ they obtained power laws degree distribution $\rho(k)$ with the same exponent ν.

2.4.2. Random walks on activity-driven model. The study of random walks in TVN is a core concept for both analytical and computational models of many real-world dynamical processes, that mostly evolve on temporal scale-free networks. To study this formalism on the activity-driven network [96] we introduce the propagator $\Pi_{i\to j}^{\Delta t}$ of the random walk as the probability that a walker moves from the node i to the node j in the time interval Δt, then we can write the master equation for the probability $P_i(t)$ that the walker is in the node i at time t

$$P_i(t+\Delta t) = P_i(t)\left[1-\sum_{j\neq i}\Pi_{i\to j}^{\Delta t}\right] + \sum_{j\neq i} P_i(t)\Pi_{i\to j}^{\Delta t}$$

Considering only the first order terms in Δt, the expression for the propagator is

$$\Pi_{i\to j}^{\Delta t} \simeq \frac{\Delta t}{N}(a_i + m a_j)$$

where m, as usual, is the number of links fired by an active nodes at each time step.

For the activity-driven framework, as we will see in detail later, it is sometimes convenient to consider groups of the same activity class a assuming that they are statistically equivalent in the limit of $N \to \infty$. If we define $W_a(t) = [NF(a)]^{-1} W \sum_{i \in a} P_i(t)$ as the number of walkers in the same activity class a at time t, we can write in the continuous time limit, $\Delta t \to 0$, the dynamical equation for this quantity

$$\partial_t W_a(t) = -aW_a(t) + amw - m \langle a \rangle W_a(t) + \int a' W_{a'}(t) F(a') da'$$

where w is the average density of walkers per node. The first two terms account for the active nodes which release all the walkers they have and are visited by the walkers traveling from all the other nodes. The last two terms account for the contribution of the inactive nodes due to the activity of the nodes in all the other classes. We are interested in the stationary state in the infinite time limit which gives

$$W_a = \frac{amw + \phi}{a + m \langle a \rangle}$$

where $\phi = \int a' W_{a'}(t) F(a') da'$ is the average number of walkers escaping from the active nodes, and it is constant in the stationary case. Hence the problem reduces to find the solutions of the self-consistency equation

$$\phi = \int a F(a) \frac{amw + \phi}{a + m \langle a \rangle} da$$

The result depends on the node activity and tends to a constant as a grows.

In case of a heavy-tailed distribution the explicit solution for ϕ can be written in term of the hypergeometric function. We can also analyze the mean first passage time T_i, or the average time needed for a walker to arrive to a vertex i starting from any other node in a network, which is given by

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

CHAPTER 3

Epidemic Models

3.1. Introduction

Infectious diseases create a significant problem for health and economic all around the world. The appearing of new diseases and the persistence of old ones, make epidemics modeling a fundamental tool to study this phenomenon and guide the health policy around the world.

Different approaches from different scientific fields have been used during the last two centuries to describe epidemics, from the Bernoulli model of the 1766 up to now, ranging from biology to computer science and mathematics [8, 58].

The standard mathematical approach to epidemic processes is the *compartmental model* [34, 56, 57]. In this case the population is divided into classes, or compartment, depending on the stage of the disease. It is possible to define a variety of compartments, but for our purpose we will focus just on three of them: the susceptible stage (S), in which the individual can be infected, the infectious stage (I), in which the individual is infected, and the recovered stage (R), in which the individual is cured and immune to a reinfection. In this work we will focus on two compartmental models the Susceptible-Infected-Susceptible (SIS) and the Susceptible-Infected-Recovered (SIR) .

The main objective in the studying of an epidemic is to establish and formalize the transitions between compartments so that it is possible to track the number of individuals in each stage. In this chapter we will consider some of the most relevant and simplest models formulated up to now [92].

In the first section we will introduce the classical mathematical approach to epidemics and define some of the fundamental parameter needed to characterize this phenomenon.

In the second section we will use static networks concepts to understand how introducing the topology affects the spreading process, and in the last section we will see the role of the time evolution of the network.

3.2. Traditional models

The traditional approach studies the epidemics by using the *mean-field approximation* [48] without introducing networks at all. In this framework each individual

interact with the whole population randomly. Under this approximation, the density of individuals N^σ in the compartment σ or its density $\rho^\sigma = N^\sigma/N$, fully describes the state of the epidemics, where σ can be S, I, or R in our case, and N is the total population.

In the simplest definition of epidemics dynamics N is fixed and all the other demographic processes can be ignored. There are two types of transitions between the compartments which completely define the epidemic evolution: the infection and the recovery processes. The recovery transition is spontaneous after a certain time. In the discrete time models an infected individual has a probability μ to recover at any time step, and the time it will spend in the infectious compartment will be μ^{-1}. In the continuous time formulation, it is generally assumed a Poisson process [32] where now μ is a probability per unit time (rate), and we can define the probability that infected individuals remain in this state for a time τ as $P_{inf}(\tau) = \mu e^{-\mu\tau}$, with average infection time $\langle\tau\rangle = \mu^{-1}$. This means that the epidemic model can be formulated in terms of a Markov process [52, 126].

The infection transition occurs only if there is an interaction between a susceptible and an infected individual and hence depends on the interaction pattern considered in the model and on several other factors. Without information about the connection between individuals, the individuals are considered in the homogeneous mixing approximation, hence randomly interacting among each others. In this case the larger is the number of infectious agents among an individual's neighbors the higher is the probability of the infection. This naturally leads to the introduction of the *force of infection* α which is the probability that an individual can contract the infection in a single time step and in the continuous time limit is defined as the rate

$$\alpha = \bar{\lambda}\rho^I$$

where $\bar{\lambda}$ depends on the specific disease and contact pattern of the population. In some cases $\bar{\lambda}$ can be split in the rate of infection per effective contacts λ and the number of contacts k with other individuals.

This approach can also be used considering the epidemic as a stochastic reaction-diffusion process, where the individuals of each compartment can be seen as different kinds of particles evolving according to specific interaction defined by the reaction rate. This framework is generally more complicated, and goes beyond the objectives of this introduction.

We will present the classical results for epidemic processes considering the dynamics in terms of deterministic ordinary differential equations obtained applying the laws of mass action, in the mean-field approximation. In this case the change of the density of the population in each compartment due to the interactions is given by the force of infection times the average population density.

Notice that the mass-action approximation is not realistic. In fact, people interact with a small fraction of the entire population and not randomly, which underline the importance of the introduction of a set of rules that define an interaction structure in the system. Nonetheless, the classical approach is useful to explore the core mathematical features of the epidemic spreading.

3.2.1. SIS. Many real diseases don't confer immunity after the recovery, which let an individual susceptible to *reinfection*. The simplest model that describes this behavior is the *SIS model*, in which only two states are possible: the infected I, and the susceptible S. The dynamics of this system can be described by the reaction scheme

$$I + S \xrightarrow{\lambda} 2I \qquad\qquad I \xrightarrow{\mu} S$$

where λ is the infection rate, and μ is the recovery rate.

The deterministic differential equation describing the process is given by

$$\partial_t \rho^S = \mu \rho^I - \lambda \rho^I \rho^S$$

$$\partial_t \rho^I = \lambda \rho^I \rho^S - \mu \rho^I$$

Considering that $\rho^I + \rho^S = 1$ for a fixed number of the total population, the set of equation can be simplified to

$$\partial_t \rho^I = (\lambda - \mu - \lambda \rho^I)\rho^I$$

of which the solution is

$$\rho^I(t) = \left(1 - \frac{\mu}{\lambda}\right) \frac{Ce^{(\lambda-\mu)t}}{1 - Ce^{(\lambda-\mu)t'}}$$

where the integration constant is determined by the initial number of infected individuals ρ_0

$$C = \frac{\lambda \rho_0}{\lambda - \mu - \lambda \rho_0}$$

In the limit of large population, small numbers of infected agents $\rho_0 \to 0$, and $C = \lambda\rho_0/(\lambda - \mu)$, leading to

$$\rho^I(t) = \rho_0 \frac{(\lambda - \mu)e^{(\lambda-\mu)t}}{\lambda - \mu + \lambda \rho_0 e^{(\lambda-\mu)t'}}$$

If $\lambda > \mu$ the population can never be totally infected, and in the long-time limit the stable state corresponds to a steady fraction of the population always infected with the disease. This fraction can be obtained imposing $\partial_t \rho^I = 0$ to give $\rho^I = (\lambda - \mu)/\mu$ which is called *endemic state*. On the other hand, when λ approaches to μ, the fraction of infected nodes in the endemic state goes to zero, while if $\lambda < \mu$ the disease will die out exponentially.

A fundamental parameter to evaluate the rising of an epidemic outbreak is the *basic reproduction number* R_0 [6]. Consider a susceptible individual who catches the disease in the early stage of an outbreak, then R_0 is defined as the average number of additional infections caused by this agent before it recovers.

If $R_0 < 1$, the relative size of the epidemics vanishes because a single individual can't generate enough secondary infection to sustain the spreading. On the other hand if $R_0 > 1$ the average fraction of infected agents grows exponentially. The last condition, while necessary and sufficient for deterministic models, is only necessary for stochastic models where fluctuations of the number of infected individuals can lead to the extinction of the infection for a small initial number of infected agents. The point $R_0 = 1$ separates the two opposite behaviors defining the *epidemic threshold.*

In the SIS model the transition between epidemic and non-epidemic regime happens at the point $\lambda = \mu$ also called *epidemic transition* point, and the basic reproduction number is given by $R_0 = \lambda/\mu$.

3.2.2. SIR. For many diseases people retain their immunity after the recovery process, preventing them from a reinfection. The simplest model to describe this behavior is the SIR model. In this framework, a susceptible individual (S) can catch the disease from an infected individual (I) which, after a certain time, can recover and be removed from the dynamics (R). The dynamical process can be described by the reaction scheme

$$I + S \xrightarrow{\lambda} 2I \qquad\qquad I \xrightarrow{\mu} R$$

where contacts with infected individuals happen with an average rate λ, while the recovery process happens with a constant average rate μ.

It is possible to define the probability to recover in a time interval $\delta\tau$ as $\mu\delta\tau$, and obtain the probability to stay infected after a total time τ as

$$\lim_{\delta\tau\to 0}(1-\mu\tau)^{\tau/\delta\tau} = e^{-\lambda\tau}$$

The probability that an infected individual recover in the interval $[\delta\tau, \tau + \delta\tau]$ is $p(\tau)d\tau = \mu e^{-\mu\tau}d\tau$, which is a standard exponential distribution, meaning that the recovery process is most likely to happen just after the infection takes place. In most cases this is quite unrealistic, considering that people may remain infected for much longer time, depending on the disease. We will see how this estimation improves introducing the network.

In terms of the fraction of individuals in each compartment, the system is described by the differential equations

$$\begin{aligned}\partial_t\rho^S &= -\lambda\rho^I\rho^S\\ \partial_t\rho^I &= \lambda\rho^I\rho^S - \mu\rho^I\\ \partial_t\rho^R &= \mu\rho^I\end{aligned}$$

For a fixed number of population we can consider the normalization condition $\rho^I + \rho^S + \rho^R = 1$, the set of equations can be simplified eliminating the ρ^I variable

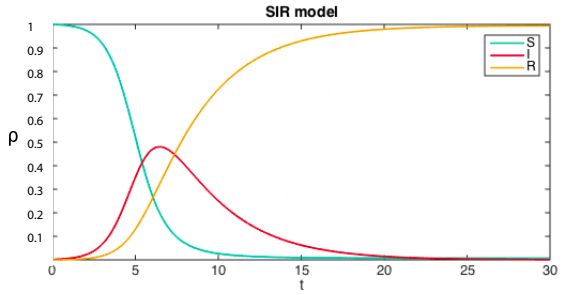

FIGURE 3.2.1. SIR epidemic processes. Density of nodes in each compartment, depending on time t.

obtaining

$$\rho^S = \rho_0^S e^{-\lambda \rho^R/\mu}$$

where ρ_0^S is the fraction of susceptible individuals at time $t = 0$, and then using the normalization condition it is possible to obtain

$$\partial_t \rho^R = \mu(1 - \rho^R - \rho_0^S e^{-\lambda \rho^R/\mu}) \tag{3.2.1}$$

The solution can be written as

$$t = \frac{1}{\mu} \int_0^{\rho^R} \frac{dx}{1 - x - \rho_0^S e^{-\lambda x/\mu}}$$

which can not be evaluated in closed form, but just numerically.

As shown in figure (3.2.1), the fraction of susceptible individuals in the population decreases monotonically and the fraction of recovered individuals increases monotonically. The fraction of infected goes up at first as people get infected, then down again as they recover, and eventually goes to zero when $t \to \infty$.

On the other hand, the fraction of susceptible individuals doesn't go to zero, because when $\rho^I \to 0$ it is not possible to have new infections. Also the fraction of recovered doesn't reach one as $t \to \infty$, and its asymptotic value represents the total number of individuals that caught the disease, hence is the total size of the outbreak, which can be useful to characterize the epidemic. This can be calculated from the eq 3.2.1 imposing $\partial_t \rho^R = 0$, which gives $\rho^R = 1 - \rho_0^S e^{-\lambda \rho^R/\mu}$.

The most common choice for the initial condition is to consider the infection to start, either from a single individual or from a small fraction r of the population. In this case the initial values of the variables are $\rho_0^R = 0$, $\rho_0^S = 1 - r/N$ and $\rho_0^I = r/N$, so that in the limit of large population $N \to \infty$, the total outbreak size is

$$\rho_\infty^R = 1 - e^{-\lambda \rho_\infty^R/\mu}$$

These results indicate that the size of the epidemic continuously goes to zero for $\lambda \leq \mu$, which means that the infected individuals recover faster than the susceptible ones become infected, so that the disease dies out.

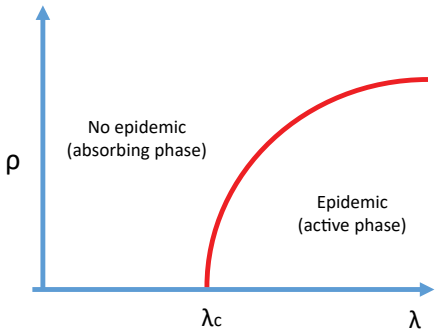

FIGURE 3.2.2. Phase diagram of a SIS-like absorbing state phase transition.

An individual that remains infected for a time τ by the same amount of time will have contacted a number $\lambda\tau$ of other individuals. By definition the reproduction number R_0 is the average number of additional people that an infected individual passes the disease to before they recover, which is

$$R_0 = \lambda\mu \int_0^\infty \tau e^{-\mu\tau} d\tau = \frac{\lambda}{\mu}$$

As for the SIS model also in the SIR model the epidemic threshold falls in the point where $\lambda = \mu$ in the long time regime.

3.2.3. Epidemics and phase transition. Epidemic processes are a typical example of critical phenomena [133, 46, 65]. In this case the phase transition is between the *non-epidemic (absorbing)* and *epidemic (active)* phases, characterized by the order parameter ρ^σ, and the control parameter λ.

In the SIS case the order parameter is the density of the infected individuals ρ^I determining the distinction between the non-epidemic and epidemic phases. This model belongs to the universality class of direct percolation, which is the paradigm of dynamical phase transitions.

In the SIR model the order parameter is the size of the outbreak, hence the density ρ^R of all the population ever being infected.

For both the SIS and SIR problem the control parameter is the infection rate λ. The critical point λ_c such that $\rho = 0$ for $\lambda < \lambda_c$ and $\rho > 0$ for $\lambda > \lambda_c$, defines the epidemic threshold of the system. The phase diagram can be expressed in terms of $\rho(\lambda)$ as shown in the figure 3.2.2.

3.3. Epidemics on static networks

Classical models of epidemic spreading consider the population to be fully connected and the individuals to randomly interact within each other, this assumption is clearly unrealistic. In general people have a regular set of acquaintances, friends, and coworkers, whom they interact with while ignoring the rest of the population. The potential contacts of an individual form a set that can be easily represented

as a network. As we will see from now on, the network structure and its evolution strongly affect the spreading of a disease.

Introducing the network with N node and considering χ number of compartments representing the stages of an epidemic process, the state of the node i at time t is given by the random variable $X_i(t)$, where $X_i(t) = \sigma$ means that the node i belongs to the compartment σ at time t. Considering the transitions between the compartments as independent Poisson processes with certain rates, the epidemic process can be studied in terms of a Markov chain [52]. At this point it is possible to study the evolution of the probability for $X_i(t)$ to be in a state $\sigma_i \in [0, \chi]$ at time t.

The other possible approach is to describe the evolution in terms of the master equation (see 1.4.2) for the probability $P(s^\sigma, t)$ to be in the compartment σ at the time t, where s^σ is the set of states $s_i^\sigma(t)$ indicating that the node i belongs to the compartment σ at time t.

3.3.1. Individual based mean field. In the *individual-based mean-field* (IBMF) model the evolution equation is written in terms of the probability ρ_i^η that node i is in the state η for each node, assuming that the dynamical state of each node is statistical independent from the ones of its nearest neighbors, i.e that the probability for a node i to be in a state η and for its neighbor j to be in a state η' is $\rho_i^\eta \rho_j^{\eta'}$ [70, 43].

This approach keeps the full structure of the networks while using the mean field approximation to neglect the correlations between neighbors. As a consequence, the solutions depend in general on the spectral properties of the adjacency matrix, they fail to describe the system when either the variable are highly correlated or when the densities in a compartment are very small.

3.3.1.1. *SIS IBMF.* The SIS epidemic process on a network can be described by a Bernoulli random variable $X_i(t) \in \{0, 1\}$, where $X_i = 0$ corresponds to the susceptible state and $X_i(t) = 1$ corresponds to the infected state of the node i at time t [70, 125]. Hence the probability for a node i to be infected at time t is given by $\rho_i^I(t) = \Pr[X_i(t) = 1]$, which for a Bernoulli variable corresponds to the expectation value $E[X_i(t)]$, while the probability to be susceptible is $1 - \rho_i^I(t)$. The general exact equation that describes the expectation of being infected for each node i is given by

$$\partial_t E[X_i(t)] = E\left[-\mu X_i(t) + [1 - X_i(t)]\, \lambda \sum_{j=1}^{N} a_{ij} X_j(t)\right] \tag{3.3.1}$$

where the second term is the expectation value that the node i recovers with rate μ, and being susceptible, $[1 - X_i(t)]$, is infected by its neighbors. In this case a_{ij} are the elements of the adjacency matrix. This formalism can be extended to both time dependent adjacency matrix $A(t)$ and asymmetric adjacency matrix. From the

formula above we can say that the time evolution of the probability to be infected is affected by two mechanisms: if the node is infected then $\partial_t E[X_i(t)]$ decreases with a rate μ, while if it is healthy it can be infected with rate λ.

For static networks, Eq. (3.3.1) reduces to the following [105]

$$\partial_t \rho_i^I(t) = -\mu \rho_i^I(t) + \lambda \sum_{j=1}^{N} a_{ij} \rho_j^I(t) - \lambda \sum_{j=1}^{N} a_{ij} E[X_i(t) X_j(t)] \tag{3.3.2}$$

Now we can apply the IBMF approximation to close the equation, assuming the statistical independence of the states of two neighboring nodes, i.e.

$$E[X_i(t)X_j(t)] = E[X_i(t)]E[X_j(t)] = \rho_i^I(t)\rho_j^I(t)$$

and Eq. (3.3.2) becomes

$$\partial_t \rho_i^I(t) = -\mu \rho_i^I(t) + \lambda \left[1 - \rho_i^I(t)\right] \sum_{j=1}^{N} a_{ij} \rho_j^I(t) \tag{3.3.3}$$

Hence the time evolution of the probability $\rho_i^I(t)$ is given by minus the probability that an infected node recovers with rate μ, plus the probability that a susceptible node gets the infection from one of its nearest infected neighbors with rate λ.

The epidemic threshold is obtained applying the linear stability analysis to (3.3.3) and studying the eigenvalues problem for the Jacobian J with elements $J_{ij} = -\delta_{ij} + \frac{\lambda}{\mu} a_{ij}$. We are in the endemic case when the largest eigenvalues Λ_1 of J is positive [70],which leads to the condition

$$\lambda \geq \lambda_c^{IBMF} = \frac{1}{\Lambda_1}$$

For heterogeneous networks with power-law degree distribution $P(k) \sim k^{-\gamma}$, $\Lambda_1 \sim \max\left\{\sqrt{k_{max}}, \left\langle k^2 \right\rangle / \left\langle k \right\rangle\right\}$ [30] where k_{max} is the maximum degree of the network, the epidemic threshold is

$$\lambda_c^{IBMF} \sim \begin{cases} 1/\sqrt{k_{max}} & \gamma > 5/2 \\ \left\langle k \right\rangle / \left\langle k^2 \right\rangle & 2 < \gamma < 5/2 \end{cases}$$

This means that for every network where the maximum degree is a growing function of the network size, the epidemic threshold vanishes in the thermodynamic limit.

3.3.1.2. *SIR IBMF.* Also in the SIR case, the approach of the IBMF approximation is to write the full master equation for the probabilities of the states for each node and to obtain from it the deterministic evolution equation for the quantities

$$\partial_t \rho_i^S(t) = -\lambda \sum_{j=1}^{N} a_{ij} \left\langle S_i I_j \right\rangle$$

$$\partial_t \rho_i^I(t) = \lambda \sum_{j=1}^{N} a_{ij} \left\langle S_i I_j \right\rangle - \mu \rho_i^I(t)$$

where S_i and I_i are Bernoulli variable such that $\rho_i^S = \langle S \rangle_i$ and $\rho_i^I = \langle I \rangle_i$ are the probability for the node i to be susceptible or infected respectively, while $\langle S_i I_j \rangle$ is the join probability of being in the state $S_i I_j$.

The IBMF approximation, $\langle S_i I_j \rangle = \langle S \rangle_i \langle I \rangle_j$, introduced to close the equation, brings out a physical interpretation problem.

$$\partial_t \rho_i^S(t) = -\lambda \sum_{j=1}^{N} a_{ij} \rho_i^S \rho_j^I$$

$$\partial_t \rho_i^I(t) = \lambda \sum_{j=1}^{N} a_{ij} \rho_i^S \rho_j^I - \mu \rho_i^I(t)$$

As shown by Sharkey [105] this approximation is equivalent to write the evolution equation of $\langle S_i I_j \rangle$, implying that a node can be susceptible and infected at the same time. In this case the IBMF approximation leads to the same result for the epidemic threshold of the SIS case giving $\lambda_c = 1/\Lambda_1$. For heterogeneous power-law distributed networks and $\gamma > 3$, this result leads to a vanishing epidemic threshold which is not correct as shown in [28].

3.3.2. Degree based mean field. In the *degree-based mean field* (DBMF) or *heterogeneous mean field* (HMF) approximation all the nodes of the same degree k are considered statistically equivalent, which means that any node with degree k is connected with probability $P(k|k')$ to every node of degree k'. In this case, the relevant variables specify the degree class of a node [91], and the dynamical equation is written in terms of the probability $\rho_k^\eta(t)$ that a node with degree k is in the compartment η at time t . The variables, which are not independent, must satisfy the condition $\sum_\eta \rho_k^\eta(t) = 1$, while the total fraction of agents in the same compartment η is given by $\rho^\eta(t) = \sum_k P(k) \rho_k^\eta(t)$ where $P(k)$ is the degree distribution.

In this framework, the network itself is considered in a mean field perspective preserving only the degree information and the correlation between two nodes [35]. In this sense, the DBMF is equivalent to use the ensemble average of the adjacency matrix in the IBMF theory

$$\bar{a}_{ij} = \frac{k_j P(k_i|k_j)}{N P(k_i)}.$$

which represents the probability that the nodes i and j are connected, this is also referred to annealed network approximation.

The solutions of DBMF models generally depend on the statistical properties of the networks, but, despite being a strong approximation, it well describes spreading processes evolving on networks for which the interactions changes on time scales much faster than the dynamics on them.

3.3.2.1. *SIS DBMF.* In the DBMF approximation the dynamical equation for the SIS process can be described by the law of mass action [91]

$$\partial_t \rho_k^I(t) = -\mu \rho_k^I(t) + \lambda k \left[1 - \rho_k^I(t)\right] \sum_{k'} P(k'|k)\, \rho_{k'}^I(t) \tag{3.3.4}$$

The first term considers the recovered nodes of degree k. The second accounts for the infection of new nodes, considering the probability that a node v_k of degree k is susceptible, $1 - \rho_k^I(t)$, times the infection rate λ and probability that a node $v_{k'}$ of degree k' is infected and connected to v_k with probability $P(k'|k)$, summed over all possible values of k'. We can divide Eq. 3.3.4 by μ and imposing it equal to one without loss of generality. The epidemic threshold is obtained studying the linear stability of the equation and performing a first order expansion in $\rho_k^I(t)$

$$\partial_t \rho_k^I(t) \simeq \sum_k J_{k,k'} \rho_{k'}^I(t)$$

where $J_{k,k'} = -\delta_{k,k'} + \lambda k P(k'|k)$ is the Jacobian matrix element. The healthy phase corresponds to a null steady state, which is stable when the largest eigenvalue of the Jacobian Λ_M, is negative. On the other the solution $\rho_k = 0$ is unstable if exist at list on positive eigenvalue of the Jacobian matrix, this means, that the epidemic phase emerges when $\lambda\Lambda_M - 1 > 0$, i.e. when

$$\lambda > \lambda_c^{DBMF} = \Lambda_M^{-1} \tag{3.3.5}$$

as shown in Ref. [19].

For uncorrelated networks

$$P(k'|k) = k' P(k) / \langle k \rangle \tag{3.3.6}$$

and Eq 3.3.4 can be written as

$$\partial_t \rho_k^I(t) = -\rho_k^I(t) + \lambda k \left[1 - \rho_k^I(t)\right] \Theta(\lambda) \tag{3.3.7}$$

where

$$\Theta(\lambda) = \sum_k k \frac{P(k)}{\langle k \rangle} \rho_k^I(t)$$

is the probability that a random chosen link leads to an infected node.

From the stationary condition it is possible to obtain an expression for the probability $\rho_k^I(t)$, given by

$$\rho_k^I(t) = \frac{\lambda k \Theta(\lambda)}{1 + \lambda k \Theta(\lambda)}.$$

This indicates that, for uncorrelated networks, the higher the node's degree is, the higher is its probability to be infected, implicating that high heterogeneity in the connectivity patterns strongly affects the spreading of a disease.

Notice that $\Theta(\lambda)$ can be computed solving the self-consistency equation

$$\Theta(\lambda) = \sum_k k \frac{P(k)}{\langle k \rangle} \frac{\lambda k \Theta(\lambda)}{1 + \lambda k \Theta(\lambda)}. \tag{3.3.8}$$

In this case, the epidemic threshold can be derived either substituting the expression 3.3.6 in the Jacobian and computing the eigenvalue or imposing that the self-consistency equation 3.3.8 admits a non-zero solution, obtaining

$$\lambda > \lambda_c^{DBMF,unc} = \frac{\langle k \rangle}{\langle k^2 \rangle}. \tag{3.3.9}$$

The critical behavior of the order parameter around the critical point can be obtained from 3.3.8 giving $\rho_k^I(t) \sim (\lambda - \lambda_c^{DBMF})^{\eta_{SIS}^{DBMF}}$ where η_{SIS}^{DBMF} is the critical exponent. Moreover for networks with power-law degree distribution $P(k) \sim k^{-\gamma}$ with exponent $2 < \gamma \leq 3$, in the limit of infinite scale networks, the epidemic threshold tends to zero, while the critical exponent is larger than 1. This means that while the disease spreads more easily, the epidemic activity grows very slowly increasing the spreading rates, making the epidemic less threatening.

Notice that for regular networks $\langle k^2 \rangle = \langle k \rangle^2$, recovering the result $\lambda_c^{DBMF} = 1/\langle k \rangle$.

3.3.2.2. *SIR DBMF.* To extend the DBMF approximation to the SIR model it is necessary to introduce also the partial densities of the recovered and of the susceptible nodes with degree k, $\rho_k^R(t)$ and $\rho_k^S(t)$ respectively, which fulfill the normalization condition $\rho_k^R(t) + \rho_k^S(t) + \rho_k^I(t) = 1$ for finite size population. The set of the equations describing the process is given by

$$\begin{aligned} \partial_t \rho_k^I(t) &= -\rho_k^I(t) + \lambda k \rho_k^S(t) \sum_{k'} P(k'|k)\, \rho_{k'}^I(t) \\ \partial_t \rho_k^R(t) &= \rho_k^I(t) \end{aligned} \tag{3.3.10}$$

Also in this case as well, the linear stability analysis leads to the value for the epidemic threshold, which is the inverse of the largest eigenvalue of the adjacency matrix.

For the SIR model the order parameter is the number of recovered individuals at the end of the epidemics, i.e. $\rho_\infty^R(t) = \lim_{t\to\infty} \sum_k P(k)\rho_k^R(t)$.

For uncorrelated networks it is possible to integrate the rate equation over time to study the whole temporal evolution of the process, introducing the function

$$\phi(t) = \sum_k k \frac{P(k)}{\langle k \rangle} \rho_k^R(t)$$

In general the solution depends on the differential equation for the function $\phi(t)$, but in the limit of infinite time it is possible to obtain the final prevalence

$$\rho_\infty^R = \sum_k kP(k) \left(1 - e^{-\lambda k \phi_\infty}\right)$$

where

$$\phi_\infty = 1 - \frac{1}{\langle k \rangle} - \sum_k k \frac{P(k)}{\langle k \rangle} e^{-\lambda k \phi_\infty} \tag{3.3.11}$$

The epidemic threshold can be obtained from 3.3.11, giving $\lambda_c = \frac{\langle k \rangle}{\langle k^2 \rangle}$. Moreover for power-law degree distributed networks with $P(k) \sim k^{-\gamma}$, the equation for the order parameter is $\rho_\infty^R \sim (\lambda - \lambda_c)^{\eta_{SIR}}$ [76].

Notice that in case of annealed networks, the results above are exact,but in case of static networks it is possible to improve the model's prediction considering that in the SIR model the reinfection of a recovered node is prohibited, and the disease can't propagate through the neighbors that have already been infected. The effect on this approximation can be included modifying the sum in the second r.h.s term of the 3.3.11 $P(k'|k) \to P(k'|k)(k'-1)/k'$, giving a new largest eigenvalue of the adjacency matrix

$$\tilde{\Lambda}_1 = \frac{\langle k^2 \rangle}{\langle k \rangle} - 1$$

which corresponds to the epidemic threshold

$$\lambda_c = \frac{\langle k \rangle}{\langle k^2 \rangle - \langle k \rangle}\,.$$

An important insight, in particular for the SIR-like models, is the time scale evolution of an epidemic outbreak, which is of order $(\lambda\tilde{\Lambda}_1)^{-1}$and in this case is given by

$$\tau = \frac{\langle k \rangle}{\lambda \langle k^2 \rangle - (\mu + \lambda)\langle k \rangle}$$

Notice that as for the epidemic threshold, the time-scale of an epidemic outbreak vanishes when the second moment of the degree distribution diverges, for example in scale-free networks. This mechanism can be extensively studied in a scale-free network with computer simulations showing that at the beginning, the infection reaches the hubs and invades the rest of the networks via a cascade process [15, 16].

3.3.3. Other Results.

3.3.3.1. *SIS.* A relevant result for the SIS model provides a lower bound for the epidemic threshold. This was introduced by Mieghem [69] considering the inequality $0 \leq \sum_{j=1}^{N} a_{ji} X_i(t) X_j(t)$ where $X_j(t)$ are the Bernoulli random variable introduced in the eq 3.3.1. In this case it is possible to write

$$\partial_t \rho_i^I(t) \leq -\mu \rho_i^I(t) + \lambda \sum_{j=1}^{N} a_{ij} \rho_j^I(t)$$

Considering the vector $W = (\rho_1^I, \ldots, \rho_N^I)$ in a network of N nodes, the solution of the inequality is

$$W(t) \leq e^{(\frac{\lambda}{\mu} A - 1)t} W(0)$$

The inequality is dominated by the term $\frac{\lambda}{\mu}\Lambda_1 - 1$, where Λ_1 is the largest eigenvalue of the adjacency matrix A. When $\frac{\lambda}{\mu}\Lambda_1 - 1 \leq 0$, $W_i(t) = \rho_i^I(t)$ tends to 0 and the fraction of infected individuals rapidly decreases ending the epidemic spreading.

This imposes a lower bound for the epidemic threshold

$$\lambda_c \geq \frac{1}{\Lambda_1}$$

which is the same result as for the IBMF model.

3.3.3.2. *SIR.* The SIR process, in the long time regime, can be mapped to a bond percolation problem [81]. In this framework, the links in a network are kept with probability $1-p$ and removed with probability p. The probability that a randomly chosen link doesn't attach to a vertex connected to a giant component is given by

$$u = 1 - p + \sum_k \frac{kP(k)}{\langle k \rangle}(1 - p + pu)^{k-1} \tag{3.3.12}$$

which is the equation for degree uncorrelated networks with no loops, in which a randomly chosen edge points to a node of degree k with probability $kP(k)/\langle k \rangle$. The probability that a randomly chosen node belongs to the giant component is

$$P_G(p) = 1 - \sum_k P(k)(1 - p + pu)^k. \tag{3.3.13}$$

Introducing the degree distribution generating function $G_0(z) = \sum_k P(k)z^k$ and the excess degree generating function $G_0(z) = \sum_k (k+1)P(k+1)z^k/\langle k \rangle$, it is possible to write the equations 3.3.13 and 3.3.12 as

$$u = 1 - p + G_1(1 - p + pu),$$

$$P_G(p) = 1 - G_0(1 - p + pu).$$

The condition for the existence of a giant component translates into the condition for the existence of a nonzero solution which is

$$p > p_c = \frac{G_0'(1)}{G_0''(1)} = \frac{\langle k \rangle}{\langle k^2 \rangle - \langle k \rangle}$$

The behavior of the order parameter can be found performing the expansion of the generating function near the critical point around the nonzero solution obtaining $P_G(p) \sim (p - p_c)^{\beta_{perc}}$, where the critical exponent in case of homogeneous networks is $\beta_{perc} = 1$. For heterogeneous networks with degree distribution $P(k) \sim k^{-\gamma}$, in the thermodynamic limit $N \to \infty$ the percolation threshold tends to zero for $\gamma < 3$ and the critical exponents take the values

$$\beta_{perc} = \begin{cases} \frac{1}{(3-\gamma)} & \text{for } \gamma < 3 \\ \frac{1}{(\gamma-3)} & \text{for } 3 < \gamma \leq 4 \\ 1 & \text{for } \gamma > 3 \end{cases}$$

As shown in [78] the probability that a link exists p is related to the probability that an infected node can transmit the disease to a connected susceptible node. Let's consider the SIR model with uniform infection time τ, i.e. the recovery time after the infection, and infection rate λ, the transmissibility T is defined as the

probability that an infected node transmits the disease to a susceptible node, and, in the continuous time limit, it is given by

$$T = 1 - \lim_{\delta t \to 0} (1 - \lambda \delta t)^{\tau/\delta t} = 1 - e^{-\tau\lambda}$$

Now the cluster of the bond percolation problem to which the initial node belongs, is the same as the set of recovery nodes of an SIR outbreak generated from a single node where the occupation probability p is now T. The correspondence is exact and for tree-like networks we have

$$T_c = \frac{\langle k \rangle}{\langle k^2 \rangle - \langle k \rangle}$$

$$\lambda_c = \frac{1}{\tau} \ln \frac{\langle k^2 \rangle - \langle k \rangle}{\langle k^2 \rangle - 2 \langle k \rangle}$$

where the behavior of the outbreak size close to the epidemic threshold is given by the exponent β_{perc}. Notice that this means that the epidemic thresholds has qualitatively different behavior for scale-free networks ($\gamma < 3$) and scale-rich ones ($\gamma \geq 3$). For scale-free networks, the threshold vanishes, meaning that this type of structures are extremely vulnerable to disease spreading.

In presence of loops and multiple spreading paths, the possible correlation can invalidate the result. However, for random graphs, which are locally treelike, this result still stands in the thermodynamic limit where the loops are infinitely long.

We can extend the result to non uniform infection times, assuming that τ_i and λ_{ij} vary between individuals. In this case, the trasmissibility T_{ij} depends on the edge (i, j) and neglecting the fluctuation we can replace T_{ij} with its mean value

$$\langle T_{ij} \rangle = 1 - \int d\tau \int d\lambda e^{-\lambda\tau} Q(\lambda) P(\tau)$$

where P and Q are the distributions of τ_i and λ_{ij} respectively. For non-degenerate infection times exponentially distributed and constant recovery rate, performing the integral, we obtain the epidemic threshold

$$\lambda_c = \frac{\langle k \rangle}{\langle k^2 \rangle - 2 \langle k \rangle}$$

Notice that this approximation fails when correlations are involved, in fact if an individual recovers quickly, the probability that it will transmit the disease to its neighbors is small, while it is much higher if it recovers slowly. When τ_i are degenerate and λ_{ij} vary, this approximation is not exact [71].

3.4. Epidemics on time evolving networks

In all the model of epidemic spreading presented until now, we have supposed a fundamental approximation: the dynamics of the network doesn't affect the dynamics of the contagion process. Indeed we have considered the limit of extreme

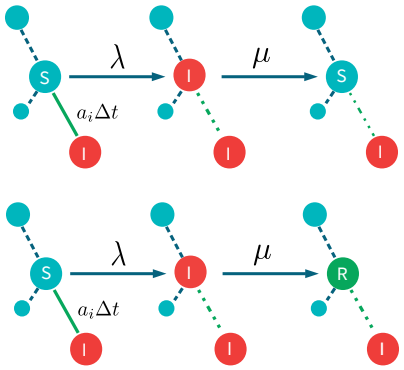

FIGURE 3.4.1. SIR and SIS processes on the activity-driven network

time scale separation between the network and the contagion process either considering the network frozen with time scales much larger than the dynamical process or considering the network rewiring on much faster time scale than the contagion process.

In reality, especially in social dynamics, the contacts change on the same time-scale of spreading phenomena. For instance, we generally interact with a small number of contacts at the time, rather than with all our friends simultaneously, hence the time evolution of the networks strongly affect the spreading process dynamics in real social systems.

3.4.1. Simple activity-driven model (AD). As described in the previous chapter, the activity driven model is one of the most versatile models of TVN that allows the analytical study of a dynamical process on the network. In this framework, the time variable is embedded in the nodes activity a which represents the probability for the node to form a number m of links at time t. The activity is drawn from the distribution $F(a)$.

In the original formulation of the epidemic spreading on the ADN, Perra et al [94] considered the evolution of the number of infected individuals I_a^t in the class of activity a at time t. They used the HMF approximation considering all the nodes with the same activity statistically equivalent.

The dynamic process for both the SIS and the SIR model is explained in the figure (3.4.1), while the equation for the SIS is given by

$$I_a^{t+1} - I_a^t = -\mu I_a^t + \lambda m(N_a - I_a^t)a \int da' \frac{I_a^t}{N} + \lambda m(N_a - I_a^t) \int da' a' \frac{I_a^t}{N}$$

where $N_a = NF(a)$ is the total number of individual in the activity class a, μ and λ are the recovery and infection probabilities respectively and $(N_a - I_a^t) = S_a^t$ is the number of susceptible individual in the activity class a at time t. The first term on r.h.s. represents the recovery nodes, the second accounts for the probability that a susceptible node activates and becomes infected contacting an infected node, while

the third is the probability that a susceptible node is infected when contacted by one of its neighbors.

Summing over all the activity classes, and neglecting the second order correlations, in the continuous time limit the equation reduces to the system of equations

$$\partial_t I = -\mu I + \lambda m \langle a \rangle I + \lambda m \Theta$$

$$\partial_t \Theta = -\mu \Theta + \lambda m \langle a^2 \rangle I + \lambda \langle a \rangle \Theta$$

where $\Theta = \int da I_a a$.

Studying the linear stability analysis the system can be exactly solved, and requiring the largest eigenvalue of the Jacobian

$$J = \begin{pmatrix} -\mu + \lambda m \langle a \rangle & \lambda m \\ \lambda m \langle a^2 \rangle & -\mu + \lambda m \langle a \rangle \end{pmatrix}$$

to be positive, it is possible to obtain the epidemic threshold

$$\frac{\lambda}{\mu} \geq \frac{1}{m} \frac{\langle a \rangle}{\langle a \rangle + \sqrt{\langle a^2 \rangle}}$$

As we can see the threshold depends only on the activity distribution and in particular the fluctuations of the activity tends to dump the threshold. The same result can be obtained for the SIR model. We will introduce other variations of the AD model and in particular we will discuss thoroughly what happens when we introduce a memory process.

3.4.2. Activity-driven model with attractiveness (ADA). The attractiveness b measures the probability for an individual to target for an interaction distributed according to $G(b)$. As we discussed, the introduction of this process in the activity driven model affect the spreading of a disease [97]. In the general case the system is described by the distribution $H(a, b)$ of both the activity and the attractiveness. Let's first study the SIS process in the HMF approximation where all the nodes with the same activity and attractiveness are statistically equivalent. Then we can write the equation for the number of infected nodes $I^t_{a,b}$ in the class of activity a and in the class of attractiveness b at time t in limit $N \gg 1$, where N is the total number of nodes.

$$I^{t+1}_{a,b} - I^t_{a,b} = -\mu I^t_{a,b} + \frac{\lambda m}{N \langle b \rangle} (N_{a,b} - I^t_{a,b}) \left[a \sum_{a',b'} I^t_{a',b'} b' + b \sum_{a',b'} a' I^t_{a',b'} \right]$$

The main difference respect to the AD case is that now the probability for a node in the class (a, b) to be contacted depends on b. In the limit of $I_{a,b} \ll N_{a,b}$, where $N_{a,b}$ is the number of nodes in the class (a, b), the equation reduces to the system

$$\partial_t I = -\mu I + \frac{\lambda m}{\langle b \rangle} [\langle a \rangle \Phi + \langle b \rangle \Theta]$$

$$\partial_t \Theta = -\mu\Theta + \frac{\lambda m}{\langle b\rangle}\left[\left\langle a^2\right\rangle \Phi + \langle ab\rangle \Theta\right]$$
$$\partial_t \Phi = -\mu\Phi + \frac{\lambda m}{\langle b\rangle}\left[\langle ab\rangle \Phi + \left\langle b^2\right\rangle \Theta\right]$$

with eigenvalue $l_0 = -\mu$, $l_\pm = \frac{\lambda m}{\langle b\rangle}\left(\langle ab\rangle \pm \sqrt{\langle a^2\rangle \langle b^2\rangle}\right) - \mu$. Imposing the condition of positive eigeinvalue for the outbreak to happen we obtain the epidemic threshold

$$\frac{\lambda}{\mu} \geq \frac{1}{m}\frac{\langle a\rangle \langle b\rangle}{\langle ab\rangle + \sqrt{\langle a^2\rangle \langle b^2\rangle}}$$

If the activity and the attractiveness are uncorrelated, then $H(a,b) = F(a)G(b)$ and the epidemic threshold become

$$\frac{1}{m}\frac{1}{1+\sqrt{\frac{\langle a^2\rangle \langle b^2\rangle}{\langle a\rangle^2 \langle b\rangle^2}}}$$

We can also consider a deterministic correlation between a and b imposing $H(a,b) = F(a)\delta(b - q(a))$, where $q(a)$ is a function determining b for a given a. Considering the relation $G(b) = F(q^{-1}(b))|dq^{-1}(b)/db|$, then if one of the variables is power-law distributed so the other is.

In the particular case of $q(a) = a^{\gamma_c}$ the epidemic threshold is given by

$$\frac{\lambda}{\mu} \geq \frac{1}{m}\frac{\langle a\rangle \langle a^{\gamma_c}\rangle}{\langle a^{1+\gamma_c}\rangle + \sqrt{\langle a^2\rangle \langle a^{2\gamma_c}\rangle}}$$

In both the choices for the distribution $H(a,b)$ we can notice that, for heterogeneous systems, the epidemic threshold is lowered by the attractiveness. This analytical results are valid also for the SIR process and can also be verified by numerical simulations.

3.4.3. Activity-driven model with modularity. In the previous chapter, we introduced a community structure of size s and distributed according a $P(s) \sim s^{-\omega}$ on the AD model. In this section we will describe the derivation of the epidemic threshold for the SIS and SIR processes obtained by Nadini et al [77] in this framework.

To write the mean-field equation governing the dynamic of the epidemic process we need to introduce the activity block variable indicating for each compartment the number of individuals with activity a and community size s at time t, hence we will have the block variable $S_{a,s}(t)$, $I_{a,s}(t)$ and $R_{a,s}(t)$. The evolution for the SIR process is governed then by the equation

$$\partial_t I_{s,a} = -\gamma I_{a,s} + \lambda S_{a,s}\left[\mu a\frac{I_s}{s} + (1-\mu)a\frac{I}{N}\right] + \lambda\sum_{a'}\left[\mu I_{a',s}\frac{S_{a,s}}{s} + (1-\mu)I_{a',s}\frac{S_{a,s}}{N}\right]$$

where I_s is the number of infected individual in a community of size s, while I is the number of infected in the whole network. Here γ is the recovery probability, λ is the

infection probability, and μ is the probability of joining a community. The second and the third terms on the r.h.s. represent the probability that a susceptible node in a community of size s contacts an infected node inside its community (I_s) or outside (I), while the fourth and the fifth terms are the probability for an infected node of activity class a' to connect with a susceptible node inside or outside its community respectively. In the approximation of small community size, i.e. $N - s \sim N$, and small initial numbers of infected individuals in each community ($I - I_s \sim I$), it is possible to write with the same considerations of the simple AD the set of dynamical equations

$$\partial_t I = -\gamma I + \lambda \langle a \rangle I + \lambda \Theta + \lambda \mu \sum_s (\langle a \rangle_s - \langle a \rangle) I_s$$

$$\partial_t \Theta = -\gamma \Theta + \lambda \langle a^2 \rangle I + \lambda \langle a \rangle \Theta + \lambda \mu \sum_s (\langle a^2 \rangle_s - \langle a^2 \rangle) I_s + (\langle a \rangle_s - \langle a \rangle) \Theta_s$$

where $\langle a^n \rangle_s$ are the moments of the activity distribution in any community of size s, $\Theta = \sum_a a I_a$, and $\Theta_s = \sum_a a I_{a,s}$.

In the limit of $\mu \to 0$ the model recovers the simple AD threshold for both the SIS and SIR processes, while for $\mu \to 1$ numerical results for the threshold show that it goes in opposite directions. In the SIR model the reinfection process is not allowed by the dynamics, in this case strongly cohesive communities, with high modularity in the connectivity patterns once recovered tends to block the spreading. On the other hand in the SIS model the reinfection mechanism promotes the spreading among the community and high modularity lower the epidemic threshold.

3.4.4. Epidemics in continuous time evolving networks. A general approach to calculate the epidemic threshold on temporal networks was proposed by Valdano et al [124, 123]. In this framework the temporal network is described by the time dependent adjacency matrix $A(t)$, in the time interval $t \in [0, T]$, which completely characterizes the epidemic threshold in the SIS case. The first approach is to discretize the time in steps of length Δt, obtaining a discrete sequence of adjacency matrices $\{A^h\}_{h=1,..,T_{step}}$.

The SIS dynamics is introduced by meaning of a discrete-time Markov chain equation for the probability $p_{h,i}$ that a node i is infected at the time step h given by

$$p_{h+1,i} = (1 - P_{h,i}) \left[1 - \prod_j (1 - \lambda \Delta t A^h_{ij} p_{h,j} \right] + p_{h,i}(1 - \mu \Delta t)$$

where $\lambda \Delta t$ and $\mu \Delta t$ are the probability to be infected and to recover respectively. Now we can introduce the infection propagator

$$P(T_{step}) = \prod_{k=1}^{T_{step}} [1 + \lambda \Delta t A_k - \mu \Delta t]$$

where the generic element $P_{ij}(T_{step})$ represents the probability that the disease propagates from a node i at time 1 to a node j at time T_{step} when λ is close to

λ_c in the quenched mean-field approximation Wang et al. [131] and Gómez et al. [43]. In the discrete time problem the epidemic threshold can be found solving the equation

$$\rho[P(T_{step})] = 1 \tag{3.4.1}$$

for the spectral radius ρ. In particular when the contagion dynamic is much faster than the network evolution, the adjacency matrix is a constant $A^h = A$ and the solution of the equation 3.4.1 gives the epidemic threshold $\lambda_c = 1/\rho(A)$, which is the same result of the quenched case. On the other hand, in the annihilated case, when the dynamics is much slower than the network evolution, temporal correlations are lost and we need to consider the number of times edges were active during the whole interval T_{step}. Considering $\mathcal{A} = \sum_h A^h$, the epidemic threshold in this case is given by $\lambda_c = T_{step}/\rho(\mathcal{A})$.

The extension to the continuous time limit can be obtained considering the evolution equation for the infection propagator

$$\dot{P(t)} = P(t)[-\mu + \lambda A(t)]$$

Introducing the rescaled trasmissibility $\gamma = \lambda/\mu$ it is then possible to solve this equation in terms of series of μ Blanes et al. [17]

$$P(t) = 1 + \sum_{j>0} \mu^j P^{(j)}(t)$$

where

$$P^{(j)}(t) = \int_0^t dx_1 \int_0^{x_1} dx_2 ... \int_0^{x_{j-1}} dx_j [\gamma A(x_j) - 1][\gamma A(x_{j-1}) - 1] \cdots [\gamma A(x_1) - 1]$$

For $t = T$ the epidemic threshold can be found solving the equation $\rho[P(T)] = 1$.

In the special case of weak commutation $\left[A(t), \int_0^t A(t')dt'\right]$, $P(T) = e^{T[-\mu+\lambda\langle A\rangle]}$ and the threshold is given by

$$\lambda_c = \mu/\rho(\langle A\rangle)$$

The epidemic threshold of the AD model can be retrieved by this method using the asymptotic form of the average adjacency matrix $\langle A\rangle_{ij} = (m\delta/N)(a_i + a_j)$ of the model, where a_i is the activity of the node i, m is the number of links generated by an active node, and δ is the lasting duration of the contacts between nodes.

CHAPTER 4

Epidemic Spreading and Aging in Temporal Networks with Memory

4.1. Introduction

In this chapter, we study the SIS and SIR epidemic processes on activity-driven time-varying networks with memory. We formulate the *activity-based mean-field* (ABMF) approach, deriving analytically a prediction for the epidemic threshold as a function of the activity distribution and of the parameter tuning the memory. The results show that memory overall reinforces the effects of activity fluctuations, leading to a lower value for the epidemics threshold.

The mean-field approach provides exact results when the epidemics start after the network has evolved for a long time In this regime, the dynamical process is equivalent to an epidemic model defined on an effective static network, explained in the dissertation. We show however that strong aging effects are present and that, in the preasymptotic regime, the epidemic threshold is deeply affected by the starting time of the epidemics. In particular, due to memory, at short times the dynamics displays correlations among the infection probabilities of the nodes which have already been in contact with. The correlations give rise to backtracking effects that cannot be neglected. In this case, typically the threshold of the SIS and SIR models are respectively smaller and larger than the mean-field prediction. We explain in detail the origin of such deviations, opening new perspectives for epidemic control of disease and information spreading on temporal networks with high correlations.

The chapter is organized as follows. In Section 2 we will refresh the activity-driven model for network topology in the presence of memory and how the SIS and SIR dynamics evolve on top of it. In Section 3, after a brief reminding of the analytical approach to epidemic dynamics on memoryless activity-driven networks, we will describe in detail the effects of the introduction of the memory to this method, deriving predictions for the epidemic threshold. In Section 4 we compare analytical predictions with numerical results, obtained by considering both an effective static network and the full time-evolution of the topology. The final Section presents some concluding remarks and perspectives for future work.

4.2. The model

4.2.1. Activity-Driven Networks with memory. In activity-driven models, each node v_i ($i = 1, \ldots, N$) of the graph G_t has an activity a_i assigned randomly according to a given distribution $F(a)$. The dynamics occurs over discrete temporal steps of length Δt. At each step, with probability $a_i \Delta t$ the vertex v_i becomes active and gets linked to m other vertices. Connections last for a temporal interval Δt. At the next time step $t + \Delta t$ all existing edges are deleted and the procedure is iterated. Notice that the activity a is a probability per unit time. Real data observations indicate that human interactions are very often characterized by skewed and long tailed activity distributions so $F(a)$ is typically assumed to be a power-law, $F(a) = Ba^{-(\nu+1)}$ with $\varepsilon \leq a_i \leq A$. Since in our simulations we will keep the time interval $\Delta t = 1$, the upper cutoff is naturally set to $A = 1$.

In order to consider the tendency of individuals to persist in their social connections, we can introduce a "reinforcement" mechanism. The nodes are endowed with a memory of their previous contacts and they contact preferably individuals belonging to their social circle. For an active node v_i, which has already contacted $k_i(t)$ different nodes at time t, this process is described by assuming that the node connects with a new node with probability

$$p[k_i(t)] = [1 + k_i(t)/c_i]^{-\beta_i}. \tag{4.2.1}$$

Complementary, with probability $1 - p[k_i(t)]$ the node establishes a connection with a previously contacted. The parameter β_i controls the memory process and the constant c_i sets an intrinsic value for the number of connections that node v_i is able to engage in before memory effects become relevant. The probability depends on the degree of the integrated network at time t, $k_i(t)$, i.e, the number of nodes that v_i has contacted up to time t. We will call $A_{ij}(t)$ the adjacency matrix of this integrated network. Empirical measures on several datasets are compatible with constant values of β_i and c_i so we will consider $c_i = 1$, and set $\beta_i = \beta$, independently of the site i. With this choice, the function $p(\cdot)$ is independent from i. For now on we will consider that the number of link m generated at each time-step is equal to one, with no loss of generality.

As shown in [121] the asymptotic form of the degree distribution for the integrated network can be derived analytically. In particular, in the regime $1 \ll k \ll N$ the degree of nodes of activity a is narrowly distributed around the average value

$$\bar{k}(a, t) = C(a) t^{1/(1+\beta)}, \tag{4.2.2}$$

i.e. the degree of each node increases sublinearly in time, with a prefactor depending on its activity. The prefactor $C(a)$ is determined by the condition

$$\frac{C(a)}{1+\beta} = \frac{a}{C^{\beta}(a)} + \int da \frac{F(a) a}{C^{\beta}(a)}. \tag{4.2.3}$$

In the memoryless case $\beta = 0$, where an active node connects always with a randomly chosen vertex, Eq. (4.2.3) gives $C(a) = a + \langle a \rangle$ recovering the result of [113]. We will denote in general with $\langle g \rangle = \int da F(a) g(a)$ the average of a function of the activity $g(a)$ over the network.

4.2.2. The epidemic process. We now turn to the spreading of infectious diseases on activity-driven temporal networks with memory. We start by considering the standard Susceptible-Infected-Susceptible (SIS) model, introduced in the previous chapter, i.e. the simplest description of a disease not conferring immunity. An infected node can turn spontaneously susceptible with rate μ, while an infected node transmits the infection over an edge to a susceptible neighbor with rate λ. The two elementary events are therefore:

$$I + S \xrightarrow{\lambda} 2I \qquad\qquad I \xrightarrow{\mu} S \tag{4.2.4}$$

In the Susceptible-Infected-Recovered (SIR) model the disease confers immunity, and the dynamics is described by the following reaction scheme:

$$I + S \xrightarrow{\lambda} 2I \qquad\qquad I \xrightarrow{\mu} R \tag{4.2.5}$$

The epidemic process on activity-driven networks is implemented by iterating discrete time steps of duration Δt:

- at the beginning of each time step there are N disconnected vertices;
- with probability $a_i \Delta t$ a vertex v_i becomes active and connects to a previously linked node with probability $1 - p(k_i)$, or with a new node v_j with probability $p(k_i)$, in this second case $k_i(t)$, $k_j(t)$ and $A_{ij}(t)$ are increased by one unit;
- if one of the nodes connected by the link is infected and the other one is susceptible, the susceptible becomes infected with probability λ;
- a vertex v_j, if infected, becomes susceptible (SIS), or recovers (SIR) with probability $\mu \Delta t$.

In activity-driven models, λ is a pure number, i.e. the probability that in a single contact the infection is actually transmitted, while μ is still the rate of recovery for a single individual. Ignoring the inhomogeneity in the activities, one can estimate the total rate for the infection process per node as $\lambda \langle k' \rangle$, where $\langle k' \rangle = 2 \langle a \rangle$ is the average degree per unit time; this is the quantity to be compared with the recovery rate per node μ.

4.3. Analytical results

Epidemics on memoryless activity-driven networks. The epidemic spreading for the memoryless case $\beta = 0$ has been studied in adopting an ABMF approach. The epidemic state of a node, when averaged over all possible dynamical evolutions,

only depends on the value of its activity a_i. In particular, one can define the probability $\rho(a_i, t)$ that a node with activity a_i is infected at time t. The corresponding evolution equation is:

$$\partial_t \rho(a_i) = -\mu\rho(a_i) + \lambda[1 - \rho(a_i)] + \frac{1}{N-1}\sum_{j\neq i}[a_i\rho(a_j) + a_j\rho(a_j)] \tag{4.3.1}$$

The first term on the right side is due to recovery events; the second term takes into account the event that a susceptible node of class a_i becomes active and contracts the disease by connecting to an infected individual, while the third term is the analogous term for the case of a susceptible node that, independently of her own activity, is contacted by an infected active individual.

The description in terms of quantities that only depend on the activity is conceptually analogous to the heterogeneous-mean-field approach for dynamical processes on static networks [89]. In that case, one assumes that the only property determining the epidemic state of a node is the degree k and then derives equations for the probabilities ρ_k. An important difference must however be stressed. Assuming the epidemic state to depend only on the degree is an approximation for static networks, because it neglects the quenched nature of the network structure that makes properties of nodes, with the same degree but embedded in different local environments, different. In practice, this assumption is equivalent to replacing the actual adjacency matrix of the network (A_{ij} equal to 0 or 1 depending on the presence of the connection between v_i and v_j) with an annealed adjacency matrix $P_{ij} = k_i k_j/(\langle k\rangle N)$ [35], expressing the probability that vertices v_i and v_j with degree k_i and k_j are connected. The annealed approach is an approximation for static networks, while it is exact for networks where connections are continuously reshuffled at each time step of the dynamics, since the reshuffling process destroys local correlations. Because in memoryless activity-driven networks connections are extracted anew at each time step, the ABMF approach provides exact results in this case.

Equation (4.3.1) can be analyzed by means of a linear stability analysis, yielding, for large N, the threshold [94]

$$\left(\frac{\lambda}{\mu}\right)_{\mathrm{ML}} = \frac{1}{\langle a\rangle + \sqrt{\langle a^2\rangle}}. \tag{4.3.2}$$

The same result can be derived for the SIR case.

Epidemics on activity-driven networks with memory.

Individual-based mean-field approach. In presence of memory, interactions occur preferably with a subset of the other nodes (the social circle) creating correlations. Therefore, we implement a different, individual-based, mean-field approach, keeping explicitly track of the evolution of social contacts (i.e. of the memory). Let us first consider the SIS model. The observable of interest is the probability $\rho_i(t)$

that node v_i is infected at time t. Its evolution can be written as

$$\begin{aligned}\partial_t \rho_i(t) = -\mu\rho_i(t) + \\ \lambda\left[1-\rho_i(t)\right]\Big\{\textstyle\sum_j a_i\left[1-p(k_i)\right]\frac{A_{ij}(t)}{k_i}\rho_j(t) + \sum_{j\nsim i} a_i p(k_i)\frac{1}{N-k_i-1}\rho_j(t) + \\ \textstyle\sum_j a_j\left[1-p(k_j)\right]\frac{A_{ij}(t)}{k_j}\rho_j(t) + \sum_{j\nsim i} a_j p(k_j)\frac{1}{N-k_j-1}\rho_j(t)\Big\}\end{aligned} \tag{4.3.3}$$

Here $j \nsim i$ indicates the sum over the nodes j not yet connected to i, $N - k_j(t) - 1$ is their number. The quantity $A_{ij}(t)$ is the adjacency matrix of the time-integrated network at time t, i.e., it is equal to 1 if v_i and v_j have been in contact at least once in the past and 0 otherwise. In Eq. (4.3.3), the only approximation made is that the dynamical state of every node is considered to be independent of the state of the partner in the interaction; in other words, we neglect the existence of dynamical correlations among nodes, which are created by the partially quenched nature of the interaction pattern due to memory. This is the same approximation that is involved by the individual-based mean-field approach for static networks [93] discussed in Chapter 3.

On the right hand side of Eq. (4.3.3), the first term is the recovery rate of $\rho_i(t)$. The second term, describing the infection process, is the product of λ times the probability for v_i to be susceptible and, in curly brackets, the fraction of infected nodes contacted by v_i per unit time. In the curly brackets, the first and the second term describe the case where v_i is active and connects to the infected node v_j taking into account that the link can be an old or a new one respectively. In the same way, the third and the fourth term represent the probabilities that v_i is contacted by an infected and active node v_j.

Since both $A_{ij}(t)$ and $k_i(t)$ depend on the evolution time t, the behavior of the epidemics can strongly depend on the starting time of the outbreak, giving rise to aging effects that will be investigated in numerical simulations. When the epidemic starts at very large times, an analytic approach can be considered. In this regime, with $1 \ll k_i(t) \ll N$, we expect that the creation of new contacts can be ignored and that the dynamical correlations are asymptotically negligible, since the connectivity of the integrated network becomes large. If the epidemic starts at very large times, therefore, we can apply an heterogeneous mean-field approximation for $A_{ij}(t)$, allowing for an analytical solution of the problem which we expect to be asymptotically exact.

The behavior for large times. Let's consider the regime of large times, where $1 \ll k_i(t) \ll N$ for all nodes. In this case each node has already had a large number of contacts, but that number is not too large, so that the integrated network cannot be considered as a complete graph, i.e., it is still sparse. In the limit of large N there is a large temporal interval such that this condition is fulfilled. The condition $1 \ll k_i(t) \ll N$ allows us to replace in Eq. (4.3.3) $N - k_i(t) - 1$ with N and $p(k_i)$

with $(k_i(t))^{-\beta}$. Considering only leading terms Eq. (4.3.3) becomes

$$\partial_t \rho_i(t) = -\mu \rho_i(t) + \lambda \left[1 - \rho_i(t)\right] \sum_j A_{ij}(t) \left(\frac{a_i}{k_i} + \frac{a_j}{k_j} \right) \rho_j(t). \tag{4.3.4}$$

The linking probability. To proceed further we perform the equivalent of the heterogeneous mean-field approximation for static networks, i.e., we replace the time-integrated adjacency matrix $A_{ij}(t)$ with its annealed form, $P_{ij}(t)$, i.e., the probability that v_i and v_j have been in contact in the past. The evolution of $P_{ij}(t)$ is described by the master equation:

$$\partial_t P_{ij}(t) = \left[\frac{a_i p(k_i)}{N - k_i - 1} + \frac{a_j p_j(k_j)}{N - k_j - 1} \right] \left[1 - P_{ij}(t)\right]. \tag{4.3.5}$$

In Eq. (4.3.5) P_{ij} grows either because the node v_i activates (probability per unit time a_i), it creates a new connection [probability $p(k_i)$] and the new partner is v_i [probability $(N - k_i - 1)^{-1}$] or because of the event with the role of v_i and v_j interchanged.

In the temporal interval of interest we can use again the relations holding for large times $p(k_i) \approx k_i^{-\beta}$ and $N - k_j - 1 \approx N$. Moreover, for large times, the degree of a node of activity a_i can be estimated by its average value $\bar{k}(a_i, t)$, given by Eq. (4.2.2). So we obtain

$$\partial_t P_{ij}(t) = [1 - P_{ij}(t)] \frac{g(a_i) + g(a_j)}{N t^{\frac{\beta}{1+\beta}}}, \tag{4.3.6}$$

where we have defined

$$g(a_i) = a_i / [C(a_i)]^{\beta}. \tag{4.3.7}$$

Eq. (4.3.6) can be readily solved, yielding

$$P_{ij}(t) = 1 - e^{-\frac{(1+\beta) t^{1/(1+\beta)}}{N} [g(a_i) + g(a_j)]} \tag{4.3.8}$$

In the regime $t^{1/(1+\beta)} \ll N$, $P_{ij}(t)$ becomes

$$P_{ij}(t) = (1+\beta) \frac{t^{1/(1+\beta)}}{N} [g(a_i) + g(a_j)]. \tag{4.3.9}$$

Notice that $P_{ij}(t)$ is a topological feature of the activity-driven network, independent of the epidemic process.

Asymptotic ABMF equation. We now introduce into Eq.(4.3.4) the annealed expression for the integrated adjacency matrix $A_{ij}(t) \approx P_{ij}(t) = P(a_i, a_j, t)$ and for the connectivity $k_i(t) = \bar{k}(a_i, t)$. In this way the equations depend on the nodes v_i and v_j only through their activities a_i and a_j. The equation for the probability $\rho(a, t)$ that a generic node of activity a is infected at time t is therefore:

(4.3.10)

$$\begin{aligned} \partial_t \rho(a,t) = -\mu \rho(a,t) + \\ \lambda \left[1 - \rho(a,t)\right] \Big\{ \tfrac{a g(a)}{g(a) + \langle g \rangle} \int da' F(a') \rho(a',t) + \tfrac{a}{g(a) + \langle g \rangle} \int da' F(a') \rho(a',t) g(a') + \\ g(a) \int da' F(a') \tfrac{a'}{(g(a') + \langle g \rangle)} \rho(a',t) + \int da' F(a') \tfrac{a' g(a')}{(g(a') + \langle g \rangle)} \rho(a',t) \Big\} \end{aligned}$$

where we have replaced the sums over nodes with integrals over the activities $1/N \sum_j \to \int da' F(a')$ and used Eq. (4.2.3), which can be rewritten as

$$C(a) = (1+\beta)\left[g(a) + \langle g \rangle\right]. \tag{4.3.11}$$

Eq.(4.3.10) is effectively an ABMF approach, since all the information on the behavior of the node v_i depends on its activity a_i. Note that, although Eqs. (4.3.3) and (4.3.4) described the dynamics of the individual node, the further approximation underlying Eq. (4.3.5) has transformed the approach into an ABMF one, conceptually analogous to the heterogeneous mean-field approximation on static networks, where all the information on node v_i is encoded in its degree k_i.

It is important to remark that in Eq. (4.3.4) the time dependencies of $P(a_i, a_j, t) \propto t^{1/(1+\beta)}$ and of the average degree $\bar{k}(a_i, t) \propto t^{1/(1+\beta)}$ cancel out, so that the right hand side of Eq. (4.3.10) does not depend explicitly on time. This suggests that in this temporal regime the epidemic can be seen as an activity-driven process taking place on an effective static graph, where the probability for nodes v_i and v_j to be linked is given by Eq. (4.3.9) and the quantity $t^{1/(1+\beta)}/N$ is a fixed quantity τ whose value only determines the average degree of the network. Performing simulations over an ensemble of these effective static networks and averaging the results one should then reproduce the predictions of the ABMF approach, Eq. (4.3.10).

From Equation (4.3.10), by performing a linear stability analysis around the absorbing state $\rho(a,t) = 0$, it is possible to compute analytically the epidemic threshold $(\lambda/\mu)_c$, for any value of the reinforcement parameter β and of the exponent of the analytical distribution ν. Since for large times the node degrees diverge and correlations can be neglected, we expect the linear stability analysis to provide the correct estimate of the epidemic threshold when the epidemics start at very long times i.e. when the degrees $k_i(t)$ have already become very large.

The results of the linear stability analysis are presented in Fig. 4.3.1 showing that the thresholds are smaller than in the memoryless case. This lower value is a consequence of the fact that memory reinforces the activity fluctuations, and in these models fluctuations clearly reduce the the epidemic threshold, as shown by Eq. (4.3.2). The effect can be simply understood since nodes with large activity have also a large degree, therefore they are easily involved in epidemic contacts not only because they are frequently activated but also because they are frequently contacted by other nodes. In this way memory reinforces the effect of activity fluctuations. In this framework, Fig. 4.3.1 also shows that at large ν i.e. for increasingly smaller fluctuations, the difference with the memoryless model vanishes. In particular, for $F(a) = \delta(a - a_0)$ i.e. when the activity does not fluctuate, one obtains from Eq.(4.3.10) $\partial_t \rho(t) = -\mu\rho(t) + 2a_0\lambda[1-\rho(t)]$ that is the same equation of the memoryless case. This also explains the quite surprising observation that the threshold is a growing function of β, converging to the memoryless case as $\beta \to \infty$.

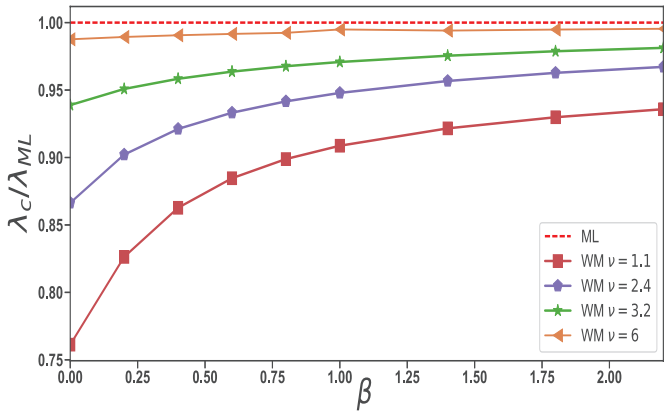

FIGURE 4.3.1. Plot of the ratio λ_c/λ_{ML} between the epidemic thresholds in the memory and in the memoryless (ML) cases, for different values of the exponent ν of the distribution $F(a) = Ba^{-(\nu+1)}$. The dashed lines are the mean-field memoryless results, while the solid lines are the outcomesnof the ABMF equations in presence of memory.

Indeed, the tail of the degree distribution decays at large k as $k^{-[(1+\beta)\nu+1]}$, therefore at large β we get a faster decay and smaller degree fluctuations. For the same reason, in the limit $\beta \to 0$ the difference with the memoryless case is maximal, since degree inhomogeneities are stronger in this case.

We remark that in Eq. (4.3.10), as in the memoryless case, dynamical correlations are ignored. However, we expect that at finite times, due to the finite connectivity of the integrated graph, the effect of correlations becomes important. The memory process leads to the formation of small clusters of mutually connected high activity vertices, which become reservoirs of the disease in the SIS model. The high frequency of mutual contacts allows for reinfection, favoring the overall survival of the epidemic spreading in the system. In this way, social circles with high activity play a role analogous to that played by the max K-core or the hub and its immediate neighbors for SIS epidemics in static networks [26, 27]. To clarify the effect of dynamical correlations at finite time, in the next Section we compare the analytical predictions with results of numerical simulations. As a final remark, we note that, in the asymptotic ABMF approach, the linear stability analysis also holds for the SIR model, implying that the epidemic threshold is the same of the SIS model. However, in the SIR model reinfection is not allowed so that the initial presence of small clusters of mutually connected high activity vertices effectively inhibits the spread of the disease. For this reason, we expect that finite connectivity (i.e. finite time) increases the epidemic threshold with respect to the mean-field result, as we will check in numerical simulations.

4.4. Linear Stability Analysis

The dynamical process is described by the ABMF equation [Eq. (4.3.10)] which we rewrite as

$$\partial_t \rho(a) \tag{4.4.1}$$
$$\lambda\left[1-\rho(a)\right]\left[A(a)g(a)\left\langle\rho(a)\right\rangle + A(a)\left\langle g(a)\rho(a)\right\rangle + g(a)\left\langle A(a)\rho(a)\right\rangle + \left\langle A(a)g(a)\rho(a)\right\rangle\right]$$

where for simplicity we have omitted the time dependencies and defined $A(a) = a/[g(a) + \langle g(a)\rangle]$.

To study the stability of the system linearized around the fixed point $\rho(a) = 0$, we introduce the following functions

$$\begin{aligned} \rho &= \langle \rho(a)\rangle \\ x &= \langle g(a)\rho(a)\rangle \\ y &= \langle A(a)\rho(a)\rangle \\ z &= \langle A(a)g(a)\rho(a)\rangle \end{aligned}$$

Integrating Eq. (4.4.1) over a and keeping only linear terms in $\rho(a)$ we obtain an equation for $\partial_t\rho$. Similarly, multiplying Eq. (4.4.1) by $g(a)$ and integrating over a we get and equation for $\partial_t x$. Doing the same for y and z we obtain a closed system of four equations for four variables

$$\begin{aligned} \partial_t\rho &= -\mu\rho + \lambda\left[\langle A(a)g(a)\rangle\,\rho + \langle A(a)\rangle\,x + \langle g(a)\rangle\,y + z\right] \\ \partial_t x &= -\mu x + \lambda\left[\left\langle A(a)g^2(a)\right\rangle\rho + \langle A(a)g(a)\rangle\,x + \left\langle g^2(a)\right\rangle y + \langle g(a)\rangle\,z\right] \\ \partial_t y &= -\mu y + \lambda\left[\left\langle A^2(a)g(a)\right\rangle\rho + \left\langle A^2(a)\right\rangle x + \langle A(a)g(a)\rangle\,y + \langle A(a)\rangle\,z\right] \\ \partial_t z &= -\mu z + \lambda\left[\left\langle A^2(a)g^2(a)\right\rangle\rho + \left\langle A^2(a)g(a)\right\rangle x + \left\langle A(a)g^2(a)\right\rangle y + \langle A(a)g(a)\rangle\,z\right] \end{aligned}$$

These equations describe the epidemic near the state $\rho(a) = 0$ and the jacobian matrix of this system of equations is

$$J = \begin{pmatrix} \lambda\langle Ag\rangle - \mu & \lambda\langle A\rangle & \lambda\langle g\rangle & \lambda \\ \lambda\left\langle Ag^2\right\rangle & \lambda\langle Ag\rangle - \mu & \lambda\left\langle g^2\right\rangle & \lambda\langle g\rangle \\ \lambda\left\langle A^2g\right\rangle & \lambda\left\langle A^2\right\rangle & \lambda\langle Ag\rangle - \mu & \lambda\langle A\rangle \\ \lambda\left\langle A^2g^2\right\rangle & \lambda\left\langle A^2g\right\rangle & \lambda\left\langle Ag^2\right\rangle & \lambda\langle Ag\rangle - \mu \end{pmatrix}$$

The state $\rho(a) = 0$ is stable provided all eigenvalues of J are negative, hence the epidemic threshold is given by the value $(\lambda/\mu)_c$ such that largest eigenvalue of the Jacobian matrix is zero. Numerical evaluation of the matrix J and of its eigenvalues can be obtained, first by solving numerically Eq. (4.2.3) to get $C(a)$ and $g(a)$ and then calculating the averages defining J.

4.5. Numerical simulations

SIS model on the effective static network. As discussed above, Eq. (4.3.10) can be interpreted as a heterogeneous mean-field equation for a SIS epidemic on an

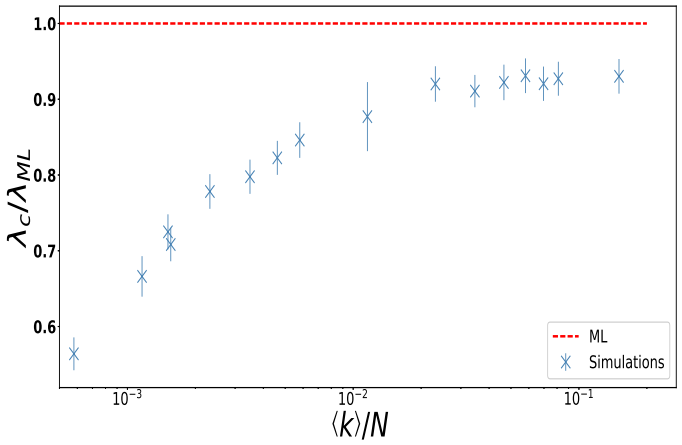

FIGURE 4.5.1. Ratio between the epidemic threshold found in simulations and the estimate given by equation Eq.(4.3.2) valid for the memoryless model, as a function of $\log(\langle k\rangle/N)$. For $\langle k\rangle/N > 0.01$, we observe practically no dependence on $\langle k\rangle$.

effective static network where the probability that v_i and v_j are connected is

$$P_{ij} = P(a_i, a_j) = \tau(1+\beta)[g(a_i) + g(a_j)]. \tag{4.5.1}$$

Here $\tau \ll 1$ is a constant, $g(a) = a/[C(a)]^\beta$ and $C(a)$ is a function that can be evaluated numerically for $\beta > 0$, while for $\beta = 0$ it takes the simple form $C(a) = a + \langle a\rangle$. The constant τ can be tuned in order to set the average degree of the network, because

$$k(a) = N\int da' F(a')P(a,a') = (1+\beta)N\tau[g(a) + \langle g\rangle], \tag{4.5.2}$$

so that

$$\langle k\rangle = \int da' F(a')k(a') = 2(1+\beta)N\tau\,\langle g\rangle\,. \tag{4.5.3}$$

We now study the SIS epidemic evolution on the effective static network.

Given the activity of each node, extracted according to the distribution $F(a)$, for each of the possible pairs of nodes we place an edge with probability given by Eq. (4.5.1). On top of this quenched topology we run a memoryless activity-driven SIS dynamics, starting with 10% of the nodes in the infected state, until the stationary state is reached and we record the fraction of infected nodes. We repeat the procedure many times for each value of λ, while μ is fixed to 0.015. We determine the threshold as the position of the maximum of the susceptibility [39] $\chi = N(\overline{\rho^2} - \overline{\rho}^2)/\overline{\rho}$, where the overbar denotes the average over dynamical realizations at fixed topology. We repeat all this for several networks obtained using different sequences of activities and different samplings of P_{ij} and we average the epidemic thresholds found for each of them.

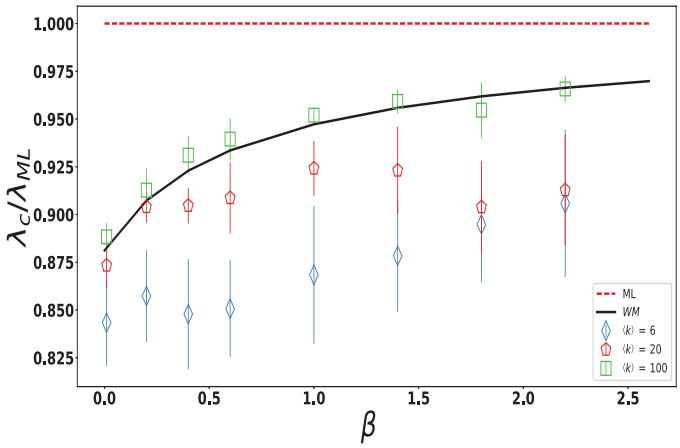

FIGURE 4.5.2. Ratio between the epidemic threshold with memory and the epidemic threshold of the memoryless case as a function of the reinforcement parameter $\beta = [0.01, 0.2, 0.4, 0.6, 1, 1.4, 1.8, 2.2]$, for simulations on the effective static network with $\nu = 2.4$, $\epsilon = 0.01$, $N = 5 \cdot 10^4$. The points are averages of different realizations of the network with different sequences of activity $a_1, a_2, \ldots, a_N$: 32 realizations for $\langle k \rangle = 6$, 16 realizations for $\langle k \rangle = 20$, 4 realizations for $\langle k \rangle = 100$

We first check that, as long as $1 \ll \langle k \rangle \ll N$, the results are independent of the exact value of $\langle k \rangle$, as predicted by the theory. Fig. 4.5.1 shows, for $\beta = 1$, that the effective threshold initially grows with $\langle k \rangle$ but then reaches a plateau, in accordance with the expectations.

In Fig. 4.5.2 we report the dependence of the effective epidemic threshold as a function of β for three values of the average degree $\langle k \rangle$, compared with the predictions of the ABMF theory with and without memory. We observe that, as $\langle k \rangle$ grows, numerical results tend to coincide with theoretical predictions.

On the other hand, for small values of $\langle k \rangle$ the value of the threshold is smaller than the one predicted theoretically. Indeed, on effective static networks with small connectivity we expect the presence of clusters of mutually interconnected nodes to be relevant, as they are able to reinfect each other several times. It is well known that for the SIS model these backtracking effects tend to lower the epidemic threshold since social circles with high activity favor the overall survival of the epidemic.

Epidemics on time-evolving networks. Let us now consider simulations of the epidemic spreading on the full time evolving network. We consider a graph of size $N = 5 \cdot 10^4$ with activity distributed according to $F(a) = Ba^{-(\nu+1)}$ ($\nu = 2.4$) and a cutoff $\varepsilon = 10^{-2}$. To extract the activities of individual nodes we perform an importance sampling so that, even in the finite size system, the moments $\langle a \rangle$ and $\langle a^2 \rangle$ coincide with their expected values.

We first start the temporal evolution of the network and at a later time t_0 we let the epidemic begin. We evaluate at t_0 the average connectivity of the nodes $\langle k \rangle_0$ which measures the evolution of the network at the starting time. In both the SIS and SIR models, we use two different initial conditions. The first is to randomly select (RA) the node to infect at time t_0, Fig. 4.5.4 and Fig. 4.5.5 while the second is to infect the most active node (MA) at time t_0. We keep the recovery rate fixed at $\mu = 1.5 \cdot 10^{-2}$ and vary the probability of infection λ to study the dependence of its critical value on the memory parameter β.

SIS model. In the SIS model, we determine the epidemic threshold using the lifespan method. We plot (see Fig. 4.5.3), as a function of the parameter λ, the average lifespan of simulations ending before the coverage (i.e. the fraction of distinct sites ever infected) reaches a preset value that we take equal to 1/2. The threshold is estimated as the value of λ for which the lifespan has a peak.

The epidemic thresholds of numerical simulations are compared with theoretical predictions in Fig. 4.5.4 (RA case) and 4.5.5 (MA case). Numerical results converge toward the analytical prediction as $\langle k \rangle_0$ becomes larger, while there are strong deviations for small $\langle k \rangle_0$.

For small $\langle k \rangle_0$ two competing effects are at work. First, infections are mediated by an effective static network with small connectivity, therefore we expect backtracking effects to enhance epidemic spreading and to lower the threshold. However, in this case, moves connecting new partners are also possible. In these moves nodes are chosen randomly in the whole system and the epidemic dynamics is memoryless, leading to a higher epidemic threshold. So there exists a competition between backtracking correlations and memoryless moves which reduce and increase the threshold, respectively. Clearly for large $\langle k \rangle_0$ both effects become negligible and the ABMF result is recovered. However, at small β the memoryless moves are more probable and indeed the threshold are larger, while for large β memory effects are more relevant. We remark that the case $\beta = 0$ coincides with the memoryless case (ML) and the dynamics never occurs on the effective static network. On the other hand, for any $\beta > 0$ at sufficiently large value of $\langle k \rangle_0$ the dynamics is dominated by memory and infections spread on the effective static network. This originates a singular behavior where the limits $\beta \to 0$ and $\langle k \rangle_0 \to \infty$ do not commute.

Finally Figs. 4.5.4 and 4.5.5 show that backtracking effects (leading to small thresholds) are strong when the evolution starts from the most active site, while they are negligible with random initial conditions. The most active node indeed has the largest degree and it forms a cluster of highly activated nodes where the high frequency of mutual contacts allows for reinfections and positive correlations. Conversely, the average site has a small connectivity and can activate new links with high probability giving rise essentially to a memoryless epidemic dynamics.

SIR model. The results of simulations of the SIR process are displayed in Fig. 4.5.6 and Fig 4.5.7 for the RA and MA case respectively. The threshold is estimated

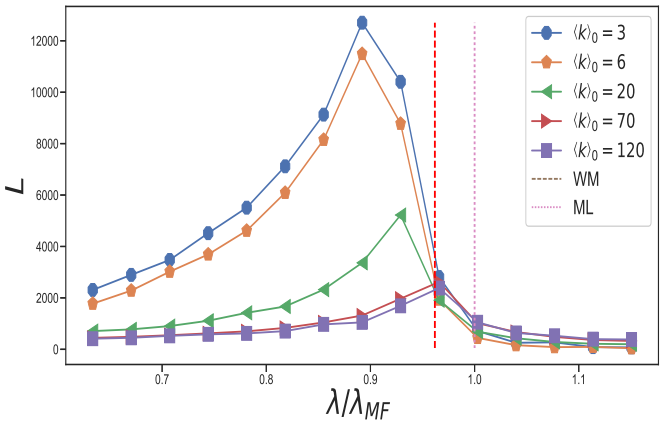

FIGURE 4.5.3. SIS epidemic process. Lifespan (L) as function of the ratio between the epidemic threshold with memory and the epidemic threshold of the memoryless for different values of $\langle k \rangle_0$. $N = 5 \cdot 10^4$, $\nu = 2.4$, $a \in [10^{-2}, 1]$. We consider $4 \cdot 10^3$ epidemic realizations for each value of λ. Dynamics starts from the most active site and at small $\langle k \rangle_0$ back-tracking effects are dominant favoring the epidemic spreading; this on one side lowers the value of the threshold (value of λ corresponding to the peak) but also increases the lifespan of the system at small λ.

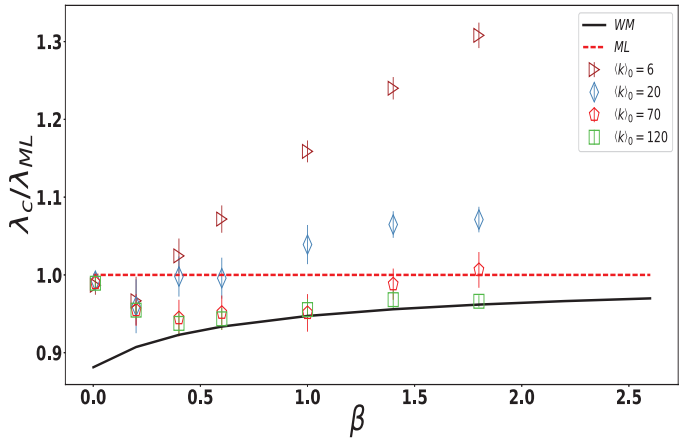

FIGURE 4.5.4. SIS epidemic process, RA. Ratio between the epidemic threshold with memory and the epidemic threshold of the memoryless case as a function of the reinforcement parameter $\beta = [0.01, 0.2, 0.4, 0.6, 1, 1.4, 1.8, 2.2]$. $N = 5 \cdot 10^4$, $\nu = 2.4$, $a \in [10^{-2}, 1]$. The dotted line represents the memoryless result (ML), the solid line is the analytical prediction obtained from Eq. (4.3.10) in the memory case (WM). We consider $4 \cdot 10^3$ epidemic realizations for each value of λ.

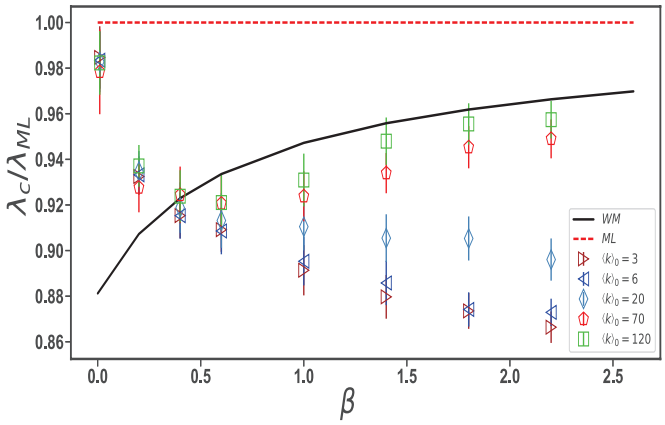

FIGURE 4.5.5. SIS epidemic process, MA. Ratio between the epidemic threshold with memory and the epidemic threshold of the memoryless case as a function of the reinforcement parameter $\beta = [0.01, 0.2, 0.4, 0.6, 1, 1.4, 1.8, 2.2]$. $N = 5 \cdot 10^4$, $\nu = 2.4$, $a \in [10^{-2}, 1]$. The dotted line represents the memoryless result (ML), the solid line is the analytical prediction obtained from Eq. (4.3.10) (WM). We consider $4 \cdot 10^3$ epidemic realizations for each value of λ.

from the peak of the variability $\Delta = \sqrt{\langle N_R^2\rangle - \langle N_R\rangle^2}/\langle N_R\rangle$, i.e., the standard deviation of the number of recovered nodes N_R at the end of the simulation. As for SIS, at large $\langle k\rangle_0$ dynamical correlations can be neglected and simulations recover the ABMF result. Simulations clearly show that now correlations at small $\langle k\rangle_0$ inhibit the epidemic spreading and the critical threshold becomes larger. As in the SIS model, at small β the memory is negligible and the dynamics is driven by the creation of new links, so that the threshold values are close to the memoryless case (ML) almost independently of $\langle k\rangle_0$. On the other hand, for larger β, the epidemics evolves on the integrated network, dynamical correlations become important and the thresholds grow even larger than in the memoryless case.

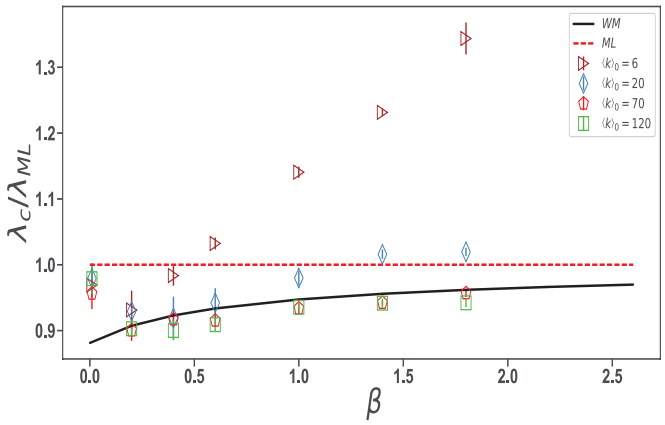

FIGURE 4.5.6. SIR epidemic process, RA. Ratio between the epidemic threshold with memory and the epidemic threshold of the memoryless case as a function of the reinforcement parameter $\beta = [0.01, 0.2, 0.4, 0.6, 1, 1.4, 1.8]$. $N = 5 \cdot 10^4$, $\nu = 2.4$, $a \in [10^{-2}, 1]$. The dotted line represents the memoryless result (ML), the solid line is the analytical prediction obtained from Eq. (4.3.10) in the memory case (WM). We consider $2{\cdot}10^3$ epidemic realizations for each value of λ.

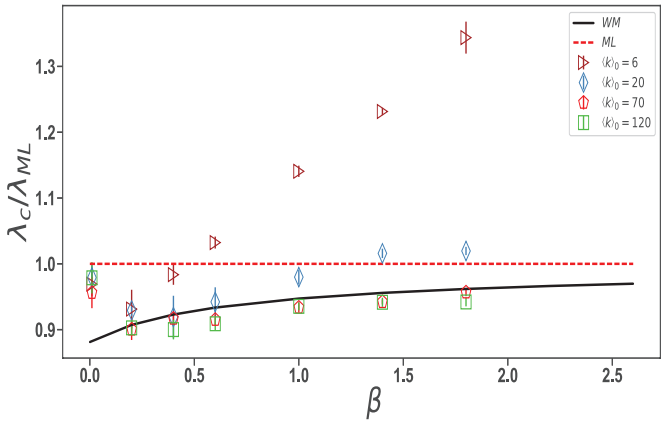

FIGURE 4.5.7. SIR epidemic process, MA. Ratio between the epidemic threshold with memory and the epidemic threshold of the memoryless case as a function of the reinforcement parameter $\beta = [0.01, 0.2, 0.4, 0.6, 1, 1.4, 1.8]$. $N = 5 \cdot 10^4$, $\nu = 2.4$, $a \in [10^{-2}, 1]$. The dotted line represents the memoryless result (ML), the solid line is the analytical prediction obtained from Eq. (4.3.10) in the memory case (WM). We consider $4{\cdot}10^3$ epidemic realizations for each value of λ.

CHAPTER 5

Conclusions

In this work we have stressed out the importance of time-varying networks when processes evolve on the same time-scale of the evolving network, in particular we consider the activity-driven model. In this framework we have considered a mechanism that emerge from the analysis of large dataset that can be interpreted as the memory that individuals have of their pears in their social circle. We have presented our analytical and numerical results providing a complete understanding of the interplay between the temporal evolution of the activity-driven network with memory and the epidemic process occurring on top of it, focusing our attention on the SIS and SIR epidemic model. To this end we have first used an individual-based mean-field approach for the epidemic process on the activity-driven network with memory, keeping explicitly track of social contacts. We have then studied the behavior for large time of the system, hence we have considered the limit in which each node has already had a large number of contacts, but the time-integrated graph can still be considered sparse. In this regime we have performed an approximation equivalent to the heterogeneous mean-field for static networks and assigned the description of the topology of the TVN to the linking probability that two nodes have been in contact in the past. At this point we have introduced the asymptotic activity-based mean-field description of the dynamics where all the information on the behavior of the node depends on its activity. In this framework the explicit time dependency of the dynamical equation disappear, and, in this temporal regime, the epidemic can be considered as an activity-driven process on the effective static network built with the linking probability. Moreover, in the case of large connectivity, correlations can be neglected and performing the linear stability analysis we have obtained the epidemic threshold as a function of the memory parameter, expected to be exact at very long times.

Comparing the analytic results with the numerical simulations on both the static and the time-varying networks, we have seen that the starting time of the outbreak has crucial consequences on the epidemic threshold.

In the long time limit the reinforcement mechanism of the topological evolution completely inhibits the formation of new connections. When the activity-driven epidemic dynamics starts at this stage, it takes place on a topology which can be considered as static. All nodes have a very large number of connections and hence the mean-field theory is asymptotically exact. The epidemic threshold, which is the

same for SIS and SIR dynamics, is lower then the memoryless case, because memory enhances the effect of activity fluctuations, as also confirmed by the simulations.

If instead the epidemic process starts before the memory has completely taken over, interesting model-dependent preasymptotic effects are observed. The fundamental observation is that at this stage nodes with large activity tend to interact with their social circles, while less active nodes still tend to explore the system creating new connections. The first type of interaction tends to facilitate the spreading in the SIS model, while the second tends to suppress it. This leads to positive or negative corrections to the asymptotic value of the threshold, depending on the initial conditions and on the reinforcement parameter β. In the SIR case instead, since reinfection is not possible, the interaction within social circles is strongly detrimental for the epidemic propagation, so that the asymptotic value of the threshold is always larger. Hence our results allow to fully understand the contrasting effects of strong ties on SIS and SIR dynamics observed in Ref. [117].

Finally in this work we have fully explained the mechanism that the memory introduces in the activity-driven network. We have introduced an agile framework to study spreading processes on highly correlated temporal network, opening new perspectives to control epidemic and information dynamics.

Several possible extensions of the model considered here are possible to make it more realistic, both in terms of the topological evolution and of the spreading process, among them probably the most interesting would be the inclusion of burstiness in agents activity. We have seen the effect on activity-driven network of tie reinforcement and non exponentially-distributed inter-event times in Chapter 2, in this framework the inclusion of a spreading dynamics is a promising and challenging path for future research.

Bibliography

[1] Lada A. Adamic and Bernardo A. Huberman. "The Web's Hidden Order". In: *Commun. ACM* 44.9 (Sept. 2001), pp. 55–60. ISSN: 0001-0782. DOI: 10.1145/383694.383707. URL: http://doi.acm.org/10.1145/383694.383707.

[2] E. Agliari et al. "Efficiency of information spreading in a population of diffusing agents". In: *Phys. Rev. E* 73 (4 Apr. 2006), p. 046138. DOI: 10.1103/PhysRevE.73.046138. URL: http://link.aps.org/doi/10.1103/PhysRevE.73.046138.

[3] William Aiello, Fan Chung, and Linyuan Lu. "A random graph model for power law graphs". In: *Experiment. Math.* 10.1 (2001), pp. 53–66. URL: http://projecteuclid.org/euclid.em/999188420.

[4] Réka Albert and Albert-László Barabási. "Statistical mechanics of complex networks". In: *Reviews of Modern Physics* 74.1 (2002), pp. 47–97. ISSN: 0034-6861. DOI: 10.1103/RevModPhys.74.47. arXiv: 0106096v1 [arXiv:cond-mat]. URL: http://dx.doi.org/10.1103/RevModPhys.74.47%7B%5C%%7D5Cnpapers2://publication/doi/10.1103/RevModPhys.74.47%7B%5C%%7D5Cnhttp://stacks.iop.org/1478-3975/1/i=3/a=006?key=crossref.7e041937ef77358d60afe44f40925c5b%7B%5C%%7D5Cnhttp://link.aps.org/doi/10.1103/RevModPhys.74.47.

[5] Reka Albert, Hawoong Jeong, and Albert-Laszlo Barabasi. "Internet: Diameter of the World-Wide Web". In: *Nature* 401.6749 (Sept. 1999), pp. 130–131. URL: http://dx.doi.org/10.1038/43601.

[6] R.M. Anderson and R.M. May. *Infectious Diseases of Humans: Dynamics and Control*. Dynamics and Control. OUP Oxford, 1992. ISBN: 9780198540403. URL: https://books.google.it/books?id=HTO--xXBguQC.

[7] Roland H. Ggkvist Armen S. Asratian. *Bipartite Graphs and Their Applications*. CAMBRIDGE UNIV PR, May 11, 2004. 272 pp. ISBN: 052159345X. URL: https://www.ebook.de/de/product/4033924/armen_s_asratian_roland_h_ggkvist_bipartite_graphs_and_their_applications.html.

[8] Norman T. J. Bailey. *Mathematical Theory of Epidemics*. Arnold, 1957. ISBN: 0852641133. URL: http://www.amazon.com/Mathematical-Theory-Epidemics-Norman-Bailey/dp/0852641133%3FSubscriptionId%3D0JYN1NVW651KCA56C102%26tag%3Dtechkie-20%26linkCode%3Dxm2%26camp%3D2025%26creative%3D165953%26creativeASIN%3D0852641133.

[9] Paolo Bajardi et al. "Dynamical Patterns of Cattle Trade Movements". In: *PLoS ONE* 6.5 (May 2011), e19869. DOI: 10.1371/journal.pone.0019869. URL: http://dx.doi.org/10.1371%2Fjournal.pone.0019869.

[10] D. Balcan et al. "Multiscale mobility networks and the spatial spreading of infectious diseases". In: *PNAS* 106 (2009), p. 21484. DOI: 10.1073/pnas.0906910106.

[11] Albert László Barabási. "The origin of bursts and heavy tails in human dynamics". In: *Nature* 435.7039 (May 2005), pp. 207–211. ISSN: 00280836. DOI: 10.1038/nature03459. arXiv: 0505371 [cond-mat]. URL: http://www.nature.com/articles/nature03459.

[12] A. Barrat, M. Barthélemy, and A. Vespignani. *Dynamical processes on complex networks*. Cambridge University Press, 2008. ISBN: 9780511791383. DOI: 10.1017/CBO9780511791383. URL: http://dx.doi.org/10.1017/CBO9780511791383.

[13] A Barrat et al. "The architecture of complex weighted networks." In: *Proceedings of the National Academy of Sciences of the United States of America* 101.11 (Mar. 2004), pp. 3747–52. ISSN: 0027-8424. DOI: 10.1073/pnas.0400087101. URL: http://www.ncbi.nlm.nih.gov/pubmed/15007165%20http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC374315.

[14] Marc Barthélémy and Luis A. Nunes Amaral. "Small-World Networks: Evidence for a Crossover Picture". In: *Phys. Rev. Lett.* 82 (15 Apr. 1999), pp. 3180–3183. DOI: 10.1103/PhysRevLett.82.3180. URL: http://link.aps.org/doi/10.1103/PhysRevLett.82.3180.

[15] Marc Barthélemy et al. "Dynamical patterns of epidemic outbreaks in complex heterogeneous networks". In: *Journal of Theoretical Biology* 235.2 (2005), pp. 275–288. ISSN: 00225193. DOI: 10.1016/j.jtbi.2005.01.011. arXiv: 0410330 [cond-mat].

[16] Marc Barthélemy et al. "Velocity and Hierarchical Spread of Epidemic Outbreaks in Scale-Free Networks". In: *Physical Review Letters* 92.17 (Apr. 2004). DOI: 10.1103/physrevlett.92.178701.

[17] S. Blanes et al. "The Magnus expansion and some of its applications". In: *Physics Reports* 470.5-6 (Jan. 2009), pp. 151–238. DOI: 10.1016/j.physrep.2008.11.001.

[18] Stefano Boccaletti et al. *Complex networks: Structure and dynamics*. 2006. DOI: 10.1016/j.physrep.2005.10.009. arXiv: 1610.01345. URL: www.elsevier.com/locate/physrep.

[19] Marián Boguñá and Romualdo Pastor-Satorras. "Epidemic spreading in correlated complex networks". In: *Physical Review E - Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics* 66.4 (2002), p. 4. ISSN: 1063651X. DOI: 10.1103/PhysRevE.66.047104. arXiv: 0205621 [cond-mat].

URL: https://journals.aps.org/pre/pdf/10.1103/PhysRevE.66.047104.

[20] B. Bollobas. *Random Graphs*. Ed. by W. Fulton et al. Cambridge University Press, 2001.

[21] Bela Bollobas. *Modern Graph Theory*. Springer, 1998.

[22] Andrei Broder et al. "Graph Structure in the Web". In: *Comput. Netw.* 33.1-6 (June 2000), pp. 309–320. ISSN: 1389-1286. DOI: 10.1016/S1389-1286(00)00083-9. URL: http://dx.doi.org/10.1016/S1389-1286(00)00083-9.

[23] Raffaella Burioni, Enrico Ubaldi, and Alessandro Vezzani. "Asymptotic theory of time varying networks with burstiness and heterogeneous activation patterns". In: *J. Stat. Mech. Theory Exp.* 2017.5 (May 2017), p. 054001. ISSN: 17425468. DOI: 10.1088/1742-5468/aa6ce7. URL: http://stacks.iop.org/1742-5468/2017/i=5/a=054001?key=crossref.e5f692cb999d11ff344d124f66866b45.

[24] G. Caldarelli et al. "Scale-Free Networks from Varying Vertex Intrinsic Fitness". In: *Phys. Rev. Lett.* 89 (25 Dec. 2002), p. 258702. DOI: 10.1103/PhysRevLett.89.258702. URL: http://link.aps.org/doi/10.1103/PhysRevLett.89.258702.

[25] Guido Caldarelli. *Scale-Free Networks*. 2007. DOI: 10.1093/acprof:oso/9780199211517.001.0001.

[26] Claudio Castellano and Romualdo Pastor-Satorras. "Competing activation mechanisms in epidemics on networks". In: *Sci. Rep.* 2 (Apr. 2012), p. 371. DOI: 10.1038/srep00371.

[27] Claudio Castellano and Romualdo Pastor-Satorras. "Relating Topological Determinants of Complex Networks to Their Spectral Properties: Structural and Dynamical Effects". In: *Phys. Rev. X* 7 (4 Oct. 2017), p. 041024. DOI: 10.1103/PhysRevX.7.041024. URL: https://link.aps.org/doi/10.1103/PhysRevX.7.041024.

[28] Claudio Castellano and Romualdo Pastor-Satorras. "Thresholds for epidemic spreading in networks". In: *Physical Review Letters* 105.21 (2010), pp. 1–4. ISSN: 00319007. DOI: 10.1103/PhysRevLett.105.218701. arXiv: 1010.1646.

[29] Ciro Cattuto et al. "Dynamics of Person-to-Person Interactions from Distributed RFID Sensor Networks". In: *PLoS ONE* 5.7 (July 2010), e11596. DOI: 10.1371/journal.pone.0011596. URL: http://dx.doi.org/10.1371%2Fjournal.pone.0011596.

[30] F. Chung, L. Lu, and V. Vu. "Spectra of random graphs with given expected degrees". In: *Proceedings of the National Academy of Sciences* 100.11 (May 2003), pp. 6313–6318. DOI: 10.1073/pnas.0937490100.

[31] Aaron Clauset, Cosma Rohilla Shalizi, and M. E. J. Newman. "Power-Law Distributions in Empirical Data". In: *SIAM Review* 51.4 (2009), pp. 661–703. ISSN: 0036-1445. DOI: 10.1137/070710111. arXiv: arXiv:0706.1062v2. URL: https://arxiv.org/pdf/0706.1062.pdf%20http://epubs.

siam.org/doi/abs/10.1137/070710111%7B%5C%%7D5Cnpapers3://publication/doi/10.1137/070710111%20https://epubs.siam.org/doi/pdf/10.1137/070710111%20http://epubs.siam.org/doi/10.1137/070710111%20http://epubs.siam.org/doi/pdf/1.

[32] D.R. Cox. *Renewal Theory*. Methuen science paperbacks. Methuen, 1970. ISBN: 9780416523805. URL: https://books.google.it/books?id=cClHYAAACAAJ.

[33] Easley David and Kleinberg Jon. *Networks, Crowds, and Markets: Reasoning About a Highly Connected World*. New York, NY, USA: Cambridge University Press, 2010. ISBN: 0521195330, 9780521195331.

[34] O. Diekmann, J. Heersterbeek, and J. Metz. "On the definition and the computation of the basic reproduction number ratio R0 in models for infectious diseases in heterogeneous populations". In: *J Math Biol* 28 (1990). DOI: 10.1007/BF00178324. URL: http://dx.doi.org/10.1007/BF00178324.

[35] S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes. "Critical phenomena in complex networks". In: *Rev. Mod. Phys.* 80 (4 Oct. 2008), pp. 1275–1335. DOI: 10.1103/RevModPhys.80.1275. URL: https://link.aps.org/doi/10.1103/RevModPhys.80.1275.

[36] R. I. M. Dunbar. "Neocortex size as a constraint on group size in primates". In: *J. Hum. Evol.* 22.6 (1992), pp. 469–493. DOI: 10.1016/0047-2484(92)90081-J. URL: http://www.sciencedirect.com/science/article/pii/004724849290081J.

[37] Paul Erdos and Alfred Renyi. "On the evolution of random graphs". In: *Publ. Math. Inst. Hungary. Acad. Sci.* 5 (1960), pp. 17–61.

[38] S. Eubank et al. "Modelling disease outbreaks in realistic urban social networks". In: *Nature* 429 (2004), p. 180. DOI: 10.1038/nature02541. URL: http://dx.doi.org/10.1038/nature02541.

[39] Silvio C. Ferreira, Claudio Castellano, and Romualdo Pastor-Satorras. "Epidemic thresholds of the susceptible-infected-susceptible model on networks: A comparison of numerical and theoretical results". In: *Phys. Rev. E* 86 (4 Oct. 2012), p. 041125. DOI: 10.1103/PhysRevE.86.041125. URL: https://link.aps.org/doi/10.1103/PhysRevE.86.041125.

[40] Laetitia Gauvin et al. "Randomized reference models for temporal networks". In: (June 11, 2018). arXiv: http://arxiv.org/abs/1806.04032v1 [physics.soc-ph].

[41] S. Gil and D. H. Zanette. "Optimal disorder for segregation in annealed small worlds". In: *The European Physical Journal B* 47.2 (Sept. 2005), pp. 265–273. DOI: 10.1140/epjb/e2005-00319-8.

[42] K.-I. Goh and A.-L. Barabási. "Burstiness and memory in complex systems". In: *EPL (Europhysics Letters)* 81.4 (2008), p. 48002. URL: http://stacks.iop.org/0295-5075/81/i=4/a=48002.

[43] S. Gómez et al. "Discrete-time Markov chain approach to contact-based disease spreading in complex networks". In: *EPL (Europhysics Letters)* 89.3 (Feb. 2010), p. 38009. DOI: 10.1209/0295-5075/89/38009.

[44] Bruno Gonçalves, Nicola Perra, and Alessandro Vespignani. "Modeling Users' Activity on Twitter Networks: Validation of Dunbar's Number". In: *Plos One* 6.8 (2011), e22656. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0022656. URL: http://dx.plos.org/10.1371/journal.pone.0022656.

[45] Mark S. Granovetter. "The Strength of Weak Ties". In: *Am. J. Sociol.* 78.6 (1973), pp. 1360–1380. ISSN: 00029602, 15375390. URL: http://www.jstor.org/stable/2776392.

[46] M. Henkel et al. *Non-Equilibrium Phase Transitions: Volume 2: Ageing and Dynamical Scaling Far from Equilibrium*. Theoretical and Mathematical Physics. Springer, 2011. ISBN: 9789048128693. URL: https://books.google.fr/books?id=AiofeEteLVcC.

[47] Malte. Henkel and Michel. Pleimling. *Non-Equilibrium Phase Transitions Volume 2: Ageing and Dynamical Scaling Far from Equilibrium*. Springer Netherlands, 2010. ISBN: 9789048128693. DOI: 10.1007/978-90-481-2869-3.

[48] Herbert W. Hethcote. "The Mathematics of Infectious Diseases". In: *SIAM Review* 42.4 (Jan. 2000), pp. 599–653. ISSN: 0036-1445. DOI: 10.1137/S0036144500371907. URL: http://www.siam.org/journals/ojsa.php%20http://leonidzhukov.net/hse/2014/socialnetworks/papers/2000SiamRev.pdf%20http://epubs.siam.org/doi/10.1137/S0036144500371907.

[49] P. Holme and J. Saramäki. "Temporal networks". In: *Phys. Rep.* 519.3 (2012), pp. 97–125. ISSN: 0370-1573. DOI: 10.1016/j.physrep.2012.03.001. URL: http://www.sciencedirect.com/science/article/pii/S0370157312000841.

[50] Petter Holme and Jari Saramäki. "Temporal Networks". In: (2011), pp. 1–28. ISSN: 03701573. DOI: 10.1016/j.physrep.2012.03.001. arXiv: 1108.1780. URL: http://arxiv.org/abs/1108.1780%7B%5C%%7D0Ahttp://dx.doi.org/10.1016/j.physrep.2012.03.001.

[51] Stanley (University of Illinois, Urbana-Champaign) Wasserman, and Katherine (University of South Carolina) Faust. *Structural Analysis in the Social Sciences*. Cambridge University Press, Nov. 25, 1994. 857 pp. ISBN: 0521387078. URL: https://www.ebook.de/de/product/3719721/stanley_university_of_illinois_urbana_champaign_wasserman_katherine_university_of_south_carolina_faust_structural_analysis_in_the_social_sciences.html.

[52] N. G. van Kampen. *Stochastic Processes in Physics and Chemistry*. Elsevier LTD, Oxford, Mar. 11, 2007. 464 pp. ISBN: 0444529659. URL: https://www.ebook.de/de/product/6071989/n_g_van_kampen_stochastic_processes_in_physics_and_chemistry.html.

[53] M. Karsai et al. "Small but slow world: How network topology and burstiness slow down spreading". In: *Phys. Rev. E* 83 (2 Feb. 2011), p. 025102. DOI: 10.1103. URL: https://link.aps.org/doi/10.1103/PhysRevE.83.025102.

[54] Márton Karsai, Kimmo Kaski, and János Kertész. "Correlated dynamics in egocentric communication networks". In: *PLoS One* 7.7 (July 2012). Ed. by Renaud Lambiotte, e40612. ISSN: 19326203. DOI: 10.1371/journal.pone.0040612. URL: http://dx.plos.org/10.1371/journal.pone.0040612.

[55] Márton Karsai, Nicola Perra, and Alessandro Vespignani. "Time varying networks and the weakness of strong ties". In: *Sci. Rep.* 4 (Feb. 2014), p. 4001. URL: http://dx.doi.org/10.1038/srep04001.

[56] Matt J. Keeling. *Modeling Infectious Diseases in Humans and Animals*. Princeton University Press, Oct. 11, 2007. ISBN: 0691116172. URL: https://www.ebook.de/de/product/6599625/matt_j_keeling_modeling_infectious_diseases_in_humans_and_animals.html.

[57] Matt J. Keeling and Ken T.D. D Eames. "Networks and epidemic models". In: *Journal of The Royal Society Interface* 2.4 (Sept. 2005), pp. 295–307. ISSN: 1742-5689. DOI: 10.1098/rsif.2005.0051. URL: http://rsif.royalsocietypublishing.org/cgi/doi/10.1098/rsif.2005.0051%20http://rsif.royalsocietypublishing.org/content/royinterface/2/4/295.full.pdf.

[58] W. O. Kermack and A. G. McKendrick. "A Contribution to the Mathematical Theory of Epidemics". In: *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* 115.772 (1927), pp. 700–721. ISSN: 0950-1207. DOI: 10.1098/rspa.1927.0118. eprint: http://rspa.royalsocietypublishing.org/content/115/772/700.full.pdf. URL: http://rspa.royalsocietypublishing.org/content/115/772/700.

[59] Hyewon Kim, Meesoon Ha, and Hawoong Jeong. "Dynamic topologies of activity-driven temporal networks with memory". In: *Phys. Rev. E* 97 (6 June 2018), p. 062148. DOI: 10.1103/PhysRevE.97.062148. URL: https://link.aps.org/doi/10.1103/PhysRevE.97.062148.

[60] Hyewon Kim, Meesoon Ha, and Hawoong Jeong. "Scaling properties in time-varying networks with memory". In: *Eur. Phys. J. B* 88.12 (Dec. 2015), p. 315. ISSN: 1434-6036. DOI: 10.1140/epjb/e2015-60662-7. URL: https://doi.org/10.1140/epjb/e2015-60662-7.

[61] Gueorgi Kossinets, Jon Kleinberg, and Duncan Watts. "The structure of information pathways in a social communication network". In: *Proceeding of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining - KDD 08*. ACM Press, 2008. DOI: 10.1145/1401890.1401945.

[62] Renaud Lambiotte, Vsevolod Salnikov, and Martin Rosvall. "Effect of memory on the dynamics of random walks on networks". In: *J. Complex Networks* 3.2 (June 2015), pp. 177–188. ISSN: 20511329. DOI: 10.1093/comnet/cnu017. URL: https://academic.oup.com/comnet/article-lookup/doi/10.1093/comnet/cnu017.

[63] Vito Latora and Massimo Marchiori. "Efficient Behavior of Small-World Networks". In: *Physical Review Letters* 87.19 (Oct. 2001). DOI: 10.1103/physrevlett.87.198701.

[64] Suyu Liu et al. "Controlling Contagion Processes in Activity Driven Networks". In: *Phys. Rev. Lett.* 112 (11 Mar. 2014), p. 118702. DOI: 10.1103/PhysRevLett.112.118702. URL: http://link.aps.org/doi/10.1103/PhysRevLett.112.118702.

[65] Roberto Livi and Paolo Politi. *Nonequilibrium Statistical Physics*. Cambridge University Press, Oct. 2017. DOI: 10.1017/9781107278974.

[66] JOSEPH J. LUCZKOVICH et al. "Defining and Measuring Trophic Role Similarity in Food Webs Using Regular Equivalence". In: *Journal of Theoretical Biology* 220.3 (Feb. 2003), pp. 303–321. DOI: 10.1006/jtbi.2003.3147.

[67] Naoki Masuda and Renaud Lambiotte. *A Guide to Temporal Networks*. WORLD SCIENTIFIC (EUROPE), Apr. 2016. DOI: 10.1142/q0033.

[68] Naoki Masuda, Mason A Porter, and Renaud Lambiotte. *Random walks and diffusion on networks*. 2017. DOI: 10.1016/j.physrep.2017.07.007. arXiv: 1612.03281. URL: www.elsevier.com/locate/physrep.

[69] P. Van Mieghem and R. van de Bovenkamp. "Non-Markovian Infection Spread Dramatically Alters the Susceptible-Infected-Susceptible Epidemic Threshold in Networks". In: *Physical Review Letters* 110.10 (Mar. 2013). DOI: 10.1103/physrevlett.110.108701.

[70] P. Van Mieghem, J. Omic, and R. Kooij. "Virus Spread in Networks". In: *IEEE/ACM Transactions on Networking* 17.1 (Feb. 2009), pp. 1–14. DOI: 10.1109/tnet.2008.925623.

[71] Joel C. Miller. "Epidemic size and probability in populations with heterogeneous infectivity and susceptibility". In: *Physical Review E* 76.1 (July 2007). DOI: 10.1103/physreve.76.010101.

[72] Giovanna Miritello et al. "Limited communication capacity unveils strategies for human interaction". In: *Sci. Rep.* 3 (2013), p. 1950. URL: http://dx.doi.org/10.1038/srep01950.

[73] Antoine Moinet, Michele Starnini, and Romualdo Pastor-Satorras. "Burstiness and Aging in Social Temporal Networks". In: *Phys. Rev. Lett.* 114 (10 Mar. 2015), p. 108701. DOI: 10.1103/PhysRevLett.114.108701. URL: http://link.aps.org/doi/10.1103/PhysRevLett.114.108701.

[74] Antoine Moinet, Michele Starnini, and Romualdo Pastor-Satorras. "Burstiness and Aging in Social Temporal Networks". In: *Physical Review Letters* 114.10 (Mar. 2015), p. 108701. ISSN: 10797114. DOI: 10.1103/PhysRevLett.114.108701. arXiv: 1412.0587. URL: https://journals.aps.org/prl/pdf/10.1103/PhysRevLett.114.108701%20https://link.aps.org/doi/10.1103/PhysRevLett.114.108701.

[75] MICHAEL MOLLOY and BRUCE REED. "The Size of the Giant Component of a Random Graph with a Given Degree Sequence". In: *Combinatorics, Probability and Computing* 7 (03 Sept. 1998), pp. 295–305. ISSN: 1469-2163. DOI: null. URL: http://journals.cambridge.org/article_S0963548398003526.

[76] Yamir Moreno, Romualdo Pastor-Satorras, and Alessandro Vespignani. "Epidemic outbreaks in complex heterogeneous networks". In: *European Physical Journal B* 26.4 (2002), pp. 521–529. ISSN: 14346028. DOI: 10.1007/s10051-002-8996-y. arXiv: 0107267 [cond-mat]. URL: https://arxiv.org/pdf/cond-mat/0107267.pdf.

[77] Matthieu Nadini et al. "Epidemic spreading in modular time-varying networks". In: *Scientific Reports* 8.1 (Feb. 2018). DOI: 10.1038/s41598-018-20908-x.

[78] M. E. J. Newman. "Spread of epidemic disease on networks". In: *Phys. Rev. E* 66 (1 July 2002), p. 016128. DOI: 10.1103/PhysRevE.66.016128. URL: http://link.aps.org/doi/10.1103/PhysRevE.66.016128.

[79] M. E. J. Newman. "The Structure and Function of Complex Networks". In: *SIAM Review* 45.2 (2003), pp. 167–256. DOI: 10.1137/S003614450342480. eprint: http://dx.doi.org/10.1137/S003614450342480. URL: http://dx.doi.org/10.1137/S003614450342480.

[80] M. E. J. Newman. "The Structure and Function of Complex Networks *". In: *SIAM Review* 45.2 (Jan. 2003), pp. 167–256. ISSN: 0036-1445. DOI: 10.1137/S003614450342480. URL: http://epubs.siam.org/doi/pdf/10.1137/S003614450342480%20http://epubs.siam.org/doi/10.1137/S003614450342480.

[81] M. E. J. Newman, I. Jensen, and R. M. Ziff. "Percolation and epidemics in a two-dimensional small world". In: *Phys. Rev. E* 65 (2 Jan. 2002), p. 021904. DOI: 10.1103/PhysRevE.65.021904. URL: http://link.aps.org/doi/10.1103/PhysRevE.65.021904.

[82] M. E. J. Newman, S. H. Strogatz, and D. J. Watts. "Random graphs with arbitrary degree distributions and their applications". In: *Phys. Rev. E* 64 (2 July 2001), p. 026118. DOI: 10.1103/PhysRevE.64.026118. URL: http://link.aps.org/doi/10.1103/PhysRevE.64.026118.

[83] Mark Newman. "Networks: An Introduction". In: *Networks: An Introduction* (2010), pp. 1–784. ISSN: 1578-1275. DOI: 10.1093/acprof:oso/9780199206650.001.0001. arXiv: 1212.2425.

[84] Vincenzo Nicosia et al. "Components in time-varying graphs". In: *journal of nonlinear ...* (2012). arXiv: arXiv:1106.2134v3. URL: https://arxiv.org/pdf/1106.2134.pdf%20http://aip.scitation.org/doi/abs/10.1063/1.3697996.

[85] Vincenzo Nicosia et al. "Temporal Networks". In: (2013). ISSN: 1860-0832. DOI: 10.1007/978-3-642-36461-7. URL: http://link.springer.com/10.1007/978-3-642-36461-7.

[86] Jae Dong Noh and Heiko Rieger. "Random Walks on Complex Networks". In: *Physical Review Letters* 92.11 (Mar. 2004). DOI: 10.1103/physrevlett.92.118701.

[87] Tore Opsahl et al. "Prominence and Control: The Weighted Rich-Club Effect". In: *Physical Review Letters* 101.16 (Oct. 2008). DOI: 10.1103/physrevlett.101.168702.

[88] Raj Kumar Pan and Jari Saramäki. "Path lengths, correlations, and centrality in temporal networks". In: *Phys. Rev. E* 84 (1 July 2011), p. 016105. DOI: 10.1103/PhysRevE.84.016105. URL: http://link.aps.org/doi/10.1103/PhysRevE.84.016105.

[89] R. Pastor-Satorras and A. Vespignani. "Epidemic spreading in scale-free networks". In: *Phys. Rev. Lett.* 86 (2001), p. 3200. DOI: 10.1103/PhysRevLett.86.3200. URL: http://dx.doi.org/10.1103/PhysRevLett.86.3200.

[90] Romualdo Pastor-Satorras, Alexei Vázquez, and Alessandro Vespignani. "Dynamical and Correlation Properties of the Internet". In: *Phys. Rev. Lett.* 87 (25 Nov. 2001), p. 258701. DOI: 10.1103/PhysRevLett.87.258701. URL: http://link.aps.org/doi/10.1103/PhysRevLett.87.258701.

[91] Romualdo Pastor-Satorras and Alessandro Vespignani. "Epidemic spreading in scale-free networks". In: *Physical Review Letters* 86.14 (Apr. 2001), pp. 3200–3203. ISSN: 00319007. DOI: 10.1103/PhysRevLett.86.3200. arXiv: 0010317 [cond-mat]. URL: https://journals.aps.org/prl/pdf/10.1103/PhysRevLett.86.3200%20https://link.aps.org/doi/10.1103/PhysRevLett.86.3200.

[92] Romualdo Pastor-Satorras et al. "Epidemic processes in complex networks". In: *Reviews of Modern Physics* 87.3 (2015), pp. 925–979. ISSN: 15390756. DOI: 10.1103/RevModPhys.87.925. arXiv: 1408.2701.

[93] Romualdo Pastor-Satorras et al. "Epidemic processes in complex networks". In: *Rev. Mod. Phys.* 87 (3 Aug. 2015), pp. 925–979. DOI: 10.1103/RevModPhys.87.925. URL: https://link.aps.org/doi/10.1103/RevModPhys.87.925.

[94] N. Perra et al. "Activity driven modeling of time varying networks". In: *Sci. Rep.* 2.1 (Dec. 2012), p. 469. ISSN: 2045-2322. DOI: 10.1038/srep00469. URL: http://www.nature.com/articles/srep00469.

[95] N. Perra et al. "Random Walks and Search in Time-Varying Networks". In: *Phys. Rev. Lett.* 109.23, 238701 (Dec. 2012), p. 238701. DOI: 10.1103/PhysRevLett.109.238701.

[96] Nicola Perra et al. "Random walks and search in time-varying networks". In: *Physical Review Letters* 109.23 (2012), pp. 1–5. ISSN: 00319007. DOI: 10.1103/PhysRevLett.109.238701. arXiv: 1206.2858. URL: https://journals.aps.org/prl/pdf/10.1103/PhysRevLett.109.238701.

[97] Iacopo Pozzana, Kaiyuan Sun, and Nicola Perra. "Epidemic Spreading on Activity-Driven Networks with Attractiveness". In: (2017). arXiv: 1703.02482. URL: http://arxiv.org/abs/1703.02482.

[98] Iacopo Pozzana, Kaiyuan Sun, and Nicola Perra. "Epidemic Spreading on Activity-Driven Networks with Attractiveness". In: *Phys. Rev. E* 96.4 (2017), p. 042310. DOI: 10.1103/PhysRevE.96.042310.

[99] Alessandro Rizzo and Maurizio Porfiri. "Innovation diffusion on time-varying activity driven networks". In: *Eur. Phys. J. B* 89.1 (Jan. 2016), p. 20. ISSN: 14346036. DOI: 10.1140/epjb/e2015-60933-3. URL: http://link.springer.com/10.1140/epjb/e2015-60933-3.

[100] Martin Rosvall et al. "Memory in network flows and its effects on spreading dynamics and community detection". In: *Nat. Commun.* 5.1 (Dec. 2014), p. 4630. ISSN: 20411723. DOI: 10.1038/ncomms5630. URL: http://www.nature.com/articles/ncomms5630.

[101] Jari Saramäki and Esteban Moro. "From seconds to months: an overview of multi-scale dynamics of mobile telephone calls". In: *Eur. Phys. Jour. B* 88.6 (2015), p. 164. URL: http://dx.doi.org/10.1140/epjb/e2015-60106-6.

[102] Jari Saramäki et al. "Persistence of social signatures in human communication". In: *PNAS* 111.3 (2014), pp. 942–947. ISSN: 0027-8424. DOI: 10.1073/pnas.1308540110.

[103] Ingo Scholtes et al. "Causality-driven slow-down and speed-up of diffusion in non-Markovian temporal networks". In: *Nat. Commun.* 5.1 (2014), p. 5024. ISSN: 2041-1723. DOI: 10.1038/ncomms6024. URL: http://www.nature.com/doifinder/10.1038/ncomms6024%20http://www.nature.com/articles/ncomms6024.

[104] M. Ángeles Serrano, Marián Boguñá, and Romualdo Pastor-Satorras. "Correlations in weighted networks". In: *Phys. Rev. E* 74 (5 Nov. 2006), p. 055101. DOI: 10.1103/PhysRevE.74.055101. URL: http://link.aps.org/doi/10.1103/PhysRevE.74.055101.

[105] Kieran J. Sharkey. "Deterministic epidemic models on contact networks: Correlations and unbiological terms". In: *Theoretical Population Biology* 79.4 (June 2011), pp. 115–129. DOI: 10.1016/j.tpb.2011.01.004.

[106] Shigeru Shinomoto, Keiji Miura, and Shinsuke Koyama. "A measure of local variation of inter-spike intervals". In: *Biosystems* 79.1-3 (Jan. 2005), pp. 67–72. DOI: 10.1016/j.biosystems.2004.09.023.

[107] American Physical Society. *APS Data Sets for Research*. URL: https://journals.aps.org/datasets.

[108] Bo Söderberg. "General formalism for inhomogeneous random graphs". In: *Phys. Rev. E* 66 (6 Dec. 2002), p. 066121. DOI: 10.1103/PhysRevE.66.066121. URL: http://link.aps.org/doi/10.1103/PhysRevE.66.066121.

[109] D. J. de Solla Price. "Networks of Scientific Papers". In: *Science* 149.3683 (July 1965), pp. 510–515. DOI: 10.1126/science.149.3683.510.

[110] Leo Speidel et al. "Temporal interactions facilitate endemicity in the susceptible-infected-susceptible epidemic model". In: *New Journal of Physics* 18.7 (July 2016), pp. 1–24. ISSN: 13672630. DOI: 10.1088/1367-2630/18/7/073013. arXiv: 1602.00859. URL: http://stacks.iop.org/1367-2630/18/i=7/a=073013?key=crossref.950461fc36231b959b6c696b8c789b92.

[111] Michele Starnini and Romualdo Pastor-Satorras. "Temporal percolation in activity-driven networks". In: *Phys. Rev. E* 89 (3 Mar. 2014), p. 032807. DOI: 10.1103/PhysRevE.89.032807. URL: http://link.aps.org/doi/10.1103/PhysRevE.89.032807.

[112] Michele Starnini and Romualdo Pastor-Satorras. "Topological properties of a time-integrated activity-driven network". In: *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics* 87.6 (June 2013), p. 062807. ISSN: 1539-3755. DOI: 10.1103/PhysRevE.87.062807. arXiv: arXiv:1304.3370v1. URL: https://link.aps.org/doi/10.1103/PhysRevE.87.062807%20https://journals.aps.org/pre/pdf/10.1103/PhysRevE.87.062807.

[113] Michele Starnini and Romualdo Pastor-Satorras. "Topological properties of a time-integrated activity-driven network". In: *Phys. Rev. E* 87 (6 June 2013), p. 062807. DOI: 10.1103/PhysRevE.87.062807. URL: http://link.aps.org/doi/10.1103/PhysRevE.87.062807.

[114] Dietrich Stauffer and Muhammad Sahimi. "Diffusion in scale-free networks with annealed disorder". In: *Physical Review E* 72.4 (Oct. 2005). DOI: 10.1103/physreve.72.046128.

[115] Juliette Stehlé, Alain Barrat, and Ginestra Bianconi. "Dynamical and bursty interactions in social networks". In: *Phys. Rev. E* 81 (3 Mar. 2010), p. 035101. DOI: 10.1103/PhysRevE.81.035101. URL: http://link.aps.org/doi/10.1103/PhysRevE.81.035101.

[116] James Stiller and R. I. M. Dunbar. "Perspective-taking and memory capacity predict social network size". In: *Soc. Netw.* 29.1 (Jan. 2007), pp. 93–104. DOI: 10.1016/j.socnet.2006.04.001. URL: http://www.sciencedirect.com/science/article/pii/S0378873306000128.

[117] Kaiyuan Sun, Andrea Baronchelli, and Nicola Perra. "Contrasting effects of strong ties on SIR and SIS processes in temporal networks". In: *Eur. Phys. J. B* 88.12 (2015), p. 326. ISSN: 14346036. DOI: 10.1140/epjb/e2015-60568-4.

[118] Michele Tizzani et al. "Epidemic Spreading and Aging in Temporal Networks with Memory". In: (July 31, 2018). arXiv: http://arxiv.org/abs/1807.11759v2 [physics.soc-ph].

[119] Jeffrey Travers and Stanley Milgram. "An Experimental Study of the Small World Problem". In: *Sociometry* 32.4 (Dec. 1969), p. 425. DOI: 10.2307/2786545.

[120] Enrico Ubaldi et al. "Asymptotic theory for the dynamic of networks with heterogenous social capital allocation". In: *Scientific Reports* 6 (2015), p. 35724. DOI: 10.1038/srep35724. arXiv: 1509.04563. URL: http://www.nature.

com/articles/srep35724%7B%5C%%7D5Cnhttp://arxiv.org/abs/1509.04563.

[121] Enrico Ubaldi et al. "Asymptotic theory of time-varying social networks with heterogeneous activity and tie allocation". In: *Sci. Rep.* 6 (Oct. 2016), p. 35724. URL: http://dx.doi.org/10.1038/srep35724.

[122] Enrico Ubaldi et al. "Burstiness and tie activation strategies in time-varying social networks". In: *Scientific Reports* 7 (2017). ISSN: 20452322. DOI: 10.1038/srep46225. arXiv: 1607.08910. URL: https://www.nature.com/articles/srep46225.pdf.

[123] Eugenio Valdano et al. "Analytical computation of the epidemic threshold on temporal networks". In: *Physical Review X* 5.2 (2015), pp. 1–9. ISSN: 21603308. DOI: 10.1103/PhysRevX.5.021005. arXiv: 1406.4815.

[124] Eugenio Valdano et al. "Epidemic Threshold in Continuous-Time Evolving Networks". In: *Phys. Rev. Lett.* 120 (6 Feb. 2018), p. 068302. DOI: 10.1103/PhysRevLett.120.068302. URL: https://link.aps.org/doi/10.1103/PhysRevLett.120.068302.

[125] Piet Van Mieghem and Eric Cator. "Epidemics in networks with nodal self-infection and the epidemic threshold". In: *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics* 86.1 (2012). ISSN: 15393755. DOI: 10.1103/PhysRevE.86.016116. URL: https://journals.aps.org/pre/pdf/10.1103/PhysRevE.86.016116.

[126] Piet Van Mieghemy, Faryad Darabi Sahnehz, and Caterina Scoglioz. "An upper bound for the epidemic threshold in exact Markovian SIR and SIS epidemics on networks". In: *Proceedings of the IEEE Conference on Decision and Control*. Vol. 2015-Febru. February. 2014, pp. 6228–6233. ISBN: 9781467360883. DOI: 10.1109/CDC.2014.7040365. arXiv: 1402.1731. URL: https://arxiv.org/pdf/1402.1731.pdf.

[127] Alexei Vázquez, Romualdo Pastor-Satorras, and Alessandro Vespignani. "Large-scale topological and dynamical properties of the Internet". In: *Phys. Rev. E* 65 (6 June 2002), p. 066130. DOI: 10.1103/PhysRevE.65.066130. URL: http://link.aps.org/doi/10.1103/PhysRevE.65.066130.

[128] Alexei Vázquez et al. "Modeling bursts and heavy tails in human dynamics". In: *Phys. Rev. E* 73 (3 Mar. 2006), p. 036127. DOI: 10.1103/PhysRevE.73.036127. URL: http://link.aps.org/doi/10.1103/PhysRevE.73.036127.

[129] A Vespignani, M Bathelemy, and A Barrat. *Dynamical Processes on Complex Networks*. 2008, p. 367. ISBN: 9780521879507. DOI: 10.1017/CBO9780511791383. arXiv: arXiv:1011.1669v3.

[130] Thomas Andrew Waigh. *The Physics of Living Processes: A Mesoscopic Approach*. PAPERBACKSHOP UK IMPORT, Oct. 20, 2014. 599 pp. ISBN: 1118449940. URL: https://www.ebook.de/de/product/22523623/thomas_andrew_waigh_the_physics_of_living_processes_a_mesoscopic_approach.html.

[131] Yang Wang et al. "Epidemic spreading in real networks: an eigenvalue viewpoint". In: *22nd International Symposium on Reliable Distributed Systems, 2003. Proceedings.* IEEE Comput. Soc. DOI: 10.1109/reldis.2003.1238052.

[132] Duncan J. Watts and Steven H. Strogatz. "Collective dynamics of /'small-world/' networks". In: *Nature* 393.6684 (June 1998), pp. 440–442. URL: http://dx.doi.org/10.1038/30918.

[133] J. M. Yeomans. *Statistical Mechanics of Phase Transitions.* OXFORD UNIV PR, May 11, 1992. 168 pp. ISBN: 0198517300. URL: https://www.ebook.de/de/product/3596629/j_m_yeomans_statistical_mechanics_of_phase_transitions.html.