UNIVERSITY OF SÃO PAULO
SCHOOL OF ARTS, SCIENCES AND HUMANITIES
GRADUATE PROGRAM IN MODELING OF COMPLEX SYSTEMS

CRISTIANE DIAS DE SOUZA MARTORELLO

Epidemiology in Complex Networks – Modified Heterogeneous Mean-Field Model

São Paulo
2019

CRISTIANE DIAS DE SOUZA MARTORELLO

Epidemiology in Complex Networks – Modified Heterogeneous Mean-Field Model

Dissertation presented to the School of Arts, Sciences and Humanities of the University of São Paulo to obtain the title of Master of Science by the Graduate Program in Modeling of Complex Systems.

Corrected version containing the changes requested by the judging committee at November, 29th, 2018. The original version is available at the Library of EACH-USP and in the Digital Library of Theses and Dissertations of USP (BDTD), in accordance with CoPGr Resolution 6018 of October 13th, 2011.

Concentration area: Complex systems

Supervisor: Masayuki Oka Hase

São Paulo

2019

CATALOGUING IN PUBLICATION
(University of São Paulo. School of Arts, Sciences and Humanities. Library)
CRB 8 -4936

Martorello, Cristiane Dias de Souza
Epidemiology in complex networks – modified heterogeneous mean-field model / Cristiane Dias de Souza Martorello ; supervisor, Masayuki Oka Hase. – 2019.
70 p. : il.

Dissertation (Master of Science) – Graduate Program in Complex Systems Modeling, School of Arts, Sciences and Humanities, University of São Paulo, 2018.
Corrected version.

1. Epidemiology - Computer simulation. 2. Complex networks. 3. Dynamic systems. I. Hase, Masayuki Oka, advisor II. Title

CDD 22.ed.– 614.4011

Name: MARTORELLO, Cristiane Dias de Souza
Title: Epidemiology in Complex Networks – Modified Heterogeneous Mean-Field Model

Dissertation presented to the School of Arts, Sciences and Humanities of the University of São Paulo to obtain the title of Master of Science by the Graduate Program in Modeling of Complex Systems.

Corrected version containing the changes requested by the judging committee at November, 29th, 2018. The original version is available at the Library of EACH-USP and in the Digital Library of Theses and Dissertations of USP (BDTD), in accordance with CoPGr Resolution 6018 of October 13th, 2011.

Concentration area: Complex systems

Approved in: 29/11/2018

Examination board

Prof. Dr. Camilo Rodrigues Neto
Universidade de São Paulo. Escola de Artes, Ciências e Humanidades.

Prof. Dr. Fabio Stucchi Vannucchi
Universidade Estadual Paulista Júlio de Mesquita Filho, Campus do Litoral Paulista. Unidade São Vicente.

Prof. Dr. Marcone Corrêa Pereira
Universidade de São Paulo. Instituto de Matemática e Estatística.

ABSTRACT

MARTORELLO, Cristiane Dias de Souza. **Epidemiology in Complex Networks – Modified Heterogeneous Mean-Field Model**. 2019. 70 p. Dissertation (Master of Science) – School of Arts, Sciences and Humanities, University of São Paulo, São Paulo, 2018. Corrected version.

The study of complex networks presented a huge development in last decades. In this dissertation we want to analyze the epidemic spread in scale-free networks through the Susceptible - Infected - Susceptible (SIS) model. We review the fundamental concepts to describe complex networks and the classical epidemiological models. We implement an algorithm that produces a scale-free network and explore the Quenched Mean-Field (QMF) dynamics in a scale-free network. Moreover, we simulate a change on the topology of the network according to the states of the nodes, and it generates a positive epidemic threshold. We show analytically that the fraction of infected vertices follows a power-law distribution in the vicinity of this critical point.

Keywords: Networks. Scale-free network. SIS model.

RESUMO

MARTORELLO, Cristiane Dias de Souza. **Epidemiologia em Redes Complexas – Modelo de Campo Médio Heterogêneo Modificado**. 2019. 70 p. Dissertação (Mestrado em Ciências) – Escola de Artes, Ciências e Humanidades, Universidade de São Paulo, São Paulo, 2018. Versão corrigida.

O estudo de redes complexas tem se desenvolvido muito nos últimos anos. Nesta dissertação queremos analisar o processo de propagação de epidemia em redes livres de escala através do modelo Suscetível - Infectado - Suscetível (SIS). Apresentamos uma revisão de redes e as principais características dos modelos epidemiológicos clássicos. Implementamos um algoritmo que produz uma rede livre de escala dado um expoente e exploramos a dinâmica do modelo Quenched Mean-Field (QMF) inserido em uma rede livre de escala. Além disso, foi simulada uma possível alteração na topologia da rede, devido aos estados dos vértices infectados, que gerou um limiar epidêmico positivo no modelo e a probabilidade de vértices infectados seguiu uma lei de potência na vizinhança desse ponto crítico.

Palavras-chave: Redes. Rede livre de escala. Modelo SIS.

List of Figures

Contents

1 Introduction

Real systems composed of many elements connected in some way can be usually modelled by a network, where the elements are the vertices and the interactions among the elements can be seen as the edges. Some examples are the Internet, where computers are linked, the social network that connects people and the brain that is composed by billions of connected neurons (1).

One of the most interesting discovery of network is that the architecture of networks emerging in various domains of science, nature, and technology are similar to each other (1). As a consequence, we can use a common set of mathematical tools to explore these networks.

Many of these networks present new characteristics through interactions of their components and exhibit emergent phenomena which can not be explained by local properties (2). Due to these emergent characteristics, it is therefore called a complex network. Complex networks have a multidisciplinary nature, being a significant area of study for the last decades.

How a virus propagate in a complex network is one of the questions this dissertation wants to explore, in other words, we want to understand the way an infection evolves over time. Besides, viral propagation is very similar to the spread of rumors and information on a social network and the spreading of computer viruses over the Internet (3), implying that the solutions for epidemics would give us insights to many others areas.

We model an epidemic in several types of networks and analyze the effects of the complex network features on the epidemic spreading. We reviewed some basic epidemiological models, such as the Susceptible - Infected (SI) model and the Susceptible - Infected - Recovered (SIR) model, besides the Susceptible - Infected - Susceptible (SIS) model (4). We considered the SIS model in different network structures to simulate an epidemic spreading and analyze its behavior in the stationary state, mainly if the diseases vanish in the network or if it is persistent for long times.

One interesting phenomenon studied in this dissertation is how the dynamic of the network is affected when the states of the nodes are changed. One example of this effect is when an infected computer with a virus has limited access to the Internet, restricting its connection to others computers and preventing of infected them. This change have an

important impact in some properties of the network, differently from the expected from a traditional heterogeneous mean-field (HMF) network. Modifying a factor in the model made possible to achieve a positive epidemic threshold, where the fraction of the infected elements vanishes in the scale-free network for a range of values of the infection rate, other than in the traditional HMF, where the epidemic threshold is null and, for a infection rate positive, the fraction of infected individuals is always positive.

Our research is supported by many works, and we reviewed the fundamental concepts related to complex networks. A brief review to describe complex networks is presented in section 2. A review of the epidemiological models is provided in section 3, including a discussion about the epidemic threshold. Heterogeneous mean-field models of SIS is provided in section 4, and we separated in two particular cases: the simplified mean-field and the homogeneous mean-field. The Quenched Mean-Field is presented in section 5. Scale-free network modeling is in section 6 and computational simulations are presented in section 7. The modified model where we studied the effects on dynamic's networks when changing the states of the nodes is in section 8 and, lastly, the conclusion is in section 9.

2 Complex Networks

Complex systems are not trivial systems due to their unpredictable collective behaviour and complex network is a graph with non-trivial topological features — properties that do not happen in regular lattice. Complex networks is used to model real networks, which usually present heavy tail in degree distribution, and one of the most important characteristic of the complex network is that it can be modeled as a scale-free network, where the degree distribution of the network follows a power-law. Examples of real networks are communication network and a social network (1).

After the classic paper of Erdős and Rényi (5), introducing the random network in 1959, the interest in networks strongly increased, which allowed the development of network theory and general properties.

We will start reviewing some properties of graph theory, which will lead to a better understanding about networks.

2.1 Graph theory

To describe networks in simple terms, the mathematical graph theory provides the tools for terminology and notations. In this section we present some concepts and notation that will be used throughout the text.

A graph is a set of vertices connected via edges and it can be directed or undirected. A directed graph has a direction between the elements; for instance, if a vertex i is connected to a vertex j, it does not mean that the vertex j has a connection to the vertex i. Still, in an undirected graph, if a vertex i is connected with a vertex j, j is connected to vertex i, and it does not have direction in the edge.

The size of the network is denoted by N, which is the number of vertices (or nodes) and can be seen as the individuals in the network. The nodes are indexed by integer numbers from 1 to N, and an edge is defined as a link between a pair of nodes.

Two nodes are said to be adjacent if there is an edge between them. The neighborhood of a vertex i is the set of all vertices that are adjacent to node i.

The number of connections of vertex i is the degree of i, denoted by k_i. It is the size of the neighborhood of i. A hub is a node with a very large degree, in other terms, a highly connected vertex.

The mean degree $\langle k \rangle$ of a network is obtained by

$$\langle k \rangle = \frac{1}{N} \sum_{i=1}^{N} k_i. \tag{2.1}$$

One important property of networks is the degree distribution P. Degree distribution refers to the probability $P(k)$ that a selected node has degree k.

The adjacency matrix $A = (a_{ij})$ is defined as a matrix that provides the complete representation of a network. An $N \times N$ matrix can be attributed to a graph of N vertices which contains the information of all connections in the network, since the connections can be represented by the elements a_{ij} of the matrix, where:

$$a_{ij} = \begin{cases} 1, \text{ if vertices } i \text{ and } j \text{ are connected} \\ 0, \text{ otherwise.} \end{cases}$$

Therefore, the degree k_i is given by $k_i = \sum_{j=1}^{N} a_{ij}$, and if L is the number of edges of the network, then $L = \frac{1}{2} \sum_{i=1}^{N} k_i$. The adjacency matrix of undirected networks is symmetric, so $a_{ij} = a_{ji}$.

A path is any sequence of linked nodes. The length of a path is the number of edges along the path, and the shortest path between nodes i and j is the least number of intermediate links in the path between nodes i and j. The shortest path length is also called intervertex distance l_{ij} and the distribution of intervertex distances $P(l)$ describes the global structure of a random network. The mean intervertex distance $\bar{l}(N)$ characterizes the compactness of a network (6). The diameter of a network is the length of the longest shortest path between two vertices.

The small-world effect is a phenomenon in networks that the mean distance is small in relation to N, typically $\bar{l} \sim \ln N$. Milgram's experiment suggests that most pairs of people in a population can be connected by six degree on average (7) . This effect has implications for the system, for example, when a disease occurs in a population it reaches susceptible individuals much faster if they are connected by only a short chain of intermediate acquaintances (8).

The moments are quantities that provide statistical properties of the network, i.e. the first moment ($m = 1$) is the mean value and the second moment ($m = 2$) is the mean square. The m-th moment of the degree distribution is obtained by

$$\langle k^m \rangle = \sum_{k=0}^{\infty} k^m P(k). \tag{2.2}$$

The clustering coefficient measures the density of triangles in the network, which is the average probability that two neighbors of a vertex are neighbors themselves.

A loop is a closed path visiting each vertex once, and trees are graphs without loops. A network can have multiple connections between vertices, besides a vertex can be connected with itself - it is a self-loop.

The networks we studied in this work are simple graphs, because they have a single edge between a pair of vertices. Besides, it is an undirected network and do not have self-loops.

2.2 Random Networks

The random network model predominated since 1959 by the paper of Erdős and Rényi (5). This model is important to explore the properties of real networks. Their work introduced two random graph models, the $G(N, p)$ and the $G(N, M)$. The $G(N, M)$ model is a graph from a set of all graphs with N nodes and M edges. A random network $G(N, p)$ is a network with N elements where each pair can be connected with probability p. This last model will be used throughout the text.

The probability of a vertex being connected to k vertices is

$$P(k) = \binom{N-1}{k} p^k (1-p)^{N-1-k}, \tag{2.3}$$

since each vertex can see $N - 1$ others vertices; the vertex connects to k other vertices with probability p, and it is not connected to $N - 1 - k$ with probability $(1 - p)$. The factor $\binom{N-1}{k}$ is the number of all possible configuration under these conditions; then the degree distribution of a random network is a binomial distribution.

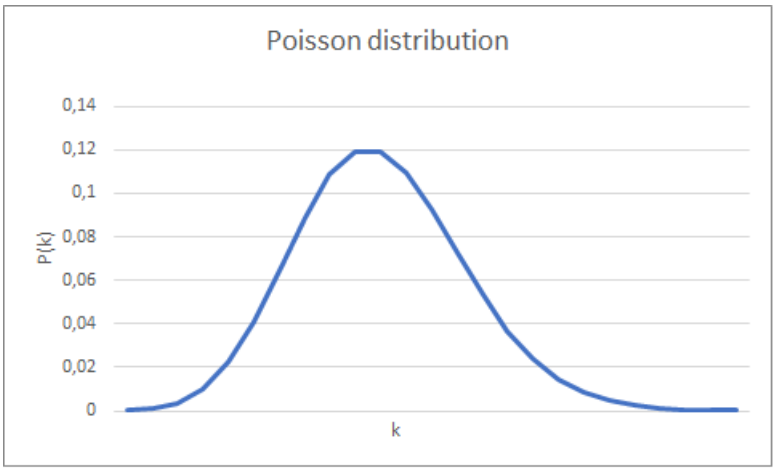
For large N and small p, such that the mean degree $\langle k \rangle = p(N - 1)$ is $O(1)$, the binomial distribution can be approximated by a Poisson distribution:

$$P(k) = \frac{e^{-\langle k \rangle} \langle k \rangle^k}{k!}. \tag{2.4}$$

It can be seen that in (2.3), the degree distribution has two parameters (N and p), while the Poisson distribution only needs $\langle k \rangle$ to be characterized.

The Erdős-Rényi network is not complex since its Poisson distribution means that the majority of the nodes has a similar number of edges. Figure 1 represents a Poisson degree distribution.

Figure 1 – Poisson distribution with $\langle k \rangle$=11.



Source: Martorello, Cristiane Dias de Souza, 2018

2.3 Scale-Free Networks

Complex networks are networks with a more complex architecture than classical random graphs with Poisson distribution (6). Real networks usually does not follow a Poisson distribution, instead, the data points form an approximate straight line on a log-log scale, indicating that the degree distribution follows a power-law (1).

A scale-free network is a network with power-law degree distributions and it has the form

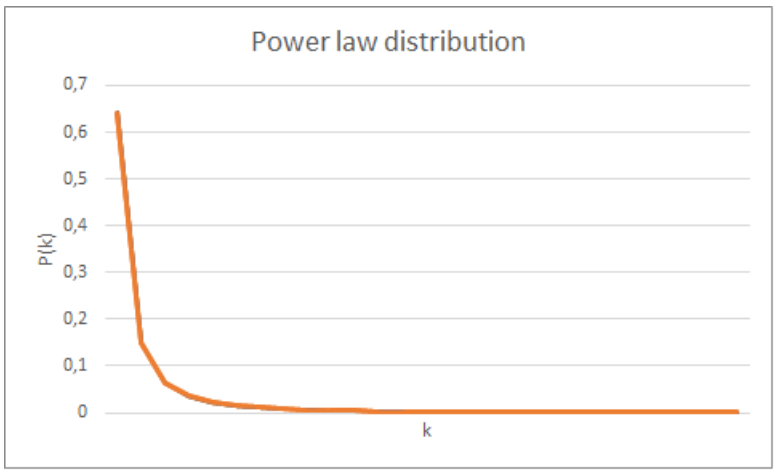
$$P(k) \sim k^{-\alpha}, \tag{2.5}$$

where α is the degree exponent. If we take a logarithm of it, we obtain

$$\log P(k) \sim -\alpha \log k,$$

where $\log P(k)$ depends linearly on $\log k$, and the slope is the degree exponent α. Figure 2 represents a power-law degree distribution. Real networks usually have exponent that ranges between 2 and 3 (3).

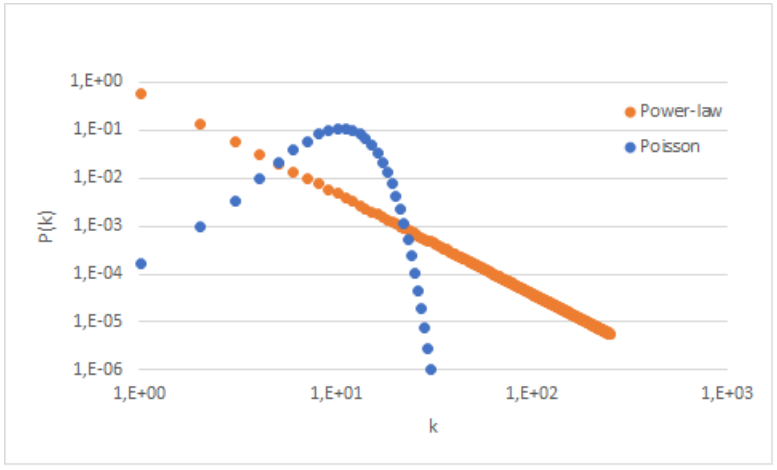
Figure 2 – Power-law degree distribution with $\alpha = 2.1$



Source: Martorello, Cristiane Dias de Souza, 2018

The emergence of many hubs is a property of scale-free degree distribution. These special nodes of the network has a larger probability to develop more connections to other nodes (9). This is the main difference between scale-free and Poisson networks: hubs in scale-free networks present a much larger number of links comparing to Poisson networks. It can be seen in the tail of the degree distribution, using WWW as a example, the probability to have a node with k=100 is about 10^{-94} in a Poisson distribution with $\langle k \rangle = 11$, while it is about 10^{-4} in a power-law with an $\alpha = 2.1$ (1). Figure 3 compares the Poisson distribution of Figure 1 and the Power-law distribution of Figure 2 in a log-log plot.

Figure 3 – A Poisson distribution with $\langle k \rangle$=11 and a power-law degree distribution with $\alpha = 2.1$ in a log-log scale.



Source: Martorello, Cristiane Dias de Souza, 2018

2.4 Barabási-Albert (BA) model

Real networks usually has the property of growth, where new nodes are added through the time. Besides, new nodes might have a characteristic of "preferring" to link to the more connected nodes, and this property is called preferential attachment (1).

The BA model (10) uses preferential attachment and growing networks to get a scale-free distribution of the network. In BA model, we start with v_0 vertices and at each time step a new node joins the network with m links to connect to existing nodes. After t time steps the model leads to a random network with $t + v_0$ nodes and mt edges. While most nodes have only a few links, some nodes gradually turn into hubs. Due to preferential attachment, new nodes are more likely to connect to the more connected nodes than to the smaller nodes, it is called the "rich-gets-richer" phenomenon.

The probability Π that a new link connects to a node i depends on the degree k_i as

$$\Pi(k_i) = \frac{k_i}{\sum k_i};$$

where the denominator is the total degree (normalization term).

Approximating the degree k_i to a continuous real variable, the rate at which a vertex obtain links is

$$\frac{dk_i}{dt} = m\frac{k_i}{\sum k_i}.$$

In the model, we add a node at equal time step (1). Making the sum, we have $\sum k_i = 2mt$, therefore

$$\frac{dk_i}{dt} = \frac{k_i}{2t}.$$

Considering a large t, if we integrate the equation above, we obtain

$$k_i(t) = m\left(\frac{t}{t_i}\right)^{\frac{1}{2}}, \tag{2.6}$$

where t_i is the time at which vertex i was added to the system.

The probability that a vertex i has a connectivity smaller than k, $P(k_i(t) < k)$, can be written as $P(t_i > \frac{tm^2}{k^2})$ from equation (2.6). Assuming that we add new vertices at equal time intervals

$$P(t_i \leq \tau) = \frac{\tau}{t},$$

where $P(t_i \leq \tau)$ is the probability of the vertex i, at time t, joins the network before τ, we have

$$P(k_i < k) = 1 - P\left(t_i \leq \frac{tm^2}{k^2}\right) = 1 - \frac{\frac{tm^2}{k^2}}{t} = 1 - \frac{m^2}{k^2}.$$

Since $P(k_i < k) = \int_0^k dk' P(k')$, then

$$P(k_i < k) = 1 - \frac{m^2}{k^2} = \int_0^k dk' P(k'),$$

which leads to

$$P(k) = \frac{d}{dk} \underbrace{\int_0^k dk' P(k')}_{P(k_i<k)} = \frac{d}{dk}\left(1 - \frac{m^2}{k^2}\right) = \frac{2m^2}{k^3} \tag{2.7}$$

for $t \to \infty$, giving $\alpha = 3$, independent of m (10).

Therefore,

$$\langle k \rangle = \int_m^\infty dk \quad kP(k) = 2m$$

and

$$\langle k^2 \rangle = \int_m^\infty dk \quad k^2 P(k) = \infty.$$

The second moment diverges in BA model.

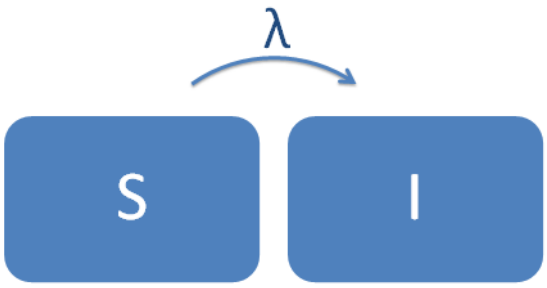
3 Epidemiological Models

The connection between network structure and dynamics of a disease spreading has been extensively studied in recent years. Many diseases spread over networks of contacts between individuals. If some infected individual is introduced into a population, describing the dynamics of infection spreading as a function of time is a problem that the epidemic models help to solve (11).

Epidemic models can assume that the total population is constant and it can be divided into different classes. The simplest model has only two different states: Susceptible and Infected. A susceptible is a healthy individual which can be infected by an illness, and an infected individual already has the disease and can pass it on if he/she has contact with a susceptible individual (4).

A Susceptible-Infected (SI) model allows that susceptible individuals get infected, but the infected individuals stay infected (there is not a change of state). An example of disease that can be modelled by a SI model is AIDS, where once a person had the syndrome, the individual stays infected and can infect other people. A diagram of SI model is in Figure 4, where S is the Susceptible state and I is the Infected state. It is assumed that λ is the infection rate.

Figure 4 – A SI model's diagram.

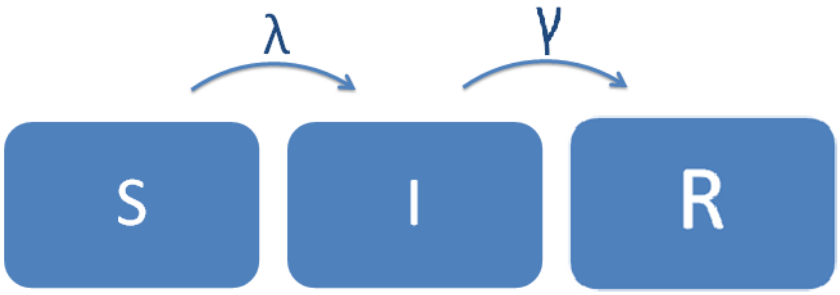


Source: Martorello, Cristiane Dias de Souza, 2018

A usual model with three states is the Susceptible-Infected-Recovered (SIR) model, and it allows infected individuals to be removed from the dynamics by immunity or death. A diagram of SIR model is in Figure 5, the symbol S stands for the Susceptible state, I for the Infected state and R is the Recovered state. It is assumed that λ is the infection rate and γ is the removal rate. An example of disease that can be modelled as a SIR model is chickenpox, where once a person had the virus, he/she stays infected and can pass it to

other people while in infected state, and then he/she gets recovered and becomes immune to the disease.

Figure 5 – A SIR model's diagram.

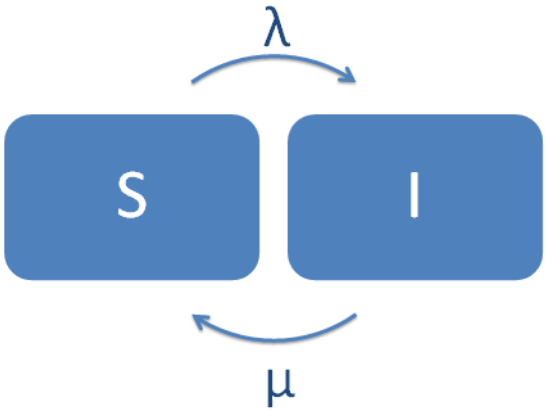


Source: Martorello, Cristiane Dias de Souza, 2018

The Susceptible-Infected-Susceptible (SIS) model has two states, however the disease is transient and the individual can recover and becomes susceptible again. Cold and Flu can be modelled as a SIS model, because a person can be infected and totally recovered several times.

A simple way to exemplify the SIS model is presented in Figure 6, where S is the Susceptible state and I is the Infected state. It is assumed that λ is the infection rate and μ is the cure rate.

Figure 6 – A SIS model's diagram.

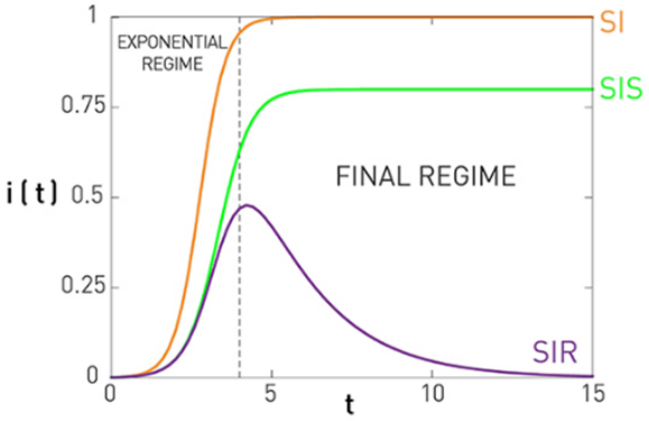


Source: Martorello, Cristiane Dias de Souza, 2018

The outcomes of the three kinds of epidemiological models cited above are different for large times: In the SI model everyone becomes infected; in the SIR model infected individuals disappears and in the SIS model there are two possible final regimes: either

reaches an endemic state, in which a finite fraction of individuals is always infected, or the infection dies out. Figure 7 shows the final regime for the three models cited above in the case of SIS having an endemic state.

Figure 7 – Fraction of infected $i(t)$ × time (t). It represents the final regime for different epidemiological models.



Source: Barabasi (2015)

The SIS model has a topological complex pattern due to its epidemic threshold, that can be positive or absent, however it keeps a simple dynamic. For regular lattice, there is a critical value λ_c that separates the endemic phase from the absorbing state - a state that, once entered, cannot be left.

In a uniformly mixed assumption, where every pairs of individuals has the same probability of having contact to each other, the model mechanism for a SIS model can be described as

$$\begin{cases} \dfrac{dS}{dt} &= -\lambda S(t)I(t) + \mu I(t) \\ \\ \dfrac{dI}{dt} &= \lambda S(t)I(t) - \mu I(t), \end{cases} \tag{3.1}$$

where $S(t)$ is the proportion of the susceptible individuals in the population and $I(t)$ is the proportion of infected individuals. The total population is assumed to be fixed, which means that $S(t) + I(t) = 1$. The term $\lambda S(t)I(t)$ means the number of infections caused by I infected individuals per unit of time and $\mu I(t)$ is the fraction of recovery individuals per unit of time.

Only non-negative solutions is interesting for us. The initial conditions are $S(0) = S_0 > 0$ and $I(0) = I_0 > 0$.

Using the relation $S(t) + I(t) = 1$ on equation (3.1):

$$\frac{dI}{dt} = (\lambda - \mu)I(t) - \lambda I(t)^2. \tag{3.2}$$

The solution of this equation is

$$I(t) = \left(1 - \frac{\mu}{\lambda}\right) \frac{1}{1 + ce^{(-\mu+\lambda)t}},$$

where $c = \frac{\lambda(1-I_0)-\mu}{\lambda I_0}$.

To find out in which situations the epidemic will spread, we considered the stationary state, where $\frac{dI}{dt} = 0$. From equation (3.2):

$$(\lambda - \mu)I(t) - \lambda I(t)^2 = 0 \begin{cases} I_\infty = 0 \rightarrow \text{the infection dies out} \\ I_\infty = 1 - \frac{\mu}{\lambda} \rightarrow \text{there are infected and susceptible individuals;} \end{cases}$$

the stability of $I_\infty = 0$ is given by

$$\frac{d}{dI}\left[(\lambda - \mu)I(t) - \lambda I(t)^2\right]_{I=I_\infty} > 0\ . \tag{3.3}$$

Then the epidemic threshold is given by an equality in (3.3), which yields $\lambda_c = \mu$.

The main focus of this work is concentrated on the SIS model. We present mean-field cases based on SIS, including annealed and quenched mean-field. The simulation developed is also based on SIS model in several different types of networks.

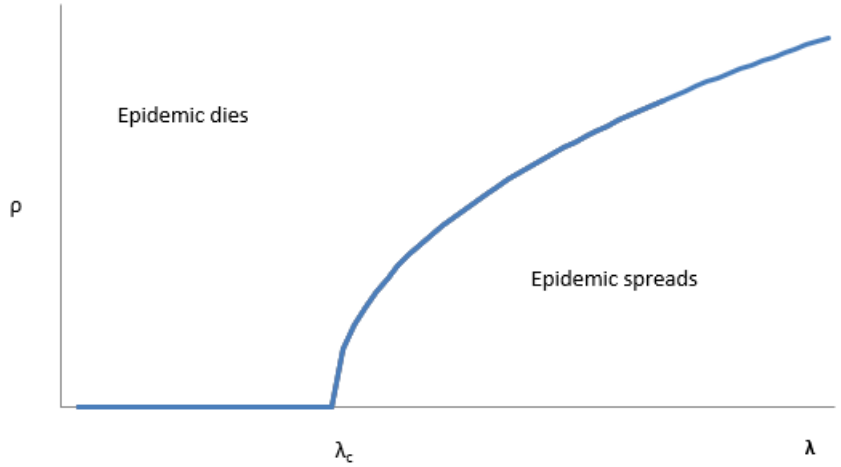
3.1 Epidemic threshold

An epidemic threshold characterizes the critical condition on which a global epidemic occurs, and it depends on the network structure. Being able to predict an epidemic threshold allows us to separate the cases when an infection spreads on a population from the case where the epidemics dies out. The determination of this value allows many practical applications, for instance, the effectiveness of a given immunization strategy because a immunization is effective when it increases the value of epidemic threshold (12).

An epidemic threshold is given by the critical value λ_c that separates the behavior of the spreading of the infection between a healthy and an endemic phase on the SIS model (Figure 8). Researchers have developed several theoretical methods for predicting

λ_c, including the heterogeneous mean-field method (HMF), also called annealed, and the quenched mean-field method (QMF). The deductions of the epidemic threshold in these cases will be made ahead in the text.

Figure 8 – The rate λ_c is the epidemic threshold that separates the healthy phase from the spreading disease phase.



Source: Martorello, Cristiane Dias de Souza, 2018

Annealed approach uses the degree distribution as a parameter, while the quenched mean-field method describes the network topology in terms of an adjacent matrix. In other words, annealed can be seem as a mean of values taken before the events occurs - it is a statistical value considering expecting values. The Quenched mean-field considers the actual structure of network, and it considers the mean value after some events occurs.

The epidemic threshold of the infection rate λ for HMF is

$$\lambda_c^{HMF} = \frac{\langle k \rangle}{\langle k^2 \rangle}, \tag{3.4}$$

where $\langle k \rangle$ and $\langle k^2 \rangle$ are the first and the second moments of the degree distribution (13). The threshold vanishes when $\langle k^2 \rangle$ diverges.

The behavior of the SIS model is relevant in the case of heterogeneous networks for which a vanishing epidemic threshold occurs (14). In (3), the authors proposed a HMF theory for complex networks, and find that the epidemic threshold of SIS model in scale-free networks, characterized by $P(k) \sim k^{-\gamma}$, with a degree distribution exponent γ, is zero for $\gamma \leq 3$. The HMF theory is exact in the annealed network. This "absence of an epidemic threshold and its associated critical behavior, which implies that scale-free networks are prone to the spreading and the persistence of infections whatever spreading rate the epidemic agents possess" (3), therefore these absence of epidemic threshold of

the SIS model transformed the analysis of epidemics on scale-free networks, because this absence makes these networks vulnerable for propagation of infections and computational viruses, for example.

In the article (14), the authors provide arguments showing that the epidemic threshold asymptotically vanishes in any network with a degree distribution decaying slower than exponentially, so it includes the QMF. However, in the study of Lee *at al* (15), the authors focuses in the QMF theory for the SIS model and they pointed that, in QMF theory, λ_c should have a finite value for an endemic phase in which a finite fraction of infected vertices is present for a random scale-free network. In the reply letter (16), the authors conclude that Lee *at al* are correct, but in some situations the result of the initial article is correct.

The epidemic threshold for QMF is given by

$$\lambda_c^{QMF} = \frac{1}{\Lambda}, \tag{3.5}$$

where Λ is the largest eigenvalue of the adjacent matrix. For scale-free networks with degree exponent large than 2.5, it was found that the largest eigenvalue Λ is determined by the maximum degree q_{max}, $\Lambda \propto \sqrt{q_{max}}$ (13). In infinite size limit, q_{max} tends to infinity, then the epidemic threshold is absent in networks with a infinite $\langle q^2 \rangle$, and the conclusion is that the epidemic threshold may be absent even in the networks with rapidly decaying degree distribution (13).

4 Heterogeneous Mean-Field

The first mathematical approach to SIS in complex networks was based on the heterogeneous mean-field (HMF), expressing the probability that two vertices connects in average, the so-called annealed network (17). The HMF can deal with heterogeneity in the degree distribution.

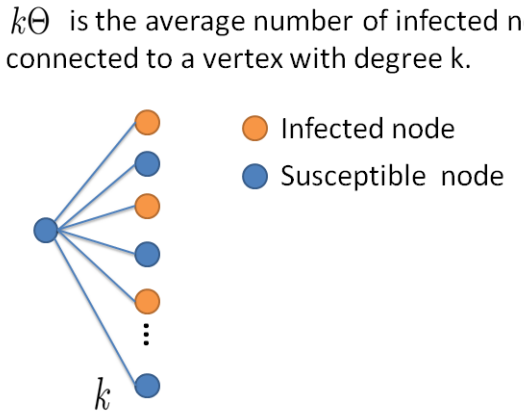
The master equation suitable for the definition of SIS model can be written as in (2),

$$\partial_t \rho_k(t) = -\mu \rho_k(t) + \lambda k (1 - \rho_k(t)) \Theta_k(t), \tag{4.1}$$

where ρ_k is the density of infected individuals with degree k, that is, the probability that a vertex with k connections is infected. The term $-\mu\rho_k$ represents individuals who become susceptible at a recovery rate μ which contributes negatively to $\partial_t \rho_k$. The second term on the right-hand side says that the positive variation of infected population is proportional to the infection rate λ, to degree k, to the density of healthy individuals with k links $(1 - \rho_k(t))$ and to the probability $\Theta_k(t)$ that a link, which belongs to a node with degree k, points to an infected vertex - see Figure 9.

The initial condition is $\rho_k(t_0) = \rho_{k0}$. Without loss of generality, we can consider the cure rate $\mu = 1$, because it is necessary only to adapt λ by a factor and scale the time accordingly. We will assume that Θ_k is independent of degree k, so $\Theta_k = \Theta$. The model considers a closed population of size N.

Figure 9 – $k\Theta$ is the average number of infected nodes connected to a vertex with degree k.



Source: Martorello, Cristiane Dias de Souza, 2018

In the case of an uncorrelated random network, the probability that an edge points to a vertex with k connections is equal to (8)

$$\frac{kP(k)}{\langle k\rangle}, \tag{4.2}$$

then,

$$\Theta = \frac{1}{\langle k\rangle}\sum_{k'} k'P(k')\rho_{k'}(t). \tag{4.3}$$

With (4.1) and (4.3) we can analyze the stability of stationary state with no infection. The stationary condition $\partial_t\rho_k(t) = 0$ implies

$$-\rho_k + \lambda k(1-\rho_k)\Theta = 0, \tag{4.4}$$

thus

$$\rho_k = \frac{\lambda k\Theta}{1+\lambda k\Theta}. \tag{4.5}$$

Substituting (4.5) in (4.3),

$$\Theta = \frac{1}{\langle k\rangle}\sum_{k} kP(k)\frac{\lambda k\Theta}{1+\lambda k\Theta}. \tag{4.6}$$

Note that $\Theta = 0$ is always a solution and corresponds to the state where the disease is absent. One should examine if the equation (4.6) allows a nontrivial solution. Representing the right-hand-side by $f(\Theta)$, we have

$$\begin{aligned} f(0) &= 0 \\ f(1) &= \frac{1}{\langle k\rangle}\sum_{k} kP(k)\frac{\lambda k}{1+\lambda k} \leq \frac{1}{\langle k\rangle}\sum_{k} kP(k)\frac{1+\lambda k}{1+\lambda k} \\ &= \frac{1}{\langle k\rangle}\underbrace{\sum_{k} kP(k)}_{\langle k\rangle} = 1 \\ f(1) &\leq 1. \end{aligned} \tag{4.7}$$

Then, considering $f'(\Theta)$,

$$f'(\Theta) = \frac{1}{\langle k\rangle}\sum_{k} kP(k)\frac{\lambda k}{(1+\lambda k\Theta)^2} > 0, \tag{4.8}$$

so $f(\Theta)$ is a crescent function and $f''(\Theta)$,

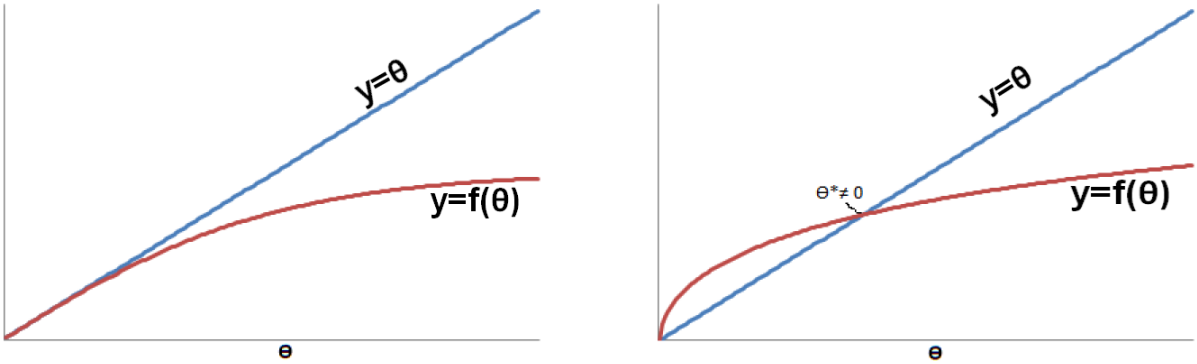
$$f''(\Theta) = -\frac{2}{\langle k\rangle}\sum_k kP(k)\frac{(\lambda k)^2}{(1+\lambda k\Theta)^3} < 0, \tag{4.9}$$

and therefore $f(\Theta)$ is a concave function. These equations give us that there is a nontrivial solution when $f'(0) > 1$, and there is only the trivial solution when $f'(0) < 1$, then the critical situation happens when $f'(0) = 1$. Using equation (4.8), $\lambda_c = \frac{\langle k\rangle}{\langle k^2\rangle}$, because

$$\begin{aligned} f'(0) &= \frac{1}{\langle k\rangle}\sum_k \lambda_c k^2 P(k) = 1 \\ 1 &= \frac{1}{\langle k\rangle}\lambda_c \underbrace{\sum_k k^2 P(k)}_{\langle k^2\rangle} \\ \lambda_c &= \frac{\langle k\rangle}{\langle k^2\rangle}. \end{aligned} \tag{4.10}$$

In Figure 10, we have the graphical solution of (4.6). There are sketches for graphs with only the trivial solution and with one nontrivial solution.

Figure 10 – Only trivial solution in the left graph and one nontrivial solution in the right graph.



Source: Martorello, Cristiane Dias de Souza, 2018

4.1 Stationary solution in Barabási-Albert model

In the Barabási-Albert model, the degree distribution, as in (2.7), is given by

$$P(k) = \frac{2m^2}{k^3}, \tag{4.11}$$

where m is the number of connections of the new incoming vertex at each time step. For nonzero values of λ, which implies $\Theta(\lambda) \neq 0$, and adopting the Barabási-Albert distribution (4.11), the equation (4.6) can be cast as

$$1 = \lambda m \int_m^\infty \frac{1}{k[1 + \lambda k \Theta(\lambda)]} dk. \tag{4.12}$$

The integral above can be evaluated by partial fraction decomposition, which leads to

$$\begin{aligned} \frac{1}{\lambda m} &= \lim_{R\to\infty} \int_m^R \left[\frac{1}{k} - \frac{\lambda\Theta(\lambda)}{1 + \lambda k \Theta(\lambda)}\right] dk \\ &= \lim_{R\to\infty} \left[\ln(1 + \lambda\Theta(\lambda)m) - \ln\left(\lambda m\Theta(\lambda) + \frac{m}{R}\right)\right]. \end{aligned} \tag{4.13}$$

For small values of λ, which implies small values of $\Theta(\lambda)$, the equation (4.13) becomes

$$\Theta(\lambda) \simeq \frac{1}{\lambda m} e^{-\frac{1}{\lambda m}}, \tag{4.14}$$

and substituting in equation (4.5), we have

$$\rho_k \simeq \frac{k}{m} e^{-\frac{1}{\lambda m}}. \tag{4.15}$$

Then, with $\lambda \sim 0$, the total density of infected nodes is

$$\begin{aligned} \rho &= \sum_k P(k)\rho_k \\ &\simeq \int_m^\infty dk P(k)\rho_k \quad \simeq \quad 2e^{-\frac{1}{\lambda m}}. \end{aligned} \tag{4.16}$$

This result indicates the absence of any threshold for epidemics, which means $\lambda_c = 0$.

4.2 "Simplified" Mean-Field

We call "simplified" a particular case in the HMF, $\Theta(t) = \theta$, where θ is a constant, in equation (4.1). Then,

$$\partial_t \rho_k(t) = -\rho_k(t) + \lambda k(1 - \rho_k(t))\theta, \tag{4.17}$$

with $\rho_k(t_0) = \rho_{k0}$. The fact that we consider θ constant means statistically that we admit that each link has the same probability of connection to an infected individuals. Although it is not a commonly explored model, it is important to investigate the equation.

The stationary solution is given by $-\rho_k + \lambda k(1-\rho_k)\theta = 0$, then

$$\rho_k = \frac{\lambda k\theta}{1+\lambda k\theta}. \tag{4.18}$$

Note that there is not a stationary solution in $\rho_k = 0$, except for $\lambda = 0$.

The equilibrium point is obtained from the condition $\partial_t \rho = 0$, and the critical λ can be obtained in equation (4.17) using the stability analysis. Denoting by $f(\rho_k)$ the right-hand side of equation (4.17), we have

$$f(\rho_k) = -\rho_k + \lambda k(1-\rho_k)\theta, \tag{4.19}$$

then

$$f'(\rho_k) = -1 - \lambda k\theta. \tag{4.20}$$

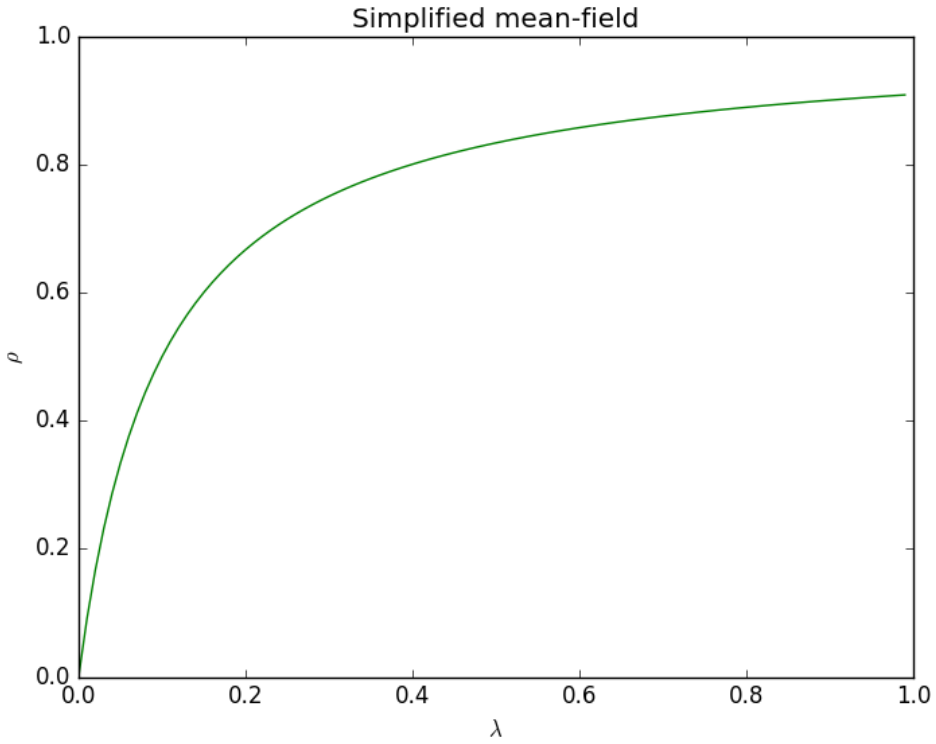
Note that f' is always negative, because $\lambda \geq 0$, $k \geq 0$ and $\theta \geq 0$. Then $\lambda_c = 0$ and the solution (4.18) is stable for $\lambda > 0$.

Using the method of variation of parameters (Appendix A) to solve differential equations, we obtain the following result from (4.17):

$$\rho_k(t) = \left(\rho_{k0} - \frac{\lambda k\theta}{1+\lambda k\theta}\right) e^{-(t-t_0)(1+\lambda k\theta)} + \frac{\lambda k\theta}{(1+\lambda k\theta)}. \tag{4.21}$$

In Figure 11 we can see the dynamics of simplified mean-field for the stationary state of infected individuals for different values of λ.

Figure 11 – Simplified mean-field. For $t = 200$, $k = 100$, $\theta = 0.1$ and $\rho_0 = 0.1$.



Source: Martorello, Cristiane Dias de Souza, 2018

4.3 Homogeneous Mean-Field

The second particular case of HMF considers the graph structure of the homogeneous mean-field. This approach assumes that each vertex has approximately the same number of connections $k \approx \langle k \rangle$ and that the probability of a susceptible individual connected to an infected node is approximately the fraction of infected individuals in the network, i.e., $\Theta = \rho$. Besides, the fraction of infected nodes with degree k corresponds to the global fraction of infected vertices, which means that $\rho_k \approx \rho$ (4). The dynamical equation for this case is

$$\partial_t \rho(t) = -\rho(t) + \lambda \langle k \rangle (1 - \rho(t))\rho(t). \tag{4.22}$$

It is assumed the homogeneous mixing hypothesis, where the strength of the disease is proportional to the density of infected individuals $\langle k \rangle \rho(t)$ (2), that is, proportional to the mean degree of the vertex $\langle k \rangle$ and proportional to the infected population.

The solution of the equation (4.22) was obtained using the partial fractions method:

$$\rho(t) = \frac{\lambda\langle k\rangle - 1}{ce^{-t(\lambda\langle k\rangle - 1)} + \lambda\langle k\rangle}, \quad \text{where} \quad c = \frac{-1 + \lambda\langle k\rangle(1-\rho_0)}{\rho_0}. \tag{4.23}$$

The equilibrium points of the equation (4.22), obtained from $\partial_t\rho(t) = 0$, are

$$\rho^* = 0 \quad \text{and} \quad \rho^* = 1 - \frac{1}{\lambda\langle k\rangle}. \tag{4.24}$$

The critical point λ_c can be obtained by equation (4.22) using the stability analysis for ρ^*, and

$$\lambda_c = \frac{1}{\langle k\rangle}, \tag{4.25}$$

therefore:

$$\begin{cases} \lambda < \frac{1}{\langle k\rangle} \rightarrow \rho^* = 0 \text{ is a stable point, the disease disappears over time,} \\ \lambda > \frac{1}{\langle k\rangle} \rightarrow \rho^* = 0 \text{ is a unstable point, the disease spreads and is persistent.} \end{cases}$$

A similar studied case considers a fully connected network, where the number of connections of each vertex is $(N-1)$ and $\Theta(\rho_k(t)) \approx \rho$. Then

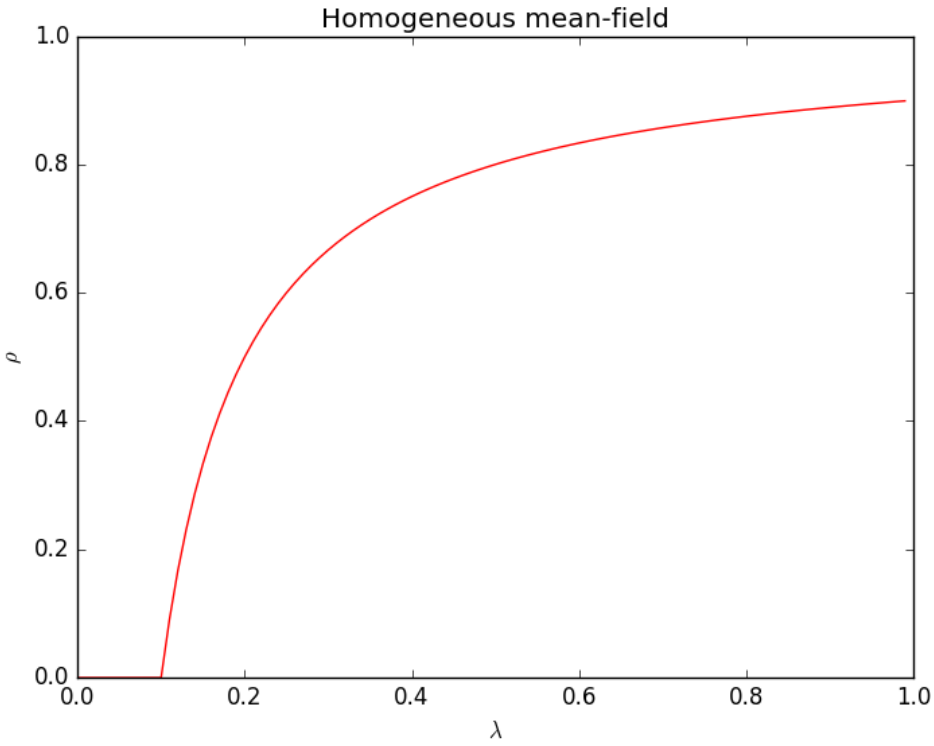
$$\partial_t\rho(t) = -\rho(t) + \lambda(N-1)(1-\rho(t))\rho(t). \tag{4.26}$$

It is only necessary to replace $\langle k\rangle$ by $N-1$ in the results of the equation (4.23):

$$\rho(t) = \frac{\lambda(N-1) - 1}{ce^{-t(\lambda(N-1)-1)} + \lambda(N-1)}, \quad \text{with} \quad c = \frac{-1 + \lambda(N-1)(1-\rho_0)}{\rho_0}. \tag{4.27}$$

In Figure 12 we show dynamics of homogeneous mean-field in the stationary state for different values of λ.

Figure 12 – Homogeneous mean-field with $\langle k \rangle = 10$ and $t = 100$. Here, $\lambda_c = \frac{1}{\langle k \rangle} = 0.1$.



Source: Martorello, Cristiane Dias de Souza, 2018

5 Quenched Mean-Field

The Quenched Mean-Field model (QMF) (18) considers the actual connectivity of the network through the adjacency matrix A. In the article (18), the authors proposed a general epidemic threshold condition that is applied to general graphs and the threshold is related to the largest eigenvalue of the adjacency matrix. The equation for QMF is

$$\partial_t \rho_i(t) = -\rho_i(t) + \lambda(1 - \rho_i(t)) \sum_{j=1}^{N} a_{ij}\rho_j, \tag{5.1}$$

where a_{ij} is the element of the adjacency matrix A, with $a_{ij} = 1$ if there is a link between vertices i and j, and $a_{ij} = 0$ otherwise, and ρ_i is the probability of the vertex i being infected - note that ρ_i has a different meaning from the function ρ_k from the previous sections. The first term in right-hand side of equation (5.1) reduces the probability of the vertex i being infected because it is the recovery rate and the second term of the right-hand side considers that the node i is susceptible and gets the infection by an infected neighboring nodes of i - $\sum_{j=1}^{N} a_{ij}\rho_j$. Analyzing the linear stability of the steady-state solution around $\rho_i = 0$ - which is the state without disease, we can linearize (5.1), getting

$$\partial_t \rho_i(t) \approx \sum_{j=1}^{N} L_{ij}\rho_j, \tag{5.2}$$

where the Jacobian matrix is

$$L_{ij} = -\delta_{ij} + \lambda a_{ij}, \tag{5.3}$$

and δ_{ij} is the Kronecker delta symbol. The solution $\rho_i = 0$ is unstable if there exists a positive eigenvalue of the matrix L. Let Λ_m be the largest eigenvalue of A and Ψ_m is the largest eigenvalue of L, so $\Psi_m = -1 + \lambda\Lambda_m$ (19). This approach (see Appendix B) predicts the existence of an epidemic threshold given by the inverse of the largest eigenvalue of the adjacent matrix (2):

$$\lambda_c = \frac{1}{\Lambda_m}. \tag{5.4}$$

5.1 Complete Graph

For the case where the adjacency matrix has $a_{ij} = 1$ for any i and j, $i \neq j$, it represents a fully connected network and the dynamical equation for the SIS model (5.1) becomes

$$\partial_t \rho_i(t) = -\rho_i(t) + \lambda(1 - \rho_i(t)) \sum_j \rho_j. \tag{5.5}$$

If we sum over all the sites, we have

$$\partial_t \rho(t) = -\rho(t) + \lambda N \rho(t)(1 - \rho(t)), \tag{5.6}$$

where

$$\rho(t) = \frac{1}{N} \sum_i \rho_i(t).$$

With the initial condition $\rho(0) = \rho_0$, the solution through the method of partial fractions yields

$$\rho = \frac{-1 + \lambda N}{ce^{-t(-1+\lambda N)} + \lambda N}, \qquad \text{where} \qquad c = \frac{-1 + \lambda N}{\rho_0} - \lambda N. \tag{5.7}$$

Calculating $\lim_{t\to\infty} \rho(t)$,

$$\lim_{t\to\infty} \rho(t) = \begin{cases} 1 - \frac{1}{\lambda N} & , \text{if } \lambda > \frac{1}{N} \\ 0 & , \text{if } \lambda < \frac{1}{N} \end{cases}$$

so $\lambda_c = \frac{1}{N}$. This result could also be obtained by the eigenvalues of A.

6 Modeling Scale-Free Complex Networks

At this point, we want to formulate an algorithm that generates a scale-free network with degree distribution $P(k) \sim k^{-\alpha}$. This algorithm should have the exponent α as the entry and the output is the adjacency matrix consistent with this degree distribution. There are two related articles that give us the idea for the algorithm cited below.

The first article is (20), where the authors construct scale-free networks, with N nodes, assigning weights $p_i = i^{-\gamma}$ to the vertex i, $1 \leq i \leq N$, and γ is a control parameter between 0 and 1. The construction process is: select two different vertices (i, j) with probabilities $p_i/\sum_k p_k$ and $p_j/\sum_k p_k$, and add a link between them until mN edges are made in system, where m are the already existing vertices in the network, and they find that the degree distribution follows a power-law, $P(k) \sim k^{-\alpha}$, where $\alpha = (1+\gamma)/\gamma$. Thus, choosing a suitable value for γ, it is possible to construct a network with desired exponent α.

This idea of weighting vertices is similar to the article (21), the second article, where the prescription examines the average distances in random graphs with a given expected degree. These conceptions was considered in this study to create a generator of a scale-free complex network with a power-law degree distribution $P(k) \sim k^{-\alpha}$. Consider a general model $G(\mathbf{w})$ with given expected degree sequence $\mathbf{w} = (w_1, w_2, ..., w_n)$. The edge between two nodes i e j is chosen with probability p_{ij}, where p_{ij} is proportional to the product $w_i w_j$:

$$p_{ij} = \frac{w_i w_j}{\sum_{k=1}^{N} w_j}. \tag{6.1}$$

Note that this equation (6.1) corresponds to the annealed adjacency matrix and the degree k_i of the vertex i is given by

$$k_i = \sum_{j=1}^{N} p_{ij} = w_i. \tag{6.2}$$

Starting from this setup, a connection is introduced to the graph according to the linking probability given by (6.1) at each time step.

For random graphs with a given expected degree sequence satisfying a power-law distribution with exponent α, the article (21) assumes that the expected degrees are $w_i = ci^{-\frac{1}{(\alpha-1)}}$ where c depends on the average degree d: $c = \frac{\alpha-2}{\alpha-1} dN^{\frac{1}{(\alpha-1)}}$.

In next subsection we show why the prescription above generates a scale-free network.

The pseudocode with the general idea for the algorithm is described below.

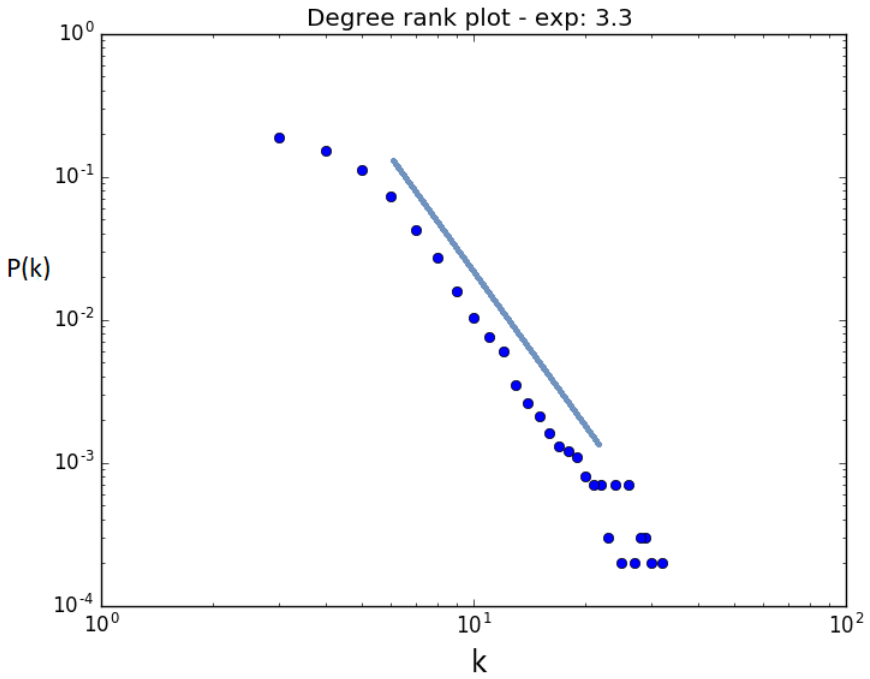
Algorithm 1: Creation of a power-law degree distribution network with a given exponent β.
Input: The number of nodes N and the power-law exponent β
Output:A power-law network with the given exponent β

```
begin
    Calculate c
    for all nodes i = 1, ...N do
        Calculate w_i
    end
    Set w_sum is the sum of w_i
    for all nodes i = 1, ...N do
        j = 1
        while j < N do
            Set a random uniform r between 0 and 1
            if r < w_i w_j / w_sum then
                create a edge between i and j
            end
        end
    end
end
```

The algorithm 1 above was developed in Python and the results were satisfactory. Figure 13 was constructed by Algorithm 1, with parameters 3.5 for the exponent and N=10.000 nodes. The exponent obtained was 3.3. Varying the exponent parameter to 3.0, we obtained by the algorithm the value of 3.1 (Figure 14), and for exponent 2.5, we obtained by the algorithm a network with exponent 2.3 (Figure 15). These values were obtained by Ordinary Least Squares in the linear piece in log-log graph.

In Appendix E, the algorithm programmed in Python is presented.

Figure 13 – Log-log graph for a power-law network with a given exponent = 3.5 and N=10000. The exponent obtained was 3.3.



Source: Martorello, Cristiane Dias de Souza, 2018

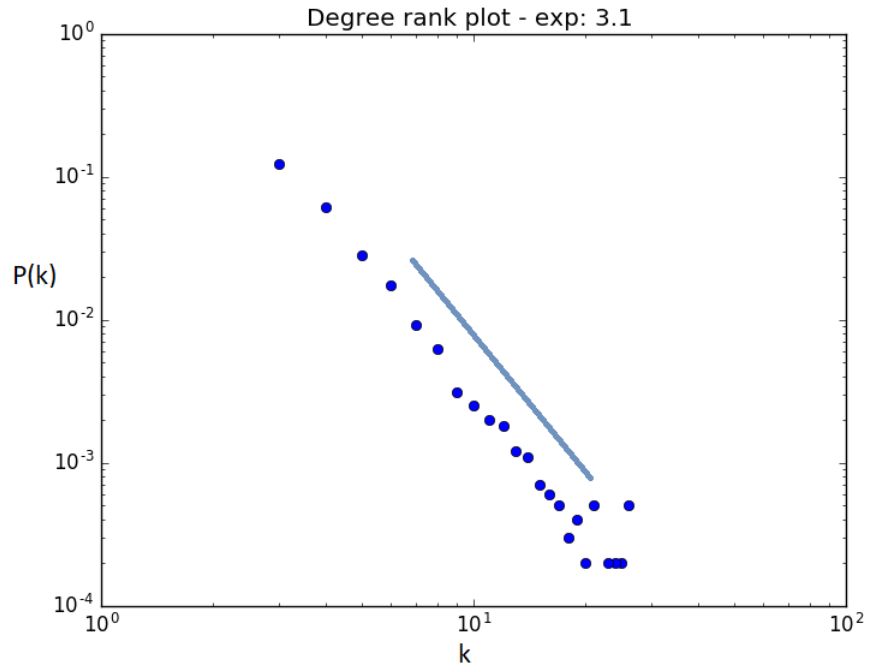
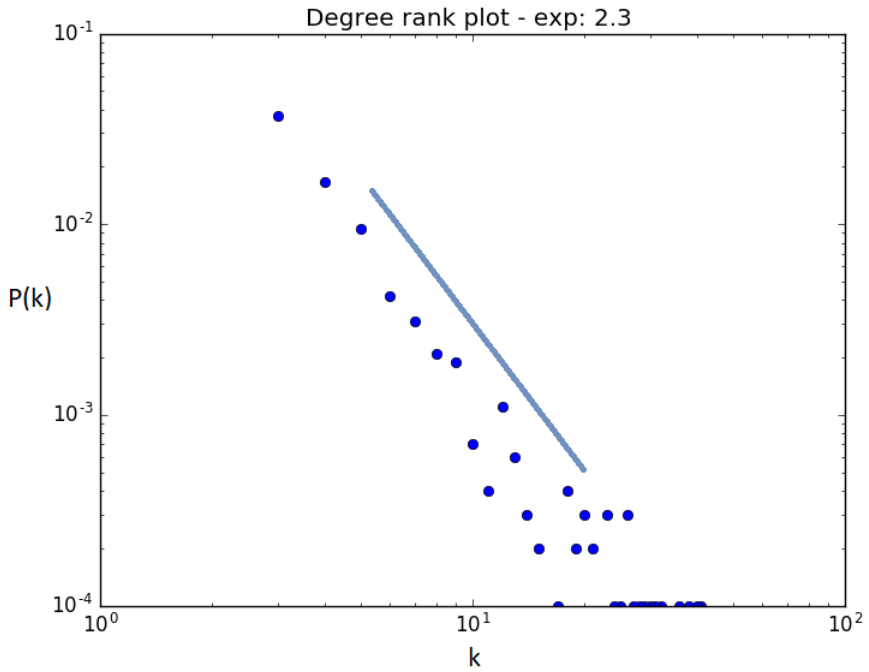


Figure 14 – Log-log graph for a power-law network with a given exponent = 3.0 and N=10000. The exponent obtained was 3.1.

Figure 15 – Log-log graph for a power-law network with a given exponent = 2.5 and N=10000.The exponent obtained was 2.3.



Source: Martorello, Cristiane Dias de Souza, 2018

6.1 Power-law degree distribution network

Suppose that the N vertices are labeled according to its degree. Let $k_N \leq k_{N-1} \leq ... \leq k_1$, then the probability that a node m has degree larger or equal than k_m is $\frac{m}{N}$. We can evaluate this probability as $\int_{k_m}^{\infty} P(k)dk$, where P is the degree distribution, therefore,

$$\int_{k_m}^{\infty} P(k)dk = \frac{m}{N}. \tag{6.3}$$

Here, the sum over the degrees is approximated by an integral and is asymptotically exact. From equation (6.2),

$$\int_{w_m}^{\infty} P(k)dk = \frac{m}{N}, \tag{6.4}$$

and this is the main result that allows us to deduce the degree distribution from a given weight choice. Deriving (6.4) leads to

$$P(k) = -\frac{1}{N}\frac{d}{dk}w_k^{-1}(k), \tag{6.5}$$

where w_k^{-1} is the inverse function of w_k.

Since $w_k = k^{-\alpha}$, we have $w_k^{-1} = k^{-1/\alpha}$, then

$$P(k) = \frac{1}{\alpha N}k^{-\frac{1}{\alpha}-1}. \tag{6.6}$$

Therefore, choosing the weight $w_m = m^{-\frac{1}{\beta-1}}$, then the generated network will have a degree distribution $P(k) \sim k^{-\beta}$.

7 Computational Simulation

The computational simulation is a fundamental tool to help analyzing and understanding epidemic spreading. This work applied the notion presented in (22) for the SIS model dynamics. The article considers the Gillespie algorithm (23) decomposing the dynamics into independent processes and performing a change of state by each time step.

The Gillespie algorithm is a Monte Carlo simulation procedure, in its original form was outlined in four steps as follows (23): Step 1 - Initialization of all variables; Step 2 - Generate random values to determine the occurrence of the next event and the time interval according to Monte Carlo; Step 3 - Increase the time step by the randomly generated time in Step 2; and Step 4 - Recalculation of the variables and finish or return to initialization.

According to (22), it says that this first algorithm can be "computationally prohibitive for large networks" mainly due the time-consuming update of all lists related to the state of the node and size of the time step. However, it proposes a more efficient simulation called Optimized Gillespie algorithm that includes "phantom processes, which do not change the states but do count for time increments". These phantom processes do not consider steps that do not change the states in the dynamic of the network, because it keeps the last position of the network. The Optimized Gillespie algorithm considers an adjacent matrix A with all connections between the nodes $i = 1, ..., N$. Vertices have two states: it can be infected (with value 1) or susceptible (with value 0). An infected vertex i can be healed with rate μ and becomes a susceptible vertex, or it can transmit the disease to a susceptible neighbor node j with rate $\lambda_{ij} = \lambda A_{ij}$.

In Gillespie model, two lists are necessary - one list has the positions of the infected vertices and one list has the address of all the neighbors of the infected vertices. Optimized Gillespie algorithm only has the necessity to create one list with the infected nodes because it uses the phantom processes. This change contributes to significantly decrease the processing time.

The total healing rate is $M = \mu N_I$, where N_I is the number of infection nodes. The infection rate is $L = \lambda N_{IS}$, where N_{IS} is the number of edges pointing from an infected to a susceptible vertex. The total rate of transitions is $R = L + M$. With probability $m = M/R$, one infected vertex is chosen and healed. With probability $l = L/R$, one

susceptible neighbor is chosen and infected. After the dynamics, the variables and list are updated.

To initiate the epidemic, we randomly select some "seed" nodes as infected and designate the others as susceptible vertices.

The pseudocode is presented in algorithm 2. The implementation code in Python language is in Appendix E.

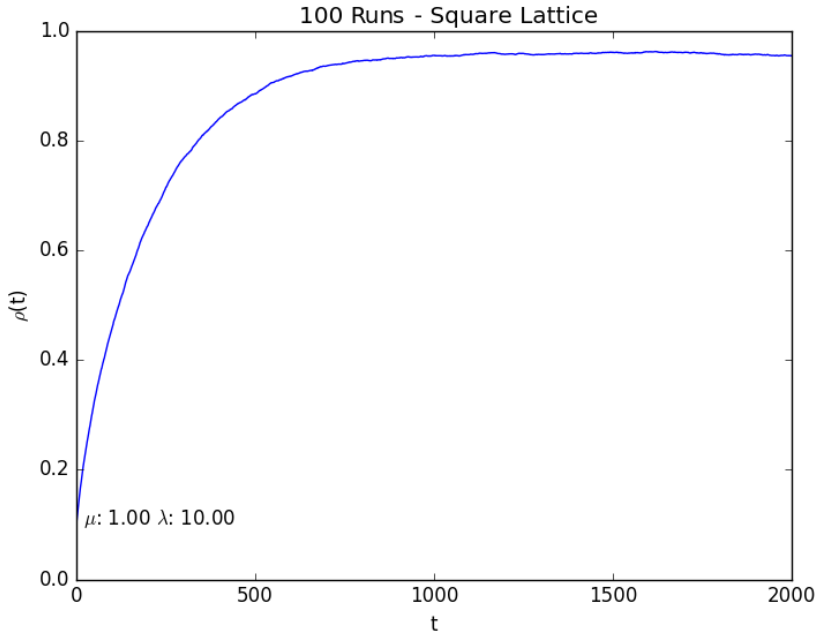
Algorithm 2: Dynamics for an epidemic spreading using the optimized Gillespie algorithm (22).
Input: a network
λ: infectious rate
μ: cure rate
Output: The dynamics for $\rho(t)$
begin
  Draw the infected positions at random given a infected rate;
  Calculate the healing rate $M = \mu N_I$
  for *all infected nodes* **do**
    Sum the edges of the infected positions through the adjacent matrix and find N_{IS}
  end
  Calculate the infection rate $L = \lambda N_{IS}$;
  Calculate the total rate of transitions $R = L + M$ and the probabilities m and l;
  for *all time step* **do**
    Choose a uniform random number between 0 and 1, *rand*
    if $rand < m$ **then**
      one vertex is chosen and healed
    else
      one vertex is chosen according to its degree
      one neighbor is chosen
      if *the neighbor is susceptible* **then**
        the neighbor gets infected
      end
    end
  end
end

The algorithm 2 above can be applied for many different networks. We applied for a square lattice 100 x 100 and change the infectious rate to obtain differents outputs. These results are in Figures 16 and 17. While in the former the infection persists, in the latter it vanishes, indicating the existence of an infection threshold.

We simulate a Homogeneous Mean-Field model for $\langle k \rangle = 4$, to verify the main differences between the Homogeneous model and the Square lattice for the same values of

the infection rate as Figures 16 and 17, and we obtained different results once the epidemic threshold is given by (4.25), $\lambda_c = \frac{1}{\langle k \rangle} = 0.25$ - Figure 18. It indicates that, in this case, the epidemic threshold for the square lattice is higher than in the homogeneous mean-field.

Figure 16 – Graph of SIS dynamics with regular lattice 100 x 100 and $\lambda = 10$, with 100 runs.

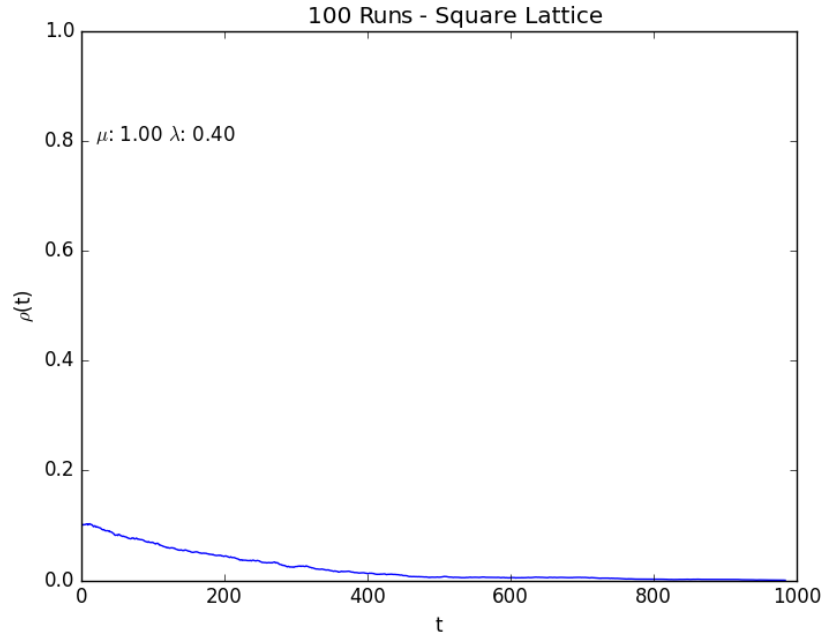


Source: Martorello, Cristiane Dias de Souza, 2018

We compare a complete graph simulation with the analytical result for a fully connected network (equation (4.27)) in Figure 19. Note that in the stationary state there are similar.

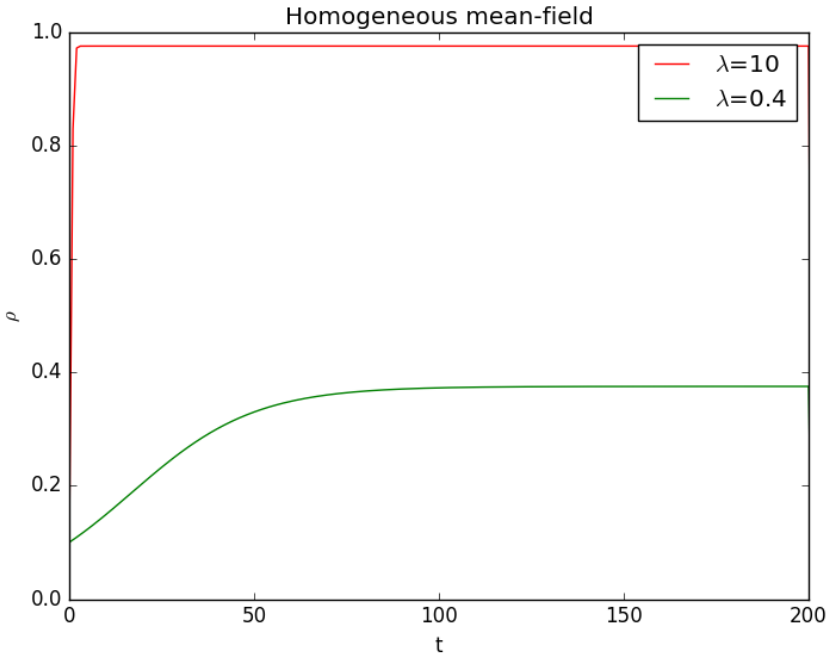
For a power-law network with degree exponent 2.5, we obtained the dynamics of the infected population $\rho(t)$ for different values of λ - Figures 20 and 21. Both simulations have a positive value for the infected state even with different behaviors of $\rho(t)$: the stationary infected population can be larger (Figure 20) or smaller (Figure 21) than the initial infected population. .

Figure 17 – Graph of SIS dynamics with regular lattice 100 x 100 and $\lambda = 0.4$, with 100 runs. The infection vanishes.



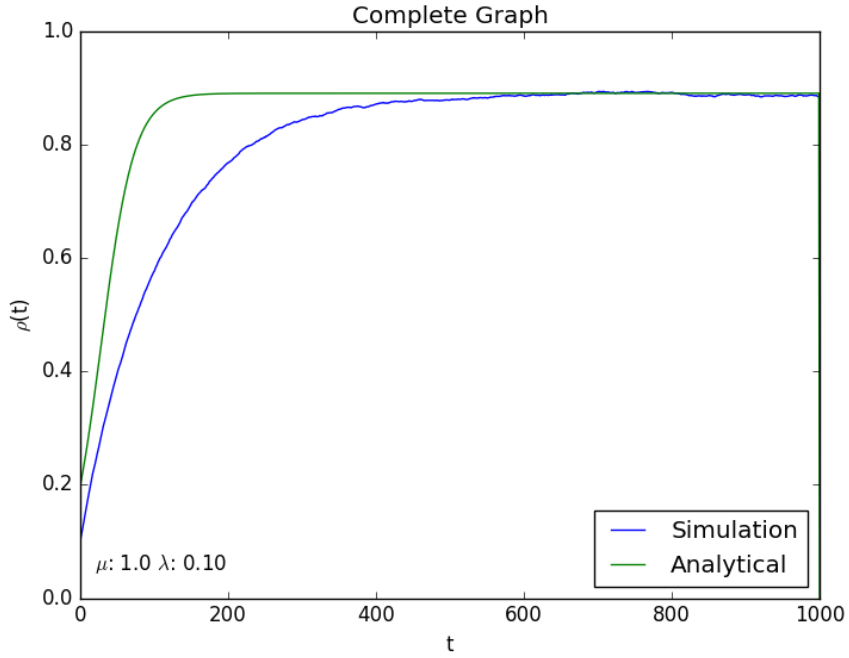
Source: Martorello, Cristiane Dias de Souza, 2018

Figure 18 – Graph of SIS dynamics with homogeneous mean-field model with $\langle k \rangle = 4$. The infection persists.



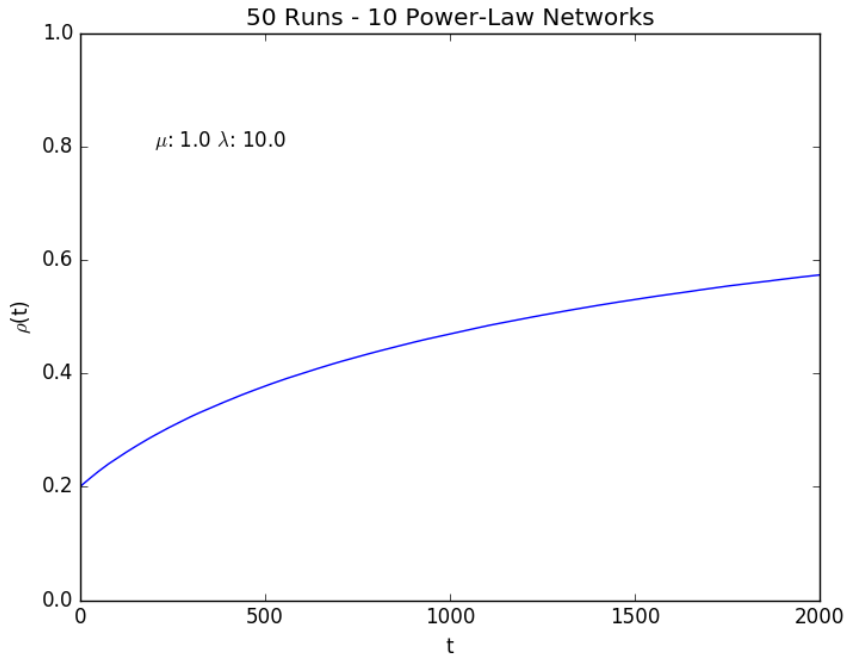
Source: Martorello, Cristiane Dias de Souza, 2018

Figure 19 – Graph with 10000 nodes, 100 runs, $\mu = 1$ and $\lambda = 0.1$.



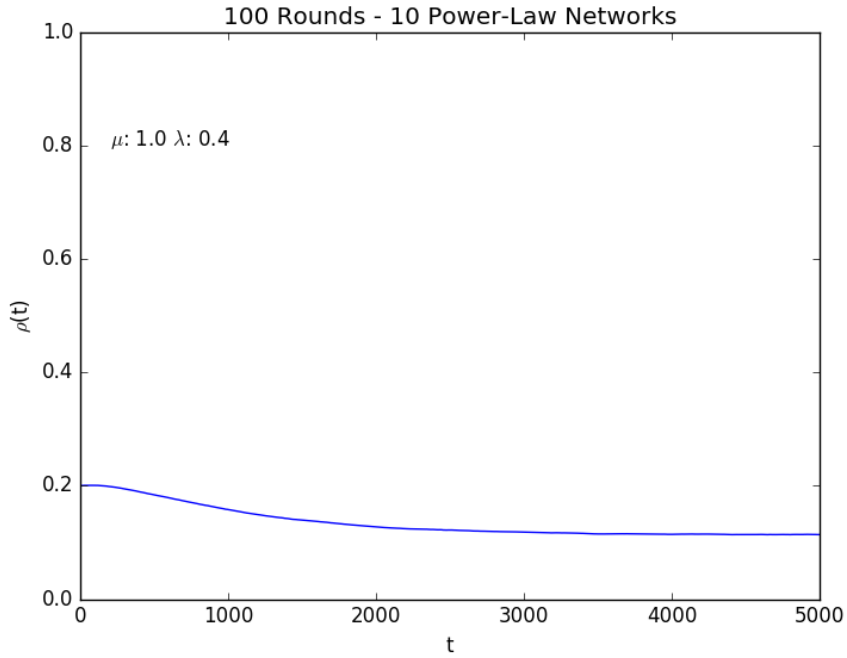
Source: Martorello, Cristiane Dias de Souza, 2018

Figure 20 – Graph with N=1000, $\mu = 1$, $\alpha = 2.5$ and $\lambda = 10$.



Source: Martorello, Cristiane Dias de Souza, 2018

Figure 21 – Graph with N=1000, $\mu = 1$, $\alpha = 2.5$ and $\lambda = 0.4$.



Source: Martorello, Cristiane Dias de Souza, 2018

8 Modified Heterogeneous Mean-Field Model

In the last chapters we present the HMF and the QMF on a network with a constant topology. Regardless of the state of the vertices (susceptible or infected), the links obtained by each node in the construction of the network are preserved: the degree distribution $P(k)$ does not change and the adjacency matrix is constant.

We want to study the effects on the dynamic of infection's propagation when the topology of the network depends on the states of the nodes. One example of it is a sick person who is at home resting due to the infection and reduces the contact with friends (social network) for a while, decreasing the probability of connection of infected individuals to susceptible ones. As a possible consequence, the epidemic threshold (λ_c) can be impacted due to the change in the network structure.

One way of decreasing the contact network for infected vertices is to change the degree distribution (in HMF case) or the adjacency matrix (in QMF case) according to the states of the vertices. However, we first simulate a decrease in the original degree (k) for the infected nodes by replacing the degree by a smaller value (k^a, with $a < 1$). And in this case, we do not change the degree distribution $P(k)$. Secondly, we simulate a change in the probability of connection of an infected node with a susceptible one in the algorithm presented in the previous section, by reducing the connection of infected nodes.

8.1 Exploring HMF with a change in the degree of the infected node

As in (4.1), the master equation for HMF is

$$\partial_t \rho_k(t) = -\mu \rho_k(t) + \lambda k(1 - \rho_k(t))\Theta_k(t), \tag{8.1}$$

where

$$\Theta_k = \frac{1}{\langle k \rangle} \sum_k kP(k)\rho_k(t). \tag{8.2}$$

As in the previous case, we assume that $\mu = 1$ and $\Theta_k = \Theta$.

In (8.2), the probability of linking to a vertex of degree k is proportional to $kP(k)$. Then, the probability of linking to a infected vertex of degree k is proportional to $kP(k)\rho_k$. To simulate a change in the network topology due to an eventual reclusion of the infected

vertices, we can suppose that the probability of linking to a infected vertex of degree k is proportional to $k^a P(k)\rho_k$, with $a \leq 1$. With this assumption, the equation (8.2) is substituted by

$$\Theta = \frac{1}{\langle k^a \rangle} \sum_k k^a P(k) \rho_k(t). \tag{8.3}$$

The stationary condition for (8.1), $\partial_t \rho_k(t) = 0$, implies

$$\rho_k = \frac{\lambda k \Theta}{1 + \lambda k \Theta}. \tag{8.4}$$

Combining (8.3) with (8.4), we have

$$\Theta = \frac{1}{\langle k^a \rangle} \sum_k k^a P(k) \frac{\lambda k \Theta}{1 + \lambda k \Theta}. \tag{8.5}$$

Note that $\Theta = 0$ remains a solution where the disease does not exist. Representing the right side of the equation (8.5) by $g(\Theta)$ we have:

$$\begin{aligned} g(0) &= 0 \\ g(1) &= \frac{1}{\langle k^a \rangle} \sum_k k^a P(k) \frac{\lambda k}{1 + \lambda k} \leq \frac{1}{\langle k^a \rangle} \sum_k k^a P(k) \frac{1 + \lambda k}{1 + \lambda k} \\ &= \frac{1}{\langle k^a \rangle} \underbrace{\sum_k k^a P(k)}_{\langle k^a \rangle} = 1 \\ g(1) &\leq 1. \end{aligned}$$

Calculating $g'(\Theta)$,

$$g'(\Theta) = \frac{1}{\langle k^a \rangle} \sum_k k^a P(k) \frac{\lambda k}{(1 + \lambda k \Theta)^2} > 0, \tag{8.6}$$

and $g''(\Theta)$,

$$g''(\Theta) = -\frac{2}{\langle k^a \rangle} \sum_k k^a P(k) \frac{(\lambda k)^2}{(1 + \lambda k \Theta)^3} < 0, \tag{8.7}$$

we can conclude that, as in the HMF, the critical situation happens when $g'(0) = 1$ (as in Figure 10). Then,

$$g'(0) = \frac{1}{\langle k^a \rangle} \lambda_c \underbrace{\sum_k k^{(a+1)} P(k)}_{\langle k^{(a+1)} \rangle} = 1, \quad \text{which implies} \quad \lambda_c = \frac{\langle k^a \rangle}{\langle k^{(a+1)} \rangle}. \tag{8.8}$$

Note that for $a = 1$, the λ_c for the traditional HMF is recovered.

Calculating $\langle k^b \rangle$, with $b < 2$, for BA model, we have

$$\begin{aligned} \langle k^b \rangle &= \int_m^\infty dk \quad k^b \frac{2m^2}{k^3} \\ &= 2m^2 \int_m^\infty dk \quad k^{(b-3)} \\ &= \frac{2m^b}{2-b}. \end{aligned} \tag{8.9}$$

Using (8.9) in (8.8),

$$\lambda_c = \frac{1-a}{(2-a)m}. \tag{8.10}$$

For $a = 1$, we have $\lambda_c = 0$, but for $a < 1$, $\lambda_c > 0$, which implies there exists a positive epidemic threshold for this modified model. In other words, there is a healthy phase in the dynamics of this modified model, for lower values of the infection rate in relation to the critical value ($\lambda < \lambda_c$). This result is different for the traditional HMF, where $\lambda_c = 0$.

To solve equation (8.5) for nontrivial solutions $\Theta > 0$, using the BA distribution for $P(k)$, (2.7), we have:

$$1 = \frac{2m^2\lambda}{\langle k^a \rangle} (\lambda\Theta)^{1-a} \int_{\lambda m\Theta}^\infty dy \frac{y^{a-2}}{1+y} \qquad (a < 1). \tag{8.11}$$

We will study equation (8.11) in the neighborhood of the epidemic threshold λ_c, when $|\Theta| \ll 1$.

The solution of (8.11) can be cast as (the details are on Appendix F)

$$1 = \Gamma(3-a)\Gamma(a-1)\lambda m(\lambda m\Theta)^{1-a} + \frac{\lambda}{\lambda_c}. \tag{8.12}$$

Simplifying and isolating Θ, the equation becomes

$$\Theta = \left\{ \frac{1}{\Gamma(3-a)(-\Gamma(a-1))(\lambda m)^{2-a}\lambda_c} \right\}^{\frac{1}{1-a}} (\lambda - \lambda_c)^{\frac{1}{1-a}}, \tag{8.13}$$

therefore

$$\Theta \sim (\lambda - \lambda_c)^{\frac{1}{1-a}}. \tag{8.14}$$

Calculating the stationary fraction of infected nodes, ρ, in the vicinity of the epidemic threshold (where $\lambda\Theta \ll 1$), we have

$$\begin{aligned}
\rho &= \sum_k P(k)\rho_k \\
&= \sum_k P(k)\frac{\lambda k\Theta}{1+\lambda k\Theta} \\
&\approx \lambda\Theta \int_m^\infty dk \frac{2m^2}{k^3}\frac{k}{1+\lambda k\Theta} \\
&\simeq \lambda\Theta \int_m^\infty dk \frac{2m^2}{k^2} \\
&= 2\lambda\Theta m,
\end{aligned} \tag{8.15}$$

using (2.7).

Replacing Θ of (8.14),

$$\rho \sim (\lambda - \lambda_c)^{\frac{1}{1-a}}. \tag{8.16}$$

For this result, the fraction of infected nodes (ρ) follows a power-law in the vicinity of the critical point as $(\lambda - \lambda_c)^{\frac{1}{1-a}}$, for $\lambda > \lambda_c$, as represented in figure 8, differently from the traditional HMF (see equation (4.16)). This power-law behaviour makes the present model critical and closer to epidemiological models defined on regular lattices, where the interactions are local. This suggests that the modified HMF simulated a change in topology that was responsible for the finiteness of the epidemic threshold.

8.2 QMF modified simulation

Using the Gillespie model described in the previous chapter, we adjusted the assumption of connection of an infected node into the network, reducing the contagious according to a factor a that changes the degree of infected nodes, altering the topology of the network.

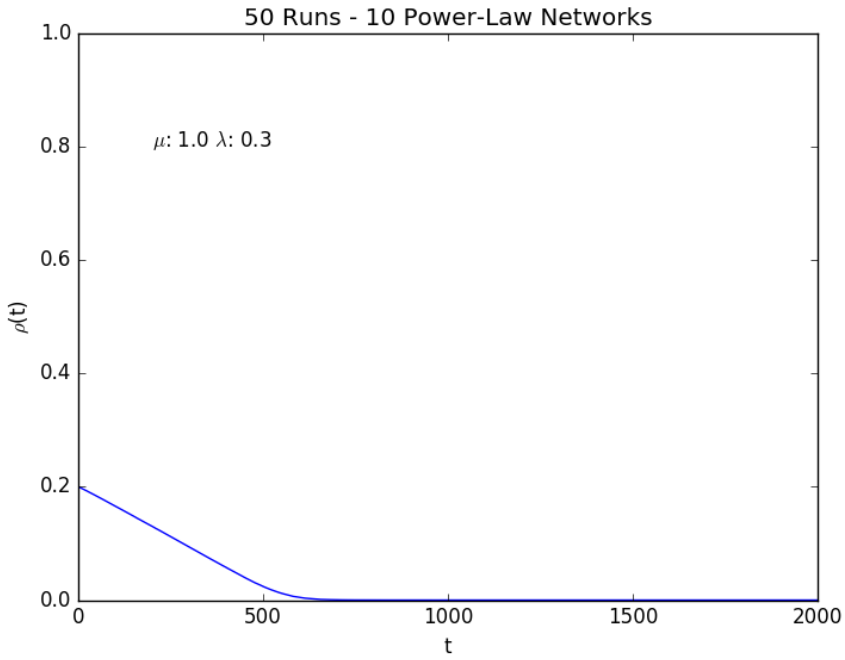
The changed algorithm is shown below (Algorithm 3):

Algorithm 3: Dynamics for an epidemic spreading modifying the probability of connection with infected nodes.
Input: a network
λ: infectious rate
μ: cure rate a: a value $0 < a <= 1$ that reduces the probability of connection between an infected node and a susceptible one.
Output: The dynamics for $\rho(t)$

```
begin
    Draw the infected positions at random given a infected rate;
    Calculate the healing rate M = μN_I
    for all infected nodes do
        Sum the edges of the infected positions through the adjacent matrix and find N_IS
    end
    Calculate the infection rate L = λN_IS;
    Calculate the total rate of transitions R = L + M and the probabilities m and l;
    for all time step do
        Choose a uniform random number between 0 and 1, rand
        if rand < m then
            one vertex is chosen and healed
        else
            one vertex is chosen according to its degree
            if rand < neighbors^a / neighbors + m then
                one neighbor is chosen
                if the neighbor is susceptible then
                    the neighbor gets infected
                end
            else
        end
    end
end
```

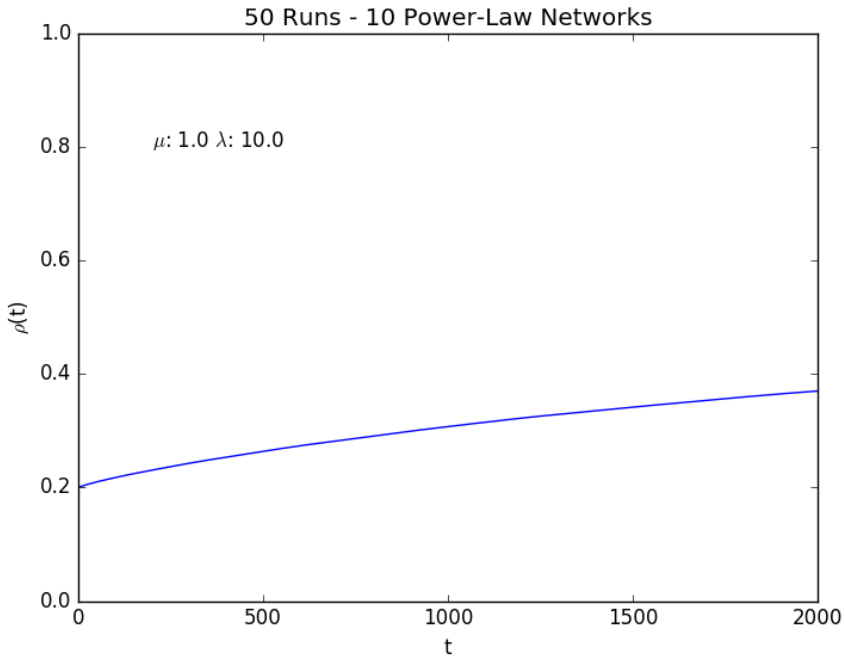
In Figure 22 we can note that the probability of infection goes to zero while in Figure 23, the probability of infection is different from zero. Therefore there is a positive epidemic threshold.

Figure 22 – Complex Network using a=$\frac{1}{2}$, $\alpha = 2.5$ and $\lambda = 0.3$.



Source: Martorello, Cristiane Dias de Souza, 2018

Figure 23 – Complex Network using a=$\frac{1}{2}$, $\alpha = 2.5$ and $\lambda = 10$.



Source: Martorello, Cristiane Dias de Souza, 2018

9 Conclusion

In this work, we reviewed the principal aspects of complex networks and the main epidemiological statistical models. Based on the SIS model, we investigated the behavior of the propagation of an infection in several kinds of networks, including regular lattices, complete graphs and complex networks with HMF and QMF approaches. Moreover, we implemented algorithms to generate a scale-free network with a given exponent, and applied the Optimized Gillespie algorithm to analyze the dynamics of an epidemic spreading.

We had a new result by studying a modification on the HMF model. We simulated an eventual change on the topology of the network, due to the states of the nodes. In the original HMF model, the probability of connecting to an infected vertex of degree k was proportional to $kP(k)$; we changed this factor to $k^a P(k)$, with $a < 1$. This modification aimed to mimic an effective decrease in the connections of the infected vertices. Although we did not changed the degree distribution, this alteration led to an appearance of a positive epidemic threshold in HMF, resulting in a great change compared with the traditional HMF that has a null critical point. This implies that the epidemic vanishes for lower values of infection rate compared to the critical value and the epidemic spreads for higher values of infection rate.

Besides, the fraction of infected nodes (ρ) follows a power-law distribution in the vicinity of the critical point as $(\lambda - \lambda_c)^{\frac{1}{1-a}}$, differently from the traditional HMF, where the fraction of infected individuals follows $2e^{-\frac{1}{\lambda m}}$.

We also implemented an adapted Optimized Gillespie algorithm and observed an identical appearance of a positive epidemic threshold in QMF modified simulation.

References

1 BARABÁSI, A.-L. *Network Science*. [S.l.]: Online on Barabasi.com, 2015.

2 BOGUÑÁ, M.; PASTOR-SATORRAS, R.; VESPIGNANI, A. Epidemic spreading in complex networks with degree correlations. *Statistical mechanics of complex networks*, Springer, p. 127–147, 2003.

3 PASTOR-SATORRAS, R.; VESPIGNANI, A. Epidemic spreading in scale-free networks. *Physical Review Letters*, APS, **86**, n. 14, p. 3200, 2001.

4 PASTOR-SATORRAS, R. et al. Epidemic processes in complex networks. *Reviews of Modern Physics*, APS, **87**, n. 3, p. 925, 2015.

5 ERDŐS, P.; RÉNYI, A. On random graphs. *Publicationes Mathematicae Debrecen*, **6**, p. 290–297, 1959.

6 DOROGOVTSEV, S. N.; GOLTSEV, A. V.; MENDES, J. F. F. Critical phenomena in complex networks. *Reviews of Modern Physics*, APS, v. 80, n. 4, p. 1275, 2008.

7 MILGRAM, S. The small world problem. *Phychology Today*, v. 2, p. 61–67, 1967.

8 NEWMAN, M. *Networks: An Introduction*. [S.l.]: Oxford University Press, 2010.

9 PASTOR-SATORRAS, R.; VESPIGNANI, A. Epidemic dynamics and endemic states in complex networks. *Physical Review E*, APS, **63**, n. 6, p. 066117, 2001.

10 BARABÁSI, A.-L.; ALBERT, R. Emergence of scaling in random networks. *Science*, American Association for the Advancement of Science, **286**, n. 5439, p. 509–512, 1999.

11 MURRAY, J. D. *Mathematical biology. I. An introduction*. [S.l.]: Springer-Verlag Berlin Heidelberg, 1993. I.

12 PASTOR-SATORRAS, R.; VESPIGNANI, A. Immunization of complex networks. *Physical Review E*, APS, v. 65, n. 3, p. 036104, 2002.

13 GOLTSEV, A. V. et al. Localization and spreading of diseases in complex networks. *Physical Review Letters*, APS, **109**, n. 12, p. 128702, 2012.

14 BOGUÑÁ, M.; CASTELLANO, C.; PASTOR-SATORRAS, R. Nature of the epidemic threshold for the susceptible-infected-susceptible dynamics in networks. *Physical Review Letters*, APS, **111**, n. 6, p. 068701, 2013.

15 LEE, H. K.; SHIM, P.-S.; NOH, J. D. Epidemic threshold of the susceptible-infected-susceptible model on complex networks. *Physical Review E*, APS, **87**, n. 6, p. 062812, 2013.

16 BOGUÑÁ, M.; CASTELHANO, C.; PASTOR-SATORRAS, R. Reply to Lee et al. comment on "nature of the epidemic threshold for the susceptible-infected-susceptible dynamics in networks"[arxiv: 1403.7913v1]. *arXiv preprint arXiv:1403.7913v1*, 2014.

17 FERREIRA, S. C.; CASTELLANO, C.; PASTOR-SATORRAS, R. Epidemic thresholds of the susceptible-infected-susceptible model on networks: A comparison of numerical and theoretical results. *Physical Review E*, APS, **86**, n. 4, p. 041125, 2012.

18 WANG, Y. et al. Epidemic spreading in real networks: An eigenvalue viewpoint. In: IEEE. *Reliable Distributed Systems, 2003. Proceedings. 22nd International Symposium on Reliable Distributed Systems*. [S.l.], 2003. p. 25–34.

19 MATA, A. S. d. *Epidemic Processes and Diffusion on Networks: Analytical and Computational approaches*. Tese (Doutorado) — Universidade Federal de Viçosa, 2015.

20 GOH, K.-I.; KAHNG, B.; KIM, D. Universal behavior of load distribution in scale-free networks. *Physical Review Letters*, APS, v. 87, n. 27, p. 278701, 2001.

21 CHUNG, F.; LU, L. The average distances in random graphs with given expected degrees. *Proceedings of the National Academy of Sciences*, National Acad Sciences, v. 99, n. 25, p. 15879–15882, 2002.

22 COTA, W.; FERREIRA, S. C. Optimized gillespie algorithms for the simulation of markovian epidemic processes on large and heterogeneous networks. *arXiv preprint arXiv:1704.01557*, 2017.

23 GILLESPIE, D. T. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *Journal of computational physics*, Elsevier, v. 22, n. 4, p. 403–434, 1976.

24 DOROGOVTSEV, S. N.; MENDES, J. F. F. *Evolution of networks*. [S.l.]: Taylor & Francis, 2002.

25 NEWMAN, M.; BARABÁSI, A.-L.; WATTS, D. J. *The structure and dynamics of networks*. [S.l.]: Princeton University Press, 2011.

26 WATTS, D. J.; STROGATZ, S. H. Collective dynamics of 'small-world' networks. *Nature*, Nature Publishing Group, **393**, n. 6684, p. 440–442, 1998.

27 LEE, H. K.; SHIM, P.-S.; NOH, J. D. Comment on" nature of the epidemic threshold for the susceptible-infected-susceptible dynamics in networks" [arxiv: 1309.5367v1]. *arXiv preprint arXiv:1309.5367v1*, 2013.

28 CASTELLANO, C.; PASTOR-SATORRAS, R. Thresholds for epidemic spreading in networks. *Physical Review Letters*, APS, v. 105, n. **21**, p. 218701, 2010.

Appendix

APPENDIX A – Method of Variation of Parameters

Consider the following equation:

$$\frac{d}{dt}y(t) = A(t)y(t) + B(t), \text{ where } y(t_0) = y_0. \tag{A.1}$$

Assume that $y(t) = u(t)v(t)$, without loss of generality, and replace it in equation (A.1):

$$\frac{d}{dt}[u(t)v(t)] = A(t)[u(t)v(t)] + B(t). \tag{A.2}$$

The derivative of $y(t)$ is

$$\frac{d}{dt}[u(t)v(t)] = \left[\frac{d}{dt}u(t)\right]v(t) + u(t)\left[\frac{d}{dt}v(t)\right], \tag{A.3}$$

and replacing it in (A.2) leads to

$$\left[\frac{d}{dt}u(t) - A(t)u(t)\right]v(t) + u(t)\frac{d}{dt}v(t) - B(t) = 0. \tag{A.4}$$

Imposing

$$\frac{d}{dt}u(t) - A(t)u(t) = 0, \tag{A.5}$$

with $u(t_0) = u_0$, the solution is

$$u(t) = u_0 e^{\int_{t_0}^{t} A(\epsilon)d\epsilon}, \tag{A.6}$$

and replacing (A.6) in (A.4) leads to

$$u_0 e^{\int_{t_0}^{t} A(\epsilon)d\epsilon}\frac{d}{dt}v(t) - B(t) = 0, \tag{A.7}$$

with $v(t_0) = v_0$. Isolating $\frac{d}{dt}v(t)$, we have

$$\frac{d}{dt}v(t) = \frac{B(t)}{u_0}e^{-\int_{t_0}^{t} A(\epsilon)d\epsilon}, \tag{A.8}$$

and, therefore:

$$v(t) = \frac{1}{u_0}\int_{t_0}^{t} B(\eta)e^{-\int_{t_0}^{\eta} A(\epsilon)d\epsilon}d\eta + \tilde{c}, \tag{A.9}$$

where $\tilde{c}$ is an integration constant.

Imposing the initial condition $v(t_0) = v_0$ in equation (A.9) leads to

$$v(t) = \frac{1}{u_0}\left[\int_{t_0}^{t} B(\eta) e^{-\int_{t_0}^{\eta} A(\epsilon)d\epsilon} d\eta\right] + v_0. \tag{A.10}$$

Then, from (A.6) and (A.10) in $y(t) = u(t)v(t)$:

$$y(t) = \left(u_0 e^{\int_{t_0}^{t} A(\epsilon)d\epsilon}\right)\left\{\frac{1}{u_0}\left[\int_{t_0}^{t} B(\eta) e^{-\int_{t_0}^{\eta} A(\epsilon)d\epsilon} d\eta\right] + v_0\right\}. \tag{A.11}$$

Thus, from $y_0 = y(t_0) = u(t_0)v(t_0) = u_0 v_0$,

$$y(t) = y_0 e^{\int_{t_0}^{t} A(\epsilon)d\epsilon} + \int_{t_0}^{t} B(\eta) e^{\int_{\eta}^{t} A(\epsilon)d\epsilon} d\eta. \tag{A.12}$$

APPENDIX B – Hyperbolic dynamical system

In order to analyze if the behaviour of an ordinary differential equation system has a stable or unstable equilibrium in hyperbolic case, it is enough to verify the behaviour near the equilibrium points. Consider the system

$$\begin{cases} \dfrac{dx_1}{dt} = & f_1(x_1, x_2, \ldots, x_n) \\ \vdots & \\ \dfrac{dx_n}{dt} = & f_n(x_1, x_2, \ldots, x_n), \end{cases} \tag{B.1}$$

the system (B.1) can be written as

$$\frac{d}{dt}\vec{x} = \vec{F}(\vec{x}), \tag{B.2}$$

where

$$\vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad \text{and} \quad \vec{F}(\vec{x}) = \begin{pmatrix} f_1(\vec{x}) \\ \vdots \\ f_n(\vec{x_n}) \end{pmatrix}.$$

The stationary solution $\vec{x^*} = \begin{pmatrix} x_1^* \\ \vdots \\ x_n^* \end{pmatrix}$ is the root of

$\frac{dx_1}{dt} = ... = \frac{dx_n}{dt} = 0$. Note that it implies that

$$f_1(x_1^*, ..., x_n^*) = ... = f_n(x_1^*, ..., x_n^*) = 0.$$

.

In the neighborhood of $\vec{x^*}$:

$$f_i(x_1, ..., x_n) = f_i(\vec{x^*}) + \underbrace{\vec{\nabla} f_i(\vec{x^*}) \cdot (\vec{x} - \vec{x^*})}_{\sum_{j=1}^{n} \frac{\partial f_i}{\partial x_j}(\vec{x^*})(x_j - x_j^*)} + \ldots.$$

Then

$$\frac{d}{dt}\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\vec{x^*}) & ... & \frac{\partial f_1}{\partial x_n}(\vec{x^*}) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1}(\vec{x^*}) & ... & \frac{\partial f_n}{\partial x_n}(\vec{x^*}) \end{pmatrix} \begin{pmatrix} x_1 - x_1^* \\ \vdots \\ x_n - x_n^* \end{pmatrix}$$

or

$$\frac{d}{dt}\vec{x} = J(\vec{x^*})(\vec{x} - \vec{x^*}), \tag{B.3}$$

where the Jacobian evaluated at $\vec{x^*}$ is

$$J(\vec{x^*}) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\vec{x^*}) & \dots & \frac{\partial f_1}{\partial x_n}(\vec{x^*}) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1}(\vec{x^*}) & \dots & \frac{\partial f_n}{\partial x_n}(\vec{x^*}) \end{pmatrix}.$$

Since we know that $\frac{d}{dt}\vec{x^*} = 0$, then (B.3) can be written as

$$\frac{d}{dt}(\vec{x} - \vec{x^*}) = J(\vec{x^*})(\vec{x} - \vec{x^*}). \tag{B.4}$$

The matrix J is assumed to be diagonalizable by

$$J = SDS^{-1}, \tag{B.5}$$

where D is the matrix of eigenvalues,

$$D = \begin{pmatrix} \Lambda_1 & & 0 \\ & \ddots & \\ 0 & & \Lambda_n \end{pmatrix},$$

and S is the matrix of the eigenvectors, where the coordinates of the i-th eigenvector is written in the i-th column as

$$S = \begin{pmatrix} | & \cdots & | \\ \vec{u_1} & \cdots & \vec{u_n} \\ | & \cdots & | \end{pmatrix}.$$

In hyperbolic case, Re $\Lambda_i \neq 0$ for $1 \leq i \leq n$.

Representing $\vec{x} - \vec{x^*}$ in the basis of the eigenvectors,

$$\vec{x} - \vec{x^*} = \sum_j c_j(t)\,\vec{u_j} = S\,\vec{c}, \tag{B.6}$$

where $\vec{c} = (c_1 \dots c_n)^T$, and substituting (B.6) into (B.4), we have

$$\frac{d}{dt}(S\,\vec{c}) = J(S\,\vec{c}),$$

and multiplying it by S^{-1}:

$$\frac{d}{dt}(\vec{c}) = (S^{-1}JS)\ \vec{c}\ .$$

By (B.5), $D = S^{-1}JS$, then

$$\frac{d}{dt}\vec{c} = D\vec{c} \rightarrow \begin{cases} \frac{d}{dt}c_1 = \Lambda_1 c_1 \rightarrow c_1 = e^{t\Lambda_1}c_1(0) \\ \vdots \\ \frac{d}{dt}c_n = \Lambda_n c_n \rightarrow c_n = e^{t\Lambda_n}c_n(0) \end{cases},$$

where $c_j(0)$ $(1 \leq j \leq n)$ is the initial condition. Then, since

$$c_j(t) = e^{t\Lambda_j}c_j(0),\ \text{for}\ 1 \leq j \leq n,$$

applying it in (B.6), we have

$$\vec{x} - \vec{x^*} = \sum_j e^{t\Lambda_j}c_j(0)\ \vec{u}_j\ . \tag{B.7}$$

If $\Lambda_j < 0$ for every j, then $\lim_{t\rightarrow\infty}\ \vec{x}\ (t) \rightarrow \vec{x^*}$. Then, by denoting by Λ_{max} the largest eigenvalue, we have

$$\Lambda_{max} < 0 = \lim_{t\rightarrow\infty}\ \vec{x}\ (t) = \vec{x^*}\ .$$

Then, the threshold is $\Lambda_{max} = 0$.

APPENDIX C – SIS Model Algorithm

Algorithm for a complete graph in Python language, for one run and with all vertices connected to each other.

```
import random
import pylab as pl

#all vertices connected

N=1000 #number of individuals

l = 0.3 #infectious rate lambda example
mu = 1 #cure rate mu
zeta = 1/((l+mu)*(N-1))

#start

I = 10 #infected individuals in the beginning
S = N-I #number of susceptible
t = 100*N #number of runs

I_list=[] #number of infected
P_list=[] #address of infected individuals

I_list.append(I)

#select at a random the infected vertex position
for i in range(I):
  i0=random.randint(0,N-1)
  while i0 in P_list: #seach if the vertex is already infected
    i0=random.randint(0,N-1)
  P_list.append(i0)

for j in range(t):
  v0=random.randint(0,N-1)
  if v0 in P_list: #seach if the vertex is infected
    pos = P_list.index(v0)
              if random.random()<=mu*zeta:
      tam = len(P_list) #number of infected
```

```
        I=I-1
        S=S+1
        #copy the last value in v0
        P_list[pos]=P_list[tam-1]

        #remove the last position
        P_list.remove(P_list[tam-1])
        if I==0:
          I_list.append(I) #last infected
          break

    else:
      for k in range(N-1):
        inf = random.random()

        #will get infected
        if inf <=l*zeta:
            #add in infected list
          P_list.append(v0)
          S=S-1
          I=I+1
          break

    I_list.append(I)

#normalization
for item in range(len(I_list)):
    I_list[item]=I_list[item]/float(N)

#plot the I_list
```

APPENDIX D – Scale-Free Network Algorithm

Based on the article (21), we developed an algorithm that generates a scale-free network given an expected degree exponent in Python.

```
import matplotlib.pyplot as plt
import networkx as nx
import math
import random
import numpy as np

G=nx.Graph()
W=[] #degree sequence
N=10000  #nodes
beta = float(input("power-law exponent: ")) #gets the exponent

d = float(input("average degree: "))
c=(beta - 2)/(beta-1)*d*N**(1/(beta-1))

x0=0
y0=0
W_sum =0.0
temp=0.0
j=0

#calculation of the weights

for i in range(N):
  W.append(c*(i+1)**(-(1/(beta-1))))
  G.add_node(i)
  W_sum=W[i]+W_sum

#including the links between nodes

for i in range(N):
  j=1
  while j<N:
    r = random.random()
    if r<(W[i]*W[j]/W_sum):
```

```
      G.add_edge(i, j)
    j=j+1

degree_sequence=sorted(nx.degree(G).values(),reverse=True)
dmax=max(degree_sequence)

x_list, y_list = [],[]
logx_list, logy_list = [],[]
regx_list, regy_list = [],[]
Sx,Sy, Sx2, Sxy = 0,0,0,0
dados_x, dados_y = [], []
cont=0

for i in range(len(degree_sequence)):
  if degree_sequence[i]<=dmax and degree_sequence[i]>0:
    if x_list.count(degree_sequence[i])==0:
      x_list.append(degree_sequence[i])
      y_list.append(degree_sequence.count(
          degree_sequence[i]))/N)
      logx_list.append(
          math.log(degree_sequence[i]))
      logy_list.append(
          math.log(degree_sequence.count(
          degree_sequence[i]))/N))
      cont = cont+1

#get values of minimum squares method

coefficients = np.polyfit(logx_list, logy_list, 1)
coef=coefficients[0]*(-1)

#plot the graph
```

APPENDIX E – Optimized Gillespie Algorithm

Based on the article (22), we developed an algorithm in Python that simulates the dynamic of a SIS model epidemics on a network.

```
import random
import math
import pylab as pl

N=1000 #number of nodes

n_degree = float(input("power-law exponent: "))

#given a scale free network G
#AM is the adjacent matrix
AM = nx.to_scipy_sparse_matrix(G)

#draw and show graph
pos = nx.spring_layout(G)

degree_sequence=sorted(nx.degree(G).values(),reverse=True)
dmax=max(degree_sequence)

lamb = 0.4 #infectious rate lambda example
mi = 1 #cure rate mi

Rod = int(input("number of rounds: ")) #number of rounds

#select at a random the infected vertex position
inf_rate =0.1

i=0

#start
t = 10000 #number of runs
I_sum = [0]*(t+1) #sum of the lists

for k in range(Rod):
  I_list=[] #number of infected
  P_list=[] #address of infected individuals
```

```
Ki_list=[]
pos_list=[]
I = inf_rate*N # infected individuals
S = N-I #number of susceptible
I_list.append(I)

#select at a random the infected vertex position
for i in range(int(I)):
  i0=random.randint(0,N-1)
  while i0 in P_list:
    i0=random.randint(0,N-1)
  P_list.append(i0)

#healing rate M
M=mi*len(P_list)

#infection rate W. Check if the neighbors are eligible.
W=0
for i in P_list:
  cont=0
  for j in range(0,N-1):
    cont+=AM[j,i]
    W = W+ AM[j,i]
    if AM[j,i]==1:
      pos_list.append(i)
  Ki_list.append(cont)

W=lamb*W
#total rate
R=M+W
#probability of being healed
m=M/R
#probability of being infected
w=W/R

for j in range(t):
  v0=random.uniform(0,1)
  if (v0 <= m): #an infected vertex is healed
    #number of infected vertices
```

```
    tam = len(P_list)
    vc = random.randint(0, tam-1)
    pos = P_list[vc]
    I=I-1
    S=S+1
    cont = pos_list.count(pos)
    for i in range(0,cont):
      pos_list.remove(pos)
    #remove the position item
    P_list.remove(pos)
    if I==0:
        #end of infected vertices
      I_list.append(I)
      break

  else:
    vis = random.randint(0, len(pos_list)-1)
    #infected vertex infects a susceptible vertex
    pos=pos_list[vis]
    cont = pos_list.count(pos)
    viz=random.randint(0,cont)
    cont_viz=0
    for j in range(0,N-1):
      if AM[j,pos]==1:
        cont_viz +=1
        if cont_viz==viz:
          pos_viz = P_list.count(j)
          #check if it is infected
          if pos_viz==0:
              #includes in infected list
          P_list.append(j)
          S=S-1
          I=I+1
          for j2 in range(0,N-1):
            if AM[j2,j]==1:
                pos_list.append(j)
          break
  I_list.append(I)
if len(I_sum)==0:
```

```
    I_sum=I_list
  else:
    for item in range(len(I_list)):
      I_sum[item]+=I_list[item]
for item in range(len(I_sum)):
  I_sum[item]=float(I_sum[item])/(Rod*N)

pl.figure()

#plot graph
```

APPENDIX F – Probability Θ in the vicinity of the epidemic threshold for the modified HMF

The starting point is the equation (8.11):

$$1 = \frac{2m^2\lambda}{\langle k^a \rangle}(\lambda\Theta)^{1-a}\int_{\lambda m\Theta}^{\infty} dy \frac{y^{a-2}}{1+y} \qquad (a<1). \tag{F.1}$$

In the neighborhood of the epidemic threshold λ_c, we have $|\Theta| \ll 1$. Renaming

$$\epsilon = \lambda m \Theta,$$

and integrating by parts the integral on the right-hand side, we have

$$\begin{aligned}\int_{\epsilon}^{\infty} dy \frac{y^{a-2}}{1+y} &= \frac{\epsilon^{a-1}}{(1-a)(1+\epsilon)} + \int_{\epsilon}^{\infty} dy \frac{1}{(1+y)^2}\frac{y^{a-1}}{a-1} \\ &= \frac{1}{1-a}\left[\frac{\epsilon^{a-1}}{1+\epsilon} + \int_0^{\epsilon} dy \frac{y^{a-1}}{(1+y)^2}\right] + \frac{1}{a-1}\int_0^{\infty} dy \frac{y^{a-1}}{(1+y)^2}.\end{aligned} \tag{F.2}$$

Calculating $\frac{\epsilon^{a-1}}{1+\epsilon}$ by Taylor series, we have

$$\begin{aligned}\frac{\epsilon^{a-1}}{1+\epsilon} &= \epsilon^{a-1}(1-\epsilon+\epsilon^2+O(\epsilon^3)) \\ &= \epsilon^{a-1} - \epsilon^a + O(\epsilon^{a+1}).\end{aligned} \tag{F.3}$$

The leading term of $\int_0^{\epsilon} dy \frac{y^{a-1}}{(1+y)^2}$ is

$$\begin{aligned}\int_0^{\epsilon} dy \frac{y^{a-1}}{(1+y)^2} &= \int_0^{\epsilon} dy y^{a-1}[1-2y+O(y^2)] \\ &= \frac{\epsilon^a}{a} + O(\epsilon^{a+1}).\end{aligned} \tag{F.4}$$

From (F.3) and (F.4), the first part of the right-hand side of equation (F.2) is

$$\begin{aligned}\frac{1}{1-a}\left[\frac{\epsilon^{a-1}}{1+\epsilon} + \int_0^{\epsilon} dy \frac{y^{a-1}}{(1+y)^2}\right] &= \frac{1}{1-a}\left[\epsilon^{a-1} - \epsilon^a + O(\epsilon^{a+1}) + \frac{\epsilon^a}{a} + O(\epsilon^{a+1})\right] \\ &= \frac{\epsilon^{a-1}}{1-a} + \frac{\epsilon^a}{a} + O(\epsilon^{a+1}).\end{aligned} \tag{F.5}$$

To solve the integral $\int_0^{\infty} dy \frac{y^{a-1}}{(1+y)^2}$, we use the Gamma Function, where, by definition, $\Gamma(z) = \int_0^{\infty} dt \quad e^{-t}t^{z-1}$. Using the relation

$$\frac{1}{c^z} = \frac{1}{\Gamma(z)}\int_0^{\infty} dt e^{-ct}t^{z-1},$$

with $z = 2$ and $c = 1 + y$, we have

$$\begin{aligned}
\int_0^\infty dy \frac{y^{a-1}}{(1+y)^2} &= \int_0^\infty dy \quad y^{a-1} \frac{1}{\Gamma(2)} \int_0^\infty dt e^{-(1+y)t} t^{2-1} \\
&= \int_0^\infty dt \quad e^{-t} t \int_0^\infty dy e^{-yt} y^{a-1} \qquad \text{(subtituting } yt = x\text{)} \\
&= \int_0^\infty dt \quad e^{-t} t \int_0^\infty dx \frac{1}{t} e^{-x} \left(\frac{x}{t}\right)^{a-1} \\
&= \int_0^\infty dt \quad e^{-t} t^{1-a} \int_0^\infty dx e^{-x} x^{a-1} \\
&= \Gamma(2-a)\Gamma(a).
\end{aligned} \tag{F.6}$$

Lastly, inserting all these parts together into (F.2), we have

$$\begin{aligned}
\int_\epsilon^\infty dy \frac{y^{a-2}}{1+y} &= \frac{\epsilon^{a-1}}{1-a} + \frac{\epsilon^a}{a} + O(\epsilon^{a+1}) + \frac{1}{a-1}\Gamma(2-a)\Gamma(a) \\
&= \Gamma(2-a)\Gamma(a-1) + \frac{\epsilon^{a-1}}{1-a} + \frac{\epsilon^a}{a} + O(\epsilon^{a+1}).
\end{aligned} \tag{F.7}$$

Substituting (F.7) in the (F.1), we have

$$\begin{aligned}
1 &= \frac{2m^2\lambda}{\langle k^a \rangle} (\lambda\Theta)^{1-a} \left[\Gamma(2-a)\Gamma(a-1) + \frac{\epsilon^{a-1}}{1-a} + \frac{\epsilon^a}{a} + O(\epsilon^{a+1}) \right] \\
&\simeq \frac{2m^2\lambda}{\langle k^a \rangle} (\lambda\Theta)^{1-a} \left[\Gamma(2-a)\Gamma(a-1) + \frac{(\lambda m\Theta)^{a-1}}{1-a} \right] \quad \text{(because} \quad |\lambda m\Theta \ll 1|\text{)} \\
&= \Gamma(2-a)\Gamma(a-1)\frac{2m^2\lambda}{2m^a}(2-a)(\lambda\Theta)^{1-a} + \frac{2m^{a+1}}{1-a}\frac{\lambda}{\langle k^a \rangle} \left(\text{note that } \frac{2m^{a+1}}{1-a} = \langle k^{a+1} \rangle \right) \\
&= \Gamma(2-a)\Gamma(a-1)(2-a)\lambda m^{2-a}(\lambda\Theta)^{1-a} + \lambda\frac{\langle k^{a+1} \rangle}{\langle k^a \rangle}.
\end{aligned} \tag{F.8}$$

Observe that $\Gamma(2-a)(2-a) = \Gamma(3-a)$, and $\frac{\langle k^{a+1} \rangle}{\langle k^a \rangle} = \frac{1}{\lambda_C}$ (see (8.8)), then

$$1 = \Gamma(3-a)\Gamma(a-1)\lambda m(\lambda m\Theta)^{1-a} + \frac{\lambda}{\lambda_c}, \tag{F.9}$$

which is (8.12) as desired.