UNIVERSITÀ DEGLI STUDI DI PADOVA

Dipartimento di Fisica e Astronomia "Galileo Galilei"

Master Degree in Physics (LM-17)

Final Dissertation

Modelling COVID-19 In a Network

Candidate	Thesis supervisor
Riccardo Milocco	Prof./Dr. Marco Baiesi

Academic Year 2020/2021

Contents

Abstract

The usual simplified description of epidemic dynamics predicts an exponential growth. This is due to the mean field character of the dynamical equations. However, a recent paper (Thurner S, Klimek P and Hanel R 2020 Proc. Nat. Acad. Sci. 117, 22684) [27] showed that in a network with fixed connectivity, the nodes become infected at a rate that increases linearly rather than exponentially. Experimental data for COVID-19 seem to validate this approach. In this thesis we plan to study this model by tuning its parameters. In particular, we monitor the effect induced by a significant presence of hubs in the network.

Motivations

Last Update: 22-08-2021"The COVID-19 pandemics has led to a dramatic loss of human life worldwide and presents an unprecedented challenge to public health, food systems and the world of work" [9]. In this scenario, monitoring and forecasting the diffusion of the virus is an essential tool for policymakers to handle the health-care resources and restrictions among individuals. To this purpose, many academics, e.g. UniPD against COVID-19[1], have been trying to tackle the COVID-19 spreading as well as its side-effects, which have to be considered on the "path to normality". This drastic shift with respect to the "path to herd immunity" comes from a recent study [14] which shed light on the fact that "herd immunity", due to delay on vaccinations, could probably not be reached for the end of 2021. On the other hand, due to few practical points [2], it is also probabile that the long-term prospects include COVID-19 becoming an endemic disease, much like influenza. Therefore, the recovery of the pre-COVID-19 situation could be slow even in the presence of effective vaccines.

So, an interesting path to cover is to describe the evolution of the COVID-19 pandemic, e.g. with the SIR model [21], on different network topologies in order to simulate the different containment policies, i.e. "lock-downs".

---

[1]https://web.unipd.it/covid19/en/

Introduction

1.1 Complexity

Since its beginning, Physics has been driven by a "Unification Principle"[1],which aims at describing the reality as a "reduction" and simplification of its constituents and interactions. Nevertheless, this optimistic gaze, is (amazingly) challenged by the complexity that emerges in the world we are living in. This complexity could have broader origins but also different spatial scales: from ants' nest to storms, from the markets to human/animal migrations, from the epidemics to ecosystems. Therefore, Nature is capable of obeying simple laws but, on the other hand, to straightforwardly generate complex phenomena. A paradigmatic example is the flocking of birds: the core of the flock assumes different shapes due to a collective interaction among the birds; while an alone bird would proceed in a defined direction. These facts were enough to establish the new branch of *Physics Of the Complex Systems* which ultimate goal is to describe the emergence of the "complex" phenomena from the simple law of physics, e.g. "gravity model" [3]. Having understood that *"More is different"* [1] and its ubiquity, it is natural to demand a formal definition of "Complexity". Nevertheless, quoting the Nobel Price Murray Gell-Mann [13], "a variety of different measures would be required to capture all our intuitive ideas about what is meant by complexity and by its opposite, simplicity" . In fact, from computational complexity to other information measures, all such quantities are context-dependent and subjective, since they depend on many aspects, such as the detail level (coarse graining) of the entity to describe. Thus, a rough but clear definition could be drawn for the Complex System Society site [23]:

Definition 1.1.1 (*"Complexity"*)**.** Complex systems are systems where the collective behavior of their parts entails emergence of properties that can hardly, if not at all, be inferred from properties of the parts.

In other words, the non-linear interactions among the parts of which a complex system is composed let emerge a behavior which could not be understood simply by applying the fundamental laws to its constituents.

---

[1]from Newton with the "gravity"-relationship between an apple and the Moon; passing to Einstein, whom the effects of gravity are equal to the ones of an accelerated rocket for; ending to the *Grand Unified Theory* as an unification of of general relativity with the particle physics

The first representation of this weaved system is by considering entities and connections among them. Thus, a powerful tool to describe it is the Network Science. In particular, the purpose of this thesis is to study the epidemic spreading of the COVID-19 over different topologies of the underlying social network of individuals. This phenomenon is regarded as a complex system through the analysis of the SIR model, which grasps the emergent features rather than the microscopic behavior of the disease.

1.1.1 Models Ecosystem

Infectious epidemics are responsible for a significant health and economic burden on society. New diseases appear frequently and old diseases persist. Therefore, many mathematical models have been used recently to guide policy makers to mitigate the impact of new infectious diseases (e.g. Polio, H1N1 influenza and Ebola) or established ones (such as HIV, cholera and seasonal influenza). On the other hand, with the phrase *All models are wrong; some models are useful* [8], George Box wanted to highlight that no model, no matter how complicated[2], is perfect. Thus, a way to evaluate a "successful" model has to be recovered.

Inside the "ecosystems" of models, there are the simple models which allow to analytically explain how the primary mechanisms influence important characteristics (e.g. epidemic threshold, epidemic size[3], how long it will last). To be more accurate, it is also possible to develop models that incorporate much more details about both the disease and individual-level interactions [15].
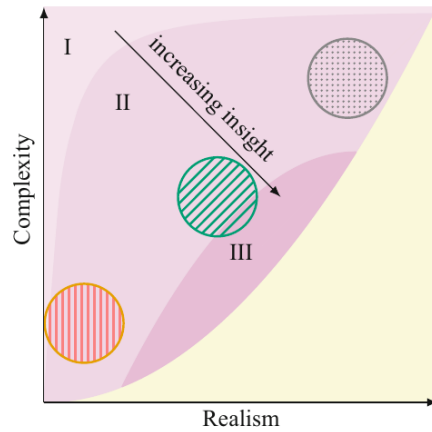
Figure 1.1: A caricature of the relation between model realism, complexity and the insight the model provides [15].

As suggested by the Figure 1.1[4] ([16]), models provide maximum insight where the right balance of realism (circle filled with dots, e.g. regarding epidemiology Graph-

---

[2]Instead, the complexity of a model could challenge its possibility to be predictive in many scenarios. A related topic is the "over-fitted model" in statistical learning or "Occam's razor" in Philosophy

[3]how large it will be

[4] *"Est modus in rebus"*

Neural-Network to predict R_t [11]) and simplicity (circle with vertical lines e.g. on epidemics the deterministic SIR model) is met[5] (central circle with oblique lines).

In this middle region, there could be found several works that use network analysis as the underlying structure for a phenomenon([27] [20] [21]); while other, more involved, that incorporate human mobility data in several hybrid models [25, 30] in order to predict its diffusion. For the sake of simplicity, the present thesis is going to deepen the paper of [27]. In particular, Thurner et al. [27], by focusing on the COVID-19 spreading in Austria and in the USA, are able to find a scenario, i.e. a combination of the SIR spreading parameters[6], such that the outbreak is (quasi-)linear for a value of the average number of contact D. This behavior of the SIR model is qualitatively comparable with the curves reported by the JHUniversity reported below. So, inspired by [27], the aim of this project is to analyze how the SIR model behaves by changing its "inner" epidemic parameters but also the topology of the underlying social network, i.e. the susceptible people.

1.2   A Technical Introduction

At 8 May 2020, none of the affected COVID-19 states have reached "herd immunity" but still they reached the "epidemic peak" due to the containment restrictions on social contacts. One paradigmatic example is the case of Austria, for which the total infected cases at the first peak were 0.16% of the total population, remarkably low with respect to the SARS-COV-2 "herd-immunity" level that are supposed to be at 0.8% ([31]) of the entire (susceptible) population. A graphical representation of the cases could be found in Figure 1.2, Figure 1.3 produced by Austria New Daily Infected.

(a) Austria until 8.5.2020   (b) USA until 8.5.2020
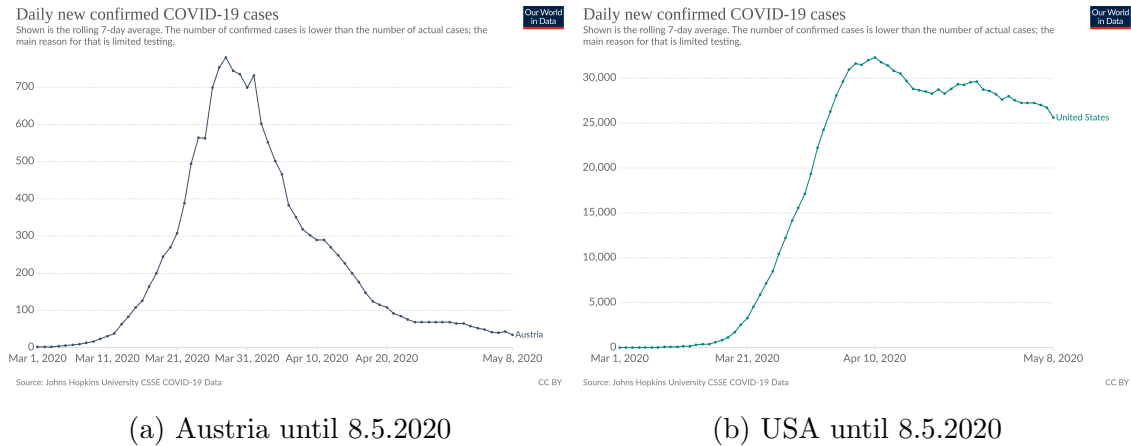
Figure 1.2: Daily Number Of Cases

The most striking observation is that the number of total cases (Figure 1.3), after early-stages when Non-Pharmaceutical-Interventions (NPIs), i.e. lockdown, were at work, exhibits a (quasi-)linear growth for an extended time interval in contrast with

---

[5]The equilibrium point may depend on the specific posed question

[6]the infection rate β, recover rate μ, the long-range attachment probability p

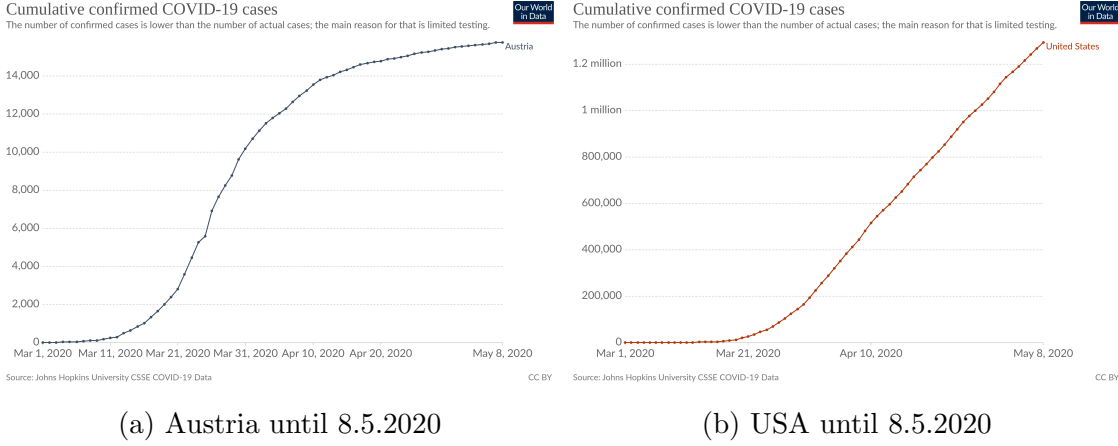(a) Austria until 8.5.2020   (b) USA until 8.5.2020

Figure 1.3: Cumulative Daily Number Of Cases

the "S-shaped" logistic curve predicted by the standard compartmental models. The extension of the linear regime depends on the onset of the measures; while for early stages, as it was the case for many countries [27] (8 May 2020), an exponential growth dominates the spreading of the disease. By taking care of the modification of the underlying social network structure, in this present thesis, we want to (qualitatively) recover the spreading trends for the states that have or have not applied the NPIs measures.

Specifically, an infection may occur for two reasons:

1. interaction between an infected and a susceptible person;

2. contact is "intense" (e.g. long, close,...) enough to lead to a disease transmission.

So, the rationale behind the social distancing is that it takes to a reduction of both of these factors. On the other hand, the analytically solvable SIR model assumes that, defined N, D as the total number of individuals in the population and the average number of contacts respectively, there is the same probability that an individual encounters an infected person ("well-mixed population") and all the nodes have the same number of neighbors ($N - 1$ or D). This approach allow to recover analytical results for an epidemic spreading but it annihilates the underlying network structure. Therefore, as claimed in [20], there is the need of studying how the network affects the spreading of a disease, but still no focus is put on the spreading below the epidemic threshold [27] as it is the case if nodes were separated through NPIs. In this thesis, the main goal is to grasp the relevant features of a complex social network in presence of NPIs and compare them with respect to a "well-mixed" population model. In detail, by changing the epidemic parameters, we let the SIR Model evolve in different network topologies such as standard models of the graph-theory ("Erdös-Rényi","Scale-Free", "Caveman Model") but also a new involved topology, i.e. a poissonian Small-World Network. As a benchmark of a "well-mixed" population, it has been used an "annealed" mean-field (see ??) which is a natural realization of a mean-field network with a fixed average number of contacts per node. Furthermore, $R_0 := \beta \cdot D/\mu$ is heuristically defined as the the average number of new infections

caused by an individual in a completely susceptible population[7] [15]. Moreover, it is well consolidated that for $R_0 > 1$ an epidemic is possible (not guaranteed); while for $R_0 < 1$ the disease will die out. Thus, these informations drive naturally to suppose that R_0 could also be a signal of the strength of an epidemic outbreak. Instead, this is not the case, since, as reported in the section ??, for equal R_0 values there could be different behavior of a disease. Thus, a new parameter has to be introduced that has to:

- grasp the different behavior of epidemics;

- depend only on the initial condition, i.e. network metrics (e.g. "long-range" parameter p) and disease parameters (β and μ).

A heuristics solution could be to characterize each spread via $R_{new} := \beta \cdot D / \mu * \mathbb{O}$ where $\mathbb{O} = \frac{D}{N-1}$, or equally $\mathbb{O} = D$. Thus, R_0 would still be considered as the secondary infected in a fully susceptible network; while R_{new} as the "strength" a spread has due network limitations.

Finally, as in [27], a "epidemic" order parameter[8] is computed for each spread to classify a light outbreak, i.e. an outbreak which infect less than the 30% of the entire population, with a stronger one. In this way, a first-order phase transition could be obtained by plotting, for sufficiently high p[9] the order parameter against the average degree of contacts.

On the top of the presentation of the work developed, it seems natural to deepen the area of the Network Science.

1.3 Network Science

Two aspects contribute to the popularity of Network Science: its universality nature of networks and the availability of large datasets[10] which map huge networks of the furthest kind, e.g. "food-webs" or the Science of Science. Its roots could be found in the Graph Theory, devoted to the analysis of the networks (or graphs), and Statistical Mechanics, committed to the formalization of the dynamical processes on network, such as self-organization and information theory. For our purpose, the Graph Theory allows to formalize a "networky" - complex system such as the social network of the individuals over which a disease spreads; while Statistical Mechanics to develop a model focusing on its emergent characteristic.

1.3.1 Graph Theory

A graph is a tuple (V,E) where the "V" stands for "vertexes" (1,2,3,4) while "E" for "edges" (the links between the vertexes). With the same notation, an edge

---

[7]The susceptible individuals are chosen according to "epidemic targets", e.g. only children for the measles

[8]Precisely (see ??), it is the standard deviation of the daily new infected

[9]p is the long-range rewiring probability.

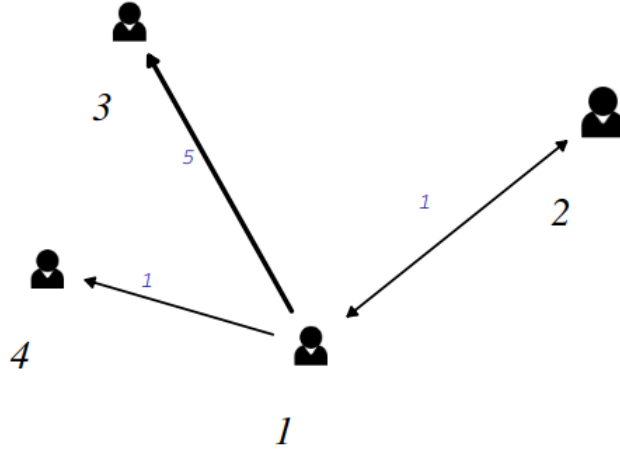[10]and the computational power to analyze them

Figure 1.4: Contacts interaction Network. A contact is defined as the droplets that reach one individual. Therefore, a symmetric edge could represent a chat among two persons; where directed edges would be a "one-way" interaction such as in a room with pre-existent droplets. The number above each edge are the weights, which could be the minutes spent in the previous quoted room

connecting the node[11] i to j is, often, identified by (i,j). Formally, the labels of the vertexes are the image of any bijective embedding from the nodes to a chosen set, e.g. integer numbers. The relationship among vertexes are accounted by the edges, which could be unsymmetrical, as for $(1,4)$, and with different strength of the connection, such as $(1,4)$ and $(1,3)$ which. If all the edges are symmetric, i.e. if (A,B) implies (B,A), the graph is said undirected, e.g. the friendship network; otherwise it is said directed, e.g. the phone contacts network. With the given definitions, it is now clear that roughly everything could be divided in vertexes connected by am arbitrary relation. The simplest example, is the *simple graph*, which is an unweighted, undirected graph containing no graph loops, i.e. (i,i) edges, or multiple edges such as multiples-(i,j).

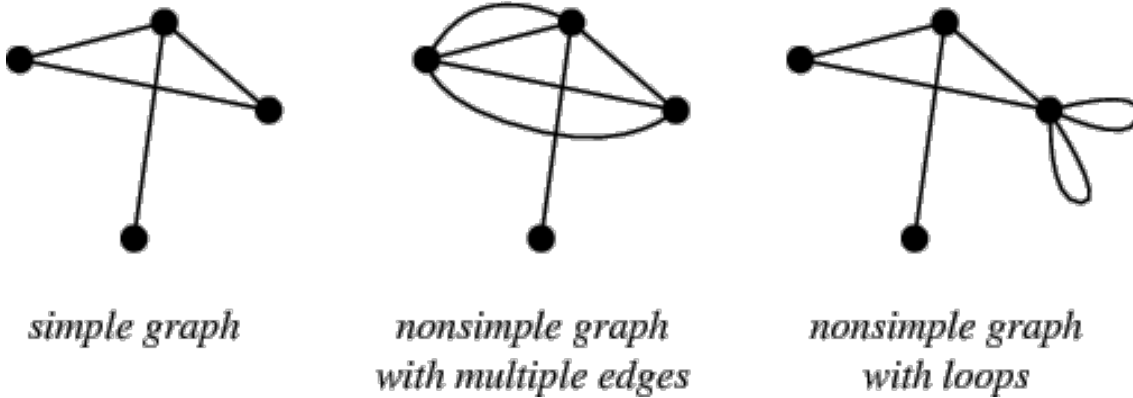*simple graph*   *nonsimple graph with multiple edges*   *nonsimple graph with loops*

Figure 1.5: Edges peculiarities. The labels are $1, 2, 3, 4$ assigned anti-clockwise starting from the bottom

---

[11]In the context of a social network, a vertex represents a person/individual

1.3.2 Adjacency Matrix

Typically for real networks, the number of involved nodes is big, e.g. the Google web-graph available at Standford University Network Dataset is of the order of the $800k$ nodes and $5milion$ edges. Thus, despite the clearness of Figure 1.4, the most suitable way of representing a graph is by grouping all its properties in a matrix called the adjacency matrix A. Starting from the basics, the (symmetric) adjacency matrix for the simple graph above is:

$$A_{ij} := \begin{cases} 1 & \text{if exist an unweighted edge among i and j} \\ 0 & \text{otherwise} \end{cases} \quad \Leftrightarrow \quad A = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix}$$

Hereby, $A \in M_4(\mathbb{N})$ holds the nature of a simple graph:

- the rows represent the direct connections between the $i-th$ row and the $j-th$ column; while, the columns the edge (j, i). Thus, there are only 4 nodes;
- unweighted: the value of the elements of A is binar;
- without loops: the diagonal is 0;
- without multi-edges: the value of the entries of A are strictly 0 or 1, account for the presence of, at maximum, one edge.

A more involved, adjacency matrix is the one of the rightest graph which present multiples loops[12]. Formally,

$$A = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & \mathbf{2} \end{bmatrix}$$

Thus, in Figure 1.4 is report a (simple) weighted and directed graph with 4 nodes connected by a "central"[13] node 1. In particular,

$$A = \begin{bmatrix} 0 & 1 & \mathbf{5} & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

where it is encoded the direct edges since A is not symmetric.

1.3.3 Degree

As quoted before, a (real) network could be very tangled. Thus, the degree distribution, i.e. the normalized number of vertexes with a certain degree, allows to check

---

[12]As far as multi-edges are concerned, the 2 would be off-diagonal

[13]The "centrality" measure of the node is another important quantity which is context-dependent. In this case, the "centrality" refers to a degree centrality for which the most central node is the one with the highest number of contacts

(a) Network measures [24]
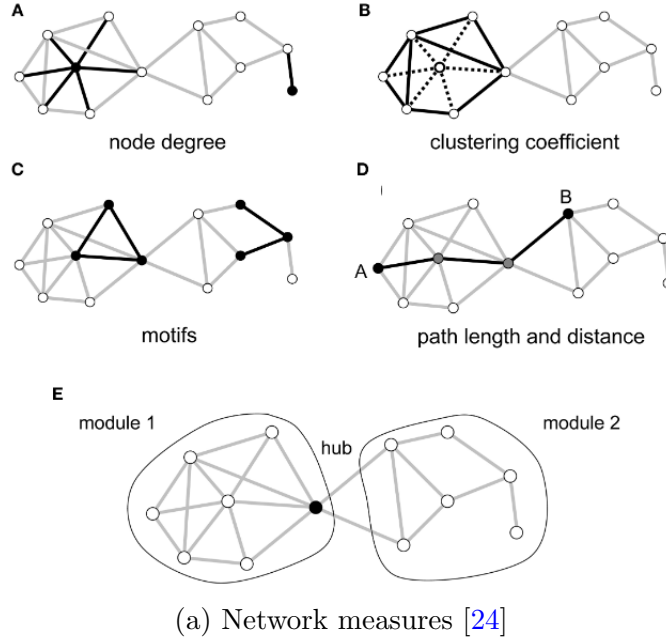
whether all the nodes have, or not, the same number of contacts even at a plain eye inspection.

Formally, the degree of a vertex, in a (undirected) graph, is the number of edges connected to it. We denote the degree of node i by k_i and, by mean of A,

$$k_i := \sum_{i=1}^{N} A_{ij},$$

where N is the number of total nodes. For directed graphs, as in Figure 1.4, a distinction between the entering and exiting edges to a node has to be defined. In particular, by a generalization of the previous k_i,

$$k_i^{in} := \sum_{j=1}^{N} A_{ij}, \quad k_j^{out} := \sum_{i=1}^{N} A_{ij}. \tag{1.1}$$

where $k_i^{in(out)}$ is the number of in-going (out-going) number of edges.

Therefore, by applying the Equation 1.1[14] to Figure 1.4, it is possible to recover the degree distribution of the vertexes as in the Figure 1.7. Thus, a possibile way to build a graph would be to draw the nodes degrees from a fixed distribution and, then, connecting the nodes according to their degree. In fact, this approach is going to be exploited in the further chapters by using a poissonian degree distribution and connecting the vertexes only to nearest neighbors via an ad-hoc algorithm.

---

[14]A practical check is to count the tails for the out-degree and arrows for the inner-degree. Ultimately, each sum has to equate the number of edges (a undirected edge has to be count doubled)
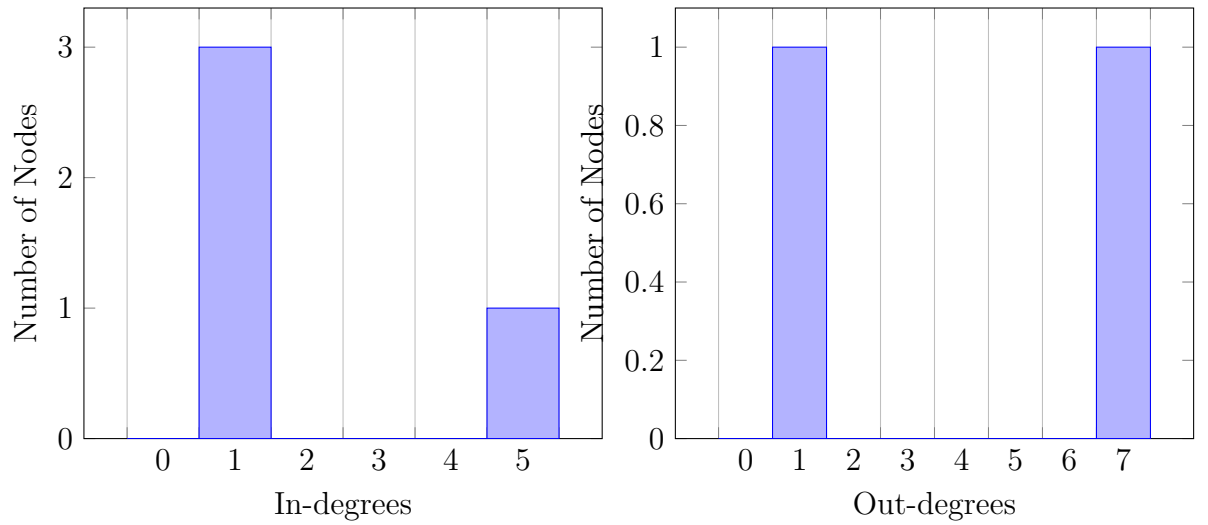
Figure 1.7: Degree Distribution of the Figure 1.4

1.4   Chapters Outline

In chapter 2, the models used in this thesis are going to be exposed: starting from a "regular" Watts-Strogatz and Caveman models, passing to the random ones (Erdös-Rényi, Watts-Strogatz, Poissonian-Small-World models); then, ending with the scale-free ones, such as the Barabási-Albert model.

In chapter 3, two kind of SIR models are presented: the mean-field SIR model and the network one, which combine the mean-field SIR with the network structure.

The Network Models

As it could be perceived, modeling a complex (social) phenomenon[1], as the epidemic spreading, in a network has two main tasks. The first is the identification of the nodes, which is a fairly simple task. The second is to pinpoint the emergent correlations in order to connect the vertexes and reproduce the phenomenon. The latter one is the real challenge in network theory.

Many networks are characterized by strong constraints on the node degree which forces a satisfying regularity to be present, e.g. the radial architecture of a spider net or the map of the main streets of Manhattan. Inspired by this, a first insight, the pathogen is left to act on a regular society, which individual contact are differing at most by one, e.g. "regular" Watts-Strogatz and Caveman model.

Most of the real networks, however, do not show a comforting orderliness. Rather, their inner structure seems to be characterized by a certain degree of randomness, e.g. the internet network [6]. Random network theory embraces this feature by constructing network models that are truly random, e.g. Erdös-Rényi. Meanwhile, they are not able to capture the presence of hubs, alias the "Scale-Free" property. Thus, new models as to be proposed to generate those highly connected degree nodes, such as the Watts-Strogatz and Barabási-Albert models. A review of these is reported in the following sections; alongside with the newly proposed "Poissonian Small World" network (see section 1.2, [27]).

2.1 Regular Networks

2.1.1 Lattice Graph

As briefly reported, the "Regular Networks" are characterized by a fixed number of neighbors per node $D = \langle k \rangle$. Thus, they are expected as the constituent graphs for the class of systems which constraints D to a certain value, e.g. solids in the physics of matter. Ultimately, by trying to study a new phenomenon, it is fundamental to test its behavior on a controlled environment as the one provided by the regular networks.

---

[1]The "social" nature has been only evoked for concreteness. Nevertheless, this in-depth presentation of models is interdisciplinary as Network Science itself.
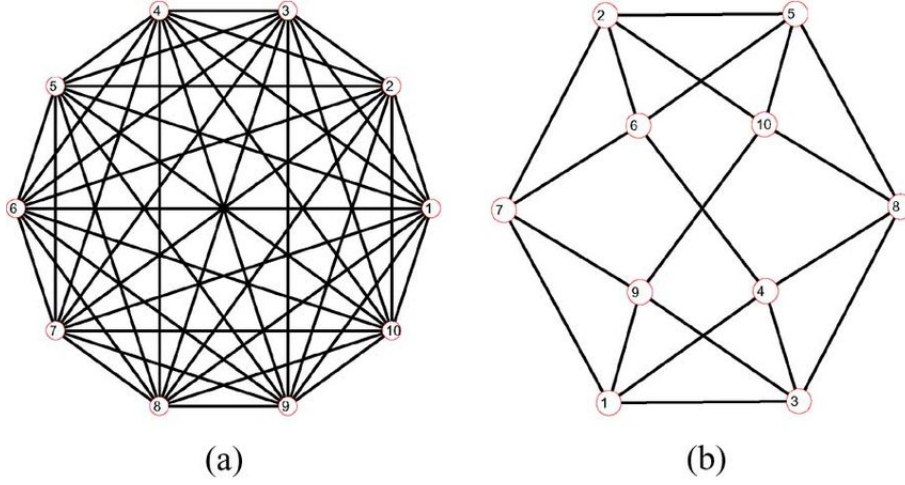
(a) In (a) - fully-connected graph; In (b) - $D = 4$ regular graph [29]

In subsection 2.2.5, there will be presented the "Watts-Strogatz" model which depending by an inner parameter p can interpolate between a regular graph for $p = 0$(Figure 2.1a) and a random network ($p = 1$). So, setting N as the number of nodes, chosen $p = 0$ and starting from a fully-connected network ((a) figure in Figure 2.1a), the average degree $D = N - 1$ is halved at each time-step. Thus, to preserve the secondary cases R_0, the infectivity of the pathogen β is doubled at each time-step.

2.1.2   Connected Caveman Graph
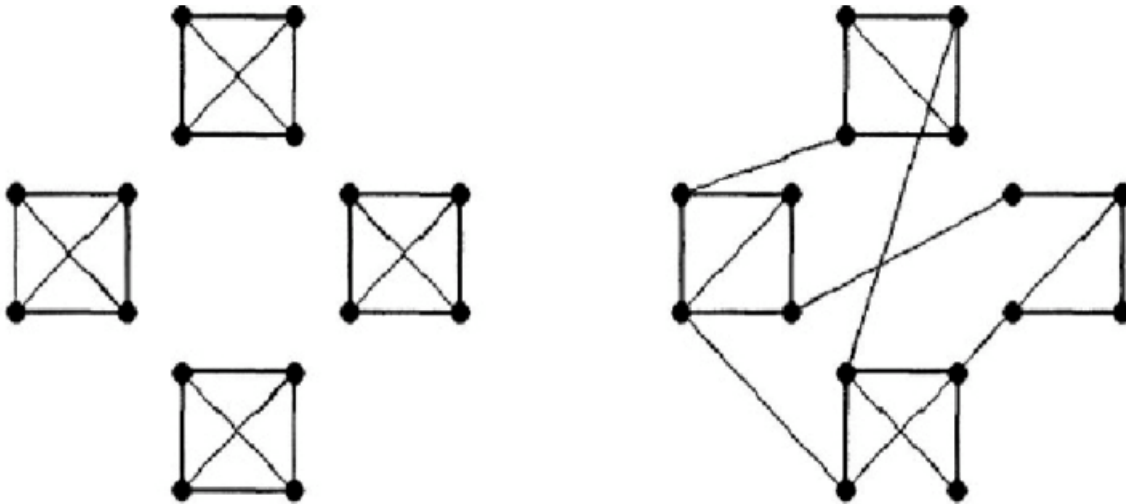
(a) Disconnected and Connected with rewiring Caveman Model [28]

By analogy with the NPIs measures which allowed only for the household interactions, it has been considered the "Caveman Model".

The rationale is to consider as if one node, of the previously modeled networks, would be a family composed by a fixed number of people. This idea drives to the "coarse-

graining" approach seen in the previous sections. In detail, the Caveman model enables to fix both the number of "caves"/families and the number of individual for each family.

In the present thesis, it is going to be developed an intermediate version among the Caveman models shown in Figure 2.2a. Indeed, connect the nearest caves of the left figure by creating a new edge on randomly choosing nodes within each cave. Then, introduce a "rewiring parameter" to enable "long-range" connections, resulting in "tunneling cuts" among distant nodes as in the right rewired graph. The ending graph is comparable to a Watts-Strogatz model, whose nodes are substituted with an entire cave.

A naive expectation is that the epidemic spread slower and more stochastically, since the nodes are poorly connected with the "outside" rather than in the family.

2.2 Random Networks

Not all the networks are characterized by hubs as WWW or the airports graphs. In fact, there could be constraints, such as cost-benefit problem, which limit the maximum degree of the bigger node. In those cases, random networks may be the best (thus, not conclusive) fit, e.g. national highways network, power grid of generators and switches, $\cdots$

2.2.1 The Erdös-Rényi Model
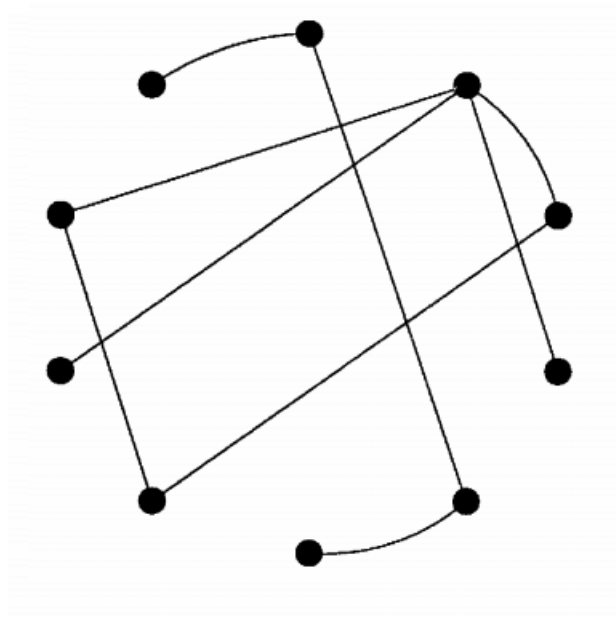
(a) Erdös-Rényi Model for $p = 0.2$ [7]

The Erdös-Rényi (ER - Figure 2.3a) model was the first attempt to formally describe a random network. In particular, these two mathematicians were the fathers of the random graph theory, obtained by merging graph theory with probability theory.

The primitive hypothesis, for a random wiring, is that the nodes are linked together with a probability p. Formally,

Definition 2.2.1 (*"Random Network"*). A random network $G(N, p)$ has N nodes where two nodes are connected with probability p.

Due to its random nature, this model may have different realizations even with N, p fixed parameters. Therefore, the number of links and the nodes degree distribution (see subsection 1.3.1) becomes vital informations to be considered. In particular, a specific realization could exhibit L links with a probability

$$p_L = p^L \cdot (1-p)^{\binom{N}{2}-L} \cdot \left(\begin{array}{c} \binom{N}{2} \\ L \end{array} \right). \tag{2.1}$$

The Equation 2.1 is the product of 3 terms: the probability to have L edges; the probability not to have the remaining ones knowing that the total possible node pairs are

$$\binom{N}{2} = \frac{N(N-1)}{2};$$

the total combinations in the choice of the L edges, since what counts is only to pick a number L of edges and not which ones. As expected, p_L could be interpreted as L successful results knowing that an edge is chosen with probability p and the total trials are $\frac{N(N-1)}{2}$ pairs. Thus, a Binomial distribution with the average number of links equal to

$$\langle L \rangle = \sum_{L=0}^{\frac{N(N-1)}{2}} p_L L = p \frac{N(N-1)}{2}$$

and, thus, the average per node contacts

$$\langle k \rangle = \frac{2\langle L \rangle}{N} = p(N-1).$$

The two equations above highlight that $\langle L \rangle$ and $\langle k \rangle$ could be obtained by multiplying p by the total number of available links and neighbors resp.[2]. Alongside, as the network becomes denser, i.e. increasing p, also $\langle L \rangle$ and $\langle k \rangle$ would be enhanced. Moreover, by multiplying the expression by q, it is possible to obtain $\sigma_{\langle x \rangle} = q \cdot \langle x \rangle$, where $x = L$ or k.

Similarly, the probability of having k number of contacts is

$$p_k = p^k \cdot (1-p)^{\binom{N-1}{2}-k} \cdot \left(\begin{array}{c} \binom{N-1}{2} \\ k \end{array} \right). \tag{2.2}$$

where $N-1$ are the total neighboring nodes differently from $\binom{N}{2}$, which were the total node pairs (cf. Equation 2.1). Another way to recover $\langle k \rangle$ and σ_k is by means of p_k.

---

[2]This is accordance with the formal definition of the "average"

Real network are sparse. In fact, by defining the "sparsity" of a network as the ratio $s := \langle k \rangle / N$, the internet-routers has $s \simeq 3.0 \cdot 10^{-5}$, since $N \simeq 200000$ while $\langle k \rangle \simeq 6$. Many others concrete example support this sparsity nature of real networks [6]. Therefore, in the $\langle k \rangle / N \ll 1$ limit[3], the Equation 2.2 becomes the poissonian distribution with average $\langle k \rangle$

$$p_k = e^{-\langle k \rangle} \frac{\langle k \rangle^k}{k!}. \tag{2.3}$$

Therefore, keeping in mind that the Equation 2.3 is only valid in the sparse regime,

- the explicit expressions of $\langle k \rangle \simeq p \cdot N$ and $\sigma_k = \langle k \rangle^{1/2}$ assume a simpler form for a poissonian distribution. Moreover, as the binomial case, $\langle k \rangle$ and σ_k increases if p increases, since the network is becoming denser and, so, more spread in the node degrees.

- it allows the same description for a pletora of networks with the same $\langle k \rangle$, since the network characteristics does not depend on its size N.

2.2.2   The "Small-World" Property

A captivating argument, network science allows to study is the "Small-World (SW) phenomenon" or "six degrees of separation". The two quoted names comes, respectively, from a theme developed by the Hungarian writer Fridgyes Karinthy in a play called "Chains" (1978) and an empirical experiment injected by the psychologist S. Milgram in 1967. In fact, the most famous version states that a person, e.g. Kevin Bacon, is linked to a random one, e.g. Andrea Scotti, through only few acquaintances: according to Kevin Bacon Number[4] only by 2 persons in common. So, the distance between two randomly chosen nodes in a graph is small. Random networks account for this phenomenon as we are going to deepen.

On average, the expected number of nodes N at a certain distance d is

$$N(d) \approx 1 + \langle k \rangle + \cdots + \langle k \rangle^d = \frac{\langle k \rangle^{d+1} - 1}{\langle k \rangle - 1},$$

where it has been taken advantage of the geometric series, after noting that "1" is the selected nodes, $\langle k \rangle$ are its neighbors, $\langle k \rangle^2 \approx \langle k \rangle(\langle k \rangle - 1)$ the next neighbors and so on. Thus, by considering that for a real network $\langle k \rangle \ll 1$ and that $N(d_{max}) \sim N$,

$$d_{max} \sim \frac{\ln(N)}{\ln(\langle k \rangle)}$$

On the other hand, the above equation is more suitable in describing $\langle k \rangle$, since d_{max} is dominated by few distant paths from the core giant component. Hence,

$$\langle d \rangle \sim \frac{\ln(N)}{\ln(\langle k \rangle)} \tag{2.4}$$

---

[3]At least for $\langle k \rangle \simeq 50$ and $N \simeq 10^3$
[4]https://oracleofbacon.org/movielinks.php

is the formulation of the Karinthy's "Small-World phenomenon" obtained by considering that most of the nodes are contained within $\langle d \rangle$. Indeed, for a real network, $\ln(N) \ll N$ drives to a "smaller" average path length compared to N, which is the typical scaling for a regular lattice[5] Moreover, the denominator of the Equation 2.4 implies that the denser it becomes the network, the closer the nodes aggregate. As a sociological example, the average distance of a random person, in the current century, would be derived by applying the Equation 2.4 by fixing $N \sim 7 \cdot 10^9$ and $\langle k \rangle \sim 10^3$ yielding $\langle k \rangle \sim 3.28$ which is a more reliable results than the "six degrees of separation" [6].

---

[5]The common notion of physical distance is based on a regular lattice which do not shows the "Small-World" effect. At first glance, this result could be perceived as an uncomfortable.

2.2.3   The Clustering Coefficient

As far as a social network is considered, it seems natural to introduce a measure of community membership (Figure 1.6a) which capture the possibility to join a local community, e.g. an association or a family. More generally, real networks are characterized by the clustering formation, which "simplify" the network by aggregating the nodes, called clusters, which present common characteristics. Proceeding as the "coarse graining" technique of Statistical Mechanics, it becomes appealing to take care of the emergent properties exhibited by the "coarse-grained" network. In the following section, it is introduced the degree of "clusterization" of a network.

More precisely, the local clustering coefficient C_i measures the chance of having a link among the neighbors of the $i-th$ node. To gain a rough intuition of its purpose, it worths to consider two opposite limits:

- $C_i = 0$ as there are no links among neighbors of i;

- $C_i = 1$ as there are fully connected neighbors of i.

Formally,

$$C_i := \frac{\text{average links of the i-th neigbors}}{\text{total links among neighbors}} = \frac{\langle L_i \rangle}{\frac{k_i(k_i-1)}{2}} = p = \frac{\langle k \rangle}{N} = \langle C \rangle \qquad (2.5)$$

where k_i and N are the amount of neighboring vertexes of i and the total nodes respectively,

$$\langle L_i \rangle = p \cdot \frac{k_i(k_i - 1)}{2} \quad \text{and} \quad \langle C \rangle := \frac{1}{N} \sum_{i=1}^{N} C_i.$$

The presence of a link among the two neighbors closes a triangle with the focal node i ("triadic closure"). Hence, C_i could also be interpreted as the normalized number of triangles centered in the selected node. In addition, $C_i = \langle C \rangle$ is independent by the $i - th$ node degree and, fixing the average degree of a network, it scales with $1/N$: for real networks $p = \langle k \rangle /N$ is small (real networks are sparse) and drives to a smaller C that the one observed. A model that naturally generates "triangle", would be a good candidate to enhance the clustering coefficient even for small p.

2.2.4   Problems in Random Networks

By looking at a real social network, e.g. at a cocktail party, there is the co-presence of highly connected nodes, called *"hubs"*, alongside with the less interacting ones. The ER graph does not account for the presence of high-degree nodes [6], since they are strongly damped by the factor term $1/k!$ in Equation 2.2. Indeed, real networks are sparse but have also strongly heterogeneous degrees, showing a high variance through node degrees.

Secondly, by looking at the evolution of the contact network, e.g. the one underlying the previous cocktail party, there could be identified 4 topological regimes starting from isolated individuals for $\langle k \rangle = 0$ and ending with a fully connected graph for

$\langle k \rangle = N - 1$, where N are the total invited persons. The ER model predicts the co-existence of a giant component, containing loops and cycles, and small components, which forms trees, for $1 < \langle k \rangle < \ln(N)$ - find a better description below. However, this is clearly at odds with reality since, as a concrete example, in a random electric power-grid some consumers should not get electric power [6]. Many network, as the actor network, are in the connected regime; most are in the supercritical one while still connected. In particular, these regimes scan the formation path of a fully connected network, passing trough a rapid emergence of a giant component. Schematically,

- Sub-critical Regime for $0 < \langle k \rangle < 1$ $(p < 1/N)$
- Critical point for $\langle k \rangle = 1$ $(p = 1/N)$
- Supercritical Regime for $\langle k \rangle > 1$ $(p > 1/N)$: Single giant component formation
- Connected Regime for $\langle k \rangle > \ln(N)$ $(p > \ln(N)/N)$

This is in an one-to-one relation with a First-Order-Phase Transition treated in Statistical Mechanics. The formation of an ordered giant component could be compared with ice formation: the passage from disordered water molecules to an ordered structure.

Thirdly, the real world networks exhibits a "Small-World" phenomenon not in agreement with the Equation 2.4. Hence, a new "Small-World" measure has to be proposed.

Fourthly, fixing N, the theoretical expectation is to grasp the independent nature of C with respect to k. Nevertheless, the results is unsatisfactory, since C decreases with the increasing of k for a pletora of the real networks. Furthermore, $C/\langle k \rangle (N)$, for high N, is higher that its theoretical expectation. Thus, a refined model is needed, to explain the high C even at small p (or high N).

2.2.5 Watts-Strogatz Model

To face the problem of the "high C at low edge probability (for now, redefined as p_{ER})", while preserving the "Small-World" property, the mathematicians Duncan Watts and Steven Strogatz proposed in the 1998 a hybrid model which interpolated between a regular lattice for $p = 0$ to a random network model for $p = 1$. The new borne model is, thus, called "Watts-Strogatz (WS) model". Practically, the idea is to start with a grid-like network, where all the nodes have $\langle k \rangle$ neighbors; then, with probability p, a link is chosen and one of the endpoints, e.g. the "rightmost", is replaced with a random vertex within the graph. The underlying degree distribution, hence, narrows the nodes degrees since it passes from an one-block histogram centered in $\langle k \rangle$ to a poissonian distribution. Thus, nodes have barely different neighbors degree and no hubs are allowed. A further constraint is that no multiple edges can connect the same pair of vertexes, i.e. no multi-edges. The presence of multiedges gives a correction of $1/N$ for big N, negligible in a real modeling scenario with big N.

The rationale of ER model is that by a little decrease of clustering (i.e. destroying
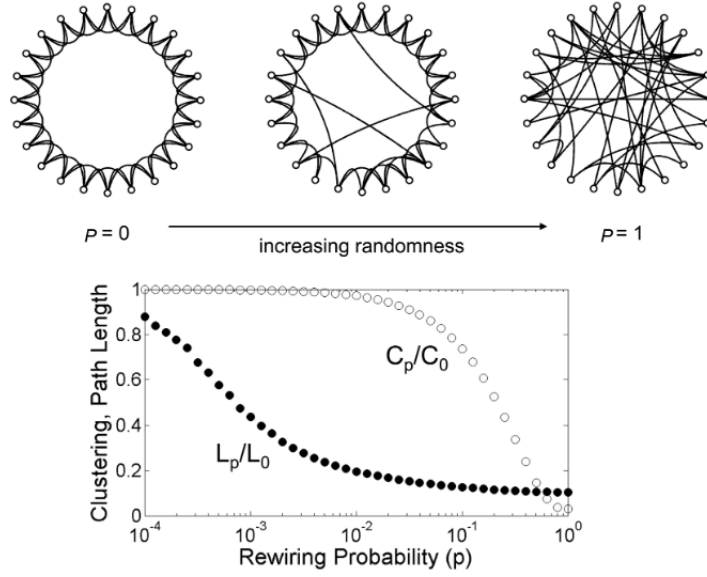
(a) In the top panel, the changing topology from lattice to random-like network; In the bottom panel, the behavior of average path length and clustering coefficient with p [24]

triangles) corresponds a significant reduction in the average path length $\langle d \rangle$. Other way around, by connecting nodes in a "long-range" fashion, the overall "triadic closures" are still relevant.

It worths to highlight that the present p, called in this context the "rewiring parameter", is the "long-range" rewiring probability, while the ER model p_{ER} is the probability of an edge formation either among neighbors ("short range" connection) or among distant nodes ("long range" connection). Thus, the "distant" connected vertexes are, on average, $p \cdot L$ where L is the total number of links. In particular, a regular lattice is characterized by a high clustering (neighboring nodes are fully connected) but a high average path length; while a random network, as seen before, exhibits manifestly the "Small-World" property alongside with a low clustering coefficient especially, since triangles are rare for small p. Therefore, the parameter p is introduce to fine tune the co-presence of the two kind of networks. As reported in [18], there is a range of rewiring probabilities $p \in (0.01, 0.1)$ in which $\langle d \rangle_p \approx \langle d \rangle_{p=0}$ and $C_p \approx C_{p=0}$ encoding the fact that the average path length and clustering coefficient of the WS model are, respectively, near the one of a random network and the one of a lattice.

Anyway, this model doesn't explain the presence of the hubs, since neglected by all the degree distributions, and the (inverse) dependence of $C(k)$ on the average network degree.

As a pedagogical spoiler, the (four) previous problems would be faced by modifying the WS model adding the emergent properties of the real networks, such as the "scale-free" property. Hence, an useful approach would be to start from the degree distribution rather than on imposing a fixed probability of wiring as for ER and WS models.

2.3 Configuration Model

To face the problems previously exposed, a practical starting point could be to reproduce the degree distribution by taking advantage of the configuration model [18]. Indeed, it is possible to built a network from an arbitrary degree sequence which either could be drawn by a distribution or by a real network of interest. Precisely,

Definition 2.3.1 (*"Degree Sequence"*)**.** The degree sequence is a list of N numbers $(k_0, \cdots, k_{N-1})$ where k_i is the degree of node i.

A degree sequence, thus, determines uniquely a degree distribution but the reverse is not true, since for a particular degree distribution there are $N!$ possibilities of labeling the nodes: $(1, 2, 3)$ has the same degree distribution of $(3, 2, 1)$. Additionally, a method to wire the nodes has to be chosen. Especially, the configuration model wires the nodes randomly; thus, building a random network. For these reasons, the generated graph is also called "random network with a pre-defined degree sequence".

As done before, a practical way of constructing a network clarifies many ideas. As a first step, consider to have a set of nodes with an assigned degree sequence. Pictorially, this could be depicted as a node with a number of dangling stubs as its degree. Thus, the number of isolated stubs have to be twice the sum of all the degrees. The network is, further, composed by the following iterative steps:

- A pair of stubs is selected at random;

- An edge is formed, by attaching together the selected stubs.

This tying technique is replicated until all the stubs are saturated and it is random by construct. Therefore, this method generates random networks with a fixed degree distribution.

Since the number of stubs equals the node degree, each node ends up having the desired degree. Thus, from the selected distribution, multiple (random) networks could be created; even, showing some peculiar properties, such as loops or multi-edges. Therefore, further constraints has to be imposed to obtain a specific topology, such as a connection only to "short-range" nodes.

The network is characterized by mainly two properties:

- the probability that a fixed node i may connect to another one j is

$$p_{ij} = \frac{k_i k_j}{2L - 1}. \tag{2.6}$$

  In fact, there are k_i stubs (or rays) trying to match the k_j ending of node j; while $2L - 1$ are the overall rays available for the chosen one which produces the $" - 1"$;

- the number of self-loops and multi-links decreases as N increases, since it is less probable to connect with a specific node as $p \sim 1/N$. Typically it is not needed to exclude them [19]. More importantly, arbitrarily rejecting self-loops or multi-links would result in a modification of the starting degree distribution, yielding the analytical calculations difficult.

The different realizations of the same degree distribution have a statistically objective. Indeed, let's suppose a certain feature, such as the presence of hubs, is independent on the other network characteristics rather than the degree distribution. By exploiting an ensemble of graphs, it would be possible to obtain an average and a standard deviation of that feature. Thus, by comparison with the real value, evaluate the unique dependence on the degree distribution.

Another method to generate a network from a fixed degree distribution is the "hidden parameter (η) method" which, in addition, do not form multi-edges and loops [6] and allows to tune the average degree $\langle k \rangle$ modifying η. On the other hand, to recover a randomized version from the degree sequence, such as of a fixed real network, the "degree-preserving randomization" is the best help the theory provides. In this way, by just swapping two endpoint nodes at a time, it is possible to build an networks ensemble with the same degree distribution. So, check whether a quantity depends only on the degree distribution itself or has a more involved structure. In the following sections, only the Configuration Model is going to be used.

2.3.1   Random Configuration Model: Poissonian Small-World Network

Social network are highly non trivial structure, by including a multilevel organization; weak ties between communities; and temporal aspects that suggest a degree of fluidity with stable social cores [27]. As seen in the section 1.2, however, the lock-down measures prevents the formation of highly connected nodes (hubs), expected in an unconstrained (social) scenario. Thus, a possible approach would be to use a Watts-Strogatz network, which has a degree sequence drawn from a poissonian distribution, regardless on the magnitude of $\langle k \rangle / N$ ratio (differently from subsection 2.2.5); while edges only connecting nearest neighboring nodes. This involved structure, called "Poissonian Small-World (PSW) Network" as in [27], enables to capture the heterogeneity in the number of contacts (node degree); the SW phenomenon and the clusterization of families alongside with their overlap. Moreover, the WS parameter p represents the fraction of individuals that are connected at "long-range" through leisure activities ($p = 0$ these are prohibited).

In order to built a PSW network, the starting point is by using the configuration model with a degree sequence drawn from a poissonian distribution. To simulate a closed population, the nodes are assumed to be put on a circle: the last and the first of nodes of the degree sequence are assumed as neighbors. Hence, both the "SW" property (see "The emergence of the Small-World property" page 15) and the heterogeneity on the node degrees could be embodied at the same time.

With low values of the poissonian parameter μ is possible to take into consideration the lock-down scenario where the average per node degree of contacts is the household size $D \sim 2.5$ [27]. The configuration model generates a random wiring of the links, producing "long-range" connections which spoil the locality of the community structures (families). Therefore, by an ad-hoc algorithm, the stubs are forced to attach to the nearest nodes in the circle according to their degree. In practice, the wiring algorithm iterates over few steps[6]:

- choose the lowest degree k_l node that has not already been saturated;

- link an even number of times $\lfloor k_l/2 \rfloor$ on the clockwise and anti-clockwise nodes, where $\lfloor * \rfloor$ is the floor function of $*$;

- if the degree of the selected node is odd, wire one last time ($k_l - \lfloor k_l/2 \rfloor = k_l \bmod (2) = 1$), e.g. anti-clockwise;

- delete, among the available nodes, both the chosen node and the target nodes which have saturated their degree.

More practically, fixed $[(1,0),(6,1),(0,2),(2,2),(4,2),(7,2),(8,2),(9,3),(5,3),(3,4)]$, i.e. a degree sequence composed by (node,degree) tuples. The edges (i,j) are created by connecting i to j and diminishing the degree of j at each step:

- node 1 is left alone as $deg(1) = 0$;

---

[6]Recall that the degree sequence is already fixed and drawn by a poissonian distribution

- $(6,7)$;

- $(0,2)$ and $(0,9)$

- $(2,0)$ (already present) and $(2,3)$;

- $(4,3)$ and $(4,5)$;

- $(7,6)$ (already present) and $(7,8)$;

- $(8,7)$ (already present) and $(8,9)$;

- $(9,3), (9,8)$ (already present) and $(9,5)$ - "tunneling" nodes ;

- $(5,4), (5,9)$ (already present) and $(5,3)$;

- node 3 is left with $deg(3) = 1$, but this issue reduces with sparsity $\langle k \rangle \ll 1$.

Hence, the algorithm preserves the poissonian degree distribution, while enhancing the local clustering of the nodes. However, it is possible to have "long-range" connections $((9,5))$, since the nearest available vertexes could be already saturated; therefore, resulting in a distant linking. This "tunneling" nodes are neither "super-spreaders" (or hubs of the infection), as they are forbidden by the poissonian degree distribution; nor "super-infectors", which would be characterized by an higher infectivity rate β.

2.4 Scale-Free Networks

2.4.1 Scale-Free Property
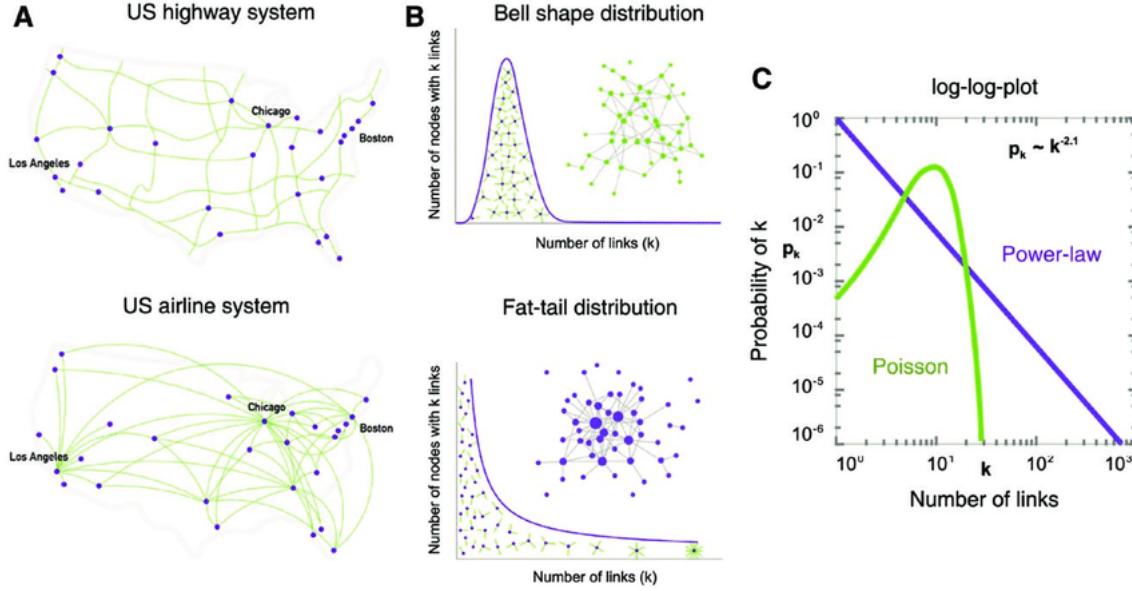
(a) On the left: the comparison of a random network(top panel) and a scale-free one(bottom panel; On the right: their degree distributions on log-log scale) [6]

Definition 2.4.1 (Scale-Free Network [4])**.** A scale-free network is a network whose degree distribution follows a power law.

By plotting the degree distribution of the real networks, due to "hubs", a better approximation is obtained via a power-law distributions (PLDs) of the kind[7] $p_k \sim k^{-\gamma}$ where γ is called the "degree exponent".

How a scale does (not) emerge

The PLDs are long-tailed functions; hence, describing a phenomenon that has not a reference scale. For example, the heights distribution of a sample population displays a peak at the average height, which is the reference scale for that population. Instead, PLDs do not exhibit a peak but allows also for outliers, e.g. 100ft tall individuals. More precisely, the variance, for a poissonian distribution, is $\sigma_k = \langle k \rangle^{1/2}$, limited as the network grows. So, the $\langle k \rangle$ scale is trustworthy.

At the same time, for a PLDs the variance diverges since [6]

$$\lim_{N \to \infty} \sigma_k^2 \approx k_{max}^{2-\gamma+1} = \infty. \tag{2.7}$$

since real networks have typically $\gamma \in (2,3)$ (WWW has $\gamma = 2.1$ [6]) and $k_{max} \sim N$.

So, a reference scale for PLDs is not naturally present as, extending the poissonian result, for the "Exponential Bounded Distributions" (EBDs) [6], i.e. when there is a exponential or faster decay for high k, such as binomial/poissonian/Gaussian distributions.

---

[7]Usually, reported in the log-log scale linear version: $\ln(p_k) \sim -\gamma \ln(k)$.

The Degree(s) of Hubs

In the following, it is developed a quantitative comparison between the maximum degree allowed by EBDs and by PLDs to gain a numerical perception of their difference. In detail, k_{max} (or the "natural cutoff") is expected to be, roughly speaking, the size of the bigger hub; while k_{min} approximates the smallest degree. So, for an exponential distribution [6],

$$k_{max} = k_{min} + \frac{\ln(N)}{\lambda}. \tag{2.8}$$

where $\lambda = 1$ for convenience. The same $\ln(N)$ dependence would be recovered of any other EBD.

On the other hand, applying the same rationale to a scale-free network,

$$k_{max} = k_{min} \, N^{\frac{1}{\gamma-1}}. \tag{2.9}$$

The results are somehow expected, since for:

- EBDs: $k_{max} \sim k_{min}$ since $\ln(N)$ strongly reduces contribution with N;

- PLs: there could be order of magnitude of difference between k_{max} and k_{min}.

Indeed, as a simple example, WWW forms a graph of on $N \sim 10^5$ documents with $\langle k \rangle \simeq 4.6$ and $\gamma = 2.1$ [6]. Knowing that $k_{min} = 1$ for EBDs, $k_{max} \sim 14$; while $k_{max} = 95.000$. This reinforces the insight that big hubs are naturally arising by increasing N only for PLDs.

Ultra Small-World

Hubs represent a one-hop bridge for many nodes, by centralizing the otherwise-distant nodes. Indeed, the average distance $\langle d \rangle \sim \ln(\ln(N))$ for $\gamma \in (2, 3)$, formally describing the so called "ultra small-world phenomenon". An heuristic example is the airport lines, which diminish the $\langle d \rangle$ among (far) cities with respect the national highways.

Surrounding the "Ultra Small-World", there are 3 other regimes [10]:

1. *Anomalous Regime ($\gamma = 2$)* producing a "wheel rim configuration", a central hub with spoke nodes. In this regime, $\langle d \rangle = constant$;

2. *Ultra-Small World Regime ($2 < \gamma < 3$)*: $\langle d \rangle \sim \ln(\ln(N))$ gives rise to the "ultra-small world phenomenon" as a slower growth in $\langle d \rangle$ than random network;

3. *Critical Point ($\gamma = 3$)* for which $\langle d \rangle \sim \frac{\ln(N)}{\ln(\ln(N))}$, marking the passage between the small-world and the ultra-small world;

4. *Small World ($\gamma > 3$)* random networks are recovered, resulting in $\langle d \rangle \sim \ln(N)$.

Thus, $\langle d \rangle$ is changing with γ, as the smaller it is, the shorter would be the average distance; but also with N, as it is the "ab ovo" hypothesis for having big hubs.

Growth and Preferential Attachment

Even with broad different origins, most real networks display two important features without which it is impossible for hubs to emerge [6]: growth and preferential attachment (see section A.2,section A.3).

In particular, growth is an active process driven by the sequential increasing of the size of the network N as per the addition of new nodes. Rather, a random network assume N to be fixed.

Furthermore, most real networks new nodes prefer to link to the more connected ones, a process called "preferential attachment". On the other hand, random networks wire randomly with other nodes.

Ça va sans dire , the derided model, developed in Figure 2.4.2, has to be dynamical and encode a "preferential" probability which changes every time ("event time") there is a modification of the topology.

Final Remarks on Scale-Free

It worths to recall that, in general, the hosted phenomenon (-a) is complex and affects the graph topology and, in turn, the degree distribution. Rebus sic stantibus , it is sufficient to establish a fair starting point, or EBDs or PLDs class, over which developing gradually ad-hoc features, rather then the meticulous fitting.

As a final remark, many interesting features of scale-free networks, from their robustness to anomalous spreading phenomena, are linked to the $\gamma \in (2,3)$; thus, due to the small average distance which strengthen the considered phenomenon. For a more detailed dissertation, see section A.1.
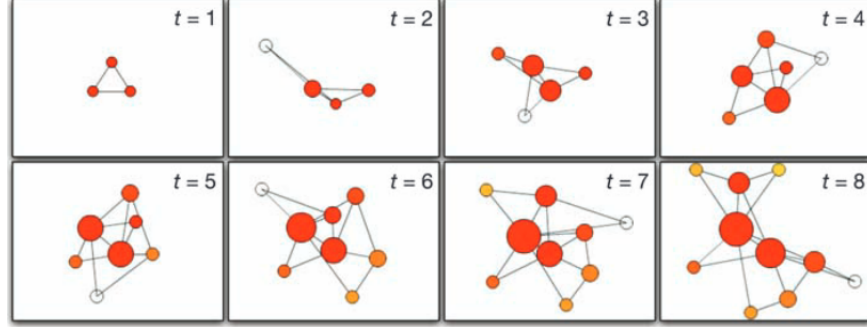
2.4.2   Barabási-Albert Model

Figure 2.6: Formation of a scale-free network [5]

To take advantage of the coexistence of both growth and preferential attachment in the real networks, it has been proposed a minimal model, the Barabási-Albert model [6]:

- start with m_0 nodes randomly connected and a degree at least equal to 1;

- **Growth:** at each event time add a new node with $m(\leq m_0)$ stubs which are linked to different nodes;

- **Preferential Attachment:** the probability to wire with a node i depends on the degree centrality, i.e. importance based on the degree:

$$\Pi(k_i) = \frac{k_i}{\sum_{j=1}^{N} k_j} \tag{2.10}$$

So, preferential attachment has a stochastic nature not a deterministic one. Nevertheless, it increases the appeal of the (already) high-degree nodes, alias "rich-gets-richer" phenomenon. As a result, while most of the nodes end up having few links, a bunch of nodes become highly connected; thus, performing the role of big hubs. In turn, this translates into a degree distribution which does not decay exponentially, i.e. a power law with $\gamma = 3$ [6].

Empirical studies on networks, such as actor and scientific collaboration networks, have found that the exponent at the numeratore of $\Pi(k_i) \sim k_i^\alpha$ could differ from 1. Indeed, there are 3 paradigmatic kinds of Preferential Attachments (PAs) by varying the exponent α: sublinear PAs $(0 < \alpha < 1)$, (linear[8]) PAs $(\alpha = 1)$, superlinear PAs $(\alpha > 1)$. Briefly, the sublinear PAs drives to an "stretched exponential distribution" (see Equation A.1); the linear to the standard scale-free network; while in the superlinear almost all nodes connects to few super-hubs (winner-takes-all phenomenon). Hence, only for $\alpha = 1$, a power-law distribution could be recovered, recalling that knowing the high-degree behavior is just a starting point to model the underlying phenomenon. In addition, after t event times, the network has gained t new nodes for a total of $N = m_0 + t$ nodes and $m_0 + mt$ links.

---

[8]Commonly, the here-called "linear" PAs is just called "preferential attachment" without further specification.

2.4.3 Linearized Chord Diagram

Having in mind the general purpose and results, a more precise description on the building process would clarify two point left open, i.e.

- the arrangement of the initial m_0 nodes;

- whether the next m links were added simultaneously or one at the time. This could drive to the formation of multi-links, i.e. conflicting edges, due to assumed independent nature of the new stubs.

These problem are solved with the idea of sequentially adding nodes to the graph(s) created at a previous times.

Fixing $m = 1$ for simplicity, start with the empty graph $G_m^t = G_1^0$. At each event time, choose an initial vertex v_i of $G_1^{(t-1)}$ and connect it to the new vertex v_t with probability p such that

$$p = \begin{cases} \frac{k_i}{2t-1} & if \quad i \in [1, t-1] \\ \frac{1}{2t-1} & if \quad i = t \end{cases}. \tag{2.11}$$

Rephrasing it, connect v_t with v_i with a probability of $\frac{k_i}{2t-1}$ if the nodes are different; otherwise, $\frac{1}{2t-1}$. This method does not forbid the formation of multi-edges and loops, but constraints their number to be negligible as the network grows. The present technique referred to $m = 1$ for an immediate perception. Whether $m > 1$ stubs are involved, add them one at the time, enabling their contribution to the hubs degree.

The Epidemiological Models

The most studied dynamical process on networks is the spreading of an agent. Rather than its common interpretation as a pathogen, this agent could also be considered as an information ("knowledge spreading"), an ask for collaboration, a mobile/computer virus, a business practice, a rumor or a meme, etc. These phenomena acts on different types of networks; are characterized by peculiar time-scales and follow various microscopic-transmission mechanisms. Despite all of these differences, they obey to a common pattern conveniently captured by analyzing their spreading on a network: the quoted graphs (cf. chapter 2) provides different insights on how links could be created and preserved. Thus, although it has been starting from the biological sector [20], it is used also in many other human-related fields, as the social sciences and digital security.

This thesis is devoted to the infective epidemiology: a paradigmatic model (SIR model) is let spread on different topologies simulating the COVID-19. In particular, infectious diseases are also called contagious, due to their transmission by a contacts with a secretion of an ill person, e.g. droplets. They account for 43% of the global burden of diseases alongside with other noninfectious diseases (for example obesity, gambling or smoking) which are equally affected by the graph of contacts [6].

The epidemiological frameworks, in general, find their roots in two fundamental hypotheses:

- Compartmentalization: Each individuals is classified according to a state, or compartment, allowed by the considered model. In particular, the simplest compartments are the following:

  - Susceptible (S): vulnerable people which have not been in contact with the disease;

  - Infectious (I): individuals, which are able the spread the pathogen to their neighbors;

  - Recovered (R): infected individuals which has are removed from the I-state. No difference among deaths or recovered has been settled[1]. Thus,

---

[1] In reality, this is not the case since death individuals are no more present in the society and leaves space for new contacts; while immunize nodes shields the more inner nodes, by arranging in a sun-like configuration (a susceptible core and spoken "recovered" nodes)

many texts use the fair word "Removed" as a definition of the "R".

Driven by some transition probabilities, individuals are able to move among compartments: $S \rightarrow I$ is a random event regulated by the transmission rate (β); while $I \rightarrow R$ by the recovery rate (μ). So, as for COVID-19, in the early stages a large fraction of the population was susceptible, e.g. 99%, called "naive population". Then, the pathogen drives some nodes into the infectious state; while, at the end, to the "Recovered" compartment.

For convenience, it is also possible to introduce many other compartments to allow for additional states, such as the "asymptomatic" which are able to spread the disease without exhibit symptoms, the first signal of the infection. The epidemiological frameworks are generally called as their constitutive compartments characterizing the model, e.g. "SIR" stands for the node path among the susceptible-infectious-recovered states. The division in compartments has to be considered at a macroscopic scale, since the aim is to improve the prediction of the number of infected; while not caring about the microscopic infection within the host body; all these details, indeed, are encoded in β and μ;

- Mixing Level: People live and mixes but has a fixed number of total individuals N, i.e. no death or birth are taken into account (see [19]). Hence, two possible hypothesis on how the individuals, restricted to their compartments, could get exposed to the disease are the:

    - *homogeneous mixing*: this hypothesis assumes that each individual has the same chance of interact with another one, e.g. an infectious. For this reason, it is also called the "fully-mixed", "mass-action" or "mean-field" approximation. Indeed, the specific degree of a single node is not taken into account; rather, the number of infectious-contacts per node is fixed. In turns, this could be seen as a "field" of mean infectious spikes which attack the susceptible. Only the "global" scale is present;

    - *heterogeneous (degree-based) mixing*: this approach hypothesizes that the individuals are grouped by their degree of contacts k and considered "statistically" equivalent within that class: this is a finer mean-field approximation and could be improved by considering an "individual-based" approximation.

It seems pretty natural that the "fully-mixed" hypothesis is false at some extent. In particular, it would be possible to observe that for small "epidemic strength" (R_0) the network structure deeply plays a role (see chapter 5), since individuals may transmit a pathogen only to (local) neighbors. Nevertheless, if a pathogen is "strong" enough the total number of infected are comparable to the one of a "mean-field" approach.

The model which may be composed with the quoted compartments are the SI, the SIS and the SIR (the case of study). In detail, the SI model accounts for the fact that it would be impossible to recover from the disease, e.g. herpes; the SIS model for diseases where it would be not possible to waning immunity, thus, returning

susceptible, e.g. the HPV, generic flu; the SIR where the last state is the gained immunity, e.g. the most strain of seasonal influenza or chickenpox.

Before gaining the details on the two mixing hypothesis, it is worth define the variables involved in the two approach. In particular, by denoting the total number of individuals N, $s(t) := S(t)/N$ is the number of susceptible (healthy) normalized with N; $i(t) := I(t)/N$ is the fraction of already infectious persons; while $r(t) := R(t)/N$ the recovered one. In this way, the "closed population" constraint translates into $s(t) + i(t) + r(t) = 1$ at any time t. Thus, at the initial time ($t = 0$), a seed of infected has to be injected to start the infection, e.g. $I(0) := 10$. So, $s(0) = 1 - i(0)$ since the recovered are $r(0) = 0$. Furthermore, let assume that the likelihood of transmitting the disease to a susceptible per unit time is β, i.e. the "transmission rate"; while the "recovery rate" is μ. Hence, it worths define the *basic reproduction number* $R_0 := \beta \langle k \rangle / \mu$, embodying the number of secondary infected in a naive population.

The final goal would be to find the evolving number of new daily infected at each time (a slight modification of $I(t)$). Moreover, since the disease-spreading is a stochastic process, it would be the average of the various scenarios–in our case 50.
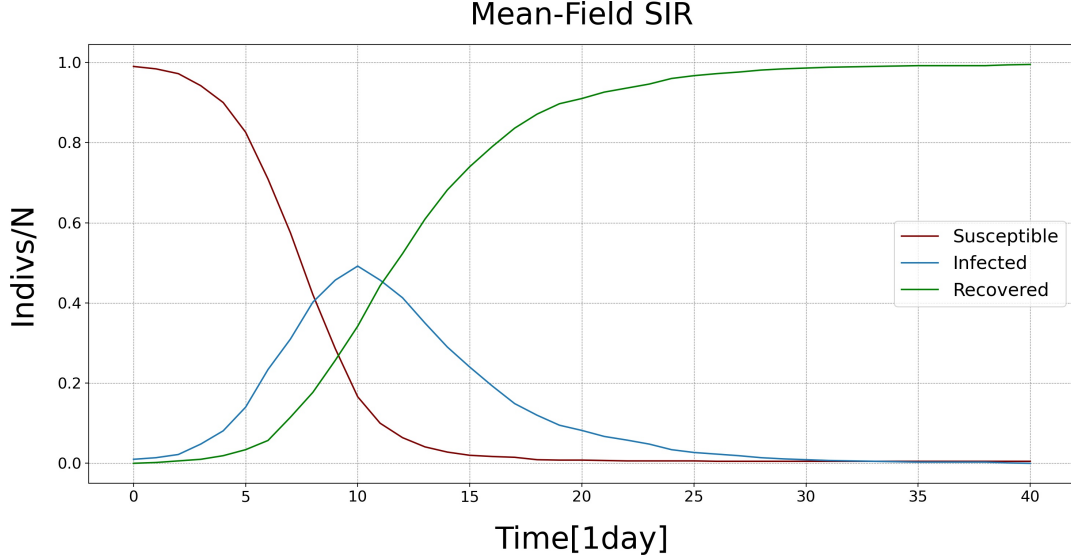
3.1   Homogeneous Mixing

Figure 3.1: Mean-Field SIR using a Complete Graph with $N = 10^3, \langle k \rangle = 999, \beta = 0.001, \mu = 0.2$

Within this class of hypotheses, people mingle and meet completely at random: an infected person can be in contact with $\frac{\langle k \rangle\, S(t)}{N}$ susceptible individuals. Taking into account that at a time t there are $I(t)$ infected which spread the disease at a rate β[2] and recovers at a rate μ, the SIR equations are the following [19]

$$\begin{cases} \frac{ds}{dt} = -\beta\langle k \rangle si \\[2ex] \frac{di}{dt} = \beta\langle k \rangle si - \mu i = \beta\langle k \rangle\,(1 - i - r)\,i - \mu i \\[2ex] \frac{dr}{dt} = +\mu i \end{cases} \qquad (3.1)$$

where it has been exploited the fact that the population is closed, i.e. no death/births are considered. This translates into $s + i + r = 1 \forall t$ or even $ds/dt + di/dt + dr/dt = 0 \forall t$. In addition, the removal of the time dependence since $s(t), i(t), r(t)$ is understood.

To solve Equation 3.1, by integrating "division" of the third equation and the first equation

$$s(r) = s_0 e^{-\beta\langle k \rangle r/\mu} \qquad (3.2)$$

where it has been required that $s(t = 0) = s_0$ arbitrarily. Then, substituting $s(r)$ into the second equation [19],

$$\frac{dr}{dt} = \mu\,(1 - r - s_0 e^{-\beta\langle k \rangle r/\mu}). \qquad (3.3)$$

---

[2]In many text books, e.g. [19], the $\beta\langle k \rangle$ total rate considered in the following is just β.

Now, solving this equation for r gives, in turn, the possibility to find s and, thus, i. In principle,

$$t = \frac{1}{\mu} \int_0^r \frac{du}{1 - u - s_0 e^{-\beta\langle k\rangle u/\mu}}.$$

The initial conditions can be chosen arbitrarily but the most suited is to fix a small number $c << N$ of infected individuals such that $s_0 = 1 - c/N \approx 1$, $i_0 := i(0) = c/N \approx 0$ and $r_0 \approx 0$.

Unfortunately, this equations have no closed form, but special insights cloud be reached by looking at the plot of the numerical solutions of Equation 3.1.

The case where $R_0 < 1$ represents the scenario when the pathogen will die out fast from the population, i.e. a limited spreading occurs.
Indeed, at early times $t \approx 0 \implies i(0) \sim 0, r(0) \sim 0$.
So, [6]

$$i(t \approx 0) \sim i_0\, e^{(\mu(R_0-1)t)}.$$

On the other hand, the SIR features are more involved if $R_0 > 1$:

1. $s(t)$ is monotonically decreasing; while $r(t)$ is always increasing;

2. $i(t)$ increases depending on the availability of susceptible: at early times the increasing is exponential; then, deflect and decreases on the way to the end of the epidemic, i.e. $i(t_{max}) = 0$;

3. $s(t_{max}) > 0$ since when $i \to 0$ there are no infected individuals left to infect the remaining susceptible. Formally, $r(\infty) := \lim_{t\to\infty} r(t)$ is the total size of the outbreak, i.e. the total number of individuals who ever catch the disease.

   Setting $dr/dt \overset{!}{=} 0$, $s_0 \simeq 1$ ("naive population") and using Equation 3.3, the solution of

   $$r = 1 - e^{-\beta\langle k\rangle r/\mu} \tag{3.4}$$

   is precisely $r(\infty)$. Therefore, whenever $r(\infty) \neq 1$ is a sign that "herd-immunity" is reached, alongside with $1 - r(\infty)$ susceptible left;

4. the characteristic time

   $$\tau = \frac{1}{\mu(R_0 - 1)}$$

   is the time needed to reach about the 36% of the population and it is the inverse of "disease velocity", which further depends both on the possibility to infect and the average nearest neighbors. With this new measure of epidemics, it would be possible to recover the previous result observing that there is a spread iff $\tau > 0$, i.e. $R_0 > 1$.

Interestingly, it would be possible to recognize $S \leftrightarrow r$ and $D := \langle k\rangle \leftrightarrow \beta\langle k\rangle/\mu$ (Equation 3.4) where the "left" quantities are describing the percolation phenomenon and the "right" ones the SIR spreading describes. In this way, the equation Equation 3.4 regulates the size S of the giant component of a Poisson random graph. In particular, the SIR infection of neighboring nodes could be seen as a bond percolation process of an agent, e.g. water, which propagates according to the same

probability distribution [19], [6]. This analogy allows to double the interpretations of the results of Equation 3.4. Indeed, in percolation theory, the giant component size S depends on the D average degree of links per node when $N \to \infty$.

So, the final outbreak size $r(\infty)$ depends on $\beta\langle k\rangle/\mu$ and the *basic reproduction number* is, conveniently, defined as $R_0 := \beta\langle k\rangle/\mu$.

Moreover, $r(\infty)$ goes continuously to $I(0)/N \sim 0$ as $1^+ \leftarrow R_0$; while it is constantly $I(0)/N$ for $R_0 \to 1^-$. Analogously of the appearance of the giant component, the point where (first-order) phase transition occurs, i.e. $R_{0c} := 1$, is called epidemic transition. Thus, R_{0c} separates the "epidemic" ($R_0 > R_{0c}$) from the "non-epidemic" ($R_0 < R_{0c}$) regime. The R_0 dependence of $r(\infty)$ has to be somehow expected since if an infection has less transmissibility than recovery ($R_0 < 1$), the pathogen would die out fast. On the way to R_{0c}, an increasing of the β/μ ratio would describe a less strong epidemic, which consecutively provide a minor $r(\infty)$.

To be precise, R_0 is the number of second infections in a naive population, i.e. at early times $t \approx 0$: $\beta/\mu = \beta\tau$ is the probability to infect over a time-interval of τ. Thus, R_0 is an estimate of the slope of the curve describing the infection at the start of the epidemic. Furthermore, it is a rough marker of an exponential growth[3]. (for $R_0 > 1$) or vanishing of the pathogen (for $R_0 < 1$) since $I(t) \approx R_0^{t/d}$. On the other hand, $R_0 = 1$ is not indicative of whether this slope is going to increase until $r = 1$ (pandemic scenario), decreases fast guiding into $r(\infty) << 1$, or both resulting in $r(\infty) < 1$. Indeed, for a more precise signal on the spreading behavior depending only by the initial conditions, higher-order features of the underlying network should be taken into account. A new reproduction number is going to be proposed in the next section.

---

[3]For SI model, $\lim_{\mu \to 0} R(t) = \infty$ since every infected could arbitrarily infect all the other nodes. In a real scenario, $R_0 (\leq N')$ is finite, as the size N' of a population is finite.
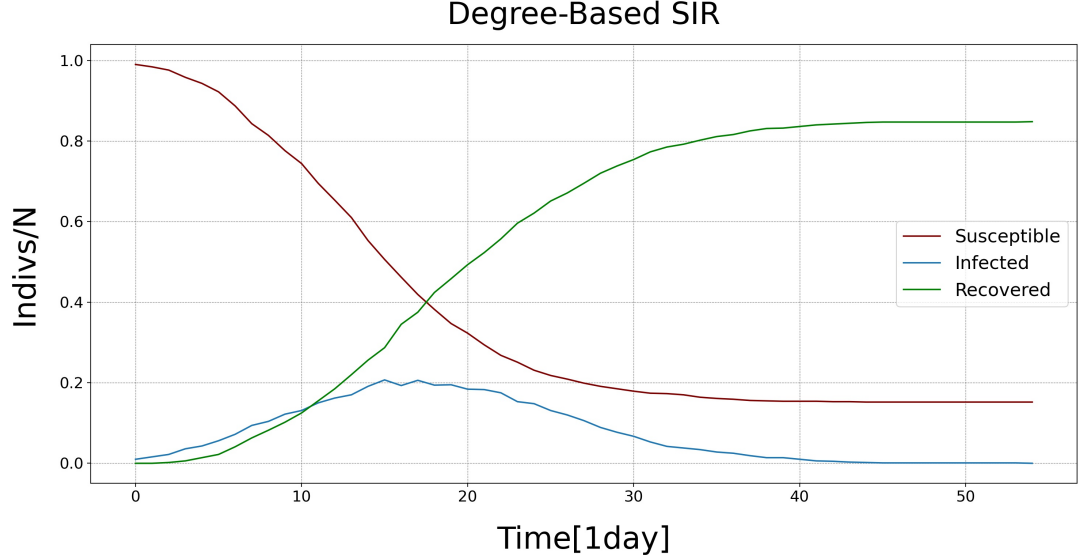
3.2 Heterogeneous Mixing

Figure 3.2: Degree-Based SIR using Erdös-Rényi Graph with $N = 10^3, p = 0.5, \langle k \rangle = 500, \beta = 0.001, \mu = 0.2$

As quoted before, within the well-mixed population it has been considered that every node could infect, at any time, a fixed amount of susceptible neighbors $\langle k \rangle$. This is not the case, e.g., for a pletora of real networks (Barabási-Albert or Linearized Chord Diagram models) as they are "complex"/"scale-free" in the sense that, for $\gamma \leq 3$, $\langle k^2 \rangle$ typically diverges. Thus, fixing the number of contacts to $\langle k \rangle$, it drives to underestimate the spreading, which could have worser effect than overestimating it.

An improvement would be to consider that if a pathogen spreads on a network, individuals with more links are more likely to be infected. Therefore, the nodes are further classified according to their degree and assumed statistically equivalent within their class. This method may, thus, called degree-based mean-field approach. Indeed, there would be $3 \cdot k_{max}$ classes, i.e. 3 compartments for each degree, but there are no individual differences among nodes with the same degree k.

The total fraction of infected $i(t) = \sum_{k=0}^{k_{max}} p_k i_k(t)$. Alongside with the other decomposition of susceptible and recovered, substituting them in the Equation 3.1 it would be possible to recover for each $k = 0, \cdots, k_{max}$

$$\begin{cases} \frac{ds_k}{dt} = -\beta k s_k \Theta_k \\\\ \frac{di_k}{dt} = \beta k s_k \Theta_k - \mu i_k = \beta k \left(1 - i_k - r_k\right) \Theta_k - \mu i_k \\\\ \frac{dr_k}{dt} = +\mu i_k \end{cases} \qquad (3.5)$$

where $\Theta_k(t) := i_0 \frac{\langle k \rangle - 1}{\langle k \rangle} e^{t/\tau}$ is the density function, representing the fraction of infected nodes surrounding a susceptible individual with degree k. It would be recov-

ered, and commented, shortly but note that, a priori, it can depend both on k and t.

The above equations depend, by construction, on k, the $\langle k \rangle$ dependence is no more present; while network structure is encoded into the $\Theta_k(t)$ factor, which in the homogeneous case was i. Before tackling the problem of solving them, it is necessary to explicitly recover the $\Theta_k(t)$ factor.

The Density Function

The density function, as reported before, provides the fraction of infected nodes neighboring a susceptible degree-k node. In fact, as we are going to prove, this estimation account for all the infectors connected with an arbitrary number of hops, since it has no upper-limit.

Formally, it has been assumed that the network lacks the degree correlation: nodes are connected at random, thus an hub could equally connect with another hub or a lower-degree node; so, called, a "neutral network". Hence, the probability of a vertex to point to a degree k' node is the excess degree

$$q_{k'} = \frac{k' p_{k'}}{\langle k \rangle}. \tag{3.6}$$

Indeed, the probability of choosing a degree-k' node is $p_{k'}$ but, since it has been k' possibilities with one stub to connect with it, this has been enhanced by a factor of k'. As is likely to be for all the probability, the denominator comes to match the $\sum_k q_k \overset{!}{=} 1$ constraint. The Equation 3.6 is local, random and linear in the "bias term" k. So, as reported in the subsection A.3.1, it is the starting point to build the (linear) preferential attachment. Moreover, Equation 3.6 grows with k' or $p_{k'}$, i.e. the "presence" of k' degree nodes.

For a disease to spread, it is required at least one of the spoke of the infected node is connected to the one target node: only $k' - 1$ links would be available. Thus,

$$\Theta_k = \frac{\sum_{k'} (k' - 1) p_{k'} k'}{\langle k \rangle} = \Theta \tag{3.7}$$

where the rightmost equation explicitly display the independence on k.

In order to capture its time-dependence, it would be take advantage of the specific model of spreading, e.g. SIR. For this purpose, substituting Equation 3.7 into Equation 3.5 for i_k, it could be obtained that [6]

$$\begin{cases} \frac{d\Theta}{dt} = \left(\beta \frac{\langle k^2 \rangle - \langle k \rangle}{\langle k \rangle} - \mu \right) \Theta \\ \\ \Theta(0) \overset{!}{=} i_0 \frac{\langle k \rangle - 1}{\langle k \rangle} \qquad \text{I.C.} \end{cases} \tag{3.8}$$

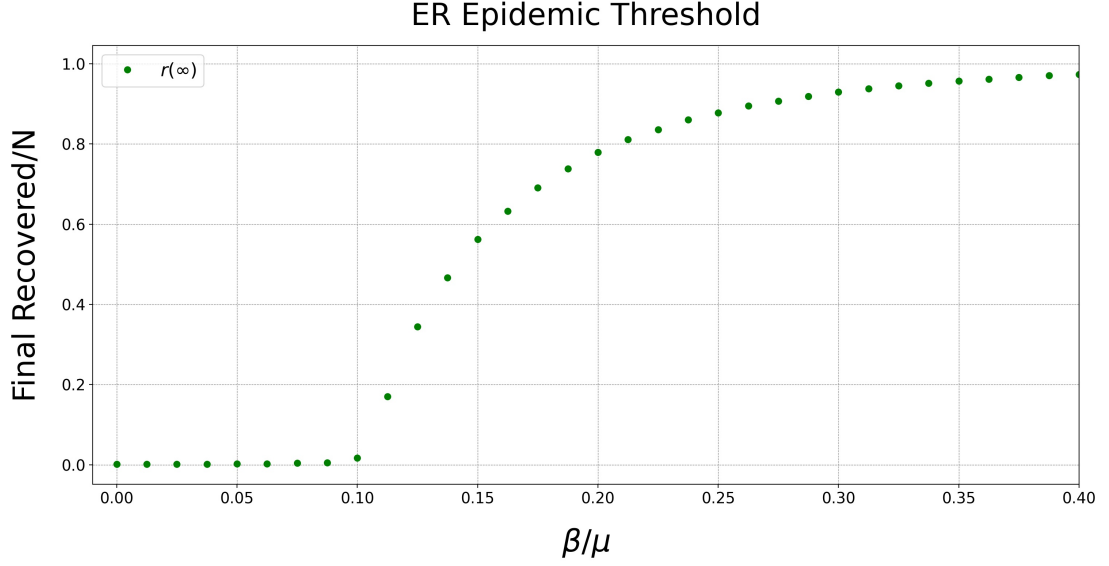which solving by Θ gives

$$\Theta(t) = C e^{t/\tau} \tag{3.9}$$

Figure 3.3: Epidemic Threshold in a ER Graph with $N = 10^4, \beta \in [0, 0.4], \langle k \rangle \sim 10, \mu = 0.8$. Note that the departure from zero happens near 0.1, due to finite size effects. Thus, $R_0 < 0.1 \cdot 10 = 1$ drives to a fast-decay of the infection

where

$$C := i_0 \frac{\langle k \rangle - 1}{\langle k \rangle} \qquad \text{and} \qquad \frac{1}{\tau} := \beta \left(\frac{\langle k^2 \rangle}{\langle k \rangle} - 1 \right) - \mu. \tag{3.10}$$

A global outbreak is possible if, and only if, $\tau > 0$, which implies (cf. Equation 3.9), that the number of infected nodes grows exponentially with time. This allows to recover the condition for a global outbreak

$$\lambda > \lambda_c := \left[\frac{\langle k^2 \rangle}{\langle k \rangle} - 1 \right]^{-1} \Rightarrow R_0 > R_{0c} := \frac{\langle k \rangle^2}{\langle k^2 \rangle - \langle k \rangle} \geq 0 \tag{3.11}$$

where it is specified the interdependence of the degree-heterogeneity ($\langle k^2 \rangle$), μ and $\langle k \rangle$ for a decay in the presence of the pathogen. For these reasons, λ is called the (biological) "spreading rate", since include all the biological factor of the disease, and λ_c the "epidemic threshold". Indeed, the number of the contagions is not linear with the enhancement of the spreading rate; rather, it is highly enhanced by passing λ_c.

By specializing Equation 3.10, it is possible to recover that, in the case of random networks, the final size of an outbreak displays a first-order-phase transition; while for scale-free network λ_c vanishes meaning that every epidemics could spread, regardless its biological weakness.

In particular, for a *poissonian network* $\langle k^2 \rangle = \langle k \rangle (\langle k \rangle + 1)$[4], resulting in the "epidemic threshold of a random network" [6]

$$\lambda_c = \frac{1}{\langle k \rangle} \Rightarrow R_{0c} = 1.$$

---

[4]As a rule of thumb, for a Poisson-distributed random variable k holds $\sigma_k^2 = \langle k^2 \rangle - \langle k \rangle^2 = <k>$. So, $\langle k^2 \rangle = \langle k \rangle^2 + \langle k \rangle$

Being $\langle k \rangle$ always finite, and the reference scale for a random network, λ_c play the role of a true threshold, which separates the epidemic regime to the non-endemic regime for as recovered in the previous section.

For a *scale-free network*, $\langle k^2 \rangle$ diverges for large networks ($N \to \infty$), since $\gamma < 3$ and the probability of having big hubs does not vanish. In this limit, $\lambda_c, \tau \to 0$ showing, respectevely, the fact that even low-transmissibility disease are successfully spreading in the population and are instantaneous. An heuristic explaination of this is that, as for the "network-distance" contraction, hubs are super-infectors of every disease. Thus, even if a pathogen has a weak load of infectiveness, when a hub is infected, it could be able to pass it to a larger number of nodes, enabling the pathogen to persist within the population. In the present context, hubs are responsible of a "time-length" contraction.

The exposed examples are at the extrema of the possible scenarios. For most networks in the middle, e.g. the ones described in chapter 2, the enanchement of $\langle k^2 \rangle > \langle k \rangle (\langle k \rangle + 1)$ implies a reduction both of the epidemic threshold and the characteristic time of the disease. Thus, in contrast with the "mean-field" predictions, it is possible to refine the calculations for a specific network.

To sum up, for $\lambda > \lambda_c$ the total epidemic outbreak size $i_\infty \neq 0$; while for $\lambda < \lambda_c$, $i_\infty = 0$, since the pathogen is not strong enough to reach an endemic state. Moreover, $\lambda_c = 0$ for a "typical"($\gamma < 3$) large scale-free network; while $\lambda_c \neq 0$ for the random networks.

About the SIR equations

No analytic solution exists for the equations regulating the SIR epidemic phenomenon. Furthermore, even the early times analysis exhibits some technical difficulties which demands for the numerical simulations of the ODEs of Equation 3.5.

Early Times

At early times ($t \approx 0 \implies r_k, s_k \approx 0$), according to Equation 3.5 and Equation 3.9, the fraction of infected is

$$\begin{cases} \frac{di_k}{dt} = -\mu i_k + i_0 \beta k \frac{\langle k \rangle - 1}{\langle k \rangle} e^{t/\tau} \\[2mm] i(0) \stackrel{!}{=} i_0 \qquad \text{I.C.} \end{cases} \tag{3.12}$$

whose, first equation, could be straightforwardly rewritten as

$$\frac{di_k}{dt} + \mu i_k - i_0 B e^{t/\tau} = 0 \tag{3.13}$$

where $i_0 B(k) := i_0 \beta k \frac{\langle k \rangle - 1}{\langle k \rangle}$.

Hence, the solution of the last equation assumes the form

$$i_k(t) = i_0 \left[C_1(k) e^{t/\tau} + C_2(k) e^{-\mu t} \right] \tag{3.14}$$

where it has been defined as

$$
\begin{cases}
C_1(k) := \frac{B(k)\tau}{1+\mu\tau} = \lambda \cdot \left[k \left(\frac{\langle k \rangle - 1}{\langle k \rangle} \right) \right] \cdot \frac{1}{\tau} \\[3mm]
C_2(k) := 1 - C_1(k)
\end{cases}
$$

where there have been isolated, in $C_1(k)$, the contribution from the pathogen ($\lambda := \frac{\beta}{\mu}$), the network and the interaction among the two (τ^{-1}). In all the previous equations, only the dependence was made explicit, since normally the epidemics parameter are fixed, while the degrees k vary.

In this way, the initial exponential-phase has two contribution: the first, which grows in time, due to an interplay among the disease and the underlying network; the second is diminishing the number of infected according to the recovery of infected individuals.

For real networks, the "degree exponent" $\gamma < 3$ allowing, in the large networks limit, for k and τ^{-1} to diverge. This, in turns, gives birth to a peculiar limit of $C_1(k)$:

$$
C_1(k) \propto \tau k \propto \frac{1}{\int k^2 p_k dk} k \to 0 \cdot \infty. \tag{3.15}
$$

where the last proportionality displays that k diverge alongside with $\langle k^2 \rangle$. More specific insights could be reached only by considering the underlying probability degree distribution p_k.

On the other hand, for many networks which find a good approximation in the random limit, i.e. when the individual contacts are limited by a cost-benefit constraint (e.g. power-grids), $\langle k^2 \rangle$ is limited and, therefore,

$$
i_k(t) \approx i(0) + (C - C_2(\langle k \rangle)\mu)\, t \tag{3.16}
$$

where $C := \frac{\beta(\langle k \rangle - 1)}{\tau}$ since $\langle k \rangle \approx k$ and $e^{at} \approx 1 + at$ as either $t \approx 0$ or $a \approx 0$. The dependence on k is, thus, lost: at the beginning of an epidemic all the nodes are statistically equivalent.

Ultimately, the degree-block approximation is itself a "mean-field" framework, since all the nodes, endowed with the same degree, are equivalent. However, this approach, although simplifying the calculations, it is not completely correct. Indeed, for the SIS model (with reinfection), the formal mathematical calculations drive to a vanishing epidemic threshold even for $\gamma > 3$ [6].

Methodology

As reported in section 1.2, the goal of this thesis is to study the "(quasi-)linear contagion" scenario, where the number of total infected individuals grows linearly in time rather than exponentially, as it could be the case for early times (cf. chapter 3). Indeed, as NPIs measures drives the social network to be sparse, there would no more be the presence of hubs but a "poorer" random model could capture the essence of the lockdown.

To this purpose, it has been sequentially selected a network model (chapter 2) over which the (degree-based) SIR simulates a disease spreading. Moreover, for every plot, it has been reported the mean-field SIR approximation to appreciate the effects of the underlying topologies (cf. chapter 3).

So, as in [27], initialize the N total individuals as susceptible and randomly pick $i(0) = 10$ vertexes as the starting infectors. Moreover, choose an "epidemic", i.e. fix β, μ, alongside with the average number of contacts $\langle k \rangle =: D$ and p[1] in the underlying network.
At every time-step t, find all the infected nodes and infect the susceptible neighbors with probability β (transmissibility or microscopic spreading rate). After the "infectious" phase, at the same time-step t, simulate their recovery with probability $\mu = 1/d$ (recovery rate), where d are the average days an individual stays infected.

> To implement the stochasticity nature of an epidemic, generate a random number between $q \in [0, 1]$ and infect the chose susceptible neighbors as $q \leq \beta$. Similarly, for the recovery transition with μ.

Proceed iterating the exposed infection chain, until the dynamic comes to a halt, i.e. $i(\infty) = 0$. In general, it could be find also that $s(\infty) \neq 0, r(\infty) \neq 0$.
In particular, for every t, collect the "new daily cases" $C(t)$ alongside with their (normalized[2]) "cumulative sum" $P(t) := \sum_{\bar{t} \leq t} C(t)/N$, as shown in the panels in section 1.2 reporting the real collected data. In details, $C(t) = N(i(t) - i(t-1) + r(t) - r(t-1))$: for this two states spreading $(I(t-1), S(t-1)) \to (R(t), S(t))$, $C(t) = 1$ as $i(t) - i(t-1) = 0$ but $r(t) - r(t-1) = 1$.

After having collected $C(t)$ and $P(t)$ also for the "Annealed Mean-Field Network

---

[1]The fraction of long-range contacts
[2]It has been chosen not to normalize the "new daily cases" since are $O(1)$ while $N \sim O(3)$

(AMFN)" (see chapter 4), it is produced a plot with the joined $C(t) and P(t)$ from a specific network, e.g. Watts-Strogatz, and the AMFN. Moreover, there would be a series of plots for each model, since there are some free parameters to be fixed: $\langle k \rangle, p, \beta, \mu$. Thanks to this method, it would be much clear to look for the departure of the network spread from the mean-field SIR.

Annealed Mean-Field Network

As a benchmark of the network results, it has been introduced the "annealed-mean-field" network. The definition of a AMFN will be given through the practical implementation of a SIR spreading over it, i.e. using a bottom-up approach. Indeed, the mean-field SIR model is obtain, at every time-step t, with the following scheme:

- select an individual I_1 over the pool of infected;

- chose $\langle k \rangle$ neighbors at random among all the nodes[3], e.g. I_2, I_3;

- try to infect them with probability β and to recover with probability μ.

With this procedure, whenever the target infected node, e.g. I_2, is chosen to pass the pathogen at a future time t', it is highly improbable to re-select its past infector I_1, since the probability goes as D/N, which typically is much less than one.

Moreover, the number of neighbors is fixed to $\langle k \rangle$, giving to this model the "annealed" characteristic name.

As promised, a AMFN is defined as a (directed) random network whose D links are randomly chosen by each node where the agent lays down. With this technique, the phenomenon is well mixed emulating a "mean-field" approximation.

The spreading on network resembles the mean-field model whenever either R_0 or the long-range connections $p \cdot N$ are high enough to allow the disease, reaching the epidemic state rapidly. Indeed, the comparison between mean-field and network spreading regards the fraction of total infected nodes and not how the nodes are infected: for the mean-field case the neighbors are chosen randomly along the circle; while they are fixed for the chosen network case. For this reason, there could an accordance in the measures considered.

(Quasi-)Linear Contagion

In order to analyze a linear behavior of $P(t)$, it worths recall that $R_{0c} := \beta D_c / \mu = 1$ is the critical basic reproduction number, which separates the exponential growth $(R_0 > 1)$ from the exponential decay $(R_0 < 1)$ of the number of infected. The critical degree $D_c = \frac{\mu}{\beta \langle k \rangle}$ and $C(t) \propto R^{t/d}$. So, as the epidemic is fixed, the above behavior are recovered, respectively, for $D > D_c$ and $D < D_c$. For $D = D_c$ [27], the number of infected will behave critically, i.e. $P(t) \propto t^\alpha$, where $\alpha \in (1, 2)$ for a simply

---

[3]A fair spreading should consider as available contacts all the nodes, i.e. the susceptible but also the already infected or recovered ones. Indeed, all nodes are possible count; while only the susceptible ones could be infected.

connected (no multiedges or loops) short-range network embedded in dimension 2. More in detail, $\alpha \sim 2$ for a network resembling a regular lattice in 2 dimensions, since $C(t) \propto t^2$ is proportional to the covered area of the infection front; while $\alpha \sim 1$ for a tree or a collection of 1-dimensional chains, as the front spreads linearly as a flame along a fuse, i.e. $C(t) \propto t$. In between these regimes, the specific α is influenced and shaped by non-pharmaceutical interventions (NPIs).

For a sub-exponential growth, it has to be expected that the network features, e.g. $\langle k^2 \rangle$, play a significant role in this context. In particular, for a fixed pathogen with $\lambda = \beta/\mu$, $R_0 \propto D$ and it would be natural to search for a critical average degree D_c ([27]), when it is informative (e.g. not in scale-free networks), such that, fixing p, β, μ,

- for $D > D_c$, the situation resemble a mean-field with $P(t) = R(t + d)$, where the shift in time is due to the fact that the average time of one infected to recover is d days. In addition, the number of infected assumes a typical bump on the immediate days after the starting, which is an evident mark of an exponential growth both in $C(t)$ and $P(t)$.

- for $D < D_c$, the sub-exponential behavior is recovered with the consequence that the final outbreak size is lower than the mean-field one.

A proper way to characterize this transition is per the definition of the order parameter [27] as the standard deviation (SD) of the new daily cases, after excluding all the days without new cases.
Formally,

$$O := SD(C(t)) = \begin{cases} 0 & \text{if } C(t) = constant \\ \neq 0 & \text{otherwise} \end{cases} \tag{4.1}$$

where SD is the "standard deviation".
As anticipated, a linear growth is possible only if the new daily cases are constants, since $P(t) \sim t^\alpha$. For the "S-shaped" curve, resembling a logistic curve, a bump in the daily infected occurs producing an enhancement of the SD. Hence, a SD deviating from zero signals the presence of a nonlinear increase of the cumulative positive cases, $P(t)$ [27]. Finally, D_c depends on the set of parameters p, β, μ [27] and, eventually N.

Also R_0 should be modified as it represents only the initial strength of the pathogen without considering the underlying topology, e.g. the "long-range" contacts pN. In other words, equal R_0s could drive to completely different scenarios (see the next chapter).

Analytical critical degree

In Statistical Mechanics, the dimensionality of a system is crucial to legitimate the mean-field approximation: the Ising Model is well modeled by a mean-field in a dimension $d \geq 4$ as it captures its phase transition.

In the following, D_c is derived for an ER model and a ring-shaped poissonian small-world network (see subsection 2.3.1) with D average degree, p rewiring probability, $D(1-p)$ short-range per node links, e.g. family, and Dp long-range connections, e.g. leisure activities. So, the probability that an infector will infect a susceptible person in the next d days is $\beta_d := 1 - (1-\beta)^d \sim \beta d = \beta/\mu$, i.e. the complementary probability of not having a "risky" interaction for d days.

Taking care of the dimensional effects, it worths to define the curvature of the front as $\mathcal{K} := \pm 1/\mathcal{R}_{front}$ where $\mathcal{R}_{front}$ is the radius of the fitting circle of the front: $\mathcal{K} > 0$ for concave case; while $\mathcal{K} < 0$ if the front surrounds an island of susceptible. For the poissonian small-world network, $\mathcal{K} = 0$, since the front is flat.

Moreover, R_0 is exact only for a naive population, i.e. at early times. In fact, if a node is later infected, it would find at least one neighbor infected, alias its infector. So, taking care of the epidemic spreading, $R := (D - n)rd$, where n is the average number of not-susceptible neighbors which, further, implies that $\nu := 1 - n/D$ is the fraction of susceptible. This latter quantity has to be interpreted as an general estimate of how infectious is a disease, i.e. for low infectivity, $\nu \sim 1$ which implies $D \lesssim D_c$. To be precise, for the "fuse-mode" ν represents only the fraction of susceptible in the neighborhood of the infection front; while for the ER model, n embodies the entire population.
Whenever, n is known,

$$D_c = n + \frac{1}{rd}. \tag{4.2}$$

Moving forward with the calculations, the network models have to be exploited. In particular, it is going to be discussed the ER model (see chapter 2) and the tree- or chain-like "fuse model" [27]. For low infection level, i.e. $\beta << 1$ or $\nu \sim 1$, the claim is that $n \sim 1$ for a random graph, while $n \sim (1 - \nu)D$ with $\nu \sim 1/2$ for a "fuse-model": $\nu \sim 1/2$ since the disease spreads on half of the tree, living the remaining nodes as susceptible.

Erdös-Rényi Graph

For a Erdös-Rényi random graph with "rewiring probability" p, the average degree of contacts is $D = Np$. So, the number of un-infectable nodes from an infected is $n \sim 1 + (D - 1)(1 - \nu)$, as 1 is the past infector of the fixed node summed with the remaining neighbors which are infected or recovered [26].
Inserting the obtained n in Equation 4.2,

$$D_c^{ER} = 1 + \frac{1}{\nu rd}. \tag{4.3}$$

which grasps the fact that for low infection levels $n \sim 1$. In addition, D_c^{ER} increases as ν drops: it has been needed more contacts to spread a disease if there are less susceptible available.
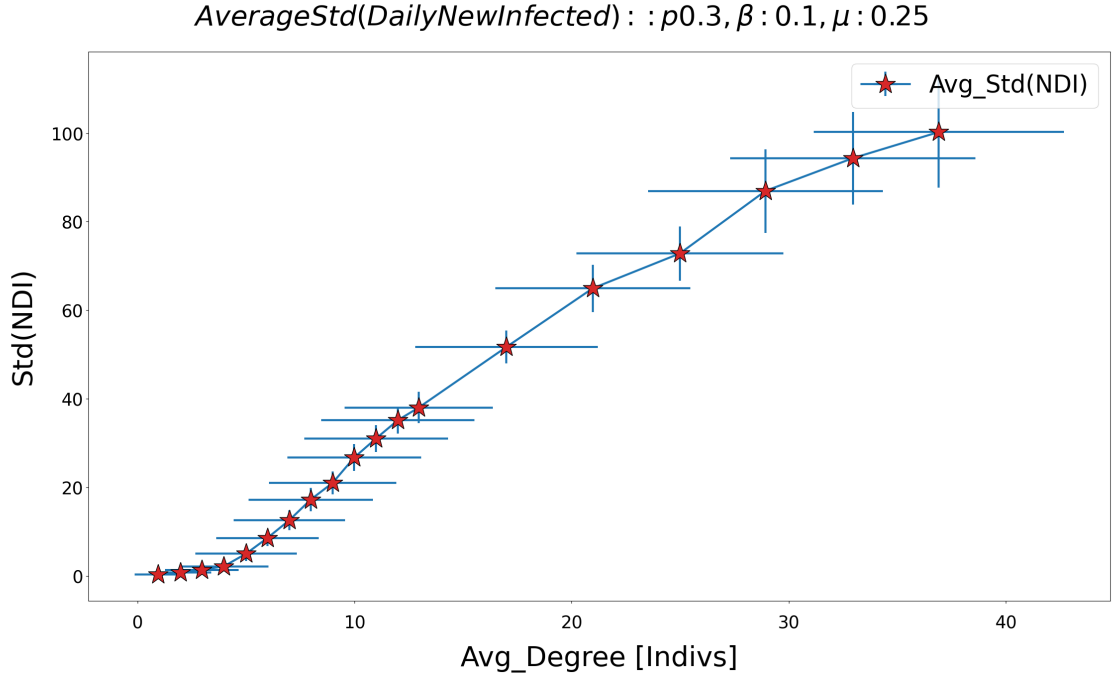
Figure 4.1: SD-Phase Transition of a Poissonian-SW Network with $N = 10^3, \beta = 0.1, D_c \sim 4.8, \mu = 0.25$. As before, the D_c of the figure differs from the analytical estimation for finite size effects.

The fuse model

In this section, it is going to be obtained, for a one-dimensional regular grid with average degree D, the critical degree [27]

$$D_c = 1 + \frac{2}{\beta d(1 + p)}. \tag{4.4}$$

The same result may be used for the poissonian small-world network.

The first step is to introduce an invariant probability function $\rho(x, t) = f(x - vt)$ that grasps the dynamics of the infection front: $\rho(x, t)$ is the probability a node in the x location is infected at time t. In particular, x is the network distance, as the number of in-between nodes from the initial infected node; t is the time passed in days after disease started infecting and v is the traveling velocity of the spreading along the x axis. A pedagogical representation of ρ could be find in [26]: roughly speaking, it has the shape of a $\arctan(x - vt)$ where, on the x-axis, there is the 1d-chain on nodes; while, on the ordinates, it has been translated to fit the $y \in [0, 1]$ interval[4]. Furthermore, it has been centered to an arbitrary node (q), which has $\rho(q) = 1/2$.

To estimate v and $\rho(x, t)$, it is taken advantage on $\rho(x, t + 1)$ assuming that v is such that all the "non-susceptible nodes" are indeed infected, not recovered. So, the

---

[4]The "past" nodes have higher ρ while the future nodes have diminishing probability of being infected.

probability of being infected at a following time $t + 1$ is

$$\rho(x, t+1) = \rho(x,t) + (1 - \nu)2D\beta\langle\rho\rangle(x,t)(1 - \rho(x,t)) \quad (4.5)$$

where $\langle\rho\rangle(x,t)$ is the (local) average probability of being infected at (x,t) over the adjacent neighbors. In other words, the Equation 4.5 considers the probability of already being infect; alongside with 2 terms:

- $1 - \nu$ and $1 - \rho$ which are the fraction of infected nodes and the probability of being susceptible;

- $\delta := 2(1 - \nu)D\beta$ is the total probability rate of having (at least) a contagion over the $2(1 - \nu)D$ infected neighbors.

For a one-dimensional front, its radius proceed at a $x = vt$ pace and, defining τ as the next infection time, the a ratio of the non-susceptible/susceptible nodes may be approximated as

$$r := \frac{1 - \nu}{\nu} = \frac{t + \tau}{t} = 1 + \frac{\tau}{t} = 1 + v\tau.$$

Moreover, defining τ as the time one single node is infected $v \cdot \tau = 1$ implying that

$$R = \frac{1}{2 + \mathcal{K}} = \frac{1}{2}$$

since for a flat "fuse model" $\mathcal{K} = 0$. The naive expectation is recovered as $\nu \sim 1/2$.

Inserting the Equation 4.5 on the invariance condition

$$\frac{\partial\rho}{\partial t} = -\frac{\partial x}{\partial t}\frac{\partial\rho}{\partial x} = -v\frac{\partial\rho}{\partial x}, \quad (4.6)$$

with the further approximation that $\langle\rho\rangle(x,t) \sim \rho(x,t)$,

$$\rho(x,t) = \frac{1}{1 + e^{\delta(\tau x - t)}} \quad (4.7)$$

with $\delta := 2(1-\nu)D\beta$. For this solution, $\rho(vt,t) = \rho(0,0) = 1/2$ marks the infection front [26].

The average probability of the nodes on the left of the front being infected is

$$\hat{\rho} \sim \frac{1}{\nu D}\int_{-\nu D}^{0} dx\rho(x,0) = 1 + \frac{1}{\delta'}\log(\frac{1 + e^{-\delta'}}{2}) \sim \frac{1}{2} + \frac{1}{8}\delta' + O(\tau^3) \quad (4.8)$$

where it has been exploited that $\delta' := \tau\delta\nu D$ and the last expansion as $\tau \ll 1$. Thus, $\hat{\rho}$ could also interpreted as the probability of the past infected nodes.

Moreover, n could be obtained considering the situation where a ν nodes are susceptible but the "past" nodes are not and viceversa: $\nu\hat{\rho}$ or $(1 - \nu)(1 - \hat{\rho})$. Formally,

$$n \sim 1 + (D - 1)(\nu\hat{\rho} + (1 - \nu)(1 - \hat{\rho})) \sim 1 + (D - 1)(\frac{1}{2} - (1 - 2\nu)\frac{1}{8}\delta') \quad (4.9)$$

which for a flat infection, where $\nu \sim 1/2$, reduces to

$$n = 1 + \frac{1}{2}(D - 1). \tag{4.10}$$

Ultimately, introducing the pD long-range ("lr") neighbors; while keeping the infection low, $n_{lr} \sim 1$.

Since n is a local estimate, removing the n_{lr} contribution,

$$n = 1 + \frac{1 - p}{2}(D - 1) \tag{4.11}$$

and, therefore,

$$D_c = 1 + \frac{2}{\beta d(1 + p)}. \tag{4.12}$$

This estimate express the critical degree, i.e. the average number of contact under which the mean-field approximation fails, as a function of $p, \beta, \mu = 1/d$. As a second order approximation in τ as been exploited in Equation 4.8, overestimates the number of infected which, in turn, drives to an overestimation of D_c. Other sources of divergence from the simulated critical degree may be the finite size effects or deviation from $= 0$.

In the "Results" chapter, D_c for different simulation is reported but only for poissonian small-world network, it approximates the point where the $SD(C(t))$ start increasing.

48

Results

50

Conclusions

52

Appendix - Network Models

A.1 Scale-Free Distributions Details

In the following part, the discrete approach, assuming $k = 0, 1, 2 \cdots$, is going to be developed; while only the continuum result would be reported. In fact, the discrete approach provides the ready-to-use formulas handle by a calculator which, by construction, performs discrete operations. On the other hand, the continuum approach is useful for analytically support the calculations and could be obtained similarly to the discrete case. Formally, by exploiting the constraint on probabilities $p_k = C \cdot k^{-\gamma}$

$$\sum_{k=1}^{\infty} p_k = 1,$$

it is possible to recover the closed form [6]

$$p_k = \frac{1 - p_0}{\zeta(\gamma)} k^{-\gamma}. \tag{A.1}$$

where $p_0 = p_{k=0}$ has been separated, since the power-law diverges for $k = 0$.

On the another hand, the continuum formalism drives to

$$p(k) = (\gamma - 1)k_{min}^{\gamma-1}k^{-\gamma},$$

since the the smallest value for the power law to hold is not 1, as in the discrete approach, but the minimum degree for a node k_{min}.

Formally, k_{min} is the degree for which

$$\int_0^{k_{min}-1} p(k)dk \overset{!}{=} 1/N \qquad \text{or} \qquad P(k_{min} - 1) \overset{!}{=} 1/N. \tag{A.2}$$

In the next part, the allowed maximum degree k_{max} would be obtained both for an exponential bounded and a power-law distributions. Through this quantitative insight, it would be shown the dependence of k_{max} on N, i.e. the size of the network.

In particular, the exponential distribution for which $\int_{k_{min}}^{\infty} p(k)dk = 1$ holds, is

$$p_k = \lambda e^{\lambda k_{min}} e^{-\lambda k}$$

where, using Equation A.2, $k_{min} = 1$. Furthermore, k_{max} (or the "natural cutoff") is defined to be such that the at most one node i could have a degree k_i larger than k_{max}.

In other words, the cumulative distribution for degrees greater than k_{max} is $1/N$; and, hence, k_{max} is expected to be the size of the bigger hub.

In formulas,

$$\int_{k_{max}}^{\infty} p(k)dk = 1/N \implies k_{max} = k_{min} + \frac{\ln(N)}{\lambda}. \tag{A.3}$$

So, k_{max} would not differ much with respect to k_{min} since $\ln(N)$ strongly reduces contribution due to the size of the graph N. For a Poisson distribution, more effort is needed but the same interpretation holds.

On the other hand, applying the same rationale to a scale-free network,

$$k_{max} = k_{min} N^{\frac{1}{\gamma-1}}. \tag{A.4}$$

In this case, there could be order of magnitude of difference between k_{max} and k_{min}. Indeed, as a simple example, WWW forms a graph of on $N \sim 10^5$ documents with $\langle k \rangle \simeq 4.6$ and $\gamma = 2.1$ [6]. By assuming $\lambda = 1$ and recovering that , $k_{min} = 1$ using Equation A.2, $k_{max} \sim 14$; while $k_{max} = 95.000$ for a scale-free network with $\gamma = 2.1$, e.g. WWW network. This short results reinforces the quoted insight that for the class of "exponentially bounded distributions" the hubs are strongly forbidden since the nodes have comparably the same degree. Instead, the highly connected nodes are naturally arising by increasing N with an underlying power-law distribution.

Not all the network are scale-free since the hubs are present only if nodes have unlimited degree capacities. In fact, there could be constraints, such as cost-benefit problem, which limit k_{max} of the bigger node. In this overview, random network may be the best fit, e.g. national highways network, power grid of generators and switches, $\cdots$

With all the previous effort a more quantitative approach to the explanation on the "scale-free" etymology could be ultimate. More precisely, defining the statistics momentums as

$$\langle k^n \rangle := \int_{k_{min}}^{k_{max}} k^n \, p(k)dk \tag{A.5}$$

it is possibile to recover the the mean and the variance for $n = 1, 2$. Thus, for a poissonian distribution, the nodes degree k belong to $\langle k \rangle \pm \langle k \rangle^{1/2}$ interval, which drives to the desired $\langle k \rangle$ scale. At the same time, for a power-law distribution holds

$$\langle k^n \rangle \sim \frac{k_{max}^{n-\gamma+1} - k_{min}^{n-\gamma+1}}{n - \gamma + 1}. \tag{A.6}$$

Therefore, knowing that real networks have typically $\gamma \in (2,3)$, e.g. WWW has $\gamma = 2.1$ [6]

$$k \in \langle k \rangle \pm \lim_{N \to \infty} \langle k^2 \rangle = \infty. \tag{A.7}$$

where the hidden limit of $k_{max} \to \infty$ is understood, since k_{max} increases with the system size. So, no scale naturally arises.

The result rapidly quoted before is surrounded by three other regimes [10]:

$$\langle d \rangle = \begin{cases} constant & \gamma = 2 \\ \ln(\ln(N)) & \gamma \in (2,3) \\ \frac{\ln(N)}{\ln(\ln(N))} & \gamma = 3 \\ \ln(N) & \gamma > 3 \end{cases} \tag{A.8}$$

Anomalous Regime ($\gamma = 2$): Due to $k_{max} \sim N$, the network is forced into a "wheel rim configuration", a central hub with spoke nodes. In this regime, the average path length is independent on the size of the graph.

Ultra-Small World Regime ($2 < \gamma < 3$): By predicting a $\ln(\ln(N))$ trend of the average distance, Equation A.8 guides to the concept of "ultra-small world phenomenon" as a slower growth than random network. Indeed, recalling the worldwide social network, $\ln \ln(N) \sim 3$ while $\ln(N) \sim 22$ more than few times higher than the expected "six degrees of separation".

Critical Point ($\gamma = 3$): Looking back on the Equation 2.7, the variance ($n = 2$) is finite, marking the passage between the small-world ($\sim \ln(N)$) and the ultra-small world ($\sim \ln(\ln(N))$) classes of graphs.

Small World ($\gamma > 3$): In this regime, the finiteness of the variance is understood, due to the light aggregation that the high-degree nodes could perform in these kind of networks.

Moreover, recalling the root of "ultra-small world" Equation A.4, for $N \lesssim 10^2$ the path length distributions overlap in all the reported regimes and, so, converges to a Poisson distribution ($\gamma > 3$) whose hubs are small to produce the "ultra-small world" phenomenon. In fact, to validate the power-law nature of a distribution, the nodes degree should be separated by 2-3 orders of magnitude. Hence, reverting the Equation A.4,

$$N = \left(\frac{k_{max}}{k_{min}} \right)^{\frac{1}{\gamma-1}},$$

which lucidly reproduce $N \sim 10^2$ for $\gamma = 2.1$ (WWW network).

Thus, $\langle d \rangle$ is changing with γ, as the smaller it is, the shorter would be the average distance; but also with N, as it is the "ab ovo" hypothesis for having big hubs.

A sociological fact has to be remarked when trying to map the "ultra-small" world over social interactions. Indeed, an online experiment which aimed to replicate the "six degrees" concept find that individuals involved in complete chains, reaching their target, were less likely to send a message to a hub than individuals involved in incomplete chains [12]. The reason may be self-imposed, since hubs are perceived as being busy, so contacted only in real need and avoid to concretely complete an "online experiment" challenge. Thus, network is not everything since actual success depends sensitively on individual incentives.

A.2   Recovering Power-Law Distribution

Due to a new $m-$stubs node, the degree of an existing node i changes at a rate

$$\frac{dk_i}{dt} = m\Pi(k_i) = m\frac{k_i}{\sum_j k_j}. \tag{A.9}$$

Hence, for $t >> 1$,

$$k_i(t) = m\left(\frac{t}{t_i}\right)^{\beta} \tag{A.10}$$

where β, alias "dynamical exponent", has the value $\beta = 1/2$ [6] and t_i is the entering time for the node i.

The remarks, coming from the Equation A.10, are multiple:

- all the nodes follow the same power law, since β is independent on the node label i. Furthermore, it is independent on m and m_0 constructing parameters;

- the inverse dependence of $k_i \sim 1/t_i$ is a clear signal of the fact that old nodes are supposed to be larger, also called "first-mover advantage" in business;

- both $\beta = 1/2 < 1$ (sublinear growth) and the rate $dk_i/dt \sim 1/\sqrt{tt_i}$ underly a competing scenarios: the existing nodes compete for the $m-$new stubs with a growing bunch of nodes, resulting in a sublinear growth, which drives to a decreasing rate of acquaintances.

To validate the power-law nature of the probability of having a certain node degree, by exploiting Equation A.10 it is possible note that the nodes i with a degree higher than k are such that

$$t_i < t\left(\frac{m}{k}\right)^{1/\beta}.$$

Thus, for $t >> m_0$, the (cumulative) probability of having a degree less than k is

$$P(k) = 1 - \left(\frac{m}{k}\right)^{1/\beta},$$

from which the probability of having a degree k could be recovered, by differentiating $P(k)$ and substituting $\beta = 1/2$, $p_k = 2m^2k^{-3}$ [6]. As desired, the distribution exhibits a long-tail nature with $\gamma = 1/\beta + 1 = 3$. Alongside with the BA model simulations, the dependence between γ and β demonstrate the deep relationship among the graph topology and the temporal degree evolution, respectively contained in γ and β parameters.

Another feature present in Equation A.10 is that the degree distribution is independent on t and N or a "stationary distribution". This result is the one expected, since the model aims at describing the real networks in general; thus, of rather different age and size.

A.3 The Origins of Preferential Attachment

Since preferential attachment plays a relevant role in guiding the dynamics of a real network, the goal is to recover the same $\Pi(k)$ as the one assumed in the Barabási-Albert model. It worths to introduce two classes of microscopic processes that generate it naturally [6]: *"local (or random) mechanisms"* representing the interplay between random events and some structural properties of the network; and *"global (or optimized) mechanisms"* which take advantages of a cost-benefit analysis on the whole graph. In the following, it would be reviewed the "Link Selection" and "Copying" Models which belongs to the random class; and, finally, the "Optimization" Model of the second one.

A.3.1 Local Mechanisms

Link Selection

Assuming growth is at work, i.e. at each event time a new node is added, select randomly an edge and connect the new node to a vertex belonging to that edge. Thus, the probability q_k that the selected node, already present in the graph, has degree k is

$$q_k = \frac{1}{\langle k \rangle} \cdot k p_k \tag{A.11}$$

where the first factor exploiting the normalization constraint $\sum_k q_k = 1$. Hence, it has the hoped form, being linear in the number of stubs (k) and in the frequency of the degree-k nodes (p_k). This displays an increasing of the chance of wiring to a degree k node, by enhancing either the number of degrees or the presence of degree-k nodes themselves.

Copying Model

The idea is compute the probability of obtaining a connection of a degree-k node by mimicking the phenomenon of authors of a new webpage which copy links from existing webpages on the related topic [17].

By introducing a new node n in the network, as before, the difference stands in the link selection. More in detail, selecting an existing node u regardless of its degree, e.g. picking an arbitrary web document which topic is related to n, there are two possible way of connections:

- Random (direct) connection: with a probability p, link n to u;

- Copying (undirect) connection: with likelihood $1 - p$, wire n to an neighboring node v of u. Reprising the web example, the new webpage n, copying the link of u, connects with its target v.

As a constraint to recover a dependence on k of $\Pi(k)$, fix that v should have degree k, by considering $p_v = k/2L$. Hence,

$$\Pi(k) = \frac{p}{N} + \frac{k}{2L}(1 - p) \tag{A.12}$$

which is linear in k, yielding the desired preferential attachment. The power of this approach is its "empirical" root which could be detected from citation networks to the social network of acquaintances.

A.3.2 Global Mechanisms

A.3.3 Optimization Mechanism

The principle of minimization of cost-benefits analysis in an economic market [22] can lead to preferential attachment whether a proper cost function is considered.

For simplicity assume that the vertexes are laying in an unit square, e.g. routers in a square continent. As before, at each event time a new router i is added, while an edge with an existing node j is created according to the cost function [6]

$$C_i = min_{j \in G(V,E)} \left(\delta d_{ij} + h_j \right).$$

where d_{ij} is the euclidean distance between the vertex i and j; while h_j is the number of hops from j to a pre-defined "center" which embodies the best network performance, e.g. net distribution hub. More clearly, d_{ij} and h_j represent the physical distance and "network-based" distance. Their summation describes the cost to be minimized for real results.

Depending on δ and N, there are 3 network topologies:

- **Star Network** $(\delta < (1/2)^{1/2})$: only δd_{ij} drives the growth, thus, the new nodes connects with the central one forming a star;

- **Random Network** $(\delta \geq N^{1/2})$: h_j is the minimum quantity to be minimized. So, the new nodes connect to the nearest node;

- **Scale-free Network** $(4 \leq \delta \leq N^{1/2})$: in this intermediate regime scale-free topologies may be recovered as a result of optimization and randomness. In particular, optimization creates a basin of attraction with size $s \sim k_j$ and centered in j, within all the appearing nodes i wire with j. On the other hand, Randomness is the chance of choosing a specific basin.

Thus, linear preferential attachment can come from both rational choice and random actions. Yet, most complex systems are driven by processes that have a bit of both luck or reason.

Bibliography

[1] P. W. Anderson. More is different. *Science*, 177(4047):393–396, 1972.

[2] Christie Aschwanden. Five reasons why covid herd immunity is probably impossible, 2021.01.07. https://www.nature.com/articles/d41586-021-00728-2.

[3] Scott Baier and Samuel Standaert. Gravity models and empirical trade, 03 2020.

[4] Albert-László Barabási and Réka Albert. Emergence of scaling in random networks. *Science*, 286(5439):509–512, 1999.

[5] Albert-Lászlo Barabási. Scale-free networks:a decade and beyond. *Science*, 325(5939):412–413, 2009.

[6] Albert-László Barabási and Márton Pósfai. *Network science*. Cambridge University Press, 2016.

[7] Andrea Baronchelli and Mikayel Poghosyan. *Epidemic spreading on complex networks - MMath Year 3 Thesis/Dissertation Project*. PhD thesis, City University London, 04 2017.

[8] George E. P Box. *Statistics for experimenters : design, innovation and discovery*. Wiley-Interscience, Hoboken, N.J., 2nd ed. edition, 2005.

[9] Kimberly Chriscaden. Impact of covid-19 on people's livelihoods, their health and our food systems, 2021.07.01. site.

[10] Reuven Cohen and Shlomo Havlin. Scale-free networks are ultrasmall. *Physical Review Letters*, 90(5), Feb 2003.

[11] Mohammad Reza Davahli, Krzysztof Fiok, Waldemar Karwowski, Awad M. Aljuaid, and Redha Taiar. Predicting the dynamics of the covid-19 pandemic in the united states using graph theory-based neural networks. *International Journal of Environmental Research and Public Health*, 18(7), 2021.

[12] Peter Sheridan Dodds, Roby Muhamad, and Duncan J. Watts. An experimental study of search in global social networks. *Science*, 301(5634):827–829, 2003.

[13] Murray Gell-Mann. Simplicity and complexity in the description of nature. *Caltech Magazine*, 1987. http://calteches.library.caltech.edu/3579/.

[14] Youyang Gu. Path to normality website, 2021.01.07. https://COVID19-projections.com/path-to-herd-immunity/.

[15] Istvan Z Kiss, Joel C Miller, and Peter Simon. *Mathematics of epidemics on networks: from exact to approximate models*. Springer, 2017.

[16] Istvan Z Kiss, Joel C Miller, and Peter Simon. *Mathematics of epidemics on networks: from exact to approximate models*. Springer, 2017.

[17] Jon M. Kleinberg, Ravi Kumar, Prabhakar Raghavan, Sridhar Rajagopalan, and Andrew S. Tomkins. The web as a graph: Measurements, models, and methods. In Takano Asano, Hideki Imai, D. T. Lee, Shin-ichi Nakano, and Takeshi Tokuyama, editors, *Computing and Combinatorics*, pages 1–17, Berlin, Heidelberg, 1999. Springer Berlin Heidelberg.

[18] Filippo Menczer, Santo Fortunato, and Clayton A. Davis. *A First Course in Network Science*. Cambridge University Press, 2020.

[19] M. E. J. Newman. *Networks: an introduction*. Oxford University Press, Oxford; New York, 2010.

[20] Romualdo Pastor-Satorras and Alessandro Vespignani. Epidemic Spreading in Scale-Free Networks. *Physical Review Letters*, 86(14):3200–3203, April 2001.

[21] Clara Pizzuti, Annalisa Socievole, Bastian Prasse, and Piet Van Mieghem. Network-based prediction of covid-19 epidemic spreading in italy. *arXiv*, 2020.

[22] Gerald Shively. An overview of benefit-cost analysis, 01 2012.

[23] Complex System Society. Site, 2021.01.07. https://cssociety.org/about-us/what-are-cs.

[24] Olaf Sporns. The non-random brain: Efficiency, economy, and complex dynamics. *Frontiers in Computational Neuroscience*, 5:5, 2011.

[25] Johannes Stübinger and Lucas Schneider. Epidemiology of coronavirus covid-19: Forecasting the future incidence in different countries. *Healthcare*, 8(2), 2020.

[26] Stefan Thurner, Peter Klimek, and Rudolf Hanel. Appendix: A network-based explanation of why most covid-19 infection curves are linear. *Proceedings of the National Academy of Sciences*, 117(37):22684–22689, 2020.

[27] Stefan Thurner, Peter Klimek, and Rudolf Hanel. A network-based explanation of why most covid-19 infection curves are linear. *Proceedings of the National Academy of Sciences*, 117(37):22684–22689, 2020.

[28] Florian Täube. *Transnational Networks and The Evolution of the Indian Software Industry:*, volume 6, pages 97–121. Springer, 01 2005.

[29] Daniel Zelazo and Mehran Mesbahi. Graph-theoretic analysis and synthesis of relative sensing networks. *Automatic Control, IEEE Transactions on*, 56:971 – 982, 06 2011.

[30] Abdelhafid Zeroual, Fouzi Harrou, Abdelkader Dairi, and Ying Sun. Deep learning methods for forecasting covid-19 time-series data: A comparative study. *Chaos, Solitons & Fractals*, 140:110121, 2020.

[31] Janaina P. Zingano, Paulo R. Zingano, Alessandra M. Silva, and Carolina P. Zingano. Herd immunity for covid-19 in homogeneous populations, 2021.