UNIVERSITY OF PALERMO

PHD JOINT PROGRAM:
UNIVERSITY OF CATANIA - UNIVERSITY OF MESSINA
XXXIV CYCLE

DOCTORAL THESIS

An Original Convolution Model to analyze Graph Network Distribution Features

Author:
Giuseppe GIACOPELLI

Supervisor:
Prof. Domenico TEGOLO

Co-Supervisor:
Prof. Michele MIGLIORE

*A thesis submitted in fulfillment of the requirements
for the degree of Doctor of Philosophy*

in

Mathematics and Computational Sciences

April 11, 2022

"Ad Maiora!"

Author lost in the ancient Rome history

UNIVERSITY OF PALERMO

Abstract

Department of Mathematics and Computer Sciences

Doctor of Philosophy

An Original Convolution Model to analyze Graph Network Distribution Features

by Giuseppe GIACOPELLI

Modern Graph Theory is a newly emerging field that involves all of those approaches that study graphs differently from Classic Graph Theory. The main difference between Classic and *Modern Graph Theory* regards the analysis and the use of graph's structures (micro/macro). The former aims to solve tasks hosted on graph nodes, most of the time with no insight into the global graph structure, the latter aims to analyze and discover the most salient features characterizing a whole network of each graph, like degree distributions, hubs, clustering coefficient and network motifs. The activities carried out during the PhD period concerned, after a careful preliminary study on the applications of the Modern Graph Theory, the development of an innovative *Convolutional Model* to model brain connections at the cellular level capable of combining exponential models and power law models. This new theoretical framework has been introduced in the first instance with an aspatial graph formulation and then proposed a spatial graph model with Convolutive connectivity able to fit the degree distributions of data driven *Connectome* reconstructions. In order to evaluate the qualities of the Convolutional Model, theoretical graphical models capable of characterizing brain activity were taken into consideration. In the specific case, the model examined characterizes the epileptic activity through a simple Hindmarsh-Rose model system of point neurons and reproduces the functional characteristics observed in the data driven model. Such a model provides insight into the deep impact of micro connectivity in macro-scale brain activity. Other evaluations have been done in different applications, in the field of image cell segmentation with Explainable Artificial Intelligence's neuronal agents in which has been used a methodology that is not only explainable but also resistant to adversarial noise and also in the field of modelling Covid-19 outbreak in gaining insight on vaccines and role of our habits as individuals in the pandemic spread. Therefore, the core of the thesis is to introduce Modern Graph Theory with a new competitive Convolutive Model and then expose some applications to real-world problems like a characterization of Brain networks, simulation and analysis of Brain dynamics with a particular focus on Epilepsy, Immunofluorescence images segmentation with neuronal based agents and modelling of Covid-19 Epidemic spread with a specific interest in human social networks. All this takes continuously into account the whole dialogue between Graph Theory and its applications.

Acknowledgements

Contents

List of Figures

List of Tables

List of Abbreviations

BA model	**B**arabasi-**A**lbert model
CGT	**C**lassic **G**raph **T**heory
ER model	**E**rdős-**R**enyi model
GCM	**G**eneralized **C**onvolutive **M**odel
HR model	**H**indmarsh-**R**ose model
LIF	**L**eaky **I**ntegrate and **F**ire
MGT	**M**odern **G**raph **T**heory
PSNR	**P**eak **S**ignal-to-**N**oise **R**atio
SCM	**S**patial **C**onvolutive **M**odel
STFT	**S**hort **T**ime **F**ourier **T**ransform

List of Symbols

A_G	Adjacency matrix of graph G
$B(N, p)$	Binomial random variable
$B_p^N(k)$	Binomial random variable distribution
$D_G^I(i)$	Indegree of the node i inside the graph G
$D_G^I(i, H_G)$	Indegree of the node i inside the graph G constrained to the subset H_G
$D_G^I(i, \Omega_G)$	Indegree of the node i inside the graph G constrained to the region Ω_G
$D_G^O(i, \Omega_G)$	Outdegree of the node i inside the graph G
$D_G^O(i, H_G)$	Outdegree of the node i inside the graph G constrained to the subset H_G
$D_G^O(i, \Omega_G)$	Outdegree of the node i inside the graph G constrained to the region Ω_G
$\mathbb{E}(I)$	Expected value of random variable I
$G = (V(G), E(G))$	General graph with nodes $V(G)$ and edges $E(G)$
N_G	Number of nodes of graph G
$P(D_G^I = k)$	Indegree distribution of graph G
$P(D_G^O = k)$	Outdegree distribution of graph G
$P(E)$	Probability of event E
X_p	Bernoulli random variable with probability p
$[u_q * v_q]_k$	Convolution between the distributions u_k and v_k
$\chi_p(k)$	Bernoulli random variable distribution with probability p
$\hat{x}(\omega)$	Fourier Transform of $x(t)$
$\hat{x}(\omega, t_0)$	Short Time Fourier Transform of $x(t)$ centered in t_0
$\bar{z}$	Conjugate of $z \in \mathbb{C}$

Introduction

1.1 What are networks and why have to be studied

A simple network form is a set of points *(vertices)* paired by lines *(edges)*; such definition is reported by Newman in (Newman, 2010) (see figure 1.1). This simple data structure can model various experimental scenarios in biology, computer science, epidemiology, sociology, etc. For example, at the beginning of the graph theory, the vertices of the graphs were geographical locations and the edges paths through the city of Konigsberg (Euler, 1736). This geographical interpretation of the network is still studied to model traffic flow in cities (Easley and Kleinberg, 2010), where the vertices of the graph are key locations of the city, and the edges are the routes between these locations.
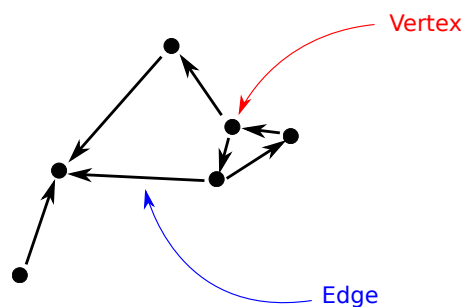
FIGURE 1.1: Simple representation of a graph.

Network analysis divides into two main branches (Steen, 2010): *Classic Graph Theory* and *Modern Graph Theory*. The example of the traffic networks is a perfect example of the Classic Graph Theory (CGT since now) approach. CGT bases its study on the idea of the graph *as it is*; therefore, the primary assumption of CGT is that *is known and wholly determined the full connectivity of the network represented*. However, this simple assumption could not be valid in every experimental setting; for example, it is the case of Connectome.

Define a Connectome as an ensemble of all connections in the brain (Jirsa and McIntosh, 2007), and it can be described in terms of networks at several scales (Bullmore and Sporns, 2009).

One of the most commonly used approaches is to assign to the graph nodes whole cerebral areas (Farahani, Karwowski, and Lighthall, 2019); where the functional connectivity of the subject is reconstructed by using correlations or model-based inferences, and also the synchronous activation of two regions determines

the functional connection of two brain regions after a given stimulus (Bullmore and Sporns, 2009). In the fMRI case, the brain is first of all subdivided in small areas and for each volume is calculated the mean activity (Farahani, Karwowski, and Lighthall, 2019) extracting the average activity. Then on these averaged signals are estimated the strength of connections using analysis like Principal Component Analysis, Transfer Entropy or Cross Correlation (Farahani, Karwowski, and Lighthall, 2019).

However, recently, the improvements of the resolution of microscopy have introduced the possibility to reconstruct the Connectome at cellular scales (Markram et al., 2015), and this has created new opportunities for the analysis of this new kind of data (Gal et al., 2017) and its modelling (Giacopelli et al., 2021). One of the main features of the connectivity data is that they are a sort of fingerprint of the individual, which means that there are no two individuals with the same Connectome. In addition, the very massive number of neurons in the brain has limited the complete reconstruction of the Connectome of each individual. This lack of knowledge about the whole network connectivity of Connectome and its uniqueness for each individual requires a new approach in studying graphs, this change of pass in Graph Theory can be called the *Modern Graph Theory* (Barrat, Barthélemy, and Vespignani, 2008; Steen, 2010).

The concept of *Modern Graph Theory* (since now MGT) for a long time has been subjected to fierce debate because many people have argued MGT as nothing of new concerning CGT (Barrat, Barthélemy, and Vespignani, 2008). However, some authors (Steen, 2010; Barrat, Barthélemy, and Vespignani, 2008) identify the change of MGT in a transformation of methodological processes of network analysis. The key difference between CGT e MGT is that CGT usually solves given problems in an opaque way; in contrast, one of the main goals of MGT is the concept of network analysis. To clarify the differences between the two approaches, let's start with an example: defining a *computer network* as a graph in which each vertex is a computer and the edges the connection among computers. A stereotypical problem of CGT is to determine the minimum path between two given machines inside the networks (Steen, 2010), and CGT has developed several algorithms (for example, Dijkstra's algorithm and Bellman-Ford algorithm (Steen, 2010)) able to solve this task. However, even after solving this task, we do not know what is inside the network of computers that we are analysing.

The MGT approach is quite different, indeed given a network of computers, it can be:

- calculated how many incoming connections and outgoing connections occur (Barabási and Albert, 1999);

- discovered their power law behavior of distributions (Barabási and Albert, 1999);

- calculated the average of its path lengths (Dorogovtsev and Mendes, 2003);

- calculated the clustering coefficient (Steen, 2010) etc.

All of these quantities are not directly linked to the solution of a specific task (like in CGT) but they give an insight of how the network is connected. The degree distributions show if there are over connected nodes in the network (the so called *hubs* (Newman, 2010)) giving a sort of ranking to nodes in terms of centrality (Page et al., 1999), the average path length quantify the degree of separation between two

general nodes (Easley and Kleinberg, 2010) and the clustering coefficient gives an insight on how much the nodes are interconnected between their selves (Steen, 2010).

The evolution of MGT made possible a change of paradigm where the target of our knowledge is not anymore the problem to solve but the network itself, introducing a series of approaches able to give a qualitative description of networks with billions of nodes (Barabási and Albert, 1999) like the internet. Then, the simplification of the theory (made sometimes at the cost of some approximations) has led to a clearer and more direct explanation of the results, bringing a deeper comprehension of the modelled phenomena.

One of the main actors of this "new science of networks" is the unseen computational power of our days (Barrat, Barthélemy, and Vespignani, 2008). It has led to the abandonment of a strict formalism for more sketched approaches, which become more powerful when coupled with numerical simulations, based on statistical physics and mathematical physics (Dorogovtsev and Mendes, 2003; Giacopelli, Migliore, and Tegolo, 2020). For this reason, the used theoretical frameworks will be discussed and given an insight into the computer science realization of the models described.

All of these advances in Graph Theory have improved the understanding of experimental networks. Key examples that will be discussed are neuronal networks, and their modelling (Markram et al., 2015; Giacopelli et al., 2021) in which has been proved a *convolutive behavior* (Giacopelli, Migliore, and Tegolo, 2020) for many reconstructed neuronal networks. Another example is computer science where it has been proved that most of the computer networks (like the internet) have a power law behavior (Barrat, Barthélemy, and Vespignani, 2008) leading to the understanding of the importance of hubs for their functionality. A contemporary example is COVID-19 pandemic where the methods of network-based computational epidemiology (Barrat, Barthélemy, and Vespignani, 2008; Giacopelli, 2021) have been extensively used to predict the epidemic spread and to better weight decisions for this pandemic.

In the following will be made a brief introduction to Classic Graph Theory, then an introduction to Modern Graph Theory and finally will be explored the concept of *Convolutive model* (Giacopelli, Migliore, and Tegolo, 2020; Giacopelli et al., 2021; Giacopelli, Migliore, and Tegolo, 2020) with its applications.

1.2 The graph

A graph is the formalization of the concept of a network. To define a graph G, a set of vertices V must be defined; thus, the vertices will be labelled with an integer number. Then, a set of edges between nodes $E \subseteq V \times V$ will be defined. Each connection is a couple $(i, j) \in V \times V$ where i is called starting node of the connection and j ending node of the connection.

Defining $V = V(G)$ and $E = E(G)$, the designed graph is a graph $G = (V(G), E(G))$ (Newman, 2010). More formally

Definition 1.2.1 *A graph $G = (V(G), E(G))$ is an object where $V(G) \subseteq \mathbb{N}$ is the set of nodes of G and $E(G) \subseteq V(G) \times V(G)$ is the set of edges of G.*

An example of graph is in figure 1.2 where the vertices are the somas of the neurons (in our case the node number will range from 1 to 6) and the edges are the synaptic connections (in this case is $E = \{(1, 2), (3, 2), (4, 2), (4, 6), (5, 4), (5, 3), (6, 5)\}$).
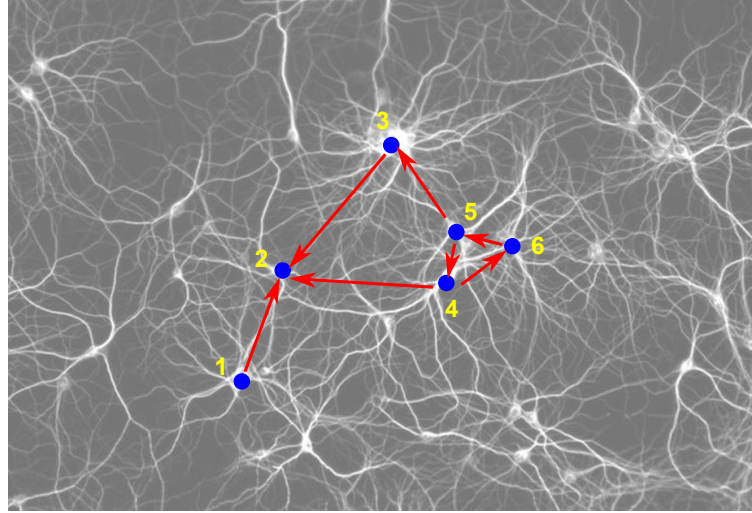
FIGURE 1.2: Graph theoretical representation of Connectome.

Another kind of representation of a graph is in terms of adjacency matrix (Newman, 2010). The Adjacency matrix A_G is defined as the binary matrix where the element $(A_G)_{ij}$ is 1 if there is a connection starting from node i and arriving to node j, and 0 otherwise. In this example the adjacency matrix is:

$$A_G = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

More formally:

Definition 1.2.2 *For a given graph $G = (V(G), E(G))$, the adjacency matrix A_G is defined as*

$$(A_G)_{ij} = \begin{cases} 1, \text{ if } (i,j) \in E(G) \\ 0, \text{ otherwise} \end{cases}$$

Moreover, each node can be assigned a position x_i, and because the graph is planar, it will be a two-dimensional coordinate on the plane.

1.3   Brief introduction to probability

Since now will be widely used the concept of probability, but what is probability? This question can appear simple, but it is not so because there are many definitions of probability. The two main currents of thoughts are:

1. Frequentist approach (Neyman, 1937);

2. Bayesian approach (Regazzini, 1987).

The first approach is based on empirical frequency; instead, the second approach is based on rationality. The first approach will be the one used since now.

In Frequentist probability theory, the fundamental concepts are the random variables and the events.

Definition 1.3.1 *A random variable is an abstract entity able to make non predictable choices inside a set of possible outcomes.*

For example a dice is a random variable D_6 and the set $\{1, 2, 3, 4, 5, 6\}$ is its set of outcomes.

Definition 1.3.2 *An event is a proposition involving one or more random variables.*

Example of propositions are $D_6 = 1$ (the dice result is face one) or $D_6 > 3$ (the face number is greater than 3) or $D_6 = 1 \vee D_6 > 4$ (the face is one or greater than four) or $D_6 < 5 \wedge D_6 \in \{2, 4, 6\}$ (the face is smaller than 5 and must be in the set $\{2, 4, 6\}$).

Given a proposition E and executing N_s independent tests involving E, and using the Frequentist probability approach, the probability $P(E)$ is defined as:

$$P(E) = \frac{\text{number of cases in which } E \text{ is true}}{N_s} \in [0, 1].$$

Given an integer random variable I (random variables with integer values), the distribution of the random variable is defined as the following sequence:

$$\iota_j = P(I = j), \text{ for } j = -\infty, ..., +\infty$$

and the expected value $\mathbb{E}(I)$ is defined as

$$\mathbb{E}(I) = \sum_{j=-\infty}^{+\infty} jP(I = j).$$

In order to simplify the notation since now will be possible to compute the expected value of a distribution ι_j as

$$\mathbb{E}(\iota_j) = \sum_{j=-\infty}^{+\infty} j\iota_j$$

where this notation can be explained saying that was computed the expected value of some random variable with distribution ι_j. In conclusion another important concept is the similarity between random variables. Defined I_1 and I_2 they are similar (in symbols $I_1 \sim I_2$) if they have the same distributions, more precisely

$$P(I_1 = j) = P(I_2 = j), \text{ for } j = -\infty, ..., +\infty.$$

1.4 The degree

The degree is the number of connections incoming or outgoing to a node. In particular, the indegree is the number of incoming connections of a node, and the outdegree is the number of outgoing connections of a node. For example, on the network of figure 1.2 can be computed graphically the indegree that is 0 for node 1, 3 for node 2, 1 for node 3, 1 for node 4, 1 for node 5 and 1 for node 6. So we have as a set of indegrees $\{0, 3, 1, 1, 1, 1\}$. Many authors (Steen, 2010) used such a method to reorder the set of the degrees creating the degree sequence, but will not be adopted this strategy. On the other hand, it can be computed the outdegree set as $\{1, 0, 1, 2, 2, 1\}$. Therefore:

Definition 1.4.1 *For a given node $i \in V(G)$ the indegree $D_G^I(i)$ can be defined as*

$$D_G^I(i) = \sum_{j \in V(G)} (A_G)_{ji}$$

and the outdegree $D_G^O(i)$ as

$$D_G^O(i) = \sum_{j \in V(G)} (A_G)_{ij}$$

therefore the indegree is the sum on columns of adjacency matrix and the outdegree the sum on rows of adjacency matrix.

Given the degree sets can be computed the probability that a node has indegree 0 (that is $\frac{1}{6}$) or indegree 1 (that is $\frac{4}{6}$). For this reason we define a random variable D_G^I such as

$$P(D_G^I = k) = \frac{\text{number of nodes with indegree } k}{\text{number of nodes of the graph}}$$

since now the number of nodes of the graph G will be N_G. Simmetrically

$$P(D_G^O = k) = \frac{\text{number of nodes with outdegree } k}{N_G}$$

The probability distribution $P(D_G^I = k)$ is called indegree distribution of the graph G, and the probability distribution $P(D_G^O = k)$ is called the outdegree distribution of the graph G.

Anyway we could be interested in calculating degrees in some particular subsets of nodes (for example $H_G = \{1, 2, 3\}$) or regions of space (for example the region Ω_G^1 of figure 1.3).
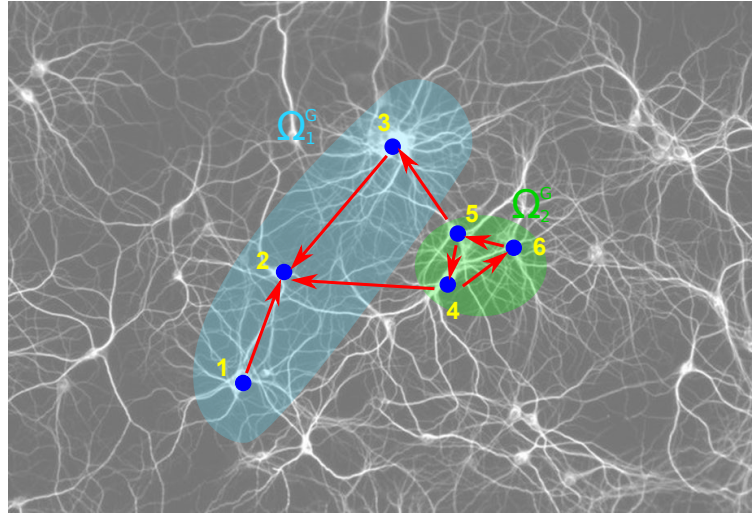
FIGURE 1.3: Example of spatial regions.

Definition 1.4.2 *For a given $H_G \subseteq V(G)$ can be defined the subset constrained degrees $D_G^I(i, H_G)$ as:*

$$D_G^I(i, H_G) = \sum_{j \in H_G} (A_G)_{ji}$$

and simmetrically $D_G^O(i, H_G)$ is defined as

$$D_G^O(i, H_G) = \sum_{j \in H_G} (A_G)_{ij}$$

Definition 1.4.3 *For a given region $\Omega_G \subseteq \mathbb{R}^2$ can be defined the region constrained degrees $D_G^I(i, \Omega_G)$ as:*

$$D_G^I(i, \Omega_G) = \sum_{j \mid x_j \in \Omega_G} (A_G)_{ji}$$

and $D_G^O(i, \Omega_G)$ is defined as

$$D_G^O(i, \Omega_G) = \sum_{j \mid x_j \in \Omega_G} (A_G)_{ij}.$$

In a straight forward way these concepts can be generalized to their degree distributions. In particular given $H_G^1, H_G^2 \subseteq V(G)$ are defined $P(D_G^I(H_G^1, H_G^2) = k)$ and $P(D_G^O(H_G^1, H_G^2) = k)$, and then, defined $\Omega_G^1, \Omega_G^2 \subseteq \mathbb{R}^2$ are defined $P(D_G^I(\Omega_G^1, \Omega_G^2) = k)$ and $P(D_G^O(\Omega_G^1, \Omega_G^2) = k)$.

1.5 Clustering Coefficient

The concept of Clustering Coefficient is based on the idea that similar nodes are more connected among them (Easley and Kleinberg, 2010). These clusters of nodes are usually called communities, and finding them in the network analysis can have a crucial role in many fields (Easley and Kleinberg, 2010; Barrat, Barthélemy, and Vespignani, 2008; Steen, 2010). The key idea in finding clusters is that take triplets of nodes with at least two edges if many of these triplets have three edges, so the node is strongly clusterized with its neighborhood (Steen, 2010). This feature is represented by a quantity called clustering coefficient, defined for a node i as

$$Cc_i = \frac{\text{number of triplets with all the nodes connected involving the node } i}{\text{number of triplets with two nodes connected with } i \text{ involving } i}$$

More formally can be defined as (Wang et al., 2017)

Definition 1.5.1 *For a node i (with two or more connections) can be defined the clustering coefficient Cc_i as*

$$Cc_i = \frac{\sum_{j,k \in V(G)} (A_G^S)_{ij}(A_G^S)_{jk}(A_G^S)_{ki}}{\sum_{j \in V(G)} (A_G^S)_{ij} \left(\left(\sum_{j \in V(G)} (A_G^S)_{ij} \right) - 1 \right)}$$

where

$$(A_G^S)_{ij} = \begin{cases} 1, \text{if } (i,j) \in E(G) \vee (j,i) \in E(G) \\ 0, \text{otherwise} \end{cases}$$

With this definition can be computed for example the clustering coefficient of figure 1.2 obtaining that node 1 has just one connection and then is not possible calculate clustering coefficient, for the other nodes $Cc_2 = 0$, $Cc_3 = 0$, $Cc_4 = \frac{1}{3}$, $Cc_5 = \frac{1}{3}$ and $Cc_6 = 1$. This result reflects the fact that nodes 4, 5 and 6 create a large cluster that is in contrast with the sparsity of the connections of the nodes 1, 2 and 3. From these local clustering coefficients can be calculated a global clustering coefficient averaging on the valid coefficients.

The clustering coefficient has a crucial role in many fields of MGT; however, this quantity will not be used extensively for two main reasons:

- There are results suggesting that mean field activity of a large-scale neuronal network can be approximated in terms of mean fields only by knowing its

degree distributions (Nykamp et al., 2017). For this reason, the analysis of the degree distributions will be preferred to the calculation of the clustering coefficient.

- The calculation of clustering coefficient (acting on triples) can be computationally expensive for large scale networks.

1.6 Network motifs

The idea behind network motifs is that exist recursive patterns inside real world networks able to model whole network activity (Gal et al., 2017). In particular, the 2-motifs are all the patterns composed of two nodes (then two nodes disconnected, two nodes connected and two nodes with a double connection) and the 3-motifs all the patterns composed by a set of three nodes (that are in total 13 (Milo et al., 2002)). However, even if in many fields of Graph Theory they have proved to be an essential quantity to take into account, they will not be taken into account for the following reasons:

- For large-scale networks has been proved that, under certain conditions of independence of the degree distributions(Nykamp et al., 2017), their number of occurrences can be approximated from degree distributions; therefore, the study of degree distributions will be preferred.

- Their counting (particularly the 3-motifs that act on triples) can be computationally expensive for large-scale networks.

1.7 Path measures

The path concept is one of the key concepts of classic graph theory and, in general, one of the more natural concepts of graph theory. Given a graph G an observer could ideally be placed in a general node i. This observer is able to walk on the edges (considering as only possible direction of crossing from starting node to ending node) and then is asked to the observer to reach the node j through the graph. For example, the observer could be placed in node 4 of the graph of figure 1.2 and the task could be to reach node 2. Then observer could follow the path $4 \to 2$, but another possible path is $4 \to 6 \to 5 \to 3 \to 2$.

Definition 1.7.1 *A path starting from node i and ending to node j can be defined as tuple*

$$(k_0 = i, k_1, k_2, ..., k_{l-1}, k_l = j)$$

such as for each $i = 0, ..., l - 1$ holds $(k_i, k_{i+1}) \in E(G)$. Moreover the number l is called length of the path.

Then the paths $(4, 2)$ and $(4, 6, 5, 3, 2)$ are valid paths from 4 to 2, but the first path has length 1 and the second path has length 4, then the first path is shorter than the first. In general, it is possible to define the shortest length between two nodes as the length of the shortest path possible between two nodes. If a path between two nodes does not exist, this value is supposed to be $+\infty$. Then can be defined (Dorogovtsev and Mendes, 2003)

Definition 1.7.2 *The Average path length is the average of the finite shortest paths length (so $< +\infty$) between the nodes of a network.*

The average path length is helpful to determine the small-worldness of a network (Easley and Kleinberg, 2010) that is, the capacity of a network to connect many nodes with short paths. However, to compute the shortest path between two nodes is a problem solved by ad hoc algorithms (for example, Dijkstra's algorithm and Bellman-Ford algorithm (Steen, 2010)); however, they can be computationally expensive for large networks.

Moreover, they require the full connectivity of the network to be computed because, taking into account or not some nodes can significantly affect the length of the paths. In fact, if there are just two disjoint paths from the node i to the node j, when just a node is not available from the shorter path then the shortest path becomes the longer path. In real world networks is rare to have a perfect and complete knowledge of the network. For these reasons in Modern Graph Theory, the estimation of the small-worldness of the network is made by analyzing the tail of the degree distributions in search of highly connected nodes (*the hubs* (Barabási and Albert, 1999)). The underlying idea is that if there are many hubs then is likely that two random nodes are connected by a short path passing through highly connected nodes.

1.8   Eigenvectors centrality

The idea behind Eigenvectors centrality is that defining $\underline{z}_0$ the column vector such as $(\underline{z}_0)_i = 1$, for $i = 1, ..., N_G$ and

$$(A_G^S)_{ij} = \begin{cases} 1, \text{if } (i,j) \in E(G) \vee (j,i) \in E(G) \\ 0, \text{otherwise} \end{cases}$$

can be defined the product

$$\underline{z}_1 = A_G^S \, \underline{z}_0$$

and in general

$$\underline{z}_n = A_G^S \, \underline{z}_{n-1} = (A_G^S)^2 \, \underline{z}_{n-2} = ... = \left(A_G^S \right)^{n-1} \underline{z}_1 = \left(A_G^S \right)^n \underline{z}_0$$

for $n > 0$. The property of this product is that, for large values of n, the more is connected the node i the more the component i-th of the vector (that is $(\underline{z}_n)_i$) is bigger in relation to the other components (Newman, 2010).

Being A_G^S symmetric and in virtue of spectral theorem (Axler, 1997) hold the following:

- the matrix A_G^S is diagonalizable;

- all Eigenvalues are all real;

- exists an eigenvector basis;

Then can be defined $\lambda(G)_j$, for $j = 1, ..., N_G$, the eigenvalues of A_G^S such as

$$|\lambda(G)_1| \geq |\lambda(G)_1| \geq ... \geq |\lambda(G)_{N_G-1}| \geq |\lambda(G)_{N_G}|.$$

Supposing $v(G)_i$ the eigenvector related to the eigenvalue $\lambda(G)_i$, then must exist v_i such as $\underline{z}_0 = \sum_{i=1}^{N_G} v_i v(G)_i$.

Then can be written

$$\underline{z}_1 = A_G^S \, \underline{z}_0 = \sum_{i=1}^{N_G} v_i A_G^S v(G)_i = \sum_{i=1}^{N_G} v_i \lambda(G)_i v(G)_i = \lambda(G)_1 \sum_{i=1}^{N_G} v_i \left(\frac{\lambda(G)_i}{\lambda(G)_1} \right) v(G)_i$$

more in general

$$\underline{z}_n = \lambda(G)_1^n \sum_{i=1}^{N_G} v_i \left(\frac{\lambda(G)_i}{\lambda(G)_1} \right)^n v(G)_i.$$

Most of applications holds the empirical assumption (Easley and Kleinberg, 2010; Newman, 2010)

$$|\lambda(G)_1| > |\lambda(G)_2| > ... > |\lambda(G)_{N_G-1}| > |\lambda(G)_{N_G}|.$$

and observing that $\lim_{n \to +\infty} \left(\frac{\lambda(G)_i}{\lambda(G)_1} \right)^n = 0$, for $\frac{\lambda(G)_i}{\lambda(G)_1} < 1$, then

$$\underline{z}_{+\infty} = \lim_{n \to +\infty} \underline{z}_n \; // \; v(G)_1$$

therefore the vectors $\underline{z}_{+\infty}$ and $v(G)_1$ are proportional.

In conclusion $\underline{z}_{+\infty}$ is the solution of the equation

$$A_G^S \, \underline{z}_{+\infty} = \lambda(G)_1 \, \underline{z}_{+\infty}.$$

In literature have been presented many variations of this equation, the most representative (Newman, 2010) are Katz equation, that is

$$A_G^S \, \underline{z}_{+\infty} = \alpha_{Katz} \lambda(G)_1 \, \underline{z}_{+\infty} + \beta_{Katz} \mathbb{1}.$$

where $\alpha_{Katz}, \beta_{Katz} \in \mathbb{R}$ and $\mathbb{1}$ is a column vector of all ones, and another version is PageRank equation

$$A_G \, \underline{z}_{+\infty} = \alpha_{PR} D_{PR}^{-1} \lambda(G)_1 \, \underline{z}_{+\infty} + \beta_{PR} \mathbb{1}.$$

where $\alpha_{PR}, \beta_{PR} \in \mathbb{R}$, $\mathbb{1}$ is a column vector of all ones and D_{PR} is the diagonal matrix with diagonal elements $(D_{PR})_{ii} = max \left(D_G^O(i), 1 \right)$. In particular the PageRank equation is used in the search engine Google (Page et al., 1999).

However Eigenvalues decomposition can be computationally expensive for a general large-scale network (Pan, Chen, and Zheng, 1998), in fact in search engines this task is solved computing $\underline{z}_n$ until it reaches the convergence (Page et al., 1999). For these reasons to the Eigenvalues analysis (the so called spectral analysis) will be preferred the tail analysis of the degree distributions of the networks (Barabási and Albert, 1999).

1.9   Spatial quantities

Since now has been mostly discussed, quantities can be calculated taking into account just the Adjacency matrix. Each node i has precise spatial information $\underline{x}_i$, where the vector position has a fixed dimension d_G that will be called graph spatial dimension. For example in the graph in figure 1.2 $d_G = 2$.

One of the most used spatial invariants is the Euclidean distance induced by Euclidean norm:

Definition 1.9.1 *Given a vector $\underline{z} = \left((\underline{z})_1, ..., (\underline{z})_{d_G} \right) \in \mathbb{R}^{d_G}$ the Euclidean norm of the vector $\underline{z}$ is defined as*

$$\|\underline{z}\| = \sqrt{\sum_{d=0}^{d_G} \left((\underline{z})_d \right)^2}$$

Definition 1.9.2 *The Euclidean distance between the two nodes i and j is defined as*

$$d(i,j) = \left\| \underline{x}_i - \underline{x}_j \right\|$$

However this concept can be generalized. For example suppose the function $L_{ij} = L(i,j) = \left\| \underline{x}_i - \underline{x}_j \right\|^2$ that is a general function defined on spatial positions of nodes. In general

Definition 1.9.3 *Given $f_L : \mathbb{R}^{d_G} \times \mathbb{R}^{d_G} \to \mathbb{R}$ can be defined the spatial function L on nodes of G as*

$$L_{ij} = L(i,j) = f_L \left(\underline{x}_i, \underline{x}_j \right)$$

It can be noted that the simplest invariant involving distance is the Connection Lengths Distribution.

Definition 1.9.4 *The Connection Lengths Distribution is the distribution of the random variable C_L such as for each $d, \epsilon > 0$*

$P \left(d \leq C_L \leq d + \epsilon \right) = $ *probability of picking a connection of G with length in range $[d, d+\epsilon]$*

Another used measure is the *probability of connection dependent from the distance* (Potjans and Diesmann, 2012)

Definition 1.9.5 *The Probability of Connection dependent from distance is a function such as for each $d > 0$*

$P_C(d) = $ *probability that two general nodes at distance d are connected*

1.10 Software implementation

Network analysis is a continuously expanding field of Graph Theory, and most of the time nowadays is performed using computers. There are two main approaches in network analysis (Finotelli, 2015):

- Visual approach: the graph is plotted and is performed visual inspection to find the key features of the network. This approach is particularly suitable for small networks that are easy to draw. Programs that perform this task are, for example, *Gephi* and *Cytoscape*;

- Quantitative approach: the graph is analyzed through quantitative measures (like degree distributions, connection length distribution, etc.) able to depict network connectivity without drawing the whole graph. That approach will be preferred because it is very suitable for large-scale networks. Software that performs this task are the Matlab toolboxes *Brain Connectivity Toolbox* and *Network Analysis package* and the program *igraph*.

However, previous software are thought for macro connectivity Connectomes (Bullmore and Sporns, 2009), networks that (representing connectivity between brain areas) are made up by order of hundreds of nodes (Farahani, Karwowski, and Lighthall, 2019). In contrast, this contribution will analyze networks with more than 30 thousand nodes (Markram et al., 2015), and this will require software highly optimized to tackle these problems, and in many cases, this software will be written in Python (*Python Language Reference, version 3.6* 2016), or Matlab (*MATLAB v. 2021b* 2021).

1.11   Computational representation of networks

In the previous section has been shown that graph connectivity is entirely determined by adjacency matrix and then has been calculated the adjacency matrix for the graph in figure 1.2. Then could be a good idea to store this matrix as an array. In the case of figure 1.2 it will be an array with shape 6×6 of binary values. Because in most of the programming language binary values are not represented by 1 bit but from 1 byte (= 8 bit) or more to be stored, the size of memory allocated by the graph of figure 1.2 will be at least $6^2 = 36$ bytes. In general

Theorem 1.11.1 *For a general graph G the memory required to allocate its adjacency matrix is proportional to* $(N_G)^2$.

However, to save in memory the adjacency matrix, even if it is the simplest way to store graph connectivity, it is not always the best choice. Suppose to add 4 nodes (then nodes 7,8,9 and 10) with no connections to graph in figure 1.2 creating the graph G'. Then its adjacency matrix will become

$$A_{G'} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

It is clear that all the connections are concentrated in the first 6 rows and columns (being outside this submatrix each element 0). However, the memory allocated to store this new adjacency matrix is more significant than before and proportional to $(N_G')^2 = 10^2 = 100$. Therefore, the matrix presented is a sparse matrix.

Definition 1.11.1 *A sparse matrix is a matrix with few non-zero elements.*

Most of the experimental large-scale networks have an adjacency matrix sparse (Markram et al., 2015; Barabási and Albert, 1999). If a graph has an adjacency matrix sparse (it has few connection between nodes) then the optimal strategy is to construct the connection list

Definition 1.11.2 *Given a graph $G = (V(G), E(G))$ and the number $|E(G)|$ the number of its edges, the Connections List is the matrix $E_L(G)$ with shape $|E(G)| \times 2$ where the rows are the edges in $E(G)$.*

For example

$$E_L(G) = \begin{bmatrix} 1 & 2 \\ 3 & 2 \\ 4 & 2 \\ 4 & 6 \\ 5 & 3 \\ 5 & 4 \\ 6 & 5 \end{bmatrix}$$

but holds also

$$E_L(G') = \begin{bmatrix} 1 & 2 \\ 3 & 2 \\ 4 & 2 \\ 4 & 6 \\ 5 & 3 \\ 5 & 4 \\ 6 & 5 \end{bmatrix}$$

because the nodes added have no connections. The Connection lists then can be stored in an integer array of size $2 \times 7 = 14$. In general

Theorem 1.11.2 *For a general graph G the memory allocated to store its Connections List is proportional to* $2\,|E(G)|$.

Then in case of figure 1.2 even if the memory used to allocate an integer is 2 bytes the memory used to allocate the Connections List is 28 bytes for both graphs G and G' less than the scenario of $A(G)$ (36 bytes) and of $A(G')$ (100 bytes).

However the advantage of storing Connections List is more visible at large scales. First of all

Definition 1.11.3 *The sparsity of a general graph G is defined as*

$$\sigma_s(G) = 1 - \frac{|E(G)|}{(N_G)^2} \in [0,1]$$

In particular $\sigma_s(G) = 1 - \frac{7}{36} \approx 0.81$ and $\sigma_s(G') = 1 - \frac{7}{100} \approx 0.93$. In general more σ_s is closer to 1 more the matrix is sparse. For experimental Connectome reconstructions this quantity range from $1 - 0.1 = 0.9$ (Potjans and Diesmann, 2012) to $1 - 0.01 = 0.99$ (Markram et al., 2015).

Theorem 1.11.3 *The memory to allocate the Connections List for a large-scale network is proportional to* $2\,|E(G)| = 2\,(1 - \sigma_s(G))\,(N_G)^2$.

This means that for $(1 - \sigma_s(G))$ very small and large N_G then $2\,(1 - \sigma_s(G))\,(N_G)^2 << (N_G)^2$, showing the advantage of allocating the Connections List instead of adjacency matrix.

1.12 Degree calculation and degree distributions

Calculating the degree (indegree or outdegree) will have a crucial role in the proposed analysis. For this reason, it is debated about the degrees defined for the indegree

$$D_G^I(i) = \sum_{j \in V(G)} (A_G)_{ji}$$

and for the outdegree

$$D_G^O(i) = \sum_{j \in V(G)} (A_G)_{ij}$$

it can be observed that the indegree is the sum of the elements on columns of adjacency matrix and the outdegree is the sum of the elements on rows of the adjacency matrix. This means that the number of sums to calculate degree with this approach is $N_G - 1$ for each row/column for the N_G columns/rows, then $N_G (N_G - 1)$.

However the number of sums $N_G (N_G - 1)$ means that the number of sums (and then the amount of time) used to calculate the degree of a fully connected graph is the same of a graph with no connection. Being real world large-scale networks most of times sparse, it is not a good approach. Instead calculating degree distributions from connections list has the cost of $|E(G)|$, because it is sufficient to count the occurencies of nodes in first column for the outdegree and in second column for the indegree. We have already observed that

$$|E(G)| = (1 - \sigma_s(G)) (N_G)^2 << N_G (N_G - 1)$$

for sparseness $\sigma_s(G) \to 1$ and large numbers of nodes N_G, concluding that the connections list structure for graph storing it is not only efficient in terms of memory used, but has an efficient algorithm for calculating degrees. For this reason will be widely used in large-scale networks analysis.

Calculated the degree set, it is possible to calculate the degree distributions, that are the distributions $P(D_G^I = k)$ and $P(D_G^O = k)$. The easiest way to calculate it is counting the number of nodes for each possible degree and then to calculate each probability. More formally

Definition 1.12.1 *Defined the indegree frequency function* $\left(F_G^I\right)_k$ *as the function defined as*

$$\left(F_G^I\right)_k = \text{number of nodes with indegree } k$$

the empirical indegree distribution can be calculated as

$$\hat{P}_E(D_G^I = k) = \frac{\left(F_G^I\right)_k}{N_G}.$$

Simetrically empirical outdegree distribution is

$$\hat{P}_E(D_G^O = k) = \frac{\left(F_G^O\right)_k}{N_G}.$$

However this definition has some problems when applied to real-world networks. Indeed some experimental networks can have significant oscillation in their degree by degree distributions, for example in figure 1.4 is shown the indegree distribution from (Markram et al., 2015). In order to avoid these oscillations, is defined the probability density function degree distribution.

Definition 1.12.2 *The probability density function indegree distribution is calculated starting from a bin* $\beta \in \mathbb{N}$ *with the formula*

$$\hat{P}_\beta(D_G^I = k) = \frac{\sum_{q=(m-1)\beta}^{m\beta-1} \left(F_G^I\right)_q}{\beta N_G}, \text{ if } k \in [(m-1)\beta, m\beta[\text{ for some } m \in \mathbb{N}^+.$$

Simmetrically the probability density function outdegree distribution is

$$\hat{P}_\beta(D_G^O = k) = \frac{\sum_{q=(m-1)\beta}^{m\beta-1} \left(F_G^O\right)_q}{\beta N_G}, \textit{if } k \in [(m-1)\beta, m\beta[\textit{ for some } m \in \mathbb{N}^+.$$

From the previous definition can be observed that the empirical indegree distribution is a particular case when the bin value $\beta = 1$ (then $\hat{P}_E = \hat{P}_1$). This generalization can be quite important because most of the theory proposed here will result in quite smooth distributions, in contrast to what is observed in experimental findings (Markram et al., 2015). Therefore, to smooth the degree distributions found, most of the time, the values of $\beta > 1$. In fact as can be seen in figure 1.4 Top-Left panel, if $\beta = 1$ the degree distribution is noisy. If $\beta = 5$ (Top-Right) or $\beta = 25$ (Bottom-Left) the histogram is smooth enough. However, if $\beta = 300$ (Bottom-Right), the histogram is over averaged and most of the degree distribution information is lost. This simple example makes evident the key role of bin size β in calculating degree distributions for real-world networks. Put in evidence that, in general, there is no standard strategy to choose the right value of β, because if β is too low the distribution is noisy, instead if it is too high most of the detail can be lost. Since now, it will be assumed that for each degree distribution shown in a figure, a reasonable value of β has been chosen.
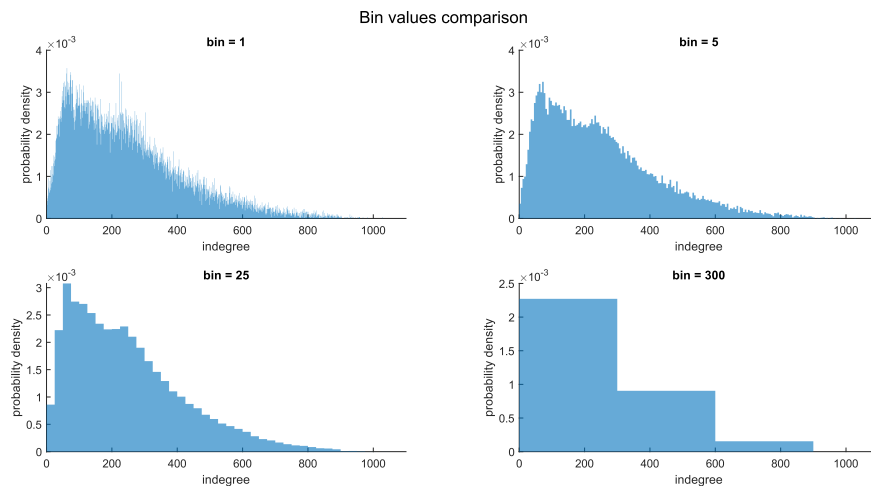
FIGURE 1.4: Indegree distribution from Neocortical model (Markram et al., 2015) with bins respectively 1, 5, 25 and 300. Can be observed that for the bin size 1 the histogram is noisy. For bins sizes 5 and 25 the histogram is correctly smoothed. For bin size 300 the histogram is too averaged.

1.13 Introduction to Modern Graph Theory

The first step in the direction of Modern Graph Theory was the ER (Erdös and Rényi, 1959).

Definition 1.13.1 *The ER model with $N(ER) = N_{ER}$ nodes and probability p_{ER} is a graph constructed connecting every couple of nodes independently with a probability p_{ER}.*

This simple model has an incredible property, if $p_{ER} \in]0, 1[$, every graph with N_{ER} nodes is the possible outcome of an ER model with N_{ER} nodes and probability p_{ER}.

Then it is legit to think that the ER model is a universal graph generator and the solution to every problem in Graph Theory. In order to understand that is not the case, suppose $p_{ER} = 10^{-1}$ and $N_{ER} = 4$, then the probability of obtaining a fully connected graph (a graph with all the values in adjacency matrix ones, except the diagonal) from such ER model is p_{ER} raised to the number of possible edges (that is $N_{ER}(N_{ER} - 1) = 12$), and then it is

$$p_{ER}^{N_{ER}(N_{ER}-1)} = 10^{-12}.$$

In order to comprehend this probability should be noted that the probability to win the multi millionaire jackpot at the Mega Millions lottery is order of 10^{-9} (*How to Play Mega Millions Lottery* 2021). This simple example makes clear why a new approach is required, in order to avoid situations where buying a lottery ticket is a better idea than trying to fit experimental data with the wrong model. This new approach is the Modern Graph Theory.

1.14 ER Model

The ER model defined above was the first random graph model (Erdös and Rényi, 1959), a model able to generates graphs randomly with the same algorithm. This model can be expressed in terms of random variables (M Evans and Peacock, 2000), but will be required some definition:

Definition 1.14.1 *A Bernoulli random variable X_p with probability $p \in [0, 1]$ is a random variable with distribution $\chi_p(k)$ defined as*

$$\chi_p(k) = \begin{cases} (1 - p), \text{ if } k = 0 \\ p, \text{ if } k = 1 \\ 0, \text{ otherwise} \end{cases}$$

The Bernoulli random variable describes the act to flip a coin that gives the value 1 with probability p and 0 with probability $1 - p$. Then in order to construct the ER model ER with N_{ER} nodes and probability p_{ER}, the adjacency matrix will be defined as

$$(A_{ER})_{ij} \sim X_{p_{ER}}.$$

For a general node i holds

$$D_{ER}^{I}(i) = \sum_{j \in V(ER)} (A_{ER})_{ji} \sim \sum_{j \in V(ER)|i \neq j} X_{p_{ER}} \sim B(N_{ER} - 1, p_{ER})$$

where the random variable is the Binomial random variable (M Evans and Peacock, 2000) defined as

Definition 1.14.2 *Given a number N and a probability $p \in [0, 1]$ the Binomial random variable $B_p^N(k)$ is the random variable with distribution $B(N, p)$ defined as*

$$B_p^N(k) = \binom{N}{k} p^k (1 - p)^{N-k}.$$

Then holds

$$P\left(D_{ER}^{I} = k\right) = B_{p_{ER}}^{N_{ER}-1}(k).$$

Symmetrically can be proved that

$$P\left(D_{ER}^{O} = k\right) = B_{p_{ER}}^{N_{ER}-1}(k).$$

This simple fact shows as even if the ER model could have as outcome any *possible* graph with N_{ER} nodes and consequently any possible degree distribution, there is just one *likely* degree distribution that it can assume and it is a Binomial bell. This simple fact in the applications implies that the degree distributions of an ER model is assumed to be always a Binomial bell even if there is the infinitesimal probability to obtain a fully connected graph (see figure 1.5).
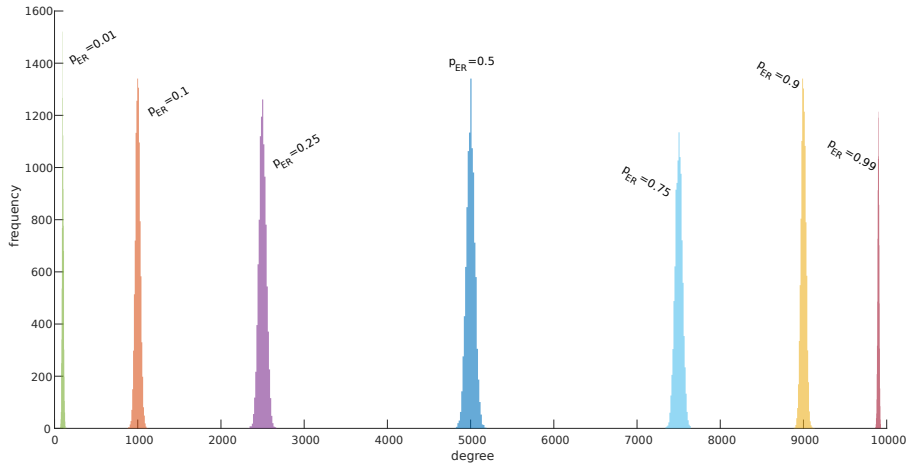
FIGURE 1.5: ER models with $N_{ER} = 10^4$ and p_{ER} respectively (from left to right) $0.01, 0.1, 0.25, 0.5, 0.75, 0.9$ and 0.99.

This model is composed by two independent parameters N_{ER} and p_{ER}, however in the applications the number of nodes is fixed by experimental constrains, then the only free parameter is the probability p_{ER}. In order to fit this general probability p_{ER} must be calculated the expected number of edges

$$\mathbb{E}\left(|E_{ER}|\right) = \mathbb{E}\left(\sum_{i,j\in V(ER)|i\neq j} (A_{ER})_{ij}\right) = \mathbb{E}\left(B(N_{ER}(N_{ER}-1), p_{ER})\right) = N_{ER}(N_{ER}-1)p_{ER}$$

then in general holds

$$p_{ER} = \frac{\mathbb{E}\left(|E_{ER}|\right)}{N_{ER}(N_{ER}-1)}.$$

From a general experimental network G than can be computed the quantity

$$\rho_G = \frac{|E_G|}{N_G(N_G-1)}$$

defined as the *probability of connection* of the graph G and the best ER model ER to fit the graph G has parameters $N_{ER} = N_G$ and $p_{ER} = \rho_G$.

1.15   The ER mixed model

The key idea behind the ER mixed model is the concept of an inhomogeneous network. An inhomogeneous network is a network where there are significant differences between nodes. In particular, the central hypothesis of ER mixed model is that exist many distinct populations with a different constant probability of connections between themselves (Potjans and Diesmann, 2012). The result (will be proved when will be introduced the convolutive approach) is that the ER mixed distributions are a convex combination of Binomial bells like distributions (Giacopelli, Migliore, and Tegolo, 2020). In practice, ER mixed models appear with distributions with several peaks less equal than the number of populations. This model is viral because to fit it is sufficient to estimate the probability of connections between populations as described in the previous section. This model is involved in the theory that claims to compute the probability of connection between populations is enough to reconstruct a network fully. However, in the sections below, it will be shown that it is not the case in general.

1.16   Power Law models

In 1999 has been introduced a new model in MGT: the Barabasi-Albert (Barabási and Albert, 1999) (since now BA). This model starts from the assumption that in social networks (Easley and Kleinberg, 2010) and in computer networks (Barabási and Albert, 1999) can be observed that the more a node is connected more is likely that this node will create new connections. This simple concept has been applied for example in search engine Google (Page et al., 1999), where a node with a large number of incoming connections (hyperlinks pointing the web page) is called an *authority* of the network. In fact the principle is that if many pages points to a specific page there is in high probability that this page has a superior probability to be safe, trustfully and it is more probable that the Google user is looking for this page. Barabasi and Albert have observed that such dynamics can be observed in a general computer network (Barabási and Albert, 1999) and in particular in such networks the degree distributions exhibit a power law tail (Giacopelli, Migliore, and Tegolo, 2020). They have also proposed a model able to produce such networks that is the BA model, and (following (Newman, 2010)) first of all will be introduced the concept of growing network and preferential attachment:

Definition 1.16.1 *A growing network* $\{G(0), G(1), ..., G(t), ...\}$ *is a sequence of graphs where each graph* $G(i + 1)$ *is created adding a node to the network* $G(i)$.

Definition 1.16.2 *A growing network* $\{G(0), G(1), ..., G(t), ...\}$ *is constructed using the preferential attachment if for each new node i added at the network* G_t *the new node has a probability to be connected to an old node* $j \in V(G(t))$

$$\frac{D^I_{G(t)}(j) + a}{\sum_{v \in V(G(t))} \left(D^I_{G(t)}(v) + a \right)}$$

with $a > 0$ *a constant.*

Definition 1.16.3 *A BA model is a preferential attachment growing network with parameter* $a > 0$ *and where every new node is connected to c old nodes.*
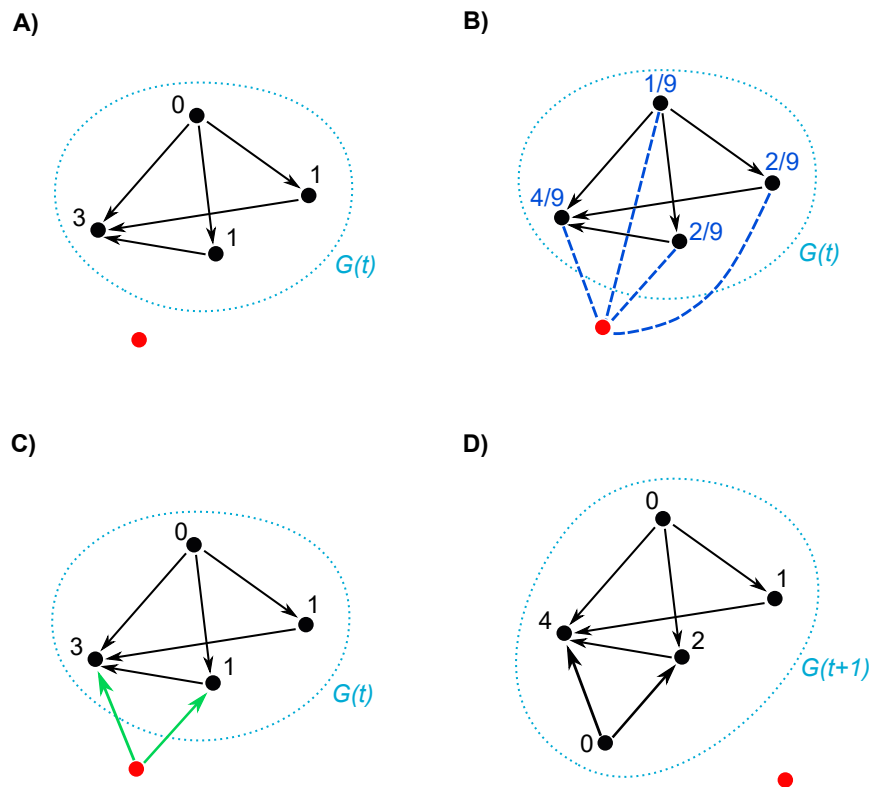
FIGURE 1.6: The process of growing network with preferential attachment in BA model ($a = 1$ and $c = 2$). A) The network at time t has the connectivity $G(t)$ (black nodes). The red node is the new node. For every black node is calculated the indegree (black numbers). B) For each black node is computed the probability of connection using the formula of preferential attachment (blue numbers). C) Following the probabilities of connection are chosen c old nodes and they are connected to the new node (green arrows). D) The new graph created is now called $G(t+1)$ and the routine can be repeated until the desired number of nodes has been reached.

In figure 1.6 is shown a schematics of the construction process. Dorogostev *et al.* (Dorogovtsev and Mendes, 2003) have proved, in an elegant way, that a BA model has a distribution with a power law tail. We will follow the proof in (Newman, 2010) that simplify the exposition:

Theorem 1.16.1 *A BA model with parameters a and c follows asymptotically the expression*

$$\lim_{t \to +\infty} P(D^I_{G(t)} = k) \approx k^{-2 + \frac{a}{c}}$$

Proof: Defining $p_k(t) = P(D^I_{G(t)} = k)$, $k_i(t) = D^I_{G(t)}(i)$ and $n(t) = N_{G(t)}$. Then the preferential attachment can be rewritten as

$$\frac{k_i(t) + a}{\sum_i (k_i(t) + a)} = \frac{k_i(t) + a}{n(t)\mathbb{E}(k(t)) + n(t)a} = \frac{k_i(t) + a}{n(t)(c + a)}$$

where $c = \mathbb{E}(k(t))$ is true because the expected value of indegree and outdegree are equal. The number of nodes with indegree k is $n(t)p_k(t)$ then the expected number of new edges is

$$n(t)p_k(t) \, c \, \frac{k + a}{n(t)(c + a)}$$

When a new node is added to $G(t)$, the nodes with indegree k increases by one each node of indegree $k - 1$ that receives a new connection, this node becomes a node of indegree k

$$\frac{c(k - 1 + a)}{c + a} \, p_{k-1}(t)$$

similary will be lost one node of degree k every time that it becomes of degree $k + 1$

$$\frac{c(k + a)}{c + a} \, p_k(t)$$

then can be written the master equation

$$n(t+1)p_k(t+1) = (n(t) + 1)p_k(t+1) = n(t)p_k(t) + \frac{c(k - 1 + a)}{c + a} \, p_{k-1}(t) - \frac{c(k + a)}{c + a} \, p_k(t).$$

However the previous equation holds for $k > 0$. $k = 0$ is a special case, because there is not a node that can become of degree 0 (Newman, 2010)

$$(n(t) + 1)p_0(t+1) = n(t)p_0(t) + 1 - \frac{ca}{c + a}p_0(t).$$

The hypothesis of master equation is that exists and it is finite the limit

$$\lim_{t \to +\infty} p_k(t) = p_k(\infty) = p_k.$$

Then the previous equations become for $k > 0$

$$p_k = \frac{c}{c + a}[(k - 1 + a)p_{k-1} - (k + a)p_k]$$

and for $k = 0$

$$p_0 = 1 - \frac{ca}{c + a}p_0.$$

The equation for $k = 0$ can be rearranged

$$p_0 = \frac{1 + a/c}{a + 1 + a/c}$$

and for $k > 0$

$$p_k = \frac{k + a - 1}{k + a + 1 + a/c} \, p_{k-1}.$$

Solving the recursive equations can be derived the solution

$$p_k = \frac{B(k + a, 2 + a/c)}{B(a, 1 + a/c)}$$

where $B(x, y)$ is the Beta function (Olver et al., 2010). In (Newman, 2010) is observed how is valid the asymptotic form

$$p_k = \frac{B(k + a, 2 + a/c)}{B(a, 1 + a/c)} \approx k^{-2 + \frac{a}{c}}.$$

$\square$

However the BA model proposed has the feature that the outdegree of each new node added is constant and equal to c. For this reason has been introduced the Price model

Definition 1.16.4 *The Price model is preferential attachment growing network with parameter $a > 0$ where for each iteration t is added a node with c_t connections with the old nodes.*

In (Newman, 2010) is observed that doing computational simulations still holds the identity

$$\lim_{t \to +\infty} P(D^I_{G(t)} = k) \approx k^{-2 + \frac{a}{c}}$$

where $c = \mathbb{E}(c_t)$. Often instead of defining the c_t sequence of the Price model, is defined a random variable Γ such that $\mathbb{E}(\Gamma) = c$ and then the c_t are defined as possible outcomes of Γ. In figure 1.7 can be seen an example of two Price models with different Γ distributions. The advantage of the last approach is shown in the following theorem

Theorem 1.16.2 *Given a Price model with parameter $a > 0$ and random variable Γ, the outdegree distribution can be approximated by the distribution of Γ.*

Proof: Because Γ controls the number of edges added when a new node is added, most of the nodes outdegree of the network are occurencies of Γ. Then the the outdegree distribution almost follows the Γ probability distribution. $\square$

This property has a key role in the usability of the Price model because a possible fitting procedure is to assign the target outdegree distribution to Γ and tune the a parameter to fit at best the indegree distribution. Since now will not be made the difference between the Price model and the BA model, using these two interchangeably terms to refer to the actual Price model. This choice is motivated because even if the actual name of this model is the Price model, most of the power law models are usually called the BA model because it was the first of this kind, and it is the most known power law model.
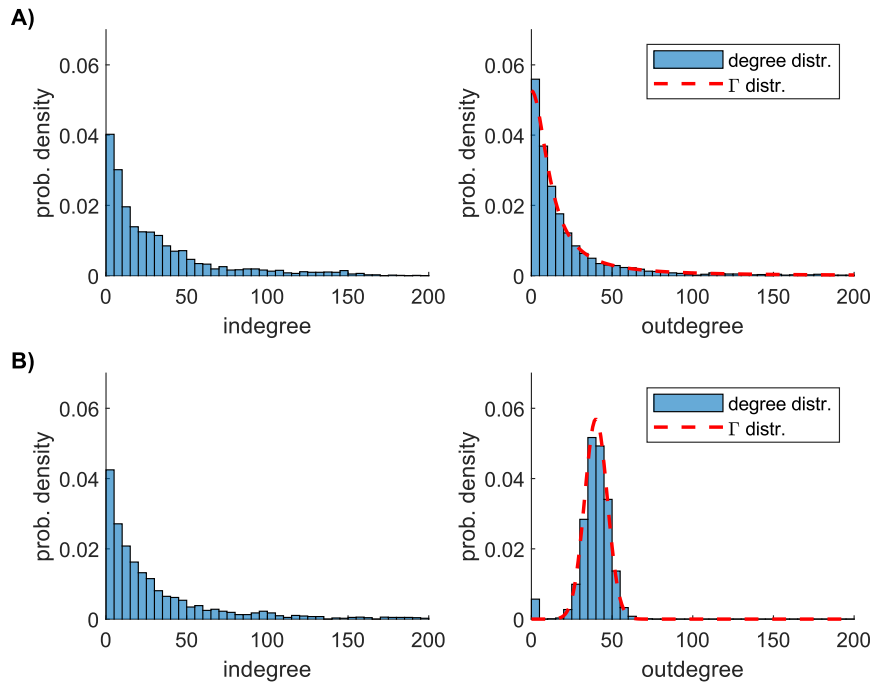
FIGURE 1.7: Degree distributions of two Price models of 2500 nodes with different Γ distributions but same $c = 40$ and $a = c$. A) The Γ distribution follows the distribution $P(\Gamma = k) = \frac{7.6226}{k^2+150}$. B) The Γ distribution is a Gaussian distribution with mean 40 and standard deviation 7 (Olver et al., 2010).

1.17   Small-World networks and Small-Communities

An experiment in 1960 (Easley and Kleinberg, 2010; Newman, 2010) made by Stanley Milgram has shown that real-world social networks exhibit a small average shortest path length. The experiment gave 96 packets to 96 voluntaries with an address and a name with the instruction to not send directly by mail the load but to send this packet to a friend who had a better chance to bring this packet to destination until the packet reaches the destination. Of 96 packets, just 18 arrived at the destination, but these packets made just 5.9 steps to arrive. This experiment is often cited as *six degrees of separations* and has given the input to a series of experiments to quantify the shortest path length between nodes in social networks (in the case of this experiment was about 6). Networks as the BA model and the ER model are small-world (Newman, 2010).

However, another feature in social networks is the creation of small communities, which is manifested in terms of networks analysis by the clustering coefficient. For example (Easley and Kleinberg, 2010) explores of the clustering coefficient can be a relevant measure to separate social groups in a school. In general, for models like ER and BA the clustering coefficient is low (Newman, 2010). For this reasons, it has been introduced the Watts-Strogatz model (Watts and Strogatz, 1998). This model starts from a circular lattice which can achieve a clustering coefficient of 3/4. The downside of this configuration is that the average path length increases. For this reason, a random rewiring with a given probability p is applied to make the configuration converge to the ER model, decreasing the average path length (at the cost of decreasing the clustering coefficient). The substantial downside of this model is that,

because of its regularity, it does not produce realistic degree distributions (Newman, 2010); in fact, the regularity of the starting lattice produces a spiky behavior in the degree distributions because many nodes have almost the same degree. However, the previous issue can be avoided by increasing the probability p, but this makes to converge the network to an ER-like configuration that will prove unsuitable for the Connectome. Last assertations have a determinant role in discarding it in our applications because exist contributions (Nykamp et al., 2017) that show the key role of degree distributions in neuronal activity modelling.

1.18 Spatial graphs models

The previous models are aspatial graphs; however, real-world networks are rarely aspatial. The spatial graphs, which will be introduced, follow a common thread: the wiring cost minimization principle.

Definition 1.18.1 *The wiring cost minimization principle states that a real-world network is built trying to minimize the connections length.*

This simple principle translates in graph theory that near nodes are more likely to be connected.

One of the models that implements this principle is the ER distance model (Billeh et al., 2020)

Definition 1.18.2 *An ER distance model is a random graph model where given two nodes i and j, they are connected with a probability dependent from the distance*

$$p_{ij} = \rho_d \left(\|\underline{x}_i - \underline{x}_j\| \right)$$

where $\rho_d : [0, +\infty[\rightarrow [0, 1]$ is a smooth function.

These kind of models often exhibit an exponential decay in ρ_d and can be combined with other geometric dependencies (Billeh et al., 2020). Will be shown with the convolutive framework that (under certain hypothesis) these models have an exponential tail.

Another spatial model is the spatial BA model (Fabrikant, Koutsoupias, and Papadimitriou, 2002). This model is constructed through the minimization of a cost function between nodes

$$C_{ij} = \delta d_{ij} + h_j$$

where d_{ij} is the spatial cost of the model (that can be the euclidean distance between nodes or its square for example) and h_j is a centrality measure respect to the first node of the network. This model will be fundamental for a truly spatial connectivity model. In (Fabrikant, Koutsoupias, and Papadimitriou, 2002) is proved that these kinds of models exhibit a power law behaviour and this theorem has also been proved numerically (Giacopelli, Migliore, and Tegolo, 2020).

Convolutive models

2.1 Observations and reflections about the neurons degree distributions

The brain Connectome is the ensemble of all the connections in the human brain. In particular, the Connectome can be described as a graph by constructing a graph whose nodes are the neurons' soma, and arches are the synaptic connections. The first experimental reconstructions of Connectome have been made through the principle *wire together, fire together* (Bonifazi et al., 2009). This principle states that if two neurons are connected when the pre-synaptic neuron fires, there is a reasonable probability that the post-synaptic will. In virtue of this principle, if an electric current stimulates a neuron, it will start to fire, making fire (and then revealing) all its post-synaptic neurons. However, this simple technique used for a long time is costly in terms of person-hours used to reconstruct the connectivity because all the neurons must be stimulated one by one (Bonifazi et al., 2009). The advances of electron microscopy have made possible a different approach where the connectivity is reconstructed from high-resolution scans of the neurons' cells, making possible reconstructions with more than a thousand neurons (Takemura et al., 2013). The previously proposed models are the connectivity of an actual brain region of a subject. However, relaxing this hypothesis is possible to construct more extensive networks from small sparse connectivity samples, thus using stochastic inference is possible to reach more than 30,000 neurons (Markram et al., 2015).

What can be observed is that experimental networks' degree distributions (shown in figure 2.1) do not follow the ER behavior (Erdös and Rényi, 1959) and either the Power Law behavior (Barabási and Albert, 1999). In particular the experimental distributions exhibit a growing behavior for low degrees (not predicted by BA model) and an heavy tail behavior for high degrees (not predicted by ER model). The situation is schematized in figure 2.1H, where can be seen how ER model is a good approximation of experimental findings for low degrees and the power law model is able to fit the tail of the experimental findings for high degrees. For this reason has been introduced a new theoretical model able to exhibit a growing behavior for low degrees (like in the ER model) and an heavy tail behavior for high degrees (like in Power Law models), this model is the *Convolutive model* (Giacopelli, Migliore, and Tegolo, 2020; Giacopelli, Migliore, and Tegolo, 2020; Giacopelli et al., 2021).

2.2 Convolution and Convolutive equations

The idea behind the Convolutive model is simple. Supposing two independent random variables R_1 and R_2. In order to compute the probability $P((R_1 + R_2) = k)$ there is a theorem able to compute this quantity, but will need a definition first
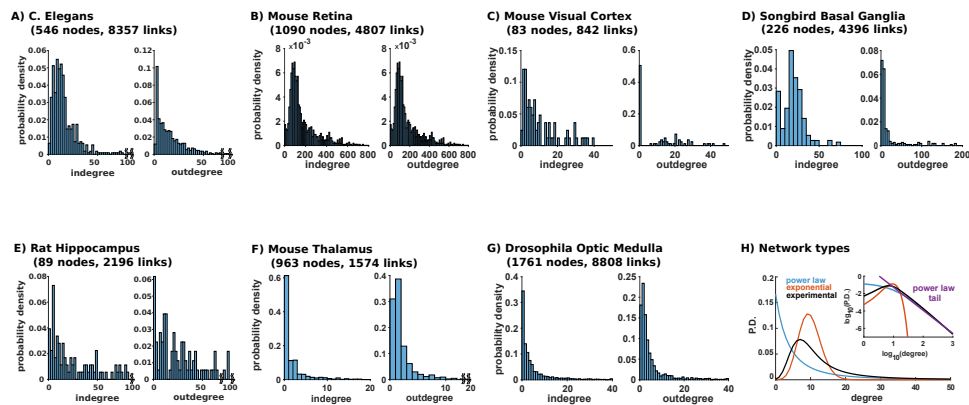
FIGURE 2.1: Degree distributions of experimental brain reconstructions. (A) 546 neurons of the C. elegans (Cook et al., 2019); (B) from a dense reconstruction of 1090 neurons from a mouse retina (Helmstaedter et al., 2013); (C) from electron microscopy data on L2/3 mouse primary visual corte (Turner et al., 2020); (D) 226 neurons from a Songbird basal ganglia (Dorkenwald et al., 2017); (E) 89 neurons from a slice from a rodent hippocampus (Bonifazi et al., 2009) (courtesy of Paolo Bonifazi); (F) mouse thalamus (Morgan et al. (Morgan and Lichtman, 2020)); (G) from electron microscopy data containing 1761 body ID's from a Drosophila (Takemura et al., 2013); (H) Schematic representation of degree distribution for diferent model networks: Power Law (blue), exponential (red) and stereotypical experimentally observed distribution (black). The inset shows a log–log plot with a (purple) line representing the Power Law tail; note how the power law tail line fails to reproduce the low degree connectivity.

Definition 2.2.1 *Given two discrete distributions u_k and v_k the convolution $[u_q * v_q]_k$ between them is defined as*

$$[u_q * v_q]_k = \sum_{q=-\infty}^{+\infty} u_q \, v_{k-q}$$

Then can be stated the Theorem

Theorem 2.2.1 *Supposing R_1 and R_2 two independent random variables then holds*

$$P((R_1 + R_2) = k) = [P(R_1 = q) * P(R_2 = q)]_k$$

Proof: The event $(R_1 + R_2) = k$ is equal to the union of the disjoint events $(R_1 = q) \wedge (R_2 = (k - q))$ for $k \in \mathbb{Z}$. Then can be written

$$((R_1 + R_2) = k) = \bigvee_{q=-\infty}^{+\infty} (R_1 = q) \wedge (R_2 = (k - q))$$

Switching to the probabilities

$$P((R_1 + R_2) = k) = P\left(\bigvee_{q=-\infty}^{+\infty} (R_1 = q) \wedge (R_2 = (k - q)) \right) =$$

using the additivity of probability for disjoint events (Olver et al., 2010)

$$= \sum_{q=-\infty}^{+\infty} P((R_1 = q) \wedge (R_2 = (k - q))) =$$

using the independence of the random variables

$$= \sum_{q=-\infty}^{+\infty} P(R_1 = q) \, P(R_2 = (k - q)) = [P(R_1 = q) * P(R_2 = q)]_k \,.$$

$\square$

Probability theory suggests that

$$\mathbb{E}(R_1 + R_2) = \mathbb{E}(R_1) + \mathbb{E}(R_2)$$

an equivalent result can be proved for convolution:

Theorem 2.2.2 *Given two distributions u_k and v_k holds*

$$\mathbb{E}([u_q * v_q]_k) = \mathbb{E}(u_k) + \mathbb{E}(v_k)$$

Proof:

$$\mathbb{E}([(r_1)_q * (r_2)_q]_k) = \sum_{k=-\infty}^{+\infty} k \sum_{q=-\infty}^{+\infty} u_q v_{k-q} = \sum_{k=-\infty}^{+\infty} \sum_{q=-\infty}^{+\infty} (q + (k - q)) u_q v_{k-q} =$$

substituting $r = k - q$

$$= \sum_{r=-\infty}^{+\infty} \sum_{q=-\infty}^{+\infty} (q + r) u_q v_r = \sum_{r=-\infty}^{+\infty} \sum_{q=-\infty}^{+\infty} q \, u_q v_r + \sum_{r=-\infty}^{+\infty} \sum_{q=-\infty}^{+\infty} r \, u_q v_r =$$

$$= \sum_{q=-\infty}^{+\infty} q\, u_q \sum_{r=-\infty}^{+\infty} v_r + \sum_{r=-\infty}^{+\infty} r\, v_r \sum_{q=-\infty}^{+\infty} u_q =$$

observing that $\sum_{q=-\infty}^{+\infty} u_q = 1$ and $\sum_{r=-\infty}^{+\infty} v_r = 1$, then

$$= \sum_{q=-\infty}^{+\infty} q\, u_q + \sum_{r=-\infty}^{+\infty} r\, v_r = \mathbb{E}(u_k) + \mathbb{E}(v_k)$$

$\square$

The convolution has several important properties

Theorem 2.2.3 *The convolution is bilinear, so that given three distributions u_k, v_k and w_k and two real numbers $\lambda, \mu \in \mathbb{R}$ hold*

$$\left[u_q * (\lambda v_q + \mu w_q) \right]_k = \lambda \left[u_q * v_q \right]_k + \mu \left[u_q * w_q \right]_k$$

and

$$\left[(\lambda v_q + \mu w_q) * u_q \right]_k = \lambda \left[v_q * u_q \right]_k + \mu \left[w_q * u_q \right]_k$$

Proof: Will be proved the linearity on the right component

$$\left[u_q * (\lambda v_q + \mu w_q) \right]_k = \sum_{q=-\infty}^{+\infty} u_q \left(\lambda v_{k-q} + \mu w_{k-q} \right) = \sum_{q=-\infty}^{+\infty} \left(\lambda u_q v_{k-q} + \mu u_q w_{k-q} \right) =$$

$$= \sum_{q=-\infty}^{+\infty} \left(\lambda u_q v_{k-q} \right) + \sum_{q=-\infty}^{+\infty} \left(\mu u_q w_{k-q} \right) = \lambda \sum_{q=-\infty}^{+\infty} u_q v_{k-q} + \mu \sum_{q=-\infty}^{+\infty} u_q w_{k-q} =$$

$$= \lambda \left[u_q * v_q \right]_k + \mu \left[u_q * w_q \right]_k$$

Simmetrically can be proved the linearity on the left component. $\square$

Theorem 2.2.4 *The convolution is commutative, so that given two distributions u_k and v_k holds*

$$\left[u_q * v_q \right]_k = \left[v_q * u_q \right]_k$$

Proof:

$$\left[u_q * v_q \right]_k = \sum_{q=-\infty}^{+\infty} u_q\, v_{k-q} =$$

defining $r = k - q$ then $q = k - r$ and the previous can be rewritten

$$= \sum_{r=-\infty}^{+\infty} u_{k-r}\, v_r = \sum_{r=-\infty}^{+\infty} v_r\, u_{k-r} = \left[v_r * u_r \right]_k =$$

the variable q is an ausiliary variable as r then can be written

$$= \left[v_q * u_q \right]_k.$$

$\square$

Theorem 2.2.5 *The convolution is associative, so that given three distributions u_k, v_k and w_k holds*

$$\left[u_q * [v_r * w_r]_q\right]_k = \left[[u_r * v_r]_q * w_q\right]_k$$

Proof:

$$\left[u_q * [v_r * w_r]_q\right]_k = \left[u_q * [w_r * v_r]_q\right]_k = \sum_{q=-\infty}^{+\infty} u_q \sum_{r=-\infty}^{+\infty} w_r v_{k-q-r} =$$

that can be rewritten

$$= \sum_{r=-\infty}^{+\infty} w_r \sum_{q=-\infty}^{+\infty} u_q v_{k-q-r} = \left[w_r * [u_q * v_q]_r\right]_k = \left[[u_q * v_q]_r * w_r\right]_k = \left[[u_r * v_r]_q * w_q\right]_k$$

$$\square$$

For these reasons can be written the expression

$$\left[u_q * v_q * w_q\right]_k$$

because wherever will be placed the parenthesis the result will be the same. Infact

$$\left[u_q * v_q * w_q\right]_k = \left[[u_r * v_r]_q * w_q\right]_k = \left[u_q * [v_r * w_r]_q\right]_k.$$

More in general can be defined for n distributions $(u_1)_k, (u_2)_k, ..., (u_n)_k$ the convolution

$$\left[*_{i=1}^n (u_i)_q\right]_k = \left[(u_1)_q * (u_2)_q * ... * (u_n)_q\right]_k$$

without ambiguity.

In general for n independent random variables holds the Theorem

Theorem 2.2.6 *Given $R_1, R_2, ..., R_n$ some n independent random variables. Then holds*

$$P\left(\left(\sum_{i=1}^n R_i\right) = k\right) = [*_{i=1}^n P(R_i = q)]_k$$

Proof: The proof is done by induction. For $n = 2$ has been already proved in theorem 2.2.1.

Supposing that is valid for $n-1$ will be proved for n. If it is valid for $n-1$ then holds

$$P\left(\left(\sum_{i=1}^{n-1} R_i\right) = k\right) = \left[*_{i=1}^{n-1} P(R_i = q)\right]_k.$$

Then can be observed that

$$P\left(\left(\sum_{i=1}^n R_i\right) = k\right) = P\left(\left(R_n + \sum_{i=1}^{n-1} R_i\right) = k\right) =$$

The value $\sum_{i=1}^{n-1} R_i$ as a random variable itself then

$$= \left[P(R_n = q) * P\left(\left(\sum_{i=1}^{n-1} R_i\right) = q\right)\right]_k =$$

for the inductive hypothesis

$$= \left[P(R_n = q) * \left(*_{i=1}^{n-1} P(R_i = q) \right) \right]_k = [*_{i=1}^{n} P(R_i = q)]_k$$

$\square$

Another important property is about delta random variable (Giacopelli, Migliore, and Tegolo, 2020)

Definition 2.2.2 *The delta random variable Δ_d is the random variable with sure outcome $d \in \mathbb{Z}$. This random variable has as distribution the delta distribution $\delta_d(k)$ defined as*

$$\delta_d(k) = \begin{cases} 1, \text{ if } k = d \\ 0, \text{ otherwise} \end{cases}$$

Theorem 2.2.7 *If f_k is a distribution then*

$$\left[f_q * \delta_d(q) \right]_k = f_{k-d}$$

Proof:

$$\left[f_q * \delta_d(q) \right]_k = \sum_{q=-\infty}^{+\infty} f_q \delta_d(k - q) =$$

then $f_q \delta_d(k - q) \neq 0$ if and only if $k - q = d$, then

$$= f_{k-d} \delta_d(d) = f_{k-d}$$

$\square$

We observe that if $d = 0$

$$\left[f_q * \delta_0(q) \right]_k = f_k.$$

Since now all the distributions will be assumed natural values:

Definition 2.2.3 *A natural values distribution f_k is a distribution such that*

$$f_k = 0, \forall k < 0$$

Theorem 2.2.8 *The convolution of two natural values distributions u_k and v_k can be rewritten as*

$$\left[u_q * v_q \right]_k = \sum_{q=0}^{k} u_q v_{k-q}$$

Proof:

$$\left[u_q * v_q \right]_k = \sum_{q=-\infty}^{+\infty} u_q v_{k-q} =$$

but can be observed that if $q < 0$ then $u_q = 0$ and if $q > k$ then $v_{k-q} = 0$. In conclusion

$$= \sum_{q=0}^{k} u_q v_{k-q}$$

$\square$

Being the convolution a binary operator is meaningful to define equations on the convolution

Definition 2.2.4 *Given two distributions f_k and h_k can be defined the convolutive equation as*

$$\left[f_q * g_q\right]_k = h_k, \forall k \geq 0$$

where the distribution g_k (if exists) is called solution of the convolutive equation.

However, the first step will be the truncated convolutive equation, which is a convolutive equation where is imposed $0 \leq k \leq k_s$ More formally:

Definition 2.2.5 *Given two distributions f_k and h_k and an integer $k_s > 0$, can be defined the truncated convolutive equation as*

$$\left[f_q * g_q\right]_k = h_k, \forall k \mid 0 \leq k \leq k_s$$

where the distribution g_k (if exists) is called solution of the truncated convolutive equation.

There is a useful theorem about truncated convolutive equations (Giacopelli, Migliore, and Tegolo, 2020)

Theorem 2.2.9 *If $f_0 \neq 0$ then the solution g_k of the truncated convolutive equation*

$$\left[f_q * g_q\right]_k = h_k, \forall k \mid 0 \leq k \leq k_s$$

exists and it is unique for every $k_s > 0$.

Proof: The convolutive equation can be rearranged in the system

$$\begin{cases} f_0 g_0 = h_0 \\ f_1 g_0 + g_1 f_0 = h_1 \\ ... \\ f_{k_s} g_0 + f_{k_s-1} g_1 + ... + f_1 g_{k_s-1} + f_0 g_{k_s} = h_{k_s} \end{cases}$$

that can be rewritten in matricial form

$$\begin{bmatrix} f_0 & 0 & ... & 0 & 0 \\ f_1 & f_0 & ... & 0 & 0 \\ ... & ... & ... & ... & ... \\ f_{k_s} & f_{k_s-1} & ... & f_1 & f_0 \end{bmatrix} \begin{bmatrix} g_0 \\ g_1 \\ ... \\ g_{k_s-1} \\ g_{k_s} \end{bmatrix} = \begin{bmatrix} h_0 \\ h_1 \\ ... \\ h_{k_s-1} \\ h_{k_s} \end{bmatrix}$$

Can be seen that if $f_0 \neq 0$ the matrix determinant is not null and the system is Kramerian, then the solution exists and it is unique. $\square$

However the previous theorem can be generalized at infinity

Theorem 2.2.10 *If $f_0 \neq 0$ then the solution g_k of the convolutive equation*

$$\left[f_q * g_q\right]_k = h_k, \forall k \geq 0$$

exists and it is unique.

Proof: Supposing that exist two distinct solutions, then taking $\bar{k}$ the smallest value of k such that the distributions differ, then the two distinct solutions should be two distinct solutions of the problem

$$\left[f_q * g_q\right]_k = h_k, \forall k \mid 0 \leq k \leq \bar{k}$$

that is a contradiction. Then the solution if exists must be unique.

In order to extend truncated solutions can be observed that for any given $\hat{k} > 0$ the value $g_{\hat{k}}$ can be computed as

$$g_{\hat{k}} = w_{\hat{k}}$$

where the distribution w_k is the solution of the problem

$$\left[f_q * w_q\right]_k = h_k, \forall k \mid 0 \leq k \leq \hat{k}$$

that must exist unique in virtue of previous theorem.   □

However, the condition $f_0 \neq 0$ in real-world applications is often not satisfied. These cases can happen that will be required some additional condition on h_k to solve the equation. For example:

Theorem 2.2.11 *The equation* $\left[\delta_d(q) * g_q\right]_k = h_k$ *(with $d > 0$) has solution if and only if* $h_0 = ... = h_{d-1} = 0$.

Proof: The system associated to the equation is

$$\begin{cases} \delta_d(0)g_0 = 0 = h_0 \\ \delta_d(1)g_0 + \delta_d(0)g_1 = 0 = h_1 \\ ... \\ \delta_d(d-1)g_0 + \delta_d(d-2)g_1 + ... + \delta_d(1)g_{d-2} + \delta_d(0)g_{d-1} = 0 = h_{d-1} \end{cases}$$

remebering that $\delta_d(q) = 0$ if $q \leq d - 1$. This imply that must be $h_0 = ... = h_{d-1} = 0$.
□

However if the previous conditions holds can be observed that

$$\left[\delta_d(q) * g_q\right]_k = g_{k-d} = h_k$$

then the solution is

$$g_k = h_{k+d}.$$

This solution is still valid for $k \geq d$ even if is not true $h_0 = ... = h_{d-1} = 0$. For this reason, this particular function will be taken as an approximate solution to the problem.

Supposing now to take an $\varepsilon \in]0, 1[$ and to define the convolutive equation

$$\left[(\varepsilon(\delta_0(q)) + (1 - \varepsilon)(\delta_d(q))) * (g^\varepsilon)_q\right]_k = h_k.$$

Now the solution of this problem exists and it is unique for every $\varepsilon \in]0, 1[$, being $(\varepsilon(\delta_0(0)) + (1 - \varepsilon)(\delta_d(0))) = \varepsilon \neq 0$. The solution is the distribution

$$(g^\varepsilon)_k = \begin{cases} \frac{h_k}{\varepsilon}, \text{ if } 0 \leq k \leq d - 1 \\ \frac{h_k}{\varepsilon} - \frac{(1-\varepsilon)h_{k-d}}{\varepsilon^2}, \text{ if } k \geq d \end{cases}$$

and can be observed that for $\varepsilon \to 0$ then $(g^\varepsilon)_k$ diverges to infinity. This simple example makes evident how small perturbations in the arguments of deconvolution can cause big changes in its result.

2.3 Convolutive model

The idea behind the convolutive model is that if the graph has a block structure with connections independent from block to block the degree will be the sum of these independent contributions that will result in a convolution switching to degree distributions. More formally:

Definition 2.3.1 *A Generalized Convolutive model with B blocks is a graph C where defined a partitions of nodes with B disjoint sets $C_1, ..., C_B \subseteq V(C)$ such that $C_1 \cup ... \cup C_B = V(C)$ and the connectivity between every couple of sets C_i and C_j (i and j can be equal) is independent from the connectivity between the other couples.*

The terms blocks and subpopulations will be used interchangeably since now. The simpler Generalized Convolutive model (since now GCM) is the convolutive model with two blocks. For a GCM with two blocks holds the theorem

Theorem 2.3.1 *Given C a GCM with two subpopulations C_1 and C_2, the degree distributions are*

$$P(D_C^I = k) = \frac{|C_1|}{|C_1| + |C_2|} \left[P(D_C^I(C_1, C_1) = q) * P(D_C^I(C_1, C_2) = q) \right]_k +$$

$$+ \frac{|C_2|}{|C_1| + |C_2|} \left[P(D_C^I(C_2, C_1) = q) * P(D_C^I(C_2, C_2) = q) \right]_k$$

and

$$P(D_C^O = k) = \frac{|C_1|}{|C_1| + |C_2|} \left[P(D_C^O(C_1, C_1) = q) * P(D_C^O(C_1, C_2) = q) \right]_k +$$

$$+ \frac{|C_2|}{|C_1| + |C_2|} \left[P(D_C^O(C_2, C_1) = q) * P(D_C^O(C_2, C_2) = q) \right]_k$$

Proof: The proof will be done for the indegree, but extends symmetrically to the outdegree. First of all must be observed that can be used the Conditional probability theorem (Olver et al., 2010)

$$P(D_C^I = k) = P(V \in C_1)P((D_C^I = k)|(V \in C_1)) + P(V \in C_2)P((D_C^I = k)|(V \in C_2))$$

where $V \in C_i$ is the event the vertex is in the set C_i. Observing that $P(V \in C_i) = \frac{|C_i|}{|C_1| + |C_2|}$ and the event $((D_C^I = k)|(V \in C_i))$ is equal to $D_C^I(C_i, C) = k$ then

$$= \frac{|C_1|}{|C_1| + |C_2|} P(D_C^I(C_1, C) = k) + \frac{|C_2|}{|C_1| + |C_2|} P(D_C^I(C_2, C) = k).$$

In general can be said that the degree of a node j is the sum of the contributions of the populations, formally

$$D_C^I(j, C) = D_C^I(j, C_1 \cup C_2) = D_C^I(j, C_1) + D_C^I(j, C_2)$$

then switching to the random variables can be said

$$D_C^I(C_i, C) = D_C^I(C_i, C_1) + D_C^I(C_i, C_2).$$

Being $D_C^I(C_i, C_1)$ and $D_C^I(C_i, C_2)$ independent then can be applied the convolution theorem 2.2.1

$$P(D_C^I(C_i, C) = k) = P((D_C^I(C_i, C_1) + D_C^I(C_i, C_2)) = k) =$$

$$= \left[P(D_C^I(C_i, C_1) = q) * P(D_C^I(C_i, C_2) = q) \right]_k.$$

In conclusion

$$P(D_C^I = k) = \frac{|C_1|}{|C_1| + |C_2|} \left[P(D_C^I(C_1, C_1) = q) * P(D_C^I(C_1, C_2) = q) \right]_k +$$

$$+ \frac{|C_2|}{|C_1| + |C_2|} \left[P(D_C^I(C_2, C_1) = q) * P(D_C^I(C_2, C_2) = q) \right]_k$$

and symmetrically holds

$$P(D_C^O = k) = \frac{|C_1|}{|C_1| + |C_2|} \left[P(D_C^O(C_1, C_1) = q) * P(D_C^O(C_1, C_2) = q) \right]_k +$$

$$+ \frac{|C_2|}{|C_1| + |C_2|} \left[P(D_C^O(C_2, C_1) = q) * P(D_C^O(C_2, C_2) = q) \right]_k$$

$$\square$$

The previous theorem can be generalized for a general GCM with $B \geq 2$ blocks

Theorem 2.3.2 *Given C a GCM with subpopulations $C_1, C_2, ..., C_B$ and $B \geq 2$, the degree distributions are*

$$P(D_C^I = k) = \sum_{j=1}^{B} \frac{|C_j|}{\sum_{l=1}^{B} |C_l|} \left[*_{i=1}^{B} P(D_C^I(C_j, C_i) = q) \right]_k$$

and

$$P(D_C^O = k) = \sum_{j=1}^{B} \frac{|C_j|}{\sum_{l=1}^{B} |C_l|} \left[*_{i=1}^{B} P(D_C^O(C_j, C_i) = q) \right]_k$$

Proof: As in the theorem 2.3.1

$$P(D_C^I = k) = \sum_{j=1}^{B} P(V \in C_j) P((D_C^I = k)|(V \in C_j)) = \sum_{j=1}^{B} \frac{|C_j|}{\sum_{l=1}^{B} |C_l|} P(D_C^I(C_j, C) = k) =$$

and applying the convolution theorem 2.2.6 (being $D_C^I(C_j, C) = \sum_{i=1}^{B} D_C^I(C_j, C_i)$ as in the previous proof of theorem 2.3.1)

$$= \sum_{j=1}^{B} \frac{|C_j|}{\sum_{l=1}^{B} |C_l|} \left[*_{i=1}^{B} P(D_C^I(C_j, C_i) = q) \right]_k.$$

Symmetrically holds for the outdegree

$$P(D_C^O = k) = \sum_{j=1}^{B} \frac{|C_j|}{\sum_{l=1}^{B} |C_l|} \left[*_{i=1}^{B} P(D_C^O(C_j, C_i) = q) \right]_k.$$

$\square$

2.4 Derivation of the degree distributions of an ER mixed model

The more straightforward of GCMs is the ER mixed model (Potjans and Diesmann, 2012), even if this model hasn't been introduced with this formalism. In fact the ER mixed model is the natural extension of ER model where a set of populations $P_1, ..., P_B$ are connected with fixed probabilities depending on the arriving and outgoing neuron types. In GCM formalism

Definition 2.4.1 *An ER mixed model M is a GCM with subpopulations $P_1, ..., P_B$ such that every couple of nodes of $P_i \times P_j$ is independently connected with a probability $(P_M)_{ij} \in [0, 1]$. The matrix P_M is called probability of connection matrix.*

Can be observed that the connectivity between each couple of populations is ER (Erdös and Rényi, 1959). Using theorem 2.3.2 can be proved

Theorem 2.4.1 *An ER mixed model M with subpopulations $P_1, ..., P_B$ has degree distributions*

$$P(D_M^I = k) = \sum_{j=1}^{B} \frac{|P_j|}{\sum_{l=1}^{B} |P_l|} \left[B_{(P_M)_{jj}}^{|P_j|-1}(q) * \left(*_{\substack{i=1 \\ i \neq j}}^{B} B_{(P_M)_{ij}}^{|P_i|}(q) \right) \right]_k$$

and

$$P(D_M^O = k) = \sum_{j=1}^{B} \frac{|P_j|}{\sum_{l=1}^{B} |P_l|} \left[B_{(P_M)_{jj}}^{|P_j|-1}(q) * \left(*_{\substack{i=1 \\ i \neq j}}^{B} B_{(P_M)_{ji}}^{|P_i|}(q) \right) \right]_k$$

Proof: Using theorem 2.3.2

$$P(D_M^I = k) = \sum_{j=1}^{B} \frac{|P_j|}{\sum_{l=1}^{B} |P_l|} \left[*_{i=1}^{B} P(D_M^I(P_j, P_i) = q) \right]_k.$$

Can be observed that for $i = j$

$$P(D_M^I(P_j, P_i) = q) = B_{(P_M)_{jj}}^{|P_j|-1}(q)$$

being an ER model subgraph.

For $i \neq j$ can be thought as the sum of $|P_i|$ Bernoulli random variables, that is the binomial

$$P(D_M^I(P_j, P_i) = q) = B_{(P_M)_{ij}}^{|P_i|}(q).$$

Substituting in the formula

$$P(D_M^I = k) = \sum_{j=1}^{B} \frac{|P_j|}{\sum_{l=1}^{B} |P_l|} \left[B_{(P_M)_{jj}}^{|P_j|-1}(q) * \left(*_{\substack{i=1 \\ i \neq j}}^{B} B_{(P_M)_{ij}}^{|P_i|}(q) \right) \right]_k.$$

Symmetrically holds for the outdegree

$$P(D_M^O = k) = \sum_{j=1}^{B} \frac{|P_j|}{\sum_{l=1}^{B} |P_l|} \left[B_{(P_M)_{jj}}^{|P_j|-1}(q) * \left(*_{\substack{i=1 \\ i \neq j}}^{B} B_{(P_M)_{ji}}^{|P_i|}(q) \right) \right]_k.$$

□

Observing that the Binomial distribution can be approximated by a Gaussian curve (Feller, 1945), the convolution of Gaussian distributions is still a Gaussian distribution (Bromiley, 2003). A Gaussian bell can still approximate the convolution of Binomial distributions (as can be seen in the schematic in figure 2.2).
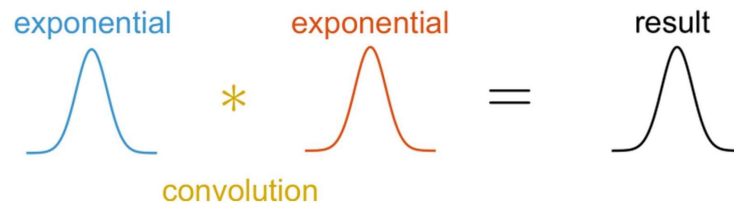
FIGURE 2.2: Schematics of convolution between Gaussian distributions (Bromiley, 2003).

The ER mixed is the convex linear combination of B like Gaussian distributions, and it is the cause of its peculiar structure of multi-modal distributions (as can be seen in figure 2.3). In general, the degree distributions of the ER mixed have at most B distinct peaks. However, this multi-peaks behavior, in general, is not present in experimental neuronal networks (Giacopelli et al., 2021) for this reason will be needed a Generalized Convolutive Model able to explain experimental findings (Giacopelli, Migliore, and Tegolo, 2020).
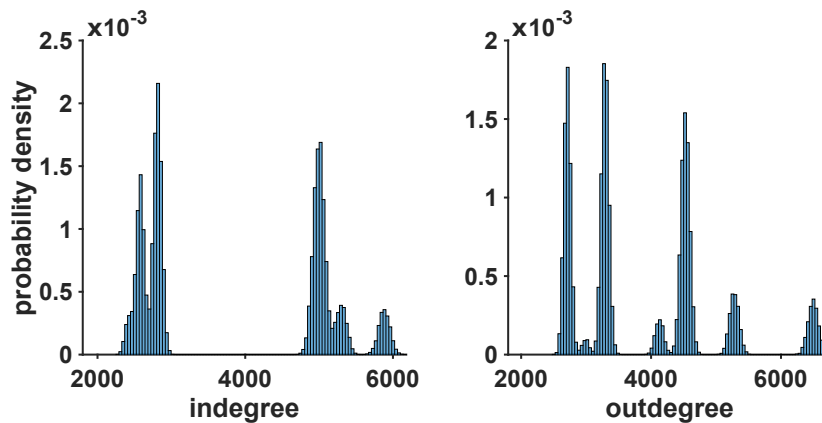
FIGURE 2.3: Degree distributions from an ER mixed model with 8 subpopulations (4 of excitatory neurons and 4 of inhibitory neurons) fitting Neocortical area connectivity (Potjans and Diesmann, 2012).

2.5 Real-world Convolutive models

As observed in the previous sections, the experimental networks (Giacopelli et al., 2021) in general does not follow the behavior of ER mixed model. For this reason, in (Giacopelli, Migliore, and Tegolo, 2020) has been introduced a simple GCM with a power-law connectivity (Barabási and Albert, 1999) inside the blocks and ER-like connectivity between blocks. Since now, the term *Convolutive model* will refer to this

type of connectivity. It classifies GCM like the ER mixed as Exponential models rather than to Convolutive models. The classification is sketched in figure 2.4.
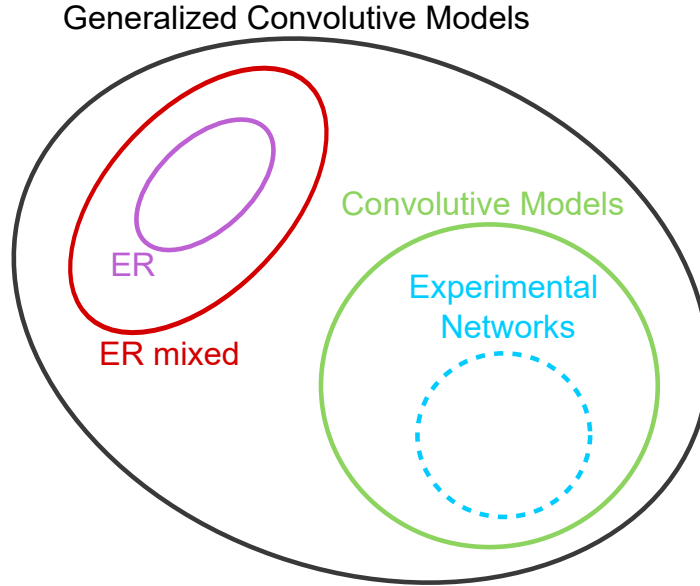
FIGURE 2.4: GCM classification, where the term *Experimental Networks* refers to the networks analyzed in (Giacopelli et al., 2021).

The first Convolutive model was the aspatial bi-population model (Giacopelli, Migliore, and Tegolo, 2020)

Definition 2.5.1 *A Convolutive aspatial bi-population model C is a convolutive model with two populations C_1 and C_2 with the same number of nodes, where the connectivity inside the blocks is a Price model (Newman, 2010) and the connectivity between the blocks is an ER-like connectivity with parameters the integer l and the probabilities $\phi^U, \phi^D, p \in [0,1]$.*

In particular each subpopulation C_1 and C_2 is subdivided in a partition with $M = \frac{|C_r|}{l}$ sets (with $r = 1, 2$) $C_r^1, C_r^2, ..., C_r^M \subseteq C_r$. Then to each couple of sets C_1^i and C_2^j is assigned a random variable $\beta_{ij} \sim X_p$. For each node $s \in C_1^i$ and $t \in C_2^j$ can be defined two random variables $\nu_{st} \sim X_{\phi^D}$ and $\eta_{st} \sim X_{\phi^U}$. In conclusion they are connected with a connection starting from s to t if the random variable

$$\eta_{st}\beta_{ij} + \nu_{st}(1 - \beta_{ij})$$

is equal to one.

Simmetrically, defining $\beta'_{ji} \sim X_p$, $\nu'_{ts} \sim X_{\phi^D}$ and $\eta'_{ts} \sim X_{\phi^U}$, then exists a connection from t to s if the random variable

$$\eta'_{ts}\beta'_{ji} + \nu'_{ts}(1 - \beta'_{ji})$$

is equal to one.

The ER-like connectivity used in the previous definition is a generalization of the ER model used to expand the standard deviation of the degree distributions. Can be proved that

Theorem 2.5.1 *For a Convolutive aspatial bi-population model C can be proved that*

$$P(D_C^I = k) = \left[f_q^I * \left(((1-p)\delta_0(q) + pB_{\phi^U}^l(q)) * (p\delta_0(q) + (1-p)B_{\phi^D}^l(q)) \right)^{*M} \right]_k$$

and

$$P(D_C^O = k) = \left[f_q^O * \left(((1-p)\delta_0(q) + pB_{\phi^U}^l(q)) * (p\delta_0(q) + (1-p)B_{\phi^D}^l(q)) \right)^{*M} \right]_k$$

Proof: As usual will be proved for the indegree and for the outdegree will be deduced by symmetry. Theorem 2.3.1 states that

$$P(D_C^I = k) = \frac{|C_1|}{|C_1| + |C_2|} \left[P(D_C^I(C_1, C_1) = q) * P(D_C^I(C_1, C_2) = q) \right]_k +$$

$$+ \frac{|C_2|}{|C_1| + |C_2|} \left[P(D_C^I(C_2, C_1) = q) * P(D_C^I(C_2, C_2) = q) \right]_k =$$

but $|C_1| = |C_2|$ by construction and $P(D_C^I(C_1, C_1) = k) = P(D_C^I(C_2, C_2) = k) = f_k^I$, then

$$= \frac{1}{2} \left[f_q^I * P(D_C^I(C_1, C_2) = q) \right]_k + \frac{1}{2} \left[f_q^I * P(D_C^I(C_2, C_1) = q) \right]_k =$$

$$= \left[f_q^I * \left(\frac{1}{2} P(D_C^I(C_1, C_2) = q) + \frac{1}{2} P(D_C^I(C_2, C_1) = q) \right) \right]_k.$$

Recalling the ER-like connectivity between blocks

$$D_C^I(C_2, C_1) \sim \sum_{j=1}^{M} \sum_{s \in C_1^j} (\eta_{st}\beta_{ij} + \nu_{st}(1 - \beta_{ij})) = \sum_{j=1}^{M} \left(\beta_{ij} \sum_{s \in C_1^j} \eta_{st} + (1 - \beta_{ij}) \sum_{s \in C_1^j} \nu_{st} \right).$$

Now can be observed that if $Z \sim X_\pi$ and W are two random variables, then

$$P(ZW = k) = \begin{cases} P(Z = 0) + P(Z = 1)P(W = 0) = (1 - \pi) + \pi P(W = 0), \text{ if } k = 0 \\ P(Z = 1)P(W = k) = \pi P(W = k). \text{ otherwise} \end{cases}$$

that can be rewritten $P(WZ = k) = (1 - \pi)\delta_0(k) + P(W = k)$. Then

$$P \left(\sum_{j=1}^{M} \left(\beta_{ij} \sum_{s \in C_1^j} \eta_{st} + (1 - \beta_{ij}) \sum_{s \in C_1^j} \nu_{st} \right) \right) = \left[*_{j=1}^{M} P \left(\left(\beta_{ij} \sum_{s \in C_1^j} \eta_{st} + (1 - \beta_{ij}) \sum_{s \in C_1^j} \nu_{st} \right) = q \right) \right]_k =$$

$$= \left[*_{j=1}^{M} \left(P \left(\beta_{ij} \sum_{s \in C_1^j} \eta_{st} = q \right) * P \left((1 - \beta_{ij}) \sum_{s \in C_1^j} \nu_{st} = q \right) \right) \right]_k =$$

$$= \left[*_{j=1}^{M} \left(\left((1-p)\delta_0(q) + p\, P \left(\sum_{s \in C_1^j} \eta_{st} = q \right) \right) * \left(p\delta_0(q) + (1-p)\, P \left(\sum_{s \in C_1^j} \nu_{st} = q \right) \right) \right) \right]_k =$$

observing that $\sum_{s \in C_1^j} \nu_{st} \sim B(l, \phi^D)$ and $\sum_{s \in C_1^j} \eta_{st} \sim B(l, \phi^U)$

$$= \left[*_{j=1}^{M} \left(\left((1-p)\delta_0(q) + p\, B_{\phi^U}^l(q) \right) * \left(p\delta_0(q) + (1-p)\, B_{\phi^D}^l(q) \right) \right) \right]_k =$$

observing that the arguments of convolution are now independent from j can be expressed in terms of convolution power

$$= \left[\left(\left((1-p)\delta_0(q) + p\, B^l_{\phi^U}(q) \right) * \left(p\delta_0(q) + (1-p)\, B^l_{\phi^D}(q) \right) \right)^{*M} \right]_k.$$

Then

$$P(D^I_C(C_2, C_1) = k) = \left[\left(\left((1-p)\delta_0(q) + p\, B^l_{\phi^U}(q) \right) * \left(p\delta_0(q) + (1-p)\, B^l_{\phi^D}(q) \right) \right)^{*M} \right]_k$$

by symmetry $D^I_C(C_2, C_1) = D^I_C(C_1, C_2)$ and then

$$P(D^I_C(C_1, C_2) = k) = \left[\left(\left((1-p)\delta_0(q) + p\, B^l_{\phi^U}(q) \right) * \left(p\delta_0(q) + (1-p)\, B^l_{\phi^D}(q) \right) \right)^{*M} \right]_k.$$

In conclusion

$$P(D^I_C = k) = \left[f^I_q * \left(\left((1-p)\delta_0(q) + p\, B^l_{\phi^U}(q) \right) * \left(p\delta_0(q) + (1-p)\, B^l_{\phi^D}(q) \right) \right)^{*M} \right]_k.$$

Symmetrically can be proved

$$P(D^O_C = k) = \left[f^O_q * \left(\left((1-p)\delta_0(q) + p\, B^l_{\phi^U}(q) \right) * \left(p\delta_0(q) + (1-p)\, B^l_{\phi^D}(q) \right) \right)^{*M} \right]_k.$$

$\square$

The previous convolutive model has degree distributions of the form

$$P(D^I_C = k) = \left[f^I_q * K^I_q \right]_k$$

and

$$P(D^O_C = k) = \left[f^O_q * K^O_q \right]_k$$

where the distributions K^I_k and K^O_k are called convolutive kernels. For construction the distributions f^I_k and f^O_k are power law and the distributions K^I_k and K^O_k are exponential. Then the shape of the convolution between them is sketched in figure 2.5. Can be observed that the result has a growing behavior for low degrees and an heavy tail behavior for high degree. However in order to mathematically prove that the result of this convolution is still power law can be proved the theorem

Theorem 2.5.2 *Consider two probability distributions u_k and v_k such that:*

- $\lim_{k \to +\infty} \frac{u_{k-s}}{u_k} = 1, \forall s \geq 0$;

- $\exists T \geq 0 | \forall t > T \Rightarrow v_t = 0$

If it is defined $c_k = \left[u_q * v_q \right]_k$ then holds

$$\lim_{k \to +\infty} \frac{c_k}{u_k} = 1$$

Proof:

$$\frac{c_k}{u_k} = \frac{[u_q * v_q]_k}{u_k} = \frac{\sum_{q=0}^{k} u_q v_{k-q}}{u_k} = \sum_{q=k-T}^{k} \frac{u_q}{u_k} v_{k-q} = \sum_{j=0}^{T} \frac{u_{k-j}}{u_k} v_j$$

since $\lim_{k \to +\infty} \frac{u_{k-j}}{u_k} = 1$ is one of the hypothesis, then

$$\lim_{k \to +\infty} \frac{c_k}{u_k} = \lim_{k \to +\infty} \sum_{j=0}^{T} \frac{u_{k-j}}{u_k} v_j = \sum_{j=0}^{T} \left(\lim_{k \to +\infty} \frac{u_{k-j}}{u_k} \right) v_j = \sum_{j=0}^{T} v_j = 1.$$

$\square$

Can be observed that the hyphothesis of v_k are often satisfied by the kernels K_k being exponential. The hypothesis of u_k are often satisfied by a power law distribution f_k, because if it is power law exists $\gamma > 0$ such as

$$\lim_{k \to +\infty} \frac{f_k}{k^{-\gamma}} = \varphi \in \mathbb{R} \setminus \{0\}.$$

Suppposing $s \geq 0$ and being $\lim_{k \to +\infty} \frac{(k-s)^{-\gamma}}{k^{-\gamma}} = \lim_{k \to +\infty} \left(\frac{k-s}{k} \right)^{-\gamma} = 1$

$$\lim_{k \to +\infty} \frac{f_{k-s}}{f_k} = \lim_{k \to +\infty} \frac{f_{k-s}}{f_k} \frac{k^{-\gamma}}{(k-s)^{-\gamma}} = \lim_{k \to +\infty} \frac{f_{k-s}}{(k-s)^{-\gamma}} \frac{k^{-\gamma}}{f_k} = \frac{\varphi}{\varphi} = 1.$$

Then the Convolutive model has the same heavy tail behavior of the distributions f_k^I and f_k^O inside the blocks.

$$P(D_B^I = k) = \left[f_q^I * \left(\left((1-p)\delta_0(q) + pB_{\phi^U}^l(q) \right) * \left(p\delta_0(q) + (1-p)B_{\phi^D}^l(q) \right) \right)^{*M} \right]_k$$

$$P(D_B^O = k) = \left[f_q^O * \left(\left((1-p)\delta_0(q) + pB_{\phi^U}^l(q) \right) * \left(p\delta_0(q) + (1-p)B_{\phi^D}^l(q) \right) \right)^{*M} \right]_k$$
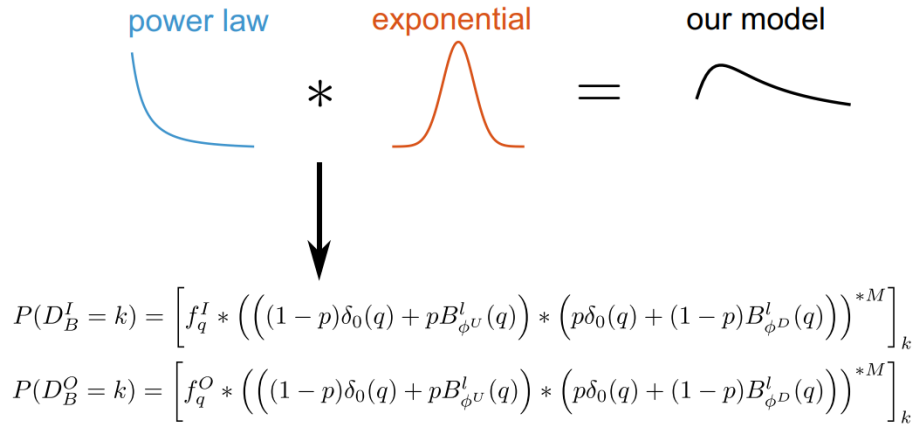
FIGURE 2.5: Convolutive Model representation from (Giacopelli, Migliore, and Tegolo, 2020).

In order to study a fitting procedure we state the theorem

Theorem 2.5.3 *Assigning the parameters m_0, ρ and l is possible to find a configuration of Convolutive Model such as*

$$P(D_C^I = k) \approx v_k^I$$

where v_k^I is the experimental network indegree distribution.

Proof: Supposing that C_1 and C_2 are graph with N nodes and Price connectivity (Newman, 2010). Then using theorem 2.5.1 can be stated that

$$P(D_C^I = k) = \left[f_q^I * K_q^I \right]_k.$$

From (Newman, 2010) can be derived that the indegree of a Price model is

$$f_k^I = \frac{N - m_0}{N}\gamma_k + \frac{m_0}{N}B_\rho^{m_0-1}(k)$$

where N is the number of nodes, the Price model starts with a seed graph with m_0 nodes connected with ER connectivity with probability $\rho \in [0, 1]$ and the Price model has a Γ random variable with distribution γ_k. The aim is to solve the convolutive equation

$$\left[\left(\frac{N - m_0}{N}\gamma_q + \frac{m_0}{N}B_\rho^{m_0-1}(q)\right) * K_q^I\right]_k = v_k^I.$$

The equation can be approximated by the equation

$$\left[\left(\frac{N - m_0}{N}\gamma_q + \frac{m_0}{N}B_\rho^{m_0-1}(q)\right) * \delta_d(q)\right]_k = v_k^I$$

now the problem is to find the parameters γ_k, d and p to solve (at least in an approximate way) the equation. Can be observed that

$$\mathbb{E}(\left[f_q^I * K_q^I\right]_k) = \mathbb{E}(f_k^I) + \mathbb{E}(K_k^I) = \mathbb{E}(v_k^I)$$

if $E_K = \mathbb{E}(K_k^I)$ and $E_V = \mathbb{E}(v_k^I)$ then

$$\mathbb{E}(f_k^I) + E_K = E_V$$

and then

$$\mathbb{E}(f_k^I) = E_V - E_K.$$

Being valid the formula

$$\left[f_q^I * \delta_d(q)\right]_k = v_k^I,$$

in previous sections has been pointed out that a possible solution is

$$f_k^I = \frac{v_{k+d}}{\sum_{s=d}^{+\infty} v_s^I}$$

then the expected value is

$$\mathbb{E}(f_k^I) = \sum_{k=0}^{+\infty} k f_k^I = \sum_{k=0}^{+\infty} k \frac{v_{k+d}^I}{\sum_{s=d}^{+\infty} v_s^I} = \frac{\sum_{k=0}^{+\infty} k v_{k+d}^I}{\sum_{s=d}^{+\infty} v_s^I} = \frac{\sum_{s=d}^{+\infty} (s - d) v_s^I}{\sum_{s=d}^{+\infty} v_s^I} =$$

$$= \frac{\sum_{s=d}^{+\infty} s v_s^I - d \sum_{s=d}^{+\infty} v_s^I}{\sum_{s=d}^{+\infty} v_s^I} = \frac{E_V - \sum_{j=0}^{d-1} v_j^I - d\left(1 - \sum_{j=0}^{d-1} v_j^I\right)}{1 - \sum_{j=0}^{d-1} v_j^I}$$

Then γ_k can be computed from the system

$$\begin{cases} \mathbb{E}(\gamma_k) = \frac{E_V - \sum_{j=0}^{d-1} v_j^I - d\left(1 - \sum_{j=0}^{d-1} v_j^I\right)}{1 - \sum_{j=0}^{d-1} v_j^I} = \psi(d, v_k^I) \\ \mathbb{E}(\gamma_k) = E_V - E_K. \end{cases}$$

Defining $\hat{k}$ the biggest integer such as $v_k^I > 0$, can be found

$$d = \arg \min_{d \in \{0,...,\hat{k}\}} |(E_V - E_K) - \psi(d, v_k^I)|.$$

Then determined d, then can be found f_k^I as

$$f_k^I = \frac{v_{k+d}}{\sum_{s=d}^{+\infty} v_s^I}$$

and observing that

$$f_k^I = \frac{N - m_0}{N} \gamma_k + \frac{m_0}{N} B_\rho^{m_0-1}(k)$$

then

$$\gamma_k = \frac{N}{N - m_0} f_k^I - \frac{m_0}{N - m_0} B_\rho^{m_0-1}(k).$$

The a value of the Price model will be set to $a = \mathbb{E}(\Gamma)$ in order to obtain a power law model with decay of k^{-3}.

In order to compute p

$$\mathbb{E}\left(\left(\left((1-p)\delta_0(q) + p\, B_{\phi^U}^l(q) \right) * \left(p\delta_0(q) + (1-p)\, B_{\phi^D}^l(q) \right) \right)^{*M} \right) =$$

$$= M\mathbb{E}\left(\left((1-p)\delta_0(q) + p\, B_{\phi^U}^l(q) \right) * \left(p\delta_0(q) + (1-p)\, B_{\phi^D}^l(q) \right) \right) =$$

$$= M\left(\mathbb{E}\left((1-p)\delta_0(q) + p\, B_{\phi^U}^l(q) \right) + \mathbb{E}\left(p\delta_0(q) + (1-p)\, B_{\phi^D}^l(q) \right) \right) =$$

$$= M(pl\phi^U + (1-p)l\phi^D) = N(p\phi^U + (1-p)\phi^D)$$

that solved for p returns

$$p = \frac{\frac{E_K}{N} - \phi^D}{\phi^U - \phi^D} \approx \frac{E_K}{N}$$

for $\phi^U \approx 1$ and $\phi^D \approx 0$. in conclusion substituting

$$P(D_C^I = k) = \left[f_k^I * \left(\left((1-p)\delta_0(q) + p\, B_{\phi^U}^l(q) \right) * \left(p\delta_0(q) + (1-p)\, B_{\phi^D}^l(q) \right) \right)^{*M} \right]_k \approx v_k^I$$

and

$$P(D_C^O = k) = \left[\frac{B\left(k+a, 2+\frac{a}{\mathbb{E}(\Gamma)}\right)}{B\left(a, 1+\frac{a}{\mathbb{E}(\Gamma)}\right)} * \left(\left((1-p)\delta_0(q) + p\, B_{\phi^U}^l(q) \right) * \left(p\delta_0(q) + (1-p)\, B_{\phi^D}^l(q) \right) \right)^{*M} \right]_k$$

$$\square$$

2.6 Applications

The first use case is about the network reconstruction of about ten CA1 hippocampal regions of rodent (Bonifazi et al., 2009). This reconstruction has been performed with *the fire together, wire together* technique, and this justifies the small number of neurons per slice (about one hundred). The data is the adjacency matrix of these slice networks. It can be observed that the degree distributions of these networks

have a great range of variability however has been found the parameters to reconstruct three stereotypical cases with the Convolutive Model (Giacopelli, Migliore, and Tegolo, 2020) as depicted in figure 2.6B in terms of survival function that is defined for the indegree as

$$S_f(D_G^I, k) = \sum_{j \geq k} P(D_G^I = j)$$

and for the outdegree as

$$S_f(D_G^O, k) = \sum_{j \geq k} P(D_G^O = j).$$

These slices have many nodes not connected, and it can be conjectured that this assertive heavy tail behavior is the result of the cutting procedure of the slice, that cutting the connections between neurons changes significantly overall connectivity of the network.
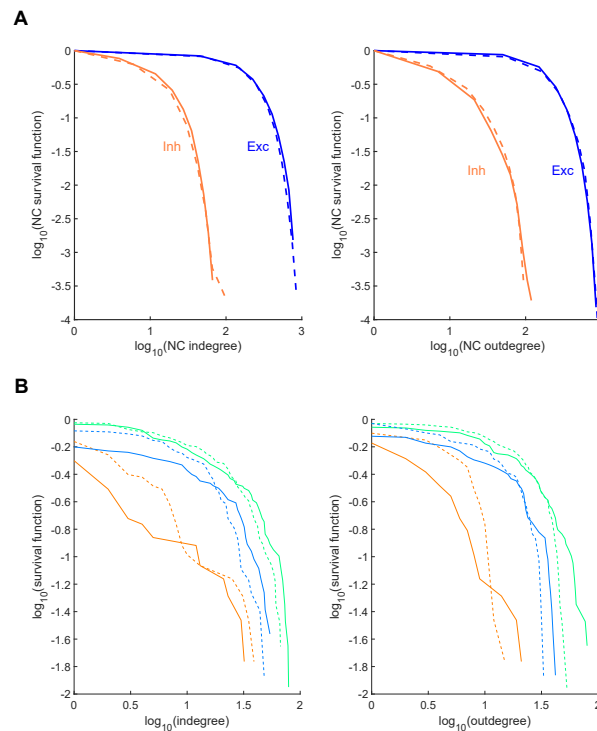
FIGURE 2.6: Result of fitting from (Giacopelli, Migliore, and Tegolo, 2020). A) The comparison between the survival functions of one of the Neocortical column (Markram et al., 2015) datasets (solid lines) and the proposed model (dotted lines). B) three representative experimental slices (solid lines) and three similar networks (dotted lines) extracted from the 5000 nodes network generated with the proposed algorithm.

The second use case is about the neocortical column reconstruction performed in (Markram et al., 2015). This stochastic model starts from small samples of neocortical area connectivity and using a data-driven algorithm (involving touch detection and pruning), generates a random network with highly accurate information about neurons' type, connectivity and displacement in the space of about 30000 neurons.

In (Giacopelli, Migliore, and Tegolo, 2020), the neurons have been divided in excitatory and inhibitory and for each population has been used a Convolutive model connectivity scheme. The result is shown in figure 2.6A, where can be seen the accordance between the model prediction and the experimental data. It must be pointed out that the information provided by the Neocortical model in (Markram et al., 2015) comprehends the neurons' type and spatial information. However, on the other hand, the model proposed in (Giacopelli, Migliore, and Tegolo, 2020), with a single-core implementation, runs in about one minute (then $1 \ core \times minutes \approx 0.017 \ core \times hours$), against the 16 $core \times hours$ required for a general touch detection model (Hjorth et al., 2020). Then we are in front of a trade-off between the specificity of networks characterization and the time required to build the networks.

Spatial models

3.1 Why spatial graphs have to be studied

Since now have been studied just the graph's adjacency matrix (and in particular the degree distributions). However, most experimental networks contain the information of soma position (the nucleus of the neuron). The question now is how the spatial information in network analysis can be relevant. In order to clarify the fundamental role of spatial information in connectivity comprehension, let us start with the example in figure 3.1.
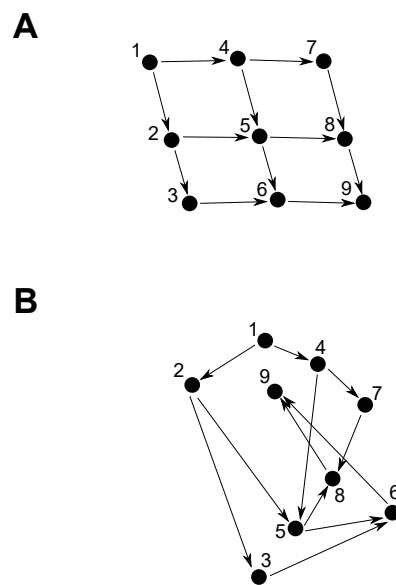
FIGURE 3.1: The graphs in figure have the same adjacency matrix. A) The position of the nodes are arranged to make a lattice structure. B) The position of the nodes does not follow any particular scheme.

In figure 3.1A and 3.1B are plotted two graphs that seems two completely different graphs. However the adjacency matrix in both cases is

$$
\begin{bmatrix}
0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}
$$

then the two graphs have the same aspatial connectivity. In the first case the regularity is evident and then if the task is to generalize and model the graph in figure 3.1A taking a decision based on the graph appearance, probably the best model would be a regular lattice (Newman, 2010). In the second case, as can be seen in figure 3.1B, the positions of the nodes are so mixed that no scheme is evident and at first glance probably the best generalization would be an ER model (Erdös and Rényi, 1959).

Then the spatial position of the nodes not only is important to experimentally characterize the network in order to make more realistic simulations (as will be seen in the following sections) but also gives an insight into real world networks connectivity schemes.

3.2   Spatial Poisson processes

In random graphs, the positions (as connectivity) can be randomly generated. The most common way to generate random positions is through an Inhomogeneous Poisson Process (Møller and Waagepetersen, 2007). Suppose the graph G a
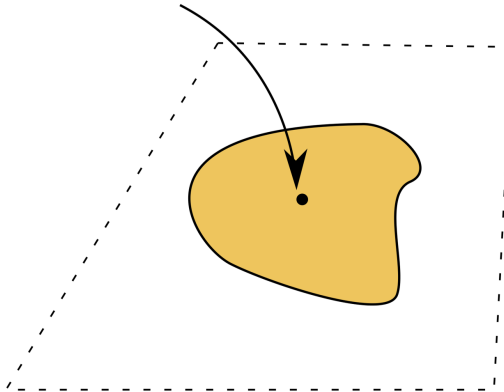
$$
E\big(N_G(\Omega)\big) \approx \rho_G(c)|\Omega|
$$

FIGURE 3.2: Schematics of Inhomogeneous Poisson processes (Møller and Waagepetersen, 2007).

d_G-dimensional spatial graphs. Then for each region $\Omega \subseteq \mathbb{R}^{d_G}$ can be defined the random variable $N_G(\Omega)$ that represents the number of nodes inside the region Ω. If

the density of nodes is constant then the expected number of nodes inside Ω is

$$\mathbb{E}(N_G(\Omega)) = \rho_0|\Omega| = \rho_0 \int_\Omega d\underline{x} = \int_\Omega \rho_0 d\underline{x}$$

for some constant $\rho_0 > 0$. However in real world networks the density is rarely constant and usually is a regular enough function $\rho_G(\underline{x}) \geq 0$, dependent from the position $\underline{x} \in \mathbb{R}^{d_G}$. Observing (as shown in figure 3.2) that in these cases for small regions $d\Omega$ holds

$$\mathbb{E}(N_G(d\Omega)) \approx \rho_G(\underline{c})|d\Omega|$$

where $\underline{c} \in d\Omega$, can be defined

Definition 3.2.1 *The positions of a graph G are generated by an Inhomogeneous Poisson process (Møller and Waagepetersen, 2007) if exists a function $\rho_G(\underline{x}) \geq 0$ such that*

$$\mathbb{E}(N_G(\Omega)) = \int_\Omega \rho_G(\underline{x})d\underline{x}.$$

Since now all the positions of the Random graphs taken into account (if not specified the contrary) will be generated by some Inhomogeneous Poisson process. A particular case of the Inhomogeneous Poisson Process will be the Region-constrained process

Definition 3.2.2 *The positions of the graph G are generated by a Region-contrained process on $\Omega_{constr} \subseteq \mathbb{R}^{d_G}$ if*

$$\rho_G(\underline{x}) = 0, \forall \underline{x} \notin \Omega_{constr}$$

This process will be very common because is the formal representation of the concept *the positions of the graph are all inside the region Ω_{constr}*.

3.3 ER spatial models

The ER spatial models are the generalizations of the ER model (Erdös and Rényi, 1959). In general an ER spatial model is:

Definition 3.3.1 *Given a function $\phi_{ES} : \mathbb{R}^{d_G} \times \mathbb{R}^{d_G} \to [0,1]$, an ER spatial model is a graph ES with position generated from some Inhomogeneous Poisson process and each couple of nodes i and j is independently connected with a probability $\phi_{ES}(\underline{x}_i, \underline{x}_j)$ dependent from the positions $\underline{x}_i$ and $\underline{x}_j$.*

This is the most general formulation for an ER spatial model, however since now will be made the following assumptions

1. The model has a constant density ρ_0;

2. Holds $\phi_{ES}(\underline{x}_i, \underline{x}_j) = p_{ES}(\|\underline{x}_i - \underline{x}_j\|)$ (Billeh et al., 2020), where the function $p_{ES}(d)$ is called probability dependent from the distance (Giacopelli, Migliore, and Tegolo, 2020), with $d \geq 0$;

3. The space is 3D, then $d_{ES} = 3$.

This particular ER spatial model is the *ER distance model* (Giacopelli, Migliore, and Tegolo, 2020). The following theorem will approximate the degree distributions

of the ER distance model, with a graph that approximate the original graph. It is made through a step function computed on a partition of $[0, +\infty[$, like in the Riemann integration. Using this step function is equivalent to subdivide the space in disjoint spherical sectors centered in the node that we are considering. If the width of the steps of the approximating probability tends to zero then the step function will tend to the actual value of probability dependent from the distance, confirming the approximation of the degree distributions.

Theorem 3.3.1 *Supposing ES an ER distance model and $R = \{r_0 = 0, r_1, ..., r_n, ...\}$ a general partition of $[0, +\infty[$ with $\Delta r = \sup_{i \in \mathbb{N}} |r_{i+1} - r_i| < +\infty$. Then in $\mathbb{R}^3$ hold the limits*

$$\lim_{\Delta r \to 0} \left[*_{i=0}^{+\infty} B_{p_i}^{N_i}(q) \right]_k = P(D_{ES}^I = k)$$

and

$$\lim_{\Delta r \to 0} \left[*_{i=0}^{+\infty} B_{p_i}^{N_i}(q) \right]_k = P(D_{ES}^O = k)$$

where $p_i = p_{ES}(r_i)$ and

$$N_i = \begin{cases} \lfloor \rho_0 \frac{4\pi}{3} r_1^3 \rfloor - 1, i = 0 \\ \lfloor \rho_0 \frac{4\pi}{3} \left(r_{i+1}^3 - r_i^3 \right) \rfloor, i > 0 \end{cases} .$$

Proof: Will be created a new graph ES' with degree distributions that approximate the ES distributions. First of all will be approximated the function $p_{ES}(d)$ with the step function

$$p_{ES'}(d) = p_{ES}(r_i), \forall d \in [r_i, r_{i+1}[.$$

If $p_{ES}(d)$ is smooth, the differences $(r_{i+1} - r_i) \to 0$ and then $p_{ES'}(d) \to p_{ES}(d)$. Fixed a node $j \in V(ES')$, can be observed that the region of the space

$$\Omega_i = \left\{ \underline{z} \in \mathbb{R}^3 | r_i \leq \left\| \underline{z} - \underline{x}_j \right\| < r_{i+1} \right\}$$

which has measure $|\Omega_i| = \frac{4\pi}{3} \left(r_{i+1}^3 - r_i^3 \right)$, and the set of $\{\Omega_i\}_{i=0}^{+\infty}$ is a partition of $\mathbb{R}^3$. Then can be stated

$$D_{ES'}^I(j) = D_{ES'}^I(j, \mathbb{R}^3) =$$

$$= \sum_{i=0}^{+\infty} D_{ES'}^I(j, \Omega_i)$$

then using theorem 2.2.6

$$P(D_{ES'}^I(j) = k) = \left[*_{i=0}^{+\infty} P\left(D_{ES'}^I(j, \Omega_i) = q \right) \right]_k .$$

Inside Ω_i the graph is ER with probability $p_i = p_{ES}(r_i)$ and expected number of nodes

$$\mathbb{E}(N_{ES}(\Omega_i)) = \rho_0 |\Omega_i| = \rho_0 \frac{4\pi}{3} \left(r_{i+1}^3 - r_i^3 \right)$$

In other to round this number to an integer has been used the floor function

$$N_i = \begin{cases} \lfloor \mathbb{E}(N(\Omega_0)) \rfloor - 1 = \lfloor \rho_0 \frac{4\pi}{3} r_1^3 \rfloor - 1, i = 0 \\ \lfloor \mathbb{E}(N(\Omega_i)) \rfloor = \lfloor \rho_0 \frac{4\pi}{3} \left(r_{i+1}^3 - r_i^3 \right) \rfloor, i > 0 \end{cases}$$

Then can be written

$$P\left(D_{ES'}^{I}(j, \Omega_i) = k\right) = B_{p_i}^{N_i}(k).$$

Then substituting in the formula

$$P\left(D_{ES'}^{I}(j) = k\right) = \left[*_{i=0}^{+\infty} B_{p_i}^{N_i}(q)\right]_k$$

because of the homogeneous network can be assumed that the nodes are all equivalent, then by symmetry

$$P\left(D_{ES'}^{I} = k\right) = \left[*_{i=0}^{+\infty} B_{p_i}^{N_i}(q)\right]_k.$$

Symmetrically can be proved that

$$P\left(D_{ES'}^{O} = k\right) = \left[*_{i=0}^{+\infty} B_{p_i}^{N_i}(q)\right]_k.$$

$\square$

However an infinite convolution is hard compute. For this reason will be introduced a sort of truncation theorem

Theorem 3.3.2 *If $p_{ES}(d)$ is such that*

$$\lim_{d \to +\infty} d^{3+\eta} p_{ER}(d) = 0$$

for some $0 < \eta < 1$.

Then the previous theorem 3.3.1 can be truncated, losing at most ϵ expected connections for each node, for every $\epsilon > 0$.

Proof: Fixing a node $j \in V(ES)$ and a region

$$\Omega_\delta = \left\{\underline{z} \in \mathbb{R}^3 \mid \delta \le \left\|\underline{z} - \underline{x}_j\right\| < (\delta + \Delta\delta)\right\} \subseteq \mathbb{R}^3$$

dependent δ that is the distance from the node j. The number of nodes in Ω_d is

$$\mathbb{E}\left(N_{ES}\left(\Omega_\delta\right)\right) = \rho_0 |\Omega_\delta| =$$

$$= \rho_0 \frac{4\pi}{3}\left((\delta + \Delta\delta)^3 - \delta^3\right) \approx \left(\rho_0 \frac{4\pi}{3} 3\Delta\delta\right)\delta^2 =$$

$$= (4\pi\rho_0\Delta\delta)\,\delta^2.$$

Then can be stated

$$\mathbb{E}\left(N_{ES}\left(\Omega_\delta\right)\right) p_{ES}(\delta) \approx (4\pi\rho_0\Delta\delta)\,\delta^2 p_{ES}(\delta)$$

Using an infinitesimal approach ($\Delta\delta \to 0$), the number of connections lost in truncation at δ_0 for each node is

$$\int_{\delta_0}^{+\infty} 4\pi\rho_0\delta^2 p_{ES}(\delta)d\delta$$

For the hypothesis exists a $M > 0$ such that for every $\delta_0 > M$ can be written

$$p_{ES}(d) < d^{-(3+\eta)}, \forall d \geq \delta_0$$

which implies

$$\int_{\delta_0}^{+\infty} 4\pi\rho_0\delta^2 p_{ES}(\delta)d\delta < \int_{\delta_0}^{+\infty} 4\pi\rho_0\delta^2\delta^{-(3+\eta)}d\delta =$$

$$= \int_{\delta_0}^{+\infty} 4\pi\rho_0\delta^{-(1+\eta)}d\delta = \frac{4\pi\rho_0}{\eta\delta_0^{\eta}}$$

In conclusion, observing that

$$\lim_{\delta_0 \to +\infty} \frac{4\pi\rho_0}{\eta\delta_0^{\eta}} = 0$$

then can be chosen $\hat{d} > M$ such that $\frac{4\pi\rho_0}{\eta\hat{d}^{\eta}} < \epsilon$.

Then

$$P(D_{ES}^I = k) \approx \left[*_{i=0}^{\iota} B_{p_i}^{N_i}(q) \right]_k$$

and

$$P(D_{ES}^O = k) \approx \left[*_{i=0}^{\iota} B_{p_i}^{N_i}(q) \right]_k$$

where ι is the smallest integer such as $r_\iota \geq \hat{d}$. In this truncation have been lost at most ϵ connection in average for each node.   □

In conclusion the condition on $p_{ES}(d)$ holds because in literature is often supposed to have an exponential decay (Billeh et al., 2020). Then can be proved that such ER distance model can be approximated by a finite convolution of binomials and this implies that (as observed in ER mixed section) its degree distributions are exponential. In general an exponential distribution does not fit experimental findings (Giacopelli et al., 2021).

3.4   Price spatial model

Since now has been introduced, just exponential spatial models. One of the first Power Law spatial models was the one introduced in (Fabrikant, Koutsoupias, and Papadimitriou, 2002). This model has been created to explain computer networks connectivity. The idea is quite simple: if the network administrator wants to connect a group of machines (s)he will try to minimize the wire required. Then done this premises can be defined the Fabrikant model

Definition 3.4.1 *The Fabrikant model is a growing spatial model such that every new node i is connected with the old node j that minimizes the cost function*

$$C_{ij} = \delta \left\| \underline{x}_i - \underline{x}_j \right\| + h_j$$

where h_j is the shortest path length from the first node of the network (called seed node) and δ is a parameter of the model.

In (Fabrikant, Koutsoupias, and Papadimitriou, 2002) has been proved that δ has a key role in network topology. In fact as can be seen in figure 3.3 for high values of δ the indegree of the network is exponential, for low values of δ is Power Law but

for $\delta < 1$ the network becomes a star network, a tree graph with one central point (the root) and many chains propagating from the center (the leaves).
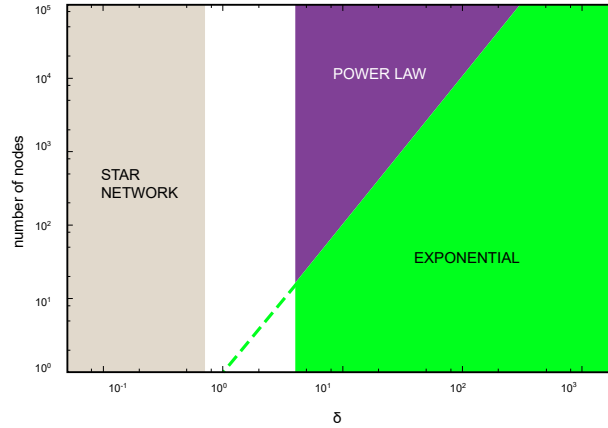
FIGURE 3.3: Schematics of Fabrikant model states (Fabrikant, Koutsoupias, and Papadimitriou, 2002)

However the outdgree of Fabrikant model (as the original BA model) is trivial, because every node added has just one outgoing connection that is created when the node is added to the network. For this reason has been introduced the Price spatial model (Giacopelli, Migliore, and Tegolo, 2020)

Definition 3.4.2 *A Price spatial model is a Fabrikant model where the network is initialized with a seed network (as in classic Price model) and every new node is connected to c old nodes, where c is the outcome of a random variable Γ (as in Price model), that minimize the cost function*

$$C_{ij} = \delta \left(\left\| \underline{x}_i - \underline{x}_j \right\| + \eta \ r_j \right) + h_j$$

where $r_j \in [0,1]$ is a random number and $\eta > 0$ a parameter.

The Price spatial model combines the cost minimization of the Fabrikant model with the fact that the outdegree follows the distribution of Γ. Since now, the adjacency matrix will be transposed (equivalently, all the connections will be inverted) such that indegree follows the distribution of Γ and the outdegree is Power Law.

3.5 Spatial Convolutive model

The natural evolution of aspatial Convolutive model is to assign to each subpopulation a spatial region of the space. Then can be defined

Definition 3.5.1 *Given a set of regions $\Omega_1, ..., \Omega_B \subseteq \mathbb{R}^3$ then a Spatial convolutive model is a Generalized Convolutive Model C with B blocks such that the positions of the nodes of each sub population C_i are inside the region Ω_i. The connectivity of a Spatial Convolutive model is a spatial Price model inside the blocks and the ER-like connectivity scheme of aspatial Convolutive model between blocks.*

The key idea of the spatial Convolutive model (since now SCM) is that the low range connectivity is Power Law and computed through cost minimization. Instead, long range connections are exponential and not space-dependent. So for a general SCM can be a theorem equivalent to theorem 2.5.1.

Theorem 3.5.1 *For a Spatial Convolutive Model C with B blocks can be proved that*

$$P(D_C^I = k) = \left[f_q^I * \left(((1-p)\delta_0(q) + pB_{\phi^U}^l(q)) * (p\delta_0(q) + (1-p)B_{\phi^D}^l(q)) \right)^{*(B-1)M} \right]_k$$

and

$$P(D_C^O = k) = \left[f_q^O * \left(((1-p)\delta_0(q) + pB_{\phi^U}^l(q)) * (p\delta_0(q) + (1-p)B_{\phi^D}^l(q)) \right)^{*(B-1)M} \right]_k$$

Proof: For theorem 2.3.2 can be written

$$P(D_C^I = k) = \sum_{j=1}^B \frac{|C_j|}{\sum_{l=1}^B |C_l|} \left[*_{i=1}^B P(D_C^I(C_j, C_i) = q) \right]_k =$$

$$= \sum_{j=1}^B \frac{|C_j|}{\sum_{l=1}^B |C_l|} \left[P(D_C^I(C_j, C_j) = q) * \left(*_{\substack{i=1 \\ j \neq i}}^B P(D_C^I(C_j, C_i) = q) \right) \right]_k =$$

being the connectivity inside the block the same spatial Price model with distributions f_k^I and f_k^O can be rewritten

$$= \sum_{j=1}^B \frac{|C_j|}{\sum_{l=1}^B |C_l|} \left[f_q^I * \left(*_{\substack{i=1 \\ j \neq i}}^B P(D_C^I(C_j, C_i) = q) \right) \right]_k =$$

using bilinearity on second component

$$= \left[f_q^I * \sum_{j=1}^B \frac{|C_j|}{\sum_{l=1}^B |C_l|} \left(*_{\substack{i=1 \\ j \neq i}}^B P(D_C^I(C_j, C_i) = q) \right) \right]_k =$$

can be observed that the terms $*_{\substack{i=1 \\ j \neq i}}^B P(D_C^I(C_j, C_i) = q)$ are all equal for every $j = 1, ..., B$ by symmetry, then

$$= \left[f_q^I * \left(*_{\substack{i=1 \\ j \neq i}}^B P(D_C^I(C_j, C_i) = q) \right) \right]_k =$$

as in theorem 2.5.1 the connectivty between blocks is

$$P(D_C^I(C_j, C_i) = k) = \left[\left(((1-p)\delta_0(q) + pB_{\phi^U}^l(q)) * (p\delta_0(q) + (1-p)B_{\phi^D}^l(q)) \right)^{*M} \right]_k$$

then substituting

$$= \left[f_q^I * \left(*_{i=1}^{B-1} \left(((1-p)\delta_0(q) + pB_{\phi^U}^l(q)) * (p\delta_0(q) + (1-p)B_{\phi^D}^l(q)) \right)^{*M} \right) \right]_k =$$

$$= \left[f_q^I * \left(\left(((1-p)\delta_0(q) + pB_{\phi^U}^l(q)) * (p\delta_0(q) + (1-p)B_{\phi^D}^l(q)) \right)^{*M} \right)^{*(B-1)} \right]_k =$$

$$= \left[f_q^I * \left(((1-p)\delta_0(q) + pB_{\phi^U}^l(q)) * (p\delta_0(q) + (1-p)B_{\phi^D}^l(q)) \right)^{*M(B-1)} \right]_k.$$

In conclusion

$$P(D_C^I = k) = \left[f_q^I * \left(\left((1-p)\delta_0(q) + pB_{\phi^U}^l(q) \right) * \left(p\delta_0(q) + (1-p)B_{\phi^D}^l(q) \right) \right)^{*(B-1)M} \right]_k.$$

Symmetrically can be proved for the outdegree that

$$P(D_C^O = k) = \left[f_q^O * \left(\left((1-p)\delta_0(q) + pB_{\phi^U}^l(q) \right) * \left(p\delta_0(q) + (1-p)B_{\phi^D}^l(q) \right) \right)^{*(B-1)M} \right]_k.$$

$\square$

Can be observed that the theorem 2.5.1 has the result of theorem 3.5.1 when $B = 2$. The fitting procedure is equivalent to the one exposed in theorem 2.5.3

Theorem 3.5.2 *Assigning the parameters m_0, ρ, l, B and the parameters of spatial Price model δ and η is possible to find a configuration of Spatial Convolutive Model such as*

$$P(D_C^I = k) \approx v_k^I$$

where v_k^I is the experimental network indegree distribution.

Proof: The proof is the same of theorem 2.5.3 except for p calculation

$$\mathbb{E}\left(\left(\left((1-p)\delta_0(q) + p\, B_{\phi^U}^l(q) \right) * \left(p\delta_0(q) + (1-p)\, B_{\phi^D}^l(q) \right) \right)^{*(B-1)M} \right) =$$

$$= M(B-1)\mathbb{E}\left(\left((1-p)\delta_0(q) + p\, B_{\phi^U}^l(q) \right) * \left(p\delta_0(q) + (1-p)\, B_{\phi^D}^l(q) \right) \right) =$$

$$= M(B-1)\left(\mathbb{E}\left((1-p)\delta_0(q) + p\, B_{\phi^U}^l(q) \right) + \mathbb{E}\left(p\delta_0(q) + (1-p)\, B_{\phi^D}^l(q) \right) \right) =$$

$$= (B-1)M(pl\phi^U + (1-p)l\phi^D) = N(B-1)(p\phi^U + (1-p)\phi^D)$$

that solved for p returns

$$p = \frac{\frac{E_K}{N(B-1)} - \phi^D}{\phi^U - \phi^D} \approx \frac{E_K}{N(B-1)}$$

for $\phi^U \approx 1$ and $\phi^D \approx 0$.

$\square$

Then is possible to fit the SCM to an arbitrary experimental indegree distribution. In next section will be exposed some examples.

3.6 Applications

Giacopelli et al. in (Giacopelli et al., 2021) introduce how to test the SCM effectiveness in just real world networks, and four fittings have been performed:

1. A *C. Elegans* (Cook et al., 2019) connectome reconstruction that has been reproduced using a two block SCM with $\delta = 1.5 \pm 0.5$, $E_k = 1 \pm 0.5$, $L = 1$, $\eta = 3 \pm 0.25$, $\phi_u = 1$ and $\phi_d = 0 \pm 0.00001$.

2. A slice of Rat Hyppocampus (Bonifazi et al., 2009) reproduced with a SCM with two blocks with unbalanced populations with parameters $\delta = 0 \pm 2$, $E_k = 0.1 \pm 0.05$, $L = 1$, $\eta = 3 \pm 0.25$, $\phi_u = 0.75$ and $\phi_d = 0 \pm 0.001$.

3. A mouse retina reconstruction (Helmstaedter et al., 2013) reproduced with a two blocks SCM (one block with connections inverted) with $\delta = 0 \pm 0.5$, $E_k = 70 \pm 0.05$, $L = 3 \pm 0.75$, $\eta = 3 \pm 0.25$, $\phi_u = 1$ and $\phi_d = 0 \pm 0.001$.

4. A stochastic reconstruction of a neocortical column (Markram et al., 2015) reproduced by a 6-layers with 9 blocks (in 3x3 scheme) for each layer model reproducing separately Excitatory and Inhibitory populations. In particular the connectivity Excitatory-Excitatory, Excitatory-Inhibitory, Inhibitory-Excitatory and Inhibitory-Inhibitory have been modeled by Spatial Convolutive models with parameters: (Excitatory-Excitatory) $\delta = 3 \pm 0.5$, $E_k = 100 \pm 0.5$, $L = 100 \pm 0.5$, $\eta = 3 \pm 0.25$, $\phi_u = 1$ and $\phi_d = 0.0001 \pm 0.00001$; (Excitatory-Inhibitory) $\delta = 4.52 \pm 1$, $E_k = 6.8 \pm 0.5$, $L = 6 \pm 1$, $\eta = 2.56 \pm 0.5$, $\phi_u = 1$ and $\phi_d = 0.0003 \pm 0.00025$; (Inhibitory-Inhibitory) $\delta = 1 \pm 0.5$, $E_k = 3 \pm 0.5$, $L = 4 \pm 1$, $\eta = 3 \pm 0.25$, $\phi_u = 1$ and $\phi_d = 0.0001 \pm 0.00001$; (Inhibitory-Excitatory) $\delta = 3.27 \pm 1$, $E_k = 7.05 \pm 0.5$, $L = 5 \pm 1$, $\eta = 3.66 \pm 0.5$, $\phi_u = 1$ and $\phi_d = 0.0006 \pm 0.00025$.

Figure 3.4 shows the results. It can be seen that the model proposed using the SCM can fit the degree distributions for the full degree range of variability. The SCM (by construction) supports parallel computing because a different worker can compute the connectivity of each block. Using this feature is possible to fully reconstruct the Neocortical model (Markram et al., 2015) in about 7 minutes on Desktop PC with 16 workers, having a computation time of 2 *core* $\times$ *hours* against the 16 *core* $\times$ *hours* required for a general touch detection based model (Hjorth et al., 2020).
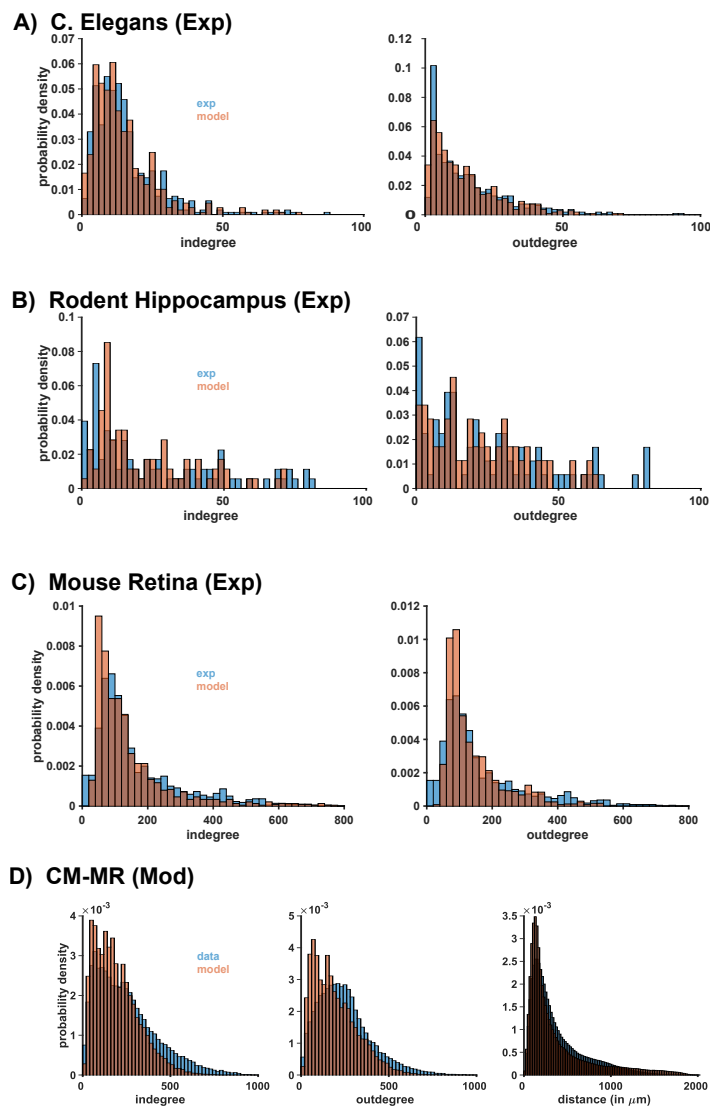
FIGURE 3.4: Applications of the Spatial Convolutive model in experimental networks fittings (Giacopelli et al., 2021).

Network dynamics

4.1 Introduction to network dynamics

Since now has been analyzed the topology of neuronal networks. However, the crucial role of neuronal electric activity in brain cognitive processes is well known. For these reasons will be important not only to describe the neuronal networks in terms of graph theory but also to determine how a network of neurons wired behaves. Then the first step is to determine a neuron model able to simulate neuronal electric activity. There are many neuronal models, but the most accurate is the Hodgking-Huxley model (Hodgkin, Huxley, and Eccles, 1952). However this model is made up by several ODE (Ordinary Differential Equations) and some PDE (Partial Differential Equations). Because of its complexity to the Hodgking-Huxley model in the applications are preferred models less accurate but simpler. A popular model in the applications is the *Leaky, Integrate and Fire model* (Dutta et al., 2017) (since now LIF). In figure 4.1 can be seen the RC circuit that composes the LIF. It is a voltage generator connected in pipeline with a resistor and a capacitor with an external source of current.
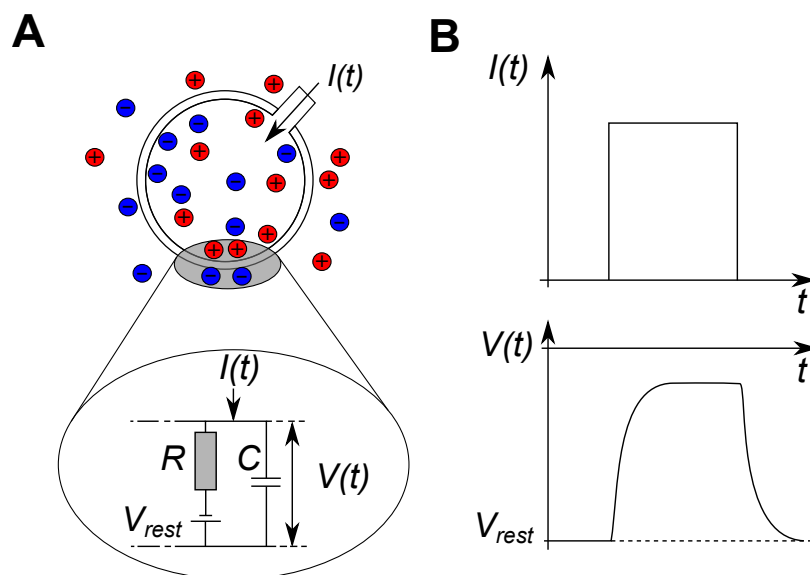
FIGURE 4.1: LIF circuit plot. (source https://neuronaldynamics.epfl.ch/online/Ch1.S3.html)

The main equation of LIF model is

$$I(t) = I_R + I_C$$

where $I(t)$ is the external current, I_R is the resistor current and I_C the capacitor current. Then can be observed that defining V the membrane potential, then $I_C = C\frac{dV}{dt}$ and $I_R = \frac{\Delta V}{R} = \frac{V - V_{rest}}{R}$. Then holds

$$I(t) = \frac{V - V_{rest}}{R} + C\frac{dV}{dt}$$

and then

$$C\frac{dV}{dt} = -\frac{V - V_{rest}}{R} + I(t)$$

multiplying by R

$$RC\frac{dV}{dt} = \tau\frac{dV}{dt} = -(V - V_{rest}) + RI(t)$$

where $\tau = RC$ is the membrane time constant. Observing that $\frac{dV}{dt} = \frac{d(V - V_{rest})}{dt} = \frac{d\,\Delta V}{dt}$ often the previous equation takes the form

$$\tau\frac{d\,\Delta V}{dt} = -\Delta V + RI(t)$$

for this reason can be assumed $V_{rest} = 0$ since now (even if experimental findings suggest that is not the case).

However, this differential equation alone cannot model neuronal activity because it is observed that if the neurons reach a threshold potential, they "fire", which means that they come back to the resting potential V_{rest}. Moreover, they keep this value for an interval of time called refractory time. For this reason, we add to the ODE model the condition that if V reaches the value V_{thr} then V is reset to V_{rest} and keeps this value for a refractory time τ_r. Since now, it seems that neurons connectivity has no role in the network activity, but this is not the case because the term $I(t)$ is still unknown. This term usually includes the synaptic connections; for example, in (Brunel, 1999) given a neuron i this term takes the form

$$I_i(t) = \tau\sum_{j=1}^{N} J_{ji}\sum_{k}\delta(t - t_j^k + D)$$

where J_{ji} is not zero if there is a connection starting from the neuron j and arriving to neuron i and in particular $J_{ji} > 0$ if the connection is excitatory and $J_{ji} < 0$ if the connection is inhibitory. The term $\delta(t - t_j^k + D)$ says that if the neuron j fire at the time t_j^k then it generates a reaction that is a delta distribution. In conclusion supposing G the connectome of the network, $J_{hj} = 0$ if and only if $(A_G)_{ji} = 0$. Then is clear the key role of connectivity of network dynamics.

4.2 Spectral analysis for physiological signals

The idea behind the spectral analysis is the concept of continuous Fourier Transform

Definition 4.2.1 *Given a signal $x(t)$ the Continuous Fourier transform of the signal is the variable $\hat{x}(\omega)$ defined as*

$$\hat{x}(\omega) = \int_{-\infty}^{+\infty} e^{2\pi i\omega\tau} x(\tau)\, d\tau$$

In general the variable $\hat{x}(\omega)$ is called spectrum of $x(t)$ and it is dependent from the frequency ω. However can be seen that the Fourier transform is an integral

dependent from every value of $x(t)$ for every time t. In the applications this could be undesired because in some cases could be not interesting to take in account times very far from experiment time. For these reasons, the concept of the Short Time Fourier Transform (since now STFT) is prevalent in applications

Definition 4.2.2 *The Short Time Fourier Transform of the signal $x(t)$, centered in the time instant t_0 and with time window T is the function $\hat{x}(\omega, t_0)$ defined as*

$$\hat{x}(\omega, t_0) = \int_{t_0 - \frac{T}{2}}^{t_0 + \frac{T}{2}} e^{2\pi i \omega \tau} x(\tau) \, d\tau$$

Usually the signal of a neuron h of LIF model can be expressed as

$$x_h(t) = \sum_k \delta(t - t_h^k)$$

where the times t_h^k are usually the firing times of the neuron h. Then the STFT of the signal is

$$\hat{x}_h(\omega, t_0) = \int_{t_0 - \frac{T}{2}}^{t_0 + \frac{T}{2}} e^{2\pi i \omega \tau} x_h(\tau) \, d\tau = \sum_{k \,|\, \left| t_h^k - t_0 \right| < \frac{T}{2}} e^{2\pi i \omega t_h^k}.$$

Can be observed that $\hat{x}_h(\omega, t_0) \in \mathbb{C}$, defined the conjugate of a complex number z as $\overline{z}$, then it is defined as in (Dummer, Wieland, and Lindner, 2014) the modulus inside the time window

$$S_h(\omega, t_0) = \frac{\sqrt{\hat{x}_h(\omega, t_0) \, \overline{\hat{x}_h(\omega, t_0)}}}{T}.$$

However, the previous measures are dependent from the neuron h, in order to define an ensemble measure will be defined the measure

$$L(\omega, t_0) = \frac{\sum_{h=1}^{N_G} \log_{10} \left(\max \left(S_h(\omega, t_0), 10^{-4} \right) \right)}{N_G}.$$

4.3 A study case on Neocortex

Up to now there are two main models of the Neocortical area:

1. An ER mixed model (Potjans and Diesmann, 2012) (since now PD) that fits the probability of connection observed between neuronal populations of Neocortex (figure 4.2A). The original network has about 80,000 neurons, however in our experiments will be used the same ER mixed model with the same probabilities, but with about 30,000 neurons.

2. A stochastic data-driven model (Markram et al., 2015) (since now MR) based on techniques of touch detection and pruning (figure 4.2B) made up by 30,000 neurons.

The model exposed in (Brunel, 1999) (since now BR) has been taken into account to investigate such two models. It is based on an ER model for connectivity and a LIF for neuronal activity. The connectivity of neuronal network of BR is independent from the LIF neuronal activity model. Then has been used the neuronal activity model of (Brunel, 1999) to simulate the activity of the networks PD and MR. These

simulations will be called since now now background activity of the network. Then has been added a sinusoidal current generator to each neuron i, making the $I_i(t)$ equation of LIF model

$$I_i(t) = \tau \sum_{j=1}^{N} J_{ji} \sum_k \delta(t - t_j^k + D) + A \sin(2\pi f t)$$

Where A is called amplitude and f is called frequency of the generator.

A)

B)
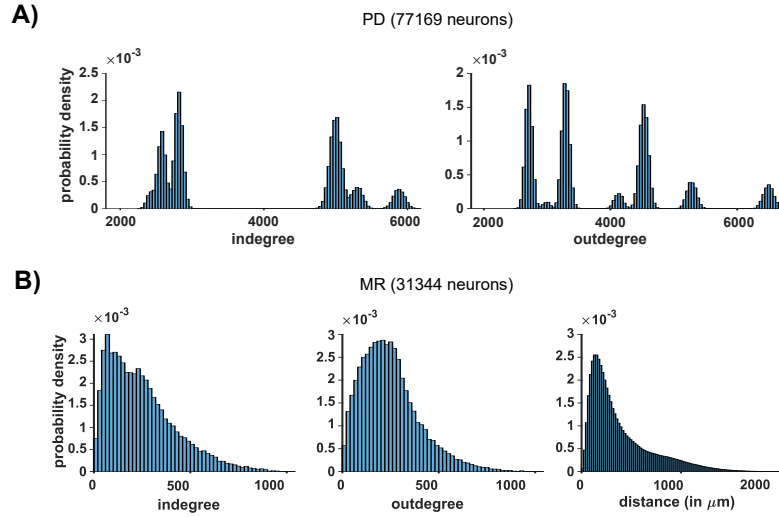
FIGURE 4.2: Two data driven models of Neocortex. A) This model is an aspatial ER mixed model with about 80,000 neurons (Potjans and Diesmann, 2012). B) This model is a touch detection based model composed by about 30,000 neurons (Markram et al., 2015).

Then has been computed (on 150 randomly chosen neurons) the measure $L(\omega, t_0)$ in different conditions. The first condition is when the generator is off (Background activity) and will be defined as $L_B(\omega, t_0)$. The second condition is when the generator is on with amplitudes A ranging from 0 to 100 pA and frequencies f ranging from 0 to 80 Hz and it will be defined as $L_S^{(f,A)}(\omega, t_0)$.

Then for each couple (f, A) has been computed the quantity

$$\Delta_S^{(f,A)}(\omega, t_0) = L_S^{(f,A)}(\omega, t_0) - L_B(\omega, t_0)$$

that represents how much the perturbed network deviates from the background activity and then this quantity has been mediated on time defining

$$D_S^{(f,A)}(\omega) = \left\langle \Delta_S^{(f,A)}(\omega, t_0) \right\rangle_{t_0}$$

where the brackets $\langle \cdot \rangle_{t_0}$ represent the average on the variable t_0. Using Parseval's Theorem the power of the response has been calculated as

$$R_S^{(f,A)} = \int_0^W 10^{2D_S^{(f,A)}(\omega)} \, d\omega$$

where W is the highest harmonic considered (in our case 250 Hz). The aim of this measure is to quantify the strength of the reaction of the network to the generator.

The results of this analysis are in figure 4.3. For example, in figure 4.3A can be seen how BR model is very reactive in the ranges δ to θ and in low γ but it has not significant reactions in the remaining frequencies. Similarly, in figure 4.3B can be seen how PD Excitatory neurons are reactive just for low frequencies, and they do not react to higher frequencies. In contrast, the MR model has a broad range of reactivity that is coherent to what is observed experimentally (Buzsáki and Draguhn, 2004). In fact, it is rare to see a neuronal network tune in small ranges of frequencies.
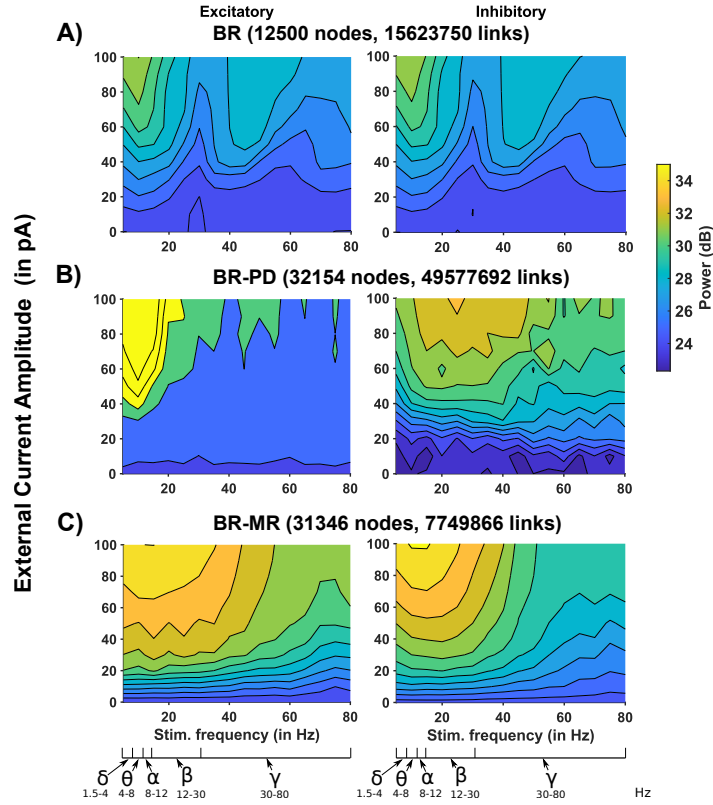
FIGURE 4.3: Power spectra as a function of external stimulation amplitude and frequency. (Left) Excitatory neurons, (Right) Inhibitory neurons. A) The network proposed in (Brunel, 1999) with about 12,500 neurons. B) The network connected as in (Potjans and Diesmann, 2012) with 30,000 neurons. C) The network connected as in (Markram et al., 2015) with 30,000 neurons. Major brain rhythms frequency range (from (Buzsáki and Draguhn, 2004)) are shown below the stimulation frequency axis

Another interesting measure is the periodicity of the response that is computed starting from a partition of the interval in Q bands. Every band has subdivided in S points creating the numbers $\{w_k^p\}_{k=1,...,S}$ for every $p = 1, ..., Q$. Then for every couple $(p, p + 1)$ has been computed a measure given by

$$C_p^{(f,A)} = \frac{\sum_{j=1}^{S} D_S^{(f,A)}\left(w_j^p\right) D_S^{(f,A)}\left(w_j^{p+1}\right)}{\sqrt{\sum_{j=1}^{S} D_S^{(f,A)}\left(w_j^p\right)^2} \sqrt{\sum_{j=1}^{S} D_S^{(f,A)}\left(w_j^{p+1}\right)^2}}$$

and averaged over p to compute the Average Cosine Distance

$$C^{(f,A)} = \frac{\sum_{p=1}^{Q-1} C_p^{(f,A)}}{Q-1}.$$

The last measure quantify the periodicity of the neuronal activity, measuring how much the network follows the generator signal frequency.

The results are shown in figure 4.4. In figure 4.4A can be seen that the BR network is more resonant for β range and high γ. In 4.4B is shown that the PD model is resonant for every frequency, and it is not a good feature because it is experimentally observed (Buzsáki and Draguhn, 2004) that exist in brain key frequencies representing different brain states. In 4.4C is shown that the excitatory population is more resonant for low frequencies; instead, the inhibitory population is resonant for low frequencies and high γ range.
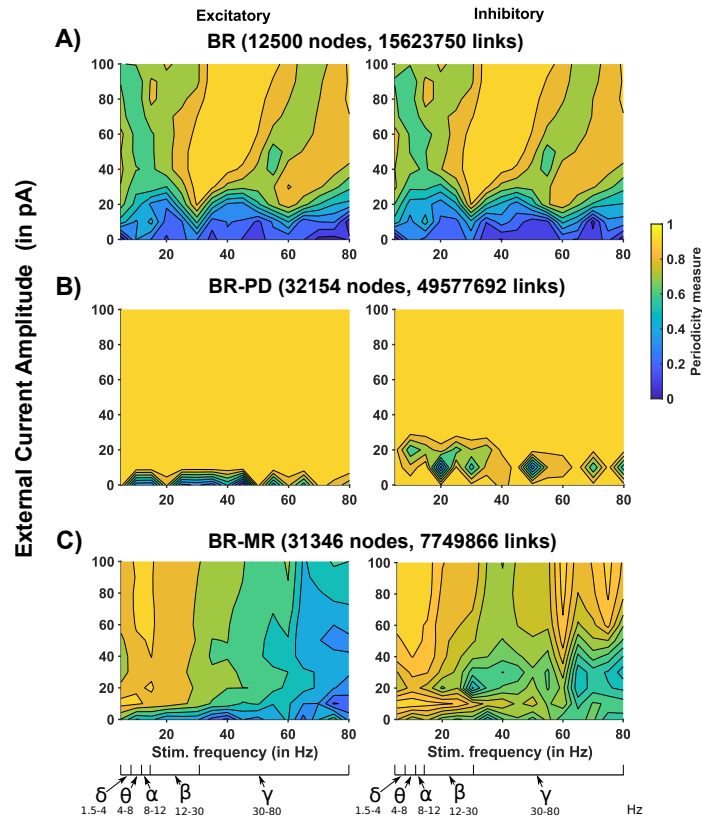
FIGURE 4.4: Periodicity as a function of external stimulation amplitude and frequency. (Left) Excitatory neurons, (Right) Inhibitory neurons. A) The network proposed in (Brunel, 1999) with about 12,500 neurons. B) The network connected as in (Potjans and Diesmann, 2012) with 30,000 neurons. C) The network connected as in (Markram et al., 2015) with 30,000 neurons. Major brain rhythms frequency range (from (Buzsáki and Draguhn, 2004)) are shown below the stimulation frequency axis

In conclusion has been performed the same tests on the SCM in figure 3.4D. The results of these analysis are shown in figure 4.5. In figure 4.5A can be seen how the SCM proposed is able to reproduce the Power Spectra of MR model (figure 4.3C) and in figure 4.5B how is able to reproduce the periodicity measure plot of MR model

(figure 4.4C). In conclusion, not only the SCM is able to reproduce the degree distributions experimentally observed, but in in this tests it was also able to reproduce the neuronal activity of a realistic network.
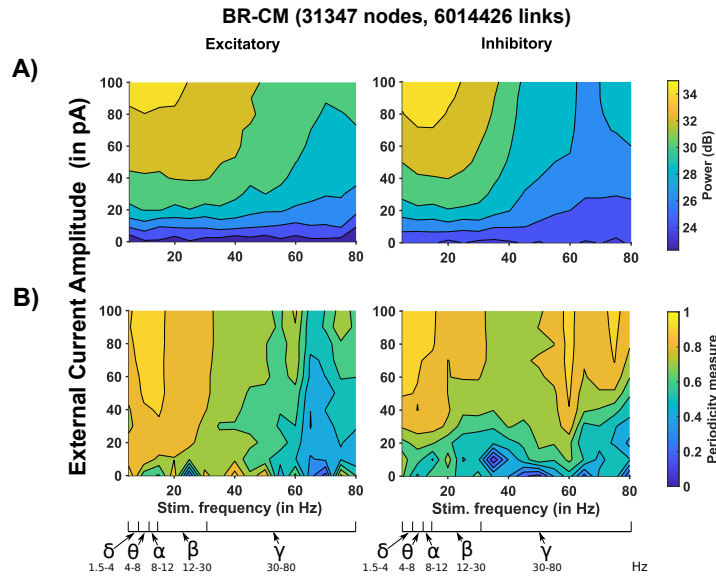
FIGURE 4.5: Network activity using the Convolutive Model connectivity. A) Power spectra as a function of external stimulation amplitude and frequency of Excitatory (Left) and Inhibitory (Right) neurons for the Convolutive model (Giacopelli et al., 2021; Giacopelli, 2021) fitting the MR network (Markram et al., 2015); B) Periodicity measure as a function of external stimulation amplitude and frequency of Excitatory (Left) and Inhibitory (Right) neurons for the Convolutive model (Giacopelli et al., 2021; Giacopelli, 2021) fitting the MR network (Markram et al., 2015).

4.4 Hindamarsh and Rose model

As been pointed out in the previous sections, even if since now the Hodgkin-Huxley model (Hodgkin, Huxley, and Eccles, 1952) is the most realistic model to simulate neuron behavior, in the applications are preferred simpler models. One of the most popular ODE based model is Hindamarsh-Rose model (since now HR model) (Hindmarsh and Rose, 1984). This model is able to explain the spkining activity through the non-linearity of the differential equations. In particular will be adopted the model exposed in (Stefanescu and Jirsa, 2008) that is composed by a set of equations able to model Excitatory and Inhibitory neurons. The equations of a general Excitatory neuron i are

$$\tau\frac{dx_{Ei}}{dt} = \nu y_{Ei} - a\frac{x_{Ei}^3}{\nu^2} + b\frac{x_{Ei}^2}{\nu} - \nu z_{Ei} + K\frac{\sum_{k=1}^{N_E} A_{ik}^{EE}(x_{Ek}-x_{Ei})}{N_E} - nK\frac{\sum_{k=1}^{N_I} A_{ik}^{EI}(x_{Ik}-x_{Ei})}{N_I} + \sigma I(t,m) + \omega(t)$$

$$\tau\frac{dy_{Ei}}{dt} = c - d\frac{x_{Ei}^2}{\nu^2} - y_{Ei}$$

$$\tau\frac{dz_{Ei}}{dt} = r\left(s\frac{x_{Ei}-x_0}{\nu} - z_{Ei}\right)$$

Analogously the equations for a general Inhibitory neuron j are

$$\tau \frac{dx_{Ij}}{dt} = v y_{Ij} - a \frac{x_{Ij}^3}{v^2} + b \frac{x_{Ij}^2}{v} - v z_{Ij} + K \frac{\sum_{k=1}^{N_E} A_{jk}^{IE}(x_{Ek} - x_{Ij})}{N_E} + \sigma I(t, m) + \omega(t)$$

$$\tau \frac{dy_{Ij}}{dt} = c - d \frac{x_{Ij}^2}{v^2} - y_{Ij}$$

$$\tau \frac{dz_{Ij}}{dt} = r \left(s \frac{x_{Ij} - x_0}{v} - z_{Ij} \right)$$

Where N_E and N_I are respectively the numbers of Excitatory and Inhibitory neurons. The parameters are $a = 1$, $b = 3$, $d = 5$, $s = 4$, $r = 0.003$, $x_0 = -1.6$, $K = 150$ and $\omega(t)$ is a white noise. The matrices A^{EE}, A^{EI} and A^{IE} are the adjacency matrices for every type of connections (Excitatory-Excitatory, Excitatory-Inhibitory and Inhibitory-Excitatory; Inhibitory-Inhibitory connections were ignored, as in (Stefanescu and Jirsa, 2008)). The terms $I(m, t)$ and n will be discussed in the next sections.

The HR model introduced in (Stefanescu and Jirsa, 2008) has been successfully used to construct a mean field model (the Epileptor) that is used to treat drug resistant Epilepsy. However, the Epileptor does not consider the micro-circuit connectivity, ignoring the role of connectivity in Epileptic activity.

4.5 Network connectivity and Simulation framework

In order to establish the role of network connectivity in Epileptic activity, two case studies have been considered:

1. In the Thick case (figure 4.6A) has been taken in account a network with the degree distributions of the *C: Elegans* (Cook et al., 2019) and the spatial distribution of neocortical column of (Markram et al., 2015), that is a parallelepiped with dimensions $500 \times 500 \times 2000 \ \mu m$.

2. In the Thin case (figure 4.6B) has been taken in account the hyppocamapal slice network of (Bonifazi et al., 2009) with a slice spatial distribution in a parallelpiped with dimensions $400 \times 300 \times 10 \ \mu m$.

These two case studies have been chosen because they represent two physiological and widely analyzed case studies in literature (Markram et al., 2015; Bonifazi et al., 2009; Cook et al., 2019; Potjans and Diesmann, 2012; Stefanescu and Jirsa, 2008).

For each case study has been created an ER model (Erdös and Rényi, 1959) (figure 4.6, red line), an ER spatial (Giacopelli, Migliore, and Tegolo, 2020) (figure 4.6, blue line) and a Convolutive model (Giacopelli et al., 2021) (figure 4.6, green line) fitting the experimental degree distributions (figure 4.6, black line).

In particular the parameters of the models are for the Thick volume case three kinds of connectivity

1. following an ER model, ignoring the distance among neurons, and with a fixed connection probability of $p = 0.026$, corresponding to the average connection probability of a C. Elegans brain (figure 4.6A red traces)

2. following an ER model but with a connection probability depending on the distance between neurons as $p(d) = Ae^{-Bd}$ (as in (Billeh et al., 2020)), where d
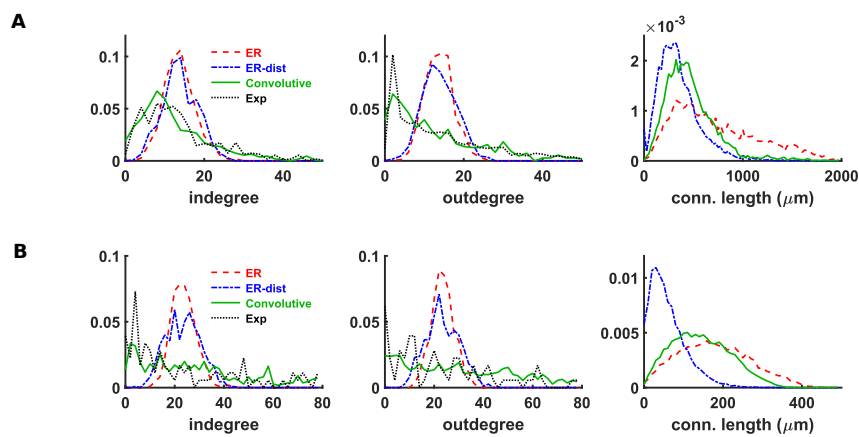
FIGURE 4.6: A) Indegree distribution (left), outdegree distribution (middle) and connection length distribution (right) for networks of Epileptors connected as an ER model (red), ER with distance (blue) and using a convolutive model (green) fitted to experimental C. elegans data (black, connection length was not available). B) Indegree distribution (left), outdegree distribution (middle) and connection length distribution (right) for networks of Epileptors connected as an ER model (red), ER with distance (blue), and using a convolutive model (green), connected as observed in an experimental hippocampal slice (black, (Bonifazi et al., 2009)), implemented with 4 blocks and 2 sub-blocks containing 1% and 99% of the neurons, respectively, and with parameters $\delta = 0.25$, $E_k = 0.1$ and $\eta = 3$.

  is the distance between two nodes, $A = 0.2$, and $B = 0.004$ (figure 4.6A blue traces)

3. using a convolutive model fitting a C. Elegans degree distributions (Cook et al., 2019), with two blocks and parameters, $\delta = 1.5$, $E_k = 1$ and $\eta = 3$ (Giacopelli, Tegolo, and Migliore, 2021) (figure 4.6A, green traces)

for the Thin volume case other three kinds of connectivity

1. following an ER model, ignoring the distance among neurons, and with a fixed connection probability of $p = 0.0438$, cconsistent with that observed in a hippocampal slice (figure 4.6B red traces)

2. following an ER model but with a connection probability depending on the distance between neurons as $p(d) = Ae^{-Bd}$ (as in (Billeh et al., 2020)), where d is the distance between two nodes, $A = 0.7$ and $B = 0.025$ (figure 4.6B blue traces)

3. using a convolutive model fitting a slice hippocampus network degree distributions (Bonifazi et al., 2009), with two blocks with respectively th 1% and 99% of neurons and parameters, $\delta = 0.25$, $E_k = 0.1$ and $\eta = 3$ (Giacopelli, Tegolo, and Migliore, 2021)(figure 4.6B, green traces)

The idea of the proposed model is to add to every neuron a space dependent current $I_i(m, t)$, where $m = (m_x, m_y, m_z)$ is the position of the neuron i, of the form

$$I_i(t, m) = S(t) \left(I_{max} e^{-\frac{m_x^2 + m_y^2 + m_z^2}{2\sigma_{xyz}^2}} + I_{GND} \right) + \zeta_i$$

where $I_{max} = \{1.5nA, 2.5nA, 5nA\}$, $I_{GND} = 1nA$, $S(t)$ is a time-dependent component of the current, with values in $[0, 1]$, σ_{xyz} is 600 μm for Thick case and 120 μm for the Thin case, and ζ_i is a random normal vector with amplitude 0.1 nA.

4.6 Epileptiform activity and Mean Amplitude

The parameter n is called inhibitory strength (Stefanescu and Jirsa, 2008) and it has the key role of modulating the Inhibitory response in the network. In figure 4.7 can be seen a set of simulated tracks with low inhibition ($n = 0.1$).

  Figure 4.7A shows how the intensity of current $I(t, m)$ is applied to the networks. In particular, at the beginning of the simulations, a low current is applied to make networks reach a steady state (that since now will be called background activity). After this, at second 5, the intensity of current increases reaching its maximum at second 9. Then the current is left steady until the gradual descent in the second 12.

  In addition, has been performed a quantitative analysis to estimate the impact of inhibitory strength n on the Epileptiform activity. In particular, the mean amplitude of the signal has been taken into account (Ma, Zheng, and Peng, 2021). This feature has a key role in Epilepsy detection because one of the most common symptoms of an Epileptic seizure is the appearance of high amplitude oscillations in the EEG patient's track (Devinsky et al., 2018). In order to perform an exhaustive quantitative analysis has been chosen the values 1.5 nA, 2 nA and 5 nA of I_{max} as representative of a small, medium and strong stimulation. Then for each value of intensity of stimulation has been computed 5 independent simulations for each value of the inhibitory
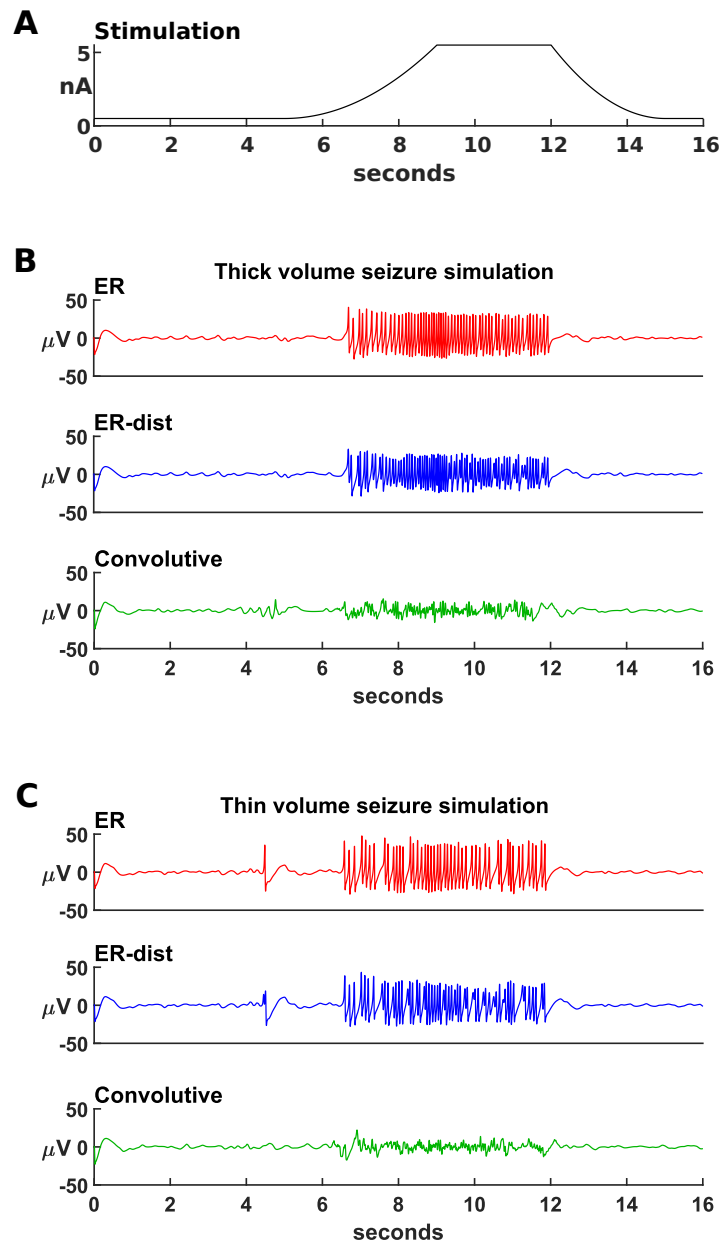
FIGURE 4.7: A) Maximum input current, in the center of the network, as a function of time; B) Simulation of epileptiform activity in networks in the Thick case study with low inhibitory interaction ($n = 0.1$). C) Same as in B) but in the Thin case study.

strength n in the values 0.05, 0.1 ,0.15, 0.2, 0.3, 0.4, 0.5, 0.75, 1.0, 1.5 and 2.0. The results are shown in figure 4.8. In the first row, there are the plots referred to the Thick network, and in the second row, there are the plots referred to the Thin network. It can be noted that in both cases, the ER model (red line of figure 4.8) shows bigger amplitudes for low inhibition. The ER distance shows amplitudes comparable to the ER model for 1.5 nA and 2 nA. However, it shows weaker mean amplitude values for 5 nA. In contrast the SCM proposed for each value of n and I_{max} are always below the threshold of 20 μV that has been considered as Epileptic threshold.
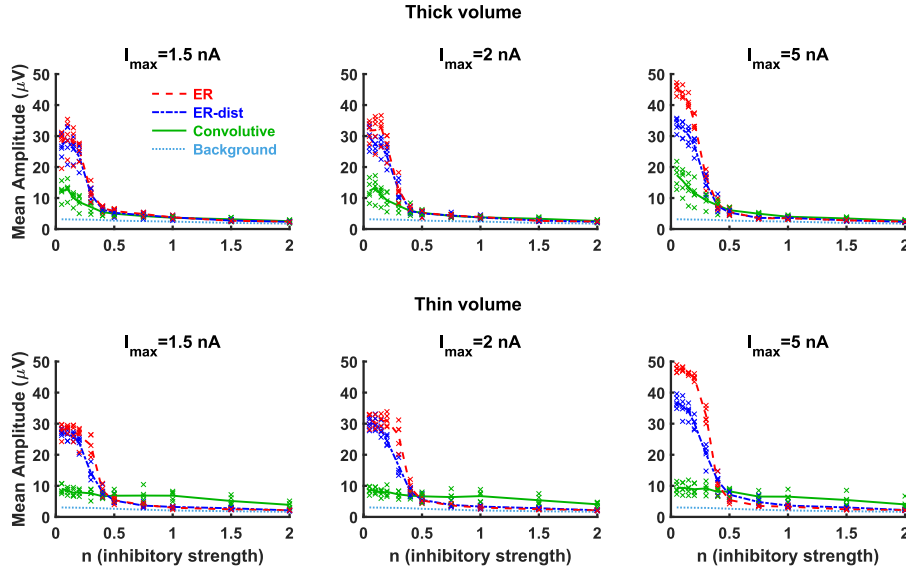
FIGURE 4.8: Average mean amplitude of the different models calculated, as a function of the inhibition strength, over the 10 sec duration of an external input of different maximum strengths, I_{max}. For any given value of I_{max} and n, symbols represent the average membrane potential obtained from five network instances. Lines represent their average value.

4.7 Spike-wave complexes

In literature have been observed during the EEG recordings 3Hz frequency discharges with variable amplitudes (Saggio et al., 2020), these discharges are often called Spike-Wave Complexes (Noachtar and Rémi, 2009). The underlying idea is that exists a mechanism of dynamic inhibition (Destexhe, Contreras, and Steriade, 2001) able to increase the inhibitory strength n when the overall activity of the network becomes too intense. In formula this mechanism can be described by two differential equations

$$\tau \frac{dn}{dt} = \rho \left(\theta - \theta_0 \right) \left(n_0 - n \right) n$$

$$\tau \frac{d\theta}{dt} = \rho_\theta \left(\frac{X(t)}{\nu} - \theta \right)$$

Where $\tau = 0.0167s$, $\nu = 20\mu V$, $n_0 = 2$, $\rho = 0.05$, $\rho_\theta = 0.15$ and θ_0 is -0.9 for the Thick case and -1 for the Thin case. The first equation is a logistic equation tuned by the variable θ. If $\theta > \theta_0$ then n converges to $n_0 = 2$ with a logistic growth. This mean

that if the variable θ (that follows the mean potential X) is over the threshold, then the network increases the inhibition to reduce the mean activity. If $\theta < \theta_0$ the value decreases to 0, then if the network activity is below the threshold then the network comes back to the minimum inhibition. The second equation rules the behavior of θ making it converge to the value $\frac{X(t)}{\nu}$ so slowly to start a delay dynamics, creating the spike-wave behavior observed in simulations. The idea behind this mechanism is to simulate the reaction of the network increasing the inhibitory strength n for high values of $X(t)$. This model is able to reproduce the behavior observed in the experimental findings (Noachtar and Rémi, 2009) in the cases of Exponential topology, in contrast the results obtained for convolutive topology are slightly affected by the phenomenon.
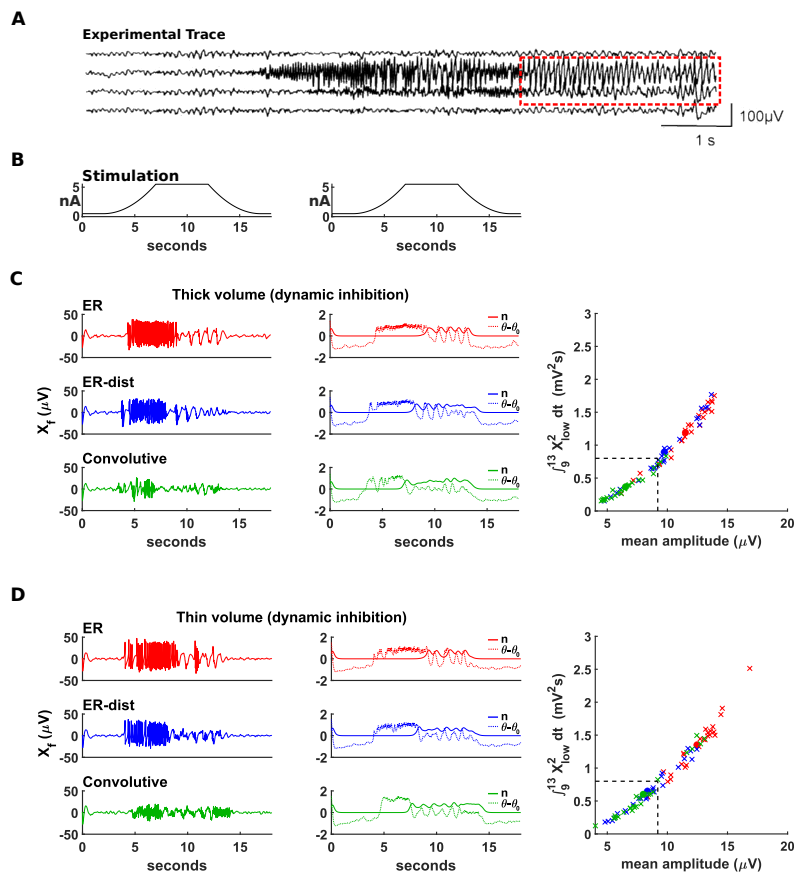
FIGURE 4.9: A) Experimental EEG traces showing a Spike-Wave occurrence during the recording (Noachtar and Rémi, 2009) B) maximum input current used in simulation in function of time; C) Most representative traces in a set of 25 simulation per topology case. (left) Filtered membrane potential track. (middle) plot of n in function of time (solid line) and of $(\theta - \theta_0)$ in function of time (dashed line). (right) Plot of all the 25 simulations in the plane with x axis mean amplitude and y axis the quantity $\int_9^{13} X_{low}^2 dt$, where $X_{low}(t)$ is a low pass filtering of the average potential $X(t)$. D) Same as in (C) but for neurons distributed in a thin volume

4.8   The role of topology in brain activity

In conclusion, it has been proved in (Giacopelli et al., 2021) that brain connectivity has a central role in brain activity. In fact, a change of network topology can cause a significant variation in the resonance frequencies of the connectome. In (Giacopelli, Tegolo, and Migliore, 2021) it is also proved that network topology has a crucial role in Epileptic seizure occurrence, leading to further reflections on current Epilepsy treatment. Current anti-epileptic drugs are based on reduction of inhibitory response (Rogawski and Löscher, 2004), which may lead to several adverse collateral effects (St Louis, 2009). However, simulations suggest that better results can be achieved by increasing inhibition selectively when needed. Furthermore, some experimental findings suggest that it could be possible that synapsin can regulate the GABA release (Song and Augustine, 2016), which is a neurotransmitter involved in the short-time plasticity of synapses. Also, anti-epileptic drugs can alter short-time plasticity (Gholmieh, Chen, and Courellis, 2007). Another promising field of study is to apply optogenetics and designer receptor technologies (Krook-Magnuson and Soltesz, 2015) showing promising results in avoiding seizure onset.

Segmentation Methods for Immunofluorescence images

5.1 Immunofluorescence images

One of modern automatic image analysis objectives is to identify predictive and prognostic elements useful for highlighting possible pathologies. The images analyzed can usually be fMRI, PET or Immunofluorescence images. The following section will focus on the last kind of image. Immunofluorescence (IF) is a technique to highlight a portion of tissue using specific antibodies that create luminescence (Im et al., 2019). The result of IF is a microscopy image that shows the area affected by the antibody brighter than the regions not affected by the antibody. Most of the algorithms used to process these images work at the grayscale level. Such images are into the dataset introduced in (Kromp et al., 2020), but that hypothesis is not valid for every dataset; for example, some datasets are RGB, and they have most of the information contained in the green channel (Nigam et al., 2015) or even on the blue channel (Gunesli, Sokmensuer, and Gunduz-Demir, 2020). For these reasons, in the proposed analysis will be implied that the starting point will be a grayscale version of every dataset containing as much information as possible.

5.2 Segmentation task and evaluation metrics

The segmentation of an image is the task to subdivide the area of an image into subregions. A classic example of segmentation is semantic segmentation, which is the task of marking the surface occupied by a particular object in an image. In the case of interest, the objects will be the cells of IF images. There are two main approaches for this task:

1. The binary approach returns a binary image where every pixel inside a cell is marked with white color otherwise it is marked with black. This approach considers all of the cells as a single object because each pixel has not the cell information that contains it.

2. The cell-by-cell approach assigns a different color to every cell. As a result, the cells are distinguishable and, in general, will be numbered with natural numbers.

Since now will be evaluated just the binary segmentations of the methods discussed, even if some of them return a cell-by-cell segmentation. In general, a bit of a binary cell segmentation can be said Positive if the segmentation says there is a cell on it or Negative otherwise. There are many metrics with slightly different interpretations, but every metric proposed is based on four sets: True Positives (TP), True Negatives (TN), False Positives (FP) and False Negatives (FN). The TP set is the

set of the pixels Positive for the ground truth and the prediction. Similarly, the TN set is the set of the pixels Negative for the ground truth and the prediction. The FP set is the set of the pixels Positive for the prediction and Negative for the ground truth. Analogously, the FN set is the set of the pixels Negative for the prediction and Positive for the ground truth. Often some metrics are used to estimate how much a prediction is close to the ground truth. The most common are:

1. Intersection over Union (IoU): is defined as $\frac{TP}{TP+FP+FN}$ and it is one of the most balanced metrics.

2. F1-score: is defined as $\frac{2\,TP}{2\,TP+FP+FN}$ and can be proved being almost proportional to the IoU.

3. Accuracy: is defined as $\frac{TP+TN}{TP+FP+TN+FN}$ and is one of the most popular metrics of Machine Learning. However, in cell segmentation tasks this metric can be biased in cases of sparse little cells, in these cases the number of Negative pixels can be very greater than the number of Positives pixels. This means that even if the prediction is fully Negative (every pixel is negative), if the ground truth ratio P/N tends to 0 then the Accuracy tends to 1.

4. Sensitivity: is defined as $\frac{TP}{TP+FN}$ and can be biased if the ground truth ratio N/P tends to 0.

5. Specificity: is defined as $\frac{TN}{TN+FP}$ and can be biased if the ground truth ratio P/N tends to 0.

5.3   Machine learning approaches

Neural networks have proved in recent years to be a valid solution to various computational problems (Deng and Yu, 2014), (Khan et al., 2020), (Tajbakhsh et al., 2016) in particular in segmentation and classification tasks (Alzubaidi et al., 2021), (Wainberg et al., 2018), (Khan et al., 2020). The basic technique consists of automatic learning from a training set and then the evaluation on a test set, to assess the performances. The Convolution Neural Network (CNN) can be identified among the most performing neuronal networks, made up of pairs of convolutional levels coupled to connected levels. The three networks analyzed in this section will be an evolution of the concept of CNN. The neural networks taken in account are:

1. UNet: The U-Net is a deep learning network (Ronneberger, Fischer, and Brox, 2015) for image processing. The idea is to scale down the information of the input image through convolution layers and then scale up the information through transposed convolutional layers to obtain an image with the exact resolution of the original with the information of the semantic segmentation in each pixel. The U-Net is simple in architecture, fast to train (the U-Net used in this paper (Ronneberger, Fischer, and Brox, 2015) has been trained on a GPU RTX 2070 with 8 GB of VRAM in about 1 hour) and requires few computational resources in prediction (Wu et al., 2019), making this network suitable for general purpose tasks. The main downside of this architecture is that in cells segmentation task returns a binary label and cannot separate the single cells by default. Therefore, this contribution will use the Keras/Tensorflow implementation available at https://github.com/zhixuhao/unet inspired to (Ronneberger, Fischer, and Brox, 2015) that takes as input grayscale images with resolution 512×512.

2. KG Network: The Keypoint Graph Network (since now KG Network) is a neural network based on the concept of Keypoint Graph (Yi et al., 2019). The network first applies a ResNet34-based features extraction. Then the network layers identify some points (called Keypoints) that discretize the input image. Then the collected key points are processed to extract the bounding boxes of the cells. Finally, the bounding boxes of the cells are taken as input for the final layers that extract the cells masks. This network (in contrast with the previous U-Net) provides a cell-by-cell segmentation. However, it has a more considerable forward time than U-Net and a training time of about 4 hours on the same machine with a GPU RTX 2070 with 8 GB of VRAM This contribution has been used a PyTorch implementation publicly available at the site https://github.com/yijingru/KG_Instance_Segmentation based on the paper (Yi et al., 2019) and takes as input grayscale images with resolution 256×256.

3. Mask R-CNN: The Mask R-CNN uses a Region-based Convolutional Neural Network (R-CNN) (He et al., 2017) to extract the masks of the single cells. This network has a more significant forward time than the previous networks, and this network requires a massive quantity of VRAM to be trained; in fact, it has been trained on a cloud node with a GPU NVIDIA K80 with 24 GB of VRAM in 2 hours and 30 minutes. Has been used a Tensorflow/Keras implementation publicly available at the site https://github.com/matterport/Mask_RCNN and takes as input grayscale images with resolution 256×256.

5.4 Deterministic approaches

There are several approaches not using Machine learning in cells segmentation. However, most of them are characterized by a single feature: they work for a combination of parameters depending on the dataset (in worst cases from the single image), but this combination must be found (most of the time using trial and error) by the operator. The proposed model will be tested on two different datasets of white cells on the black background being competitive with neural networks, so it is not affected by this issue. For this reason, the approaches below described will not be tested and will be described to have a historical point of view. However, some of them will be used in the pipeline of the proposed model. Will be made a brief list of the most common methods:

1. Otsu's method: it was introduced in (Otsu, 1979). It is based on the simple idea that if a white object is placed on a dark background, then in the grayscale pixel values histogram (Stockman and Shapiro, 2001) there must be two peaks: the first one is the most common background color end the second one is the most common object color. The method finds the gray scale value that maximizes the variance between the background class (dark pixels) and the object class (bright pixels) and then performs a thresholding on the image at the threshold value found by the method to separate the two classes. The Otsu method is a very powerful tool because it does not require any tuning by the user. However, it has many downsides that make it unpractical for real-world images. For example in real-world images rarely the histogram has a perfectly bimodal distribution (Kittler and Illingworth, 1985) being very noisy. Then Otsu's method is compelling but used alone can lead to poor performances. This is evident in the proposed model and will be solved using neuronal agents.

2. Watershed: Otsu's method returns a binary segmentation. To overcome this problem has been introduced the watershed (Beucher and Mathmatique, 1979). This transform can distinguish homogeneous objects through the gradient of the image. Meyer's version (Meyer, 1992) of this transform starts with a set of markers established by the user (most of the time extracted in an automated way using mathematical morphology). Then, the algorithm performs a "flooding" of the image to find the optimal "basins". It is one of the simplest algorithms for splitting purposes and with good quality images. However, it has some downsides. Starting from an Otsu's thresholding, it could be affected by the same problems previously described. Another downside is that in noisy images, a watershed can be affected by over-splitting (Kornilov and Safonov, 2018), which means that are present more clusters than expected. On the other side for very close and merged objects (like cells clusters), the mathematical morphology methods could fail in separating the single objects. For these reasons, the proposed model has some steps of pre-processing in the markers individuation and post-processing of watershed masks to obtain optimal results.

3. Active Contour Model: the concept of the Active Contour Model (ACM since now) was introduced in 1988 in (Kass, Witkin, and Terzopoulos, 2004). The idea of ACM is to build a discrete contour with key points that minimize through the gradient descent the energy potential

$$E_{snake}^* = \int_0^1 \left(E_{int}(s) + E_{img}(s) + E_{con}(s) \right) ds$$

where E_{int} is called internal energy and takes into account the derivative of the contour, E_{img} takes into account the gray scale values of the image and E_{con} adds further constraints to the contour. The ACM has two main problems. The first problem is that for complex shapes if the weights of the functionals E_{int} and E_{img} are not properly balanced, this could lead to a poor convergence to the actual contour. The algorithm's convergence causes the second kind of problem to local minima. Both issues imply that in real-world applications, the ACMs cannot be applied in a fully automated way (Akbari, Ziaei, and Azarnoush, 2021) because of all these fine tunings (that in worst cases involve every single contour of an image) required to obtain optimal results. For these reasons, the classic functionals will be replaced by neuronal agents, leading to a better generalization without parameter tuning.

5.5   Dataset description

The datasets used in the experiments are two:

1. Neuroblastoma dataset: This dataset introduced in (Kromp et al., 2020) is composed of 4 samples of Tumor and four samples of the bone marrow of Neuroblastoma patients. The dataset has been created with the aid of the Children's Cancer Research Institute (CCRI) biobank (EK.1853/2016) to establish a benchmark for the experimentations on automatic cells segmentation. The dataset consists of 41 train images and 38 test images in format .jpg with a resolution variable of about 1200×1000 of IF cells. The images are already in grayscale, and this means that brighter zones are white and darker zones are black, so they don't need preprocessing. The segmentation has been created by the authors in a manual way (Kromp et al., 2020), and it distinguishes between cells.

The cells segmentation is stored in text-based files. The dataset has proven to be a hard enough benchmark to test the models on real world images (Kromp et al., 2021). For these reasons, it will be taken as the main benchmark of the paper.

2. NucleusSegData dataset: this dataset has been introduced in (Koyuncu, Cetin-Atalay, and Gunduz-Demir, 2018) and used in (Gunesli, Sokmensuer, and Gunduz-Demir, 2020), it is composed of 61 RGB images with a resolution variable of about 1000×700 of cancer cells taken from the Huh7 and HepG2 regions. The cells have been stained with nuclear Hoechst 33258 (Koyuncu, Cetin-Atalay, and Gunduz-Demir, 2018), and it has been activated by U.V. light. This procedure has caused the bright color it is not white, but it is blue because of U.V. light. For this reason, this dataset has been preprocessed to extract the blue channel of the image to obtain a grayscale image representative of fluorescence. Moreover, This dataset will not be used to train the model because it has not a variety of Neuroblastoma datasets. However, it will be used as a test set to examine the generalization capacity of the algorithms.

5.6 Why a new approach is needed

The previous algorithms are all Neural Networks, and they have good performances in optimal conditions. However, it has been shown in the literature that Deep Neural Network (since now DNN) can be affected by adversarial attacks (Goodfellow, Shlens, and Szegedy, 2014). An adversarial attack is a small perturbation (often called adversarial noise) introduced and tuned by a Machine Learning algorithm to induce a misclassification of the Network. Segmentation neural networks are not immune from these kinds of attacks. For example, in (Bar et al., 2021) is proved as an adversative attack is able to fool the ICNet (Zhao et al., 2018) that provides the semantic segmentation controlling an autonomous drive car. The changes of the Adversative algorithm applied to the input image are so subtle that a real world light imperfection or a camera sensor not working properly could be able to recreate the adversative pattern, leading to an accident. In this unlucky (but possible) scenario emerges the second big problem of these algorithms: they are black-box algorithms, and then in many cases after the accident, the best that can be done is to train the Network again and hope that no similar issues will affect the Network in future again. For these reasons, the concept of Explainable AI is becoming more and more popular (since now XAI). In brief, an XAI is an algorithm such as a human can explain its actions and eventually correct them. This topic has resulted in crucial importance for a real-world application of deep learning in fields like robotics, automation, and medicine because an XAI can be fixed after an error, and a human can guarantee its errors.

5.7 The proposed model

In this contribution will be proposed an XAI algorithm based on a mask transformed by neuronal based agents. The algorithm is composed of six main parts:

1. Pre-Processing phase: prepares the images for the next steps applying, first of all, a sharpening filter, then a Gaussian filter and in conclusion, a custom filter to increase the contrast.

2. Watershed Analysis: splits the cells using the well known watershed transformation (Beucher and Mathmatique, 1979).

3. Two steps of splitting and merging: in the first instance, splits every object in the mask computed in the previous steps mask that has an average area bigger than the average, but this process reduces the area of the masks, and for this reason, this new mask is merged to the mask computed in previous steps. This process is repeated twice.

4. Extraction Phase: from the cells masks extracts the contour to run the neuronal model.

5. Neuronal Method: manipulates the contour of the mask using neuronal agents that move on the segment that connects the centre of mass of the cell to their input contour point. In figure 5.1 can be seen a sketch of the agent model. It is composed of 8 neurons (6 Excitatory and 2 Inhibitory) and 3 Layers. The first layer is an input layer and takes into account an attractive signal (figure 5.1, green line) extracted from the mask computed in previous steps, and a repulsive signal (figure 5.1, red line) extracted from the grayscale values of the image. The second layer is inhibitory and applies the repulsive signal. Finally, the third layer sums up all the contributions of the previous layers, moving the agent. The equilibrium between attraction and repulsion approximates the actual contour point.

6. Post-processing phase: Each cell mask is post-processed using an Otsu's segmentation (Otsu, 1979) inside the masks, which has good performances being the content inside the cell mask bimodal. However, some cells are still clustered, and for this reason, has been performed the last splitting and merge cycle. This last splitting differs from the previous from the fact that it is based on the distance transform L2 (Ye, 1988).

The proposed method can be considered an XAI method because each step correspond to a human explainable procedure:

1. The first step is done to remove the noise of the image, then a human can easily check by sight if this has been performed properly.

2. The second step separate the clusters and again a human can easily understand if this has been performed properly and then figure out how to improve performance or fix errors.

3. The third step improve the quality of the subdivision and then is very related to the second step.

4. The fourth step is a simple contour extraction that can be easily understood by a human inspector.

5. The fifth step (the Neuronal method) improve the results of the fourth step and it is also easily explainable (does the neuronal agent stop at the edge? if not there is a problem)

6. For the sixth step hold the same considerations about the second and the third, being a "final cut" segmentation.
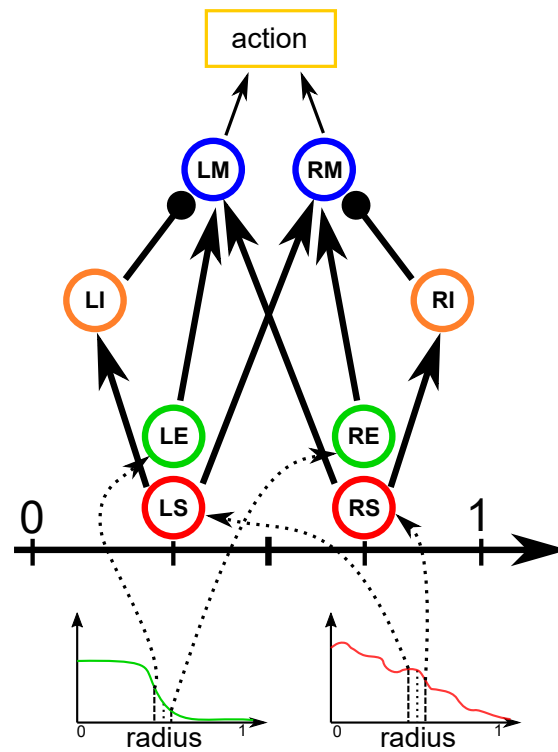
FIGURE 5.1: Sketch of Neuronal agent. Each circle is a neuron and the two bottom curves are the mask attractor (green line) and the grayscale repulsor (red line).

5.8 Analysis of the models

In the next sub sections will be exposed some tests and analysis performed on the algorithms exposed using the datasets introduced.

5.8.1 Adversative noise

The FGSM methodology used in (Goodfellow, Shlens, and Szegedy, 2014) has been performed an adversative attack towards the network U-Net. The FGSM epsilon value has been set respectively to 0, 0.01, 0.025, 0.05, 0.1 and 0.2. In figure 5 are shown some sample images related to Neuroblastoma dataset. The behavior of the U-Net of an image taken from the test set for different quantities of adversative noise. It can be seen that the U-Net, in the absence of noise, performs very well (figure 5.2A first row). However, in the presence of a slight noise (figure 5.2B first row), the U-Net segmentation starts to exhibit large holes in the cells and the phenomena get worse if the noise becomes consistent (figure 5.2C first row). In these cases, the segmentation almost disappears, not showing any cell. One of the most common critiques moved towards the adversative noise is that the ML model that generates it is trained on the network and is strongly focused on the analyzed network. Then, they have been evaluated by the KG network and the Mask R-CNN network the images with the adversative noise computed using the FGSM involving the U-Net. The result is that KG network is very resistant to intermediate noise (figure 5.2B second row) but often returns no segmentation with high noise values (figure 5.2C second row). The Mask R-CNN has shown good performance in every noise condition, even if some

TABLE 5.1: Intersection over Union values of the algorithms for different PSNR values of adversative noise (Neuroblastoma).

IoU	100.0	40.1	32.7	26.9	21.1	15.7
Neuronal Alg.	0.695	0.700	**0.695**	**0.673**	**0.614**	**0.549**
U-Net ResNet34	**0.718**	0.559	0.503	0.406	0.192	0.060
KG network	0.712	**0.704**	0.664	0.593	0.319	0.025
Mask R-CNN	0.682	0.645	0.576	0.489	0.348	0.207

TABLE 5.2: F1-score values of the algorithms for different PSNR values of adversative noise (Neuroblastoma).

F1-score	100.0	40.1	32.7	26.9	21.1	15.7
Neuronal Alg.	0.805	**0.808**	**0.802**	**0.785**	**0.739**	**0.683**
U-Net ResNet34	**0.815**	0.678	0.623	0.523	0.297	0.111
KG network	0.796	0.787	0.754	0.688	0.435	0.047
Mask R-CNN	0.787	0.755	0.695	0.606	0.457	0.287

cells have been lost with strong noise (figure 5.2C third row). However, the best performances have been achieved by the proposed model. Indeed, the segmentation is steady for low and intermediate noise (figure 5.2A and 5.2B fourth row) and exhibits just some holes if the noise is strong (figure 5.2C fourth row).

5.8.2 Results on Neuroblastoma Dataset

Judge the algorithms on a few images is not a good practice, so have been conducted a quantitative investigation on the results of the algorithms with adversative noise with PSNR values of 100.0, 40.1, 32.7, 26.9, 21.1 and 15.7. The algorithms were evaluated on the test set of 38 images; the metrics Intersection on Union, F1 score, precision, Sensitivity and Specificity were evaluated, taking into account their ground truth. Thus, the results are reported in figure 5.3 and summarized in table 5.1-5.5. It can be observed that in terms of IoU for PSNR 100.0 (no noise), the U-Net has the best performance with an IoU of 0.718, then the KG network with an IoU of 0.712, then the proposed Neuronal algorithm with an IoU of 0.708 and finally the Mask R-CNN with an IoU of 0.682. However, adding the adversative noise, the situation changes because the U-Net shows the steepest descent, but these results can be explained by saying that the adversative noise has been created ad hoc for this

TABLE 5.3: Accuracy values of the algorithms for different PSNR values of adversative noise (Neuroblastoma).

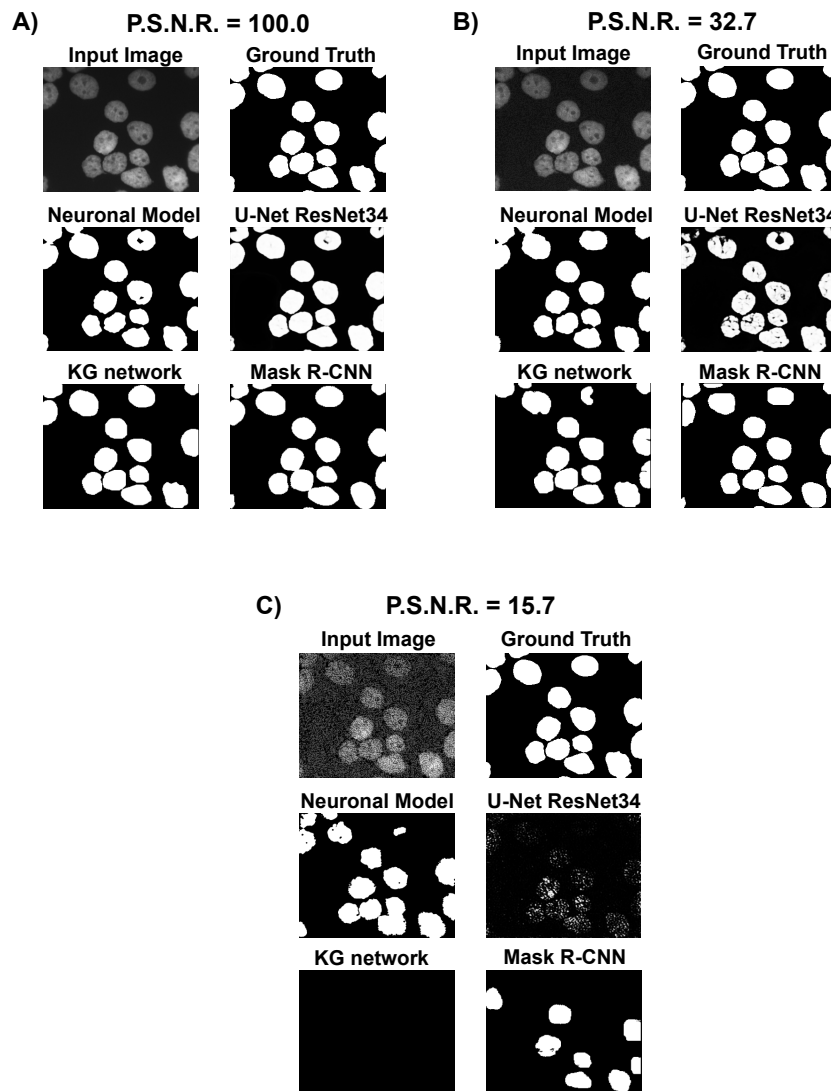| Accuracy	100.0	40.1	32.7	26.9	21.1	15.7
Neuronal Alg.	0.938	0.938	0.936	0.929	**0.910**	**0.879**
U-Net ResNet34	**0.955**	0.926	0.915	0.895	0.845	0.811
KG network	0.953	**0.952**	**0.945**	**0.932**	0.873	0.804
Mask R-CNN	0.943	0.936	0.923	0.908	0.881	0.849

FIGURE 5.2: Comparison of the performances of the algorithms for different intensities of adversative noise. A) Results for an input image without noise. B) Results for an input image with a 32.7 PSNR. C) Results for an input image with a 15.7 PSNR.

TABLE 5.4: Sensitivity values of the algorithms for different PSNR
values of adversative noise (Neuroblastoma).

Sensitivity	100.0	40.1	32.7	26.9	21.1	15.7
Neuronal Alg.	**0.848**	**0.855**	**0.853**	**0.849**	**0.816**	**0.778**
U-Net ResNet34	0.751	0.576	0.518	0.416	0.194	0.061
KG network	0.801	0.779	0.732	0.637	0.331	0.026
Mask R-CNN	0.751	0.699	0.626	0.530	0.377	0.221

TABLE 5.5: Specificity values of the algorithms for different PSNR
values of adversative noise (Neuroblastoma).

Specificity	100.0	40.1	32.7	26.9	21.1	15.7
Neuronal Alg.	0.953	0.954	0.952	0.942	0.926	0.903
U-Net ResNet34	**0.988**	**0.993**	**0.993**	**0.995**	**0.998**	0.999
KG network	0.973	0.976	0.976	0.981	0.991	**0.9996**
Mask R-CNN	0.975	0.979	0.979	0.980	0.981	0.989

network. The KG network seems to keep good performances for a low amount of
adversarial noise, but with PSNR less than 25, the IoU drops to less than 0.32. The
other algorithm outperforms the Mask R-CNN for low noise, but by adding a solid
noise, IoU has a slower decay than the other DNN, even if at the maximum adver-
sarial noise reaches the IoU of 0.21. The proposed model outperforms the DNN in
terms of noise resistance because PNSR 15.7 (the highest noise evaluated) reaches
the minimum IoU value of 0.55, which is more than double all other algorithms
performance. The F1-score (figure 5.3b and table 5.2) has a behavior equivalent to
the IoU, then all the considerations did for the IoU still holds for the F1-score. The
Accuracy (figure 5.3c and table 5.3) and all the other following metrics must be care-
fully analyzed. It has been observed that the DNN algorithm analyzed (U-Net, KG
network and Mask R-CNN) if they produce a false classification are more prone to
produce false negative (FN), instead, the Neural algorithm proposed if produces a
false classification is more prone to produce a false positive (FP). This event trans-
lates into the observation that the DNN usually have more background (negatives)
than the ground truth, and then they are under-segmenting (because cells parts are
cut by the algorithm being classified as background). On the other hand, the pro-
posed model has shown to be over-segmenting because part of the background has
been classified as cells parts. This simple observation has slightly conditioned the
Accuracy because, in under segmenting algorithms, the number TN is more signifi-
cant than in an over-segmenting algorithm. However, the whole graph still confirms
that the proposed model is more resistant to adversative noise. The previous event
becomes clear looking at Sensitivity (figure 5.3d and table 5.4) and Specificity (fig-
ure 5.3e and table 5.5). Indeed, the proposed Neuronal algorithm outperforms the
DNN algorithms in terms of Sensitivity (because in this case, FN is very small); on
the contrary, the DNN algorithms outperform the proposed algorithm in terms of
Specificity (because in these cases, FP is very small). However, the Specificity case is
very curious because it seems that the bigger the noise, the bigger the Specificity. The
answer is hosted in figure 5, in which many DNN predictions with intense noise are
almost Positives-free (black segmentation), in this case, TP=0 or FP=0, which means

TABLE 5.6: Intersection over Union values of the algorithms for different PSNR values of adversative noise (NucleusSegDataset).

IoU	100.0	43.0	34.5	28.7	22.6	16.6
Neuronal Alg.	0.798	**0.799**	**0.801**	**0.794**	**0.740**	**0.510**
U-Net ResNet34	0.778	0.722	0.680	0.622	0.516	0.337
KG network	**0.802**	0.790	0.760	0.659	0.400	0.081
Mask R-CNN	0.712	0.672	0.622	0.510	0.235	0.037

TABLE 5.7: F1-score values of the algorithms for different PSNR values of adversative noise (NucleusSegDataset).

F1-score	100.0	43.0	34.5	28.7	22.6	16.6
Neuronal Alg.	0.881	**0.886**	**0.888**	**0.883**	**0.838**	**0.608**
U-Net ResNet34	0.873	0.834	0.802	0.758	0.669	0.491
KG network	**0.889**	0.882	0.863	0.789	0.550	0.142
Mask R-CNN	0.830	0.801	0.761	0.664	0.357	0.070

that it holds Sensitivity = 1 is the value TN. Then in these cases, the stronger is the noise, the more likely it will be a Positive-free segmentation, and then higher is the Specificity. For these reasons have been preferred more balanced metrics (like IoU and F1-score) in results interpretation.

5.8.3 Results On NucleusSegData Dataset

The principal critique that moved towards the not deep learning approaches is that they must be tuned on the single dataset (in worst cases on the single image) to achieve performances comparable to neural networks. For this reason, has been performed a further test. This test tries to quantify the generalization capability of the previously exposed algorithms. For example, the DNN (U-Net, KG network and R-CNN) have segmented the images of dataset NucleusSegData (Gunesli, Sokmensuer, and Gunduz-Demir, 2020) without fine tuning. Such an effect could appear a limit, but it is the typical pipeline in real-life applications. Indeed, in some cases, the DNN algorithm is connected to a camera that directly streams the image to the algorithm (Bar et al., 2021), and then the algorithm is applied to images that could differ substantially from the training set and test set used by the authors. In other cases, the fine tuning should follow strict rules (AMIA, 2021), and such an approach makes

TABLE 5.8: Accuracy values of the algorithms for different PSNR values of adversative noise (NucleusSegDataset).

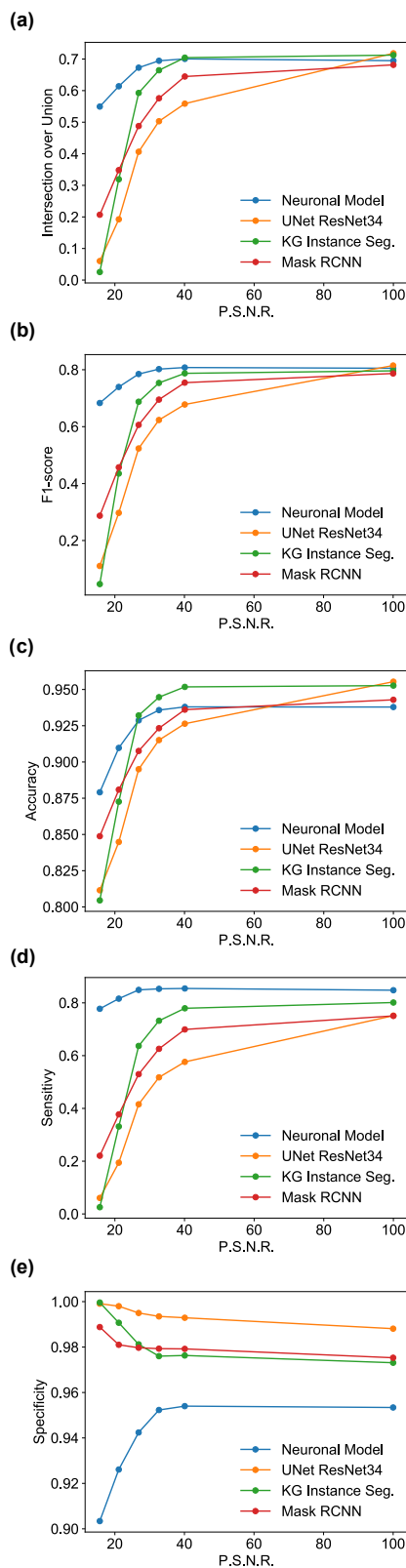| Accuracy	100.0	43.0	34.5	28.7	22.6	16.6
Neuronal Alg.	0.982	**0.982**	**0.982**	**0.982**	**0.966**	0.864
U-Net ResNet34	0.982	0.978	0.975	0.970	0.962	**0.948**
KG network	**0.983**	**0.982**	0.980	0.973	0.953	0.928
Mask R-CNN	0.977	0.973	0.969	0.959	0.938	0.917

FIGURE 5.3: performances in terms of IoU (a), F1-score (b), Accuracy (c), Sensitivity (d) and specificity (e) for each algorithm on Neuroblastoma dataset.
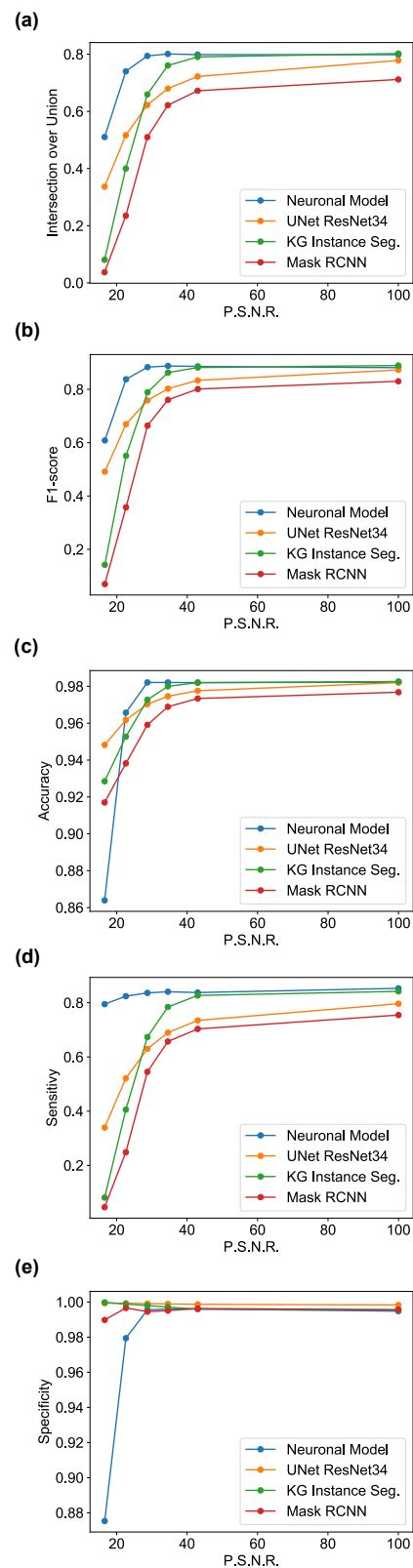
FIGURE 5.4: performances in terms of IoU (a), F1-score (b), Accuracy (c), Sensitivity (d) and specificity (e) for each algorithm on NucleusSegData dataset.

TABLE 5.9: Sensitivity values of the algorithms for different PSNR values of adversative noise (NucleusSegDataset).

Sensitivity	100.0	43.0	34.5	28.7	22.6	16.6
Neuronal Alg.	**0.853**	**0.838**	**0.841**	**0.836**	**0.824**	**0.795**
U-Net ResNet34	0.796	0.735	0.690	0.630	0.521	0.340
KG network	0.842	0.827	0.784	0.637	0.406	0.081
Mask R-CNN	0.755	0.703	0.657	0.545	0.249	0.046

TABLE 5.10: Specificity values of the algorithms for different PSNR values of adversative noise (NucleusSegDataset).

Specificity	100.0	43.0	34.5	28.7	22.6	16.6
Neuronal Alg.	0.995	0.996	0.996	0.996	0.979	0.875
U-Net ResNet34	**0.998**	**0.999**	**0.999**	**0.999**	**0.999**	0.999
KG network	0.996	0.996	0.997	0.998	**0.999**	**1.0**
Mask R-CNN	0.996	0.996	0.995	0.995	0.997	0.990

itself unpractical. Similarly, the Neuronal model does not need tuning parameters on the same dataset (NucleusSegData). In table 6 can be seen the performances of the four models with the dataset NucleusSegData. To this dataset has been added an adversative noise with PSNR 100.0, 43.0, 34.5, 28.7, 22.6 and 16.6. tables 5.6-5.10 and figure 5.4 depict their results. figures 7a and 7b show how the proposed model has outperformed the NNs for high noise values. figure 5.4c depicts that the proposed model has high Accuracy for PSNR values greater than 16.6. However, for PSNR equal to 16.6, the Accuracy value decreases suddenly. This evidence is caused by the super segment of the proposed model for high noise values; instead, the other networks are more prone to under segmentation. The same phenomenon is the cause of the values of Sensitivity and Specificity. It can be seen (as in the previous dataset) that the proposed model has, in general, higher sensitivity and the Neural Networks higher specificity, with the feature that specificity increases when the noise PSNR increases. In general, these simulations have shown better values than the Neuroblastoma case, which can be caused by the similarity of the Neuroblastoma dataset to real-world data (Kromp et al., 2021). In contrast, NucleusSegData (Gunesli, Sokmensuer, and Gunduz-Demir, 2020) is a less complex dataset. Finally, the proposed model has shown generalization capability comparable to the current state-of-the-art Deep Learning algorithms, breaking the dogma that the explainable method implies manual parameter tuning.

Applications to Epidemiology

6.1 Social networks and Covid-19 pandemic

Until now have been analyzed just applications in neuronal network modelling. However, graph theory has a wide spectrum of applications in modelling (Easley and Kleinberg, 2010). One of them is in modelling social networks. A social network is a graph where the people are the nodes, and their relations are the edges. The most common relations used as edges are "the node A is a friend of the node B" (Easley and Kleinberg, 2010) and "the node A has recently met the node B" (Mahdizadeh Gharakhanlou and Hooshangi, 2020). Because the previous relations are symmetric (if "the node A has recently met the node B" then it is also true that "the node B has recently met the node A"). This means that the graph representing the social networks will now be undirected (the adjacency matrix is a symmetric matrix). The concept of social network is fundamental in some branches of Epidemiology because the physical contact of two subjects can transmit some diseases or if they are so close to breathing, the same air (Giordano et al., 2020). The Covid-19 is one of these diseases, in fact it is a respiratory virus that diffuses through the air. The first big Epidemic of Covid-19 was in Italy in February 2020 (Giordano et al., 2020), and since then, the modelling of Covid-19 spread has been performed with several approaches (Bertozzi et al., 2020). The most used approach is through Ordinary Differential Equations (since now ODE) (Shapiro et al., 2021) to describe the Epidemic dynamics at the populations level (just a set of few equations describe the behaviour of many people) (Fernández-Villaverde and Jones, 2020). However, these kinds of models have some downsides. The main problem is that these models have abstract parameters that must be fitted on experimental data. However, this approach does not lead to a complete comprehension of the phenomenon but to a sort of black box predictor that works in virtue of the fitting and not because it uses experimental findings in the model. This issue is not very relevant when the aim is to predict Epidemic behavior for short times. However, this problem becomes crucial if the aim is to take in account alternative scenarios (Giacopelli, 2021). These reasons have introduced models based on simpler and more explainable assumptions: the Agent Based Models (since now ABM) (Son Woo-Sik, 2020). The idea behind the ABM is to simulate the behavior of a single node of a social network (a single person) to simulate the Epidemic spread. The interesting feature of the ABM's is that simulating at a single node level can be used all the estimation done from experimental data without any fitting, and an example is the probability of infection per contact that has been estimated to be 1/40500 (Bhatia and Klausner, 2020). The main downside of ABM is that because of their large number of agents involved, and their more computationally expensive than ODE models. Such a matter is one of the reasons because these kinds of models are very popular for small scale tasks (Cuevas, 2020). However, the increasing of Computer performances of the last years has made possible bigger and

bigger models, like the model with 750 thousand agents reproducing the city of Urmia (Mahdizadeh Gharakhanlou and Hooshangi, 2020) and the ten million agents of the proposed model reproducing Covid-19 spread in Lombardy (Giacopelli, 2021). The following sections will be described the state of the art ODE and ABM models and then describe the ten million inhabitant ABM model (Giacopelli, 2021) with some applications in alternative scenarios evaluation like lockdown, social distancing, and vaccine impact on the Epidemic spread.

6.2   SIR ODE based models

The SIR model is an ODE based model that takes into account the numbers of Susceptible (a person that has not already contracted the disease), Infected (a person who has contracted the disease) and Recovered (a person who has healed from the disease) as time dependent continuous variable ruled by the differential equations (Baldé, 2020)

$$\begin{cases} \frac{dS}{dt} = -\frac{\beta S(t)I(t)}{N} \\ \frac{dI}{dt} = \frac{\beta S(t)I(t)}{N} - \gamma I(t) \\ \frac{dR}{dt} = \gamma I(t) \end{cases} \tag{6.1}$$

Where $\beta > 0$ is the infection factor and $\gamma > 0$ the healing factor. For this system holds the condition

$$N = S(t) + I(t) + R(t)$$

which implies that the number of subjects taken in account is constant. Unfortunately some diseases can be mortal, then the straight forward generalization of SIR model is the SIRD model, where is added a new variable $D(t)$ representing the number of deaths. This model is ruled by the set of differential equations (Baldé, 2020):

$$\begin{cases} \frac{dS}{dt} = -\frac{\beta S(t)I(t)}{N} \\ \frac{dI}{dt} = \frac{\beta S(t)I(t)}{N} - \gamma I(t) - \mu I(t) \\ \frac{dR}{dt} = \gamma I(t) \\ \frac{dD}{dt} = \mu I(t) \end{cases} \tag{6.2}$$

where $\mu > 0$ is the mortality factor. In the case of SIRD model holds the condition

$$N = S(t) + I(t) + R(t) + D(t)$$

that takes in account the number of deaths. It can be observed that the SIRD model (but also the SIR model) has parameters that are meaningful because they can be easily interpreted but are abstract at the same time because they are very far from experimental observations. For this reason, most of the time, the free parameters of SIRD models are fitted on the past values of the variables $S(t)$, $I(t)$, $R(t)$ and $D(t)$ (Baldé, 2020) to predict the future behavior. It must be noted that the SIRD model has shown in different scenarios to be a valid predictor (Fernández-Villaverde and Jones, 2020), however being a sort of abstract black box predictor is not able to evaluate alternative scenarios based on experimental observations. For these reasons, more and more studied models have been based on the single people dynamics. They are called Agent Based Models (ABM) (Giacopelli, 2021).

6.3 ABM models

An Agent Based Model (ABM) is a model that simulates the epidemic spread simulating the behavior of each individual using a digital counterpart called agent. An agent of an ABM has a very simple mathematical description

Definition 6.3.1 *An agent is an object with a position in a 2D space and a state value. In SIRD like ABM's the state of an agent can be one of the following*

1. *Susceptible: the agent has never contracted the disease*

2. *Infected: the agent is ill and then contagious*

3. *Recovered: the agent has healed and then is now immune from the disease*

4. *Deceased: the agent has dead because of the disease and it is not contagious anymore*
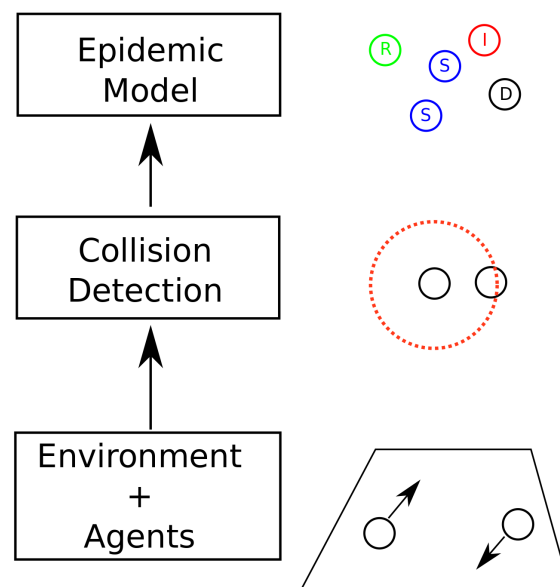

FIGURE 6.1: The 3 layers that constitute the model (Giacopelli, 2021). The first layer simulate the environment and agents' motion. The second layer simulates social interaction between the agents in terms of collision detection. The third layer simulates the virus Epidemic behavior.

These simple agents move in a simulated environment. The realism of the environment (and then of the agent motion) can vary; for example, in (Mahdizadeh Gharakhanlou and Hooshangi, 2020) the environment model is very detailed, taking into account geographic and social data. However, to make a realistic environment, a considerable quantity of data about the population is needed, opening privacy issues for these models. For this reason, in (Giacopelli, 2021) has been proposed a model that uses little and publicly available information about the 10 million inhabitants of Lombardy (for example, the density of inhabitants (*slides Lombardia 2018*) and the Italian citizen average path length (*Osservatorio UnipolSai 2018* 2018)) through the use of random walk (Rayleigh, n.d.). The agents' motion inside the environment is just the first layer of the simulation framework, as shown in figure 6.1.

A collision detection system makes up the second layer of the simulation framework; it is an algorithm that detects if two agents are closer than a given radius. The model presented in (Giacopelli, 2021) has a collision radius of about 1 Km, which is greater than the 1 m suggested by the World Health Organization (*Coronavirus disease (COVID-19) advice for the public* 2020). This change is because the population is large, so the model is very computationally expensive. The agents' behaviour at 6 Frames Per Day has been simulated to solve such an issue. It means that the radius of this model mediates on a time window of 4 hours. The result of this estimate is the radius of 1 Km, which is one of the few fitted parameters of this model. However, checking the distance for every couple of points of a set of $N \approx 10^7$ elements is a problem of complexity $N^2 \approx 10^{14}$ that is a vast number of iterations. For this reason, has been subdivided the environment into 20-Km-side square cells with an average number of nodes m. To compute the distances among the nodes inside the cells, are needed m^2 steps, and the process must be repeated N/m times, so the complexity of this task is not N^2 anymore but

$$\frac{N}{m}m^2 = Nm$$

if m is very smaller than N, it can be said that Nm has the same order of magnitude as N. This approach made it possible to run a 14 days long simulation at 6 Frames Per Day in 20 minutes of computation on a commercial computer with a CPU AMD 3900X 12-cores and 64 GB of RAM (Giacopelli, 2021). It must be noted that this approach ignores some detection when two nodes are close enough, but they are on different cells; nevertheless, numerical investigations have proved that the number of lost contacts is negligible (Giacopelli, 2021).

The third and last layer is the Epidemic model that is a SIRD model (Giordano et al., 2020). This model is based on four states:

1. Susceptible: a agent that has not already been infected by the disease. It can be become infected with a probability p_I for each contact with an infected agent

2. Infected: a agent that has been infected by the disease, it can then infect susceptible agents. After E days, it will become recovered or to deceased, with a probability p_D to die and a probability $1 - p_D$ to heal

3. Recovered: a agent that has recovered from the disease and then it cannot contract it again or infect susceptible agents anymore

4. Deceased: a agent that has died because of the disease and hence cannot infect other agents anymore

The parameters p_I, p_D and E are the probability of infection (Bhatia and Klausner, 2020), the probability of death (*Italy COVID* 2020) and the duration of the disease (*Isolation and Precautions for Adults with COVID-19* 2020) that will be set on the experimental values to reproduce experimental findings and then changed (in accordance to experimental data) to reproduce alternative scenarios.

To validate the ABM, its prediction has been compared to the prediction of a classical SIRD ODE based model (Baldé, 2020) showing equivalent performances in fitting experimental data extracted from (*COVID-19 Situazione Italia* 2020) about the first 8 days of Covid Epidemic tracking in Lombardy.
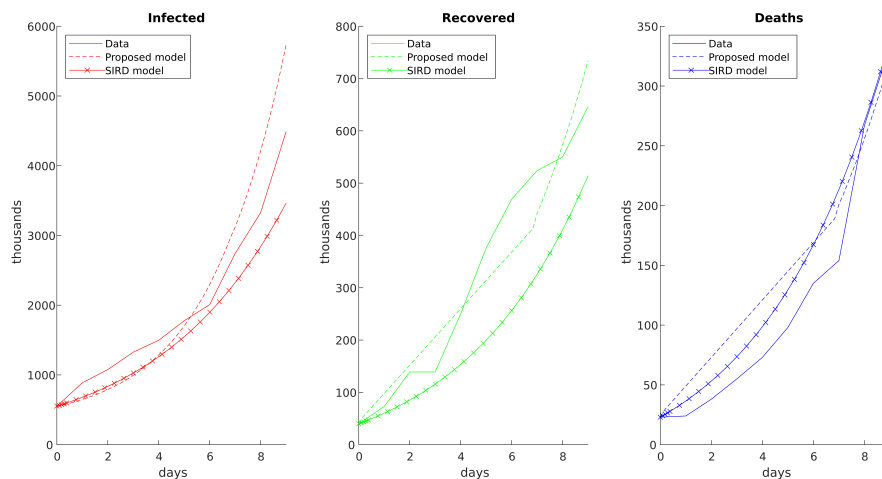
FIGURE 6.2: Comparison in the outbreak scenario between the proposed model (Giacopelli, 2021) and a classical SIRD model (Baldé, 2020).

6.4 The Outbreak scenario

The first scenario analyzed is the outbreak in Lombardy of March 2020 (*COVID-19 Situazione Italia* 2020). The parameters for this scenario are the interaction radius of 1 Km, $p_I = 1/40500$ (Bhatia and Klausner, 2020), $p_D = 0.3$ (estimated from Worldometer data (*Italy COVID* 2020)), $E = 7$ days (that is coherent with the $E < 10$ suggested by (*Isolation and Precautions for Adults with COVID-19* 2020)) and the daily average path length of an agent is about 43 Km (*Osservatorio UnipolSai 2018* 2018). The results can be seen in figure 6.3. Can be seen that the model explains well the behavior of Covid spread until the 9th day of simulation. This 9th day of simulation is the 9th of March, the day of the beginning of the lockdown (*Lockdown decree* 2020). The behavior of the model after this date seems to forecast what could have happened if the lockdown has not begun.

On this setting has also been tested the hypothesis that keeping a distance of 2 m between people can halve the risk of contracting the disease (Chu et al., 2020). For this reason, it has been evaluated the scenario where the probability p_I has been halved, and the results are in figure 6.4. It shows that the contagion spread has been significantly limited in this alternative scenario. A simulation in a lockdown scenario has also been performed, reducing the daily average path length of an agent from 43 km to 5 km and reducing the interaction radius from 1 km to 100 m. The results of this simulation are shown in figure 6.5. It can be seen that these simple actions were enough to control the virus in the simulation.

Has also been performed a simulation in a lockdown scenario, reducing the daily average path length of an agent from 43 km to 5 km and reducing the interaction radius from 1 km to 100 m. The results of this simulation are shown in figure 6.5. Can be seen that these simple actions were enough to control the virus in simulation.

6.5 The Descent and the Vaccine scenario

This scenario has reproduced the situation between May 31, 2020, and June 14, 2020. In that time, Italy concluded its lockdown period, and the number of infected was
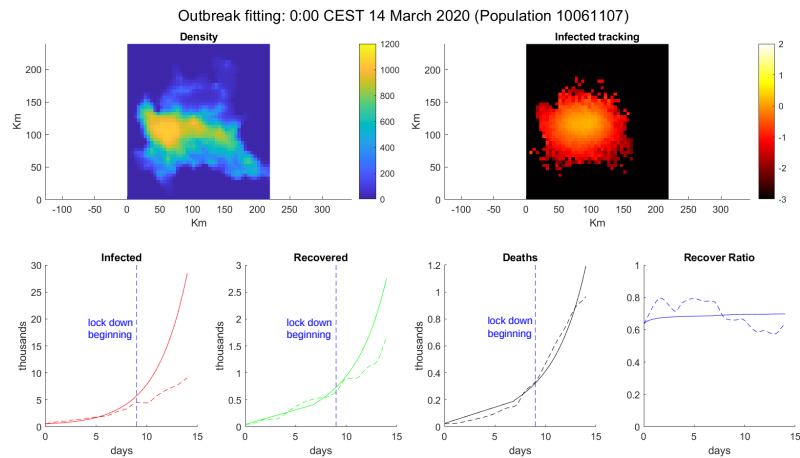
FIGURE 6.3: COVID-19 outbreak simulation. Top-left: population density of Region Lombardy. Top-right: log_{10} value of the infected percentage per cell. Bottom, from left to right: number of infected, recovered, deceased, and recovered ratio (recovered/deaths). The solid line is the model simulation, the dotted line is extracted from the Ministry of Health/Civil Protection Department data (*COVID-19 Situazione Italia* 2020) for Lombardy, and the vertical dotted blue line marks the date March 9, 2020 (*Lockdown decree* 2020).
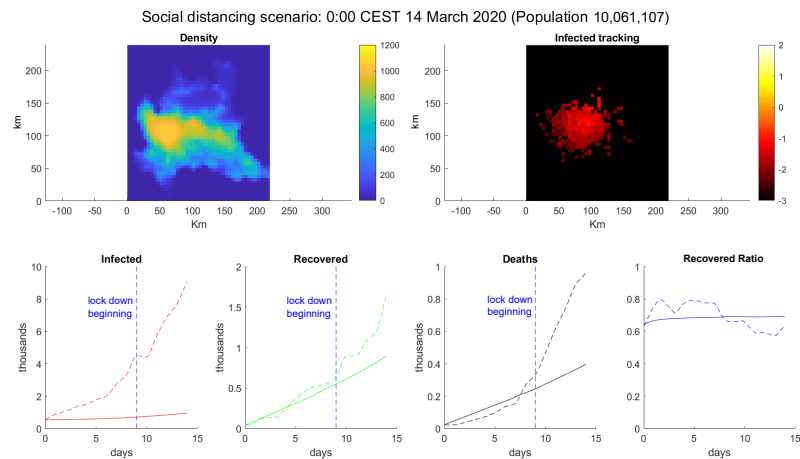

FIGURE 6.4: Social distancing simulation. Top-left: population density of Region Lombardy. Top-right: log_{10} value of the infected percentage per cell. Bottom, from left to right: number of infected, recovered, deceased, and recovered ratio (recovered/deaths). The solid line is the model simulation, the dotted line is extracted from the Ministry of Health/Civil Protection Department data (*COVID-19 Situazione Italia* 2020) for Lombardy, and the vertical dotted blue line marks the date March 9, 2020 (*Lockdown decree* 2020).

FIGURE 6.5: Lockdown simulation. Top-left: population density of Region Lombardy. Top-right: log_{10} value of the infected percentage per cell. Bottom, from left to right: number of infected, recovered, deceased, and recovered ratio (recovered/deaths). The solid line is the model simulation, the dotted line is extracted from the Ministry of Health/Civil Protection Department data (*COVID-19 Situazione Italia 2020*) for Lombardy, and the vertical dotted blue line marks the date March 9, 2020 (*Lockdown decree 2020*).
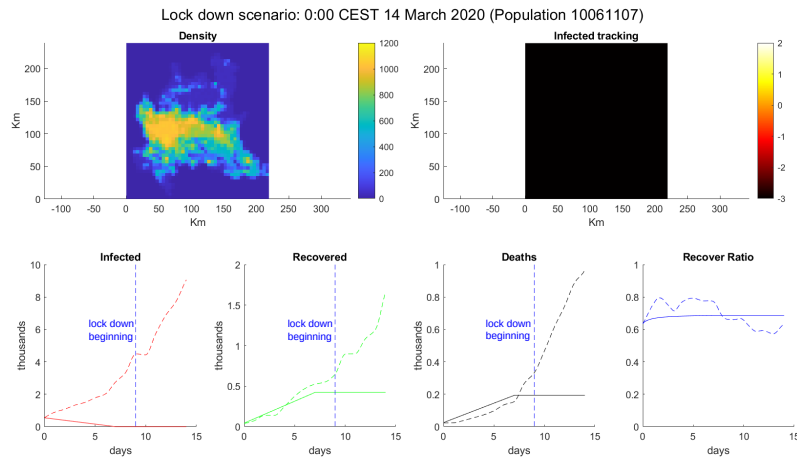
decreasing. The daily average path length has been set to 15 km in this setting because of the lack of additional information in this period's mobility. Furthermore, the probability p_I was halved to consider social distancing. Finally, the interaction radius and the disease duration were tuned to reproduce the experimental data. The value for the radius has been set to 300 m, and the disease duration was 5 weeks ($E = 35$). This last value (which is higher in comparison to that of the outbreak) could be influenced by more and more accurate clinical protocols and the waiting line for the Covid tests created by many patients. It could have slowed down the healing certification process. The qualitative fitting can be seen in figure 6.6.

In this setting has been tested an ideal vaccine experiment where the 70% of population has been set as recovered (then unable to infect their selves and the other agent) to simulate the behavior of a vaccinated agent. The results can be seen in figure 6.7. This paper (Giacopelli, 2021) has been written (and preprinted on MedRxiv) in September 2020, some months before the approval of the vaccines today available. At the time some estimates have set the herd immunity percentage at 62% (Park and Kim, 2020). However, contemporary data suggests that the actual herd immunity percentage should be higher than 70%, for example in December 2021 the percentage of vaccination in Italy is about 73% (*Coronavirus (COVID-19) Vaccinations 2021*) and the number of infected is slightly increasing (*Italy COVID 2020*). This may be caused by two main reasons:

1. The efficacy of the vaccines (Tartof et al., 2021) can range from 80% to about 90% in the first month and can decay in 5 months to 50%.

2. The vaccines could lose efficacy against Covid mutations (Tartof et al., 2021).

The model proposed could take in account this new experimental information by construction, however it is out from the aims of this thesis.
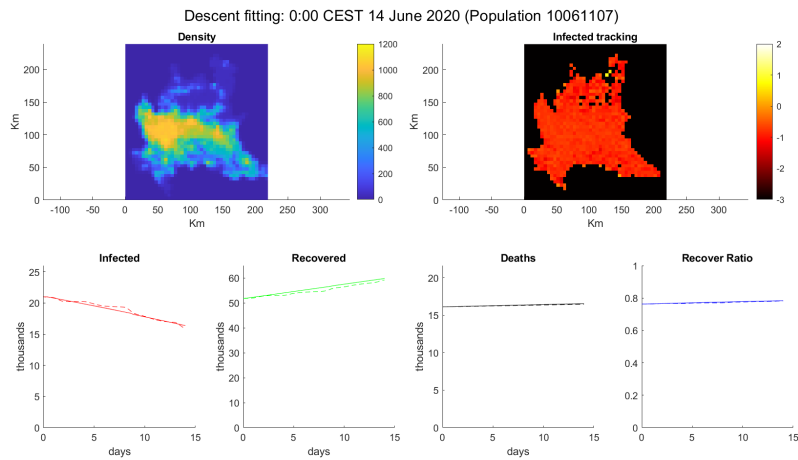
FIGURE 6.6: Simulation of a decline in cases. Top-left: population density. Top-right: log_{10} of the infected percentage per cell. Bottom, from left to right: infected number, recovered number, deceased number, and recovered ratio (recovered/deaths). The solid line is the model simulation and the dotted line is extracted data from Italian Government (*COVID-19 Situazione Italia* 2020).
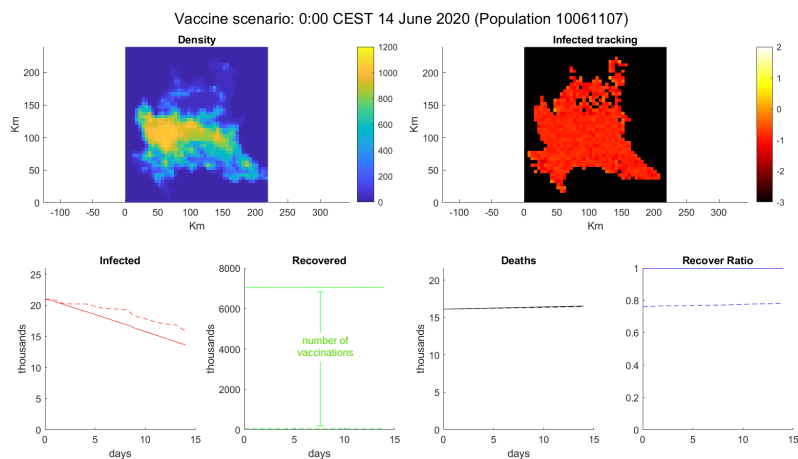

FIGURE 6.7: Simulation of vaccination. Top-left: population density. Top-right: log_{10} of the infected percentage per cell. Bottom, from left to right: infected number, recovered number, deceased number, and recovered ratio.

FIGURE 6.8: COVID-19 outbreak simulation connectivity. Top-left: population density. Top-right: log_{10} of the number of tracked agents per cell. Bottom-left: degree distribution of the tracked group. Bottom-right: daily degree distribution of the tracked group.
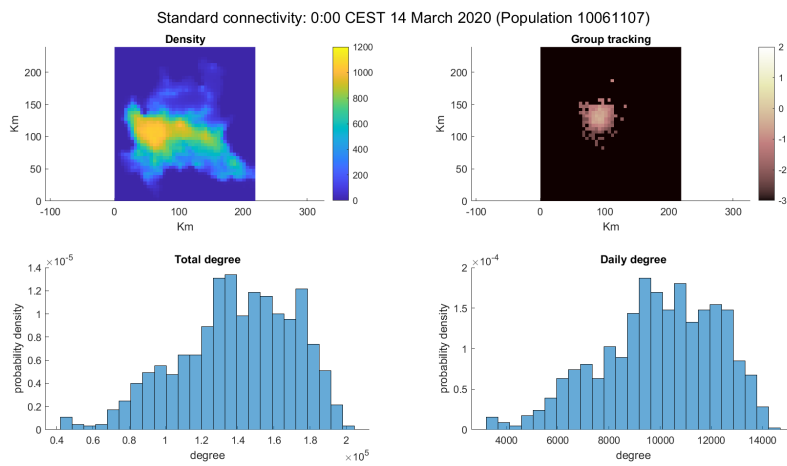

FIGURE 6.9: Lockdown simulation connectivity. Top-left: population density. Top-right: log_{10} of the number of tracked agents per cell. Bottom-left: degree distribution of the tracked group. Bottom-right: daily degree distribution of the tracked group.
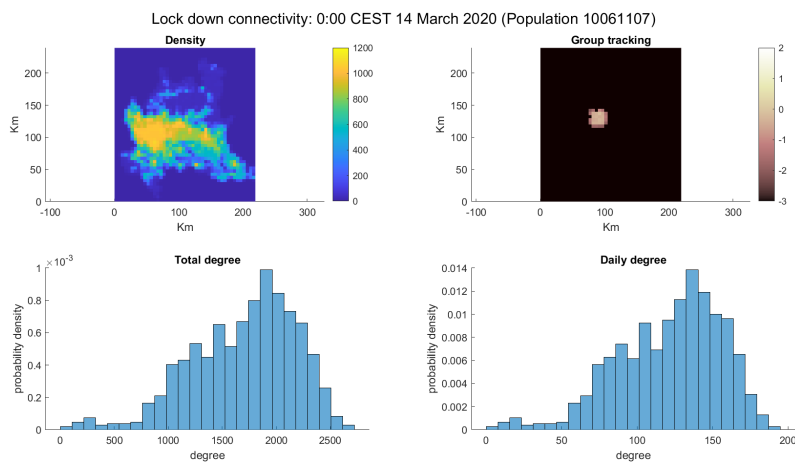
6.6 Network Topology

The impact of topology in social networks is an argument of fierce debate in the community (Manzo, 2020). Then has been evaluated the connectivity in terms of the degree distribution of 1000 agents randomly picked. The first scenario was the COVID-19 outbreak scenario in figure 6.8. It can be seen that the distribution has an evident left tail (in contrast with the right tail of the Barabási-Albert models (Barabási and Albert, 1999)). This was probably due to the very short simulation time of 14 days (in contrast with human social networks that usually take years to build). The lockdown scenario was also interesting. In figure 6.9 can be observed a substantial decline in terms of a degree from thousands of average connections per day to hundreds. It shows the importance of lockdowns in COVID-19 containment.

Conclusions

One of the leitmotifs of the thesis is to explore and characterize the micro-connectivity of Brain networks (as explained in Chapter 1). The first step in this direction was the *Convolutive hypothesis* (Giacopelli, Migliore, and Tegolo, 2020; Giacopelli et al., 2021; Giacopelli, Tegolo, and Migliore, 2021). The idea underlying this hypothesis argues that the most likely trend for the indegree and outdegree distrubtions, that are the probability distributions associated to the indegrees and outdegrees of the network, of a microconnectivity brain network seemed to have a Convolutive behavior, but no model satisfied such an observation (as observed in Chapter 2). Indeed, the specialized literature introduced two main theoretical models adopted by the scientific community. The exponential models (Erdös and Rényi, 1959) that has a growing behavior for low degrees, but they decay in an exponential way for high degrees, and the Power law models (Barabási and Albert, 1999) that have a power law tail for high degrees but they have high probability densities in the degree distributions for low degrees. The concept of Convolutive behavior tries to reproduce the considerations extracted by experimental data and data-driven models, which suggest a behavior for the degree distributions growing for low degrees and a heavy tail behavior for high degrees. A theoretical model able to satisfy this observation has been exposed for the first time in (Giacopelli, Migliore, and Tegolo, 2020) observing Neocortical network of (Markram et al., 2015), Hyppocampal slices of (Bonifazi et al., 2009) and C. Elegans connectome (Cook et al., 2019), but only in (Giacopelli et al., 2021) has become more than a simple hypothesis. Moreover, in (Giacopelli et al., 2021) and in Chapter 3 have been analyzed the degree distributions of the following eight different realistic micro-connectivity networks and can be observed the recurrent scheme of Convolutive degree distributions:

1. The 546 neurons of the C. elegans adult male nervous system (Cook et al., 2019)
2. A dense reconstruction of 1090 neurons from a mouse retina inner plexiform layer (Helmstaedter et al., 2013)
3. From electron microscopy data on a region of L2/3 mouse primary visual corte (Turner et al., 2020)
4. From electron microscopy data on 226 neurons from a Songbird basal ganglia (Dorkenwald et al., 2017)
5. 89 neurons from a slice from a rodent hippocampus (Bonifazi et al., 2009)
6. A mouse thalamus network of 963 neurons (Morgan and Lichtman, 2020)
7. From electron microscopy data containing 1761 body ID's from a Drosophila Optic Medulla (Takemura et al., 2013)
8. A neocortical column of rodent (Markram et al., 2015)

Such an analysis suggested that (to the best of our knowledge) brain at micro-scale is usually connected with a Convolutive model scheme (Giacopelli, Migliore, and Tegolo, 2020; Giacopelli et al., 2021; Giacopelli, Tegolo, and Migliore, 2021) and does not follow an exponential scheme as previously supposed by (Potjans and Diesmann, 2012; Billeh et al., 2020; Brunel, 1999). However, it must be noted that most

of the papers supposing an exponential connectivity (Potjans and Diesmann, 2012; Brunel, 1999) have been written in a period where most of the networks previously cited were not yet experimentally available.

In order to study and validate the model in the simulations field for brain networks on a cellular scale, in Chapter 4 several tests were performed on the Convolutional model. The first test was proposed in (Giacopelli, Migliore, and Tegolo, 2020), where a simple Excitatory-Inhibitory model (Brunel, 1999), with convolutive degree distributions, was tested on a classic ER (Brunel, 1999), a mixed ER (Potjans and Diesmann, 2012) is a data driven (Markram et al., 2015) model. The result was that the exponential models shown a not physiological activity, in contrast with the data driven model, that has shown many expected physiological markers experimentally observed (Buzsáki and Draguhn, 2004). Then, the Convolutive Model (Giacopelli et al., 2021) has been tested to fitting the degree distributions of the data driven model (Markram et al., 2015), showing an activity very close to the data driven model. This result suggests that matching at least the degree distributions of a network implies matching the brain activity qualitatively. A similar result was stated by (Nykamp et al., 2017) in terms of mean field, but in (Giacopelli et al., 2021) has been proved to be valid in a more realistic setting. A second test has been performed in (Giacopelli, Tegolo, and Migliore, 2021), where in two study cases (Thick volume and Thin volume) has been tested three different network topologies (ER classic (Erdös and Rényi, 1959), ER spatial (Billeh et al., 2020) and Convolutive model (Giacopelli, Tegolo, and Migliore, 2021)). The neuron model in (Giacopelli, Tegolo, and Migliore, 2021) is the Hindmarsh-Rose model (Hindmarsh and Rose, 1984), that is the starting model of Epileptor (Jirsa and McIntosh, 2007). In (Giacopelli, Tegolo, and Migliore, 2021) is shown how exponential connectivity schemes are more prone to Epileptic seizures, in contrast with Convolutive connectivity. In detail have been performed two primary tests. In the first test has been simulated a population of HR neurons very similar to the one theorized by Stefanescu in (Stefanescu and Jirsa, 2008). In (Stefanescu and Jirsa, 2008) is supposed an exponential connectivity scheme with an HR starting model; on this assumption, every neuron is influenced by the other neurons, and also given the HR starting model, a mean field approximation is computed (commonly known as Epileptor (Stefanescu and Jirsa, 2008)). The exciting part is that wiring the HR networks of (Giacopelli, Tegolo, and Migliore, 2021) with an exponential scheme, the mean field of the network is very close to the Epileptor prediction, and then to an actual Epileptic EEG (Noachtar and Rémi, 2009). However, wiring the network with a Convolutive connectivity scheme, the mean amplitude decreases significantly. In detail, when exposed to an external current, all ER model networks (with or without distance dependent connectivity) start to have an epileptic behavior, with an average value of mean amplitude in a 10 seconds time window of 46.6, 34.6, 47.3 and 37.1 μV, for the ER and ER-dist in thick volume and for the ER and ER-dist in thin volume, respectively. In contrast, networks connected as in real brain networks, show an average value o the mean amplitude in the same time window of 9.9 and 7 μV, for thick and thin volumes, respectively, very close to the value obtained from normal traces. This suggests suggesting that exponential connectivity could not properly model a healthy individual's brain activity. Therefore, in the second test has been introduced a further equation that links the inhibition to mean field potential. In particular, the higher is the mean field higher is the inhibition. This simple equation was able to reproduce the complex behavior of spike wave complexes (Noachtar and Rémi, 2009). This second test suggests that the phenomenon of Spike wave complexes could be more connected to the rapidity increasing of inhibition

than to its actual strength. Also in this case spike-wave complexes were clearly evident in the ER models, following a period of a seizure-like pattern. In this case the average mean amplitude values during the entire stimulation period (in a time window of about 14 seconds), were 32.2 and 20.9 μV. Spike-wave complexes were also present for the convolutive model but they were of much smaller amplitude, and the overall average activity during the same stimulation period they keep their average value within the range for a normal activity (9.4 μV). The encouragement of this study can be found in the analysis of current anti-epileptic drugs that are based on reduction of inhibitory response (Rogawski and Löscher, 2004), which may lead to several adverse collateral effects (St Louis, 2009). Modern Graph Theory has shown to be able to achieve good results in Connectome modeling and interpretation, with the creation of new theories and techniques. But in Chapter 5 has been also shown how a simple neuronal network based agent has been used to successfully solve segmentation tasks in an explainable way and in Chapter 6 how a spatial connectivity pattern combined with an Agent Based Model was able to simulate Covid-19 outbreak from publicly available data (Giacopelli, 2021). In the first study case has been proposed a model based on a neuronal agent that is able to segment Immunofluorescence images. This contribution seems promising because this algorithm is very resilient adversarial noise, in contrast with Deep learning algorithms. In detail, the algorithms in absence of adversative noise are very close in terms of Intersection over Union (IoU), achieving the values 0.695 for the proposed model, 0.718 for the U-Net (Ronneberger, Fischer, and Brox, 2015), 0.712 for the KG Network (Yi et al., 2019) and 0.682 for the Mask R-CNN (He et al., 2017), but in presence of adversative noise the IoU keeps a reasonable value of 0.549 for the proposed model and instead drops significantly for the Neural Networks with values 0.060 for the U-Net, 0.025 for the KG Network and 0.207 for the Mask R-CNN. This feature combined with its Explainability of the algorithm makes it very suitable for contexts where to each prediction is associated a concrete responsibility, like in medical applications (AMIA, 2021). In second study case has been proposed an Agent Based Model of 10 millions agents (Giacopelli, 2021). The proposed model was validated against a classical SIRD (Susceptible-Infected-Recovered-Deceased) model (Baldé, 2020) fitted with a parameter exploration scheme on outbreak data showing comparable results in terms of the rooted mean square error of the data: the SIRD model had an error of 150 for the infected, 71 for the recovered, and 18 for the deceased; and the proposed model exhibited an error of 535 for the infected, 58 for the recovered, and 34 for the deceased. However, most of the parameters of the proposed models are realistic and empirically measurable, making the propsed model very suitable for evaluating alternative scenarios. One of them is the efficacy of Vaccines in fighting Covid-19 pandemic, where the model has predicted an herd immunity percentage of 70% in September 2020, when a vaccine was not available yet and when so many mutations of the virus were not present yet (Tartof et al., 2021).

In conclusion Graph Theory is becoming more and more Modern, being contaminated by the wide spectrum of its applications (Barrat, Barthélemy, and Vespignani, 2008; Easley and Kleinberg, 2010). However, is opinion of the author that this contamination of Classic Graph Theory is a good thing, because every new solution in a new field changes and improves the tools of Graph Theory. This contamination is just the demonstration of the unbreakable link between Theory and its applications.

List of Contributions

Journals

1. G. Giacopelli, M. Migliore and D. Tegolo . "A Segmentation Method for Immunofluorescence images: an approach without machine learning influence". *Submitted to: IEEE Access*.

2. G. Giacopelli, D. Tegolo and M. Migliore (2021). "The role of network connectivity on epileptiform activity", *Scientific Reports*, DOI: 10.1038/s41598-021-00283-w.

3. S. Coppolino, G. Giacopelli and M. Migliore (2021). "Sequence Learning in a Single Trial: A Spiking Neurons Model Based on Hippocampal Circuitry", *IEEE transactions on neural networks and learning systems*, DOI: 10.1109/TNNLS.2021.3049281

4. G. Giacopelli, D. Tegolo, E. Spera and M. Migliore (2021). "On the structural connectivity of large-scale models of brain networks at cellular level", *Scientific reports*, DOI: 10.1038/s41598-021-83759-z.

5. G. Giacopelli, M. Migliore and D. Tegolo (2020). "Graph-theoretical derivation of brain structural connectivity", *Applied Mathematics and Computation* 377, DOI: 10.1016/j.amc.2020.125150.

Conferences

1. G. Giacopelli, M. Migliore and D. Tegolo (2020). "Spatial graphs and Convolutive Models". In: *2020 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)*, CIBCB 2020, pages 1–7, 27–29 October 2020, Vina del Mar, Chile, DOI: 10.1109/CIBCB48159.2020.9277722.

2. G. Giacopelli (2020). "The introduction of an Oracle in a simple finance inspired dynamic system leads to instability". In: *PervasiveHealth: Pervasive Computing Technologies for Healthcare*, , ACM proceedings of ComSysTech'20, pages 156–161, 19–20 June 2020, Ruse, Bulgaria, DOI: 10.1145/3407982.3408017.

Preprints

1. G. Giacopelli (2021). "A Full-Scale Agent-Based Model to Hypothetically Explore the Impact of Lockdown, Social Distancing, and Vaccination During the COVID-19 Pandemic in Lombardy, Italy: Model Development". In: *JMIRx Med* 2.3, e24630. ISSN: 2563-6316. DOI: 10.2196/24630.

2. G. Giacopelli, (2019). "Studying Topology of Time Lines Graph leads to an alternative approach to the Newcomb's Paradox". In: *ArXiv* abs/1910.09311.

Bibliography

Akbari, Parastoo, Atefeh Ziaei, and Hamed Azarnoush (2021). "Deep Active Contours Using Locally Controlled Distance Vector Flow". In: *ArXiv* abs/2105.08447.

Alzubaidi, Laith et al. (2021). "Review of deep learning: concepts, CNN architectures, challenges, applications, future directions". In: *Journal of Big Data* 8.

AMIA, American Medical Informatics Association (2021). *"AMIA supports, encourages further refinement of FDA AI/Machine Learning Regulatory Framework" available at https://amia.org/news-publications/amia-supports-encourages-further-refinement-fda-aimachine-learning-regulatory*.

Axler, Sheldon Jay (1997). *Linear Algebra Done Right*. Undergraduate Texts in Mathematics. New York: Springer. ISBN: 0387982582. URL: http://linear.axler.net/.

Baldé, Mouhamadou A.M.T. (2020). "Fitting SIR model to COVID-19 pandemic data and comparative forecasting with machine learning". In: *medRxiv*. DOI: 10.1101/2020.04.26.20081042.

Bar, Andreas et al. (2021). "The Vulnerability of Semantic Segmentation Networks to Adversarial Attacks in Autonomous Driving: Enhancing Extensive Environment Sensing". In: *IEEE Signal Processing Magazine* 38.1, pp. 42–52. DOI: 10.1109/MSP.2020.2983666.

Barabási, Albert-László and Réka Albert (1999). "Emergence of Scaling in Random Networks". In: *Science* 286.5439, pp. 509–512. ISSN: 0036-8075. DOI: 10.1126/science.286.5439.509.

Barrat, Alain, Marc Barthélemy, and Alessandro Vespignani (2008). *Dynamical Processes on Complex Networks*. Cambridge University Press. DOI: 10.1017/CBO9780511791383.

Bertozzi, Andrea L. et al. (2020). "The challenges of modeling and forecasting the spread of COVID-19". In: *Proceedings of the National Academy of Sciences* 117.29, pp. 16732–16738. ISSN: 0027-8424. DOI: 10.1073/pnas.2006520117.

Beucher, Serge and Centre Mathmatique (1979). "The Watershed Transformation Applied To Image Segmentation". In: *Scanning. Microsc.* 6.

Bhatia, Rajiv and Jeffrey Klausner (2020). "Estimation of Individual Probabilities of COVID-19 Infection, Hospitalization, and Death From A County-level Contact of Unknown infection Status". In: *medRxiv*. DOI: 10.1101/2020.06.06.20124446.

Billeh, Yazan N. et al. (2020). "Systematic Integration of Structural and Functional Data into Multi-scale Models of Mouse Primary Visual Cortex". In: *Neuron* 106.3, 388–403.e18. ISSN: 0896-6273. DOI: https://doi.org/10.1016/j.neuron.2020.01.040.

Bonifazi, P. et al. (2009). "GABAergic Hub Neurons Orchestrate Synchrony in Developing Hippocampal Networks". In: *Science* 326.5958, pp. 1419–1424. ISSN: 0036-8075. DOI: 10.1126/science.1175509.

Bromiley, P (Jan. 2003). "Products and Convolutions of Gaussian Distributions". In.

Brunel, Nicolas (1999). *Dynamics of Sparsely Connected Networks of Excitatory and Inhibitory Spiking Neurons*.

Bullmore, Ed and Olaf Sporns (2009). "Complex brain networks: graph theoretical analysis of structural and functional systems". In: *Nature Reviews Neuroscience* 10.3, pp. 186–198. ISSN: 1471-003X. DOI: `10.1038/nrn2575`.

Buzsáki, Gyorgy and Andreas Draguhn (July 2004). "Neuronal Oscillations in Cortical Networks". In: *Science (New York, N.Y.)* 304, pp. 1926–9. DOI: `10.1126/science.1099745`.

Chu, Derek K. et al. (2020). "Physical distancing, face masks, and eye protection to prevent person-to-person transmission of SARS-CoV-2 and COVID-19: a systematic review and meta-analysis". In: *Lancet (London, England)* 395, pp. 1973 –1987.

Cook, Steven J. et al. (July 2019). "Whole-animal connectomes of both Caenorhabditis elegans sexes". English (US). In: *Nature* 571.7763, pp. 63–71. ISSN: 0028-0836. DOI: `10.1038/s41586-019-1352-7`.

Coronavirus (COVID-19) Vaccinations (2021). Our World in Data. URL: `https://ourworldindata.org/covid-vaccinations?country=ITA`.

Coronavirus disease (COVID-19) advice for the public (2020). World Health Organization. URL: `https://www.who.int/emergencies/diseases/novel-coronavirus-2019/advice-for-public`.

COVID-19 Situazione Italia (2020). Ministero della Salute, Dipartimento della Protezione Civile. URL: `https://opendatadpc.maps.arcgis.com/apps/dashboards/b0c68bce2cce478eaac82fe38d4138b1`.

Cuevas, Erik (2020). "An agent-based model to evaluate the COVID-19 transmission risks in facilities". In: *Computers in biology and medicine* 121, p. 103827. ISSN: 0010-4825. DOI: `10.1016/j.compbiomed.2020.103827`.

Deng, Li and Dong Yu (2014). "Deep Learning: Methods and Applications." In: *Found. Trends Signal Process.* 7.3-4, pp. 197–387.

Destexhe, Alain, Diego Contreras, and Mircea Steriade (2001). "LTS cells in cerebral cortex and their role in generating spike-and-wave oscillations". In: *Neurocomputing* 38-40. Computational Neuroscience: Trends in Research 2001, pp. 555–563. ISSN: 0925-2312. DOI: `https://doi.org/10.1016/S0925-2312(01)00348-4`.

Devinsky, Orrin et al. (2018). "Epilepsy". In: *Nature Reviews Disease Primers* 4.1, p. 18024. ISSN: 2056-676X. DOI: `10.1038/nrdp.2018.24`.

Dorkenwald, Sven et al. (2017). "Automated synaptic connectivity inference for volume electron microscopy". In: *Nature Methods* 14, pp. 435–442.

Dorogovtsev, S.N. and J.F.F. Mendes (2003). *Evolution of Networks: From Biological Nets to the Internet and WWW*. Oxford: Oxford University Press.

Dummer, Benjamin, Stefan Wieland, and Benjamin Lindner (2014). "Self-consistent determination of the spike-train power spectrum in a neural network with sparse connectivity". In: *Frontiers in Computational Neuroscience* 8, p. 104. ISSN: 1662-5188. DOI: `10.3389/fncom.2014.00104`.

Dutta, Sangya et al. (Dec. 2017). "Leaky Integrate and Fire Neuron by Charge-Discharge Dynamics in Floating-Body MOSFET". In: *Scientific Reports* 7. DOI: `10.1038/s41598-017-07418-y`.

Easley, David A. and Jon M. Kleinberg (2010). *Networks, Crowds, and Markets - Reasoning About a Highly Connected World*. Cambridge University Press. ISBN: 978-0-521-19533-1.

Erdös, P. and A. Rényi (1959). "On Random Graphs I". In: *Publicationes Mathematicae Debrecen* 6, p. 290.

Euler, Leonhard (1736). "Solutio problematis ad geometriam situs pertinentis". In: *Commentarii Academiae Scientiarum Imperialis Petropolitanae* 8, pp. 128–140.

Fabrikant, Alex, Elias Koutsoupias, and Christos H. Papadimitriou (2002). "Heuristically Optimized Trade-Offs: A New Paradigm for Power Laws in the Internet". In: *Proceedings of the 29th International Colloquium on Automata, Languages and Programming*. ICALP '02. Berlin, Heidelberg: Springer-Verlag, 110–122. ISBN: 3540438645.

Farahani, Farzad V., Waldemar Karwowski, and Nichole R. Lighthall (2019). "Application of Graph Theory for Identifying Connectivity Patterns in Human Brain Networks: A Systematic Review". In: *Frontiers in Neuroscience* 13, p. 585. ISSN: 1662-453X. DOI: `10.3389/fnins.2019.00585`.

Feller, W. (1945). "On the Normal Approximation to the Binomial Distribution". In: *The Annals of Mathematical Statistics* 16.4, pp. 319 –329. DOI: `10.1214/aoms/1177731058`.

Fernández-Villaverde, Jesús and Charles I Jones (2020). *Estimating and Simulating a SIRD Model of COVID-19 for Many Countries, States, and Cities*. Working Paper 27128. National Bureau of Economic Research. DOI: `10.3386/w27128`.

Finotelli, Paolo (July 2015). "Graph Theoretical Analysis of the Brain. An Overview". In: *Scienze e Ricerche*.

Gal, Eyal et al. (2017). "Rich cell-type-specific network topology in neocortical microcircuitry". In: *Nature neuroscience* 20.7, 1004—1013. ISSN: 1097-6256. DOI: `10.1038/nn.4576`.

Gholmieh, Ghassan I., L.S. Chen, and Spiros H. Courellis (2007). "Screening for the Effects of Antiepileptic Drugs on Short Term Plasticity Using A Time Efficient Bioassay". In: *2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 2247–2252.

Giacopelli, G., M. Migliore, and D. Tegolo (2020). "Graph-theoretical derivation of brain structural connectivity". In: *Applied Mathematics and Computation* 377, p. 125150. ISSN: 0096-3003. DOI: `https://doi.org/10.1016/j.amc.2020.125150`.

Giacopelli, G., M. Migliore, and D. Tegolo (2020). "Spatial graphs and Convolutive Models". In: *2020 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)*, pp. 1–7. DOI: `10.1109/CIBCB48159.2020.9277722`.

Giacopelli, Giuseppe (2021). "A Full-Scale Agent-Based Model to Hypothetically Explore the Impact of Lockdown, Social Distancing, and Vaccination During the COVID-19 Pandemic in Lombardy, Italy: Model Development". In: *JMIRx Med* 2.3, e24630. ISSN: 2563-6316. DOI: `10.2196/24630`.

Giacopelli, Giuseppe, Domenico Tegolo, and Michele Migliore (2021). "The role of network connectivity on epileptiform activity". In: *Scientific Reports* 11.

Giacopelli, Giuseppe et al. (2021). "On the structural connectivity of large-scale models of brain networks at cellular level". eng. In: *Scientific reports* 11.1, pp. 4345–4345. ISSN: 2045-2322.

Giordano, Giulia et al. (June 2020). "Modelling the COVID-19 epidemic and implementation of population-wide interventions in Italy". In: *Nature Medicine* 26, pp. 1–6. DOI: `10.1038/s41591-020-0883-7`.

Goodfellow, Ian, Jonathon Shlens, and Christian Szegedy (Dec. 2014). "Explaining and Harnessing Adversarial Examples". In: *arXiv 1412.6572*.

Gunesli, Gozde Nur, Cenk Sokmensuer, and Cigdem Gunduz-Demir (2020). "AttentionBoost: Learning What to Attend for Gland Segmentation in Histopathological Images by Boosting Fully Convolutional Networks". In: *IEEE Transactions on Medical Imaging* 39, pp. 4262–4273.

He, Kaiming et al. (2017). "Mask R-CNN". In: *2017 IEEE International Conference on Computer Vision (ICCV)*, pp. 2980–2988.

Helmstaedter, Moritz et al. (2013). "Connectomic reconstruction of the inner plexiform layer in the mouse retina". In: *Nature* 500.7461, 168—174. ISSN: 0028-0836. DOI: 10.1038/nature12346.

Hindmarsh, James and R Rose (Apr. 1984). "A Model of Neuronal Bursting Using Three Coupled First Order Differential Equations". In: *Proceedings of the Royal Society of London. Series B, Containing papers of a Biological character. Royal Society (Great Britain)* 221, pp. 87–102. DOI: 10.1098/rspb.1984.0024.

Hjorth, J. J. Johannes et al. (2020). "The microcircuits of striatum in silico". In: *Proceedings of the National Academy of Sciences* 117.17, pp. 9554–9565. ISSN: 0027-8424. DOI: 10.1073/pnas.2000671117.

Hodgkin, Alan Lloyd, Andrew Fielding Huxley, and John Carew Eccles (1952). "Propagation of electrical signals along giant nerve fibres". In: *Proceedings of the Royal Society of London. Series B - Biological Sciences* 140.899, pp. 177–183. DOI: 10.1098/rspb.1952.0054.

How to Play Mega Millions Lottery (2021). Mega Millions Lottery. URL: https://www.megamillions.com/how-to-play.

Im, Kyuseok et al. (Jan. 2019). "An Introduction to Performing Immunofluorescence Staining: Methods and Protocols". In: vol. 1897, pp. 299–311. ISBN: 978-1-4939-8933-1. DOI: 10.1007/978-1-4939-8935-5_26.

Isolation and Precautions for Adults with COVID-19 (2020). Centers for Disease Control and Prevention. URL: https://www.cdc.gov/coronavirus/2019-ncov/hcp/duration-isolation.html.

Italy COVID (2020). Worldometer. URL: https://www.worldometers.info/coronavirus/country/italy/\#graph-deaths-daily.

Jirsa, V. K. and A. R. McIntosh (2007). *Handbook of brain connectivity*. Springer.

Kass, Michael, Andrew P. Witkin, and Demetri Terzopoulos (2004). "Snakes: Active contour models". In: *International Journal of Computer Vision* 1, pp. 321–331.

Khan, Asifullah et al. (2020). "A survey of the recent architectures of deep convolutional neural networks". In: *Artificial Intelligence Review*, pp. 1 –62.

Kittler, J. and J. Illingworth (1985). "On threshold selection using clustering criteria". In: *IEEE Transactions on Systems, Man, and Cybernetics* SMC-15.5, pp. 652–655. DOI: 10.1109/TSMC.1985.6313443.

Kornilov, Anton S. and Ilia Vladimirovich Safonov (2018). "An Overview of Watershed Algorithm Implementations in Open Source Libraries". In: *J. Imaging* 4, p. 123.

Koyuncu, Can, Rengul Cetin-Atalay, and Cigdem Gunduz-Demir (Sept. 2018). "Object-Oriented Segmentation of Cell Nuclei in Fluorescence Microscopy Images". In: *Cytometry Part A* 93. DOI: 10.1002/cyto.a.23594.

Kromp, Florian et al. (2020). "An annotated fluorescence image dataset for training nuclear segmentation methods". In: *Scientific Data* 7.

Kromp, Florian et al. (Mar. 2021). "Evaluation of Deep Learning Architectures for Complex Immunofluorescence Nuclear Image Segmentation". In: *IEEE Transactions on Medical Imaging* PP, pp. 1–1. DOI: 10.1109/TMI.2021.3069558.

Krook-Magnuson, Esther and Ivan Soltesz (2015). "Beyond the hammer and the scalpel: selective circuit control for the epilepsies". In: *Nature Neuroscience* 18, pp. 331–338.

Lockdown decree (2020). Gazzetta Ufficiale. URL: https://www.gazzettaufficiale.it/eli/gu/2020/03/09/62/sg/pdf.

M Evans, N Hastings and B Peacock (2000). "Statistical Distributions, Third Edition". In: *Measurement Science and Technology* 12.1, pp. 117–117. DOI: 10.1088/0957-0233/12/1/702.

Ma, Debiao, Junteng Zheng, and Lizhi Peng (2021). "Performance Evaluation of Epileptic Seizure Prediction Using Time, Frequency, and Time–Frequency Domain Measures". In: *Processes* 9.4. ISSN: 2227-9717. DOI: `10.3390/pr9040682`.

Mahdizadeh Gharakhanlou, Navid and Navid Hooshangi (2020). "Spatio-temporal simulation of the novel coronavirus (COVID-19) outbreak using the agent-based modeling approach (case study: Urmia, Iran)". In: *Informatics in Medicine Unlocked* 20, p. 100403. ISSN: 2352-9148. DOI: `https://doi.org/10.1016/j.imu.2020.100403`.

Manzo, Gianluca (May 2020). "Complex Social Networks are Missing in the Dominant COVID-19 Epidemic Models". In: *Sociologica* 14, pp. 31–49. DOI: `10.6092/issn.1971-8853/10839`.

Markram, Henry et al. (2015). "Reconstruction and Simulation of Neocortical Microcircuitry". In: *Cell* 163.2, pp. 456–492. ISSN: 0092-8674. DOI: `https://doi.org/10.1016/j.cell.2015.09.029`.

MATLAB v. 2021b (2021). The MathWorks, Inc.

Meyer, F. (1992). "Color image segmentation". In.

Milo, R. et al. (2002). "Network Motifs: Simple Building Blocks of Complex Networks". In: *Science* 298.5594, pp. 824–827. ISSN: 0036-8075. DOI: `10.1126/science.298.5594.824`.

Morgan, Josh L. and J. Lichtman (2020). "An Individual Interneuron Participates in Many Kinds of Inhibition and Innervates Much of the Mouse Visual Thalamus". In: *Neuron*.

Møller, Jesper and Rasmus P. Waagepetersen (2007). "Modern Statistics for Spatial Point Processes". In: *Scandinavian Journal of Statistics* 34.4, pp. 643–684. DOI: `https://doi.org/10.1111/j.1467-9469.2007.00569.x`.

Newman, M. E. J. (2010). *Networks: an introduction*. Oxford; New York: Oxford University Press. ISBN: 9780199206650 0199206651.

Neyman, J. (1937). "Outline of a Theory of Statistical Estimation Based on the Classical Theory of Probability". In: *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences* 236.767, pp. 333–380. ISSN: 00804614.

Nigam, Ishan et al. (2015). "Revisiting HEp-2 Cell Image Classification". In: *IEEE Access* 3, pp. 3102–3113.

Noachtar, Soheyl and Jan Rémi (Mar. 2009). "The role of EEG in epilepsy: A critical review". In: *Epilepsy & behavior : E&B* 15, pp. 22–33. DOI: `10.1016/j.yebeh.2009.02.035`.

Nykamp, Duane et al. (Mar. 2017). "Mean-Field Equations For Neuronal Networks With Arbitrary Degree Distributions". In: *Phys. Rev.* DOI: `10.1101/118463`.

Olver, Frank W. et al. (2010). *NIST Handbook of Mathematical Functions*. 1st. USA: Cambridge University Press. ISBN: 0521140633.

Osservatorio UnipolSai 2018 (2018). UnipolSai Assicurazioni. URL: `https://www.unipolsai.com/sites/corporate/files/pages\_related\_documents/cs\_osservatorio-unipolsai-2018.pdf`.

Otsu, Nobuyuki (1979). "A threshold selection method from gray level histograms". In: *IEEE Transactions on Systems, Man, and Cybernetics* 9, pp. 62–66.

Page, Lawrence et al. (1999). *The PageRank Citation Ranking: Bringing Order to the Web*. Technical Report 1999-66. Previous number = SIDL-WP-1999-0120. Stanford InfoLab. URL: `http://ilpubs.stanford.edu:8090/422/`.

Pan, Victor Y., Zhao Q. Chen, and Ailong Zheng (1998). *The Complexity of the Algebraic Eigenproblem*.

Park, Hojeong and Songhee H. Kim (2020). "A Study on Herd Immunity of COVID-19 in South Korea: Using a Stochastic Economic-Epidemiological Model". In: *Environmental & Resource Economics*, pp. 1 –6.

Potjans, Tobias C. and Markus Diesmann (Dec. 2012). "The Cell-Type Specific Cortical Microcircuit: Relating Structure and Activity in a Full-Scale Spiking Network Model". In: *Cerebral Cortex* 24.3, pp. 785–806. ISSN: 1047-3211. DOI: 10.1093/cercor/bhs358.

Python Language Reference, version 3.6 (2016). Python Software Foundation. URL: http://www.python.org.

Rayleigh (n.d.). "The Problem of the Random Walk". In: *Nature* 72 (), pp. 318–318.

Regazzini, Eugenio (1987). "De Finetti's Coherence and Statistical Inference". In: *The Annals of Statistics* 15.2, pp. 845–864. ISSN: 00905364.

Rogawski, Michael A. and Wolfgang Löscher (2004). "The neurobiology of antiepileptic drugs". In: *Nature Reviews Neuroscience* 5, pp. 553–564.

Ronneberger, Olaf, Philipp Fischer, and Thomas Brox (2015). "U-Net: Convolutional Networks for Biomedical Image Segmentation". In: *MICCAI*.

Saggio, Maria Luisa et al. (2020). "A taxonomy of seizure dynamotypes". In: *eLife* 9. Ed. by Frances K Skinner et al., e55632. ISSN: 2050-084X. DOI: 10.7554/eLife.55632.

Shapiro, Mark B et al. (2021). "Adaptive Susceptible-Infectious-Removed Model for Continuous Estimation of the COVID-19 Infection Rate and Reproduction Number in the United States: Modeling Study". In: *J Med Internet Res* 23.4, e24389. ISSN: 1438-8871. DOI: 10.2196/24389.

slides Lombardia (2018). blendspace. URL: https://www.blendspace.com/lessons/RP1R6l5ZB05CMA/lombardia-settore-secondario-terziario-vie-di-comunicazione-e-popolazione.

Son Woo-Sik, et al. (2020). "Individual-based simulation model for COVID-19 transmission in Daegu, Korea". In: *Epidemiol Health* 42.0, e2020042–0. DOI: 10.4178/epih.e2020042.

Song, Sang-Ho and George J. Augustine (2016). "Synapsin Isoforms Regulating GABA Release from Hippocampal Interneurons". In: *Journal of Neuroscience* 36.25, pp. 6742–6757. ISSN: 0270-6474. DOI: 10.1523/JNEUROSCI.0011-16.2016.

St Louis, Erik (June 2009). "Minimizing AED Adverse Effects: Improving Quality of Life in the Interictal State in Epilepsy Care". In: *Current neuropharmacology* 7, pp. 106–14. DOI: 10.2174/157015909788848857.

Steen, Maarten van (2010). *Graph Theory and Complex Networks An Introduction*. Published by Maarten van Steen. ISBN: 978-90-815406-1-2.

Stefanescu, Roxana A. and Viktor K. Jirsa (Nov. 2008). "A Low Dimensional Description of Globally Coupled Heterogeneous Neural Networks of Excitatory and Inhibitory Neurons". In: *PLOS Computational Biology* 4.11, pp. 1–17. DOI: 10.1371/journal.pcbi.1000219.

Stockman, George C. and Linda G. Shapiro (2001). "Computer Vision". In.

Tajbakhsh, Nima et al. (2016). "Convolutional Neural Networks for Medical Image Analysis: Full Training or Fine Tuning?" In: *IEEE Transactions on Medical Imaging* 35, pp. 1299–1312.

Takemura, Shin-Ya et al. (Aug. 2013). "A visual motion detection circuit suggested by Drosophila connectomics". In: *Nature* 500.7461, pp. 175–181. ISSN: 0028-0836. DOI: 10.1038/nature12450.

Tartof, Sara Y. et al. (2021). "Effectiveness of mRNA BNT162b2 COVID-19 vaccine up to 6 months in a large integrated health system in the USA: a retrospective cohort study". In: *Lancet (London, England)* 398, pp. 1407 –1416.

Turner, Nicholas L. et al. (2020). "Multiscale and multimodal reconstruction of cortical structure and function". In: *bioRxiv*. DOI: 10.1101/2020.10.14.338681.

Wainberg, Michael et al. (2018). "Deep learning in biomedicine". In: *Nature Biotechnology* 36, pp. 829–838.

Wang, Yu et al. (Feb. 2017). "Comparison of Different Generalizations of Clustering Coefficient and Local Efficiency for Weighted Undirected Graphs". In: *Neural Computation* 29.2, pp. 313–331. ISSN: 0899-7667. DOI: 10.1162/NECO_a_00914.

Watts, Duncan J. and Steven H. Strogatz (1998). "Collective dynamics of 'small-world' networks". In: *Nature* 393.6684, pp. 440–442. DOI: 10.1038/30918.

Wu, Guangming et al. (May 2019). "A Stacked Fully Convolutional Networks with Feature Alignment Framework for Multi-Label Land-cover Segmentation". In: *Remote Sensing* 11. DOI: 10.3390/rs11091051.

Ye, Q.-Z. (1988). "The signed Euclidean distance transform and its applications". In: *[1988 Proceedings] 9th International Conference on Pattern Recognition*, 495–499 vol.1. DOI: 10.1109/ICPR.1988.28276.

Yi, Jingru et al. (2019). "Multi-scale Cell Instance Segmentation with Keypoint Graph based Bounding Boxes". In: *MICCAI*.

Zhao, Hengshuang et al. (2018). "ICNet for Real-Time Semantic Segmentation on High-Resolution Images". In: *ArXiv* abs/1704.08545.