UCL
Université catholique de Louvain

École polytechnique de Louvain (EPL)

Diffusion and epidemics on networks

Dissertation presented by
Julien Di Giovanni

for obtaining the Master's degree in
Mathematical Engineering

Supervisor(s)
Jean-Charles Delvenne

Reader(s)
Pierre-Antoine Absil, Philippe Chevalier

Academic year 2017-2018

Contents

Acknowledgement

First of all, I would like to thanks my advisor professor Jean-Charles Delvenne for his continuous support and help during this whole year. I express him my deepest gratitude.

Then I thank my friends and my family for helping by giving me inspirations.

Julien Di Giovanni

Chapter 1

Introduction

Diffusion problems in networks are a recurrent phenomenon that present many applications since either population or city network are modelized by graph. Therefore, in many sectors one would like to understand the mechanism of the diffusion, and more precisely if the whole network will be concerned by a given phenomenon or if just a few vertices will be. Applications are also various. Indeed, in 2010, the events of the Arab Spring highlighted the important role played by social medias in the propagation of the main ideas[11]. Another concrete example of this is the disease diffusion in human being networks.

This document is thus about diffusion in network, in the particular case of disease diffusion. Epidemics are phenomena appearing nearly on an annual basis (think about the flu). And on larger time interval, we see the apparition of a new disease that could possibly become an epidemics (like the avian influenza or the swine fever, for example). But there are also diseases that go on spreading year after year, like HIV. Therefore, epidemics are recurrent and curent facts. In the case of small and dense communities, many models that can predict the outbreak of a disease exist. However, nowadays, many people travel all around the world, so that their contacts are not limit to a small community anymore. We thus have to consider larger networks. Moreover, these network are not dense since most of people do not know all other people but only a small number of them (of the same community). Another observation is that, in a given population, an individual often has generally more contact with the same people, even in small population. Therefore, the structure of the population will play a role in the diffusion of the disease.

We will thus look at how the structure of a society impacts the diffusion of a disease in the network. More precisely, we will look after a parameter called the global reproductive number, which gives an indication on the diffusion speed in a network, so that we can define if the disease will lead to an epidemic or not. This number is obtained thanks to the adjacent matrix of the graph. It is important to note that we consider that all people are the same (at least in a community from the whole population), which means that we do not care about individuality, everyone gets infected or cured in the same way. This document is divided into four chapters that are described here below.

- ◇ **Classical epidemic model in homogeneous population**
  In the first chapter, we will briefly describe two epidemic models and explain their mechanism. These models are true for homogeneous population (i.e. small communities), so there are not really interesting in the frame of this work. But they present the great advantage that we can consider them as dynamical systems, so we can easily compute their equilibria and

make a stability study of these. This in turn will lead us to the expression of a parameter that characterizes the diffusion in a network. The main goal of this chapter is to give an intuitive notion of these parameter.

◇ **Case of two communities**

In the second chapter we will analyze the case where the whole population is divided into two communities of the same size. We will consider a directed graph which means that the relation between vertices happens through transmitted contamination and not through a contact relation. We will find an expression of a global reproductive number. We will thus compute this parameter in two kind of populations. In the first population, there are two distinct communities linked to each other with a given rate. The other kind of population we will analyze is a population with a dense community which is linked to a nearly empty community. This kind of networks is called a core-periphery network.

◇ **Generalization to several communities**

The main goal of this chapter is to develop a tool to find the global reproductive number in the cases where the population is composed of more than two communities. Moreover this tool will allow us to consider networks with communities of different sizes. We will illustrate this tool with examples of networks of four communities of an equal size. We will also see that this tool might be powerful but presents many disadvantages.

◇ **Statistical study of random networks**

In chapter two and three, we focus on dense networks, which means that they can be approximate of a smaller network (less number of vertices) of homogeneous communities. These models become thus less good for low density networks. This chapter tries to understand a network with low density by looking at the global reproductive number. Since we do not want to study a specific network, we will work on random networks that share the same characteristics, like the average degree for the vertices. We will thus make numerical simulations to get estimations of the global reproductive number. This time, this number will not be an exact expression but an approximation base on statistical study of the random networks.

Chapter 2

Classical epidemics model in homogeneous populations

2.1 Introduction

We will begin with the classical case of the epidemic, which means that we look at how the disease spreads in a single homogeneous population where everyone knows each other and sees everyone at each time. This is for example the case in small communities like families, class rooms or offices. That way, we can approximate the network by a continuous and homogeneous population, forgetting that we speak of people. So we will replace the discrete number of people by a continuous quantity.

Our main approach consists in studying two epidemic models as dynamical systems and study the stability on its equilibrium points. We will see that the stability essentially depends on a parameter called the reproductive number called R_0. We will consider a fast diffusion so that the epidemic process has shorter time than life time of individuals, so that we can neglect natural death and birth of individual.

2.2 Description of epidemic models

We are now focusing on the different existing "traditional" models. Actually, there exists a lot of epidemic models: SI, SIS, $SIER$, $SIRS$,... but we will just explain the SIS and SIR models, that may at first sight look similar, but that present a major difference. In these two models, individuals can be in different states. The description of the models and their equation are found internet[1] :

- ◇ S: susceptible individuals, i.e.that are not infected but can become infected.
- ◇ I : infected individuals.
- ◇ R : removed individuals i.e. individuals that are either dead or cured. In both cases they don't play any role in the spreading process.

2.2.1 SIS model

In this model, individuals are susceptible (S) and can become infected with rate β at the contact with an infected individual, and infected (I) individuals get cured with rate μ and become susceptible again, like shown in figure 2.1. At a small time interval, a susceptible individual becomes infected by contact with an infected individual with probability β and an infected individual becomes susceptible again with probability μ. The mean time before getting cured is given by μ^{-1}.

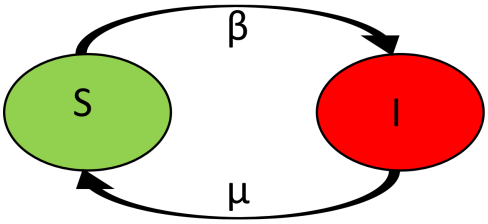


Figure 2.1: Transition state for the SIS model

In fact, we can describe the model with the following reaction :

$$\begin{aligned} \mathrm{I} + \mathrm{S} &\longrightarrow 2\,\mathrm{I} \\ \mathrm{I} &\longrightarrow \mathrm{S} \end{aligned}$$

The equations that describe the model are thus the following: let $S(t)$ the number of susceptible people in the population and $I(t)$ the number of infected people in the population. Note that since the population is constant,$S(t) + I(t) = N$, the number of individuals. The dynamic of the system is then given by equations (2.1) and (2.2).

$$\dot{S} = -\beta \frac{SI}{N} + \mu I \tag{2.1}$$

$$\dot{I} = \beta \frac{SI}{N} - \mu I \tag{2.2}$$

Note that since we make the hypothesis that all individuals are the same, the infectious rate β is the same for all people, and so is the recovery rate μ .

2.2.2 SIR model

This model is quite the same as the SIS model, the only difference being that after being infected, individuals are removed from the system (get cured or die) so they don't participate to the dynamic of the system anymore, like shown in figure 2.2:

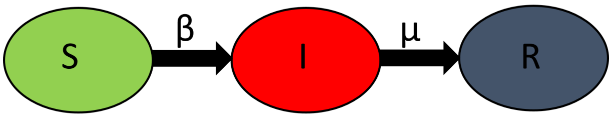


Figure 2.2: Transition state for the SIR model

The reactions we have to consider here are thus:

$$\begin{aligned} \mathrm{I} + \mathrm{S} &\longrightarrow 2\,\mathrm{I} \\ \mathrm{I} &\longrightarrow \mathrm{R} \end{aligned}$$

and the dynamic is given by equation (2.3), (2.4) and (2.5). Let $R(t)$ the number of removed individuals, every time, t : $S(t) + I(t) + R(t) = N$.

$$\dot{S} = -\beta \frac{SI}{N} \tag{2.3}$$

$$\dot{I} = \beta \frac{SI}{N} - \mu I \tag{2.4}$$

$$\dot{R} = \mu I \tag{2.5}$$

2.3 Epidemic threshold and reproductive number

We are now interested in the value of the β and μ parameters, to find the way the disease spreads in all the population. In that way, we will use the reproductive number that is defined as the average number of people an infected individual will contaminate before healing. So we know that an epidemic occurs when this reproductive number (R_0) is larger than one (= epidemics thresholds). In this section, we will show to find this parameter.

2.3.1 SIS model

If we compute the equilibria of the dynamic, we find two cases: the first one is if $I = 0$ and $S = N$ and by computing the jacobian matrix J at this equilibrium we find

$$J = \begin{pmatrix} 0 & \mu - \beta \\ 0 & \beta - \mu \end{pmatrix}$$

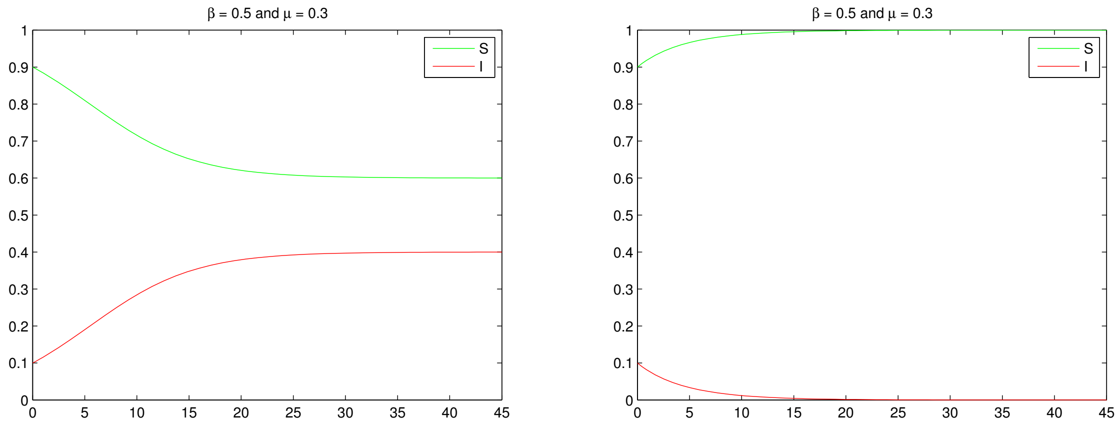
the eigenvalue of J are 0 and $\beta - \mu$, so this equilibrium is unstable if $\beta > \mu$, which in this case means that the presence of an infected individual in the system will make the disease diffuse. And in this case, we are sure that every disease that has parameter $\beta > \mu$ will lead to an epidemic.

The other equilibrium point corresponds to : $S = \frac{\mu}{\beta} N$ and $I = \left(1 - \frac{\mu}{\beta}\right) N$. So it is obvious that this equilibrium exists only if $\beta > \mu$. Here, the jacobian becomes:

$$J = \begin{pmatrix} \mu - \beta & 0 \\ \beta - \mu & 0 \end{pmatrix}$$

The eigenvalue are 0 and $\mu - \beta < 0$, so this equilibrium is always stable when it exists.

Therefore, depending on the value of the parameter, the disease will die if $\beta < \mu$ and will grows to reach a fraction $\left(1 - \frac{\mu}{\beta}\right)$ of the population when $t \to \infty$ if $\beta > \mu$. We can thus set $R_0 = \frac{\beta}{\mu}$. Such a behaviour is shown in figure 2.3, where we have plot the solution of equation (2.1) and (2.2) with different value of R_0. We can see that if $R_0 < 1$, then there is no epidemic.



(a) Solution with $\beta = 0.5$ and $\mu = 0.3$ ($R_0 > 1$) (b) Solution with $\beta = 0.3$ and $\mu = 0.5$ ($R_0 < 1$)

Figure 2.3: Numerical solution of the SIS equations and illustration of the equilibria stability

2.3.2 SIR model

By looking at equation (2.3), (2.4) and (2.5), we find that the equilibrium is only reached if $I = 0$ and so $S + R = N$. We thus have to discuss the value of S and R. Let's first assume that $R = 0$ and $S = N$. This can only be seen at the beginning of the epidemic since once an individual is infected, he will be removed. In this case the jacobian matrix becomes :

$$J = \begin{pmatrix} 0 & -\beta & 0 \\ 0 & \beta - \mu & 0 \\ 0 & \mu & 0 \end{pmatrix}$$

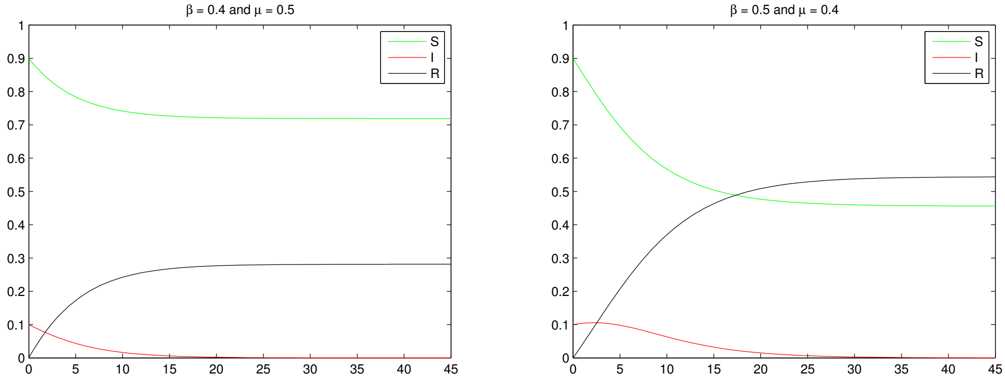
The eigenvalues are $0, 0$ and $\beta - \mu$, so like in the SIS model the disease will diffuse if $\beta > \mu$. The other extreme case is : $S = 0$ and $R = N$. It means that all the population were infected and then cured, the jacobian matrix is then given by:

$$J = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -\mu & 0 \\ 0 & \mu & 0 \end{pmatrix}$$

Here the eigennvalues are $0, 0$ and $-\mu$. This is indeed a stable equilibrium. Finally the general case is $R = \bar{R}$ and $S = N - \bar{R}$. In this case, the jacobian matrix is given by

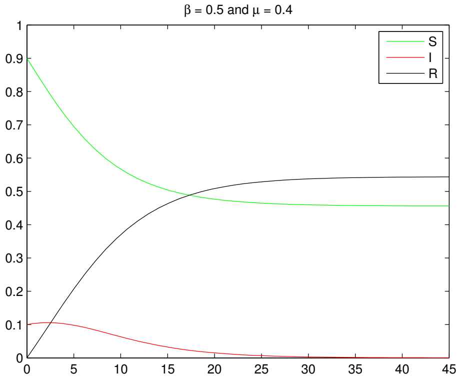
$$J = \begin{pmatrix} 0 & -\beta\dfrac{N-\bar{R}}{N} & 0 \\ 0 & \beta\dfrac{N-\bar{R}}{N} - \mu & 0 \\ 0 & \mu & 0 \end{pmatrix}$$

In this case, the eigenvalues are $0, 0$ and $\beta\left(1 - \dfrac{\bar{R}}{N}\right) - \mu$. Thus, there exists a fraction of the population that needs to become R to stop the diffusion of the disease. Therefore, again, here a value of R_0 might be $R_0 = \dfrac{\beta}{\mu}$. We can indeed see on figures 2.4 that concerning the R_0 parameter the system stabilizes differently.



(a) Solution with $\beta = 0.4$ and $\mu = 0.5$ ($R_0 < 1$)

(b) Solution with $\beta = 0.5$ and $\mu = 0.4$ ($R_0 > 1$)



(c) Solution with $\beta = 0.9$ and $\mu = 0.1$ ($R_0 < 1$)

Figure 2.4: Numerical solution of the SIR equations and illustration of the equilibria stability

Indeed, when $R_0 > 1$, $\beta > \mu$. People thus get easily ill than cured so, intuitively, the disease spreads faster. The intermediate equilibrium point (figure 2.4b) comes from the fact that we consider an homogeneous population. Therefore, infected people get surrounded from enough R-people to not infect more people.

2.4 Conclusion

In this chapter, we have seen two epidemic models and computed their stability. We have seen that both models depend on a parameter called the reproductive number. This parameter is determinant in the diseasespreading. However, these models only remain correct in the case of a single homogeneous population, which is limitd to small communities. Indeed, in a larger community like a town, a city or a country, since the homogeneous hypothesis of the population no longer holds, we can expect that the disease will spread differently. For example, consider that a disease with $R_0 > 1$ appears in a town. If there is no contact between this town and another one, then the epidemic will be limitd to that particular town. We can eventually speak about a local spreading, but this disease will not reach a large proportion of the whole population.

In next chapter we will look at how the disease will spread in a more complex network where the homogeneous hypothesis is not valid anymore. We will more specifically try and find another parameter that can characterize the disease diffusion : the global reproductive number

Chapter 3

Case of two communities

3.1 Introduction

Previous chapter studied the behaviour of the diseasespreading on a single homogeneous population (i.e. where everyone knows each other and has contact with everyone at each time). We will now look at what happens in the case where this population is not homogeneous anymore, but is divided into two smaller communities that have a particular shape. To make it simple, we will study the case of a directed graph where the two communities have the same size. The directed graph can be seen as a network in which an edge from an individual i to an individual j means that i infects j and not that there is a contact between the two of them. Inside each community, individuals have the same probability to give the disease to someone else. As shown in the previous section, we will get an epidemic if the reproductive number R_0 is above the epidemic threshold ($= 1$). Intuitively, the condition $R_0 > 1$ will not be sufficient to conclude that there is an epidemic or not. Indeed, let's for example consider the case where we have an individual that is infected. If he has no contact before healing, the disease will not spread even if $R_0 > 1$. We will thus need to define a second epidemic number R_* that will characterize the spreading of the disease within the network : the global reproductive number.

To find this R_* we thus need to study the form of the network and, more precisely, its adjacent matrix A. To know how to find R_*, we define this model in discrete time found in a master thesis from 2017 [2]: Let S_k,I_k, R_k be binary vectors where entry i corresponds to individual i. Therefore, if $I_k(i)$ the i^{th} entry of vector I_k is equal to one, then the individual i is infected at time k. S_k and R_k denotes respectively the binary vectors of the susceptible people and the removed people at time k. Finally, let binary vector V_k be the people being infected at time k. The discrete time model is then defined, in the SIR model, by (on average) :

$$\begin{aligned}
V_{k+1} &= AV_k \\
I_{k+1} &= \max\left(\min\left(\sum_{l=k-\mu^{-1}}^{k} V_l, 1\right) - R_k, 0\right) \\
R_{k+1} &= \sum_{l=0}^{k-\mu^{-1}} I_l \\
S_{k+1} &= \max\left(S_k - I_k - R_k, 0\right)
\end{aligned}$$

and in the SIS model :

$$\begin{aligned} V_{k+1} &= AV_k \\ I_{k+1} &= \min\left(\sum_{l=k-\mu^{-1}}^{k} V_l, 1\right) \\ S_{k+1} &= \max\left(S_k - I_k, 0\right) + V_{k-\mu^{-1}} \end{aligned}$$

Remember that the mean time before healing is μ^{-1}. Note that V_l with $l < 0$ is a zero vector. We thus see that all states depend on the vector V_k which is obtained by a recurrence equation so we can write : $V_k = A^k V_0$. Therefore, the evolution of V_k will be determined by the eigenvalues of A and more specifically by the spectral radius $\rho(A)$. Indeed, if $\rho(A) < 1$ then $V_k \to 0$ when $k \to \infty$. So we can define that $R_* = \rho(A)$ the condition to have epidemic are then :

$$\begin{aligned} R_0 &> 1 \\ R_* &> 1 \end{aligned}$$

Note that this definition is essentially good for an SIR model and a bit less for an SIS model. Indeed, for the SIS model, when $R_0 > 1$ we get a stationary number of infected people, we can be sure that the disease will diffuse in the whole network. Consequently, the condition $R_* > 1$ only holds in an SIR model. This chapter and the following one will thus be the study of the spectral radius of the adjacent matrix. Moreover, the population will be approximate by a graph of two dense and homogeneous communities in which all individuals are the same. Note that the density is define by the expected degree of a vertex in divided by the number of vertices.

3.2 Properties

We begin this chapter by giving two properties on the characteristic polynomial of a square matrix of size n with every entry equals to each other. These properties will be useful since we will work on average cases. Indeed, we will work on random networks. That means that we place an edge between two vertices with a given probability p. If this probability is the same for all edges then every entry of the adjacent matrix will be one with probability p and so the average adjacent matrix will be a matrix with every entry equals to p.

Property

Let $C_n(\lambda) = \det\left(a\mathbb{1}_n - \lambda\,\mathrm{I}_n\right)$ the characteristic polynomial of the matrix $a\mathbb{1}_n$, with I_n be the identity matrix of dimension n, $a \in \mathbb{R}$ and let $B_n(\lambda)$ be the determinant of matrix r of size $n \times n$. $\mathbb{1}_n$ is a square matrix of size n with every entry equal to 1 :

$$r = \begin{pmatrix} a & a & a & \cdots & a \\ a & a-\lambda & a & \cdots & a \\ a & a & a-\lambda & \cdots & a \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a & a & a & \cdots & a-\lambda \end{pmatrix}$$

We have the following recurrsive relation :

$$\begin{aligned} C_n(\lambda) &= (a-\lambda)C_{n-1}(\lambda) - a(n-1)B_{n-1}(\lambda) \\ B_n(\lambda) &= aC_{n-1}(\lambda) - a(n-1)B_{n-1}(\lambda) \end{aligned}$$

Proof. Let's first compute $C_n(\lambda) = \det\left(a\mathbb{1}_n - \lambda\,\mathrm{I}_n\right)$

$$\begin{aligned} C_n(\lambda) &= \det\begin{pmatrix} a-\lambda & a & a & \cdots \\ a & a-\lambda & a & \cdots \\ a & a & a-\lambda & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \\ &= (a-\lambda)C_{n-1}(\lambda) - a\det\begin{pmatrix} a & a & a & \cdots \\ a & a-\lambda & a & \cdots \\ a & a & a-\lambda & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} + a\det\begin{pmatrix} a-\lambda & a & a & \cdots \\ a & a & a & \cdots \\ a & a & a-\lambda & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} + \cdots \\ &= (a-\lambda)C_{n-1}(\lambda) - a\det\begin{pmatrix} a & a & a & \cdots \\ a & a-\lambda & a & \cdots \\ a & a & a-\lambda & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} - a\det\begin{pmatrix} a & a & a & \cdots \\ a & a-\lambda & a & \cdots \\ a & a & a-\lambda & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} + \cdots \\ &= (a-\lambda)C_{n-1}(\lambda) - aB_{n-1}(\lambda) - aB_{n-1}(\lambda) + \cdots \end{aligned}$$

The i^{th} term of the sum ($i = 2, ..., n$) is on the form $a(-1)^{1+i}$ times the determinant of a matrix that has elements equal to a if they are off-diagonal and equal to $a - \lambda$ else, except on the i^{th} line where all elements are equal to a (on the diagonal too). So using $i - 2$ permutations we can bring this matrix to the form of matrix r. Using properties of determinant :

$$\begin{aligned} C_n(\lambda) &= (a-\lambda)C_{n-1}(\lambda) + a\sum_{i=2}^{n}(-1)^{1+i}(-1)^{i-2}B_{n-1}(\lambda) \\ &= (a-\lambda)C_{n-1}(\lambda) + a\sum_{i=2}^{n}(-1)^{2i-1}B_{n-1}(\lambda) \\ &= (a-\lambda)C_{n-1}(\lambda) - a(n-1)B_{n-1}(\lambda) \end{aligned}$$

Let's now compute $B_n(\lambda) = \det(r)$, in a similar way we find:

$$
\begin{aligned}
B_n(\lambda) &= \det \begin{pmatrix} a & a & a & \cdots \\ a & a-\lambda & a & \cdots \\ a & a & a-\lambda & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \\
&= a \det \begin{pmatrix} a-\lambda & a & a & \cdots \\ a & a-\lambda & a & \cdots \\ a & a & a-\lambda & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} + \cdots + a(-1)^{1+i}(-1)^{i-2} \det \begin{pmatrix} a & a & a & \cdots \\ a & a-\lambda & a & \cdots \\ a & a & a-\lambda & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} + \cdots \\
&= aC_{n-1}(\lambda) + a\sum_{i=2}^{n}(-1)^{1+i}(-1)^{i-2}B_{n-1}(\lambda) \\
&= aC_{n-1}(\lambda) + a(n-1)B_{n-1}(\lambda)
\end{aligned}
$$

□

Property

Let $C_n(\lambda) = \det(a\mathbb{1}_n - \lambda \, \mathrm{I}_n)$ the characteristic polynomial of the matrix $a\mathbb{1}_n$, and let $B_n(\lambda)$ be the determinant of matrix r :

$$
r = \begin{pmatrix} a & a & a & \cdots & a \\ a & a-\lambda & a & \cdots & a \\ a & a & a-\lambda & \cdots & a \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a & a & a & \cdots & a-\lambda \end{pmatrix}
$$

Then :

$$
\begin{aligned}
C_n(\lambda) &= (-\lambda)^{n-1}(an-\lambda) \\
B_n(\lambda) &= a(-\lambda)^{n-1}
\end{aligned}
$$

Proof. Let's first compute $C_2(\lambda)$ and $B_2(\lambda)$:

$$
\begin{aligned}
C_2(\lambda) &= \det \begin{pmatrix} a-\lambda & a \\ a & a-\lambda \end{pmatrix} \\
&= (a-\lambda)^2 - a^2 \\
&= a^2 - 2a\lambda + \lambda^2 - a^2 = \lambda(\lambda - 2a) \\
B_2(\lambda) &= \det \begin{pmatrix} a & a \\ a & a-\lambda \end{pmatrix} \\
&= a(a-\lambda) - a^2 \\
&= a^2 - a\lambda - a^2 = -a\lambda
\end{aligned}
$$

Then let's suppose that the $C_n(\lambda)$ and $B_n(\lambda)$ have the form we want to show, and let's compute

$C_{n+1}(\lambda)$ and $B_{n+1}(\lambda)$ using the recurrsive property found before. Let's first consider that n is even.

$$\begin{aligned}
C_{n+1}(\lambda) &= (a-\lambda)C_n(\lambda) - anB_n(\lambda) \\
&= (a-\lambda)(-\lambda)^{n-1}(an-\lambda) - a^2n(-\lambda)^{n-1} \\
&= (a-\lambda)\lambda^{n-1}(\lambda - an) + a^2n\lambda^{n-1} \\
&= a\lambda^n - a^2n\lambda^{n-1} - \lambda^{n+1} + an\lambda^n + a^2n\lambda^{n-1} \\
&= \lambda^n(a(n+1)-\lambda) = (-\lambda)^n(a(n+1)-\lambda) \\
B_{n+1}(\lambda) &= aC_n(\lambda) - anB_n(\lambda) \\
&= a(-\lambda)^{n-1}(an-\lambda) - a^2n(-\lambda)^{n-1} \\
&= -a^2n\lambda^{n-1} + a\lambda^n + a^2n\lambda^{n-1} \\
&= a\lambda^n = a(-\lambda)^n
\end{aligned}$$

Let's now consider the case where n is odd.

$$\begin{aligned}
C_{n+1}(\lambda) &= (a-\lambda)(-\lambda)^{n-1}(an-\lambda) - a^2n(-\lambda)^{n-1} \\
&= (a-\lambda)\lambda^{n-1}(an-\lambda) - a^2n\lambda^{n-1} \\
&= a^2n\lambda^{n-1} - a\lambda^n - an\lambda^n + \lambda^{n+1} - a^2n\lambda^{n-1} \\
&= -\lambda^n(a(n+1)-\lambda) = (-\lambda)^n(a(n+1)-\lambda) \\
B_{n+1}(\lambda) &= a(-\lambda)^{n-1}(an-\lambda) - a^2n(-\lambda)^{n-1} \\
&= a\lambda^{n-1}(an-\lambda) - a^2n\lambda^{n-1} \\
&= a^2n\lambda n - 1 - a\lambda^n - a^2n\lambda^{n-1} = a(-\lambda)^n
\end{aligned}$$

□

We have thus shown that the characteristic polynomial has a particular form and thus that the eigenvalues of such a matrix are all equal to zero. Except one that is equal to an, with $a \in \mathbb{R}$ So all element of the matrix are equal to a, and n the dimension of the matrix.

3.3 Link between two communities

Here we will study the evolution of the global reproductive number in the case when we have two communities of density P_1 and P_2 (probability to transmit the disease) and both of size n that is that a vertex from the first community has probability P_1 to get an edge with every vertices and each vertex in community 2 has probability P_2 to get an edge with every other vertex. Moreover, we define the parameter $0 \leq \gamma \leq 1$ called the linking coefficient that is the average fraction of link between the communities. The average adjacent matrix is then on the form presented in figure 3.1. We will thus here compute the eigenvalues of this matrix. Here a two examples of such population:

- ◇ For the case of a sexually transmitted disease the community are the men and women and the linking coefficient is the proportion of heterogeneous sexual relations in the population. When $\gamma = 0$ there are thus only homosexual relations and in the case where $\gamma = 1$ there are just heterosexual relations
- ◇ In the case of information diffusion, we consider two kind of people, those who speak a lot and those who speak less. The linking coefficient will then be the fraction of contact a person in a community will have in the second one.

.

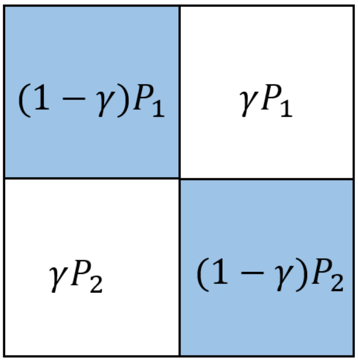


Figure 3.1: Structure of the adjacent matrix

Note that in figure 3.1, all blocks matrix are on the shape $a\mathbb{1}_n$ with $a \in \mathbb{R}$, but the whole adjacent matrix is not on the form $a\mathbb{1}_{2n}$ but is a block matrix. Let A be the adjacent matrix, matrix $A - \lambda\,\mathrm{I}_{2n}$ is thus also a 4 blocks matrix with every block of size $n \times n$, on the form:

$$A - \lambda\,\mathrm{I}_{2n} = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

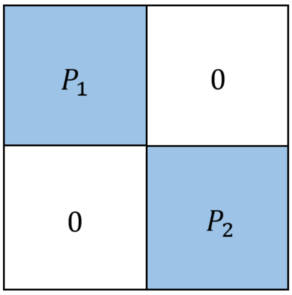
Since all block commute with each other, $\det(A - \lambda\,\mathrm{I}_{2n}) = \det(A_{11}A_{22} - A_{12}A_{21})$ [3]. In addition, matrix $A_{11}A_{22} - A_{12}A_{21}$ has particular form $a\mathbb{1}_n$, we can thus use the properties we find in the previous section to find the characteristic polynomial :

$$\begin{aligned}
\det(A - \lambda\,\mathrm{I}_{2n}) &= \det\begin{pmatrix} (1-\gamma)P_1\mathbb{1}_n - \lambda\,\mathrm{I}_n & \gamma P_1\mathbb{1}_n \\ \gamma P_2\mathbb{1}_n & (1-\gamma)P_2\mathbb{1}_n - \lambda\,\mathrm{I}_n \end{pmatrix} \\
&= \det\Big(((1-\gamma)P_1\mathbb{1}_n - \lambda\,\mathrm{I}_n)((1-\gamma)P_2\mathbb{1}_n - \lambda\,\mathrm{I}_n) - \gamma^2 P_1P_2 n\mathbb{1}_n\Big) \\
&= \det\Big(((1-2\gamma)P_1P_2n - \lambda(1-\gamma)(P_1+P_2))\mathbb{1}_n + \lambda^2\,\mathrm{I}_n\Big) \\
&= \det\Big(-\Big[(\lambda(1-\gamma)(P_1+P_2) - (1-2\gamma)P_1P_2n)\mathbb{1}_n - \lambda^2\,\mathrm{I}_n\Big]\Big) \\
&= (-1)^n \det\Big((\lambda(1-\gamma)(P_1+P_2) - (1-2\gamma)P_1P_2n)\mathbb{1}_n - \lambda^2\,\mathrm{I}_n\Big) \\
&= (-1)^n(-\lambda^2)^{n-1}((\lambda(1-\gamma)(P_1+P_2) - (1-2\gamma)P_1P_2n)n - \lambda^2) \\
&= -\lambda^{2n-2}((\lambda(1-\gamma)(P_1+P_2) - (1-2\gamma)P_1P_2n)n - \lambda^2)
\end{aligned}$$

The eigenvalues are thus the roots of this polynomial, we define $P = P_1 + P_2$ and $\bar{P} = |P_1 - P_2|$. Let's compute the nonzero roots of the polynomial:

$$\begin{aligned}
0 &= -\lambda^2 + (1-\gamma)Pn\lambda - (1-2\gamma)P_1P_2n^2 \\
\varDelta(\gamma) &= (1-\gamma)^2P^2n^2 - 4(1-2\gamma)P_1P_2n^2 \\
&= P^2n^2\gamma^2 - 2n^2(P^2 - 4P_1P_2)\gamma + (P^2 - 4P_1P_2)n^2 \\
&= P^2n^2\gamma^2 - 2n^2\bar{P}^2\gamma + \bar{P}^2n^2 \\
\lambda &= \frac{(1-\gamma)Pn}{2} \pm \frac{\sqrt{\varDelta(\gamma)}}{2}
\end{aligned}$$

Moreover since $\varDelta(\gamma) > 0\ \forall\gamma$ the eigenvalues are real. Therefore, the spectral radius $\rho_\gamma(A)$ is then given by the dominant eigenvalue so we choose the sign $+$. All we have to do now is to find for which value of γ the spectral radius of the average adjacent matrix is smaller than 1. Let's first look at three particular cases that are $\gamma = 0, 0.5, 1$ these cases are presented respectively in figure 3.2, 3.3 and 3.4.



P_1	0
0	P_2

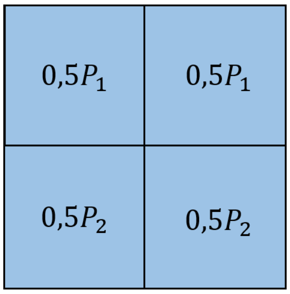
Figure 3.2: Disjoint communities

$\underline{\gamma = 0}$

In this case, we have two disjoint communities, and the two nonzero eigenvalues are :

$$\begin{aligned}\Delta(0) &= P^2n^2 - 4P_1P_2n^2 = \bar{P}^2n^2\\ \lambda_1 &= P_1n\\ \lambda_2 &= P_2n\end{aligned}$$

The spectral radius is given by $\rho_0(A) = \max(P_1, P_2)n$. But in this case the disease will never completely diffuse since there is no connection between the communities.



$0{,}5P_1$	$0{,}5P_1$
$0{,}5P_2$	$0{,}5P_2$

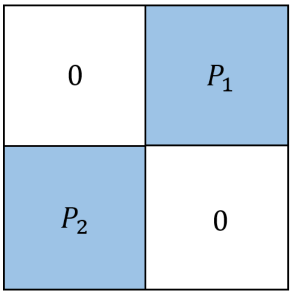
Figure 3.3: Homogeneous community

$\underline{\gamma = 0.5}$

Here, there are as many connections remaining inside the community than connections going outside. So there is only one nonzero eigenvalue.

$$\begin{aligned}\Delta(0.5) &= 0.25P^2n^2\\ \lambda &= \frac{P}{2}n = \frac{P_1 + P_2}{2}n\end{aligned}$$

The spectral radius is given by $\rho_{0.5}(A) = \dfrac{P_1 + P_2}{2}n$.



0	P_1
P_2	0

Figure 3.4: Case with an antidiagonal adjacent matrix

$\underline{\gamma = 1}$

Here, we are in the case of a bipartite graph, that is individuals have only contact with individuals of the other community.

$$\begin{aligned}\Delta(1) &= 4P_1P_2n^2\\ \lambda &= \pm\sqrt{P_1P_2}n\end{aligned}$$

The spectral radius is given by $\rho_1(A) = \sqrt{P_1P_2}n$.

Note that in the particular case of a symmetric matrix $(P_1 = P_2)$ we have that the spectral radius is always equal to $\rho_\gamma(A) = P_1n$ and is thus independent of the linking coefficient. Let's go back to the general case and let's compute the derivative of ρ with respect to γ, we find :

$$\begin{aligned}\frac{d\rho}{d\gamma} &= \frac{-Pn}{2} + \frac{n^2(P^2\gamma - \bar{P}^2)}{2\sqrt{\Delta(\gamma)}}\\ &= \frac{n}{2\sqrt{\Delta(\gamma)}}\left(-P\sqrt{\Delta(\gamma)} + P^2n\gamma - \bar{P}^2n\right)\end{aligned}$$

Which is always negative, indeed:

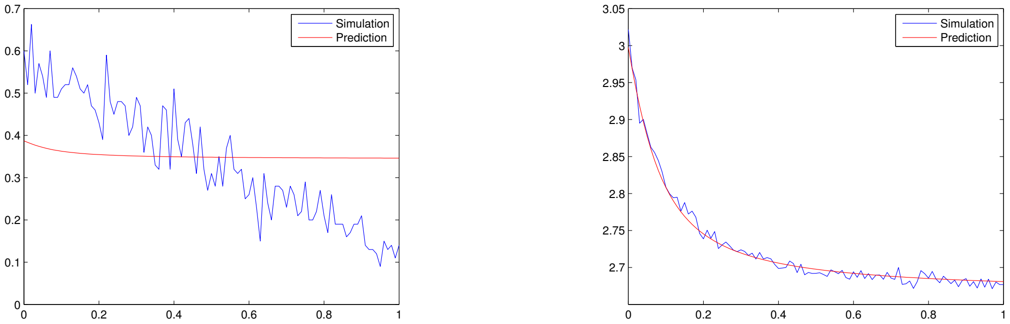
$$\begin{aligned}
& -P\sqrt{\Delta(\gamma)} + P^2 n\gamma - \bar{P}^2 n \leq 0 \\
\Longleftrightarrow & P^2\Delta(\gamma) \geq P^4 n^2\gamma^2 + \bar{P}^4 n^2 - 2p^2\bar{P}^2 n^2\gamma \\
\Longleftrightarrow & P^4 n^2\gamma^2 - 2P^2\bar{P}^2 n^2\gamma + \bar{P}^2 P^2 n^2 \geq P^4 n^2\gamma^2 + \bar{P}^4 n^2 - 2P^2\bar{P}^2 n^2\gamma \\
\Longleftrightarrow & P^2 \geq \bar{P}^2
\end{aligned}$$

This implies that in our example of a sexually transmitted disease, the bigger the fraction of homosexual relations, the faster the disease will spread in the network under the hypothesis that we have as many women as men, that all men are the same, all the women are the same and that the probability for a man to transmit the disease is different from the probability for a woman to transmit it. In the case where these probabilities are equal, then the diffusion of the disease is independent of the fraction of homosexual relations. In our second example, it means that the more people who speak a lot are in relation with each other the fastest the diffusion will be, in the case where there is the same number of people in the communities

The fact that $\rho_\gamma(A)$ decreases can be explained as follow: in the case where $\gamma = 0$, we have two dense communities, so the spectral radius will then be high. On the other side, when $\gamma = 1$, we have a bipartite graph, so the disease can only be transmitted from an individual of one community to someone of the other community. Since $P_1 \neq P_2$, let's suppose that $P_2 < P_1$, then someone from community one will easily contaminate someone from community two while people of community two will contaminate less easily someone from the first community. Therefore the diffusion will slow down. Since $\rho_\gamma(A)$ is a decreasing function of γ, we can define γ_c as the critical value of the linking coefficient above which the spectral radius of the adjacent matrix is less than one. So :

$$\begin{aligned}
& \rho_{\gamma_c}(A) = 1 \\
\Longleftrightarrow & (1-\gamma_c)Pn + \sqrt{\Delta(\gamma_c)} = 2 \\
\Longleftrightarrow & \Delta(\gamma_c) = 4 - 4Pn + 4Pn\gamma_c + P^2n^2 - 2P^2n^2\gamma_c + P^2n^2\gamma_c^2 \\
\Longleftrightarrow & \gamma_c = \frac{1 - Pn + P_1P_2n^2}{2P_1P_2n^2 - Pn}
\end{aligned}$$

Figures 3.5 show the evolution of the spectral radius with γ. We see that it is well a decreasing function. We plot the theorical result we found in this chapter (prediction) in red and the result of numerical simulations in blue. These simulations were made by computing the mean of 100 spectral radius of random binary matrices of size 1000 (two communities of size 500).



(a) $P_1 = 0.0008$ $P_2 = 0.00064$

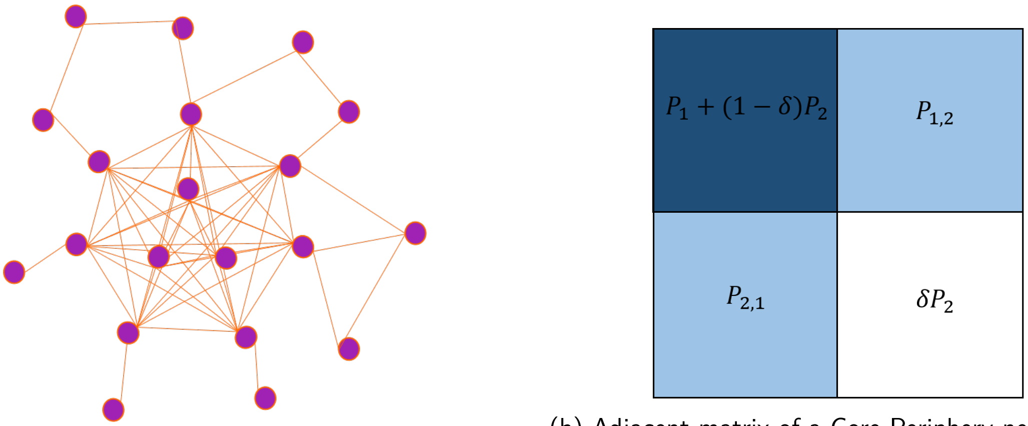
(b) $P_1 = 0.006$ $P_2 = 0.0048$

Figure 3.5: Comparison between the prediction and numerical simulation

We can notice on figure 3.5 that the predictions are correct for large enough probabilities (see figure 3.5a). But for low density, the predictions are no longer valid (see figure 3.5b). This phenomenon can be explained by the fact that the predictions are made on the average case of the adjacent matrix and as the more a network is dense, the closest it is to its average case. So it is quite logical that for dense network the predictions are more correct than for less dense network.

3.4 Core-periphery network

Core-periphery networks correspond to a population with a dense community (core) and a second less dense community (periphery) linked to the first one but less to itself, see figure 3.6a. For example, such a network can be a hospital where the core is composed with the doctors, nurses and the other people working in the hospital and the periphery is composed by the patients and people visiting them. The average adjacent matrix has the particular form described by figure 3.6b. We build this matrix in the following way: We start from a two communities model like in previous section, we then take the edges inside the periphery to put it inside the core P_1 and P_2 are called the natural density of respectively the core and the periphery.



(a) Example of Core-Periphery network

(b) Adjacent matrix of a Core-Periphery network

Figure 3.6: Core-periphery

Thus, in this section we will again compute the eigenvalues of A the average adjacent matrix that will depend on the parameter δ called the concentration coefficient. We will again use the properties found above.

$$\begin{aligned}
\det(A-\lambda \mathrm{I}_{2n}) &= \det\begin{pmatrix} (P_1+(1-\delta)P_2)\mathbb{1}_n - \lambda \mathrm{I}_n & P_{1,2}\mathbb{1}_n \\ P_{2,1}\mathbb{1}_n & \delta P_2\mathbb{1}_n - \lambda \mathrm{I}_n \end{pmatrix} \\
&= \det\left(\left((P_1+(1-\delta)P_2)\,\mathbb{1}_n - \lambda \mathrm{I}_n\right)\left(\delta P_2\mathbb{1}_n - \lambda \mathrm{I}_n\right) - P_{1,2}P_{2,1}n\mathbb{1}_n\right) \\
&= \det\left(\left(P_2^2 n\delta(1-\delta) + P_1P_2n\delta - P_{1,2}P_{2,1}n - \lambda(P_1+P_2)\right)\mathbb{1}_n + \lambda^2 \mathrm{I}_n\right) \\
&= \det\left[-\left(\left(\lambda(P_1+P_2) + P_{1,2}P_{2,1}n - P_1P_2n\delta - P_2^2 n\delta(1-\delta)\right)\mathbb{1}_n - \lambda^2 \mathrm{I}_n\right)\right] \\
&= -\lambda^{2n-2}\left(P_{1,2}P_{2,1}n^2 - P_1P_2n^2\delta - P_2^2n^2\delta(1-\delta) + \lambda n(P_1+P_2) - \lambda^2\right)
\end{aligned}$$

Looking at the non-zero roots of this polynomial to find the non-zero eigenvalue of the adjacent matrix and thus at the global reproductive number ($P_1 + P_2 = P$).

$$0 = P_{1,2}P_{2,1}n^2 - P_1P_2n^2\delta - P_2^2n^2\delta(1-\delta) + \lambda nP - \lambda^2 \tag{3.1}$$

$$\mathit{\Delta}(\delta) = n^2(4P_2^2\delta^2 - 4P_2P\delta + 4P_{1,2}P_{2,1} + P^2) \tag{3.2}$$

$$\lambda = \frac{Pn}{2} \pm \frac{\sqrt{\mathit{\Delta}(\delta)}}{2} \tag{3.3}$$

In the particular the case of two communities with same density and with a symmetric adjacent matrix ($P_1 = P_2$, $P_{1,2} = P_{2,1}$), we found that the spectral radius is:

$$\rho_\delta(A) = P_1n + n\sqrt{P_1^2(1-\delta)^2 + P_{1,2}^2} \tag{3.4}$$

Note that if we consider that $P_{1,2} \ll P_1$, by computing the first degree expansion of the function define by $f(x) = \sqrt{a^2 + x^2}$ around $x = 0$ where $a = P_1(1-\delta)$ and $x = P_{1,2}$, we find:

$$\begin{aligned} f(0) &= a \\ f'(x) &= \frac{x}{\sqrt{a^2+x^2}} \Longrightarrow f'(0) = 0 \\ f(x) &= a + \mathbf{o}(x^2) \end{aligned}$$

And thus, in this particular case:

$$\rho_\delta(A) = P_1n(2-\delta) + \mathbf{o}\left(P_{1,2}^2\right)$$

So the spectral radius will decrease linearly and reach a minimum value when $\delta = 1$. In fact, the spectral radius will indeed decrease nearly linearly depending on the value of $P_{1,2}$ and $P_{2,1}$. Such a behaviour is also present in the general case. In the general case, the spectral radius is given by expression 3.3 by taking the "plus" sign. By computing the derivative of the spectral radius with respect to δ, we find:

$$\frac{d\rho}{d\delta} = \frac{P_2n^2(2P_2\delta - P)}{\sqrt{\mathit{\Delta}(\delta)}}$$

which is negative in the case where:

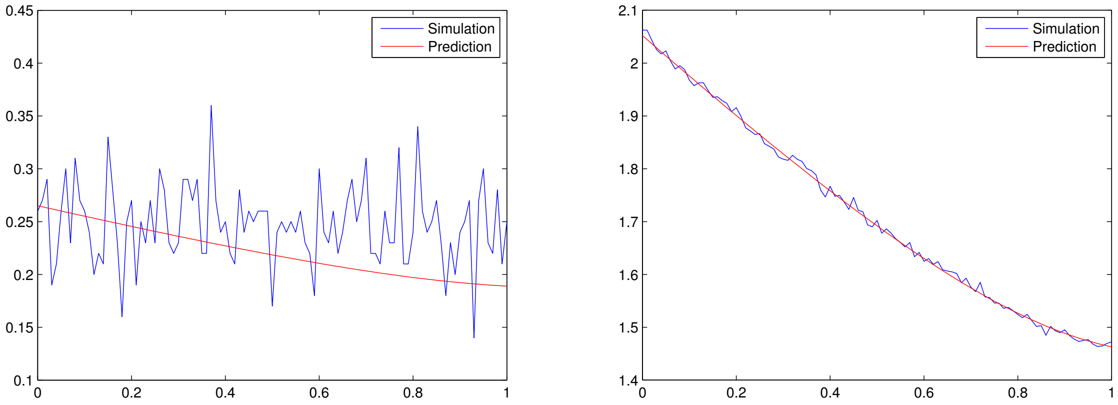
$$\begin{aligned} &2P_2\delta - P < 0 \\ \Longleftrightarrow &\delta < \frac{P}{2P_2} = \frac{1}{2} + \frac{P_1}{2P_2} \end{aligned}$$

So there might exist an interval such that $\rho_\delta(A) < 1$. Note that in the case of a symmetric adjacent matrix, we see that it is decreasing until $\delta = 1$. Let's compute this interval. First of all, if $\frac{Pn}{2} > 1$, then $\rho_\delta(A) > 1$.Otherwise, we compute the two critical values of the concentration coefficient, δ_{c1} and δ_{c2}, such that $\rho_\delta(A) < 1 \Longleftrightarrow \delta \in [\delta_{c1}; \delta_{c2}]$:

$$\begin{aligned} &\rho_\delta(A) = 1 \\ \Longleftrightarrow &\frac{Pn}{2} + \frac{\sqrt{\mathit{\Delta}(\delta)}}{2} = 1 \\ \Longleftrightarrow &\sqrt{\mathit{\Delta}(\delta)} = 2 - Pn \\ \Longleftrightarrow &\mathit{\Delta}(\delta) = 4 - 4Pn + P^2n^2 \\ \Longleftrightarrow &n^2P_2^2\delta^2 - P_2Pn^2\delta + P_{1,2}P_{2,1}n^2 + Pn - 1 = 0 \\ \Longleftrightarrow &\delta_{c1} = \frac{P}{2P_2} - \frac{1}{n}\sqrt{P^2n^2 - 4n^2P_{1,2}P_{2,1} - 4Pn + 4} \\ \text{and} \quad &\delta_{c2} = \frac{P}{2P_2} + \frac{1}{n}\sqrt{P^2n^2 - 4n^2P_{1,2}P_{2,1} - 4Pn + 4} \end{aligned}$$

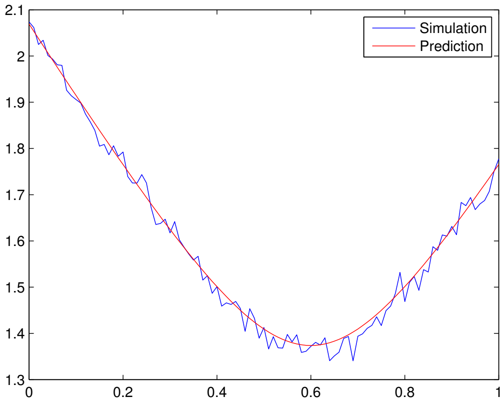
In fact, in the case where $P_1 > P_2$, which means that the vertices from the core get more easily linked to another vertex than a vertex from the periphery to get linked to another vertex, $\frac{P}{2P_2} > 1$. $\rho_\delta(A)$ is thus decreasing for $0 \leq \delta \leq 1$. The fact that the spectral radius decreases is explained in that way: when $\delta = 0$, the core is very dense and the disease will spread fast inside it and will reach periphery at the end. When δ increases, the core will become less dense, while the periphery will become dense, so we will have a population with lower density than the one of the core when $\delta = 0$, so the disease will spread less fast (if $P_1 > P_2$). In the case where $P_2 > P_1$, as the periphery becomes dense for a particular value of δ, the density of the periphery is large enough to allow a fast diffusion of the disease. This is the reason why the spectral radius increases again as shown on figure 3.7c.

Figures 3.7 show the evolution of the spectral radius with δ. We plot the theorical result we found in this chapter (prediction) in red and the result of numerical simulations in blue. These simulations were made by computing the mean of 100 spectral radius of random binary matrices of size 1000 (two communities of size 500).



(a) $P_1 = 0.0003$ $P_2 = 0.00024$
$P_{1,2} = 0.00015$ $P_{2,1} = 0.00012$

(b) $P_1 = 0.0022$ $P_2 = 0.00176$
$P_{1,2} = 0.0011$ $P_{2,1} = 0.00088$



(c) $P_2 = 0.0167$ $P_2 = 0.00334$
$P_{1,2} = 0.00167$ $P_{2,1} = 0.00835$

Figure 3.7: Comparison between the prediction and numerical simulation

As in the previous section, we see that our model is well defined for dense networks, but less for disconnected networks, and this for the same reasons.

3.5 Conclusion

We began this chapter by giving a definition of the global reproductive number that is determinant in the diffusion of diseases in non-homogeneous community. This number is in fact the spectral radius of the adjacent matrix of the network. Since the adjacent matrix of such network are binary random matrices, we make the calculation based on the average of these adjacent matrices. We thus gave two properties to the eigenvalues of such a matrix to help us in the calculation. The first property tells that most of the eigenvalues are equal to zero and the other is proportional to the dimension of the matrix.

We then studied the average case of a two communities model and we found an expression of their spectral radius. But simulations show us that when the probabilities of the entries of the average adjacent matrix are too small (the matrix is then disconnected), then the expression we find for the spectral radius is no longer valid (we will explain why in the last chapter). It is important to note that the calculations of this chapter were made under the hypothesis that the population is divided into two communities of exactly the same size and with all individuals the same inside it. We can then wonder what will happen in the case where the communities do not have same size. This again will be explained in the last chapter.

Chapter 4

Generalization to several communities

4.1 Introduction

In this chapter, we will extend what we discovered in the previous chapter. We will thus look at the behaviour of the global reproductive number (i.e. the spectral radius of the adjacent matrix) of directed graph with four communities of the same size. But in order to avoid a too large number of parameters, we will work on simple networks. Again we will work on the average adjacent matrix in order to get an expression of the spectral radius.

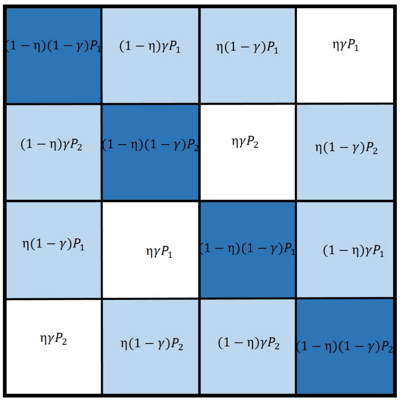
Here, we will generalize the case where we have four communities of the same size that are linked to each other, or in the case where they are linked to only another or two other communities. We then generalize the case where we can split a core-periphery into two smaller core-periphery. Finally, we will modelize a core-periphery where we can split each community into a core-periphery model.

The main goal of this chapter is to develop a tool to allow the calculation of the eigenvalues of the adjacent matrix of a more complex network. Again, we will not build a network but analyze a network. We start from a given network, then identify the communities and how they are linked, and finally compute the eigenvalues. We will limit our self in the case of 4 communities of the same size (number of vertices).

4.2 Four linked community

4.2.1 Description of the network

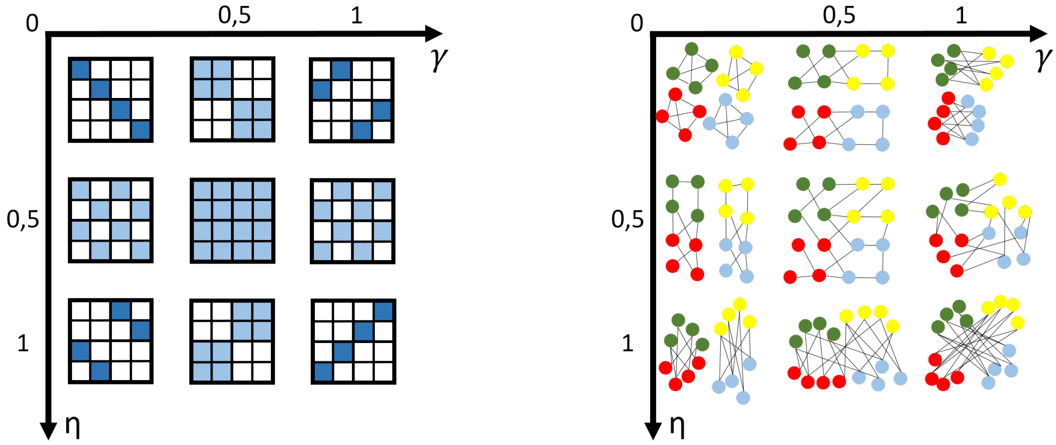
Here, we will look at a network with corresponding adjacent matrix on the form of figure 4.1. Such a network is formed by two communities linked with a parameter $0 \leq \eta \leq 1$. Both communities are each divided in two communities linked with parameter γ (linking coefficient).



$(1-\eta)(1-\gamma)P_1$	$(1-\eta)\gamma P_1$	$\eta(1-\gamma)P_1$	$\eta\gamma P_1$
$(1-\eta)\gamma P_2$	$(1-\eta)(1-\gamma)P_2$	$\eta\gamma P_2$	$\eta(1-\gamma)P_2$
$\eta(1-\gamma)P_1$	$\eta\gamma P_1$	$(1-\eta)(1-\gamma)P_1$	$(1-\eta)\gamma P_1$
$\eta\gamma P_2$	$\eta(1-\gamma)P_2$	$(1-\eta)\gamma P_2$	$(1-\eta)(1-\gamma)P_2$

Figure 4.1: General adjacent matrix of the four communities model

In fact, we will look at how R_* will evolve in the plan (γ, η), as shown in figure 4.2. We can already notice that in the cases $(\gamma, \eta) = (1, 0)$, $(\gamma, \eta) = (0, 1)$ and $(\gamma, \eta) = (1, 1)$, we have a bipartite graph, so we can expect that, in these cases, the spectral radius will be similar.



(a) Evolution of the adjacent matrix in the plan (γ, η)

(b) Corresponding network

Figure 4.2

4.2.2 Construction of the adjacent matrix and calculation of the eigenvalues

In order to find the adjacent matrix, we use the Kronecker product in the following way : let A the adjacent matrix of the network:

$$A = \begin{pmatrix} (1-\eta) & \eta \\ \eta & (1-\eta) \end{pmatrix} \otimes \begin{pmatrix} (1-\gamma)P_1\mathbb{1}_n & \gamma P_1\mathbb{1}_n \\ \gamma P_2\mathbb{1}_n & (1-\gamma)P_2\mathbb{1}_n \end{pmatrix} \tag{4.1}$$

Thus A is of dimension $4n$ and its $4n$ eigenvalues are the product of the eigenvalues of the two matrices. Remember that the eigenvalues λ of the second matrix in (4.1) are all equal zero except, two that are :

$$\begin{aligned}\lambda &= \frac{(1-\gamma)Pn}{2} \pm \frac{\sqrt{\Delta(\gamma)}}{2} \\ \Delta(\gamma) &= P^2n^2\gamma^2 - 2n^2\bar{P}^2\gamma + \bar{P}^2n^2 \\ P &= P_1 + P_2 \\ \bar{P} &= |P_1 - P_2|\end{aligned}$$

On the other hand, the eigenvalues of the first matrix in (4.1) are 1 and $1-2\eta$ $(\forall\eta)$. So the four nonzero eigenvalues of A, $\lambda(A)$ are :

$$\begin{aligned}\lambda_1(A) &= \frac{(1-\gamma)Pn}{2} + \frac{\sqrt{\Delta(\gamma)}}{2} \\ \lambda_2(A) &= \frac{(1-\gamma)Pn}{2} - \frac{\sqrt{\Delta(\gamma)}}{2} \\ \lambda_3(A) &= (1-2\eta)\left(\frac{(1-\gamma)Pn}{2} + \frac{\sqrt{\Delta(\gamma)}}{2}\right) \\ \lambda_4(A) &= (1-2\eta)\left(\frac{(1-\gamma)Pn}{2} - \frac{\sqrt{\Delta(\gamma)}}{2}\right)\end{aligned}$$

The spectral radius is then given by $\rho(A) = \max_i |\lambda_i(A)|$. Since $0 \leq \gamma \leq 1$, we know that $|\lambda_1(A)| \geq |\lambda_2(A)|$. Moreover, since $|1-2\eta| \leq 1$, $|\lambda_3(A)| \geq |\lambda_4(A)|$. Finally, $|\lambda_1(A)| \geq |\lambda_3(A)|$, so $\rho(A) = \lambda_1(A)$. We do notice that it remains independent of η . Remember however that in the case of two communities with a symmetric adjacent matrix, we saw that the spectral radius was independent of the linking coefficient. In this case, we can consider that we are also in the scenario of two communities with a symmetric adjacent matrix. It is, therefore, normal that the spectral radius does not depend on the value of η. Note that in the symmetric case, when all communities have the same density, the spectral radius will not depend on γ anymore, and will be equal to : $\rho(A) = P_1n$, where P_1 is the density of each of the four community, and n the size of said communities. The condition to avoid an epidemic thus remains:

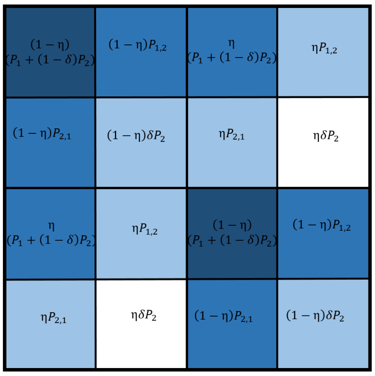
$$\begin{aligned}&\rho(A) \leq 1 \\ \iff &\gamma \geq \frac{1 - Pn + P_1P_2n^2}{2P_1P_2n^2 - Pn}\end{aligned}$$

this does not depends on η as announced.

4.3 Two linked core-periphery network

4.3.1 Description of the model

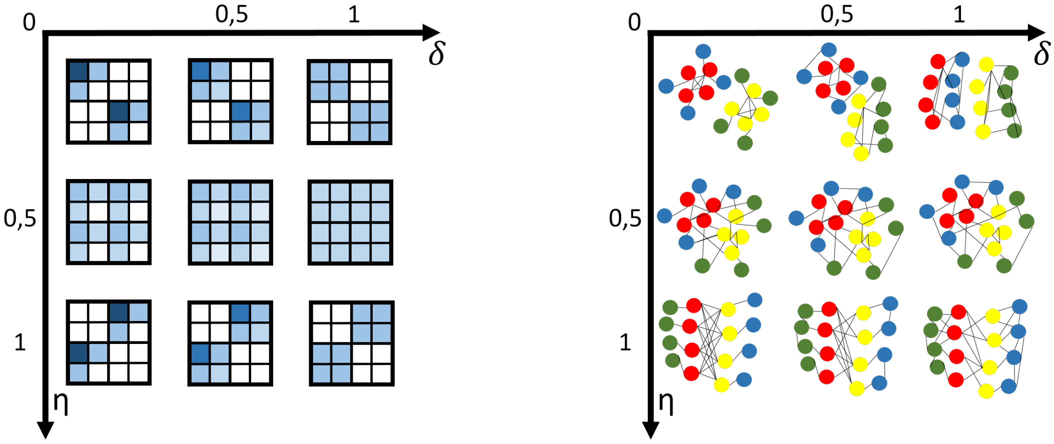
Here we will look at how the global reproductive number behaves when we consider two core-periphery networks that we link to each other. For example, we can consider each network as a hospital service. We then look at what the spectral radius will be, knowing the link between these two hospital services. So, as in chapter 3, we take the parameter δ defined as the proportion of edges that we take from the periphery to give it to the core. Then, the η parameter will qualify the link between the two core-periphery network. The size of the two cores are the same just like the size of the two periphery, in order to make our calculation clear. Again, we will reduce the number of parameters. Thus, the δ parameter will be the same for both communities. The adjacent matrix of such a matrix is given by figure 4.3.



$(1-\eta)$ $(P_1+(1-\delta)P_2)$	$(1-\eta)P_{1,2}$	η $(P_1+(1-\delta)P_2)$	$\eta P_{1,2}$
$(1-\eta)P_{2,1}$	$(1-\eta)\delta P_2$	$\eta P_{2,1}$	$\eta\delta P_2$
η $(P_1+(1-\delta)P_2)$	$\eta P_{1,2}$	$(1-\eta)$ $(P_1+(1-\delta)P_2)$	$(1-\eta)P_{1,2}$
$\eta P_{2,1}$	$\eta\delta P_2$	$(1-\eta)P_{2,1}$	$(1-\eta)\delta P_2$

Figure 4.3: General adjacent matrix of the two linked core-periphery model

The phase-plan is given by figure 4.6. The evolution along the δ axis was already explained in chapter 3. We just recall that the more δ is close to 1, the more homogeneous the network will be. Along the η axis, the two core-periphery system will progressively lose the connections inside them to be linked to the other core-periphery system. So the core of the first system will be linked to the core of the second system and its periphery, but not to the periphery and the core of the first system anymore.



(a) Evolution of the adjacent matrix in the plan (δ, η)

(b) Corresponding network

Figure 4.4

4.3.2 Calculation of the adjacent matrix and its eigenvalue

As in the previous section the adjacent matrix A is formed by the Kronecker product defined as :

$$A = \begin{pmatrix} (1-\eta) & \eta \\ \eta & (1-\eta) \end{pmatrix} \otimes \begin{pmatrix} (P_1+(1-\delta)P_2)\mathbb{1}_n & P_{1,2}\mathbb{1}_n \\ P_{2,1}\mathbb{1}_n & \delta P_2\mathbb{1}_n \end{pmatrix} \tag{4.2}$$

Recall that the eigenvalues λ of the second matrix in the product in equation 4.2 are all equal to 0 except two that are given by:

$$\begin{aligned}\lambda &= \frac{Pn}{2} \pm \frac{\sqrt{\Delta(\delta)}}{2} \\ \Delta(\delta) &= n^2(4P_2^2\delta^2 - 4P_2P\delta + 4P_{1,2}P_{2,1} + P^2) \\ P &= P_1 + P_2\end{aligned}$$

So the 4 nonzero eigenvalues of the adjacent matrix A, $\lambda(A)$ are :

$$\begin{aligned}\lambda_1(A) &= \frac{Pn}{2} + \frac{\sqrt{\Delta(\delta)}}{2} \\ \lambda_2(A) &= \frac{Pn}{2} - \frac{\sqrt{\Delta(\delta)}}{2} \\ \lambda_3(A) &= (1-2\eta)\left(\frac{Pn}{2} + \frac{\sqrt{\Delta(\delta)}}{2}\right) \\ \lambda_4(A) &= (1-2\eta)\left(\frac{Pn}{2} - \frac{\sqrt{\Delta(\delta)}}{2}\right)\end{aligned}$$

The spectral radius is given by $\rho(A) = \max_i |\lambda_i(A)|$. Using the same arguments as in the previous section, we find that $\rho(A) = \lambda_1(A)$. Here again, the spectral radius is independent of the η parameter. This is again the consequence of the fact that this model can be seen as a two-communities model where both communities have same density so the spectral radius is the size of a community times the density. The condition to avoid an epidemic (i.e. $\rho(A) < 1$) remains the same as for a core-periphery developed in chapter 3:

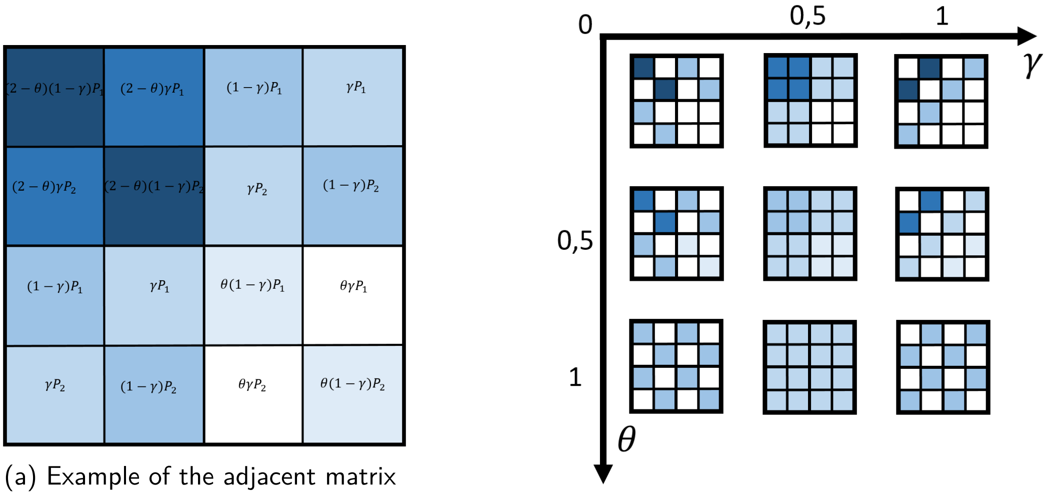
$$\begin{aligned}&\rho_\delta(A) \leq 1 \\ \iff &\delta \geq \frac{P}{2P_2} - \frac{1}{n}\sqrt{P^2n^2 - 4n^2P_{1,2}P_{2,1} - 4Pn + 4} \\ \text{and} \quad &\delta \leq \frac{P}{2P_2} + \frac{1}{n}\sqrt{P^2n^2 - 4n^2P_{1,2}P_{2,1} - 4Pn + 4}\end{aligned}$$

Again, in the case where $P_1 > P_2$, the spectral radius will be a strictly decreasing function of δ. This behaviour is explained in chapter 3. Since the two core-periphery have same total density and same parameter δ, the two cores will always have the same density and so will be the periphery then diffusion will be as fast from a core-periphery to the other as in the other way. For example, an infected person from one of the two core will infect another individual from the other core as easily as someone from the second core will infect an individual from the first core. Consequently, transmission from a core-periphery to the other will be independent of η. The diffusion will then be totally determined by the core-periphery structure. That is as more the core is dense, as faster will the diffusion be.

4.4 Core-periphery divided into two communities

4.4.1 Description of the model

We will now look at what happens in the case where we have a great core-periphery network that is divided into two communities of the same size. For example, let's consider that we have a core composed of two different kind of people (men-women, children-adult, ...), and so is the periphery too. We will work with an adjacent matrix in the form of figure 4.5a and the phase plan is given by figure 4.5b. The corresponding networks are present in figure 4.4b. Indeed, this case is similar to the previous case we studied. Indeed, if we invert the second and third community (i.e. invert the second line with the third one and the second column with the third one in figure 4.5a), the only difference is the parameterization.



(a) Example of the adjacent matrix

(b) Phase plan

Figure 4.5

In this section, we denote by θ the proportion of edges we take from communities 3 an 4 (between and inside them) to give it to communities 1 and 2 (between and inside them).

4.4.2 Calculation of the adjacent matrix and its eigenvalue

Again the adjacent matrix A is given by a Kronecker product, but this time the first factor of the kroncker product will be different since we want to apply a core-periphery model to a model with two linked communities, thus :

$$A = \begin{pmatrix} (2-\theta) & 1 \\ 1 & \theta \end{pmatrix} \otimes \begin{pmatrix} (1-\gamma)P_1\mathbb{1}_n & \gamma P_1\mathbb{1}_n \\ \gamma P_2\mathbb{1}_n & (1-\gamma)P_2\mathbb{1}_n \end{pmatrix} \tag{4.3}$$

The eigenvalues of the first factor in the Kronecker product (4.3) are $1+\sqrt{\theta^2-2\theta+2}$ and $1-\sqrt{\theta^2-2\theta+2}$. Since $\theta^2-2\theta+2>0$, these eigenvalues are always real. The 4 nonzero eigenvalues of A $\lambda(A)$ are given by ($\varDelta(\gamma) = P^2 n^2 \gamma^2 - 2n^2 \bar{P}^2 \gamma + \bar{P}^2 n^2$):

$$\begin{aligned}
\lambda_1(A) &= \left(1 + \sqrt{\theta^2 - 2\theta + 2}\right) \left(\frac{(1-\gamma)Pn}{2} + \frac{\sqrt{\varDelta(\gamma)}}{2}\right)\\
\lambda_2(A) &= \left(1 + \sqrt{\theta^2 - 2\theta + 2}\right) \left(\frac{(1-\gamma)Pn}{2} - \frac{\sqrt{\varDelta(\gamma)}}{2}\right)\\
\lambda_3(A) &= \left(1 - \sqrt{\theta^2 - 2\theta + 2}\right) \left(\frac{(1-\gamma)Pn}{2} + \frac{\sqrt{\varDelta(\gamma)}}{2}\right)\\
\lambda_4(A) &= \left(1 - \sqrt{\theta^2 - 2\theta + 2}\right) \left(\frac{(1-\gamma)Pn}{2} - \frac{\sqrt{\varDelta(\gamma)}}{2}\right)
\end{aligned}$$

The spectral radius $\rho(A)$ defined as the maximum eigenvalue in absolute value of A is then :

$$\rho(A) = \left(1 + \sqrt{\theta^2 - 2\theta + 2}\right) \left(\frac{(1-\gamma)Pn}{2} + \frac{\sqrt{\varDelta(\gamma)}}{2}\right) \tag{4.4}$$

Apart from both previous cases, R_* now depends on two parameters. We thus need to find the zone in which the spectral radius is smaller than 1. But, first, we will look at how the spectral radius evolves by computing its partial derivatives.

$$\begin{aligned}
\frac{\partial \rho}{\partial \gamma} &= \left(1 + \sqrt{\theta^2 - 2\theta + 2}\right) \frac{\partial}{\partial \gamma} \left(\frac{(1-\gamma)Pn}{2} + \frac{\sqrt{\varDelta(\gamma)}}{2}\right)\\
&\leq 0
\end{aligned}$$

Remember that we showed this partial derivative is always non-positive. It is equal to 0 in two communities with same density. We will further set $\varGamma(\gamma) = \frac{(1-\gamma)Pn}{2} + \frac{\sqrt{\varDelta(\gamma)}}{2}$ for simplicity.

$$\begin{aligned}
\frac{\partial \rho}{\partial \theta} &= \varGamma(\gamma) \frac{\partial}{\partial \theta} \left(1 + \sqrt{\theta^2 - 2\theta + 2}\right)\\
&= \varGamma(\gamma) \frac{\theta - 1}{\sqrt{\theta^2 - 2\theta + 2}}\\
&\leq 0 \qquad \iff \qquad \theta \leq 1
\end{aligned}$$

Therefore, in our range of θ values, the spectral radius decreases when θ increases. Therefore, $\rho(A)$ is a decreasing function of γ and θ, the zone of the phase plan (γ, θ), where $\rho(A) \leq 1$ is the zone above the cure $\rho(A) = 1$ (i.e. γ and θ are larger than their value on the curve). So let's find θ_c, the critical value of θ, that is the value of θ along curve $\rho(A) = 1$. Since this value depends on γ, we set : $\theta_c(\gamma)$.

$$\begin{aligned}
&\rho(A) = 1\\
\iff &\left(1 + \sqrt{\theta_c(\gamma)^2 - 2\theta_c(\gamma) + 2}\right) \varGamma(\gamma) = 1\\
\iff &\sqrt{\theta_c(\gamma)^2 - 2\theta_c(\gamma) + 2} = \frac{1 - \varGamma(\gamma)}{\varGamma(\gamma)}\\
\iff &\theta_c(\gamma) = 1 - \frac{\sqrt{1 - 2\varGamma(\gamma)}}{\varGamma(\gamma)}
\end{aligned}$$

By solving this, we find that we have actually two curves. Since the second curve corresponds to value of $\theta > 1$, we do not take this curve into account. Moreover, we see that the derivative of θ_c with respect to γ is non-negative:

$$\frac{d\theta_c}{d\gamma} = -\left(\frac{\Gamma(\gamma) - 1 + 2\Gamma(\gamma)}{\Gamma(\gamma)^2\sqrt{1-2\Gamma(\gamma)}}\right)\frac{d\Gamma}{d\gamma}$$

$$= \frac{-(\Gamma(\gamma) - 1)}{\Gamma(\gamma)^2\sqrt{1-2\Gamma(\gamma)}}\frac{d\Gamma}{d\gamma}$$

This is non-positive if $\Gamma(\gamma) \leq 1$. However, looking at the expression of $\theta_c(\gamma)$, we see it exists if and only if $\Gamma(\gamma) \leq 0.5$. This is due to the fact that $1 + \sqrt{\theta^2 - 2\theta + 2}$ reaches its minimum value when $\theta = 1$. Its value is then two, and so $\Gamma(\gamma)$ must be smaller than one half to have that $\rho(A) < 1$, and so $\theta_c(\gamma)$ is a decreasing function. Therefore the larger the value γ is, the more possible values of θ there are, such that R_* is smaller than one. In fact, the larger the value of γ is, the smaller the value of θ can be such that the global reproductive number is smaller than one. To see why, we need to remember that in a two linked communities model with two communities with different probabilities to transmit the disease, then the diffusion slows down. Indeed, if γ have value near to one, then inside the core an infected person form the first community, will easily infect someone from the second. But this second infected individual will less easily contaminate someone from the first community. Therefore, the diffusion process will slow down and the density of the core can become a bit larger. There is thus the impact of two effect:

- The community effect: in a population of two communities where each one is characterize by a different probability to infect someone. Then the more these communities are connected to each other, the smaller will the global reproductive number be.
- The core-periphery effect: in a population of two communities where one is more dense than the other, then the more dense this community is, the larger will the global reproductive number be.

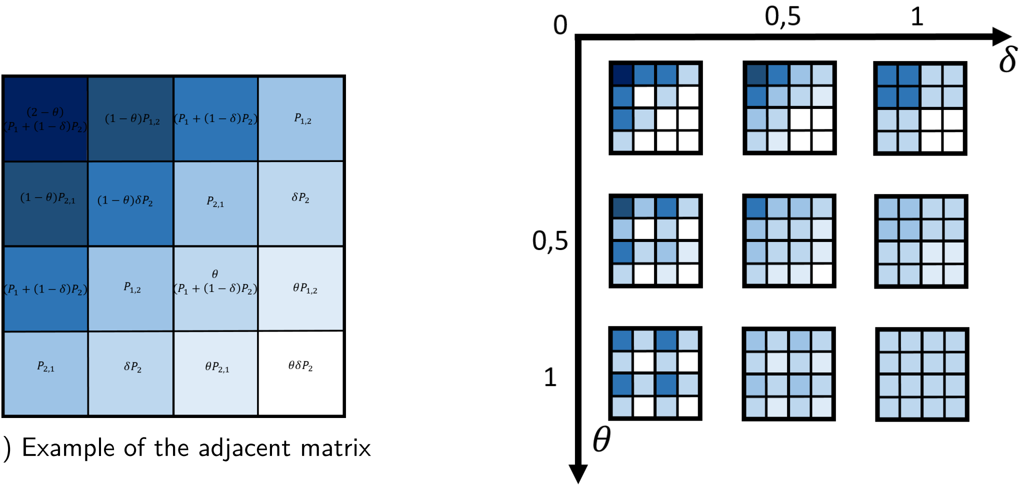
Note that in the symmetric case ($P_1 = P_2$), we find that $\Gamma(\gamma) = nP_1$, and so the curve we just found is a straight line independent of γ.

As explained in the description of this model, we do expect here the same results as in the two linked core-periphery network section. But it is not the case. In fact in the case where we have $P_1 = P_2$ in this section and $P_1 = P_2 = P_{1,2} = P_{2,1}$ in the previous section, we see that (taking back equation (3.4)) the two zones where $\rho(A) \leq 1$ are the same. That is to say that the only parameter that will influence the spectral radius is the core-periphery parameter. However, in the case of asymmetric adjacent matrix, the order in which we analyze the network will have an impact on the spectral radius.

4.5 Hierarchical core-periphery

4.5.1 Description of the model

The last case we will consider is a core-periphery where we can split the core into a smaller core-periphery network, and so can we with the periphery too. This allows us to introduce a hierarchy in the communities. That is to say four communities linked among each other with a given connection rate. Therefore, like in the previous section, we use the θ parameter with the same meaning and we use the δ parameter as in chapter three, with the same meaning too. The adjacent matrix is then on the form given in figure 4.6a and the corresponding phase plan is given by figure 4.6b.



$(2-\theta)$ $(P_1+(1-\delta)P_2)$	$(1-\theta)P_{1,2}$	$(P_1+(1-\delta)P_2)$	$P_{1,2}$
$(1-\theta)P_{2,1}$	$(1-\theta)\delta P_2$	$P_{2,1}$	δP_2
$(P_1+(1-\delta)P_2)$	$P_{1,2}$	θ $(P_1+(1-\delta)P_2)$	$\theta P_{1,2}$
$P_{2,1}$	δP_2	$\theta P_{2,1}$	$\theta\delta P_2$

(a) Example of the adjacent matrix

(b) Phase plan

Figure 4.6

4.5.2 Calculation of the adjacent matrix and its eigenvalues

The adjacent matrix A is given by this Kronecker product :

$$A = \begin{pmatrix} (2-\theta) & 1 \\ 1 & \theta \end{pmatrix} \otimes \begin{pmatrix} (P_1 + (1-\delta)P_2)\mathbb{1}_n & P_{1,2}\mathbb{1}_n \\ P_{2,1}\mathbb{1}_n & \delta P_2 \mathbb{1}_n \end{pmatrix} \tag{4.5}$$

Its four nonzero eigenvalues $\lambda(A)$ are then ($\Delta(\delta) = n^2(4P_2^2\delta^2 - 4P_2P\delta + 4P_{1,2}P_{2,1} + P^2)$):

$$\lambda_1(A) = \left(1 + \sqrt{\theta^2 - 2\theta + 2}\right)\left(\frac{Pn}{2} + \frac{\sqrt{\Delta(\delta)}}{2}\right)$$

$$\lambda_2(A) = \left(1 - \sqrt{\theta^2 - 2\theta + 2}\right)\left(\frac{Pn}{2} + \frac{\sqrt{\Delta(\delta)}}{2}\right)$$

$$\lambda_3(A) = \left(1 + \sqrt{\theta^2 - 2\theta + 2}\right)\left(\frac{Pn}{2} - \frac{\sqrt{\Delta(\delta)}}{2}\right)$$

$$\lambda_4(A) = \left(1 - \sqrt{\theta^2 - 2\theta + 2}\right)\left(\frac{Pn}{2} - \frac{\sqrt{\Delta(\delta)}}{2}\right)$$

The spectral radius is then $\rho(A) = \left(1 + \sqrt{\theta^2 - 2\theta + 2}\right)\left(\frac{Pn}{2} + \frac{\sqrt{\Delta(\delta)}}{2}\right)$. Let's now compute the partial derivatives of the spectral radius, by the definition of $\rho(A)$ and use what we found out in chapter three and in the previous section. This leads us to the following result:

$$\begin{aligned} \frac{\partial \rho}{\partial \delta} \leq 0 \qquad &\Longleftrightarrow \qquad \delta \leq \frac{P}{2P_2} \\ \frac{\partial \rho}{\partial \theta} \leq 0 \qquad &\Longleftrightarrow \qquad \theta \leq 1 \end{aligned}$$

Like in the previous section, we will compute$\theta_c(\delta)$, the critical value of θ, on the curve $\rho(A) = 1$. The area of the phase plan corresponding to $\rho(A) \leq 1$ is the area, if it exists, located between this curve and the axis $\theta = 1$. Let's pose $\mathit{\Gamma}(\delta) = \frac{Pn}{2} + \frac{\sqrt{\mathit{\Delta}(\delta)}}{2}$ for simplicity reasons, and we find as in the previous section that :

$$\begin{aligned} &\rho(A) = 1 \\ \Longleftrightarrow &\theta_c(\delta) = 1 - \frac{\sqrt{1 - 2\mathit{\Gamma}(\delta)}}{\mathit{\Gamma}(\delta)} \end{aligned}$$

What is definite only if $\mathit{\Gamma}(\delta) \leq 0.5$, which means that for a given value of the concentration coefficient, there might not exist any value of θ, such that the spectral radius is smaller than one. The disease will therefore diffuse in the network. This critical value of θ decreases with δ until $\delta = \frac{P}{2P_2}$ and then increases. Indeed,

$$\begin{aligned} \frac{d\theta_c}{d\delta} =& \frac{-(\mathit{\Gamma}(\delta) - 1)}{\mathit{\Gamma}(\delta)^2 \sqrt{1 - 2\mathit{\Gamma}(\delta)}} \frac{d\mathit{\Gamma}}{d\delta} \\ \frac{d\theta_c}{d\delta} \leq& 0 \qquad \Longleftrightarrow \qquad \delta \leq \frac{P}{2P_2} \end{aligned}$$

Note that in the case of symmetry, $P_1 = P_2$ and $P_{1,2} = P_{2,1}$, $\frac{P}{2P_2} = 1$, so $\theta_c(\delta)$ is a stricly increasing function, and is, by taylor series, nearly a straight line. But in the general case if $P_1 > P_2$ then $\frac{P}{2P_2} > 1$ and so θ_c is a stricly decreasing function. Indeed, if $\delta = 0$ then θ has to be as large as possible to avoid the creation of a very dense community linked to periphery totally disconected to theirself (see figure 4.6b). When δ increases, some communities will become less dense while other will become more dense, so the population will become more uniform. θ can then have values smaller since there will not be a community with very large density anymore. If $P_2 > P_1$ the same holds but we have to note that in this case, for value of $\delta > \frac{P}{2P_2}$ there will be a community with increasing density again, so θ cannot take to low value anymore.

4.6 General cases

In the general case, we can also use the Kronecker product to compute the global reproductive number. Indeed, it can be used on models with as many communities we want. For example, if we want a model with three interlinked communities, with density a_1, a_2 and a_3 , let α_{ij} the fraction of contact between community i and j, we have to compute the spectral radius of:

$$\begin{pmatrix} a_1\alpha_{11} & a_1\alpha_{12} & a_1\alpha_{13} \\ a_2\alpha_{21} & a_2\alpha_{22} & a_2\alpha_{23} \\ a_3\alpha_{31} & a_3\alpha_{32} & a_3\alpha_{33} \end{pmatrix} \otimes \mathbb{1}_n$$

Moreover, we can manage communities with different sizes. Indeed, if we notice that, in a network, we have a community twice as large as the other one and that they are linked, then the adjacent matrix is:

$$\begin{pmatrix} P_1(1-\gamma) & P_1\gamma & P_1\gamma \\ P_2\gamma & P_2(1-\gamma) & P_2(1-\gamma) \\ P_2\gamma & P_2(1-\gamma) & P_2(1-\gamma) \end{pmatrix} \otimes \mathbb{1}_n$$

Finally, we can also mix the model by adding factors in the kronceker product. But we have to do it in the right order, since it is not commutative.

Therefore, the Kronecker product is a useful tool to compute eigenvalues of a more complex network. But this approach presents the following inconvenient:

- ◇ We must have a great knowledge of the network in order to identify its structure. This is a general difficulty for the study of network.
- ◇ We must identify the different communities and the kind of relationship among them, and then identify the value of the parameter (γ, δ, η, θ,...).
- ◇ We have to formulate the model the right way. For example, if we have ten Kronecker products, we have to do it in the correct order.
- ◇ We don't know if it is always possible to formulate the model.
- ◇ The number of eigenvalues can be large, in particular in the case of communities of different size (for example two communities of size 500 and 501). This is also true when the number of Kronecker products increases. Moreover, the number of parameters can become very large.
- ◇ Depending on the accuracy we want - more accurate models containing small communities -, we will have a more complex model. So when should we stop?

4.7 Conclusion

In this chapter, we developed a tool allowing us to compute the global reproductive number. This tool has the great advantage that it is easy to use. Indeed, in order to find R_*, we have to compute the maximal eigenvalue, in absolute terms, which is the product of the eigenvalues of a large matrix and those of a smaller matrix. The eigenvalues are equal to zero except four of them. But, as shown in sections 4.3 and 4.4, this model is dependent of the way of interpreting the network. Indeed, if we analyze a network as the result of two interlinked core-periphery or as a core-periphery divided into two communities, this tool might give us two different results.

We analyze four cases of four communities of same size behaving differently:

- ◇ The case where the four communities are linked to each other such that increasing the linking coefficient between them will make these communities less dense
- ◇ The case where we have two linked core-periphery network
- ◇ A core periphery divided into two communities
- ◇ A hierarchical core periphery. A core-periphery where the core is also a core-periphery and so is the periphery.

For each of the four examples, we compute the spectral radius of the adjacent matrix and we define the value of the parameters such that it is smaller than one. The spectral radius were different from a case to the other. This brings to light a major difficulty of the modelization. Indeed, in real life, the kind of network is not necessarily obvious, so we need to identify the structure of the network and a mistake will lead to wrong global reproductive number. For example, once the four are identified, and we conclude to a two core-periphery divided into communities while it is a hierarchical core-periphery network. To avoid an epidemic outbreak we will thus try making the population the more bipartite possible, for example by avoiding as more as possible contact between the same communities. While we should avoid to closed communities. However we can notice that the more homogeneous a population is, the smaller will the global reproductive number be. The problem is then : how to make a population homogeneous? This again depend on the kind of population we have :

- ◇ In a four linked community, it is by having the more contact possible from people of other communities.
- ◇ In a two linked core-periphery network, it is by having more contact in the peripheries and less in the cores
- ◇ In a core periphery divided into two communities, it is by increasing the number of contact in the peripheries and decreasing in the cores. In addition increasing contact between the two cores and the two peripheries.
- ◇ In a hierarchical core-periphery, it is by making the communities having the same density.

Finally, as in the previous chapter, the models found in this chapter are built thanks to the average case, dense network. But, like in chapter three, if the network is not dense, then the models are no longer valid. Moreover, sparse networks are very common for large networks.

Chapter 5

Statistical study of random networks

5.1 Introduction

Up to now, we have studied the average case of random networks, which means : let a random network with n vertices. Each edge in this network appears with a given probability, so let A_ω a realization of this network, and A_ω is symmetric. The average network is then given by $A = \mathbb{E}[A_\omega]$. In the previous chapter, calculations were made on this average matrix. So in the present chapter, we will look at what happens in the case of a realization A_ω. More specifically, we want to find out a bound on the spectral radius of A_ω, $\rho(A_\omega)$.

In what follows, we will use a generalization of the Erdos-Rényi random graphs. We will first look at the distribution of the eigenvalues of random network and use an important result known as the Wigner semicircle law. Then we will look at the distribution of the spectral radius of random matrices. We will the give three theorems to find the spectral radius of random matrices We will illustrate these theorems with the networks from chapter 3 and see why the predictions does not work for too low density.

5.2 Case of a single community

5.2.1 Erdos-Rényi graphs and Wigner semicircle law

Erdos-Rényi random graphs are graphs of n vertices in which an edge is built between two vertices with a certain probability independently of the two vertices [4]. Such a network thus owns an adjacent matrix that is symmetric ($A_\omega^T = A_\omega$) and has average adjacent matrix $A = p\mathbb{1}_n$. In a particular realization of an Erdos-Rényi, the adjacent matrix A_ω has eigenvalues which follow a given distribution. In view of finding that distribution, we will use an important result known as the Wigner semicircle law.

First, let's define a Wigner matrix. A Wigner matrix is a symmetric (hermitian in complex cases) random matrix whose entries are independent (except for the symmetry constraint), have same distribution and have zero mean [5]. Let also suppose that they have unit variance. If W_n is a Wigner matrix of size n, and if $X_n = \frac{1}{\sqrt{n}} W_n$, and the eigenvalues of matrix X_n, λ_j $j = 1, ..., n$ and the normalized eigenvalue counting function (NCF) [5] :

$$\sigma_{X_n}(x) = \frac{1}{N}|\{\lambda_j \leq x\}|$$

Then:

$$\lim_{n \to \infty} \sigma_{X_n}(x) = \sigma(x)$$

where $\sigma(x)$ is called the Wigner semicircle law and is equal to:

$$\sigma(x) = \begin{cases} \sqrt{4 - x^2} & \text{if} \quad |x| \leq 2 \\ 0 & \text{else} \end{cases}$$

Such a result is shown in figure 5.1, where the red curve is function $\sigma(x)$. To design this figure, we compute the eigenvalues of ten random symmetric matrices of size 1000, where each entry follows a Bernoulli distribution of parameter $p = 0.01$. We then normalized these matrices. We see indeed that the eigenvalues follow the semicircle law as predicted by the previous property.

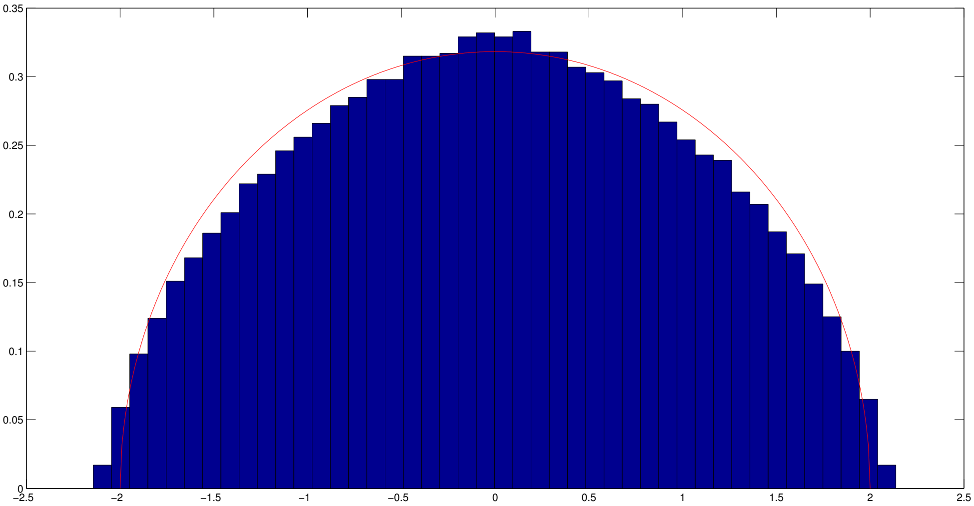


Figure 5.1: Numerical simulation made with ten matrices of 1000 vertices

5.2.2 Eigenvalues distribution in General Erdos-Rényi graphs

In this section, we will look at how the eigenvalues of the adjacent matrix A_ω of an Erdos-Rényi graph are distributed. We have already seen that the eigenvalues of the average matrix $A = \mathbb{1}_n$ are all equal to zero, except one that is equal to pn. Thus, there is a dominant eigenvalue linked to n which is much larger than the other one. In fact, such a result also appears in a random network as we can see on figure 5.2. To build this histogram we have simulated a random matrix and we have then computed its eigenvalues. Therefore, it can be interesting to know how this distribution will evolve, in particular concerning the largest eigenvalue (i.e.the spectral radius of this matrix).

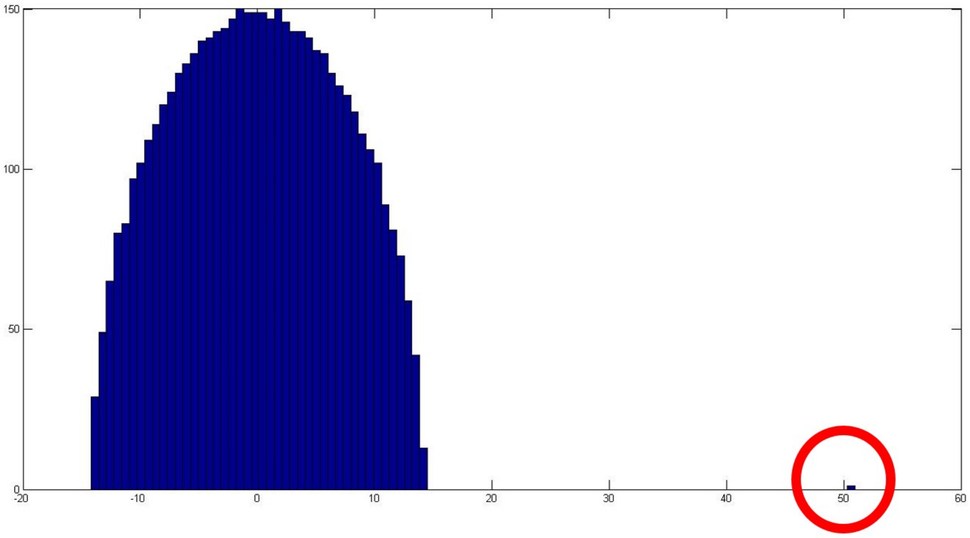


Figure 5.2: Distribution of the eigenvalues of a random matrix with $n = 5000$ and $p = 0.01$

Figure 5.2 shows the distribution of the eigenvalues of a matrix of size $n = 5000$ and $p = 0.01$. We see that most of the eigenvalues are concentrate around the value zero and have the form of a semicircle. But there is also a large eigenvalue in the red circle, which is much larger than the other. In the classical case of random matrices, this larger eigenvalue is equal to 50 and we see that, in this particular case, this largest eigenvalue is a bit larger than 50 (in fact this largest eigenvalue is equal to 50.95 so near to the average case). On Figure 5.3, we see that for random symmetric matrix with parameter $n = 500$ and $p = 0.01$, the spectral radius seems to follow a normal distribution. Figure 5.3 is drawn by making a simulation with 5000 matrices of size 5000. The empirical variance of the spectral radius is 0.02 while the empirical mean is 51.01. This way, the spectral radius of a random graph seems to be a bit larger than the one of the mean case.

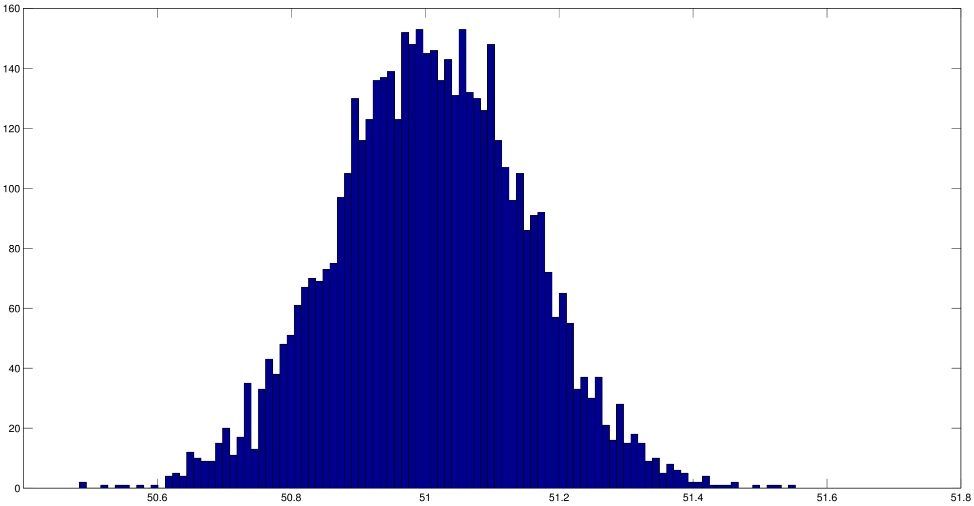
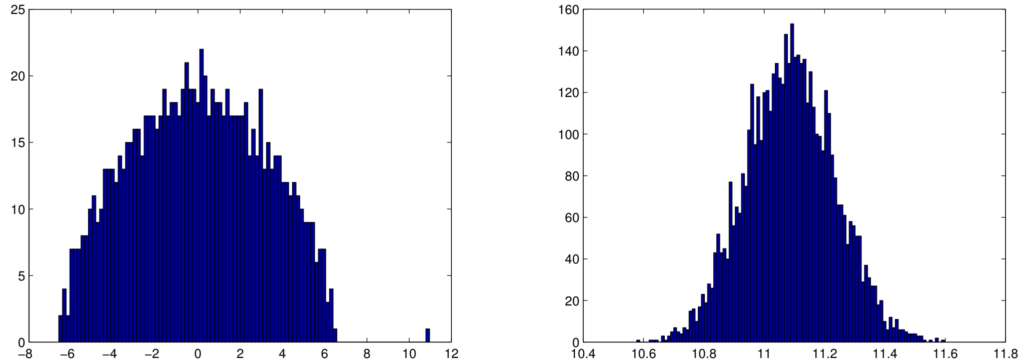


Figure 5.3: calculation of the spectral radius of 5000 matrices of size 5000

We can wonder about the validity limit of these simulations. In fact, as shown in figures 5.4, 5.5

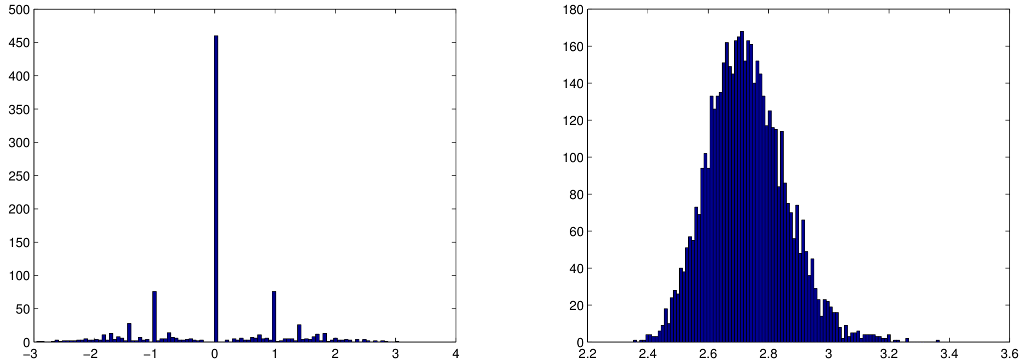
and 5.6, we see that for some value of the n and p parameters, we do not observe the semicircle law for the eigenvalues distribution anymore, while for other value of these parameters, the spectral radius does not seem to follow the normal distribution anymore. It is quite intuitive, since at its smallest, it is p, and at its most, there are entries of the adjacent matrix that are equal to zero. These simulations were made with 5000 matrices of size $n = 1000$. We then compute the 5000 spectral radius and draw the histograms.



(a) Distribution of the eigenvaluess (b) Distribution of the spectral radius

Figure 5.4: Simulation with $n = 1000$ and $p = 0.01$

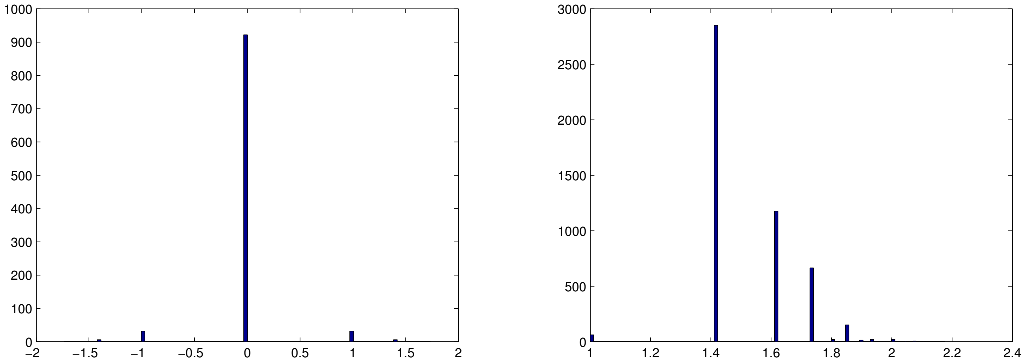
Figure 5.4a shows the distribution of the eigenvalues of a random network of $n = 1000$ vertices where each edge appears with a probability $p = 0.01$. Figure 5.4b shows the distribution of the spectral radius of such a network. Theorically, the spectral radius is equal to ten. On figure 5.4a, we see that the largest eigenvalues is a bit larger than ten (exactly 10.97). On figure 5.4b, we see that the distribution of the spectral radius is normal too, the empirical variance is 0.02 and the empirical mean is 11.09 (thus larger than ten). So we can suppose that the spectral radius is well of order np and even a bit larger.



(a) Distribution of the eigenvaluess (b) Distribution of the spectral radius

Figure 5.5: Simulation with $n = 1000$ and $p = 0.001$

Figure 5.5 shows the result of the simulation with a matrix of size $n = 1000$ and $p = 0.001$. The theorical spectral radius is given by $np = 1$, but this also means that we are in the critical case where there might be vertices that are isolated, which imply that the adjacent matrix has a line just with zeros. Indeed figure 5.5a shows that there are much more eigenvalues equal to zero, moreover in this particular case the spectral radius is near to the other eigenvalues (it is equal to 3.06 in this case). Finally, we see that this distribution does not look like the Wigner circle anymore. Figure 5.5b shows the distribution of the spectral radius still follows a normal distribution of empirical mean and variance respectively 2.73 and 0.02. Thus, in this case the spectral radius is much larger than the theorical spectral radius. Note also that in some cases, the maximum eigenvalue (in magnitude) is negative.



(a) Distribution of the eigenvaluess

(b) Distribution of the spectral radius

Figure 5.6: Simulation with $n = 1000$ and $p = 0.0001$

Here, in figure 5.6, we are in the case $np = 0.1 < 1$. We can thus expect that the number of eigenvalues equal to zero will be very large, since there are on average 9 vertices on 10 that are disconnected. Such a behaviour is well represented in figure 5.6a. The maximum eigenvalue is, in this case, one. While the distribution of the spectral radius is not normal anymore but looks like the one present in figure 5.6b where the empirical mean and variance are respectively 1.52 and 0.02. Thus, here the empirical spectral radius is approximately equal to ten times the prediction.

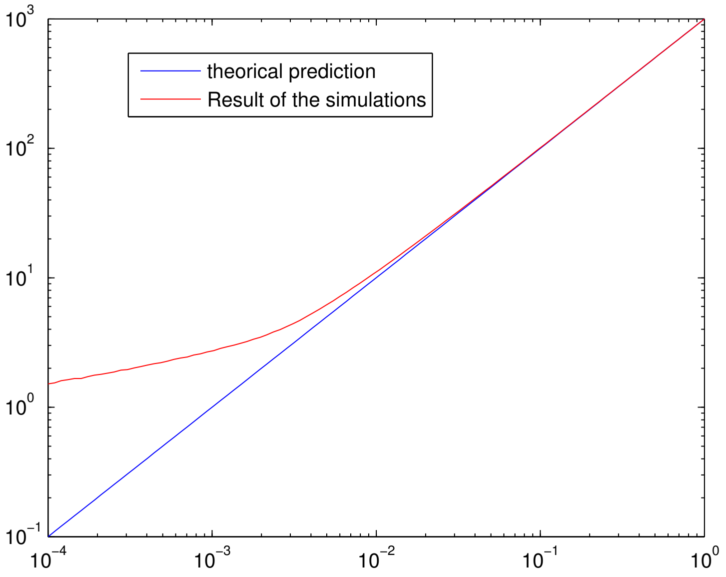


Figure 5.7: spectral radius versus the probability for matrx of size n

So it seems that, as long as the network is in average fully connected, the empirical spectral radius fits less or more to the theorical predictions, while if the graph is disconnected, the theorical prediction is not true anymore. An important result is that $\rho(A_\omega)$ Follow np for large value of p as expected from chapter three. This is illustrated in figure 5.7 . Therefore, in the study of diffusion, it seems that when the disease should theorically spread, then it indeed spreads. But when it should not spread, it still spreads, if we only look at the spectral radius criterion. But remember that if $np < 1$, the graph is disconnected. So even if the spectral radius is larger than one, then the disease will not reach all vertices, but it will well spread in the smallest community.

5.2.3 Bounds of the spectral radius

Like described in the previous section, the spectral radius is a random variable, but we would like to find some bounds to it. Using the Perron Frobenius theorem, we know that the spectral radius $\rho(A_\omega)$ satisfies the following property:

$$e \leq \rho(A_\omega) \leq E$$

where e and E are respectively the minimum and the maximum of the sum of all elements for each line. That is:

$$e = \min_i \sum_{j=1}^{n} (A_\omega)_{ij}$$
$$E = \max_i \sum_{j=1}^{n} (A_\omega)_{ij}$$

Therefore, we find that for every A_ω: $0 \leq \rho(A_\omega) \leq n$. Nevertheless, even if these bounds are deterministic, they are very large. So we can make the following changes:

$$d_{min} \leq \rho(A_\omega) \leq d_{max}$$

where d_{min} and d_{max} are respectively the minimum degree and the maximum degree present in the graph. It is important to notice that they are random variables too. Their distribution will thus depend on the distribution of the degrees. We first notice that since $(A_\omega)_{ij} \sim$ Be(p), the degree of the i^{th} vertex d_i defined in equation (5.1) follows a binomial distribution Bi(n, p).

$$d_i = \sum_{j=1}^{n} (A_\omega)_{ij} \tag{5.1}$$

Let's now compute the distribution of d_{max} and d_{min}:

$$\begin{aligned} P(d_{max} \leq k) &= P(\max_i d_i \leq k) \\ &= P(d_1 \leq k, ..., d_n \leq k) \\ &= [P(d_1 \leq k)]^n \end{aligned}$$

$$\begin{aligned} P(d_{min} \leq k) &= P(\min_i d_i \leq k) \\ &= 1 - P(d_1 \geq k, ..., d_n \geq k) \\ &= 1 - [P(d_1 \geq k)]^n \\ &= 1 - [1 - P(d_1 \leq k)]^n \end{aligned}$$

Where we used the fact that all degree are independent from each other. So we find that the upper and lower bound of the spectral radius are random variable that have a specific distribution. We will not go further in the calculation of these distribution for the reason that there are no explicit result for $P(d_1 \leq k)$. It is important to notice the following :

- ◇ In the asymmptotical case where $n \to \infty$ the distribution of the degree becomes normal with mean np and variance $np(1-p)$. That is a result from the central limit theorem.
- ◇ Some real-life graph exhibits a power law distribution in degree, which means that the number of vertices of degree d is proportional to $d^{-\beta}$ with $\beta > 1$ (in general $2 < \beta < 3$). This fact does not change the calculation of the distribution of d_{max} and d_{min}, since it is a general result [6].
- ◇ We can improve the lower bound by taking d_{mean}, which is the average degree of the vertices [2]. Indeed :

$$\begin{aligned} \rho(A_\omega) &= \max_v \frac{v^T A_\omega v}{v^T v} \\ &\geq \frac{1^T A_\omega 1}{1^T 1} \\ &= \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} (A_\omega)_{ij} \\ &= \frac{1}{n} \sum_{i=1}^{n} d_i \\ &= d_{mean} \end{aligned}$$

5.3 Case with several communities

Here, we consider that the network has several community like in chapter 3 and 4. Recall that in these sections we did not impose the symmetry of the adjacent matrix, so the probabilities P_1 or P_2 where the probability to get one in a specific entry of the adjacent matrix. But like in the previous section, we will impose symmetry of the adjacent matrix. Here is the general approach we will use to such a matrix. We describe here the general approach to compute the spectral radius. This is taken from the literature [6]

We begin to define the w vector of size n, where the i^{th} component w_i is the expected degree of node i. The probability to get an edge between node i and j is then given by equation (5.2) in the previous section, we will modify this approach but the main idea remain the same.

$$p_{ij} = p_{ji} = w_i w_j \xi \tag{5.2}$$

Where ξ is such that $\xi^{-1} = \sum_{l=1}^{n} w_l$. Note that:

$$\sum_{i=1}^{n}\sum_{j=1}^{n} p_{ij} = \frac{1}{\xi} \tag{5.3}$$

In view of getting $p_{ij} < 1$, we have to impose that $\forall i, w_i^2 < \sum_{j=1}^{n} w_j$. So we define the class of random graph $G(w)$ drawn from the w vector. We will give some theorems that allow us to have an idea of the value of the largest eigenvalue, but first we will define second order average degree $\tilde{d}$

$$\tilde{d} = \frac{\sum_{i=1}^{n} w_i^2}{\sum_{i=1}^{n} w_i}$$

These theorems are found in the literature [6]. They are given without any proof. They give an approximation of the largest eigenvalue of the random graph.

Theorem 1 [6]

For a graph G on $G(w)$, let its adjacent matrix A_ω, if the maximum degree d_{max} and the second order average degree $\tilde{d}$ satisfy: $\tilde{d} > \sqrt{d_{max}}\log(n)$ then the spectral radius is given by:

$$\rho(A_\omega) = (1 + o(1))\tilde{d}$$

Theorem 2 [6]

For a graph G on $G(w)$, let its adjacent matrix A_ω, if the maximum degree d_{max} and the second order average degree $\tilde{d}$ satisfy: $\sqrt{d_{max}} > \tilde{d}\log^2(n)$ then the spectral radius is given by:

$$\rho(A_\omega) = (1 + o(1))\sqrt{d_{max}}$$

We can thus expect that $\rho(A_\omega) = (1+o(1))\max(\sqrt{d_{max}}, \tilde{d})$. But this is in general not true. However, we will see on numerical simulations that the spectral radius behaves nearly like this. Note that $\log$ denotes the 10-base logarithm. Concerning the case of the Erdos-Rényi graph, we have that :

$$w_i = \mathbb{E}[d_i] = np$$
$$\tilde{d} = \frac{\sum_{i=1}^{n} w_i^2}{\sum_{i=1}^{n} w_i} = \frac{n^3p^2}{n^2p} = np$$

Finally we use a last theorem called the May-Wigner stability theorem for connected matrices. It gives a condition on the entries of the adjacent matrix so that it is connected, this last theorem is also given without proof and is found here [10].

Theorem 3 [10]

Let A be the adjacent matrix of a network of size n, let $0 < C < 1$ and $\epsilon > 0$. If the matrix owns Cn^2 nonzero entries then the corresponding network is :

- ◇ almost surely connected if :

$$C \geq (1+\epsilon)\frac{\log(n)}{n}$$

- ◇ almost surely disconnected if :

$$C \leq (1-\epsilon)\frac{\log(n)}{n}$$

In the case of an Erdos-Rényi random graph, we have that $C = p$. In the general case, we have that, in average:

$$\sum_{j=1}^{n}\sum_{i=1}^{n} p_{ij} = Cn^2 = \frac{1}{\xi}$$
$$\Longleftrightarrow C = \frac{1}{n^2\xi}$$

Consequently, the network will be almost surely disconnected if $\xi \geq \dfrac{1}{(1-\epsilon)n\log(n)}$.

We plot on figure 5.8 the result on the theorem 1 and 2. We make a simulation of 100 samples of matrices of dimension $n = 1000$ and varying p. For each value of p we computed the empirical mean of the spectral radius and the empirical mean of the maximum degree. We can notice that theorem 1 and 2 gives a good approximation for the spectral radius (red curve) for small value of p when the cyan curve is under the green curve for $p \leq 1.918 \times 10^{-4}$. But also for large value of p when the black curve is under the blue curve for $p \geq 0.01831$. Finally, theorem 3 tells us that if $p < \dfrac{\log(n)}{n} = 0.003$ then the graph is almost surely disconnected. We see that for this value the spectral radius is larger than 1. Therefore, even if for small value of p the spectral radius tends to 1 we have to remember that in this case the disease diffusion process will not be able to reach the whole network.

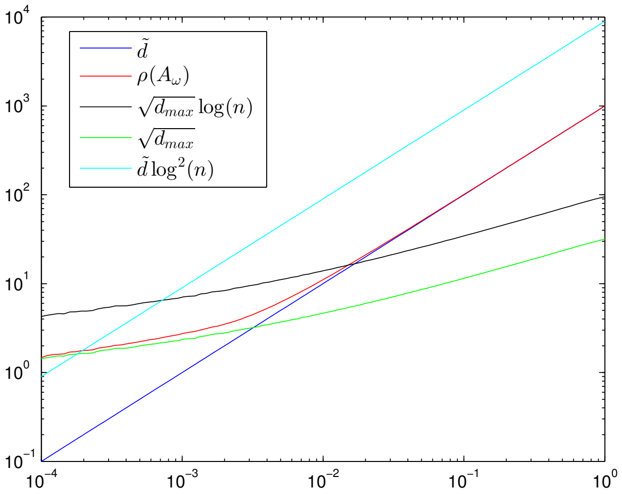


Figure 5.8: Illustration of theorem 1 and 2 on a simulation of 100 matrices of size $n = 1000$ with a varying p

5.3.1 Case of two communities

Here we will generalize the study of R_* (i.e. the spectral radius) for the case where we have two distinct communities in the network. Unlike what was done in the second chapter, we add the constraint of symmetry on the adjacent matrix. Moreover, we can look at what happens in the case where the two communities have different sizes. Thus, let's n_1 the size of the first community and n_2 the size of the second community and $n = n_1 + n_2$. The approach used here is that we consider two Erdos-Rényi graphs that we will progressively link to each other while we remove connections inside the two communities thanks to a parameter γ, the same used in chapter 3. Since we add a symmetry constraint of the adjacent matrix, the two networks must have the same parameter p (probability to get an edge between two vertices). Indeed, since every edge leaving community 1 must go to community 2. Let c_i $i = 1, 2$ the set of vertices in community i, for counter example, let's suppose that the two networks have different parameter $p_1 and p_2$, then (remind the form of the adjacent matrix of section 3.3, figure 3.1):

$$\begin{aligned}
\sum_{i\in c_1}\sum_{j\in c_2}\mathbb{E}[A_{ij}] &= \sum_{i\in c_2}\sum_{j\in c_1}\mathbb{E}[A_{ij}] \\
\sum_{i\in c_1}\sum_{j\in c_2}\gamma p_1 &= \sum_{i\in c_2}\sum_{j\in c_1}\gamma p_2 \\
n_1 n_2 \gamma p_1 &= n_2 n_1 \gamma p_2 \\
p_1 &= p_2 = p
\end{aligned}$$

The expected degree of vertex i belonging to community $j = 1, 2$ is given by w_{c_j} so we find for it:

$$\begin{aligned}
w_{c_j} &= \sum_{k\in c_j} p(1-\gamma) + \sum_{k\notin c_j} p\gamma \\
&= p(n_j(1-\gamma) + n_l\gamma) \qquad\qquad l \neq j
\end{aligned}$$

. Note that the expected degree of a node depends on the parameter γ that is logical. Indeed, let's consider for example a graph with ten vertices on a side and one on the other side. So the expected degree of a vertex from the first side is p while in the case where there is no connection between the two communities then the expected degree of the first side is $10p$. Let the probability p_{ij} to get an edge between vertex i and j. We modify equation (5.2) by adding the parameter γ to make the probability varying as we want.

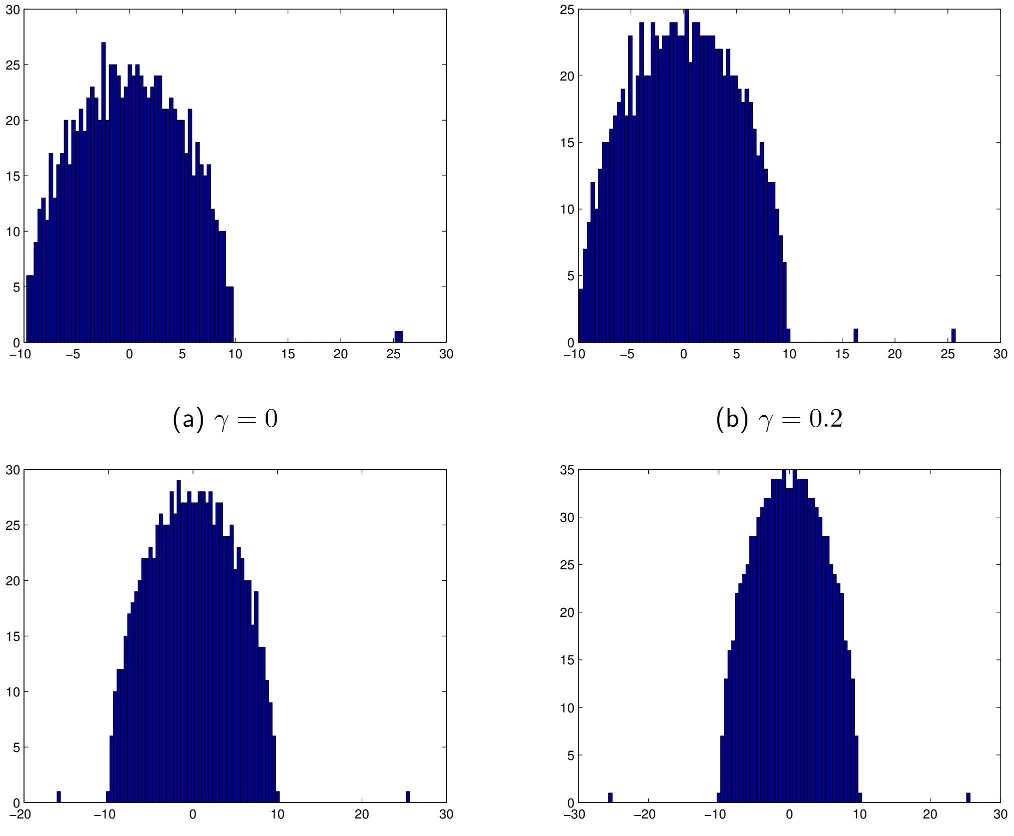
$$p_{ij} = \begin{cases} w_{c_1}^2 \xi(1-\gamma) & \text{if} \quad i, j \in c_1 \\ w_{c_2}^2 \xi(1-\gamma) & \text{if} \quad i, j \in c_2 \\ w_{c_1} w_{c_2} \xi\gamma & \text{if} \quad i \in c_1 \quad j \in c_2 \end{cases} \tag{5.4}$$

We must now find the value of ξ that may depend on γ. Remember that by definition, $w_i = \sum_{j=1}^{n} p_i j$.

So we have:

$$\begin{aligned}
\frac{1}{\xi} &= \sum_{i \in c_1} w_i + \sum_{i \in c_2} w_i \\
&= n_1 w_{c_1} + n_2 w_{c_2} \\
&= n_1 \sum_{j \in c_1} w_{c_1}^2 \xi (1-\gamma) + n_1 \sum_{j \in c_2} w_{c_1} w_{c_2} \xi \gamma + n_2 \sum_{j \in c_2} w_{c_2}^2 \xi (1-\gamma) + n_2 \sum_{j \in c_1} w_{c_1} w_{c_2} \xi \gamma \\
&= \left(n_1^2 w_{c_1}^2 + n_2^2 w_{c_2}^2 - \gamma (n_1 w_{c_1} - n_2 w_{c_2})^2 \right) \xi \\
\Longleftrightarrow \frac{1}{\xi} &= \sqrt{n_1^2 w_{c_1}^2 + n_2^2 w_{c_2}^2 - \gamma (n_1 w_{c_1} - n_2 w_{c_2})^2}
\end{aligned}$$

Let's first look at what happens when the two communities have the same size. As expected in the third chapter, since we are in the case of two communities with the same density, the global reproductive number will be independent of the linking coefficient γ. By looking at the distribution of the eigenvalues obtained with a random matrix with two communities of 500 people and probability $p = 0.05$, we can see on figure 5.9 that the largest eigenvalue remains indeed constant and is equal to 25.67 (We found that theorically, the spectral radius should be equal to $0.05 \times 500 = 25$.

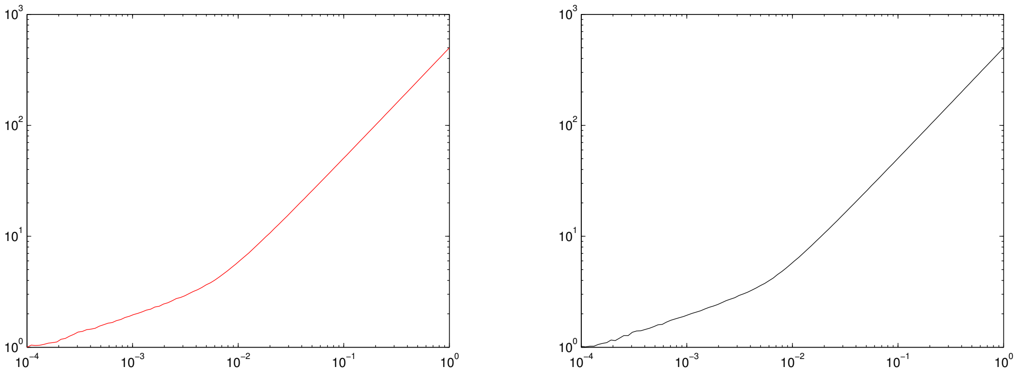


(a) $\gamma = 0$ (b) $\gamma = 0.2$

(c) $\gamma = 0.8$ (d) $\gamma = 1$

Figure 5.9: Simulation with $n = 500$ and $p = 0.05$

However, we can see that γ still has an influence on the distribution of the eigenvalues, in particular in the second larger eigenvalue. Remember that theorically the second larger eigenvalue of such a graph, in the symmetric case, is $P_1 n(1 - 2\gamma)$ so it is a decreasing function γ. Such a behaviour is well present in figures 5.9.Like in the case of a single community we note that for small value of p the spectral radius does not behave like the average case anymore. So we can wonder if for small value of p the spectral radius will depend on the parameter γ. In fact, simulations show that even for small value of p the spectral radius does not depend on the linking coefficient as shown in figure 5.10.

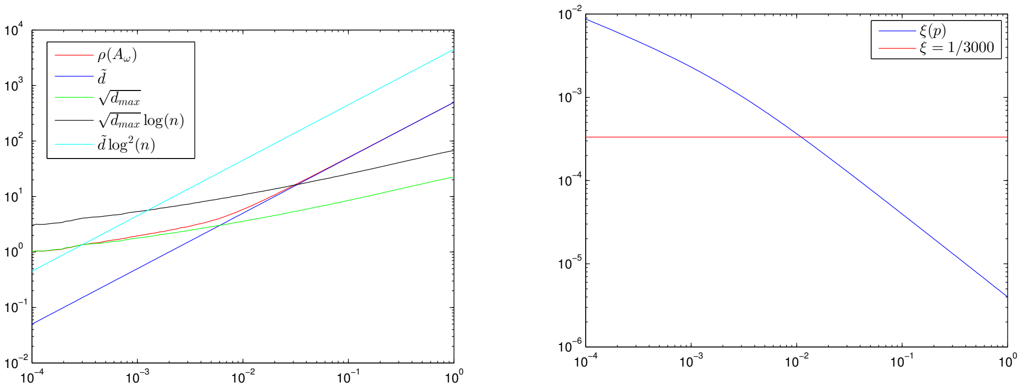


(a) $\gamma = 0$ (b) $\gamma = 1$

Figure 5.10: Evolution of the spectral radius for different value of γ

We see on figure 5.11a that the bound given in theorem 1 and 2 are still valid in this case. We found in chapter 3 that for a disconnected graph the spectral radius could not be predicted by the model, which is also shown in figure 5.11a. Notice that the prediction of the spectral radius we found in chapter 3 is equal to $\tilde{d}$. Therefore, for graph with community of low density, the spectral radius will be of the order of $\sqrt{d_{max}}$ which was not predicted by the model of chapter 3. The distribution of d_{max} is given by (d_{c_i} is the degree in community i).

$$\begin{aligned} P(d_{max} \leq k) &= P(d_{c_1} \leq k)^{n_1} P(d_{c_2} \leq k)^{n_2} \\ d_{c_1} &\sim Bi(n_1, (1-\gamma)p) + Bi(n_2, \gamma p) \\ d_{c_2} &\sim Bi(n_2, (1-\gamma)p) + Bi(n_1, \gamma p) \end{aligned}$$



(a) Evolution of the spectral radius with the density of the network

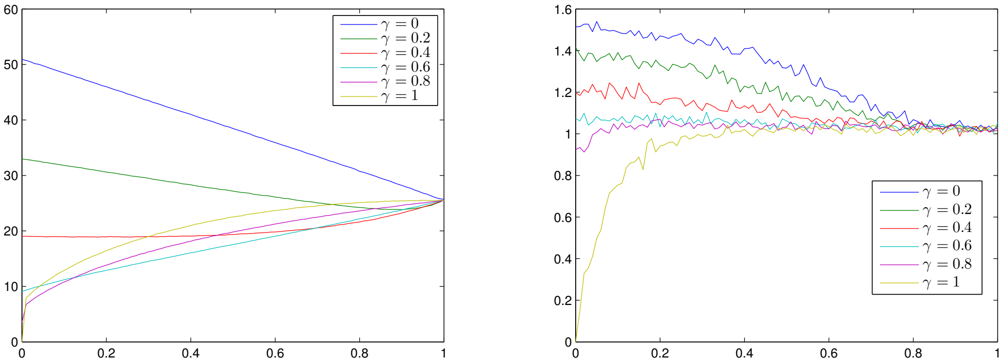
(b) Evolution of ξ for different value of p and the limit of having an almost surely connected graph

In this case we have that $n\log(n) = 3000$ and so if $\xi > 1/3000$ then the graph becomes disconnected which correspond to a value of p of approximately 0.01.

As announced, we can also look at what happens when the two communities have a different size. We work with a population of 1000 people, then we split the population into two communities. Then we use a coefficient $\in [0;1]$. Then we find the size of the two communities as :

$$n_2 = round(500\alpha)$$
$$n_1 = 1000 - n_2$$

In that way, we can analyze all case of the ratio between the size of the two communities. Indeed, $n_2 \in [0;500]$ and $n_1 \in [1000, 500]$, the case $n_2 > n_1$ will not be discussed, since by symmetry this case is obtained by reversing community 1 with community 2. Figures 5.12 show the evolution of the spectral radius of the adjacent matrix with parameter α for different values of γ. We compute it by taking the average of 100 spectral radius of random matrices.



(a) $p = 0.05$

(b) $p = 0.0001$

Figure 5.12: Evolution of the spectral radius with the difference of the two communities for different values of γ

Figure 5.12a and 5.12b show the evolution of the spectral radius for two values of the parameter p. We see on both figures that the spectral radius is in fact independent of γ only in the case where we have two communities of the same size ($\alpha = 1$). The other limit case when $\alpha = 0$ is the case where we have a single community of size 1000. The parameter γ has no sense here since it is the proportion of connection between this great community and a community that actually does not exist, but it can be interpreted as the proportion of edge we remove from this large community. We see indeed that for large value of γ that the spectral radius is near to 0. In the intermediate case we have to consider two cases: the strong linked case ($\gamma > 0.5$) and the weak linked case ($\gamma < 0.5$). In the first case, we start when $\alpha = 0$ with a graph that is nearly disconnected, by increasing α, we will add a larger and larger community and so the expected degree of the vertices from the first community will increase. The graph will thus become more connected (less and less isolated vertex) and the spectral radius will also increase. In the second case, we start with a dense community and we progressively add a community with same density in average, but since these two communities have weak links the global density will decrease and so will the spectral radius.

Finally we see that such a behaviour is independent of the density. The values will of course change but the qualitative behaviour will not. Indeed, we see that figures 5.12a and 5.12b behave in a similar way.

5.3.2 Core-periphery

We will now look at how the eigenvalues of the adjacent matrix will be distributed in the case of a core-periphery model. remember that in such a model we had 4 different probabilities P_1, P_2, $P_{1,2}$ and $P_{2,1}$ (see figure 3.6b, in chapter 3). But this time we will not limit ourselves in the case where the communities have same size, so let n_c the size of the core and n_p the size of the periphery, $n = n_p + n_c$.Let co and pe the sets of vertices respectively in the core and in the periphery. In fact the probabilities $P_{1,2}$ and $P_{2,1}$ are the same since every edge leaving the core must go to the periphery and so is the reverse. Indeed, let A the adjacent matrix

$$\begin{aligned}\sum_{i\in co}\sum_{j\in pe}\mathbb{E}[A_{ij}] &= \sum_{i\in pe}\sum_{j\in co}\mathbb{E}[A_{ij}]\\ \sum_{i\in co}\sum_{j\in pe}P_{1,2} &= \sum_{i\in co}\sum_{j\in pe}P_{2,1}\\ n_cn_cP_{1,2} &= n_cn_cP_{2,1}\\ P_{1,2} &= P_{2,1} = p_{12}\end{aligned}$$

Let w_c and w_p the average degree respectively of the vertices in the core and in the periphery let $p_1 = P_1$ and $p_2 = P_2$. In a similar way as in previous section, we find for w_c and w_p:

$$\begin{aligned}w_c &= \sum_{k\in co} p_1 + (1-\delta)p_2 + \sum_{k\in pe} p_{12}\\ &= n_c(p_1 + (1-\delta)p_2) + n_pp_{12}\\ w_p &= \sum_{k\in pe}\delta p_2 + \sum_{k\in co} p_{12}\\ &= n_p\delta p_2 + n_cp_{12}\end{aligned}$$

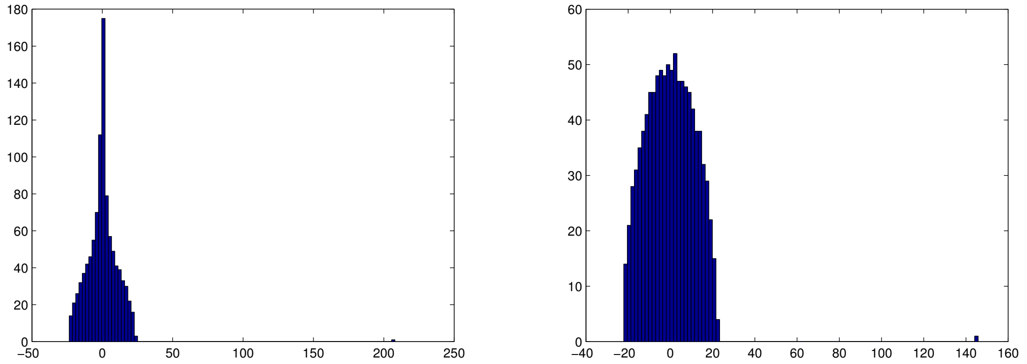
Again, the expected degrees depend on the parameter δ. Remember that δ determines the proportion of edge we take from the periphery to give it to the core. And since the degree of a vertex is defined as the number of neighbors of this vertex, when we add edges inside the core, we increase the expected number of neighbors in the core. We can thus define the probability to get an edge between two vertices i and j by taking again equation (5.2) without giving a definition to ξ yet, we have :

$$p_{ij} = \begin{cases} w_c^2\xi + (1-\delta)w_p^2\xi & \text{if} \quad i,j\in co\\ w_p^2\xi\delta & \text{if} \quad i,j\in pe\\ w_cw_c\xi & \text{if} \quad i\in co \quad j\in pe\end{cases} \tag{5.5}$$

We have to change equation (5.2) into equation (5.5) to make sure the definition of δ is still valid. Indeed, if we keep equation (5.2) in this model the only impact of the δ parameter will be on the expected degree of the vertices. For example setting δ to 0 will change the expected degree but the probability of having two vertices from the periphery linked to each other will not be equal to 0. So we make the change of equation (5.5). Taking equation (5.3) we will compute ξ.

$$\begin{aligned}\frac{1}{\xi} &= \sum_{i=1}^{n_c+n_p}\sum_{j=1}^{n_c+n_p} p_{ij}\\ \frac{1}{\xi^2} &= n_c^2w_c^2 + n_c^2w_p^2 + w_p^2\delta(n_p^2 - n_c^2) + 2n_cn_pw_cw_p\\ \frac{1}{\xi} &= \sqrt{n_c^2w_c^2 + n_c^2w_p^2 + w_p^2\delta(n_p^2 - n_c^2) + 2n_cn_pw_cw_p}\end{aligned}$$

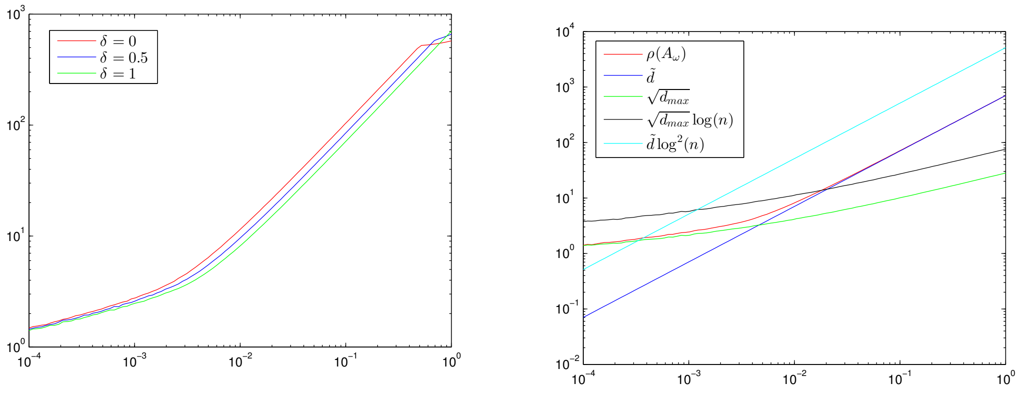
Note that in the case where we have two communities of the same size, ξ is independent of δ. We begin to analyze the behaviour in the case where the two communities have the same size ($n_1 = n_2 = 500$ in the numerical simulations). Let's begin with the distibution of the eigenvalues of the adjacent matrix, we see on figure 5.13 that we have again the semicircular distribution around the origin and a larger eigenvalue which decreases when δ increases. Here the eigenvalues are 207.21 in figure 5.13a and 144.54 in figure 5.13b, while the prediction are respectively 192.96 and 193.14. So it seems that the more the periphery is disconnected with itself, the worst the predictions are. This can be explained by the fact that the probability to get an edge inside the periphery becomes low and thus the eigenvalues fit less well with the theorical model. We can also see on figure 5.13a that the eigenvalues distribution is similar to the semicircular law, but with high concentration on 0. This is again explained by the fact that there is no edge inside the periphery and so there cannot be any direct diffusion from a vertex in the periphery to another vertex in the periphery. This simulation was made with $p_1 = 0.2$, $p_2 = 0.16$ and $p_{12} = 0.1$.



(a) $\delta = 0$ (b) $\delta = 0.9$

Figure 5.13: Evolution of the spectral radius for different value of δ

Again we wonder about the impact of the density on the behaviour of these models when δ varies. Figure 5.14a and5.14b were made by a simulation on a sample of 100 random matrices with parameter p_1 as variable, $p_2 = 0.8 * p_1$ and $p_{12} = 0.5 * p_1$. Figure 5.14a shows that the spectral radius is indeed decreasing when δ increases except for small probabilities, where the spectral radius is independent from the δ parameter as expected in chapter 3. We see on the same figure that for large probabilities the spectral radius is increasing when δ increases, such a phenomenon is also noticed in the theorical case. This is a consequence of the fact that for a too high density the core is totally connected to itself and reducing the density will not disconnected the core but only creating connection inside the periphery. Figure 5.14b shows that the bounds are also valid in this case (the plot is made with $\delta = 1$, but it gives the same result for all values of δ).



(a) Evolution of the spectral radius with the probability for different value of δ

(b) Bound on the spectral radius

Figure 5.14

Figure 5.15 shows the evolution of ξ with p, again the threshold is $\xi = 1/3000$ thus if $\xi > 1/3000$ then the graph is disconnected, this bound is the same as the threshold since it only depends on the number of vertices in the network. But this time the network will be disconnected if the probability $p_1 < 0.002$ with $p_2 = 0.8p_1$ and $p_{12} = 0.5p_1$.

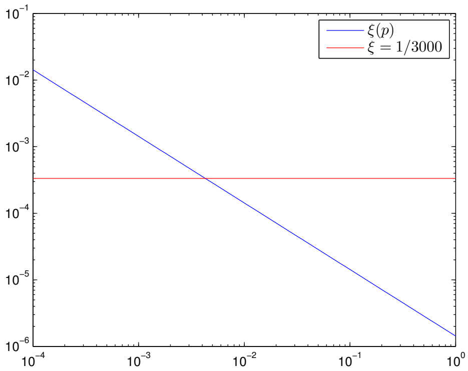
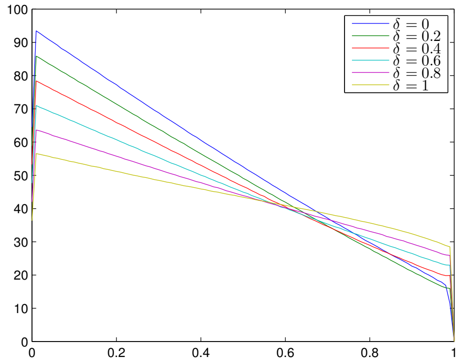


Figure 5.15: Evolution of ξ for different value of p and the limit of having an almost surely connected graph

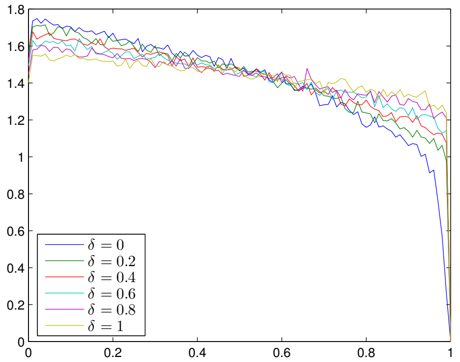
Let's look at what happens when the two communities have a different size. We define again the parameter $0 \leq \alpha \leq 1$. Contrary to the previous section case, there is no symmetric relation between the core and the periphery since the two play different roles. We will thus look at what happens if the periphery is smaller or larger than the core. The size of the communities are then given by for a population with 1000 people:

$$n_p = round(1000\alpha)$$
$$n_c = 1000 - n_2$$

Figures 5.16a 5.16b are made using a numerical simulation on a sample of 100 random matrices of size 1000 for different values of δ, and with parameter $p_2 = 0.8p_1$ $p_{12} = 0.5p_1$ $p_1 = 0.05$ in figure 5.16a and $p_1 = 0.0001$ in figure 5.16b.



(a) Evolution of the spectral radius with α for different value of δ, $p_1 = 0.05$



(b) Evolution of the spectral radius with α for different values of δ, $p_1 = 0.0001$

We notice on figures 5.16a and 5.16b the same behaviour which means that the impact of the size of the two communities will not qualitatively depend on the density of the network. We see that when $\alpha = 0$ there is just a core and no periphery so δ will not have the same meaning as usual. Indeed, δ will just impact the density of the core and so for large values of δ, the core will be less dense. When α increases a bit, which means we have a very small periphery, this last one will be absorbed by the core, which will then increase its size and for very small values of α the spectral radius will thus begin by growing. Then when δ increases, the periphery becomes larger and so the global density decreases. The diminution of the spectral radius is much more important for small values of δ since the periphery is nearly disconnected with itself, the impact of the density will thus be more important for small values of δ. When $\alpha = 0.5$, we are in the case with two communities of same size which was already explained before. When $\alpha > 0.5$, the periphery becomes larger than the core and in this case the spectral radius of the adjacent will increase when δ increases, we then observe an inversion of the pattern. Indeed, since the core becomes smaller and smaller, and the global density of the network becomes nearer to the density of the periphery which increases with δ.

5.3.3 Generalization to several communities

In the case where we have several communities linked to each other in a network of n vertices and let's suppose we have $r < n$ communities each of size n_i $i = 1, ..., r$. Then the procedure to find the spectral radius of the adjacent matrix, is the following:

- Identify the r communities in the network. i.e.the vertices that have same properties (strongly connected to each other, weakly connected to each other, connected to another community,...) and that are linked to vertices of a second community.
- Let w_{c_i} the expected degree of all vertices in community i. We then compute w_{c_i} as the empirical mean of the degree of vertices in the community i.
- Write the probability p_{ij} to get an edge between node i and j that will be the same for each pair of vertices into two communities. This probability eventually depends on a parameter that quantifies the link between two communities found empirically. For example the parameter γ between two communities is the number of edges between these two communities divided by the total number of edges between at least one vertex of the communities and another vertex that could be outside the communities. These probabilities depend on the parameter ξ.

- ◇ Compute ξ thanks to relation: $\sum_{i=1}^{n}\sum_{j=1}^{n} p_{ij} = \xi^{-1}$
- ◇ Using theorems 1 and 2 we can have an idea of the spectral radius. Or we can also build a numerical simulation to find an approximation of the spectral radius.

Using this approach presents the main advantage that we get an approximation of the vector w even if we consider again that all individuals in a community are the same. Indeed, in general, network vector w might be difficult to find, then taking an approximation, will allow us to build a model that can be applied in another population sharing the same characteristic. For example, we studied a given community that we know exactly, so we know the value of R_* exactly, and we can conclude about the diffusion of a disease. Then comes another population that we don't know. All we can say a priori is that it has the same shape as the first population. It is quite hazardous to say that the global reproductive number of the second population is the same as the one of the first population. However, by introducing the concept f community in the first population we find an approximation of R_* that is also valid for the second population.

5.4 Conclusion

The main goal of this chapter was to study the spectral radius of random symmetric binary matrices, for the purpose of applying this to the study of random networks. We first found in the literature that the eigenvalues distribution of a random matrix follows the Wigner semicircle law. And we illustrated it with the case of Erdos-Rényi random graph which corresponds to the case of a population without communities and where the individuals behave in the same way. In this case, we found that a dominant eigenvalue exists in the case of an Erdos-Rényi graph which coincides with what we found in chapter 3. We also noted that, when the random matrix becomes sparse, the spectral radius does not behave as predicted anymore.

Then we found two theorems that describe the spectral radius. These theorems said that for a dense matrix, the spectral radius are nearly equal to the second order average degree, what is close to the predictions made in the previous chapter. In the case of a sparse network, the spectral radius will be nearly equal to the square roots of the maximal degree which was not predicted in the previous chapters.

We then adapted a model found in the literature to be able to make it work on what we want, that is to introduce the notion of communities, core and periphery. Thanks to this, we extended the models of chapter 3. In particular in the case of network of low density. But remember that in this last case, we could have a disconnected graph so the disease will not be able to diffuse everywhere. Finally, we showed the impact of the size of the communities on the spectral radius. Indeed, since the models of chapter 3 are made with two communities of the same size, we wanted to know if the case of two communities of different sizes corresponds to our models. In fact it corresponds in the case where the two communities have approximately the same size.

We finished this chapter by proposing a procedure to build a model for many communities. Unless what was made in the previous chapter, this approach is empirical so there is nearly no modelization mistakes.

Chapter 6

Conclusion

Let's remember what we saw in the four chapters :

- **Classical epidemic model in homogeneous population**
  In first chapter, we saw two epidemic models, SIS and SIR that work in the case of a single homogeneous population. Then we defined the reproductive number that is a determinant parameter in epidemic dynamism. Finally we noticed that this parameter was not sufficient in a more complex population.

- **Case of two communities**
  Here, we studied the case of a population divided in two communities of the same size. We defined the global reproductive number that is an indicator of the disease diffusion in the network. We discovered that this global reproductive number was in fact the spectral radius of the adjacent matrix of the network. Then we studied two cases:

  - When we have a population with two linked communities the more the link between the communities is strong, the slower the diffusion will be, if the two communities have different density. In the reverse case, the diffusion will be independent of the link between the communities
  - When we have a core periphery network, the diffusion will be slower when the network is more homogeneous. Except if in this case the density of the periphery is higher than the one of the core, in which case, the speed of diffusion will reach a minimum value.

  But we also noticed that for too low densities, the models did not fit anymore to reality.

- **Generalization to several communities**
  Here we generalized what we saw in the previous chapter to the case where we have more than 2 communities than could possibly have different sizes. Therefore, we used the Kronecker product to compute the global reproductive number. We found thus a simple expression for the four examples given in the section. But even if the tool developed here might be powerful, it is not easy to compute when the number of communities increases.

- **Statistical study of random networks**
  In this chapter, we looked at the distribution of the eigenvalues and we noted that they follow the Wigner semicircle law and that there is a much larger eigenvalues than the other one, (two

when the population is divided into two communities). This large eigenvalue becomes less large if the network is less dense. This was also noted in the previous chapter. We found in the literature that the spectral radius is approximately equal to the average second order degree in dense network and is approximately equal to the square root of the maximal degree in network with low density.

The main conclusion is that for dense networks, the global reproductive number is easily computed and is much larger than 1, since it is proportional to the number of individuals. But in the case of low density networks, the spectral radius of the adjacent matrix becomes nearly equal to the square root of the maximal degree. In our simulation, this last number was always larger than 1 in low density networks, consequently, if we look only at the criterion : $R_* > 1$ we conclude to an epidemic outbreak in the network. However, in the case of low density networks, the network becomes disconnected and so the disease will only reach a part of the network and not the entire network. And there will not be epidemic outbreak even if $R_* > 1$. The condition of having a connected network is also given a network. Remember that for $0 < C < 1$ and that the number of nonzero entries in a adjacent matrix $n \times n$ is given by Cn^2.

Finally we can say that an epidemic outbreak will occur if :

$$\begin{aligned} &R_0 > 1 \\ &R_* > 1 \qquad \text{if} \qquad C > (1+\epsilon)\frac{\log(n)}{n} \end{aligned}$$

If the graph is disconnected then we cannot conclude on an epidemic outbreak. We know that it will not reach the whole network but we cannot say how many people will be infected by the disease, we thus need another criterion.

Chapter 7

Codes

7.1 Classical epidemics model in hpmogeneous populations

Here are the codes that are used to make the plot of chapter 2. It is the solution of the epidemic model.

```
function [dy] = SIR(t,x)
% Return the derivative value of the SIR model

dy = x;
b = 0.9; % beta
a = 0.1; % mu

dy(1) = -b*x(1)*x(2);
dy(2) = b*x(1)*x(2)-a*x(2);
dy(3) = a*x(2);
end
```

```
function [dy] = SIS(t,x)
% Return the derivative value for the SIS model

dy = x;
b = 0.3; % beta
m = 0.5; % mu
dy(1) = -b*x(1)*x(2)+ m*x(2);
dy(2) = b*x(1)*x(2)- m*x(2);
end
```

```
% This script compute the solution for the SIR model
%
SI0 = [0.9 0.1 0];
[T,S] = ode45(@SIR, [0  60],SI0);
```

```
figure
plot(T,S(:,1),'g', T,S(:,2),'r',T,S(:,3),'k')
leg = legend('S','I','R');
ylim([0 1])
title('\beta = 0.9 and \mu = 0.1');

%%
% This script compute the solution for the SIS model
%

SI0 = [0.9 0.1];;
[T,S] = ode45(@SIS, [0  60],SI0);

figure
plot(T,S(:,1),'g', T,S(:,2),'r')
leg = legend('S','I');
ylim([0 1])
title('\beta = 0.3 and \mu = 0.5');
```

7.2 Case of two communities

This function is used to create the plot of chapter 3. It is a numerical simulation to get an approximation of the spectral radius of adjacent radius.

```
function [ lambda ] = MakeSimulation2C(P1,P2,P12,P21,n,sim,type,h,m)
% This function computes the spectral radius of a matrix of size 2n in the
% case where it has two distinct community one of size n. To compute the
% spectral radius, we make m simulations of random matrices and take the
% average of thei spectral radius. type can either be 'Community' or 'Core'
% following the kind of network we are intereting in. If type = 'Community'
% then P1 and P2 are the density of the two community P12 and P21 are
% meaningless and we make the link between the two community vary from 0 to
% 1 with step h. If type = 'Core' then P1 and P2 are respectively the
% density of the core and of the periphery when the network is homogeneous
% P12 and P21 are respectively the edges leaving the core to periphery and
% the edges leaving the periphery to core. lambda is a two line vector with
% at the first line the result of the siulation and at second line the
% predictions of the spectral radius.
%

% Compute the sum and the difference of the two community densities
P = P1+P2;
Pb = abs(P1-P2);

g = {'Community' 'Core'};
str = strcmp(type,g);
t = find(str);

if t==1 % Two linked communities
```

```
gamma = 0:h:1; % Proportion of link between the two communities
lambda = zeros(2,length(gamma));
if sim % Symmetric case (P2 = P1)
    for i = 1:m % Loop on the simulations
        k = 1;
        for g = gamma % Loop trying all values of gamma
            B = rand(n); % Random matrix
            B = 0.5*(B+B'); % Make the random matrix symmetric
            % Create the submatrix for community 1
            A1 = B<(1-g)*P1;
            % Create submatrix linking the two communities
            A2 = rand(n)<g*P1;
            B = rand(n); % Random matrix
            B = 0.5*(B+B'); % Make the random matrix symmetric
            % Create the submatrix for community 2
            A3 = B<(1-g)*P1;
            A = [A1 A2;A2' A3]; % Create the general adjacent matrix
            A = double(A);
            l = eig(A); % Compute the eigenvalue of the adjacent matrix
            % Store the spectral radius
            lambda(1,k) = lambda(1,k)+max(l);
            k = k+1;
        end
    end
    lambda = lambda/m; % Compute the empirical mean
    lambda(2,:) = P1*n; % Predictions
else % Non symmetric case
    for i = 1:m % Loop on the simulations
        k = 1;
        for g = gamma % Loop trying all values of gamma
            % Create the submatrix for community 1
            A1 = rand(n)<(1-g)*P1;
            % Create submatrix linking from community 1 to community 2
            A2 = rand(n)<g*P1;
            % Create submatrix linking from community 2 to community 1
            A3 = rand(n)<g*P2;
            % Create the submatrix for community 1
            A4 =  rand(n)<(1-g)*P2;
            A = [A1 A2;A3 A4]; % Create the adjacent matrix
            A = double(A);
            o = eig(A); % Compute the eigenvalue of the adjacent matrix
            % Store the spectral radius
            lambda(1,k) = lambda(1,k)+max(abs(o));
            k = k+1;
        end
    end
    lambda = lambda/m; % Empirical mean
    % Compute the predictions
    D =  P^2*n^2*gamma.^2-2*n^2*Pb^2*gamma+Pb^2*n^2;
    r = (1-gamma)*P*n*0.5+sqrt(D)*0.5;
    lambda(2,:) = r;
```

```
    end
elseif t==2 % Core-periphery
    % Proportion of link between taken form to periphery and give to core
    delta = 0:h:1;
    lambda = zeros(3-1+sim,length(delta));
    if sim % Symmetric case (P2 = P1, P12 = P21)
        for i = 1:m % Loop on the simulations
            k = 1;
            for d = delta % Loop trying all values of delta
                B = rand(n); % Random matrix
                B = 0.5*(B+B'); % Make the random matrix symmetric
                % Create the submatrix for the core
                A1 = B<P1+(1-d)*P1;
                % Create submatrix linking the core with the periphery
                A2 = rand(n)<P12;
                B = rand(n); % Random matrix
                B = 0.5*(B+B'); % Make the random matrix symmetric
                % Create the submatrix for the periphery
                A3 =  B<d*P1;
                A = [A1 A2;A2' A3]; % Create the general adjacent matrix
                A = double(A);
                o = eig(A);% Compute the eigenvalue of the adjacent matrix
                % Store the spectral radius
                lambda(1,k) = lambda(1,k)+max(o);
                k = k+1;
            end
        end
        lambda = lambda/m; % Compute the empirical mean
        % Compute the predictions
        r = P1*n+n*sqrt(P1^2*(1-delta).^2+P12^2);
        lambda(2,:) = r;
        lambda(3,:) = P1*n*(2-delta);
    else % Non symmetric case
        for i = 1:m % Loop on the simulations
            k = 1;
            for d = delta % Loop trying all values of delta
                % Create the submatrix for the core
                A1 = rand(n)<P1+(1-d)*P2;
                % Create submatrix linking the core to the periphery
                A2 = rand(n)<P12;
                % Create submatrix linking the periphery to the core
                A3 = rand(n)<P21;
                % Create the submatrix for the periphery
                A4 =  rand(n)<d*P2;
                A = [A1 A2;A3 A4]; % Create the general adjacent matrix
                A = double(A);
                o = eig(A);% Compute the eigenvalue of the adjacent matrix
                % Store the spectral radius
                lambda(1,k) = lambda(1,k)+max(abs(o));
                k = k+1;
            end
```

```
        end
        lambda = lambda/m; % Empirical mean
        % Compute the predictions
        D = n^2*(4*P2^2*delta.^2-4*P2*P*delta+4*P12*P21+P^2);
        r = 0.5*P*n+0.5*sqrt(D);
        lambda(2,:) = r;
    end
end
end
```

7.3 Statistical study of random networks

These functions are used in chapter 5 to find the distribution of the eigenvalues. This function return a symmetric random matrix .

```
function [ A ] = GiveA(W,param , n1, n2,model)
% This function returns an random adjacent matrix A on size n( =n1+n2).
% Given the vector W of size n which contains the expected dergree of all
% vertices. param, n1,n2,model are facultative input. model is either
% 'Core' or 'Community'. if model = 'Core' then param is a value of the
% core-periphery parameter; If model = 'Community' the param is a value of
% the linking coefficient between the communities. n1 and n2 are the size
% of the two communities if model = 'Community' and n1 is the size of the
% core and n2 the size of the periphery if model = 'Core'.
%

n = length(W); % Number of vertices


if nargin  == 1 % Case of a single community Erdos-R[U+FFFD]i graph

    xi = 1/sum(W); % Definition of vector xi
    A = rand(n); % random matrix
    % Matrix of the probability P(i,j) is the probability
    % to get an edge between vertex i and j
    P = W'*W*xi;
    A = A<P; % Create the adjacent matrix
    double(A);
    % Make the adjacent matrix symmetric
    A = triu(A);
    A = A+A'-diag(diag(A));

else %Case of two communities

    g = {'Community' 'Core'};
    str = strcmp(model,g);
    t = find(str);

    % Make sure to have the right number of individuals in the population
```

```
if n1+n2~=n
    print('Error on the size of the communities');
end
if t==1 % Two linked communities
    % Definition of vector xi
    xi = 1/sqrt(n1^2*W(1)^2+n2^2*W(end)-param*(n1*W(1)-n2*W(end)));
    % Create the probability matrix of getting an edge between two
    % vertices of the first community
    P1 = W(1:n1)'*W(1:n1)*xi;
    % Create the probability matrix of getting an edge between two
    % vertices of the second community
    P2 = W(n1+1:n)'*W(n1+1:n)*xi;
    % Create the probability matrix of getting an edge between two
    % vertices of distinc communities
    P3 = W(1:n1)'*W(n1+1:n)*xi;
    % Create the submatrices of the adjacent matrix
    A1 = rand(n1);
    A1 = A1<P1*(1-param);
    A2 = rand(n2);
    A2 = A2 <(1-param)*P2;
    A3 = rand(n1,n2);
    A3 = A3<param*P3;
    A = [A1 A3;A3' A2]; % Create the general adjacent matrix
    double(A);
    % Make the adjacent matrix symmetric
    A = triu(A);
    A = A+A'-diag(diag(A));
else % Core-periphery network
    % Definition of vector xi
    xi = 1/sqrt(n1^2*W(1)^2+n1^2*W(end)^2 ...
        +W(end)^2*param*(n2^2-n1^2)+2*n1*n2*W(1)*W(end));

    % Create the probability matrix of getting an edge between two
    % vertices of the core
    P1 = W(1:n1)'*W(1:n1)*xi;
    % Create the probability matrix of getting an edge between two
    % vertices of the periphery
    P2 = W(n1+1:n)'*W(n1+1:n)*xi;
    % Create the probability matrix of getting an edge between two
    % vertices of the core and of the periphery
    P3 = W(1:n1)'*W(n1+1:n)*xi;
    % Create the submatrices of the adjacent matrix
    A1 = rand(n1);
    if n1==0
        p1 = 0;
    else
        p1 = P1(1,1);
    end
    if n2==0
        p2 = 0;
    else
```

```
            p2 = P2(1,1);
        end
        A1 = A1<(p1+p2*(1-param))*ones(size(A1));
        A2 = rand(n2);
        A2 = A2 <param*P2;
        A3 = rand(n1,n2);
        A3 = A3<P3;
        A = [A1 A3;A3' A2]; %Create the general adjacent matrix
        double(A);
        % Make the adjacentmatrix symmetric
        A = triu(A);
        A = A+A'-diag(diag(A));
    end
end
end
```

This script is used in the case of Erdos-Rényi graph. This script is used in the case of two linked communities.

```
% This script make the plots present in chapter 5 when we consider
% a two linked community graph. Run this first selection Before any other.
%

n1 = 500; % Size of the first community
n2 = 500; % Size of the second community
n = n1+n2;
p = 0.05; % Probability to get an edge between two vertices
gamma  = 0:0.1:1; % Linking coefficient between the two cmmunities
%%
% Compute the distribution of the eigenvalues of the adjacent matrix
%


i = 1;
for g = [0 0.2 0.8 1] % Value of gamma we try
    % Compute the expected degree of the vertices
    w1 = p*(n1*(1-g)+n2*g);
    w2 = p*(n2*(1-g)+n1*g);
    W = [w1*ones(1,n1) w2*ones(1,n2)];
    A = GiveA(W,g,n1,n2,'Community'); % Create the adjacent matrix
    l = eig(A); % Compute its eigenvalues
    figure % Create an histograms with the eigenvalues
    hist(l,100);
    i = i+1;
end

%%
% Compute the spectral radius for different value of p
%
```

```
p = logspace(-4, 0, 100);
nbrsimu = 100; % Number of spectral radius computed to get their mean
l = p;
dmax = p;
dtil = l;
g = 0; % value of gamma we test

for i = 1:length(p) % Loop on p
    % Expected degree of the vertices
    w1 = p(i)*(n1*(1-g)+n2*g);
    w2 = p(i)*(n2*(1-g)+n1*g);
    W = [w1*ones(1,n1) w2*ones(1,n2)];
    l3 = zeros(1,nbrsimu);
    d3 = l3;
    for j = 1:nbrsimu % Compute the average of the spectral radius
        % Create an rando adjacent matrix
        A = GiveA(W,g,n1,n2,'Community');
        A = sparse(A);
        l3(j) = eigs(A,1); % Spectral radius
        d3(j) = max(sum(A,2)); % Maximum degree
    end
    l(i) = mean(abs(l3)); % Empirical mean of the spectral radius
    dmax(i) = mean(d3); % Mean of the maximal degree
    dtil(i) = sum(W.^2)/sum(W); % Second order average degree
end

figure % Plot the spectal radius versus p
loglog(p,l,'r');
hold on
loglog(p,dtil,'b');
loglog(p,sqrt(dmax),'g');
loglog(p,sqrt(dmax)*log10(n),'k');
loglog(p,dtil*log10(n)^2,'c');

%%
% Compute the evolution of the spectral radius with the size of the
% communities
%

n = 500; % Half of the whole population size
nbrsimu = 100; % Number of spectral radius computed to get their mean
r0 = zeros(length(0:0.1:1),length(0:0.01:1));
i = 1;
for g = 0:0.2:1; % Tested value of gamma
    j = 1;
    for a = 0:0.01:1 % Proportion of the second community size
        n2 = round(a*n); % Second comunity size
        n1 = 1000-n2; % First community size
        % Expected degree of the vertices
        w1 = p*(n1*(1-g)+n2*g);
        w2 = p*(n2*(1-g)+n1*g);
```

```
        W = [w1*ones(1,n1) w2*ones(1,n2)];
        l3 = zeros(1,nbrsimu);
        for k = 1:nbrsimu % Compute the average of the spectral radius
            % Create an rando adjacent matrix
            A = GiveA(W,g,n1,n2,'Community');
            A = sparse(A);
            l3(k) = eigs(A,1); % Spectral radius
        end
        r0(i,j) = mean(abs(l3)); % Empirical mean of the spectral radius
        j = j+1;
    end
    i = i+1;
end

% Plot the evolution of teh spectral radius with the communities size
figure
a = 0:0.01:1;
plot(a,r0)
leg = legend('$\gamma = 0$','$\gamma = 0.2$', '$\gamma = 0.4$',...
    '$\gamma = 0.6$','$\gamma = 0.8$','$\gamma = 1$');
set(leg,'Interpreter','latex')
set(leg,'FontSize',12);
%%
% Compute the evolution of xi with p in the case of two communities of same
% size
%

g = 0;
p = logspace(-4,0,100);
w1 = p.*(n1*(1-g)+n2*g);
w2 = p.*(n2*(1-g)+n1*g);
xi = 1./sqrt(n1^2*w1.^2+n2^2*w2-g*(n1*w1-n2*w2));

figure % Plot the evolution of xi
loglog(p,xi,'b,')
hold on
plot([p(1) p(end)], 1/(n*log10(n))*ones(1,2));

leg = legend('$\xi(p) $','$\xi = 1/3000$');
set(leg,'Interpreter','latex')
set(leg,'FontSize',12);
```

This script is used in the case of a core-periphery graph.

```
% This script make the plots present in chapter 5 when we consider
% a coreperiphery graph. Run this first selection Before any other.
%

n1 = 500; % Size of the core
n2 = 500; % Sizeof the periphery
```

```
n = n1+n2;
p1 = 0.2; % Density of the core
p2 = 4*p1/5; % Density of the periphery
p12 = p1/2; % Density of edge between the core and the periphery
delta = 0:0.01:1 ; % Coefficient of the core periphery
%%
% Compute the distribution of the eigenvalues of the adjacent matrix
%

for d = [0 0.9] % Value of gamma we try of delta
    % Compute the expected degree of the vertices
    w1 = n1*(p1+(1-d)*p2)+n2*p12;
    w2 = n2*d*p2+n1*p12;
    W = [w1*ones(1,n1) w2*ones(1,n2)];
    A = GiveA(W,d,n1,n2,'Core'); % Create the adjacent matrix
    l = eig(A); % Compute its eigenvalues
    figure % Create an histograms with the eigenvalues
    hist(l,100);
end


%%
% Compute the spectral radius for different value of p
%

p = logspace(-4, 0, 100);
nbrsimu = 100; % Number of spectral radius computed to get their mean
l = p;
dmax = p;
dtil = p;
d =  1; % value of delta we test
for i = 1:length(p) % Loop on p1
    p1 = p(i);
    p2 = 4*p1/5;
    p12 = p1/2;
    % Expected degree of the vertices
    w1 = n1*(p1+(1-d)*p2)+n2*p12;
    w2 = n2*d*p2+n1*p12;
    W = [w1*ones(1,n1) w2*ones(1,n2)];
    l3 = zeros(1,nbrsimu);
    d3 = l3;
    for j = 1:nbrsimu % Compute the average of the spectral radius
        % Create an rando adjacent matrix
        A = GiveA(W,d,n1,n2,'Core');
        A = sparse(A);
        l3(j) = eigs(A,1); % Spectral radius
        d3(j) = max(sum(A,2)); % Maximum degree
    end
    l(i) = mean(abs(l3)); % Empirical mean of the spectral radius
    dmax(i) = mean(d3); % Mean of the maximal degree
    dtil(i) = sum(W.^2)/sum(W); % Second order average degree
```

```
end

figure % Plot the spectal radius versus p1 , p2 = 4*p2/5, p12 = p1/2
loglog(p,l,'r');
hold on
loglog(p,dtil,'b');
loglog(p,sqrt(dmax),'g');
loglog(p,sqrt(dmax)*log10(n),'k');
loglog(p,dtil*log10(n)^2,'c');


%%
% Compute the evolution of the spectral radius with the size of the
% communities
%

n = 1000; % Size of the population
nbrsimu = 100; % Number of spectral radius computed to get their mean
r0 = zeros(length(0:0.1:1),length(0:0.01:1));
i = 1;
for d = 0:0.1:1; % Tested value of delta
    j = 1;
    for a = 0:0.01:1 % Proportion of the second community size
        n2 = round(a*n); % Second comunity size
        n1 = 1000-n2; % First community size
        % Expected degree of the vertices
        w1 = n1*(p1+(1-d)*p2)+n2*p12;
        w2 = n2*d*p2+n1*p12;
        W = [w1*ones(1,n1) w2*ones(1,n2)];
        l3 = zeros(1,nbrsimu);
        for k = 1:nbrsimu % Compute the average of the spectral radius
            % Create an rando adjacent matrix
            A = GiveA(W,g,n1,n2,'Core');
            A = sparse(A);
            l3(k) = eigs(A,1); % Spectral radius
        end
        r0(i,j) = mean(abs(l3)); % Empirical mean of the spectral radius
        j = j+1;
    end
    i = i+1;
end

% Plot the evolution of teh spectral radius with the communities size
figure
a = 0:0.01:1;
plot(a,r0)
leg = legend('$\delta = 0$','$\delta = 0.2$', '$\delta = 0.4$',...
    '$\delta = 0.6$','$\delta = 0.8$','$\delta = 1$');
set(leg,'Interpreter','latex')
set(leg,'FontSize',12);
%%
```

```
% Compute the evolution of xi with p1 in the case of two communities of
% same size
%


i = 1;
d = 1;

p1 = logspace(-4,0,100);
p2 = 4*p1/5;
p12 = p1/2;
w1 = n1*(p1+(1-d)*p2)+n2*p12;
w2 = n2*d*p2+n1*p12;
xi = 1./sqrt(n1^2*w1.^2+n1^2*w2.^2 ...
    +w2.^2*d*(n2^2-n1^2)+2*n1*n2*w1.*w2);

figure % Plot the evolution of xi
loglog(p,xi,'b,')
hold on
plot([p1(1) p1(end)], 1/(n*log10(n))*ones(1,2));

leg = legend('$\xi(p1) $','$\xi = 1/3000$');
set(leg,'Interpreter','latex')
```

Chapter 8

References

Bibliography

[1] *Compartmental models in epidemiology* wikipedia, `https://en.wikipedia.org/wiki/Compartmental_models_in_epidemiology`

[2] Cheng, Sibo., *random diffusion on social network with application of disease spreading* , Ecole polytechnique de Louvain, Université catholique de Louvain, 2017. Prom. : Delvenne, Jean-Charles.

[3] Silvester John R., *Determinants of Block Matrices*, Department of Mathematics, King's College, Strand, London, 1999

[4] *Erdős–Rényi model* wikipedia , `https://en.wikipedia.org/wiki/Erd%C5%91s%E2%80%93R%C3%A9nyi_model`

[5] Boutet de Monvel Anne, Khorunzhy Alexei *Some Elementary Results around the Wigner Semicircle Law*, p7-9, 2001

[6] Fan Chung, Linyuan Lu, Van Vu, *Eigenvalues of Random Power law Graphs*, Department of mathematics, Univesity of California, San Diega, 2003

[7] Linh V. Tran, Van H. Vu, Ke Wang, *Sparse random graphs: Eigenvalues and Eigenvectors*, Department of Mathematics, Rutgers, Piscataway, p1-5, 2010

[8] Bai Z., Silverstein J.W., *Spectral Analysis of Large Dimensional Random Matrices*, Second Edition, Springer Series in Statistics, p15, 2010

[9] Feier Adina Roxana, *Methods of Proof in Random Matrix Theory*, Harvard University Department of Mathematics, p 1-8, 2012, Prom. : Brenner, Michael P.

[10] Hastings Harold M. *The May-Wigner Stability theorem for connected matrices* , Americanmathematical society, volume 7, number 2, p 387-388,1982

[11] *Les réseaux sociaux, nerfs de la guerre du Printemps arabe*, le Parisien, 14 juin 2013

Rue Archimède, 1 bte L6.11.01, 1348 Louvain-la-Neuve www.uclouvain.be/epl