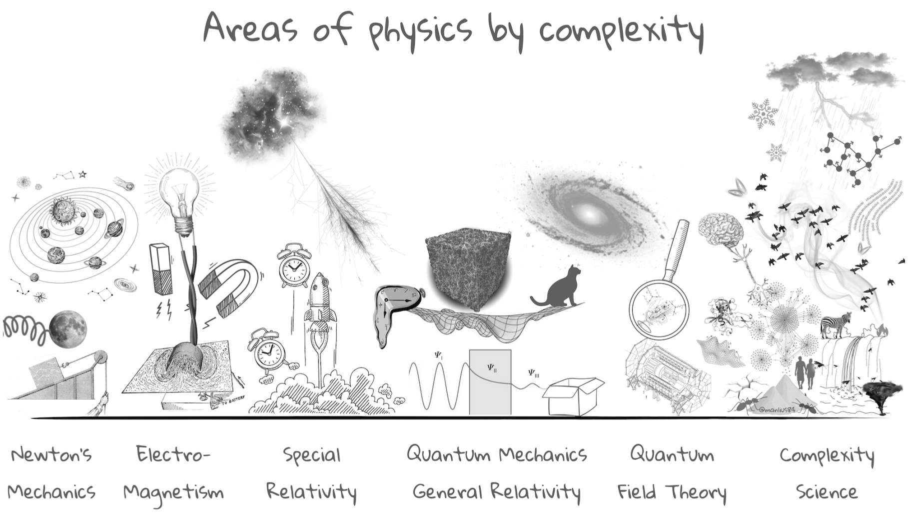
Final Report

Physics of Complex Networks: Structure and Dynamics

Università degli Studi di Padova



Project 10: Sociophysics II

Guillermo Benito Calviño, Claudia Lorenzetti, Leon Mengoni, Sofía Pacheco García

Group 10: SocioNets

Last update: July 23, 2023

Contents

1 | Schelling segregation model

Task leader(s): *Guillermo Benito Calviño and Sofía Pacheco García*

Schelling model [4] studies how individual preference and perception of difference can lead collectively to segregation. The model divides the population into two different groups assigning an attribute to each person. The attribute is permanent and recognizable for all the rest of the individuals. Everyone is conscious and cares about their nearest neighbours (neighbourhood). Each individual is also assigned a level of tolerance, which is defined as the ratio of people different from oneself in one's neighbourhood.

The main model explained in [4], and the one we will later on study as a social network, is the **Spatial Proximity Model**. Agents, each with a given attribute, are distributed in space. All the population has the same level of tolerance for others with attributes different to their own. Each person defines its neighbourhood as a fixed number of nearest neighbours. If an agent's neighbourhood numbers overcome their level of tolerance they are allowed to move to the closest point in space where their conditions are satisfied. Movements are performed one by one, an order in which individuals move has to be chosen. In this model it can be observed that clusters of similar people (with equal attributes) are created in the stable state. If the distribution of the attributes is not equal, segregation is more extreme.

Nevertheless Schelling considers also different models such as **Bounded - Neighbourhood**, where neighbourhoods are fixed and common for everyone; or **Tipping**, which is a special case of the previous model taking into account minorities.

In [2] the effect of segregation is studied within the social networks framework. To do so the Spatial Proximity Model is adapted to a social network where there is a fixed number n of agents and a fixed number m of edges. Everyone has their own neighbourhood only made of their nearest neighbours . There are still some differences between the network and the model: now the tolerance is represented with a different parameter called aversion p, comprehended between 0 and 1, and the attributes are continuous and described as a number also between 0 and 1.

The main objective of this task is to recreate the results shown in [2]. Firstly we have created a simulation of their proposed network in order to create data. Later, we have compared them to the analytical results. Moreover we have generalized the network altering some fixed conditions such as the uniqueness of p.

1.1 Description and evolution of the network

We shall start by describing the algorithm that makes our network evolve. The initial network is defined as $G(V, E)_{t_0}$. At each step a random pair of connected nodes is selected and the distance between their attributes is computed. The distance multiplied by p is the probability that this edge breaks. If the edge does not break the network proceeds to the next step without changing. If instead it breaks, another edge between two random not connected agents is created. In this way the number of edges does not change. To understand the results we have to go one step further and discretize the distances (d_K). We do it by initially fixing a whole number K which defines the number of possible values for the distance. Then we approximate the distance to the nearest top possible value[1].

The distance is calculated as

$$d(x, y) = min\{2 \cdot |x - y + u| : u \in \{-1, 0, 1\}\}^{2}. \tag{1.1}$$

We need a way to measure the segregation in our results. To do so we introduce the parameters $E_i(t)$ where $i \in \{1, 2, 3, ..., K\}$. $E_i(t)$ is the amount of edges where the distance between their nodes is equal to i/K. If segregation is present we shall expect a big amount of "short" edges and few "long" ones. This means that $E_1 \gg E_K$.

To obtain an analytical expression for E_i we need to solve its master equation. The rate of change in the expected value of E_i between a certain step *t+1* and its previous step t will have two contributions: the first term corresponds to the probability that a new edge with distance i/K is created; the second term is the probability that an edge with distance i/K is destroyed. Its complete mathematical expression is

$$\mathbb{E}\{(E_i(t+1) - E_i(t)\} = \frac{F(t)p}{Km}(1 + o(1)) - \frac{E_i(t)ip}{Km} \tag{1.2}$$

where $F(t) = \sum d_K(u, v)p$, noting that the summation is over all the edges in the graph. Finally we shall move to the continuum spectrum defining $e_i(x) = \frac{1}{m}E_i(t)$, where x equals t/m and is continuous. In doing so the LHS becomes $de_i(t)/dt$. Equation 1.2 converges to a stable state (as expected from Schelling model) with solution

$$e_i(x) = \frac{1}{i\sum_{j=1}^{K}\frac{1}{j}}. \tag{1.3}$$

[1] We want to impose that there always exists the possibility of breaking the edge, thus if the distance is equal to 0 we change it to $1/K$.

[2] Toroidal distance, the 0 is close to the 1. We need to use this definition of distance so it will follow a uniform distribution instead of a Gaussian when the attributes follow also a uniform distribution. Note that it can only yield values between 0 and 1.

1.2 | Analysis of the simulated networks

We have compared the solutions obtained from the simulation with the analytical solution and, apart from some fluctuations, we can see in figure 4.1 that the results are compatible. In this first case the attributes follow a random uniform distribution. As we can observe, the stable state does not depend on the initial state nor on the parameter p. This proves that segregation will always appear in this model except if p is fixed to 0. The convergence is faster for higher values of p.

We also have plotted the network in different stages of the evolution to see the segregation process. For visual reasons we have used a network with $n = 100$ and $m = 300$. The distribution of the attributes has to be changed since now we want two differentiated groups. The attributes will then follow a bimodal distribution [3]. We have also fixed K to a high value to be closer to the continuum since we are not longer interested in visualizing E_i. In figure 1.1 we can see at all points two differentiated groups (blue circles and green squares), which each correspond to agents whose attributes are higher or lower than the mean respectively. As time evolves the groups become more differentiated and, most importantly, more distant until the system arrives to the stable and final state.

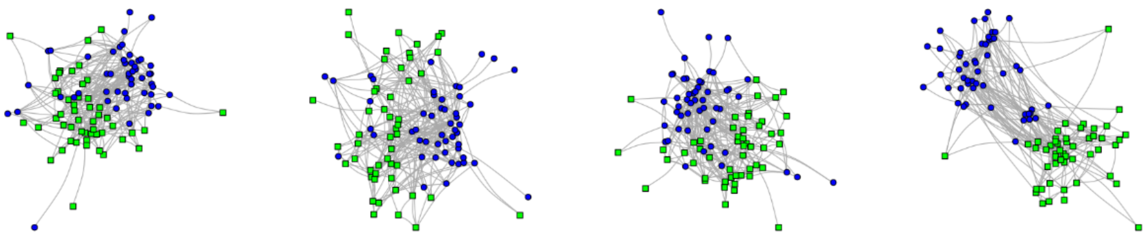


Figure 1.1: Social network evolution for $n = 300$, $m = 500$ and same p for every agent. From initial state (left) to stable state (right).

In [2] the same p is also used for all the nodes. We have decided to assign a different value of aversion p to each individual in an attempt to resemble closer to reality. Doing so we can appreciate how at figure 4.2 small clusters within each group are created. They are usually governed by an individual with high aversion.

Finally we have studied the case where one group is a minority, 4.3. The network starts more mixed than in the previous cases. It can be seen that a highly concentrated center of the majority is created and the minority stands widely in the periphery at end. This leads to a more notable segregation, as Shelling had predicted.

[3] In practice we have used a uniform distribution from 1/8 to 3/8 and another one from 5/8 to 7/8.

2 | Deffaunt's Bounded Confidence Model

Task leader(s): *Leon Mengoni*

In this work on opinion dynamics, opinions are considered as a continuous variable between 0 and 1. Readjustment occurs when similarity between opinions is within a fixed threshold parameter. As outlined in [1], we will perform analyses by first considering *complete mixing*, where social actors (agents) can randomly exchange opinions with every other agent. Next, we will allow for opinion adjustment to occur only for agents who are socially connected, and to this end we will use different types of networks. Finally, we will consider binary vectors of opinions.

2.1 | Complete mixing

2.1.1 Model

We will consider here a population of N agents i, with continuous opinions $x_i \in [0, 1]$. If two agents meet, they readjust their opinions if the difference in their opinions is less than a *threshold parameter* d. Therefore, for two agents i and j, if $|x_i - x_j| < d$:

$$\begin{cases} x_i = x_i + \mu(x_j - x_i) \\ x_j = x_j + \mu(x_i - x_j) \end{cases} \tag{2.1}$$

where μ is the *convergence parameter*, fixed at a value between 0 and 0.5.

In this section, all calculations will be done considering *complete mixing*, i.e. all agents are connected to all agents.

2.1.2 Analysis and results

First of all, we are interested in verifying the convergence of opinions for varying values of d and μ.

For varying μ, the only observable difference is the convergence time, which is shorter for larger values of μ. Qualitative differences can be observed, on the other hand, for varying d: uni-modal convergence of opinions (single peak) occurs only for larger values ($\gtrsim 0.5$), while for $d < 0.5$, there can be multiple peaks. As a matter of fact, d controls the number of peaks in the final landscape of opinions. We can see clearly in figure 2.2 that the value of d determines the number of final peaks. In our simulation, we generated 200 samples, which we let evolve according to our previous

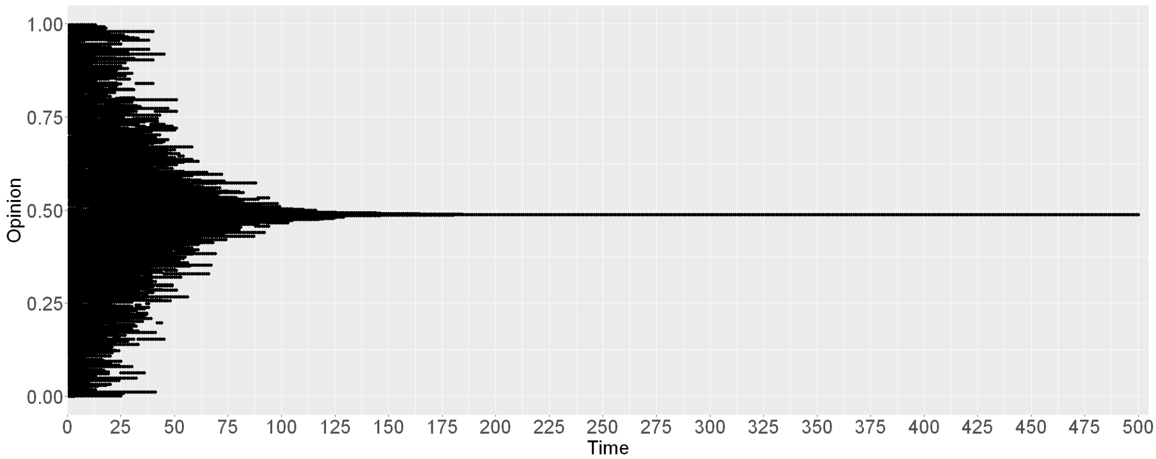


Figure 2.1: Converging opinions, for $d = 0.5$ and $\mu = 0.5$

rule. We find that the number of peaks is, generally, inversely proportional to the value of d.

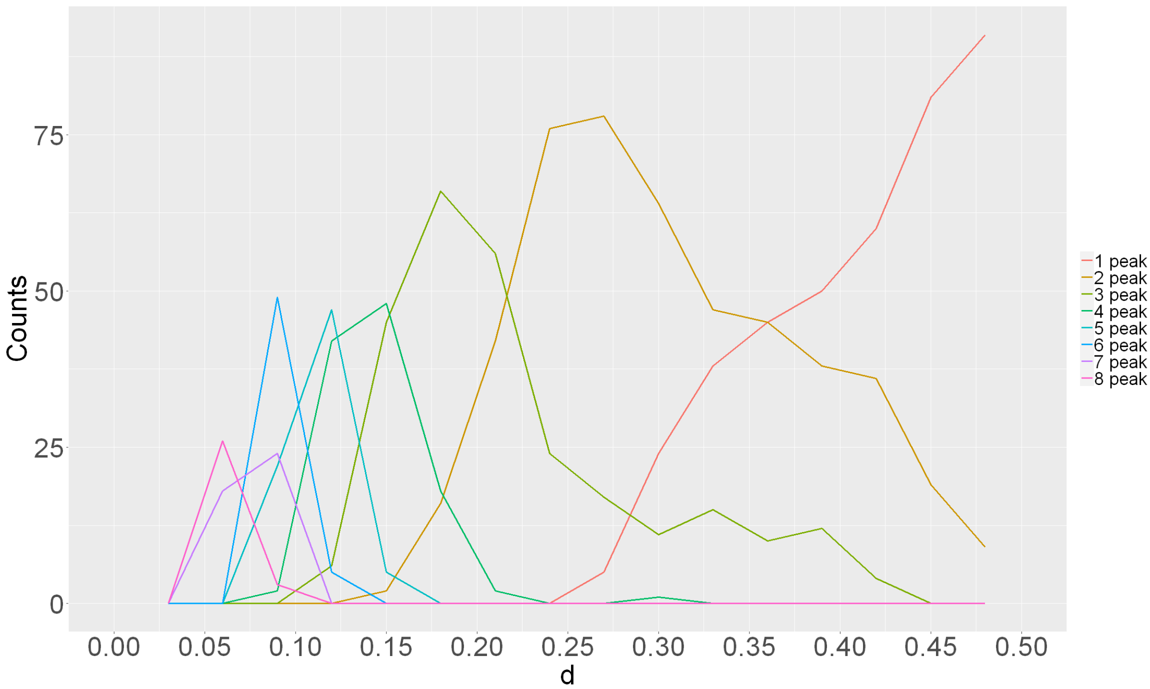


Figure 2.2: Histograms of number of peaks as a function of threshold parameter d

2.2 | Social networks

2.2.1 Model

In this section, readjustment of opinions will occur only if two agents are connected via a social connection, i.e. an edge between two nodes of a graph, according to the previously stated rule. The analysis will be undertaken on a square lattice first, then on more complex networks.

2.2.2 Square lattice

On a square $n \times n$ lattice, each node is connected with its four N, S, E and W neighbors. Again, qualitative differences arise when we vary the value for d. Here, we identify different regimes:

- for $d > 0.3$, the whole population tends to converge to a single opinion, around 0.5, except for some isolated nodes, dubbed *extremists*;
- for $d < 0.3$, converging opinions occur only on connected clusters. In addition, there is only one percolating cluster which spans the whole lattice, while all others are localized and disjoint.

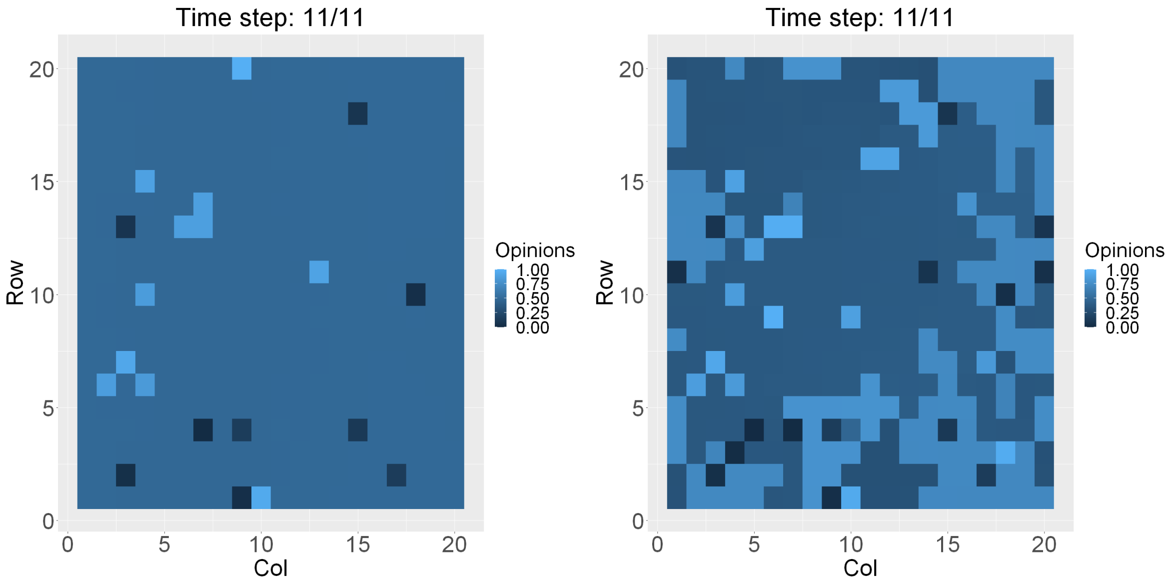


Figure 2.3: Converging opinions on a square lattice with $d = 0.31$ (left) and $d = 0.23$ (right)

The histogram of final opinions for the latter case ($d = 0.23$) shows that clusterization of opinions is not entirely homogeneous. Nodes within the same cluster have similar opinions, but can slightly fluctuate for long periods of time, since homogenization depends on the connections between clusters and the way these have evolved.

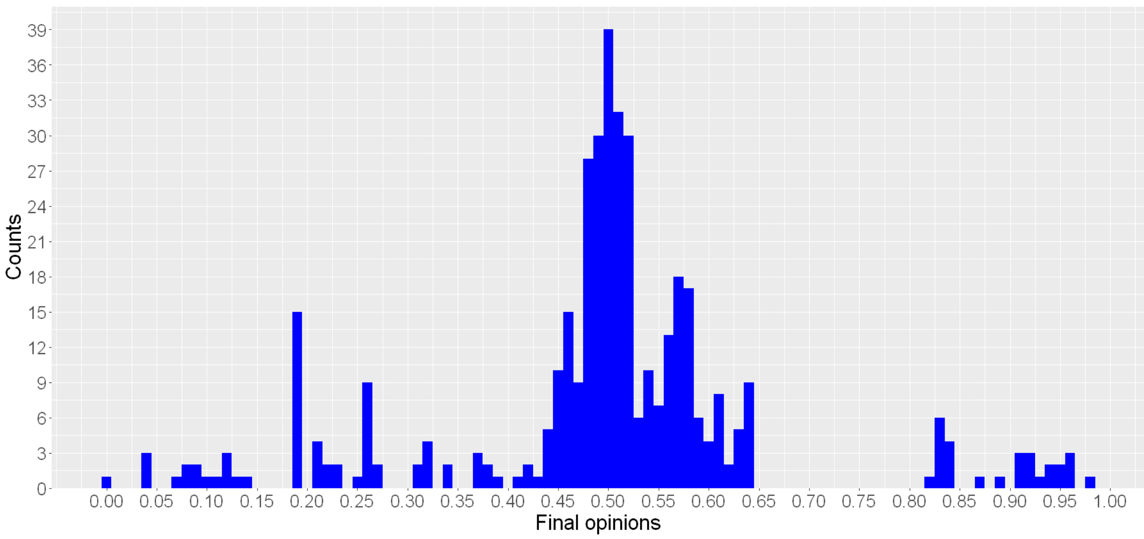


Figure 2.4: Histogram of final opinions

2.3 Vector opinions

In this section, a given agent i has binary opinions $x_i \in \{0, 1\}$ on m different subjects. Opinion readjustment occurs when two agents agree on at least $m - d$ subjects: if this is the case, all identical opinions are conserved, while differing opinions are changed with probability μ for one of the agents, chosen randomly.

We are thus interested to see whether convergence of opinions occurs also in this case. Simulations here were done by setting $N = 100$ and $m = 13$, while μ and d were free to vary. We note that changing μ only modifies the convergence times, while for d we have different behaviors:

- $d = 2$: almost no clusterization occurs for small times;
- $d = 3$: large number of small clusters is observed;
- $d = 4$: small number of large clusters, with a few isolated opinions. The distance of minor clusters from the main peak is around 6;
- $d = 5, 6, \ldots, 13$: convergence occurs.

3 | Granovetter's Bounded confidence model

Task leader(s): *Claudia Lorenzetti*

The Granovetter's model of collective behaviour deals with those situations where the actors have two different options (*to act or not to act*), mutually exclusive, and the decision of taking action depends on how many individuals have already joined it. A crucial concept in his model is the *threshold*: the number or proportion of others who must make one decision before a given actor does the same [3]. The innovation occurs in the importance that this model gives to the interaction and aggregation of individuals' needs and preferences: the decision that most of the individuals make is determined by an analysis of costs and benefits and it is not controlled by some *institutionalized norms and values*. Therefore this model can be used to describe situations where unexpected outcomes occur. It can be applied to many different situations: the number of people joining a riot, the diffusion of innovations, the spread of rumors, ecc. So, to sum up, a threshold model is an agent-based model where each individual has a binary decision to take, not only based on their needs and their preferences, but also strongly influenced by other's decisions. All those considerations conveyed in the *threshold*: that point where one's benefits become higher than the costs. Of course different agents will have different thresholds. Let's take as example the model of the riot: there will be people with a 0% or low values of thresholds, the *instigators*, and people with very high thresholds (like 90-100 %), the *conservatives*. But, what is more important, is that all the other people will have different values of thresholds that will follow a certain probability distribution. Therefore, the aim of this model is to foresee the ultimate proportion of people making each of the two decisions, by just taking into account the initial frequency distribution of thresholds. Mathematically, let $x \in [0, 1]$ be the thresholds, $f(x)$ their frequency distribution and $F(x)$ the cumulative distribution function, the percentage of population having threshold less or equal to x. So, the mathematical idea is:

$$r(t+1) = F[r(t)] \tag{3.1}$$

Where $r(t)$ is the proportion of the population activated at a certain time t. Therefore, an equilibrium has been reached when

$$r(t+1) = r(t) \tag{3.2}$$

We want to study how the model reacts to different simple probability distributions and how much the equilibrium point of a certain threshold distribution is vulnerable to perturbations. Then, we will do some implementations to the simple model. The first implementation considers as impossible the idea of describing a certain sociological model without taking into account that people interacts with each other and that the social structure has an effect in the outcome of a particular phenomenon. Returning to the example of the riot: it is more probable that a person joins the riot if a certain proportion of the rioters is their friends. Therefore, once we introduce this social structure of *friendship* we will see also if this will have an effect on the final outcome. Lastly, we will consider the *spatial effect*: it is really improbable that a person that lives in a place very far from the riot will have the chance to join it, despite their threshold. We will start to study this implementation, using the segregation model defined in Section 1: the segregation, in this case, will be spatial.

3.1 | Equilibrium outcomes in simple threshold models

Uniform distribution

Let us consider 100 people with the threshold distributed in a uniform way: one person has 0%, one person has 1%... one person has 99%, ecc. From such a distribution the equilibrium will be reached when everyone will join the riot, resulting in $r_e = 100\%$: this outcome is known as the *Band-wagon effect* 4.8. Now, if we perturb a little this distribution, by changing, for instance, the second element, we will have two people with 2% threshold, but no one with threshold 1%. The outcome will be completely different and the riot will not start(4.8). Note that the two situations are basically the same, except for one person: this small difference changes completely the equilibrium r_e. This kind of distribution in really unstable.

Normal distribution

Now we consider a normal thresholds' frequency distribution, which is the optimal choice for describing a realistic population where there is not a polarization of ideas. As in [3], let us take a normal distribution with mean $\mu = 25\%$. We want to study how much the equilibrium point changes by varying the standard deviation σ of the distribution.

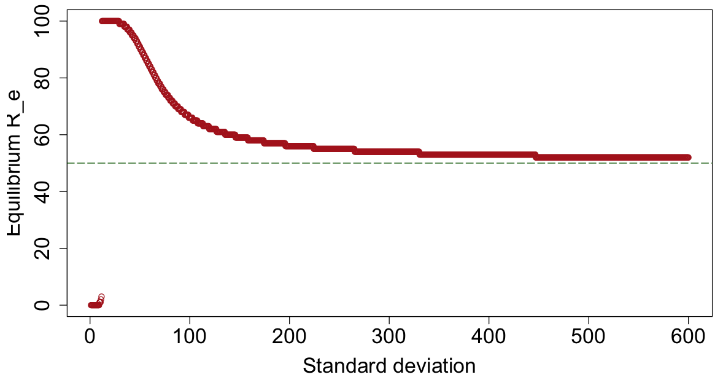


Figure 3.1: Change of the equilibrium point r_e in function of the standard deviation σ of the frequency distribution, $\mu = 25\%$

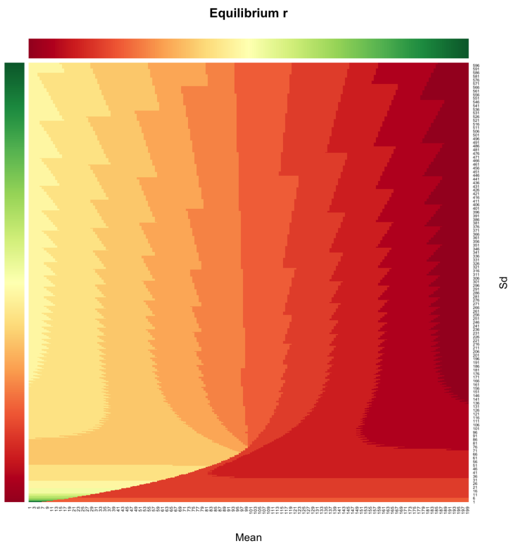


Figure 3.2: Heatmap for different values for the equilibrium point r_e by varying both the mean and the standard deviation

As we can see from the figure 3.1 the equilibrium increases in a constant way but with low values $\approx 0\%$ up until a certain value and then, when $\sigma \approx 12.2\%$ it jumps to 100%, then r_e decreases until it becomes more stable for high value of σ. The idea is that, by increasing the standard deviation, the distribution becomes more similar to the uniform one we have analyzed in 3.1.1. From the plot we can understand that those values close to $\sigma = 12.2$ are very unstable for the equilibrium point. In the figure 3.2 we can see that this unstable equilibrium point occurs for other values of the mean and, consequently, for other values of σ.

3.2 Social structures

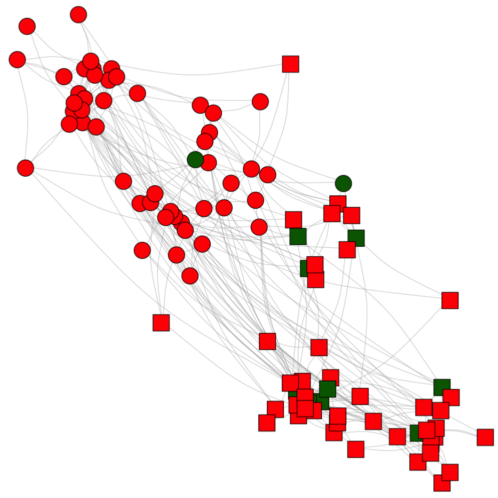


Figure 3.3: Graph of the Granovetter model originally obtained by applying a segregation process to simulate the distance between the people of the net

---

We generate a Gilbert net G(N,p) with N being the number of nodes and p being the probability that, given two nodes, those are connected. The social structure, we talk about *friendship* in this case (obviously any kind of social structure would be okay), is simulated by the edges. In fact, if two nodes (i, j) are connected by an edge, when i joins the riot, the node j will feel a greater influence due to their friendship. Let us consider a normal distribution for the thresholds. The value of r will be greater for the node j, if those two nodes are connected and one of them , i, is already active: the probability that its threshold $x[j]$ will be lower or equal to r increases, so j will have more chance to get activated. This situation will not affect so much the outcome if we are in a stable point, but will change the results around the critical value of r_e: of course the critical point will be a little bit lower, but the trend of r_e in function of σ seems basically the same (4.9 4.10, 4.11).

3.3 | Spatial segregation effects

Finally, we use the segregation model derived in Section[1] to discuss a little bit about the *spatial effect*. We consider the final graph obtained from this model as the starting point. From there, we apply the thresholds (normally distributed) and the edges have been considered as a relationship between to vertices. Also, the different shape of the vertices represents the segregation: two nodes with different shape are far one another, so the influence they have on each other is not so relevant. As it can be seen from 3.3, the green colour states that that vertex is active. We can notice how basically each *rioter* belongs to the squared edges: the spatial segregation seems to have certain effect on the evolution of the network.

4 | Annex

4.1 | Schelling segregation model

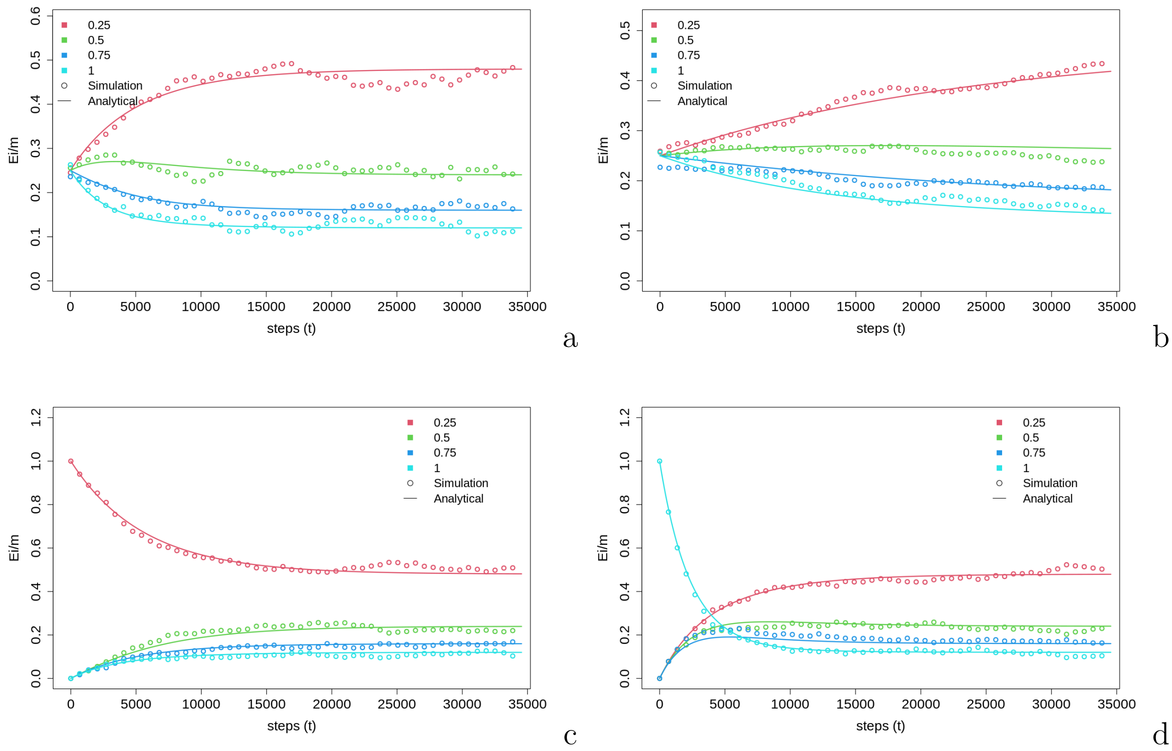


Figure 4.1: Comparison between analytical and simulated results. Simulated results belong to a network of $n = 300$, $m = 500$, random uniformly distributed attributes and (clearly) $K = 4$. a) Random $G(V, E)_{t_0}$ and $p = 0.5$. b) Random $G(V, E)_{t_0}$ and $p = 0.1$. c) Only short edges $G(V, E)_{t_0}$ and $p = 0.5$. d) Only long edges $G(V, E)_{t_0}$ and $p = 0.5$.

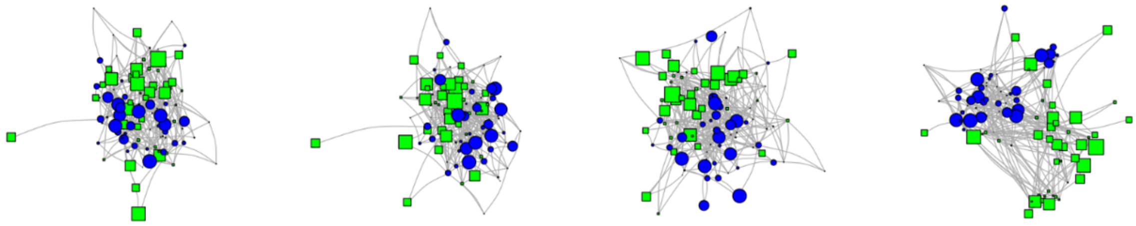


Figure 4.2: Social network evolution for $n = 300$, $m = 500$ and different p for every agent. The bigger the size the higher the aversion. From initial state (left) to stable state (right).

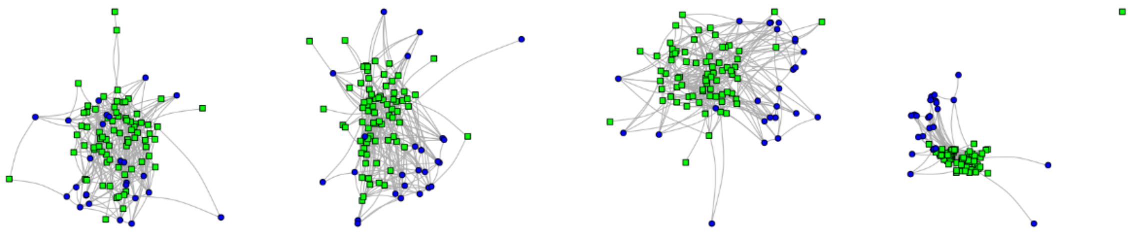


Figure 4.3: Social network evolution for $n = 300$, $m = 500$, same p for every agent and a three times bigger group. From initial state (left) to stable state (right).

4.2 | Deffaunt's bounded confidence model

4.2.1 Complete mixing

By plotting final opinions as a function of initial opinions, it results that for large μ there can be a significant overlap between distinct final opinions. This overlap disappears for smaller μ, as can be seen in figures **??** and **??**.

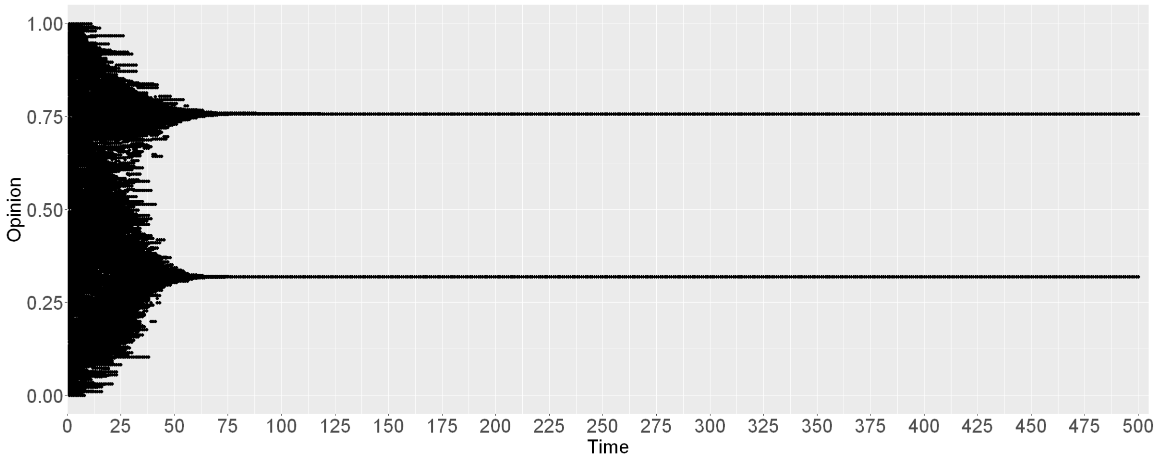


Figure 4.4: Converging opinions, for $d = 0.2$ and $\mu = 0.5$

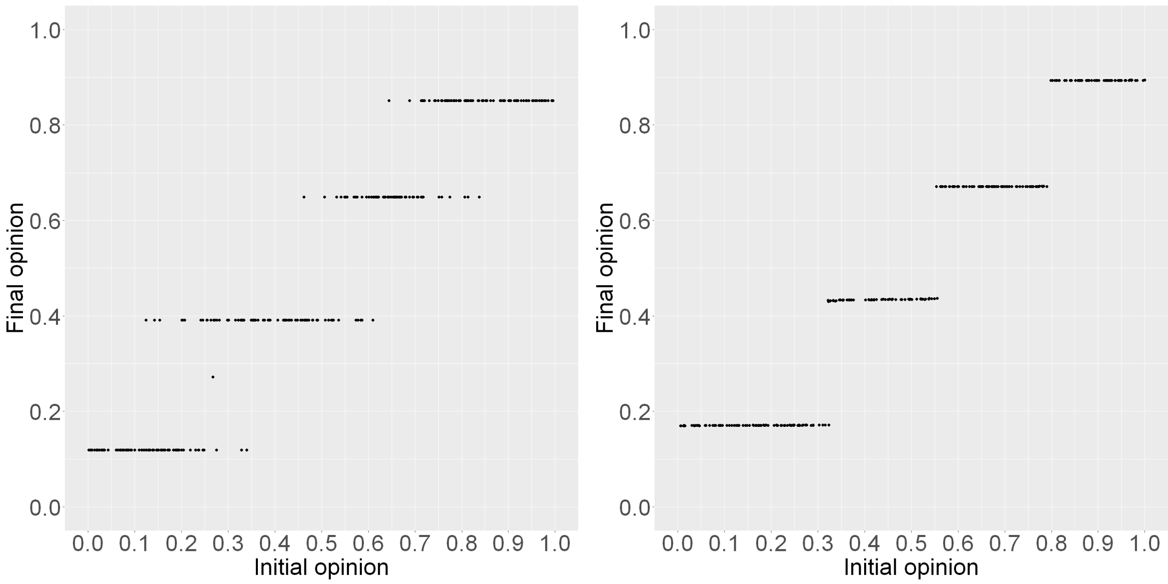


Figure 4.5: Final vs initial opinions with $\mu = 0.5$ (left) and $\mu = 0.01$ (right)

4.2.2 Social networks

Similar results are obtained by using randomly generated networks, according to the Barabasi-Albert model. We let vary M, the constant number of edges each incoming node links up with existing nodes. We see that for long times the same behavior is obtained, for $d = 0.25$: there is one percolating cluster of nodes which share the same mid-ranging opinion, while there are also localized and disjoint clusters of differently opinionated nodes. Again, the critical value of d which identifies a phase transition is around $d \simeq 0.3$.

Figure 4.6: Converging opinions on a network generated according to the Barabasi-Albert model; from left to right: $M = 1$, $M = 2$, $M = 3$

4.2.3 Vector opinions

We can also see how distant the peaks are located, in terms of the number of differing opinions. We can perform the simulation by varying the number of timesteps. From figure 4.7 we gather that the distance between clusters is around the diameter of the network, i.e. around half the value of m, which in this case corresponds to 6-7, for a small number of iterations. For longer times, the peaks tend to get closer to each other, as we can see from the fact that the histogram peak has shifted to the left, towards 0.

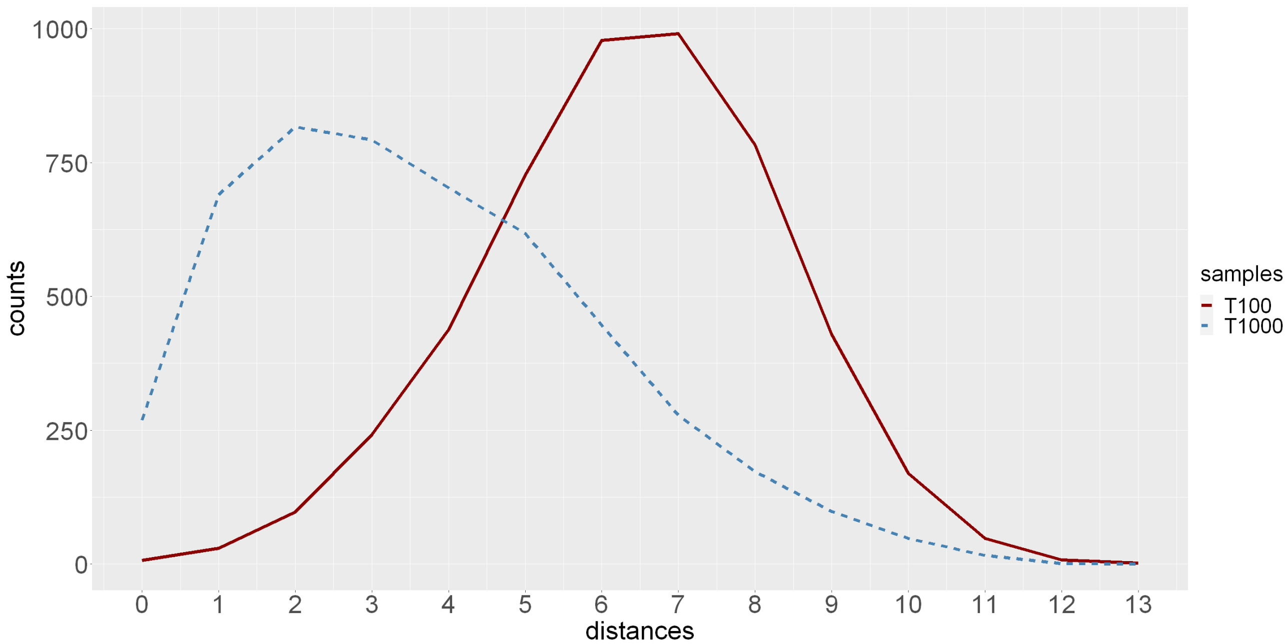


Figure 4.7: Histogram of distances between all possible pairs of nodes: red line - 100 time steps, blue dotted line - 1000 time steps

4.3 | Granovetter's bounded confidence model

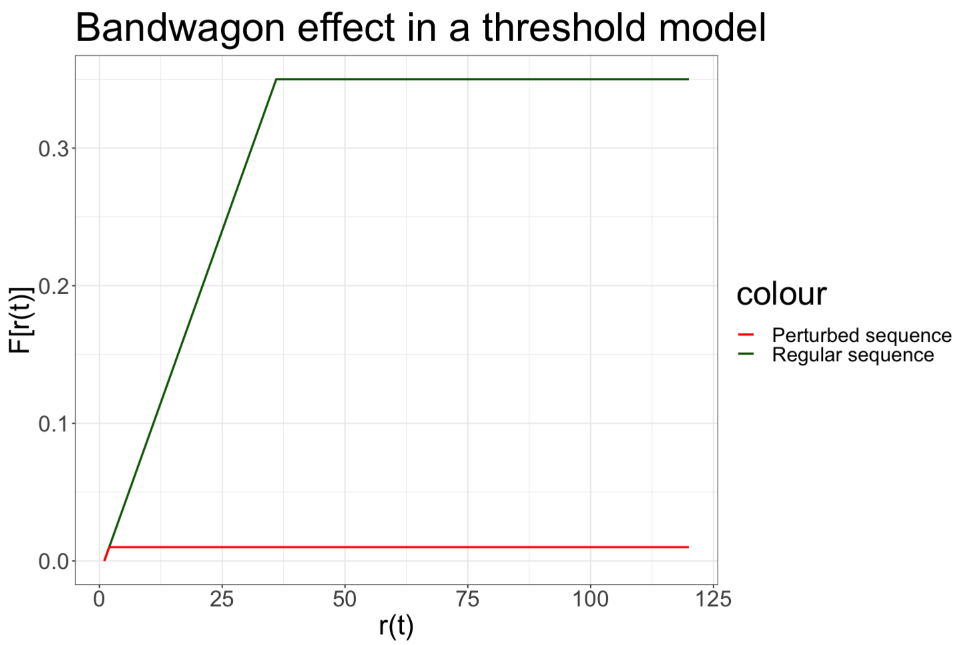


Figure 4.8: Plot of $r(t)$ for a uniform distribution of thresholds with and without perturbation

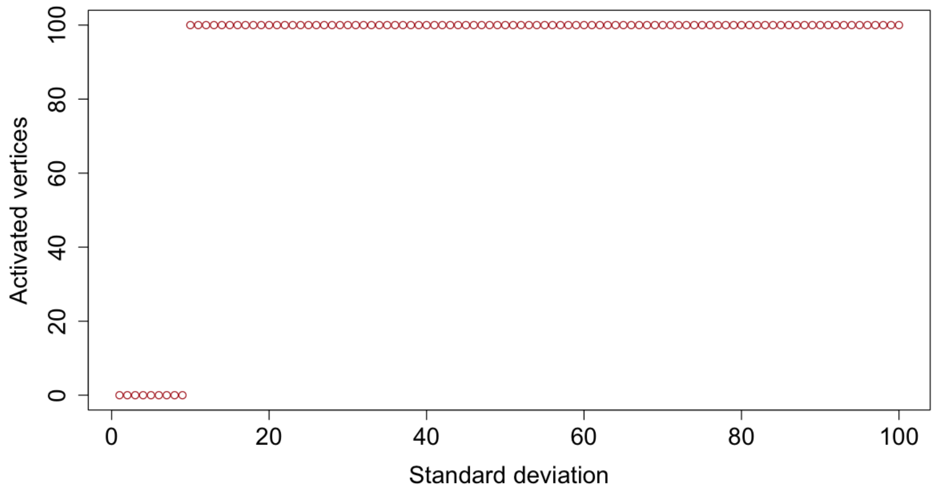


Figure 4.9: Plot of the equilibrium point in function of the standard deviation

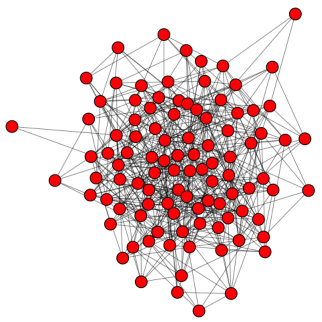


Figure 4.10: Plot of the Erdos-Renyi graph considering edges for $\mu = 25\%$ and $\sigma = 7\%$

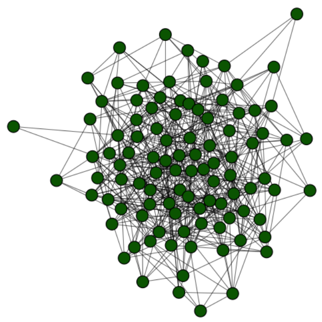


Figure 4.11: Plot of the Erdos-Renyi graph considering edges for $\mu = 25\%$ and $\sigma = 10\%$

5 | Bibliography

[1] Guillaume Deffuant, David Neau, Frédéric Amblard, and Gérard Weisbuch. Mixing beliefs among interacting agents. *Advances in Complex Systems*, 3:87–98, 01 2000. doi: 10.1142/S0219525900000078.

[2] Adam Douglas, Pawel Pralat, and Cun-quan Zhang. Emergence of segregation in evolving social networks. *Proceedings of the National Academy of Sciences of the United States of America*, 108:8605–10, 05 2011. doi: 10.1073/pnas.1014486108.

[3] Mark Granovetter. Threshold models of collective behaviour. *American journal of sociology*, pages 1420–1443, 1978.

[4] Thomas C. Schelling. Dynamic models of segregation. *The Journal of Mathematical Sociology*, 1(2):143–186, 1971. doi: 10.1080/0022250X.1971.9989794.