POLITECNICO DI TORINO

Corso di Laurea Magistrale in Ingegneria Matematica

Tesi di Laurea Magistrale

Learning models and the wisdom of crowds

Relatore:
Prof. Fabio Fagnani

Candidato:
Giulia Peluffo

Anno Accademico 2018/2019

Contents

Abstract

The wisdom of crowds is a sociological theory claiming that a mass of relatively inexperienced individuals may express a wiser behavior than a single one. In this thesis, we consider this issue in the context of the DeGroot learning model, assuming that the various agents are heterogeneous in their capability to access the true information. We study the effect of the network and of the self-confidence in the consensus value reached by the community. Particular focus is reserved to the case when the population is split into two classes, the regular people and the expert ones: we analyze to which extent the self-confidence of the experts can be beneficial to the entire community.

Introduction

The wisdom of crowds is a sociological theory claiming that a mass of relatively inexperienced individuals may express a wiser behavior than a single one. The spirit of the wisdom of crowds inspires many sites and projects on Internet, such as Wikipedia and Yahoo! Answers. It finds application even in the Sciences, as said Trisha Gura, *"Citizen science can help researchers to address previously insoluble problems"* [13].

The term *crowd* refers to a group of individuals not necessarily cohesive, who may not know each other or share the same ideas. According to James Surowiecki [16], there are four criteria that must be respected for the theory to work:

- **Diversity of opinion**, each person must have a different opinion;
- **Independence**, people's opinions should not be influenced by those of others;
- **Decentralization**, nobody must be able to control them;
- **Aggregation**, opinions must be able to be aggregated in order to obtain a final result.

According to Aristotle, who is considered to be the first one to mention the wisdom of the crowds, *"it is possible that the many, though not individually good men, yet when they come together may be better, not individually but collectively, than those who are so, just as public dinners to which many contribute are better than those supplied at one man's cost"* [1].

Since the beginning of the last century, several scientists have treated the theme of collective intelligence and the first one who talked about it was the statistician Francis Galton [11]. In 1906, at a cattle fair in Plymouth, people were asked about the weight of the exposed ox and writing it on a piece of paper. The most accurate in the prediction would have won the ox. About 800 individuals participated in the estimation competition. Each attendee had to buy stamped and numbered cards at a price of 6 pence each, on which they had to write their name, address and estimate. The six-penny fee discouraged jokes, while the hope of a prize and the joy of the competition drove each competitor to do their best. Competitors included very experienced people in judging the weight of livestock, such as butchers, farmers, and common citizens. Galton collected all the individual answers and found that the average of the estimates provided was more precise than the estimate provided by the individual experts.

Over time, the idea of the wisdom of crowds has been integrated in the learning models. An important model of network influence has been created thanks to Morris H. DeGroot. *"In the model that has been presented here, it is assumed that there is no possibility of learning whether the opinion of one individual is closer to the truth than that of another. In other words, it is assumed that no outside data, observations, or information about the value of truth is available. It is assumed that at the beginning, each individual i chooses the weights pij that he is going to use and he then continues to use these weights throughout the process."* [6].

In this work, we study the conditions in which DeGroot learning model holds the effect of wisdom of the crowds. These conditions concern the way the self-confidence connects to centrality. Accordingly, since groups are heterogeneous, we might exploit the presence of expert individuals. In fact, the use of social information is advantageous when individuals copy good performers. We focus on giving the right self-confidence to individuals so that learning leads to make a good decision. Initially, we make analitical studies, next, we find a way to determine agents' self-confidence, after they came to know the beliefs of neighbors.

The outline of the thesis is the following. In Chapter 1, we shortly present the graph theory needed in this thesis. Chapter 2 is mainly devoted to a formal description of the learning model and to a discussion of how to increase wisdom asymptotically. We present the details of the effect of adding self-weights, that represent the self-confidence of individuals. In Chapter 3, we analyze the effect of self-confidence on a social network where agents are split into two parts, expert and regular ones. Chapter 4 contains a set of examples on known networks where there are just some nodes with self-confidence, these examples verify what has been analytically addressed. In Chapter 5, we report some numerical simulations based on different graph architectures: Complete and Erdős-Rényi. Since nodes are human beings and not electronic devices, we try to simulate realistic scenarios. To this aim, we consider a social network where nobody knows its variability and, therefore, we estimate it based on the neighbors' responses. Finally, the last Chapter concludes the work, summarizing achieved results.

Chapter 1

Basic Concepts

In this chapter, we review all basic mathematical concepts needed in the formalization of our model. More precisely, we overview some basic concepts from topological and algebraic graph theory.

1.1 Graphs

Graphs are mathematical structures used to describe networks. A graph is made up of a non-empty set of nodes and a set of edges. There are three key aspects in modeling a network that are captured by the concept of graph:

1. The set of nodes $\mathcal{V}$ represents the units participating in the network, (e.g. people, biological entities, economic individuals, computers, sensors).

2. The set of edges $\mathcal{E}$ describes how nodes are linked. These connections are represented by a choice of ordered pairs sets (i, j), where $i, j \in \mathcal{V}$. In our case, the presence of a specific link (i, j) means that i can communicate with j.

3. In order to express in a quantitative way the activity level of a certain link (i, j) it is possible to associate a value W_{ij}, also called weight.

Definition 1.1. *A directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ consists in a pair of sets (always finite in this thesis) where the elements of $\mathcal{V}$,*

$$\mathcal{V} = \{v_1, ..., v_n\},$$

are called vertices (or nodes) of $\mathcal{G}$, while the elements of $\mathcal{E}$,

$$\mathcal{E} = \{e_1, ..., e_m\},$$

are called edges (or arcs) of $\mathcal{G}$. Each edge is a set of two ordered vertices $e = (v, w)$, $e \in \mathcal{E}$ and $v, w \in \mathcal{V}$.

We call it **undirected graph**, if the existence of the link (i, j) implies the existence of the link with reversed direction (j, i). For undirected graphs, one can consider undirected links described as unordered pairs $\{i, j\}$, each of which corresponds to the pair of directed links (i, j) and (j, i). Note that, by doing so, every undirected link corresponds to two directed links. In this work we use undirected graphs.
In general, in the graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, the set $\mathcal{V}$ is non-empty, whereas $\mathcal{E}$ might be empty. Of course, if $\mathcal{V} = \emptyset$, then also $\mathcal{E} = \emptyset$ and $\mathcal{G} = (\emptyset, \emptyset)$ is the trivial graph with no vertices and no edges.

Definition 1.2. *Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be an undirected graph. Two vertices $v, w \in \mathcal{V}$ are called adjacent if $(v, w) \in \mathcal{E}$, that is, if they are connected by an edge.*

Definition 1.3. *Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be an undirected graph. For a vertex $v \in \mathcal{V}$, its degree is the number of edges attached to v. Namely, $deg(v)$.*

The degree of each vertex represents the number of individuals with whom it is in communication.

1.1.1 Some graphs with special structures

Regular graph

Definition 1.4. *A regular graph is a graph where each vertex has the same number of neighbors, i.e., every vertex has the same degree.*

Definition 1.5. *A regular graph with vertices of degree r is called a $r-regular$ graph.*

It is well known that the necessary and sufficient conditions for a $r-regular$ *graph* of order n to exist are that $n \geq r+1$ and that nr is even. In Figure 1.1 it is possible to understand the creation of regular graphs by increasing their number of nodes and their degree. In particular, for any number of nodes a $0-regular$ *graph* is the empty graph, i.e., it consists of isolated nodes with no edges. A $1-regular$ *graph* consists of disconnected edges, and a $2-regular$ *graph* consists of one or more (disconnected) cycles.

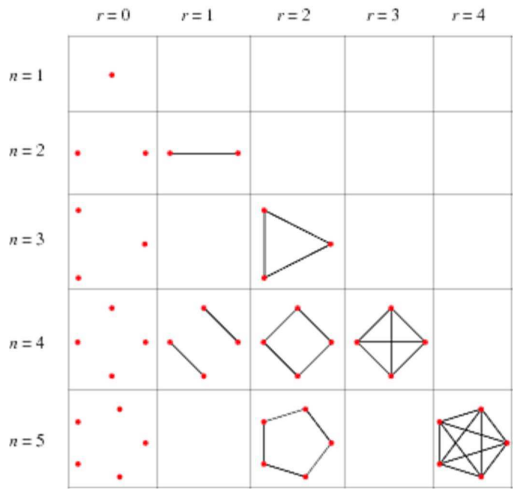


Figure 1.1: Regular graphs with n nodes ($1 \leq n \leq 5$) and with degree r ($0 \leq r \leq 4$).

Complete graph

Definition 1.6. *For every $n \geq 1$, we denote by K_n the complete graph with n vertices, that is, the graph with set of vertices $\{1, ..., n\}$ and all possible edges.*

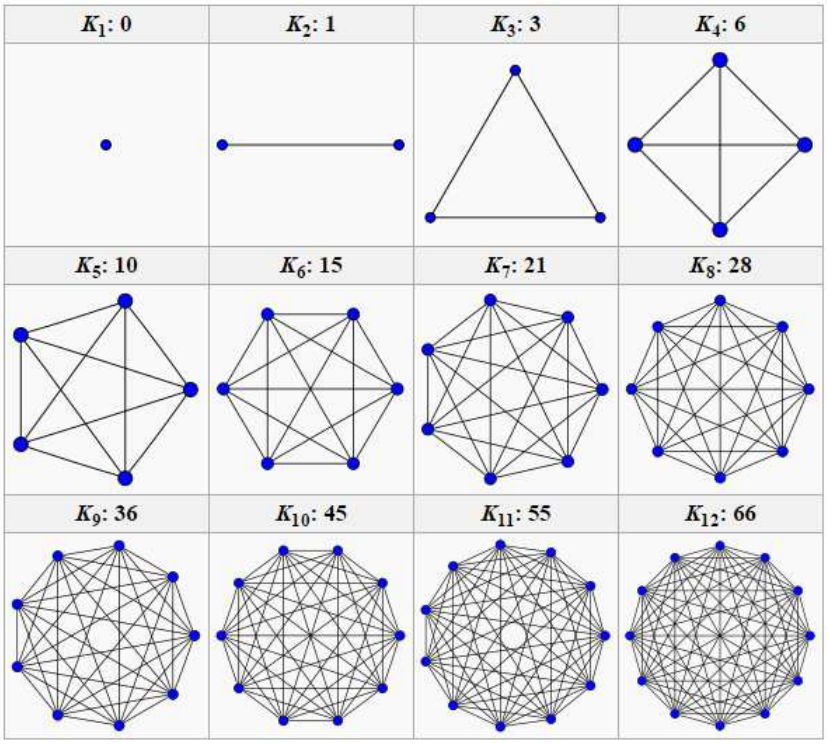


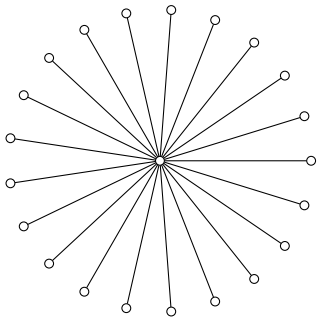
Figure 1.2: Complete graphs without self-loops with n vertices, $1 \leq n \leq 12$, and their respective number of edges.

Proposition 1.7. *Let $K_n = (\mathcal{V}, \mathcal{E})$ be the complete graph with n vertices. Then*

$$deg(v) = n - 1 \quad \forall v \in \mathcal{V}, \qquad |\mathcal{E}| = \binom{n}{2} = \frac{n(n-1)}{2}.$$

Here, we are considering a complete graph without self-loops, therefore the complete graph K_n is a regular graph of degree $n-1$. On the other hand, a complete graph K_n with self-loops is a regular graph of degree n. In Figure 1.2, it is possible to observe some examples of complete graphs without self-loops.

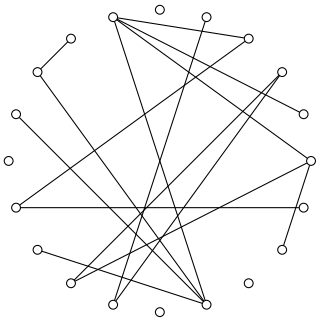
Star graph



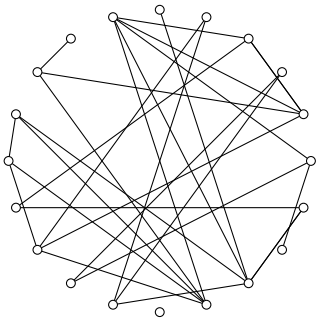
A star graph is a graph in which a node, called root, is adjacent to all the others, called leaves, and the latter are adjacent only to the center. A star with n nodes is normally referred to as S_n. In a star graph S_n, the root, labeled with index 1, has degree $(n-1)$, while the leaves $\{2, \ldots, n\}$ have degree 1.

Erdős-Rényi graph

An Erdős-Rényi graph (ER Graph), $\mathcal{G} = (n, p)$, is a random graph with n nodes in which each arc has probability p to exist. Therefore, the average number of edges is given by $\frac{n(n-1)}{2}p$. It is possible to observe that n and p cannot be univocally deduced by the graph which is the result of a random process [8]. Consequently, many different outputs may be obtained given the same values of n and p. The degree of connection in these types of graphs depends on the parameter p.



(a) $p = 0.1$



(b) $p = 0.2$

Firstly, let us consider the borderline cases:

- $p = 0$, every node has degree equal to 0, thus all nodes are isolated,
- $p = 1$, every node has degree equal to $n - 1$, thus the network is a complete graph.

Secondly, in [8], it has been proven that in order to not have isolated nodes, p must be

$$p > \frac{(1+\varepsilon)\ln n}{n}.$$

Thus $\frac{\ln n}{n}$ is a sharp threshold for the connection of $\mathcal{G} = (n, p)$.

1.1.2 The adjacency matrix

Definition 1.8. *Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a graph with $\mathcal{V} = \{1, ..., n\}$, where $n = |\mathcal{V}|$. The $n \times n$ matrix $A = (a_{ij})$ where*

$$a_{ij} = \begin{cases} 1 & \text{if } (i,j) \in \mathcal{E} \\ 0 & \text{if } (i,j) \notin \mathcal{E} \end{cases}$$

is called the adjacency matrix of $\mathcal{G}$.

The element a_{ii} on the diagonal can be $a_{ii} = 0$ if node i does not have a self-loop, otherwise $a_{ii} = 1$ if it has a self-loop. Since we consider undirected graphs, it is immediate to check that:

1. the adjacency matrix A is symmetric;
2. the structure of the graph $\mathcal{G}$ is completely determined by A (because, from A, one can reconstruct the edges of $\mathcal{G}$).

If a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ has n vertices but its vertices are not labelled as the integers $\{1, ..., n\}$ (as in our case which are individuals), it is always possible to relabel them in this way, and still consider the adjacency matrix. Of course, the resulting adjacency matrix depends on how we have labelled the vertices. Consequently,

the matrices that we obtained are similar and it is possible to get one from the other by applying a permutation.

Example 1.1. *The adjacency matrix of the complete graph K_n, without self-loops, is given by*

$$\begin{pmatrix} 0 & 1 & \dots & \dots & \dots & 1 \\ 1 & 0 & 1 & \dots & \dots & 1 \\ \vdots & 1 & \ddots & \ddots & & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & & \ddots & \ddots & 1 \\ 1 & 1 & \dots & \dots & 1 & 0 \end{pmatrix}$$

It is a $n \times n$ matrix with 1 everywhere except along the main diagonal where all entries are zero.

Example 1.2. *The adjacency matrix of the star graph S_n, where the root is labeled by 1, is given by*

$$\begin{pmatrix} 0 & 1 & 1 & \dots & \dots & 1 \\ 1 & 0 & 0 & \dots & \dots & 0 \\ 1 & 0 & \ddots & \ddots & & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & & \ddots & \ddots & 0 \\ 1 & 0 & \dots & \dots & 0 & 0 \end{pmatrix}$$

It is a $n \times n$ matrix with all 1 in the first row and in the first column except the term $(1,1)$ and all the others entries are zero.

1.1.3 Weighted graphs

A weighted graph is defined as a triple $(\mathcal{V}, \mathcal{E}, W)$, where $\mathcal{V}$ is the set of nodes, $\mathcal{E}$ is the set of links and $W \in \mathbb{R}_+^{\mathcal{V}\times\mathcal{V}}$ is the weigth matrix. In many applications, each edge may have an associated numerical value, called weight. Therefore, the

weight matrix W is a generalization of the adjacency matrix A described above. Consequently, the weight matrix is then a non-negative square matrix whose entries satisfy $W_{ij} > 0$ if $(i,j) \in \mathcal{E}$ and $W_{ij} = 0$ if $(i,j) \notin \mathcal{E}$. In many situations, it is convenient to identify $\mathcal{V} = \{1, ..., n\}$ so that $W \in \mathbb{R}_+^{n \times n}$. Let $\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$ be a graph, as said before, we focus on undirected graphs in which $W' = W$, i.e., each link and the one with reverse direction have the same weight $W_{ij} = W_{ji}$. If $W_{ij} \in \{0, 1\}$ for all nodes $i, j \in \mathcal{V}$, the graph $\mathcal{G}$ is called unweighted. In this case the graph is often described by the pair $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. The matrix W, which can be univocally deduced from the set $\mathcal{E}$, it is simply the adjacency matrix A of $\mathcal{G}$.

Besides the weight/adjacency matrix, other matrices are associated to a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$, for instance the normalized weight matrix P, also called transition matrix. The normalized weight matrix is obtained as follows

$$P = D^{-1}W, \tag{1.1}$$

where $D = diag(w)$ and $w = W\mathbb{1}$. All entries of P are non-negative, thus P is a non-negative matrix. Since $P\mathbb{1} = \mathbb{1}$, P is a stochastic matrix[1]. Any stochastic matrix P over a set $\mathcal{V}$, can be thought as the normalized weight matrix of a graph. It is sufficient to consider $\mathcal{G}_P = (\mathcal{V}, \mathcal{E}, P)$ where $\mathcal{E} = \{(i,j) | P_{ij} > 0\}$. We call $\mathcal{G}_P$ the graph associated with P. Sometimes we use the following terminology: P is called irreducible if $\mathcal{G}_P$ is strongly connected, and it is called aperiodic if $\mathcal{G}_P$ is such.

Theorem 1.9. *Let W be a non-negative matrix. There exists an eigenvalue λ_W of W real and positive and two non-negative vectors $x \neq 0$ and $y \neq 0$ such that*

1. *$Wx = \lambda_W x$, $W'y = \lambda_W y$;*

2. *every eigenvalue μ of W is such that $|\mu| \leq \lambda_W$.*

[1] A matrix is said to be *stochastic* whether it is a square matrix where all entries are non-negative and the sum of entries in each row is 1.

The eigenvalue λ_W is called the dominant eigenvalue of W. Important properties of stochastic matrices are shown below.

Proposition 1.10. *Let $\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$ be a graph and let $P = D^{-1}W$ be its normalized weight matrix. Then*

1. *$\lambda_P = 1$;*

2. *there exists a non-negative vector π such that $\pi'\mathbb{1} = 1$ and $P'\pi = \pi$.*

Proof. See Lecture notes on Dynamics over Networks, [9].

□

The probability distribution $\pi = (\pi_1, \ldots, \pi_n)$ is the invariant measure of P, besides it denotes its centrality.

1.2 Social network

A social network consists of any group of individuals connected by different social ties. For humans, ties range from casual knowledge, to working relationships, family ties and so on. The analysis of social networks, i.e., the mapping and measurement of social networks, can be conducted with a mathematical formalism by graph theory. In fact, a social network is a graph where the nodes represent people and the edges represent social connections between them. In the following of this thesis the term social network will be used to denote the graph of social interactions through which agents communicate. A social network, defined as a graph, is characterized by its density. The density of a network can give an idea of how efficient is the relational interchange between the various elements of the network. If all the elements of the network establish links between them, then the network densities would be equal to one (complete graph). Otherwise, if the network elements did not communicate with each other, the network density would be zero (null graph). The analysis of social networks has shown how structure and

density contribute to determining its potential utility for individuals. Small and dense networks may sometimes prove to be less useful than larger networks and with weak links. In fact, the latter would lend themselves more to the exchange of new ideas and opportunities, thus favoring the processes of innovation. To simulate how a social network forms, mathematicians use random graphs that model how people make connections as they enter the network. In fact, in some of our simulation we use the simplest type of random graph, the Erdős-Rényi graph.

Chapter 2

The DeGroot Model

Opinion dynamics study the way opinions of individuals modify through pairwise or group interactions. A celebrity model considered in opinion dynamics is the DeGroot model. Assume a community of people $\mathcal{V} = \{1, \dots, n\}$ to be connected to a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$, where W is the interaction matrix. As seen in section 1.1.3, the corresponding stochastic matrix P is obtained by the formula $P = D^{-1}W$. Furthermore, assume that the community has to estimate a parameter. Each individual can state its own estimation $x_i(0)$ of the parameter, where i denotes the individual and the value in brackets represents the time of estimation/revision (in this case $t = 0$). The update process is based on the weighted average of the neighbors' beliefs at the previous time. The updating rule of the DeGroot Model is the following

$$X(t+1) = PX(t), \tag{2.1}$$

where $X(t) \in \mathbb{R}^n$ is the beliefs vector at time t and P is the interaction matrix. The matrix P is a $n \times n$ non-negative matrix and it is stochastic, so that the sum of its entries in each row is equal to one. In particular, p_{ij} is the trust that agent i places on agent j. In the learning process, each member's revision is a linear combination that takes into account the latest changes in the opinion of oneself

and others. By iterating the process (2.1) we get

$$X(t) = PX(t-1) = P^t X(0). \tag{2.2}$$

It is assumed that individuals continue to make these revisions indefinitely or until $X(t+1) = X(t)$ for some value t, i.e., further revisions no longer lead to changes. The initial opinion,

$$x_i(0) = \mu + N_i, \qquad \forall i \in \mathcal{V}, \tag{2.3}$$

is defined as the sum of the truth μ and a noisy term N_i, which we assume is normally distributed with mean 0 and variance $\sigma_i^2 \geq 0$. Noisy terms are independent one from the other.

We start to define the average group belief because it is a fundamental element in the comparison between the wisdom of the crowds and the learning model; at the time t it is given by

$$m(t) = \frac{1}{n}\sum_{i=1}^{n} x_i(t),$$

and at time $t = 0$, we have

$$m(0) = \frac{1}{n}\sum_{i=1}^{n} x_i(0) = \mu + \frac{1}{n}\sum_{i=1}^{n} N_i. \tag{2.4}$$

We notice that $\frac{1}{n}\sum_{i=1}^{n} N_i$ is normally distributed with mean 0 and variance

$$\sigma_{ave}^2 = \frac{1}{n^2}\sum_{i=1}^{n} Var(N_i) = \frac{1}{n^2}\sum_{i=1}^{n} \sigma_i^2.$$

2.1 The asymptotic behavior

In a society, a consensus is reached if and only if the value of all n members converges to the same limit as $t \to \infty$.

Definition 2.1. *Let $X(t) \in \mathbb{R}^n$ be the opinions vector. The estimations of the n individuals converge to each other if and only if there is a value x^* such that*

$$\lim_{t\to\infty} x_i(t) = x^*, \quad \forall i = 1, ..., n. \tag{2.5}$$

It follows from (2.2) that a consensus is reached if and only if there exists an invariant vector $\pi = (\pi_1, ..., \pi_n)$ such that, for $i = 1, ..., n$ and $j = 1, ..., n$

$$\lim_{t\to\infty} p_{ij}^{(t)} = \pi_j, \tag{2.6}$$

where $p_{ij}^{(t)}$ denotes the element in row i and column j of the matrix P^t. When (2.6) is satisfied for every value of i and j, then $\pi_1, ..., \pi_n$ are non-negative and such that $\sum_{i=1}^{n} \pi_i = 1$.

Assume P to be irreducible and aperiodic and let us consider the invariant probability vector π related to the stochastic matrix P

$$\pi' P = \pi'$$

where $\sum_{i=1}^{n} \pi_i = 1$ and π denotes the *centrality* of P, see Proposition 1.10. Then, if $X(t)$ is the solution of the DeGroot Model (2.1), we have that

$$\lim_{t\to+\infty} X(t) = \mathbb{1}(\pi' X(0)),$$

this means that the beliefs vector $X(t)$ converges to $\mathbb{1}(\pi' X(0)) = \mathbb{1}\bar{X}_\infty$ when $t \to \infty$. More specifically we get

$$\bar{X}_\infty = \pi' X(0) = \mu + \sum_{i=1}^{n} \pi_i N_i.$$

We want to analyze the difference between the average group belief at time $t = 0$ and at time $t = \infty$:

$$m(0) = \mu + \frac{1}{n} \sum_{i=1}^{n} N_i,$$
$$m(\infty) = \mu + \sum_{i=1}^{n} \pi_i N_i.$$

We compare the wisdom of crowds with the asymptotic behavior. In addition, we notice that $\mathbb{E}(m(0)) = \mathbb{E}(m(\infty)) = \mu$. We evaluate the distance with respect to

the truth value μ by using the variance[1]:

$$\mathbb{E}|\mu - m(0)|^2 = \frac{1}{n^2}\sum_{i=1}^{n}\sigma_i^2,$$
$$\mathbb{E}|\mu - m(\infty)|^2 = Var\left(\sum_{i=1}^{n}\pi_i N_i\right) = \sum_{i=1}^{n}\pi_i^2\sigma_i^2.$$

To sum up, the first value represents the variance of the initial error and the second one represents the variance of the asymptotic error.

2.2 Wisdom of crowds and wise societies

When we talk about wisdom of crowds, we refer to a collective intelligence. Many social scientists describe the wisdom of crowds as the phenomenon for which *"large groups of people are smarter than an elite few"* [16]. It does not matter how clever people are or how much they know about the specific topic, because the group, rather than the individual, is always better at solving problems, making wise decisions or predicting the future. As reported by Golub and Jackson, *"A society is wise if and only if the influence of the most influential agent is vanishing as the society grows"* [12]. Let us consider a society as a sequence of graphs with increasing dimension n and assume that the true state μ is kept constant.

Definition 2.2. *(Wisdom notion). Given a sequence of stochastic matrices of increasing dimensions $\{P^{[n]} \in \mathbb{R}^{n\times n}\}_{n\in\mathbb{N}}$, define a sequence of opinion dynamics problems with initial state $\{X^{[n]}(0) \in \mathbb{R}^n\}_{n\in\mathbb{N}}$ satisfying (2.3). The sequence $\{P^{[n]} \in \mathbb{R}^{n\times n}\}_{n\in\mathbb{N}}$, is wise if*

$$\operatorname*{plim}_{n\to+\infty} \pi'^{[n]} X^{[n]}(0) = \mu. \tag{2.7}$$

In this work plim is meant as probability limit.

Assume $0 < \underline{\sigma}^2 \leq \sigma_i^2 \leq \bar{\sigma}^2 < +\infty$, $\forall i \in \mathcal{V}$. Definition 2.2 is equivalent to say that

[1] By definition $Var(X) = \mathbb{E}\big[(X - \mathbb{E}[X])^2\big]$.

the limiting belief of all agents converges to the truth as $n \to +\infty$ if and only if the influence of the most influential agent vanishes.

Proposition 2.3. *Consider a sequence of stochastic matrices of increasing dimensions* $\{P^{[n]} \in \mathbb{R}^{n\times n}\}_{n\in\mathbb{N}}$. *The sequence is wise if and only if*

$$\operatorname*{plim}_{n\to+\infty} \max_i \pi_i^{[n]} = 0.$$

Proof. By Definition 2.2, $\operatorname{plim}_{n\to+\infty} \pi'^{[n]} X^{[n]}(0) = \mu$, where $\pi'^{[n]} X^{[n]}(0) = \mu + \pi'^{[n]} N^{[n]}$. Therefore, we have to prove that

$$\operatorname*{plim}_{n\to+\infty} \pi'^{[n]} N^{[n]} = 0. \tag{2.8}$$

Let $\pi'^{[n]} N^{[n]} = \sum_{i=1}^n \pi_i^{[n]} N_i^{[n]}$. Then $Var\left(\sum_{i=1}^n \pi_i^{[n]} N_i^{[n]}\right) = \sum_{i=1}^n \pi_i^{2[n]} \sigma_i^2$.
First, suppose $\max_i \pi_i^{[n]} \to 0$.

$$\begin{aligned} Var\left(\sum_{i=1}^n \pi_i^{[n]} N_i^{[n]}\right) &= \sum_{i=1}^n \pi_i^{2[n]} \sigma_i^2 \\ &\le \sum_{i=1}^n (\max_i \pi_i^{[n]}) \pi_i \sigma_i^2 \\ &\le (\max_i \pi_i^{[n]}) \bar{\sigma}^2 \underbrace{\sum_{i=1}^n \pi_i}_{=1} \\ &= (\max_i \pi_i^{[n]}) \bar{\sigma}^2 \to 0, \qquad \forall n \in \mathbb{N} \end{aligned}$$

By Chebychev's inequality, fixing any $\epsilon > 0$, it is proven (2.8).

For the converse, suppose $\max_i \pi_i^{[n]} \not\to 0$.

$$Var\left(\sum_{i=1}^n \pi_i^{[n]} N_i^{[n]}\right) \ge (\max_i \pi_i^{[n]})^2 \sigma_i^2 \ge (\max_i \pi_i^{[n]})^2 \underline{\sigma}^2 \not\to 0,$$

which means that the $\sum_{i=1}^n \pi_i^{[n]} N_i^{[n]}$ can not converge to 0 in probability.

□

When noisy terms are equally distributed, $\sigma_i^2 = \sigma^2$, the variance of the asymptotic error is thus

$$Var\left(\sum_{i=1}^{n} \pi_i N_i\right) = \sigma^2 \sum_{i=1}^{n} \pi_i^2.$$

It is well known that if there exists an invariant vector π, then $\pi_i > 0$ for all i and thus $\sum_{i=1}^{n} \pi_i^2 < \sum_{i=1}^{n} \pi_i = 1$. Using this consideration as a starting point, it may be observed

$$Var\left(\sum_{i=1}^{n} \pi_i N_i\right) < \sigma^2,$$

namely, the crowd is wiser than a single.

The wisdom we have defined above refers to the asymptotic behavior of a crowd. It is possible to talk about wisdom even in the finite time.

2.3 Increasing wisdom

Given the variances σ_i^2 we are interested in analyzing for which π it holds that

$$\sum_{i=1}^{n} \pi_i^2 \sigma_i^2 < \frac{1}{n^2} \sum_{i=1}^{n} \sigma_i^2, \tag{2.9}$$

so that the learning process improves the collective estimation of the truth value μ.

We start by analyzing the following minimum problem:

$$arg \min_{\substack{\pi:\pi_i \geq 0 \\ \sum_{i=1}^{n} \pi_i = 1}} \sum_{i=1}^{n} \pi_i^2 \sigma_i^2. \tag{2.10}$$

In order to solve this problem we use the method of Lagrange multipliers. Accordingly, the Lagrange function is defined by

$$\Lambda(\sigma_1^2, ..., \sigma_n^2, \lambda) = \pi_1^2 \sigma_1^2 + ... + \pi_n^2 \sigma_n^2 \ + \lambda(\pi_1 + ... + \pi_n - 1),$$

and, the partial derivatives are

$$\begin{cases} \dfrac{\partial \Lambda}{\partial \pi_i} = 2\pi_i \sigma_i^2 + \lambda, & \forall i \quad i = 1, ..., n \\ \dfrac{\partial \Lambda}{\partial \lambda} = \pi_1 + ... + \pi_n - 1 \end{cases}$$

By setting the gradient to zero we obtain the following system:

$$\begin{cases} \dfrac{\partial \Lambda}{\partial \pi_i} = 0 \qquad \forall i,\ i = 1, ..., n \qquad \Leftrightarrow \qquad \pi_i = -\dfrac{\lambda}{2\sigma_i^2} \qquad \forall i,\ i = 1, ..., n \\ \dfrac{\partial \Lambda}{\partial \lambda} = 0 \end{cases}$$

which amounts to solving $n + 1$ equations in $n + 1$ unknowns. By replacing $\pi_i = -\frac{\lambda}{2\sigma_i^2}$ in the original constraint we find λ

$$\left(\sum_{i=1}^{n} -\frac{\lambda}{2\sigma_i^2} \right) - 1 = 0 \qquad \Leftrightarrow \qquad \frac{1}{\lambda} = -\sum_{i=1}^{n} \frac{1}{2\sigma_i^2}.$$

To minimize $\sum_{i=1}^{n} \pi_i^2 \sigma_i^2$, each probability π_i must be inversely proportional to its variance:

$$\pi_i = \frac{1}{\sigma_i^2} \left(\frac{1}{\sum_{j=1}^{n} \frac{1}{\sigma_j^2}} \right), \qquad \forall i,\ i = 1, ..., n. \tag{2.11}$$

This is the π achieving the minimum variance.

We now analyze in more detail the improvement of the wisdom of crowds starting from the simple case $n = 2$. In this case, the inequality (2.9) corresponds to

$$\pi_1^2 \sigma_1^2 + \pi_2^2 \sigma_2^2 < \frac{1}{4}(\sigma_1^2 + \sigma_2^2), \tag{2.12}$$

where $\pi_1 + \pi_2 = 1$.

Assuming that $\sigma_1^2 > \sigma_2^2$ and $\pi_1 < \pi_2$, let us take $\pi_1 = (\frac{1}{2} - x)$ and $\pi_2 = (\frac{1}{2} + x)$. Therefore, (2.12) is as follows

$$(\frac{1}{2} - x)^2 \sigma_1^2 + (\frac{1}{2} + x)^2 \sigma_2^2 < \frac{1}{4}(\sigma_1^2 + \sigma_2^2)$$

$$x^2(\sigma_1^2 + \sigma_2^2) - x(\sigma_1^2 - \sigma_2^2) < 0.$$

It is a convex parabola, which is null in $x = 0$ and $x = \frac{\sigma_1^2 - \sigma_2^2}{\sigma_1^2 + \sigma_2^2}$. It is thus clear that (2.12) holds true if $0 < x < \frac{\sigma_1^2 - \sigma_2^2}{\sigma_1^2 + \sigma_2^2}$.

For the general case, we have the following partial result.

Proposition 2.4. *Let σ_i^2 be ordered as $\sigma_1^2 \geq \cdots \geq \sigma_n^2$. Then, $\forall \pi$ probability such that*

$$\pi_i \leq \pi_{i+1} \tag{2.13}$$

$$\pi_i \sigma_i^2 \geq \pi_{i+1} \sigma_{i+1}^2, \tag{2.14}$$

it holds,

$$\sum_{i=1}^{n} \pi_i^2 \sigma_i^2 \leq \frac{1}{n^2} \sum_{i=1}^{n} \sigma_i^2. \tag{2.15}$$

Proof. Let us consider the following functional

$$F(\pi) = \frac{1}{n^2} \sum_{i=1}^{n} \sigma_i^2 - \sum_{i=1}^{n} \pi_i^2 \sigma_i^2.$$

We need to show that for any probability π satisfying (2.13) and (2.14) we have that $F(\pi) \geq 0$. This derives from the following considerations. Given $i > j$ we have that

$$\frac{\partial F(\pi)}{\partial \pi_i} - \frac{\partial F(\pi)}{\partial \pi_j} = 2\big[\pi_j \sigma_j^2 - \pi_i \sigma_i^2\big].$$

We fix π satisfying (2.13) and (2.14). Starting from $\pi^0 = (1/n, 1/n, \ldots, 1/n)$ we can imagine to move towards the direction of π by movements of type $x(e_i - e_j)$ where $i > j$ and $x > 0$. The partial derivative condition insures that as long as conditions (2.14) are not violated, the functional F will increase along these movements. This yields the result.

□

2.4 The effect of adding Self-Weights

We now modify the model introducing a new parameter $\alpha \in [0,1]^n$. The parameter α_i is called self-weight and it represents the weight that the individual i places on its own belief. Using self-weights, the original transition matrix P becomes the following

$$P_\alpha = diag(\alpha) + (I - diag(\alpha))P. \tag{2.16}$$

The new centrality vector $\pi'_\alpha = \pi'_\alpha P_\alpha$ can be computed as follows:

$$\pi'_\alpha = \pi'_\alpha diag(\alpha) + \pi'_\alpha (I - diag(\alpha))P,$$
$$\underbrace{\pi'_\alpha (I - diag(\alpha))}_{\pi'} = \underbrace{\pi'_\alpha (I - diag(\alpha))}_{\pi'} P.$$

Therefore,

$$\pi'_\alpha (I - diag(\alpha)) = \pi' \lambda,$$

for some positive scalar λ. By carrying out calculations with respect to a single component i:

$$(\pi_i^\alpha)(1 - \alpha_i) = \pi_i \lambda,$$
$$\pi_i^\alpha = \frac{\pi_i}{1 - \alpha_i} \lambda.$$

By summing on both sides we can calculate λ as follows:

$$1 = \left(\sum_{i=1}^{n} \frac{\pi_i}{1 - \alpha_i} \right) \lambda,$$
$$\lambda = \frac{1}{\sum_{i=1}^{n} \frac{\pi_i}{1 - \alpha_i}}.$$

The value of the new centrality vector is thus given by

$$\pi_i^\alpha = \frac{\dfrac{\pi_i}{1 - \alpha_i}}{\displaystyle\sum_{j=1}^{n} \frac{\pi_j}{1 - \alpha_j}}. \tag{2.17}$$

We notice that any new possible centrality vector $\tilde{\pi}$ can be achieved as follows. To obtain that $\pi_i^\alpha = \tilde{\pi}_i$ for every i, it is sufficient to put

$$\alpha_i = 1 - \frac{\pi_i / \tilde{\pi}_i}{\max_j \pi_j / \tilde{\pi}_j}.$$

In particular, setting

$$\alpha_i = 1 - \frac{\pi_i \sigma_i^2}{\max_j \pi_j \sigma_j^2}, \tag{2.18}$$

we obtain the optimal choice (2.11).

Chapter 3

Social network with experts. The effect of self-confidence

De Marzo, Vayanos, and Zweibel [7] note that a rational actor should place more weight on those estimates which it considers to be more reliable. Thus, if agents have information about the accuracy of their own estimates, therefore the self-weight is correlated with accuracy. For this reason, we want to analyze the effect of self-confidence on the wisdom.

We now assume the population to be split into two parts $N = R \cup E$ where E are the expert nodes, while R are the regular ones. We assume that $\sigma_i^2 = \sigma_R^2$ for every $i \in R$, while $\sigma_i^2 = \sigma_E^2 < \sigma_R^2$ for every $i \in E$. We also assume that there are n_1 expert nodes and n_2 regular ones, where $n_1 + n_2 = n$. In this case the minimum problem (2.10) becomes

$$\begin{aligned} arg \min_{\substack{\pi:\pi_i \geq 0 \\ \sum_{i=1}^n \pi_i = 1}} & \sum_{i=1}^{n} \pi_i^2 \sigma_i^2 \\ = arg \min_{\substack{\pi:\pi_i \geq 0 \\ \sum_{i=1}^n \pi_i = 1}} & \sum_{i=1}^{n_1} \pi_i^2 \sigma_E^2 + \sum_{i=n_1+1}^{n} \pi_i^2 \sigma_R^2, \end{aligned}$$

and the solution takes the form of

$$\begin{aligned} \pi_i &= \frac{\sigma_R^2}{n_1\sigma_R^2 + n_2\sigma_E^2}, \qquad \forall i \in E, \\ \pi_i &= \frac{\sigma_E^2}{n_1\sigma_R^2 + n_2\sigma_E^2}, \qquad \forall i \in R. \end{aligned} \tag{3.1}$$

Proposition 2.4 in this context takes the following special form.

Proposition 3.1. *Assume that $\sigma_i^2 = \sigma_E^2$, $\forall i \in E$, and $\sigma_i^2 = \sigma_R^2$, $\forall i \in R$, where $\sigma_E^2 < \sigma_R^2$. $\forall \pi$ probability such that $\pi_i = \pi_E$, $\forall i \in E$, $\pi_i = \pi_R$, $\forall i \in R$ and*

$$\begin{aligned} &\pi_E \geq \pi_R, \\ &\pi_E\sigma_E^2 \leq \pi_R\sigma_R^2. \end{aligned}$$

It is verified that

$$\sum_{i=1}^{n} \pi_i^2\sigma_i^2 \leq \frac{1}{n^2}\sum_{i=1}^{n} \sigma_i^2.$$

Notice how the above result requires the centrality vector π to be constant over the expert nodes and the regular ones, respectively.

3.1 The effect of self-confidence

We now assume the social network to possess a general centrality vector π. In order to achieve the optimal centrality (3.1), it is necessary for the agents to implement self-weights according to (2.18). However, the implementation of (2.18) is rather complicated because of the information that requires to agents, in fact, they must have full knowledge of the network.

Anyway, a natural feature is the following: to be expert is related to self-confidence, as pointed out in [2], [14]. For this reason, below, we investigate what happens when expert nodes introduce a self-weight α (the same for all of them) while the regular nodes do not change their updating rule. We want to see if such modification is always beneficial for the system in the sense that improves the

asymptotic wisdom.

From (2.17) we obtain that the new centrality vector is given by

$$\pi_i^\alpha = \begin{cases} \frac{1}{\mu}\pi_i & \text{if } i \in R \\ \frac{1}{\mu(1-\alpha)}\pi_i & \text{if } i \in E \end{cases} \tag{3.2}$$

The normalization constant μ is thus

$$\mu = \frac{\pi(E)}{1-\alpha} + \pi(R) = \frac{1-\alpha\pi(R)}{1-\alpha},$$

where we are denoting

$$\pi(E) = \sum_{i\in E} \pi_i, \qquad \pi(R) = \sum_{i\in R} \pi_i.$$

The variance of the asymptotic error as a function of the parameter α is thus given by

$$S(\alpha) = \sum_i (\pi_i^\alpha)^2 \sigma_i^2 = \frac{1}{\mu^2}\left[\frac{\sigma_E^2}{(1-\alpha)^2}\sum_{i\in E}\pi_i^2 + \sigma_R^2 \sum_{i\in R}\pi_i^2\right]. \tag{3.3}$$

It is convenient to introduce the symbols:

$$\pi^{(2)}(E) = \sum_{i\in E} \pi_i^2, \qquad \pi^{(2)}(R) = \sum_{i\in R} \pi_i^2.$$

Therefore, we can rewrite the form of $S(\alpha)$ as follows

$$S(\alpha) = \frac{\sigma_E^2 \pi^{(2)}(E) + (1-\alpha)^2 \sigma_R^2 \pi^{(2)}(R)}{(1-\alpha\pi(R))^2}. \tag{3.4}$$

Of course, in $\alpha = 0$ we have the variance of the asymptotic error of the original DeGroot model. We want to understand if $S(\alpha)$ is less than $S(0)$ for some α. This would show that some self-confidence would be helpful in the learning process. To this aim, we need to understand the monotonicity of $S(\alpha)$ near 0. Accordingly, we study the derivative in zero:

$$S'(0) = 2\left[\pi(R)\sigma_E^2 \pi^{(2)}(E) - \pi(E)\sigma_R^2 \pi^{(2)}(R)\right].$$

The only way to improve the asymptotic error of the original DeGroot model is that $S(\alpha)$ is decreasing in zero. Therefore, we have that $S'(0) < 0$ if and only if

$$\sigma_E^2 \frac{\pi^{(2)}(E)}{\pi(E)} < \sigma_R^2 \frac{\pi^{(2)}(R)}{\pi(R)}. \tag{3.5}$$

The terms

$$\omega_E = \frac{\pi^{(2)}(E)}{\pi(E)}, \qquad \omega_R = \frac{\pi^{(2)}(R)}{\pi(R)}, \tag{3.6}$$

measure how much the centrality is diffused among the expert and regular nodes, respectively. We thus proved the following Theorem.

Theorem 3.2. *There exists $\alpha > 0$ such that $S(\alpha) < S(0)$ if and only if*

$$\frac{\sigma_E^2}{\sigma_R^2} \frac{\omega_E}{\omega_R} < 1. \tag{3.7}$$

This is the condition under which, the addition of a small self-weight to all expert nodes is beneficial in terms of the asymptotic wisdom. Notice the combination of two ratios: the one of variances and the one of the diffusion rate of the centrality of experts and regulars.

3.1.1 Centrality diffusion rates

We now analyze in better detail the terms entering in the condition (3.7). It is necessary that the ratio between experts' centrality diffusion rate and regulars' centrality diffusion rate $\left(\frac{\omega_E}{\omega_R}\right)$ does not explode, otherwise it is not possible to reduce the asymptotic error of the original DeGroot model. The ratio depends not only on how much the experts are central, but also on how much they are concentrated, roughly.

We recall that $\sum_{i \in \mathcal{V}} \pi_i^2 \geq (\pi_{max})^2$, and, in addition, $\sum_{i \in \mathcal{V}} \pi_i^2 \leq \pi_{max} \sum_{i \in \mathcal{V}} \pi_i$. Using these estimates it is possible to determine a lower and upper bound for the ratio between experts' centrality diffusion rate and the regulars' one.

Lemma 3.3. *If W is a subset of $\mathcal{V}$ the following holds true:*

$$\omega_W \geq \frac{(\max_{i\in W} \pi_i)^2}{\pi(W)},$$

$$\omega_W \leq \max_{i\in W} \pi_i.$$

Proof. Since $\pi(W) = \sum_{i\in W} \pi_i$, $\pi^{(2)}(W) = \sum_{i\in W} \pi_i^2$ and $\omega_W = \frac{\pi^{(2)}(W)}{\pi(W)}$, we have

$$\pi^{(2)}(W) \geq (\max_{i\in W} \pi_i)^2,$$

$$\pi^{(2)}(W) \leq (\max_{i\in W} \pi_i)\pi(W).$$

This yields the result.

□

Proposition 3.4. *Given any subset W of $\mathcal{V}$, $\mathcal{V} = \{E, R\}$, the ratio between the centrality diffusion rates is bounded as follows:*

$$\frac{\|\pi_E\|_\infty^2}{\pi(E)\|\pi_R\|_\infty} \leq \frac{\omega_E}{\omega_R} \leq \frac{\|\pi_E\|_\infty \pi(R)}{\|\pi_R\|_\infty^2}. \tag{3.8}$$

Proof. Proof follows from Lemma 3.3, taking $W = E$ and $W = R$.

□

Suppose now that we let the size of network going to infinity, $n \to +\infty$. Let us show when the lower bound of (3.8) breaks out and, consequently, it is impossible to decrease the asymptotic error of the original DeGroot model with the introduction of self-weight. In addition, we show when the upper bound of (3.8) is bounded and, consequently, it could be possible to decrease the asymptotic error of the original DeGroot model with the introduction of self-weight.

Remark 3.1. *Assume that*

$$\pi(E) = e_n,$$

accordingly, $\pi(R) = 1 - e_n$. We focus on cases when e_n does not go to zero and to one, that means $0 < \delta \leq e_n \leq 1 - \delta < 1$[1]*. In addition, we assume that the*

[1]This condition represents a situation in which the experts are not negligible fraction of the population, but not even the whole fraction of population.

invariant measures of the most relevant agents among experts and regulars are respectively:

$$\|\pi_E\|_\infty = b_n,$$
$$\|\pi_R\|_\infty = a_n.$$

In view of this, we observe that the lower bound of (3.8) tends to $+\infty$ if and only if the following condition holds true

$$a_n = o\big(b_n^2\big).$$

Moreover, the upper bound of (3.8) is bounded if and only if the following condition holds true

$$b_n = O\big(a_n^2\big).$$

Summarizing the results, we obtain

- *if $a_n = o\big(b_n^2\big)$, then $\left(\frac{\omega_E}{\omega_R}\right) \to +\infty$,*
- *if $b_n = O\big(a_n^2\big)$, then $\left(\frac{\omega_E}{\omega_R}\right) < k$, $k \in \mathbb{R}_+$.*

This is a situation in which neither of them, at the centrality level, would take the upper hand, consequently it depends on the dispersion level. Therefore,

- if the expert agents are too concentrated compared to the regular ones, then the ratio between the centrality diffusion rates bursts,
- if expert agents are sufficiently widespread compared to the regular ones, then the ratio between the centrality diffusion rates is bounded.

3.1.2 The best self-weight

Of course, we are also interested in understanding which extent we can introduce self-confidence without implement the overall behavior. To this aim, we complete

the analysis of our function $S(\alpha)$ below. We recall that

$$S(\alpha) = \frac{\sigma_E^2 \pi^{(2)}(E) + (1-\alpha)^2 \sigma_R^2 \pi^{(2)}(R)}{(1-\alpha\pi(R))^2}.$$

Firstly, we determine the function zeros, under the condition $\alpha \neq \frac{1}{\pi(R)}$

$$\sigma_E^2 \pi^{(2)}(E) + (1-\alpha)^2 \sigma_R^2 \pi^{(2)}(R) = 0.$$

Since $\Delta = -\sigma_R^2 \sigma_E^2 \pi^{(2)}(R)\pi^{(2)}(E) < 0$, the function $S(\alpha)$ is always positive and in zero its value is equal to

$$S(0) = \sigma_E^2 \pi^{(2)}(E) + \sigma_R^2 \pi^{(2)}(R) > 0.$$

To define the stationary points, we examine the first derivative

$$S'(\alpha) = \frac{2\big(1-\alpha\pi(R)\big)\Big[-(1-\alpha)\pi(E)\sigma_R^2\pi^{(2)}(R) + \pi(R)\sigma_E^2\pi^{(2)}(E)\Big]}{\big(1-\alpha\pi(R)\big)^4}. \tag{3.9}$$

The study reveals that there is only one stationary point which is

$$\alpha = 1 - \frac{\sigma_E^2}{\sigma_R^2}\frac{\omega_E}{\omega_R},$$

and it is the absolute minimum of the function $S(\alpha)$.

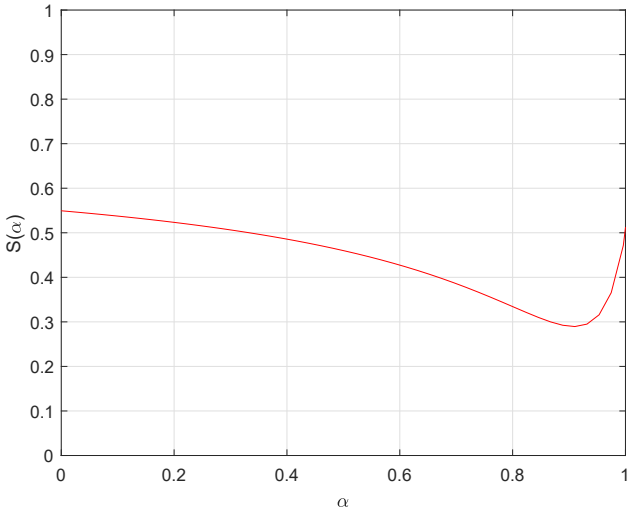


Figure 3.1: Plot of $S(\alpha)$ of a star network with 10 individuals (two expert leaves and the other nodes are regular), between 0 and 1, since α can only represent these values.

When the condition (3.7) occurs, the addition of self-weight confers benefits if $\alpha \in \left(0, 1 - \frac{\sigma_E^2}{\sigma_R^2}\frac{\omega_E}{\omega_R}\right]$. Accordingly, $\alpha = 1 - \frac{\sigma_E^2}{\sigma_R^2}\frac{\omega_E}{\omega_R}$ is the maximum self-weight for which there is a benefit, precisely the maximum benefit. The optimum alpha, denoted α^*, is thus

$$\alpha^* = 1 - \frac{\sigma_E^2}{\sigma_R^2}\frac{\omega_E}{\omega_R}.$$

In order to evaluate the extent of the improvement over the variance of the asymptotic error of the original DeGroot model, i.e., $S(0)$, we first calculate the value that $S(\alpha)$ takes in α^*

$$S(\alpha^*) = \frac{\sigma_E^2\sigma_R^2\pi^{(2)}(E)\pi^{(2)}(R)}{\sigma_R^2(\pi(E))^2\pi^{(2)}(R) + \sigma_E^2(\pi(R))^2\pi^{(2)}(E)}. \tag{3.10}$$

$S(\alpha^*)$ represents the smallest asymptotic variance value that can be obtained. Consequently, the maximum profit that we get from the introduction of alpha is given by

$$S(0) - S(\alpha^*) = \frac{\left[\sigma_R^2\pi(E)\pi^{(2)}(R) - \sigma_E^2\pi(R)\pi^{(2)}(E)\right]^2}{\sigma_R^2(\pi(E))^2\pi^{(2)}(R) + \sigma_E^2(\pi(R))^2\pi^{(2)}(E)}. \tag{3.11}$$

Namely it represents the greater asymptotic improvement with respect to the original DeGroot model when the condition (3.7) holds true.

3.1.3 Comparison with initial variance

Up to this point, we evaluated the performance of the self-weight adjustment compared to the case without self-loops. Now, we want to compare the variance of the asymptotic error by self-weight adjustment to the initial variance.

By assuming that there are n_1 experts, the initial variance is well-known and it is defined as follows:

$$V = \frac{1}{n^2}\sum_{i=1}^{n}\sigma_i^2 = \frac{1}{n^2}\left[n_1\sigma_E^2 + (n - n_1)\sigma_R^2\right].$$

We want to understand under which conditions there is an improvement over time. Since the study is relative to alpha, the initial variance is a constant. We

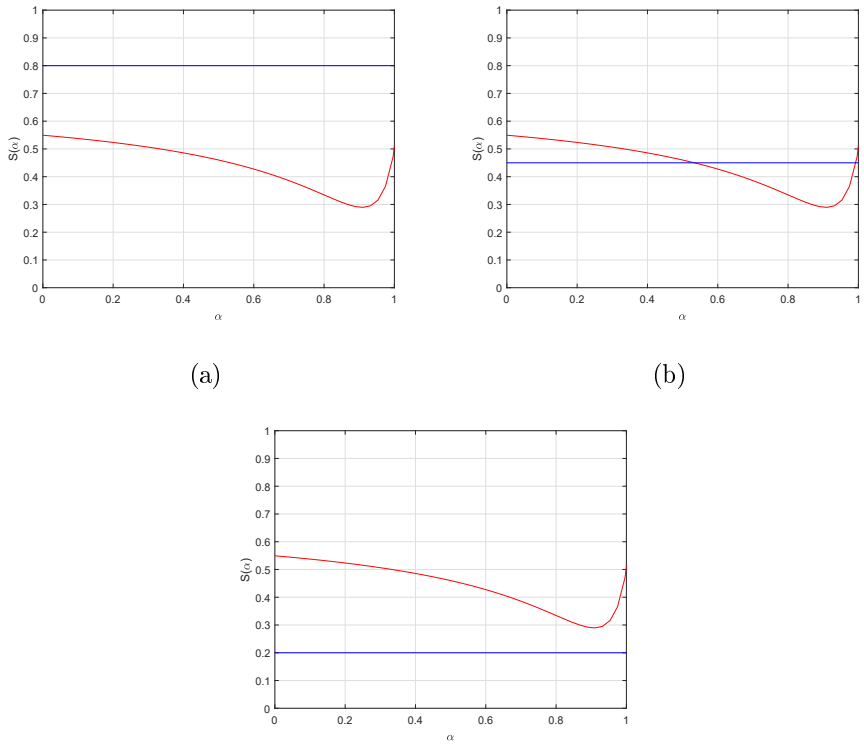


Figure 3.2: Comparison between the variance of the asymptotic error (red) and the initial variance (blue).

can obtain three different situations, depicted in Figure 3.2.

Figure 3.2(a) shows a case where the initial variance is always greater than the variance of the asymptotic error, therefore any alpha we insert it always lead to an improvement. Precisely, $\alpha = 0$ is enough for a better result. Figure 3.2(b) shows a case where the initial variance intersects the variance of the asymptotic error. This means that if we take α appropriately we can obtain an improvement. Finally, Figure 3.2(c) shows a case where the initial variance is always lower than the variance of the asymptotic error, therefore for any α introduced there is nothing that can be done, we can never break down the initial value.

Chapter 4

Examples with some self-confident nodes

In this Chapter, we analyze in detail the consequences of the theoretical results of previous section for specific examples of experts' distributions and, afterwards, specific networks. In each example, we use Theorem 3.2 in order to restate the improvement condition (3.7). In addition, we evaluate the optimum alpha α^* and the minimum variance of the asymptotic error $S(\alpha^*)$. Finally, we examine the comparison with the asymptotic error variance of the original DeGroot model.

4.1 Particular cases

In this section, we assume that no particular information is available regarding the network, but for the centrality vector.

Case 1. Only one expert

In this example, we assume there is only one expert. This yields

$$\pi(E) = \pi_E, \qquad \pi^{(2)}(E) = \pi_E^2, \qquad \text{and} \qquad \omega_E = \pi_E.$$

The variance of the asymptotic error is given by

$$S(\alpha) = \frac{\sigma_E^2 \pi_E^2 + (1-\alpha)^2 \sigma_R^2 \pi^{(2)}(R)}{(1-\alpha\pi(R))^2},$$

and, by Theorem 3.2, $S(\alpha) < S(0)$ if and only if

$$\frac{\sigma_E^2}{\sigma_R^2} \frac{\pi_E}{\omega_R} < 1. \tag{4.1}$$

If this condition occurs, then the optimum alpha is thus

$$\alpha^* = 1 - \frac{\sigma_E^2}{\sigma_R^2} \frac{\pi_E}{\omega_R},$$

and, by using this value, we obtain the minimum variance of the asymptotic error

$$S(\alpha^*) = \frac{\sigma_E^2}{1 + \frac{\sigma_E^2}{\sigma_R^2} \frac{(\pi(R))^2}{\pi^{(2)}(R)}}.$$

Accordingly, the maximum profit is equal to

$$S(0) - S(\alpha^*) = \frac{\left[\sigma_R^2 \pi^{(2)}(R) - \sigma_E^2 \pi(R) \pi_E\right]^2}{\sigma_R^2 \pi^{(2)}(R) - \sigma_E^2 (\pi(R))^2}$$

This means the asymptotic error variance of the original DeGroot model is larger than $S(\alpha^*)$, which represents the smallest asymptotic value that can be obtained with the introduction of a self-weight. In addition, we obtain an improvement over the initial variance if and only if the following inequality is verified

$$\left[1 + \frac{\sigma_E^2}{\sigma_R^2} \frac{(\pi(R))^2}{\pi^{(2)}(R)}\right] \left[\frac{1}{n^2} + \frac{\sigma_R^2}{\sigma_E^2} \frac{(n-1)}{n^2}\right] > 1.$$

Case 2. n_1 experts and $\pi_i = \pi_E$, $\forall i \in E$, and $\pi_i = \pi_R$, $\forall i \in R$

In this instance, the number of experts is more than one and we have further information on the invariant probabilities. In fact, the centrality vector is made up of only two values, one for experts π_E and one for regulars π_R. This yields

$$\pi(E) = n_1 \pi_E, \qquad \pi^{(2)}(E) = n_1 \pi_E^2, \qquad \text{and} \qquad \omega_E = \pi_E,$$
$$\pi(R) = (n - n_1) \pi_R, \qquad \pi^{(2)}(R) = (n - n_1) \pi_R^2, \qquad \text{and} \qquad \omega_R = \pi_R.$$

The variance of the asympotic error is given by

$$S(\alpha) = \frac{n_1\sigma_E^2\pi_E^2 + (n-n_1)(1-\alpha)^2\sigma_R^2\pi_R^2}{(1-(n-n_1)\alpha\pi_R)^2},$$

and, by Theorem 3.2, $S(\alpha) < S(0)$ if and only if

$$\frac{\sigma_E^2}{\sigma_R^2}\frac{\pi_E}{\pi_R} < 1. \tag{4.2}$$

It is noticeable that this decreasing condition remains unchanged even when the number of experts is one ($n_1 = 1$). This means that, if $\pi_i = \pi_E$, $\forall i \in E$, and $\pi_i = \pi_R$, $\forall i \in R$, no matter how many experts there are, the improvement depends only on the variances and the invariant probabilities. If the condition (4.2) occurs, then the optimum alpha is thus

$$\alpha^* = 1 - \frac{\sigma_E^2}{\sigma_R^2}\frac{\pi_E}{\pi_R},$$

and we get

$$S(\alpha^*) = \frac{\sigma_E^2\sigma_R^2}{n_1\sigma_R^2 + (n-n_1)\sigma_E^2}.$$

Accordingly, the maximum profit is equal to

$$S(0) - S(\alpha^*) = \frac{n_1(n-n_1)\left[\sigma_R^2\pi_R - \sigma_E^2\pi_E\right]^2}{n_1\sigma_R^2 + (n-n_1)\sigma_E^2}.$$

The introduction of alpha leads to an improvement over the initial variance if and only if the following inequality is verified

$$\left[\frac{n_1\sigma_R^2 + (n-n_1)\sigma_E^2}{n^2\sigma_E^2\sigma_R^2}\right]\left[n_1\sigma_E^2 + (n-n_1)\sigma_R^2\right] > 1.$$

By imposing $n_1 = 1$, we obtain the results for the more specific case.

Case 3. Uniform $\boldsymbol{\pi}$, $\boldsymbol{\pi} = \frac{\mathbf{1}}{\boldsymbol{n}}\mathbb{1}$

In this case, all invariant probabilities are equal, therefore, assuming n_1 experts, the diffusion rates of centrality are given by

$$\pi(E) = n_1\frac{1}{n} \qquad \pi^{(2)}(E) = n_1\frac{1}{n^2} \qquad \text{and} \qquad \omega_E = \frac{1}{n},$$
$$\pi(R) = (n-n_1)\frac{1}{n} \qquad \pi^{(2)}(R) = (n-n_1)\frac{1}{n^2} \qquad \text{and} \qquad \omega_R = \frac{1}{n}.$$

The variance of the asympotic error is given by

$$S(\alpha) = \frac{n_1\sigma_E^2 + (n-n_1)(1-\alpha)^2\sigma_R^2}{[n-(n-n_1)\alpha]^2},$$

and, by Theorem 3.2, $S(\alpha) < S(0)$ if and only if

$$\frac{\sigma_E^2}{\sigma_R^2} < 1. \tag{4.3}$$

Since by definition $\sigma_E^2 < \sigma_R^2$, we can conclude that the addition of self-weight is always suitable. The optimum alpha is thus

$$\alpha^* = 1 - \frac{\sigma_E^2}{\sigma_R^2},$$

and we get

$$S(\alpha^*) = \frac{\sigma_E^2\sigma_R^2}{n_1\sigma_R^2 + (n-n_1)\sigma_E^2}.$$

Accordingly, the maximum profit is equal to

$$S(0) - S(\alpha^*) = \frac{n_1(n-n_1)\left[\sigma_R^2 - \sigma_E^2\right]^2}{n^2\left[n_1\sigma_R^2 + (n-n_1)\sigma_E^2\right]}.$$

By introducing alpha, there is an improvement over the initial variance if and only if the following inequality is verified

$$\left[\frac{n_1\sigma_R^2 + (n-n_1)\sigma_E^2}{n^2\sigma_E^2\sigma_R^2}\right]\left[n_1\sigma_E^2 + (n-n_1)\sigma_R^2\right] > 1.$$

As can be seen, $S(\alpha^*)$ and the condition that determines if it is possible to break down the initial variance are equal to the above case. In fact, we can consider the Case 3 as a particular case of the more generic Case 2.

Let n_1 be a fraction of the total number of agents, i.e., $n_1 = \theta n$, for $\theta \in [0,1]$. In this scenario, the variance of the asymptotic error can be written as

$$S(\alpha) = \frac{\theta\sigma_E^2 + (1-\theta)(1-\alpha)^2\sigma_R^2}{n[1-(1-\theta)\alpha]^2}.$$

For large scales, i.e., $n \to \infty$, $S(\alpha)$ vanishes. This means that the variance of the asymptotic error in the limit vanishes. Consequently, we get

$$S(\alpha^*) = \frac{\sigma_E^2\sigma_R^2}{n\left(\theta\sigma_R^2 + (1-\theta)\sigma_E^2\right)},$$

and

$$S(0) - S(\alpha^*) = \frac{\theta(1-\theta)\left[\sigma_E^2 - \sigma_R^2\right]^2}{n\left[\theta\sigma_R^2 + (1-\theta)\sigma_E^2\right]}.$$

4.2 Known networks cases

In this section, we know the real structure of the network, consequently we have information about the form of the centrality vector.

4.2.1 Regular Graph

A regular graph is a graph where each vertex has the same number of neighbors, that is to say every vertex has the same degree, see Section 1.1.1. In general, in a r-regular graph, every vertex has degree r, so there are $\frac{rn}{2}$ edges. The general centrality vector π is proportional to the degree. Specifically, it is given by

$$\pi = \left(\frac{r}{rn}, ..., \frac{r}{rn}\right),$$

which is the degree divided by the total number of links.

A complete graph K_n is a regular graph of degree $n-1$, i.e., a simple undirected graph with n nodes in which every pair of distinct vertices is connected by a unique edge. Consequently, the general centrality vector π is given by

$$\pi = \left(\frac{n-1}{n(n-1)}, ..., \frac{n-1}{n(n-1)}\right).$$

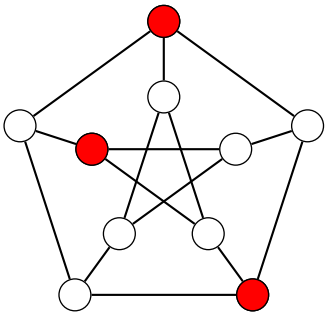
It is evident that the invariant vector in both cases is

$$\pi = \frac{1}{n}\mathbb{1}, \tag{4.4}$$

which is the uniform invariant probability.
Since the invariant probability is uniform, if we assume there are n_1 experts and $(n - n_1)$ regular agents, then we can see the detailed analysis in the Case 3 of section 4.1.

Petersen Graph

Let us consider a Petersen graph with 10 nodes, where each of them has degree three.

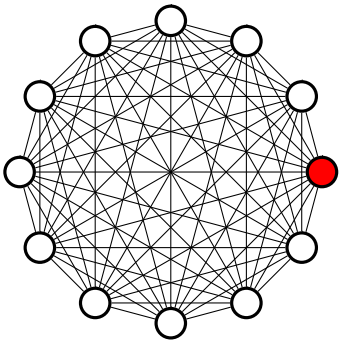


We assume there are 3 experts, identified by the red color, and $\sigma_i^2 = \sigma_E^2$, $\forall i \in E$, $\sigma_i^2 = \sigma_R^2$, $\forall i \in R$. In addition, based on the above explanation, the centrality vector is $\pi = \frac{1}{10}\mathbb{1}$. We set $\sigma_E^2 = 1$ and $\sigma_R^2 = 2$, accordingly we get that the optimum self-weight is $\alpha^* = \frac{1}{2}$. The variance improvement (over time) is given by

$$S(0) - S(\alpha^*) = \frac{17}{100} - \frac{2}{13} \approx 1.6\%.$$

Complete Graph

Let us consider a complete graph with 12 nodes, where each of them has degree $r = 11$.



We assume there is just one expert, identified by the red color, with variance σ_E^2, and $\sigma_i^2 = \sigma_R^2$, $\forall i \in R$. In addition, the general centrality vector is $\pi = \frac{1}{12}\mathbb{1}$. We set $\sigma_E^2 = 1$ and $\sigma_R^2 = 2$, accordingly we get that the optimum self-weight is $\alpha^* = \frac{1}{2}$. The variance improvement (over time) is given by

$$S(0) - S(\alpha^*) = \frac{23}{144} - \frac{2}{13} < 1\%.$$

In order to compare this example with the previous one, let us assume there are three experts and the variances remain unchanged:

$$S(0) - S(\alpha^*) = \frac{21}{144} - \frac{2}{15} \approx 1.3\%.$$

We can conclude by saying that, in regular graphs, setting the variances (σ_E^2 and σ_R^2) and the number of experts, the extent of the improvement decreases when the number of the individuals increases.

4.2.1.1 Applications of the modified DeGroot model

We apply the DeGroot Model to concretely observe the improvement, assuming that the true state is 0. In Figure 4.1, the learning model is applied to the example described above. It is possible to notice that the consensus, in the model with self-weights, is better than the original DeGroot model because it is closer to the truth, $\mu = 0$. The original dynamic would have led to a consensus $\bar{X}_\infty = -0.2905$, which is improved with the introduction of self-weight, $\bar{X}_\infty^\alpha = -0.1346$.

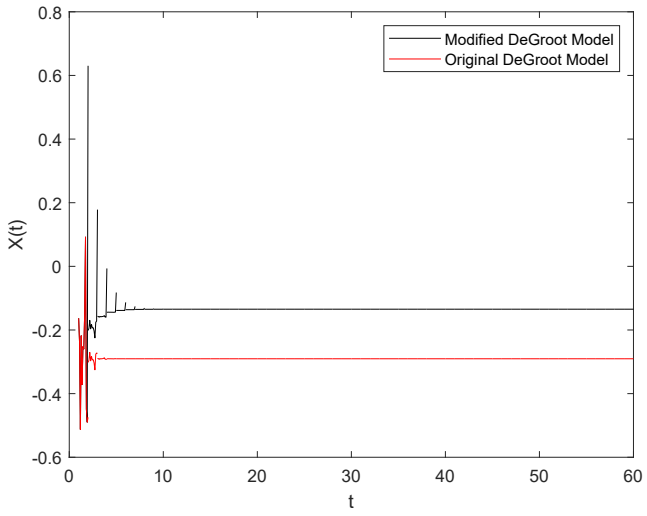


Figure 4.1: Comparison between the modified DeGroot model and the original DeGroot one on Complete Graph without self-loops, $n = 12$, $n1 = 1$, $\sigma_E^2 = 1$, $\sigma_R^2 = 2$.

This improvement is visible even on large networks. Therefore, we report two applications of DeGroot model on networks with 10000 agents. In the first case, illustrated in Figure 4.2, there are 2000 expert individuals. Expert individuals and regular ones have very different variances: $\sigma_E^2 = 1$ and $\sigma_R^2 = 11$, respectively. The consensus achieved by the original DeGroot model is $\bar{X}_\infty = 0.1111$, but the addition of self-weight to experts leads to a remarkable improvement, i.e., $\bar{X}_\infty^\alpha = 0.0471$. In the second case, illustrated in Figure 4.3, two aspects change: the percentage of experts increases to 8000, and self-loops are added to the original graph. Also in this example there is an improvement: $\bar{X}_\infty^\alpha = -0.0085$ compared to $\bar{X}_\infty = -0.0656$. In this case, it is worth noting that the DeGroot model converges to consensus in 1 step. Hence, the plot is a line.

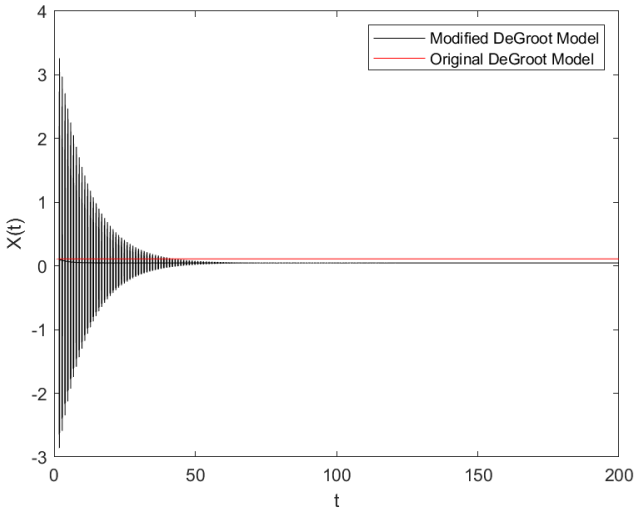


Figure 4.2: Comparison between the modified DeGroot model and the original DeGroot one on Complete Graph without self-loops, $n = 10000$, $n1 = 2000$, $\sigma_E^2 = 1$, $\sigma_R^2 = 11$.

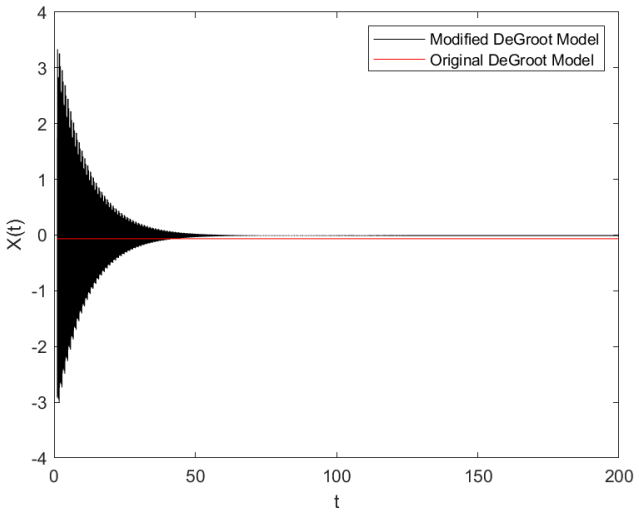


Figure 4.3: Comparison between the modified DeGroot model and the original DeGroot one on Complete Graph with self-loops, $n = 10000$, $n1 = 8000$, $\sigma_E^2 = 1$, $\sigma_R^2 = 11$.

4.2.2 Star Graph

The star graph, see Section 1.1.1, has a centralized topology. The centrality vector π is given by

$$\pi_i = \begin{cases} \frac{1}{2} & \text{if } i \text{ root} \\ \frac{1}{2(n-1)} & \text{if } i \text{ leaf} \end{cases}$$

We focus on the introduction of self-weights in a network where individuals are split in experts and regulars. We make use of system (3.2) to determine the new centrality vector and we analyze three particular cases.

Case 1. One expert: the root

The invariant vector influenced by α is

$$\pi_i^{\alpha} = \begin{cases} \frac{1}{2\mu(1-\alpha)} & \text{if } i \text{ root} \\ \frac{1}{2\mu(n-1)} & \text{if } i \text{ leaf} \end{cases}$$

As before, this yields

$$\begin{aligned}
&\pi(E) = \frac{1}{2}, && \pi^{(2)}(E) = \frac{1}{4}, && \text{and} && \omega_E = \frac{1}{2},\\
&\pi(R) = \frac{1}{2}, && \pi^{(2)}(R) = \frac{1}{4(n-1)}, && \text{and} && \omega_R = \frac{1}{2(n-1)}.
\end{aligned}$$

The variance of the asymptotic error is given by

$$S(\alpha) = \frac{(n-1)\sigma_E^2 + (1-\alpha)^2\sigma_R^2}{(n-1)(2-\alpha)^2},$$

and, by Theorem 3.2, $S(\alpha) < S(0)$ if and only if

$$\frac{\sigma_E^2}{\sigma_R^2}(n-1) < 1. \tag{4.5}$$

This condition, for large n, is never verified, therefore, for $n \to \infty$, there is never the possibility of an improvement. This is due to the fact that the expert is too concentrated. If the condition (4.5) occurs, then the optimum alpha is given by

$$\alpha^* = \frac{\sigma_R^2 - (n-1)\sigma_E^2}{\sigma_R^2},$$

and, by using this self-confidence, we obtain

$$S(\alpha^*) = \frac{\sigma_E^2\sigma_R^2}{\sigma_R + (n-1)\sigma_E^2}.$$

In addition, the maximum profit is given by

$$\begin{aligned}
S(0) - S(\alpha^*) &= \frac{(n-1)\sigma_E^2 + \sigma_R^2}{4(n-1)} - \frac{\sigma_E^2\sigma_R^2}{\sigma_R + (n-1)\sigma_E^2}\\
&= \frac{\left[\sigma_R^2 - (n-1)\sigma_E^2\right]^2}{4(n-1)\sigma_R^2 + 4(n-1)^2\sigma_E^2}.
\end{aligned}$$

Case 2. n_1 experts, only leaves

The invariant vector influenced by α is

$$\pi_i^\alpha = \begin{cases} \frac{1}{2\mu(n-1)(1-\alpha)} & \text{if } i \text{ expert leaf}\\ \frac{1}{2\mu(n-1)} & \text{if } i \text{ regular leaf}\\ \frac{1}{2\mu} & \text{if } i \text{ root} \end{cases}$$

Of course, if $n_1 = 1$, there is only one expert that is a leaf. As before, we make our considerations starting by the diffusion rates of centrality.

$$\pi(E) = \frac{n_1}{2(n-1)}, \qquad \pi^{(2)}(E) = \frac{n_1}{4(n-1)^2}, \qquad \text{and } \omega_E = \frac{1}{2(n-1)},$$
$$\pi(R) = \frac{2n-n_1-2}{2(n-1)}, \quad \pi^{(2)}(R) = \frac{n^2-n-n_1}{4(n-1)^2}, \quad \text{and } \omega_R = \frac{n^2-n-n_1}{2(n-1)(2n-n_1-2)}.$$

The variance of the asymptotic error is given by

$$S(\alpha) = \frac{n_1\sigma_E^2 + (n^2-n-n_1)(1-\alpha)^2\sigma_R^2}{\big[2(n-1)-(2n-n_1-2)\alpha\big]^2},$$

and, by Theorem 3.2, $S(\alpha) < S(0)$ if and only if

$$\frac{\sigma_E^2}{\sigma_R^2}\frac{2n-n_1-2}{n^2-n-n_1} < 1. \tag{4.6}$$

We observe that, in the limit case $n_1 = 1$, the condition simplifies to

$$\frac{\sigma_E^2}{\sigma_R^2}\frac{2n-3}{n^2-n-1} < 1.$$

It is interesting to note that, for large n, such condition is always verified, therefore the introduction of alpha always decreases the asymptotic error of the original DeGroot model. If the condition (4.6) occurs, then the optimum alpha is thus

$$\alpha^* = \frac{(n^2-n-n_1)\sigma_R^2 - (2n-n_1-2)\sigma_E^2}{(n^2-n-n_1)\sigma_R^2},$$

and we obtain

$$S(\alpha^*) = \frac{\sigma_E^2\sigma_R^2(n^2-n-n_1)}{n_1(n^2-n-n_1)\sigma_R^2 + (2n-n_1-2)^2\sigma_E^2},$$

In addition, we can conclude that the maximum profit is given by

$$\begin{aligned} S(0) - S(\alpha^*) &= \frac{n_1\sigma_E^2 + (n^2-n-n_1)\sigma_R^2}{4(n-1)^2} - S(\alpha^*) \\ &= \frac{n_1\big[(n^2-n-n_1)\sigma_R^2 - (2n-n_1-2)\sigma_E^2\big]^2}{4(n-1)^2\big[n_1(n^2-n-n_1)\sigma_R^2 + (2n-n_1-2)^2\sigma_E^2\big]}. \end{aligned}$$

Case 3. n_1 experts: root and $(n_1 - 1)$ leaves

The invariant vector influenced by α is

$$\pi_i^\alpha = \begin{cases} \frac{1}{2\mu(1-\alpha)} & \text{if } i \text{ root} \\ \frac{1}{2\mu(n-1)(1-\alpha)} & \text{if } i \text{ expert leaf} \\ \frac{1}{2\mu(n-1)} & \text{if } i \text{ regular leaf} \end{cases}$$

As before, we make our considerations starting by the diffusion rates of centrality.

$$\pi(E) = \frac{n + n_1 - 2}{2(n-1)}, \quad \pi^{(2)}(E) = \frac{n^2 - 2n + n_1}{4(n-1)^2}, \quad \text{and } \omega_E = \frac{n^2 - 2n + n_1}{2(n-1)(n + n_1 - 2)},$$
$$\pi(R) = \frac{n - n_1}{2(n-1)}, \qquad \pi^{(2)}(R) = \frac{n - n_1}{4(n-1)^2}, \qquad \text{and } \omega_R = \frac{1}{2(n-1)}.$$

The variance of the asymptotic error is given by

$$S(\alpha) = \frac{(n^2 - 2n + n_1)\sigma_E^2 + (n - n_1)(1-\alpha)^2\sigma_R^2}{\big[2(n-1) - (n-n_1)\alpha\big]^2},$$

and, by Theorem 3.2, $S(\alpha) < S(0)$ if and only if

$$\frac{\sigma_E^2}{\sigma_R^2}\frac{n^2 - 2n + n_1}{n + n_1 - 2} < 1. \tag{4.7}$$

It is possible to observe that, for large n, such condition is never verified, therefore, for $n \to \infty$, there is never the possibility of an improvement. Anyway, in cases where the condition (4.7) occurs, the optimum alpha is

$$\alpha^* = \frac{(n + n_1 - 2)\sigma_R^2 - (n^2 - 2n + n_1)\sigma_E^2}{(n + n_1 - 2)\sigma_R^2},$$

and the asymptotic variance by using this self-confidence is the following

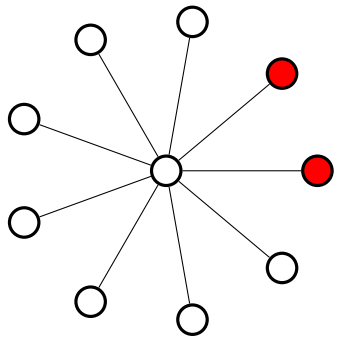
$$S(\alpha^*) = \frac{\sigma_E^2\sigma_R^2(n^2 - 2n + n_1)(n - n_1)}{(n - n_1)(n + n_1 - 2)^2\sigma_R^2 + (n - n_1)^2(n^2 - 2n + n_1)\sigma_E^2}.$$

Accordingly, the maximum profit is given by

$$\begin{aligned} S(0) - S(\alpha^*) &= \frac{(n^2 - 2n + n_1)\sigma_E^2 + (n - n_1)\sigma_R^2}{4(n-1)^2} - S(\alpha^*) \\ &= \frac{(n - n_1)\big[(n - n_1 - 2)\sigma_R^2 - (n^2 - 2n + n_1)\sigma_E^2\big]^2}{4(n-1)^2\big[(n + n_1 - 2)^2\sigma_R^2 + (n - n_1)(n^2 - 2n + n_1)^2\sigma_E^2\big]}. \end{aligned}$$

4.2.2.1 Specific numerical example

Let us suppose to have a star network with 10 individuals and just two of them are experts. Moreover, the experts are two leaves, so they can only get in touch with the central agent (Case 2). In the graph below, experts are identified by the red color.



We also assume that $\sigma_E^2 = 1$ and $\sigma_R^2 = 2$. We can easily calculate the centrality diffusion rates

$$\omega_E = \frac{1}{18} \qquad \omega_R = \frac{11}{36}$$

Since the condition (4.6) is respected, there is an improvement with the introduction of α. The value of the asymptotic variance, using the optimum alpha $\alpha^* = \frac{10}{11}$, is $S(\alpha^*) = \frac{11}{38}$. By applying α^*, the error decreases by 26% compared to the case without self-weights.

Chapter 5

Learning our expertise

We now want to study a different and more interesting situation that is when agents are not aware to which class (regular or expert) they belong. In this case they are not able to implement the self-weight adjustment as introduced in the previous sections. In this Chapter, we propose a simple learning method to overcome this difficulty and we show that, under certain assumptions, it produces a positive effect on the estimation of the truth μ. For the purpose of the simulation, we assume that the truth value μ is equal to 0.

We consider, as before, $N = R \cup E$ and that the initial individual variance is $\sigma_i^2 = \sigma_R^2$, $\forall i \in R$, and $\sigma_i^2 = \sigma_E^2$, $\forall i \in E$. For each agent i, we consider the dispersion of its measure with respect to those of its neighbors:

$$\tilde{\sigma}_i^2 = \big(x_i(0) - \frac{1}{|N_i|}\sum_{j \in N_i} x_j(0)\big)^2.$$

This means that only one measurement of the empirical variance of each agent is made and this is based on the initial opinions of its neighbors. In addition, we consider the following model for the assigment of the self-weight:

$$\alpha = \frac{0.9}{1 + k\tilde{\sigma}^2}, \tag{5.1}$$

where $k \geq 0$ is a non-negative coefficient. We observe that, since k is non-negative, α is a decreasing function. For this reason, we can express the following concept:

"The closer I get to the truth value, the more I will be self-confident in future decisions."

In order to find the best alpha values, we modify the coefficient k of the bell curve, see Figure 5.1. In the subsequent simulations the angular coefficient k

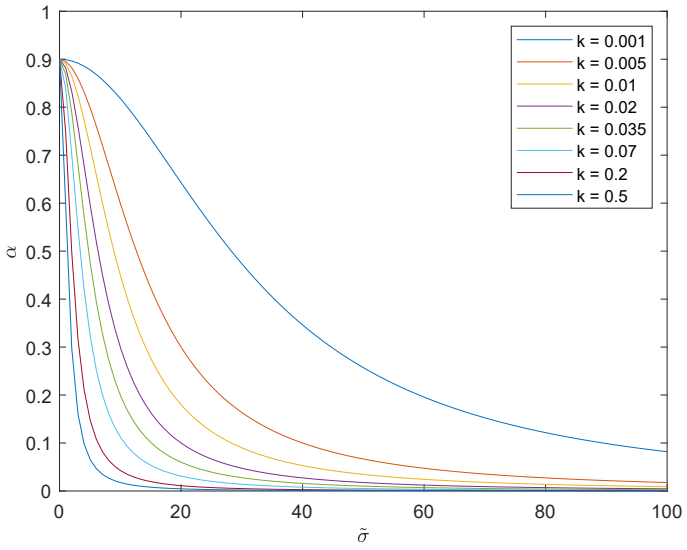


Figure 5.1: Several functions used to determine the self-weights.

assume values ranging between 0.002 and 2.5.

An analogous version of this problem is considered in [10] in the context of sensor networks. There, agents are sensors that can be faulty (regular nodes) or non-faulty (expert nodes). The proposed solution was a distributed implementation of the maximum-a-posteriori estimator. This type of solution is however not feasible in our social context for its complexity. Here we have preferred to consider and study a much simpler learning model.

5.1 Simulations

We implement these simulations on two different networks. First, we work on a Complete graph. Then, we consider an Erdős-Rényi random graph. In both cases, the number of agents and experts is known. We apply the Monte Carlo simulation to have a general idea of the goodness of the estimate. The number of runs r in each simulation is set to $r = 1000$. Hence, we get r initial conditions by generating a random number from the normal distribution, each independent of the others, and we save them in the rows of a matrix. In this way we can always use the same initial conditions when the self-weights assignment function varies and the comparison of the results is not affected by noise. The different initial conditions are studied at different times. Then, for each agent, we calculate the mean of the initial values of its neighbors which it will be used to estimate its variance. In fact, the variance of each individual is defined as the squared difference between the average of initial values of its neighbors and its own initial condition. Next, we define α_i according to the bell function (5.1). As said above, we change the coefficient k of the function to identify the one that produces the most improvement. As seen in Section 2.4, the use of self-weights modifies the invariant measures. In order to identify the new centrality vector, we have to compute the normalization constant. Finally, we can calculate the consensus point.

Performances are studied through the analysis of the following indices: $\bar{X}$, the mean of the initial condition averages in absolute value; $\bar{X}_\infty$, the mean of the consensus points in absolute value; $\bar{X}_\infty^\alpha$, the mean of the consensus points in absolute value when we use the self-confidence; and M, the magnitude of improvement between the two consensus value. We report these performance indices in each simulation. First of all, we remember from the theory that for the generic initial

condition $x \in \mathbb{R}^n$, we get the following values:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i(0), \qquad \bar{x}_\infty = \sum_{i=1}^{n} \pi_i x_i(0), \qquad \bar{x}_\infty^\alpha = \sum_{i=1}^{n} \pi_i^\alpha x_i(0).$$

The first one, i.e., $\bar{x}$, is the initial conditions mean, the other two cases concern the consensus points. The former is the consensus reached without self-weights, while the latter is the consensus reached by introducing the self-confidence. Now we can introduce the indices mentioned above, they are given by

$$\bar{X} = \frac{1}{r}\sum_{j=1}^{r} |\bar{x}_j|, \qquad \bar{X}_\infty = \frac{1}{r}\sum_{j=1}^{r} |(\bar{x}_\infty)_j|, \qquad \bar{X}_\infty^\alpha = \frac{1}{r}\sum_{j=1}^{r} |(\bar{x}_\infty^\alpha)_j|$$

and the magnitude of improvement is thus

$$M = \frac{\bar{X}_\infty - \bar{X}_\infty^\alpha}{\bar{X}_\infty}.$$

The $\bar{X}$ measures the average of the initial opinions without social influence. On the contrary instead, $\bar{X}_\infty$ and $\bar{X}_\infty^\alpha$ imply the presence of social influence. We are interested in positive values of M because they represent an improvement on the consensus point. In more detail, this means that the consensus obtained through the self-weight adjustment is better than the original consensus, i.e., $\bar{X}_\infty^\alpha < \bar{X}_\infty$. The greater M is, the bigger is the improvement got with respect to the original consensus. That is to say the greater M is, the better is the consensus obtained by the introduction of α. Therefore, we are looking for the biggest values of M. The idea is to determine r random initial conditions and to apply the various bell functions to identify the parameter α. We analyze different situations by changing the variances assigned to experts and regulars, used to determine the initial conditions. In Table visible on each simulation, the improvement compared to the case without self-weight is shown in correspondence with the number of experts n_1 and the coefficient k used. Regarding the graphs, on the horizontal axis there is the value of the coefficient k and on the vertical one there is the corresponding value of M.

5.1.1 Complete Graph

In a complete graph with n individuals, each agent communicates with all the others. Here, we consider the complete graph with self-loops. For this reason, we estimate the variance as the squared difference between the individual opinion and the overall average of the initial opinions:

$$\tilde{\sigma}_i^2 = \big(x_i(0) - \bar{x}\big)^2, \qquad \forall i = 1, ..., n,$$

where $x_i(0)$ is the initial opinion of the agent i and $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i(0)$. Since in the complete graph the centrality vector is $\pi = \frac{1}{n}\mathbb{1}$, see (4.4), then $\bar{x}$ also represents the consensus in the absence of self-weights, i.e., $\bar{x}_\infty$. For that reason $\bar{X} = \bar{X}_\infty$.

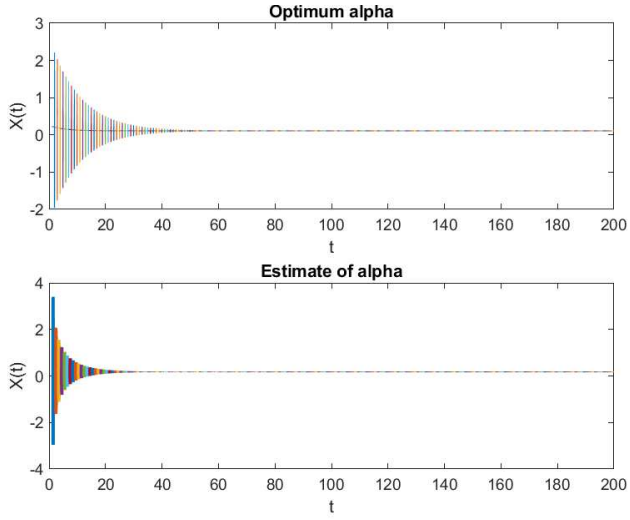


Figure 5.2: DeGroot Model on Complete Graph, $n = 400$, $n1 = 40$, $\sigma_E^2 = 1$, $\sigma_R^2 = 11$ and $\bar{x}_\infty^{\alpha^*} = 0.1091$, $\bar{x}_\infty^{\tilde{\alpha}} = 0.1752$, $\bar{x} = 0.2207$.

Figure 5.2 illustrates the comparison between the study case in Chapter 3 and the one of the following simulations, which we now analyze in detail. We consider a society where the number of individuals is 400, while the number of experts varies. We use the following values: 40, 120, and 250.

Simulation 1. The variances used to determine the random initial conditions of experts and regulars are considerably different, i.e., $\sigma_E^2 = 0.2$ and $\sigma_R^2 = 11$, respectively. As expected, $\bar{X}_\infty$ is monotonically decreasing with respect to the number of experts. It is possible to observe this phenomenon from the first row of Table 5.1: when the number of experts increases, then $\bar{X}_\infty$ decreases.

	$n_1 = 40$		$n_1 = 120$		$n_1 = 250$	
$\bar{X}_\infty$	0.4003		0.3672		0.2777	
	$\bar{X}_\infty^\alpha$	M	$\bar{X}_\infty^\alpha$	M	$\bar{X}_\infty^\alpha$	M
k=0.002	0.3687	0.0789	0.2617	0.2873	0.1361	0.5098
k=0.005	0.3493	0.1273	0.2291	0.3761	0.1069	0.6149
k=0.01	0.3335	0.1668	0.2075	0.4348	0.0906	0.6737
k=0.03	0.3131	0.2177	0.1844	0.4979	0.0762	0.7255
k=0.04	0.3097	0.2261	0.1814	0.5060	0.0750	0.7299
k=0.05	0.3079	0.2306	0.1801	0.5094	0.0748	0.7306
k=0.07	0.3066	0.2339	0.1801	0.5095	0.0758	0.7270
k=0.1	0.3072	0.2324	0.1827	0.5025	0.0787	0.7166
k=0.2	0.3141	0.2153	0.1953	0.4681	0.0897	0.6769
k=0.3	0.3211	0.1977	0.2072	0.4358	0.0995	0.6415
k=0.4	0.3271	0.1828	0.2173	0.4082	0.1080	0.6111
k=0.6	0.3364	0.1596	0.2333	0.3646	0.1217	0.5617
k=0.9	0.3461	0.1354	0.2506	0.3174	0.1372	0.5058

Table 5.1: Simulations on Complete Graph where $\sigma_E^2 = 0.2$ and $\sigma_R^2 = 11$.

Compared to the classic DeGroot model, we get remarkable improvements at the asymptotic level. In the best three cases of $n_1 = 40$, $n_1 = 120$ and $n_1 = 250$, we improve the consensus respectively of 23.39%, 50.95% and 73.06%.

As we can observe from Figure 5.3, it is clearly visible a maximum value, which

is obtained by using k of the order of magnitude of 10^{-1} in all three cases.

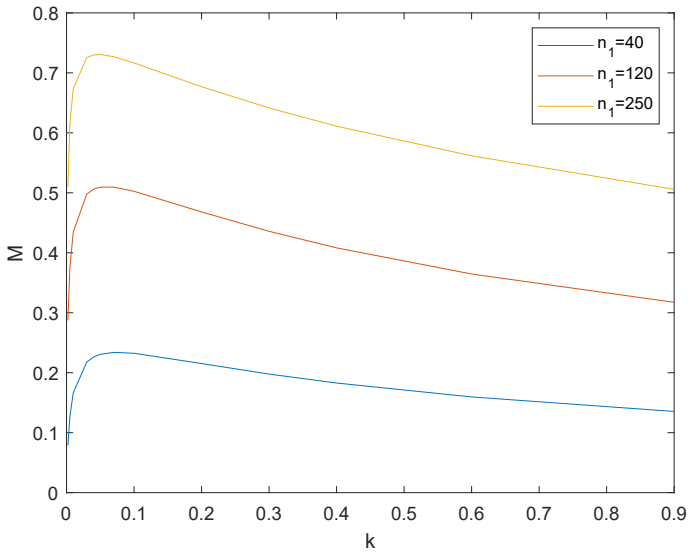


Figure 5.3: Simulations on Complete Graph with $n = 400$. Comparison between the values that M assumes in the three different situations ($n_1 = 40$, $n_1 = 120$, $n_1 = 250$), when the coefficient k varies. ($\sigma_E^2 = 0.2$ and $\sigma_R^2 = 11$).

Simulation 2. Here, for the experts, we use the same variance of the previous simulation, i.e., $\sigma_E^2 = 0.2$. The variance used to determine the initial random conditions of regular agents is clearly lower than the previous case, $\sigma_R^2 = 4$. This has a remarkable effect on the average of the initial conditions. In fact, the values of $\bar{X}_\infty$ are much lower than in Simulation 1. By comparing Figures 5.3 and 5.4, we observe a similar behavior. Furthermore, even in this case there is a maximum value, which, as before, is obtained for k of the order of magnitude of 10^{-1} in all three cases. However, we notice that these values of k are a little bit larger than in the previous simulation.

	$n_1 = 40$		$n_1 = 120$		$n_1 = 250$	
$\bar{X}_\infty$	0.1481		0.1311		0.0983	
	$\bar{X}^\alpha_\infty$	M	$\bar{X}^\alpha_\infty$	M	$\bar{X}^\alpha_\infty$	M
k=0.002	0.1448	0.0225	0.1168	0.1092	0.0757	0.2300
k=0.01	0.1386	0.0643	0.0997	0.2397	0.0547	0.4436
k=0.03	0.1314	0.1126	0.0865	0.3402	0.0425	0.5680
k=0.05	0.1279	0.1364	0.0813	0.3802	0.0386	0.6077
k=0.07	0.1257	0.1510	0.0784	0.4018	0.0368	0.6257
k=0.1	0.1237	0.1647	0.0761	0.4195	0.0357	0.6367
k=0.2	0.1212	0.1816	0.0742	0.4342	0.0361	0.6332
k=0.3	0.1208	0.1846	0.0748	0.4296	0.0378	0.6159
k=0.4	0.1210	0.1832	0.0760	0.4202	0.0397	0.5968
k=0.7	0.1226	0.1725	0.0802	0.3884	0.0447	0.5456
k=1.2	0.1254	0.1534	0.0860	0.3443	0.0508	0.4830
k=1.8	0.1281	0.1353	0.0911	0.3055	0.0560	0.4307
k=2.2	0.1295	0.1257	0.0937	0.2853	0.0586	0.4039

Table 5.2: Simulations on Complete Graph where $\sigma^2_E = 0.2$ and $\sigma^2_R = 4$.

Compared to the classic DeGroot model we get improvements at the asymptotic level. In the best three cases of $n_1 = 40$, $n_1 = 120$ and $n_1 = 250$, we improve the consensus respectively of 18.46%, 43.42% and 63.67%. Here, the original DeGroot model works better than in Simulation 1 because the regulars' variance decreased considerably. For this reasons, the improvements are lower than in the previous case.

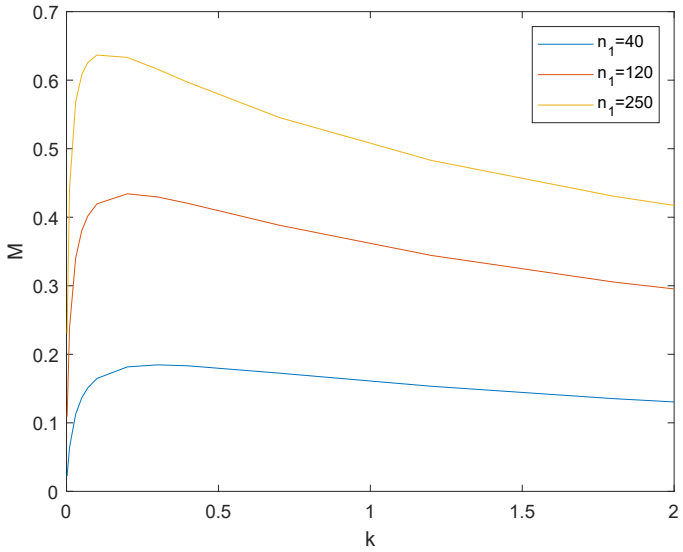


Figure 5.4: Simulations on Complete Graph with $n = 400$. Comparison between the values that M assumes in the three different situations ($n_1 = 40$, $n_1 = 120$, $n_1 = 250$), when the coefficient k varies. ($\sigma_E^2 = 0.2$ and $\sigma_R^2 = 4$).

Simulation 3. Now let us try to increase the variance used to determine the initial random conditions of expert agents compared to the previous case, by setting $\sigma_E^2 = 1.5$, while for regular ones it remains the same, $\sigma_R^2 = 4$.
This worsening of experts variance, $\sigma_E^2 = 1.5$ compared to $\sigma_E^2 = 0.2$, has two main consequences. First, as we can observe from Table 5.3, the consensus obtained by the original DeGroot model, i.e., $\bar{X}_\infty$, is slightly worse than the one obtained in Simulation 2. Second, the improvements are less significant. This last aspect is probably due to the fact that the difference between the two variabilities (experts and regulars) is reduced.

	$n_1 = 40$		$n_1 = 120$		$n_1 = 250$	
$\bar{X}_\infty$	0.1545		0.1371		0.1093	
	$\bar{X}^\alpha_\infty$	M	$\bar{X}^\alpha_\infty$	M	$\bar{X}^\alpha_\infty$	M
k=0.002	0.1522	0.0152	0.1290	0.0589	0.0993	0.0915
k=0.01	0.1510	0.0226	0.1249	0.0893	0.0946	0.1346
k=0.015	0.1510	0.0230	0.1245	0.0917	0.0944	0.1364
k=0.02	0.1510	0.0226	0.1245	0.0919	0.0945	0.1351
k=0.03	0.1512	0.0212	0.1248	0.0899	0.0951	0.1298
k=0.05	0.1517	0.0180	0.1256	0.0841	0.0963	0.1190
k=0.07	0.1521	0.0154	0.1263	0.0787	0.0973	0.1100
k=0.1	0.1526	0.0124	0.1272	0.0719	0.0985	0.0992
k=0.2	0.1534	0.0073	0.1291	0.0581	0.1008	0.0775
k=0.4	0.1539	0.0042	0.1311	0.0439	0.1030	0.0577

Table 5.3: Simulations on Complete Graph where $\sigma^2_E = 1.5$ and $\sigma^2_R = 4$.

The best improvement is obtain by using k of the order of magnitude of 10^{-2} in all three cases. Compared to the classic DeGroot model we get slight improvements at the asymptotic level. In the best three cases of $n_1 = 40$, $n_1 = 120$ and $n_1 = 250$, we improve the consensus respectively of 2.3%, 9.19% and 13.64%.

Simulation 4. Here, the variances used to determine the initial random conditions are similar to Simulation 1, but the experts one has got worse. Thus we have $\sigma^2_E = 1.5$, while for regular agents there are no changes, $\sigma^2_R = 11$. The fact that this is the worst situation that we have studied so far can be seen from the values of $\bar{X}_\infty$, worse than all the other cases, and from the value $\bar{X}^\alpha_\infty$ that is much higher than the previous cases.

	$n_1 = 40$		$n_1 = 120$		$n_1 = 250$	
$\bar{X}_\infty$	0.4156		0.3672		0.2819	
	$\bar{X}^\alpha_\infty$	M	$\bar{X}^\alpha_\infty$	M	$\bar{X}^\alpha_\infty$	M
k=0.002	0.3829	0.0785	0.2834	0.2284	0.1630	0.4220
k=0.005	0.3734	0.1016	0.2677	0.2711	0.1524	0.4594
k=0.01	0.3698	0.1103	0.2649	0.2788	0.1552	0.4493
k=0.015	0.3697	0.1104	0.2677	0.2711	0.1610	0.4288
k=0.02	0.3706	0.1082	0.2713	0.2611	0.1667	0.4086
k=0.03	0.3732	0.1021	0.2785	0.2415	0.1765	0.3740
k=0.05	0.3780	0.0905	0.2901	0.2101	0.1906	0.3237
k=0.07	0.3818	0.0814	0.2985	0.1873	0.2005	0.2889
k=0.1	0.3860	0.0712	0.3074	0.1628	0.2108	0.2523

Table 5.4: Simulations on Complete Graph where $\sigma_E^2 = 1.5$ and $\sigma_R^2 = 11$.

The best improvement is obtained by using k of the order of magnitude of 10^{-2} in all three cases. Compared to the classic DeGroot model we get improvements at the asymptotic level. In the best three cases of $n_1 = 40$, $n_1 = 120$ and $n_1 = 250$, we improve the consensus respectively of 11.04%, 27.88% and 45.94%.

Comment.

The highest improvement rates, compared to the classical DeGroot, are when the variability of the experts is very small.

Furthermore, in each simulation the order of magnitude of the best k is the same: it does not depend on the number of experts, but on the their variance. When $\sigma_E^2 = 0.2$, k is of the order of 10^{-1}, while, when $\sigma_E^2 = 1.5$, it is of the order of 10^{-2}.

5.1.2 Erdős-Rényi Graph

In this section, we perform our simulations on Erdős-Rényi graphs, which have been presented in Section 1.1.1.

In contrast to the complete graph, in Erdős-Rényi graph, individuals are not in communication with all the others. Furthermore, not all agents have the same number of neighbors. For this reason, we estimate the variance as the squared difference of the individual opinion with respect to the average response of neighbors as

$$\tilde{\sigma}_i^2 = \big(x_i(0) - \frac{1}{|N_i|}\sum_{\substack{j\in N_i \\ j\neq i}} x_j(0)\big)^2$$

The idea of establishing each individual's self-confidence based on the opinions of its neighbors has proved to be a success, in spite of reducing the comparison to just a little part of the group. In order to assign the values of the parameter α we use the decreasing function (5.1). The experiment has been carried out similarly to the one that we have analyzed for the complete case. In the implementation, the differences are mainly two:

- the network structure,
- the centrality vector π.

Regarding the structure of the network, we consider an Erdős-Rényi graph, so the structure is not defined. The centrality vector is not known a priori as in the case of the complete graph, since degrees are random variables. In this case it is possible to determine it once the network is fixed:

$$\pi_i = \frac{d_i}{2|\mathcal{E}|}, \qquad \forall i \in \mathcal{V},$$

where d_i is the degree of the individual i and $|\mathcal{E}|$ is the number of connections in the network. Once determined π, we can calculate the consensus in the case without self-weights as seen above $\bar{x}_\infty = \sum_{i=1}^n \pi_i x_i(0)$.

Firstly, we apply the standard DeGroot Model to observe that consensus is reached when agents are connected through an Erdős-Rényi graph. Secondly, we note that the Erdős-Rényi graph is wise because the limiting beliefs converge jointly to the true state $\mu = 0$. In Figure 5.5 we examine the learning model in a group of 400 individuals, with 40 experts. The variances used to obtain the random initial condition are $\sigma_E^2 = 0.2$ for expert individuals and $\sigma_R^2 = 11$ for regular ones.

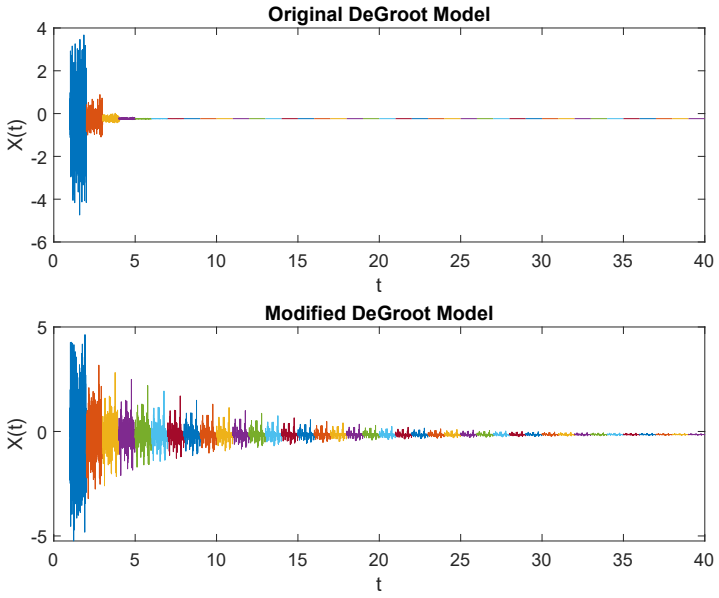


Figure 5.5: DeGroot Model on Erdős-Rényi Graph, $n = 400$, $n1 = 40$, $\sigma_E^2 = 0.2$, $\sigma_R^2 = 11$.

The consensus reached by the original DeGroot Model is $\bar{X}_\infty = -0.2367$, while the one reached with the self-weight adjustment is $\bar{X}_\infty^\alpha = -0.1469$. The self-weight adjustment loses in convergence time, but gains on the quality of the consensus point as estimation of the truth.

In the next simulations, we see in detail the effects of the self-confidence introduction on networks with $n = 400$ agents.

Simulation 1. In this scenario, the variances used to determine the random initial conditions differ greatly from experts ($\sigma_E^2 = 0.2$) to regular agents ($\sigma_R^2 = 11$), and the probability for each link to be present is equal to $p = 0.1$. Here, we generate a random network and on it we apply the different initial conditions. In Table 5.5, it is possible to observe the trend of the improvements when k varies. The improvement grows until it reaches a maximum and then it begins to decrease.

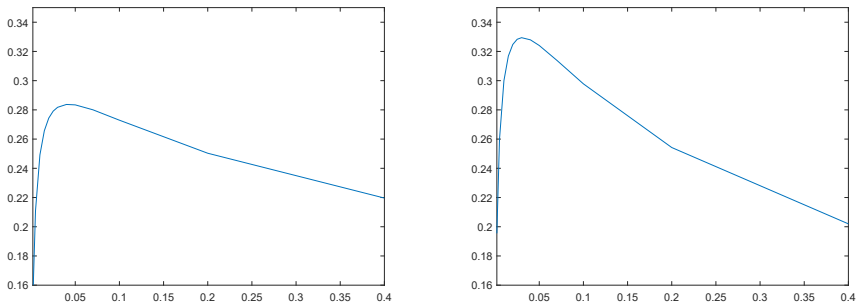
	$n_1 = 40$		$n_1 = 120$		$n_1 = 250$	
$\bar{X}$	0.4161		0.3697		0.2806	
$\bar{X}_\infty$	0.4186		0.3736		0.2823	
	$\bar{X}_\infty^\alpha$	M	$\bar{X}_\infty^\alpha$	M	$\bar{X}_\infty^\alpha$	M
k=0.002	0.3850	0.0804	0.2727	0.2700	0.1441	0.4893
k=0.01	0.3567	0.148	0.2255	0.3963	0.0997	0.6467
k=0.02	0.3482	0.1684	0.2136	0.4283	0.0899	0.6814
k=0.03	0.3452	0.1754	0.2096	0.4390	0.0868	0.6925
k=0.04	0.3442	0.1779	0.2082	0.4428	0.0857	0.6964
k=0.05	0.3439	0.1786	0.2078	0.4437	0.0854	0.6973
k=0.07	0.3442	0.1777	0.2086	0.4415	0.0860	0.6954
k=0.09	0.3452	0.1755	0.2102	0.4373	0.0870	0.6916
k=0.1	0.3457	0.1742	0.2111	0.4349	0.0877	0.6894
k=0.2	0.3512	0.1611	0.2203	0.4103	0.0941	0.6665
k=0.4	0.3589	0.1427	0.2343	0.3729	0.1044	0.6301

Table 5.5: Simulations on Erdős-Rényi Graph where $\sigma_E^2 = 0.2$ and $\sigma_R^2 = 11$.

Of course, we could generate a new random network for each initial condition. The trend would remain largely the same, but the $\bar{X}_\infty$ value for each k would be affected by noise.

From Table 5.5, we see that the social influence in the classical DeGroot model leads to a slight worsening compared to the average of the initial opinions, see $\bar{X}_\infty$ and $\bar{X}$. In all cases, the best improvement is obtained by using $k = 0.05$, i.e., k of the order of magnitude of 10^{-1}. Compared to the classic DeGroot model we get good improvements at the asymptotic level. In the best three cases of $n_1 = 40$, $n_1 = 120$ and $n_1 = 250$, we improve the consensus by 17.86%, 44.37% and 69.73%, respectively. It is interesting to note that the improvement rates are lower than the ones in the Simulation 1 of the complete graph. This is due to the fact that, here, we have just local information, therefore we have a limited knowledge.

If we had a comparison with the average opinions of all the agents (we denote this Method as Method 2), the improvements would be more significant. In Figure 5.6, we considered a network with $n_1 = 40$ experts, on the left we estimated $\tilde{\sigma}^2$ based on the average response of the neighbors, while on the right based on the average response of all individuals in the network. Between the first maximum and the second one there is an increment of 16.66%.



(a) (b)

Figure 5.6: Simulations on Erdős-Rényi Graph. Two different ways of evaluating $\tilde{\sigma}^2$, fixed $n_1 = 40$ and $p = 0.1$.

As the density of the graph decreases, the difference between the two methods increases.

Of course, Method 2 requires more information than Method 1. In a network

where agents are not connected with everyone, the global mean of the initial conditions must be given by an external system and/or individual.

Simulation 2. Now let us compare the two methods for calculating empirical variance from another point of view. Given the parameters, we wonder how many experts are needed to obtain the same consensus point. Assume that the variances used to get the initial conditions of the experts and the regulars are respectively $\sigma_E^2 = 0.2$ and $\sigma_R^2 = 4$. The number of individuals is $n = 400$, and the probability for each link to be present is equal to $p = 0.1$.

	$n_1 = 40$		$n_1 = 34$	
$\bar{X}$	0.1528		0.1531	
$\bar{X}_\infty$	0.1560		0.1542	
	$\bar{X}_\infty^\alpha$	M	$\bar{X}_\infty^\alpha$	M
k=0.05	0.1343	0.1387	0.1360	0.1181
k=0.07	0.1325	0.1508	0.1338	0.1321
k=0.1	0.1308	0.1612	0.1317	0.1454
k=0.2	0.1289	0.1739	0.1290	0.1634
k=0.5	0.1288	0.1743	0.1288	0.1642
k=0.7	0.1294	0.1703	0.1299	0.1576
k=1.1	0.1307	0.1621	0.1320	0.1440
k=1.8	0.1326	0.1502	0.1349	0.1247
k=2.1	0.1332	0.1460	0.1359	0.1182
k=2.5	0.1340	0.1412	0.1371	0.1107

Table 5.6: Simulations on Erdős-Rényi Graph by using two different ways of evaluating $\tilde{\sigma}^2$, where $\sigma_E^2 = 0.2$ and $\sigma_R^2 = 4$.

When k varies, the two methods respond similarly. We dwell on the situation where they achieve the same consensus result, that is shown in the box. These two values are really close, but the number of experts differs. By using Method 1 (comparison with neighbors) we have 40 experts, but the number of experts needed to attain that result using Method 2 is lower: $n_1 = 34$ experts are sufficient. There was a significant decrease, approximately 15% less.

Simulation 3. We report the case where there is a margin of error for expert agents. Therefore, we consider their variability equal to $\sigma_E^2 = 2.5$. In a society of $n = 400$ individuals, we focus on three particular cases: the first, where the 10% of the society is expert, the second, the 25%, and the last, the 40%. On average, each agent has 40 neighbors, this means that $p = 0.1$.

	$n_1 = 40$		$n_1 = 100$		$n_1 = 160$	
$\bar{X}$	0.4350		0.3873		0.3540	
$\bar{X}_\infty$	0.4415		0.3935		0.3570	
	$\bar{X}_\infty^\alpha$	M	$\bar{X}_\infty^\alpha$	M	$\bar{X}_\infty^\alpha$	M
k=0.002	0.4180	0.0531	0.3300	0.1614	0.2774	0.2230
k=0.005	0.4122	0.0663	0.3242	0.1761	0.2708	0.2413
k=0.01	0.4101	0.0710	0.3255	0.1728	0.2734	0.2341
k=0.015	0.4099	0.0716	0.3284	0.1654	0.2774	0.2231
k=0.02	0.4103	0.0707	0.3312	0.1583	0.2809	0.2131
k=0.04	0.4126	0.0654	0.3394	0.1375	0.2910	0.1847
k=0.07	0.4154	0.0590	0.3467	0.1191	0.2998	0.1603
k=0.1	0.4175	0.0542	0.3513	0.1072	0.3052	0.1450

Table 5.7: Simulations on Erdős-Rényi Graph where $\sigma_E^2 = 2.5$ and $\sigma_R^2 = 11$.

In the case of a limited group of experts, the asymptotic consensus decreases, but not in a significant way. Nevertheless, it is possible to obtain an improvement not only with respect to the original DeGroot, but also with respect to the average of the initial beliefs.

Simulation 4. In this section, we focus on the analysis of society with very limited number of experts. We work on a group of $n = 400$ individuals.

		$\sigma_E^2 = 0.2$, $\sigma_R^2 = 11$, $p = 0.1$			
	k	$\bar{X}$	$\bar{X}_\infty$	$\bar{X}_\infty^\alpha$	M
$n_1 = 2$	0.1	0.4377	0.4452	0.4407	0.0102
$n_1 = 1$	102	0.4491	0.4572	0.4562	0.0022

		$\sigma_E^2 = 0.2$, $\sigma_R^2 = 11$, $p = 0.2$			
	k	$\bar{X}$	$\bar{X}_\infty$	$\bar{X}_\infty^\alpha$	M
$n_1 = 2$	0.2	0.4419	0.4450	0.4417	0.0074
$n_1 = 1$	0.7	0.4475	0.4496	0.4491	0.0012

Normally, we keep getting slight improvements even when experts represent a very small part of the group. Of course, a coefficient $k = 102$ does not make much sense. This is because it assigns significant self-weights only if the empirical variance is extremely small.

Simulation 5. The self-weight adjustment does not always lead to an improvement. There could be cases in which particular values of k worsen the consensus point with respect to the standard DeGroot model. In Table 5.8 it is possible to observe an example in which this phenomenon occurs. In this simulation there are 400 agents: 40 of them are experts and the probability for each link to be present is equal to $p = 0.1$. The variances used to determine the random initial conditions are the following: $\sigma_E^2 = 1.5$ and $\sigma_R^2 = 4$. The mean of the initial opinion averages in absolute value is $\bar{X} = 0.1592$ and the mean of the consensus points in absolute value is $\bar{X}_\infty = 0.1622$.

	$n_1 = 40$	
	$\bar{X}_\infty^\alpha$	M
k=0.02	0.1613	0.0059
k=0.03	0.1609	0.0080
k=0.04	0.1607	0.0094
k=0.05	0.1605	0.0104
k=0.07	0.1604	0.0112
k=0.1	0.1604	0.0111
k=0.2	0.1607	0.0095
k=0.5	0.1613	0.0057
k=0.7	0.1616	0.0040
k=1.1	0.1619	0.0017
k=1.8	0.1622	-0.0001
k=2.1	0.1623	-0.0005
k=2.5	0.1624	-0.0008

Table 5.8: Simulation on Erdős-Rényi Graph where we observe some deteriorations compared to the original DeGroot model. ($\sigma_E^2 = 1.5$ and $\sigma_R^2 = 4$).

Simulation 6. In this simulation we test a learning model in which the update takes place at each step. The idea is to update the empirical variance and, consequently, the self-weight at every turn. Of course, this requires that also the transition matrix is updated every time step by using the new α weight.

Now we go into details. We generate an Erdős-Rényi graph and one thousand initial conditions each one independent of the others, whose behaviors are then mediated. In this way, we might get a general idea of the trend. We set, at the beginning, the maximum number of revisions, in this case 12. For the group of initial conditions, we analyze the behavior when the number of revisions varies. In particular, on these values we test, first, just one revision, then two and so on, up to testing twelve revisions.

We considered a society of $n = 400$ individuals of which $n_1 = 40$ are experts and the probability of inclusion of an edge is $p = 0.1$. We assumed that the variances used to get the initial condition of experts and regulars are respectively $\sigma_E^2 = 0.2$ and $\sigma_R^2 = 11$. For the assignment of self-weights we used the following function:

$$\alpha = \frac{0.9}{1 + 0.05\tilde{\sigma}^2}.$$

Table 5.9 shows the value of the consensus reached by the self-weight adjustment, $\bar{X}_\infty^\alpha$, and the degree of improvement with respect to the consensus without self-confidence, M, when the number of revisions varies.

The mean of the initial opinion averages in absolute value is $\bar{X} = 0.4096$ and the mean of the consensus points in absolute value is $\bar{X}_\infty = 0.4144$.

Number of revisions	$\bar{X}^{\alpha}_{\infty}$	M
1	0.3406	0.1780
2	0.3187	0.2310
3	0.3155	0.2386
4	0.3179	0.2329
5	0.3221	0.2228
6	0.3268	0.2114
7	0.3315	0.2001
8	0.3360	0.1893
9	0.3401	0.1793
10	0.3439	0.1702
11	0.3473	0.1619
12	0.3474	0.1617

Table 5.9: Continuous update on Erdős-Rényi Graph.

As can be seen from Table 5.9, in all cases we get an improvement on the consensus point of the classical DeGroot model. Consider that, doing more revisions is computationally more complex than just one at the beginning. For that reason it is useful to determine if the improvements are greater than the single initial revision. Under these conditions, from revision 2 to revision 9, we are able to obtain better results than with the only one initial revision. Nevertheless, the greatest approach to the truth is reached with 3 continuous updates. In this case, we improve the single one revision at the first step of 7.37%.

Conclusions

In this thesis we have considered the DeGroot learning model in situation where agents are heterogeneous with different capabilities to estimate the "truth". We have analyzed how modifications of the self-confidence parameter can improve the final consensus result.

At the beginning we analyzed an unknown network and our aim was to find the ideal combination between π_i and σ_i^2 in order to have an asymptotical improvement in the learning model. We obtained that the optimum π_i is inversely proportional to the variance of each individual σ_i^2. In addition, given σ_i^2, we found a right combination of π_i that allows to break down the error of the wisdom of the crowd. The largest part of the work was about the introduction of self-confidence in a society split in expert agents and regular ones. In particular, we dealt with introducing a certain self-weight α only to experts. We determined the optimal form of α with a view to minimize the error with respect to the truth value. We have seen that the possibility of improving the consensus point of the original DeGroot model depends on how the centrality is diffused among expert agents and regular ones. In fact, in order to obtain improvements by self-weight introduction, the ratio between experts centrality diffusion rate and the one of regulars must not explode. The idea of self-confidence introduction has been shown even in simulations that attempt to simulate a realistic scenario, in which nobody knows its own error with respect the truth, but tries to minimize it through the comunication with neighbors. Based on self-weights adjustments, we decided to modify

the DeGroot learning model by attributing weights to each individual. This modification was made to the first iteration, in which each agent estimates its own empirical variance based on neighbors' responses. We found a function able to assign self-weights in relation to empirical variance of individuals:

$$\alpha = \frac{0.9}{1 + k\tilde{\sigma}^2}$$

Thanks to the simulations we found the k that lead to the best consensus point. The order of magnitude of k depends on the variances used for the initial conditions, but not on the number of experts. In particular, if the variability of experts is really small, the order of magnitude of k is 10^{-1}, instead, if it grants a range of errors to expert agents, the order of magnitude of k is 10^{-2}. We managed to reduce the error of the consensus point of the original DeGroot model and we got also improvements on the wisdom of crowds obtained by the mean of the initial conditions. Of course, under the same setting conditions, in the Complete graph we got better improvements rather than the Erdős-Rényi one. This because, in Erdős-Rényi graph, we have local informations.

Clearly the study is not finished, it is possible to deepen the update at each step, i.e., updating weights at each step. Furthermore, it would be interesting to design experiments to understand how people really act in a similar dynamic and which form takes alpha in reality. We worked on DeGroot model, but it is possible to think about similar applications on other opinion dynamics model, such as the discrete-time Altafini model. Moreover, it would be interesting compare these results with statistical methods or analyze time-varying networks.

Bibliography

[1] Aristotele. *Politica, trattato sull'economia.* Politica III. 1281b.

[2] J. Becker, D. Brackbill, and D. Centola. *Network dynamics of social influence in the wisdom of crowds.* Proceedings of the National Academy of Sciences, Vol. 114(26), 2017.

[3] R.L. Berger. *A Necessary and Sufficient Condition for Reaching a Consensus Using DeGroot's Method.* Journal of the American Statistical Association, Vol. 76(374), pp. 415-418, 1981.

[4] F. Bullo, F. Fagnani, and B. Franci. *Finite-Time Influence Systems and the Wisdom of Crowd Effect.* Computing Research Repository, 2019.

[5] G. Como, F. Fagnani, and S. Zampieri. *I sistemi multi-agente e gli algoritmi di consenso.* La matematica nella societá e nella cultura: rivista della Unione Matematica Italiana, Serie I, Vol. 5(1), pp. 1-29, 2012.

[6] M.H. DeGroot. *Reaching a Consesus.* Journal of the American Statistical Association, Vol. 69, No. 345, pp. 118-121, 1974.

[7] P.M. DeMarzo, D. Vayanos, and J. Zweibel. *Persuasion Bias, Social Influence, and Unidimensional Opinions.* The Quarterly Journal of Economics, Vol. 118(3), pp. 909-968, 2003.

[8] P. Erdös and A. Rényi. *On the evolution of random graphs.* Publications of the Mathematical Institute of the Hungarian Academy of Sciences, Vol. 5, pp. 17-61, 1960.

[9] F. Fagnani. *Lecture notes on Dynamics over Networks.* Politecnico di Torino, Academic Year 2016/2017.

[10] F. Fagnani, S.M. Fosson, and C. Ravazzi. *A distributed classification/estimation algorithm for sensor networks.* SIAM Journal on Control and Optimization, Vol. 52(1), pp. 189-218, 2014.

[11] F. Galton. *Vox Populi.* Nature, Vol. 75, pp. 450-451, 1907.

[12] B. Golub, M.O. Jackson. *Naive learning in social networks and the wisdom of crowds.* American Economic Journal: Microeconomics, Vol. 2(1), pp. 112–149, 2010.

[13] T. Gura. *Citizen science: Amateur experts.* Nature, Vol. 496, pp. 259-261, 2013.

[14] J. Lorenz, H. Rauhut, F. Schweitzer, and D. Helbing. *How social influence can undermine the wisdom of crowd effect.* Proceedings of the National Academy of Sciences, Vol. 108(22), 2011.

[15] S. M. Ross. *Stochastic Processes.* United States of America, John Wiley & Sons, 2nd ed., Chapter 4, pp. 163-230, 1996.

[16] J. Surowiecki. *The Wisdom of Crowds.* New York, Anchor, pp. xi-xxi and pp. 3-22, 2004.

[17] A. Yeredor. *The Joint MAP-ML Criterion and its Relation to ML and to Extended Least-Squares.* IEEE Transactions on Signal Processing, Vol. 48(12), pp. 3484 – 3492, 2000.

List of Figures

List of Tables

A mamma e papá che nel bene e nel male ci sono sempre.
Che gioiscono per le mie vittorie, mi supportano nelle
sconfitte e mi spronano nei momenti di debolezza.