Communicate and Vote:
Collective Truth-tracking in Networks

MSc Thesis *(Afstudeerscriptie)*

written by

Nicolien S. Janssens
(born December 3st, 1996 in Arnhem, The Netherlands)

under the supervision of **Dr Davide Grossi** and **Prof Dr Giuseppe Dari-Mattiacci**, and submitted to the Examinations Board in partial fulfillment of the requirements for the degree of

MSc in Logic

at the *Universiteit van Amsterdam.*

Date of the public defense:	**Members of the Thesis Committee:**
March 24th, 2022	Prof Dr Sonja Smets
	Dr Adrian Haret
	Prof Dr Yde Venema
	Dr Spyros Terovitis

INSTITUTE FOR LOGIC, LANGUAGE AND COMPUTATION

Abstract

From different angles of science, there has been a growing interest in the abilities of groups to track the truth. The Condorcet Jury Theorem (1785) states that without communication, infinitely big groups will reach a correct majority opinion with certainty. Coughlan (2000), meanwhile formulated a model in which all agents communicate with each other, showing that majorities are only just as good as fully-communicating individuals. In reality, communication is usually between these two extremes: some agents communicate with some of the others, but not with all others. We refer to this as *partial communication.* This thesis provides a Bayesian framework to study the influence of partial communication on individual as well as group accuracy, thereby generalising Condorcet's as well as Coughlan's setting. We obtain results for individual and group accuracy in three type of networks. Firstly, we study the extreme case where there is either no communication or everyone communicates with everyone. Secondly, we determine accuracy for regular networks, in which all agents communicate with equally many other agents. Thirdly, we derive a formula to express the expected accuracy in random networks, in which agents can communicate with various numbers of other agents. The formula enables us to determine the effect of various parameters on the individual and group accuracy in a random network with partial communication. Finally, we show that in random networks, despite correlation between agents, we can still obtain accurate majorities under some constraints.

Acknowledgments

I want to thank Davide and Giuseppe for guiding and inspiring me in this process. I enjoy how many different kinds of research areas we combined and explored in the one and a half year we worked together. Starting from analysing collective decision processes ranging from the Kibbutz to the Pope, I am glad I can now say to you: “Habemus Thesis”.

I want to thank all members of the Thesis Committee: Sonja, Adrian, Yde and Spyros. Some of you I already know from previous projects, and I am happy to see you again. Thank you all for taking the time to participate in my defense.

I also want to thank Tanja and Ronald, for always being there to give advice on the MoL.

I am grateful for all the friends I made during this master program: the people from the Logic X Loosdrecht group, the Post-Study Fun Group and the Complexity Trauma Therapy group. Ultimately, I learned the most from collaborating with you. I want to thank *Branko* in particular. Although the MoL, as well as our half-marathon running, surf and skate sessions sometimes almost killed me, they all made me ultimately feel more alive and happy. And so does being with you.

Finally, thank you to all the people that always stand by my side, motivate and inspire me. My best friends and, in particular, Mum, Dad, Gladys, Brian (your attempts to help me with algebra), Loïs and *Knabbel*.

Now, I must add that writing this thesis would not have been as much fun if it was not for the music that often accompanies my writing sessions. I have played Satori’s newest album *Re:Imagned* over and over again, as well as Sarah Wild’s set *Plan:et C Beta 2021 | Turmbune | Fusion Festival* . It gives me good memories, a good energy and a lot of *Floor-freude*.

Contents

Summary of Notation

We give an overview of the most commonly used notation in the thesis. We remark that whenever we write a subscript, it concerns the property of a particular agent. For example, s_i is the signal of an agent i and δ_i is the degree of a vertex i, where the vertex represents an agent. Meanwhile, when we use brackets, it usually concerns a total amount. For example $v(x)$ is the total amount of votes that a signal x receives, and $\delta(x)$ is the total amount of times the signal x is observed. Keeping this in mind might ease the reading of this thesis.

$N = \{1, ..., n\}$	set of agents
$A = \{a, b\}$	set of alternatives, where θ is the true alternative
θ	true state of the world, where $\theta = a$ or $\theta = b$
r, $1 - r$	prior probability of a and b respectively
s_i	signal of agent i where $s_i \in \{a, b\}$
p	competence of an agent : $P(s_i = x \mid \theta = x)$ where $x \in \{a, b\}$
c	probability that two agents are connected with each other
v_i	vote of agent i
$\boldsymbol{v} = (v_1, ..., v_n)$	voting profile
$v(x)$	$\lvert\{v_i \in \boldsymbol{v} \mid v_i = x\}\rvert$: total amount of votes for $x \in \{a, b\}$
h	total amount of signals an agent receives
k, $h - k$	amount of a-type signals that an agent receives
$h - k$	amount of b-type signals an agent receives
π	$P(\theta = a \mid k)$: posterior belief about a, given evidence k
t	threshold such that an agent believes a iff $\pi > t$
δ_i	degree of a vertex i
$G(n, \delta)$	regular network, size n, all vertices have degree δ
$G(n, c)$	random network, size n, vertices linked with probability c
$p_{n,h}$	expected individual accuracy in regular network with $\delta = h$
$p_{n,c}$	expected individual accuracy in a UCRN $\langle N, c, p, r, t\rangle$
$p_{maj}(n)$	accuracy of a group N determined via majority rule
w_i	weight of agent's i signal s_i
$\delta(x)$	total amount of times signal x is observed
$w(x)$	total weight of all signals of type x
$\bar{\delta}$	mean degree

Chapter 1

Introduction

In the overheated jury room of the New York County Courthouse, a jury prepares to deliberate the case of a 19-year-old boy accused of stabbing his father to death. The judge instructs them that if there is any reasonable doubt, the jurors are to return a verdict of not guilty; if found guilty, the boy will receive a death sentence. The verdict must be unanimous. This is the story of the movie *Twelve Angry Men* by Lumet (1957) - a movie that beautifully illustrates the power of communication and voting.

The jurors collectively have to track the truth. They have to decide whether or not the boy is guilty. Everyone agrees on what decision is optimal in each case: the boy should be acquitted if innocent and convicted if guilty. But of course, no one knows what the true state of the world actually is. So, that is exactly why this jury is put together, which, by communicating will collectively try to make the best decision.

Initially, the group does not agree, as individuals composing the group have different beliefs. If we assume that all jurors reveal their private information truthfully and every other juror revises her belief correctly according to the newly received information, then the probability that the group will arrive at the correct conclusion will clearly increase as a result of communication. After all, due to communication, the collective decision will be based on more correct information than the decision of any individual could have been. However, in reality things are not that simple. One important reason for this is the fact that communication usually continues "outside office hours", off the record and unbridled by public scrutiny. As a result, information is not always publicly shared. The question that we will consider in this thesis therefore is: what is the probability that the group, as well as each individual that composes the group, will arrive at the correct conclusion as a result of communication with *some but not all* other group members?

We call the situation in which group members communicate with some but not all other group members *partial communication.* The occurrence of partial communication is beautifully illustrated in *Twelve Angry Men.* During the break, some jurors go the bathroom, and the conversation about whether the boy is to be convicted or acquitted continues, as can be read in the following passage from the movie.

> #7 exits, letting the door slam. #8 slowly dries his face. A moment later the door opens. We hear a loud laugh from outside. #6 enters the bathroom. The door closes. #6 walks over to the sink, turns on the water. During this next exchange he lets it run over his wrists. #6(sarcastically): Nice bunch of guys. #8: I guess they're the same as any. #:6 That loud, heavy set guy, the one who was tellin' us about his kid...the way he was talking...boy, that was an embarassing thing. #6(smiling): Yeah. What a murderous day. You think we'll be much longer? #8: I don't know. #6: He's guilty for sure. There's not a doubt in the whole world. We shoulda been done already. #8 doesn't answer him. #6: Listen, I don't care, 'y' know. It beats working'. He laughs and #8 smiles. Then #6 pointedly looks at #8. His smile vanishes. #6: You think he's not guilty? #8: I don't know. It's possible. #6: I don't know you, but I'm betting' you've never been wronger in your life. Y'ougta wrap it up. You're wastin' your time. #8: Supposing you were the one on trial?

Partial communication falls within two extremes: no communication and public communication. Public communication occurs when *all* people in a group talk to *all* other group members. There already exists a result for group accuracy under no communication, which is the famous Condorcet Jury Theorem by de Condorcet (1785). Meanwhile, Coughlan (2000) modeled accuracy of individual agents during public communication. Exactly because the models of the two authors are on opposite ends - the former doesn't allow for any, the latter only for public communication - we will use both their frameworks to investigate the case when communication is between these two extremes.

Group accuracy with no communication

The Condorcet Jury Theorem (henceforward denoted by "CJT") by de Condorcet (1785) shows that large groups of people are better at finding the truth than single individuals. The theorem consists of two components. Firstly, it shows that the accuracy of the majority is at least as good as the average accuracy of each individual agent. Secondly, it shows that as the group grows, the accuracy of the majority approaches 1. In practice, this means that when you ask a large group of people to give an answer regarding a certain factual question, by taking the majority of all their answers, it is very likely that you end up with the correct answer. The proof of the CJT theorem rests on some assumptions, one of which is that voters must vote independently of each other. This implies that people should not communicate with each other about what they think is the correct answer.

Accuracy with public communication

Coughlan (2000) formulated a formal model specifically adapted to jury trials. His model is an extension of work by Feddersen and Pesendorfer (1998), who model jury trials with the assumption that jurors do not communicate with each other. Coughlan extends this model by incorporating limited communication among jurors. In particular, he assumes that a jury takes a single nonbinding straw vote before taking the final binding vote for conviction or acquittal. Since the number of preliminary votes during the straw vote is publicly announced prior to the final vote, this can be seen as a form of minimal communication among jurors. Coughlan ultimately shows with his extended model that the unanimous voting rule minimise the probability of convicting innocent defendants and acquitting guilty ones. In this thesis, we will abstract away from this specific setting. Instead, we will solely adopt Coughlan's communication and voting model and generalise this by studying settings in which the single nonbinding straw vote is shared only with a subset of the agents.

Accuracy with partial communication

The goal of this thesis is to determine individual and group accuracy after partial communication. As such, we will generalise Condorcet's and Coughlan's models - the one for no communication and the other for public communication - to the case in-between: partial communication. We will do so in two steps. Firstly, we will determine individual and group accuracy in regular networks. This means that agents talk with only a subset of agents in the network, but every agent talks to a subset of agents of an equal size. In the second part, we will generalise to random networks, thus allowing for cases where some agents are more connected than others.

1.1 Contribution

The setting in which we will study individual and group accuracy is the following. We assume agents are in a network where they observe individual signals. The goal of agents is to track the true state of the world and signals represent the information that agents receive about the state of the world. These signals have a given accuracy, which we assume to be the same for all signals. In Condorcet's scenario, agents vote based solely on the signal they observe. In Coughlan's scenario, agents exchange their signals with all other agents, and then vote. The main goal of this thesis is to study the scenario in-between: agents exchange their signals with *some* other agents, and then vote.

As a result of signal exchange, agents receive new information. We assume that agents reveal their signals truthfully whenever they are matched with another agent in the network. Furthermore, we will assume that agents are Bayesian. Thus, they have a certain prior belief that they update using Bayes' rule and the incoming signals. We then assume agents are non-strategic and thus vote according to their posterior belief. We will determine the probability that an agent's posterior belief is correct: this is the agent's *individual accuracy*. Furthermore, we will determine the probability that asymptotically, the majority of agents is correct: this we will call the *group accuracy*.

Our contribution is a framework for individual and group accuracy as a result of partial communication. Regarding individual accuracy, we derive a general formula to determine the individual accuracy in any random network. This enables us to determine the effect of various parameters on the individual accuracy in a random network with partial communication. We show that as the size of the network and the amount of communication increases, also the individual accuracy increases, as one would intuitively expect. As the competence increases, often the accuracy increases too, but this is not always the case. Furthermore, we show that whenever the prior and threshold for belief are equal, then the formula for individual accuracy, in which agents try to determine the truth via Bayesian updating, reduces to a formula in which agents try to determine the truth via maximum likelihood estimation. Finally, from our general formula of individual accuracy, a formula for individual accuracy in regular networks can be derived. Since we show that Condorcet's and Coughlan's models are special cases of regular networks, it follows that the accuracy in these networks can be determined using this formula too.

There are several ways communication can influence group accuracy. Our focus in this thesis will be on the *amount of communication connections* agents have in the network and how this affects group accuracy. Therefore, we first study group accuracy in regular networks, in which each agent has equally many connections. We show that for any regular network when competence is uniformly distributed and above 0.5, in the limit we obtain a correct majority with probability 1. We therefore determine, using Chernoff Bounds (1952), the probability that in a random network, agents have close to equally many communication connections.

1.2 Related literature

Modelling partial communication and its effect on individual as well as group accuracy is the aim of this thesis. Our group accuracy result builds on the CJT and its follow-up literature in which the assumption of dependence is relaxed. We firstly discuss literature in this field. Our individual accuracy result builds on the model of Coughlan (2000). This is an epistemic, Bayesian model. We will therefore discuss literature related to this approach. Then, we discuss literature related to the epistemic benefits and risks of communication. Finally, we discuss some literature in the field of dynamic epistemic logic, as it also provides an approach to formalise information exchange in groups.

Literature on group accuracy and communication

The CJT is an example of *the wisdom of the crowds*, which is the idea that in general big groups of people are better at guessing facts than a small group of experts. It is for this reason that the CJT is often mentioned in one line with the wisdom of the crowds. For a reading on the wisdom of the crowds, we recommend the book *The Wisdom of Crowds: Why the Many are Smarter than the Few* by Surowiecki (2004).

Jury theorems are at the heart of the idea of the wisdom of the crowds. As Dietrich and Spiekermann (2019, p.386) formulate it: "A *jury theorem* is a mathematical theorem about the probability of correctness of majority decisions between two alternatives". The CJT (1785) was the very first jury theorem. Because it assumes independence between agents, it has been argued by Anderson (2006) that the CJT implies that communication might harm group accuracy. Yet, communication and voting are often seen as two important characteristics of information pooling in democratic decision making. Anderson (2006, p.11) thus argues that "an adequate model should show how they work together".

Several such models have been formulated. Dietrich and Spiekermann (2019) for example argue that we should not assume that agents have a uniform competence above 0.5, but that they instead have tendency to competence, meaning that the competence of voters *tends* to exceed 0.5, but is not always above 0.5. Furthermore, they argue that we should assume conditional independence by holding the state fixed. Under these assumptions, Dietrich and Spiekermann (2019) show that the group accuracy increases as the group size increases. Moreover, under these assumptions, they show that communication is beneficial as it makes agents more competent.

Also Hahn et al. (2019) investigate how communication has an effect on the accuracy of a group. In their Condorcet-like simulation model, the issue of the voters is to decide on the truth or falsity of a proposition. At each time step of the simulation, agents have a certain probability of receiving a signal that supports the truth or falsity of the proposition. At every time step each agent also has a probability of communicating with another agent. What agents communicate to each other is their belief of the truth value of the proposition. However, an agent only speaks when this belief exceeds a certain fixed threshold, to simulate that agents express their belief only when they are convinced up to a certain degree of what they think. Based on their simulation results, Hahn et al. (2019) conclude that when a moderate degree of communication is allowed, groups are better at tracking the truth than the individuals that composed them. However, when they allowed a lot of communication among agents, the accuracy of the group was significantly lower compared to groups with little or no communication allowed.

These results are in line with the work of Ladha (1992), who showed that when individuals in a group communicate with each other, the group can only be good at making a decision as long as the average level of dependence does not go too high and the mean competence of the individuals is good enough.

A closely related work that is also concerned with the truth-tracking potential of a group and also builds on the CJT is written by Michelini (2021). Michelini's goal is to investigate the effect on the truth-tracking ability of the group if being competent is costly. This represent the fact that acquiring information may take effort for agents. Michilini mostly considers networks in which each agent is connected with every other member of the group, but in the last part of his thesis relaxes this assumption and studies networks in which every agent is connected with a strict subset of the group. He therefore, like us, introduces an underlying network that connects the agents. However, other than us, he does not study regular or random graphs, but very specific categories of graphs: stars, rings and wheels. Our works can thus be considered as supplements of each other: while Michelini studies group accuracy with partial communication in specific categories of networks, we do so for regular and random graphs, without looking at the specific network.

Literature on individual accuracy and communication

Coughlan's model (2000) is an extension of work by Feddersen and Pesendorfer (1998), who simulate jury trials with the assumption that jurors do not communicate with each other. Their model consists of a set of Bayesian agents that have to decide on a binary issue - either the defendant is guilty or it is innocent - based on private information only. The agents are assumed to vote strategic, which means that they vote according to Nash equilibrium behavior. This means that agents do not simply vote their signal, but also take into account how to maximize utility. In this setting, Feddersen and Pesendorfer (1998) show that when agents vote strategically, the unanimity rule results in a strictly positive probability to acquit the guilty and convict the innocent.

By extending the model of Feddersen and Pesendorfer (1998) with communication, Coughlan (2000) claims that unanimity rule does not increase the risk of faulty judgements. In particular, with this extended model, Coughlan shows that the unanimity rule minimises the probability of convicting innocent defendants and acquitting guilty ones. Our work will not built on this specific result of Coughlan. Rather, we will adopt his Bayesian communication and non-strategic voting model and study individual accuracy of agents.

Another Bayesian approach is provided by Ding and Pivato (2021), who present a model of communication in which information sharing is costly. Like our model, they consider Bayesian agents in an epistemic framework. That is, there is a true state of the world that the agents collectively want to determine. Although they do make a distinction between private and public information of agents, there is no distinction between public and semi-public information sharing. That is, once an agent decides to disclose some of their private information, it immediately becomes public. Furthermore, they solely focus on group accuracy and not on individual accuracy.

Literature on the benefits and risks of communication

Communication, and specifically deliberation, is nowadays often seen as a key element in democratic decision making. Deliberation is form of communication where the goal is to exchange arguments and possibly come to an agreement. Proponents of deliberative democracy argue that deliberation and voting together justify collective decisions (see e.g. Chambers (2003)). As a result of this deliberative turn, various models of deliberation have been put forward. For example, Chung and Duggan (2020) recently modelled three different forms of deliberation - myopic discussion, constructive discussion and debate. Hereby, their focus was on the formal theory of arguments, rather than the structure of information exchange. Goldbach (2015) combined a formal model of preference formation with a model transformer that represents democratic deliberation. This model transformer imposes that all agents share their information with each other and thus we have public communication yet again. As we can see, in the literature on deliberative democracy, the communication always happens publicly, and in this respect the literature in this field thus differs from ours.

Next to the benefits of communication, there is also a field of study on the risks of communication - sometimes called the *madness of the crowds* due to Mackay (1962). Phenomena that can be related to this thesis are the ones that arise due to an asymmetry in communication. In random networks, after all, not everyone talks to the same number of other people and this can influence the group accuracy. One phenomenon that is of particular interest for our purposes is that of the majority illusion effect (see e.g. Lerman et al. (2016)). In this phenomenon, a belief that is relatively rare in a network is over-represented in the signals agents receive from other agents during communication. A related phenomenon is the Friendship Paradox formulated by Feld (1991), which states that most people have strictly fewer friends than their friends have.

Another phenomenon in which asymmetries can influence the flow of information in a network is that of information gerrymandering, recently modelled by Stewart et al. (2019). Information gerrymandering happens when the structure of a social network influences the voting outcome towards one party, even when both parties have equal sizes and each person communicates with equally many other agents. Other than in the majority illusion effect, with information gerrymandering every person talks to the same number of other persons. However, an asymmetry arises since people from one party are predominately exposed to beliefs of members from their own party, whereas in the other party agents are predominately exposed to beliefs of the members from the other party.

Literature on the logic of information exchange in groups

Another approach to study information exchange in groups and the consequences thereof is provided by the field of dynamic epistemic logic (DEL). DEL uses modal logics to describe knowledge change. A general introduction to the field is given in the book by van Ditmarsch et al. (2008). Related to this thesis is the formalisation of public communication by public announcement logic

(PAL) (van Ditmarsch et al., 2008, Chapter 4). PAL is used to reason about knowledge and belief and the changes thereof due to public announcements: an informational statement of an agent towards the whole group.

Another part of DEL that is related to this thesis are the logics that incorporate probabilistic updates. Standard DEL models belief qualitatively: an agent either believes a proposition p, or it does not believe proposition p. In contrast, probabilistic DEL, like the model that we formulate in this thesis, models belief quantitatively: an agent attaches a certain probability to a proposition p, which is interpreted as an agent's degree of belief in p. For example, van Benthem et al. (2009) generalize standard dynamic epistemic logics to a probabilistic setting. They model three probabilistic aspects of incoming information: the prior probability, the occurrence probability and the observation probability. Another example of probabilistic DEL is the work of Baltag and Smets (2008), who have developed a logic that connects DEL, belief revision theory and the Bayesian approach. They define a probabilistic product update that permits updating on events of probability zero, which is usually not possible in the Bayesian framework. Finally, in the recent work by Baltag et al. (2021), a model is formulated for forming and revising beliefs about unknown probabilities.

1.3 Overview

The goal of Chapter 2 is to determine individual accuracy and group accuracy for agents in regular networks. To do so, we start off by introducing the formal framework of this thesis in Section 2.1. We then formulate the Condorcet Jury Theorem in this framework in Section 2.2 and Coughlan's model in Section 2.3. We then introduce regular communication networks in Section 2.4 and observe that both Condorcet's as well as Coughlan's model both apply to regular networks. Therefore, in Section 2.5 we spell out a general formula to determine the individual accuracy in a regular communication network and we show how individual accuracy in the Condorcet's and Coughlan's model can be derived from this. Finally, we generalise the asymptotic part of the CJT to regualar communication networks of any size.

In Chapter 3 we shift our attention from regular networks to random networks and focus on individual accuracy. In Section 3.1 we give a brief introduction to random networks. In Section 3.2 we then generalise our theorem for individual accuracy in regular networks to random networks. The main part of the remaining chapter will spell out the exact effect of each of the parameters in our model on the individual accuracy. Section 3.2.1 discusses the influence of the size of the network, Section 3.2.2 determines the effect of the amount of communication, Section 3.2.3 spell out the effect of the competence of agents and Section 3.2.4 handles the effect of the prior and the threshold value. Finally, in Section 3.3 we show that in case the threshold value and the prior are equal, our formula for individual accuracy that is Bayesian nature, boils down to the more simple maximum likelihood estimation approach.

Chapter 4 is about group accuracy in random networks. This is non-trivial, because firstly, beliefs after communication are not independent anymore and secondly, because in a random network the number of agents other agents talk to is variable. We will therefore in Section 4.1 introduce *weights* as measures of the spread of communication in random networks and show that if each signal has the same weight, then asymptotically we reach a correct majority with probability 1. Therefore, in Section 4.2 we discuss in which networks communication is evenly spread. We show this is the case for the extreme cases when there is either no or full communication, and for the case when the number of agents tends to go to infinity. For all other cases, we propose to use the Chernoff Bound to determine the spread of communication.

Chapter 2

Individual and Group Accuracy in Regular Networks

This thesis will be a generalization of Condorcet's and Coughlan's work in two respects. In this chapter we present the first aspect, which is a generalization to regular networks. To do so, we proceed as follows. Section 2.1 sets out the formal framework that we will use in this thesis. Section 2.2 introduces the formal setting of the Condorcet Jury Theorem, while section 2.3 introduces the formal setting of Coughlan's work. In section 2.4 we will introduce regular networks and show that, although at first sight very far apart, both theorems apply to regular networks. Finally in section 2.5 we will generalize Coughlan's and Condorcet's model to regular communication networks, in which each agent communicates with an equal number of other agents and determine the individual accuracy as well as the group accuracy in this general setting.

2.1 The formal framework

Let $N = \{1, \dots, n\}$ be the set of agents with an individual agent denoted by i. There is a set $A = \{a, b\}$ of two alternatives, one of which is the true alternative, denoted by θ. Thus either $\theta = a$ or $\theta = b$. The prior probability of a is given by r, and the prior probability of b is given by $1 - r$. Before voting, each agent i receives a private signal $s_i \in \{a, b\}$. The competence of an agent is given by p, which is the conditional probability that the signal of an agent i is x given that the state of the world is x. The competence expresses the probability that an agent receives a correct signal. Since we will always assume that all agents are equally competent, we do not need to use an index i to refer to the competence p of a given agent i. For every agent i, let h_i be the number of signals an agent receives. Let k_i be the number of a-type signals out of h signals that an agent i

observes of *other* people. From this it follows that the number of b-type signals an agent i receives is equal to $h_i - k_i$. For simplicity, we sometimes will omit the subscript i in case it is clear from context which agent we are addressing.

On the basis of the newly acquired information, agents will update their beliefs according to Bayes' Theorem. Let $\theta = a$ be the event that the state of the world is a and let k denote the event that agent receives k a-type signals out of a total of h signals, such that $\mathbb{P}(k) \neq 0$, then according to Bayes' Theorem:

$$\mathbb{P}(\theta = a|k) = \frac{\mathbb{P}(k|\theta = a)\mathbb{P}(\theta = a)}{\mathbb{P}(k)} \tag{2.1}$$

In the context of beliefs, Bayes' theorem describes what probability an agent attaches to the event $\theta = a$ given new evidence k based on a prior belief in $\theta = a$. In particular, using Bayes' Theorem and the notation just introduced, we can calculate the posterior belief π in alternative a based on k a-type signals out of a total of h signals as follows:

Observation 2.1.1. *Let* $\pi = \mathbb{P}(\theta = a|k)$. *Then:*

$$\pi = \frac{r \cdot p^k(1-p)^{h-k}}{r \cdot p^k(1-p)^{h-k} + (1-r) \cdot p^{h-k}(1-p)^k}$$

Proof. By Bayes' theorem it follows that the posterior $\pi = \mathbb{P}(\theta = a|k)$ is given by:

$$\mathbb{P}(\theta = a|k) = \frac{\mathbb{P}(\theta = a)\mathbb{P}(k|\theta = a)}{\mathbb{P}(\theta = a)\mathbb{P}(k|\theta = a) + \mathbb{P}(\theta = b)\mathbb{P}(k|\theta = b)}$$

We fill in this equation as follows. $\mathbb{P}(\theta = a)$ is the prior probability that an agent attaches to a, we know this is given by r. Furthermore, $\mathbb{P}(k|\theta = a)$ is the probability of observing k signals in favour of a out of a total of h signals, given that $\theta = a$. This probability is given by the binomial distribution $B(h, p)$, where p refers to the competence of an agent. So, $\mathbb{P}(k|\theta = a) = \binom{h}{k}p^k(1-p)^{h-k}$. In this way we have the numerator of the fraction.

The denominator contains the numerator, and we add $\mathbb{P}(\theta = b)$, which is the prior probability an agents attaches to b; we know this is equal to $1 - r$. Finally, $\mathbb{P}(k|\theta = b)$ is the probability of observing k signals in favour of a out of a total of h signals, given that $\theta = b$. This probability is again given by the binomial distribution such that $\mathbb{P}(k|\theta = b) = \binom{h}{k}p^{h-k}(1-p)^k$.

So, we get the following equation.

$$\begin{aligned}\pi =& \frac{r \cdot \binom{h}{k}p^k(1-p)^{h-k}}{r \cdot \binom{h}{k}p^k(1-p)^{h-k} + (1-r) \cdot \binom{h}{k}p^{h-k}(1-p)^k} \\ =& \frac{r \cdot p^k(1-p)^{h-k}}{r \cdot p^k(1-p)^{h-k} + (1-r) \cdot p^{h-k}(1-p)^k}\end{aligned}$$

□

In words, $\pi = P(\theta = a|k)$ is the probability that an agent attaches to a being the state of the world, given the evidence k signals in favour of a, out of a total of h signals. Note that $1 - \pi$ is now the posterior for an agent believing b to be the true state of the world: the probability that b is the true state of the world, given that an agent observes k signals in favour of a out of a total of h signals.

Now, we have to determine when an agent is convinced that a is the true state of the world. After all, so far we only worked out what probability an agent attaches to a being the true state of the world. We have to determine when an agent will actually *believe* a to be the state of the world. To do so, we set a threshold $t \geq \frac{1}{2}$ for belief such that an agent i will believe a if and only if $\pi_i > t$, otherwise the agent believes b. That is, an agent will believe b if and only if $\pi \leq t$. In the literature, saying that a belief is equal to a degree of belief bigger than some threshold is called the *Lockean thesis* (see e.g. Foley (1993); Leitgeb (2014)). The Lockean thesis enables us to identify quantitative belief with qualitative belief. [1] Note that letting an agent believe a if and only if $\pi > t$ and not $\pi \geq t$ is ultimately an arbitrary choice, that does imply a small bias towards believing b. We will discuss the later in more detail. Depending on the context though, it can be argued that π needs to be strictly bigger than t, as for example in law it is common to convict someone only *beyond reasonable doubt*.

After collecting information and updating the posterior belief - either by communicating with other agents or sometimes only by consulting one's own private information - each agent i is asked to submit a vote $v_i \in A$. We will assume that agents always vote sincerely, meaning that agents vote according to their belief. That is, an agent will vote for a if and only if $\pi > t$ and an agent will vote for b if and only if $\pi \leq t$. Votes are gathered in a profile, which is a vector $\boldsymbol{v} = (v_1, ..., v_n)$. We denote the number of times an alternative $x \in \{a, b\}$ is voted for by $v(x)$ where $v(x) = |\{v_i \in \boldsymbol{v} | v_i = x\}|$.

The collective opinion is obtained by applying a voting rule f that we describe by a threshold $\hat{k}$, which is an integer between 0 and n:

$$f(\boldsymbol{v}) = \{x \in A | v(x) \geq \hat{k}\}$$

We highlight two voting rules in particular. Firstly, the majority rule, which selects the alternative that is voted on by a strict majority of the voters. That is $\hat{k} = \lfloor \frac{n}{2} \rfloor$:

[1] There are some objections against the Lockean thesis of belief. In particular, no threshold such that $t \neq 1$ will make the standard axioms for belief - KD45: seriality, transitivity, Euclideanness - sound. More importantly, the Lockean thesis makes the conjuctivity of belief fail. Leitgeb (2014) for example therefore argues that the Lockean thesis need to be expanded with some extra conditions. In particular, he argues that an agent believes x if it has a *stable* high probability in x under revision with any evidence that is consistent with x. Whether our account is compatible with this approach will have to be determined by future work. Note, however, that the fact that beliefs are not closed under conjunction is not a problem for this framework, as we are concerned with solely two states of the world that are incompatible with each other.

$$majority(\boldsymbol{v}) = \{x \in A | v(x) \geq \lfloor\frac{n}{2}\rfloor\}$$

Keep in mind that although we write $majority(\boldsymbol{v})$, the majority rule is dependent on more than just $\boldsymbol{v}$. Furthermore, notice that the majority rule allows for ties, in case the number of voters is even. For example, suppose we have 20 votes in total, 10 for each alternative. Then as $\hat{k} = 10$, both alternatives will satisfy the threshold and thus we have a tie. In that case we will break the tie by a fair coin toss, so basically a (fair) random selection of a winner. The second voting rule we highlight is the unanimity rule, in which $\hat{k} = n$:

$$unanimity(\boldsymbol{v}) = \{x \in A | v(x) \geq n\}$$

Again, we point out that the unanimity rule is dependent on more than $\boldsymbol{v}$, even though we denote it by $unanimity(\boldsymbol{v})$. Finally, we let $p_{maj}(n)$ be the accuracy of a group $N = \{1, ..., n\}$ where the collective choice is determined via $majority(\boldsymbol{v})$ and we let $p_{una}(n)$ be the accuracy of a group $N = \{1, ..., n\}$ where the collective choice is determined via $unanimity(\boldsymbol{v})$.

2.2 Condorcet's model

The Condorcet Jury Theorem (1785) shows that large groups of people are good at finding the truth. In particular, the theorem consists of three statements, explicated in the following theorem. The proof of this theorem is based on Grofman et al. (1983) and (Dietrich and Spiekermann, 2013, Appendix C).

Theorem 2.2.1 (Condorcet's Jury Theorem)**.** *Let* $\boldsymbol{v}$ *be a vector of* n *Bernoulli trials with a fixed probability* $0.5 < p \leq 1$*, where* n *is odd. Then:*

$$p_{maj}(n) \leq p_{maj}(n+2) \tag{2.2}$$

$$p \leq p_{maj}(n) \tag{2.3}$$

$$\lim_{n\to\infty} p_{maj}(n) = 1 \tag{2.4}$$

Proof. We give a proof sketch of each claim one by one.

(2.2) This claim follows from the following formula:

$$p_{maj}(n+2) = p_{maj}(n) + (2p-1) \cdot \binom{n}{\frac{n+1}{2}} \cdot (p(1-p))^{\frac{n+1}{2}} \tag{2.5}$$

Let $(2p-1) \cdot \binom{n}{\frac{n+1}{2}} \cdot (p(1-p))^{\frac{n+1}{2}} = \phi$. Since $0.5 < p \leq 1$, it follows that $\phi \geq 0$. Consequently $p_{maj}(n) \leq p_{maj}(n+2)$ as desired. We give a sketch of the derivation of formula 2.5. Let E_{n+2} be the event that the majority of the $n+2$ votes are correct and E_n be the event that a majority of the

n votes are correct. Furthermore, let $E_{n+2} \setminus E_n$ be the event the majority of $n+2$ are correct, but less than a majority of the n votes are correct. Similarly, $E_{n+2} \cap E_n$ is the event that a majority of the $n+2$ and of the n votes are correct. We can then rewrite $p_{maj}(n+2)$ as follows:

$$p_{maj}(n+2) = p_{maj}(n) + \mathbb{P}(E_{n+2} \setminus E_n) - \mathbb{P}(E_n \setminus E_{n+2}) \tag{2.6}$$

Now it can be derived that $\mathbb{P}(E_{n+2} \setminus E_n) = p\binom{n}{\frac{n+1}{2}}(p(1-p))^{\frac{n+1}{2}}$ and $\mathbb{P}(E_n \setminus E_{n+2}) = (1-p)\binom{n}{\frac{n+1}{2}}(p(1-p))^{\frac{n+1}{2}}$. Filling in these values in equation 2.6 gives us equation 2.5 as desired.

(2.3) This follows from the previous claim 2.2 and the observation that $p = p_{maj}(1)$.

(2.4) As n approaches infinity, it follows by the weak law of large numbers that the average number of correct votes converges to p. Now since $p > 0.5$ it follows that the probability of a correct majority also approaches 1.

□

In words, the CJT tells us three things. Firstly, equation (2.2) tells us that larger groups are better. More specifically, the larger the group, the higher the probability that the majority gets it right. Secondly, equation (2.3) tells us that groups are better than individuals. In particular, it says that the probability that more than half of the voters get the right answer is greater than each individual's competence p. Finally, equation (2.4) says that asymptotically, the accuracy of the group converges to 1. That is, in the limit, that is for very big groups, the group will be correct for sure. Although this statement talks about infinitely many agents, it is important to realize that the group size does not even have to be very big to get a high probability of getting it right. For example, for individuals with a competence $p = 0.8$, the probability that the majority of a group of just 13 individuals getting the right answer is greater than 0.99.

We point out that there are three main assumptions for the theorem to hold, as stated formally in the theorem. The first, *competence*, states that each individual should be more likely than a fair coin toss to get the right answer. This means that $p_i > 0.5$ for every voter i. The second assumption is that individual competences must be *homogeneous*. Formally, $p_i = p$, for some p and for all voters i. The last assumption, *independence*, states that the probability of one agent choosing one alternative does not affect any other agents' choice. This means that people should not communicate with each other. It is the last assumption that will be of our concern in this thesis and that we will challenge. We will investigate to what extent the CJT holds for networks in which agents do communicate with each other.

2.3 Coughlan's model

At the other end of the spectrum, we have Coughlan (2000). The overall goal of his paper is to determine under which voting rule jury trials perform best. In this thesis, however, we will focus on a specific part of Coughlan's work, which is his communication model[2]. The formal framework is as introduced in section 2.1. So, agents have to decide on a binary issue. In the communication model, in contrast to in the CJT model, agents vote two times. Firstly, they vote according to their own private signal. The signal is private in the sense that they share it with no one else. The results of this voting round are publicly announced. This can be compared to what is called a *public announcement* in Dynamic Epistemic Logic (van Ditmarsch et al., 2008, Chapter 4). Secondly, agents vote while taking into account the information of the previous voting round. Since the results of the first voting round are publicly announced, we will see that as a form of public communication.

It is easy to see that if all agents communicate with each other, meaning that everyone discloses their signal to one another, then each agent receives the exact same signals. Since all agents have the same prior and use the same update rule, it follows that all agents will vote the same. In other words, agents will vote unanimously. As a result, given that all signals are common knowledge, the probability that the majority is correct is the same as the probability that the unanimous decision is correct. We spell out this fact in the following observation.

Observation 2.3.1. *Given a network $N = \{1, .., n\}$ such that $h = n$ for all agents $i \in N$, after signal disclosure it follows that:*

$$p_{maj}(n) = p_{una}(n) \tag{2.7}$$

Proof. Firstly, each agent will have the same posterior belief π. Namely, π is a function of r, p, k, h. It is given that r, p are the same for all agents. Furthermore $h = n$ for all agents, which means that all agents observe the signals of all other agents. As agents observe the whole network, it also means that all agents observe the same number k a-type signals.

Since π_i is the same for all agents i, and $v_i = a$ if $\pi_i > t$ and $v_i = b$ if $\pi \leq t$ it follows that $v_i = v_j$ for all $i, j \in N$. Consequently, $p_{maj}(n) = p_{una}(n)$. □

Coughlan's framework is designed specifically for jury trials, where a small group of experts has to decide over the fate of a defendant. Therefore, it is very important that as many individuals as possible get it right, in order for the group to increase the chance of getting it right. So, we have very accurate individuals, but that due to public communication are completely correlated with each other. Consequently, as Observation 2.3.1 shows, a voting rule has basically no effect. Whatever voting rule you use, agents all hold the same belief, so any voting rule, as defined in section 2.1, will yield the same outcome.

[2]Coughlan makes a distinction between different strategies that agents can have while voting. However, in this thesis we will focus only on the non-strategic element of Coughlan's work

2.4 Regular communication networks

At first sight, Condorcet's and Coughlan's frameworks couldn't be further apart. Condorcet allows for no communication, and Coughlan assumes full information exchange between agents. Yet, what Condorcet and Coughlan have in common is that their networks can both be seen as regular graphs, as will be shown in this section.

We will be using graphs as representations of communication networks. Therefore, we will interpret vertices as agents and every edge as a connection between agents that enables communication. Following van Steen (2010), we set out the following formal concepts and notations from graph theory.

Definition 2.4.1 (Graph)**.** A graph $G = (V, E)$ consists of a collection of vertices V and edges E. Each edge $e \in E$ is said to join two vertices, which are called its *end points*. If e joins $i, j \in V$ we write $e = \langle i, j \rangle$. In this case, i, j are *adjacent* and e is *incident* with i and j respectively.

The terms *network* and *graph* will be used interchangeably in the sequel, as well as the terms *vertex* and *node* and the terms *edge* and *link*. What we just defined as a graph is also called an *undirected graph*, because edges have no direction. This is because we use graphs to represent communication networks, and we assume that if there is a connection between two agents, information flows in both ways. Thus, for the remainder of the thesis when we talk about *graphs* we will always be talking about undirected graphs.

An important property of a vertex is the number of edges that are incident with it. This number is called the *degree* of a vertex.

Definition 2.4.2 (Degree)**.** The number of edges incident with a vertex i - denoted δ_i- is called the *degree* of i.

Using only the definition of a graph and degree, we can define the following special classes of graphs: empty graphs, complete graphs and regular graphs.

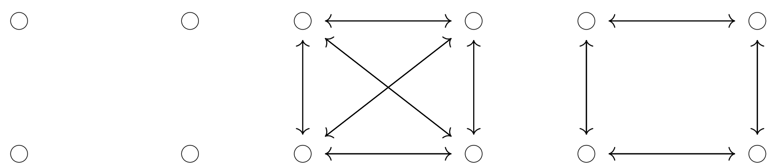
Definition 2.4.3 (Empty graph)**.** A graph $G = (V, E)$ such that $\delta_i = 0$ for all $i \in V$ is called an empty graph of size $|V|$.

Definition 2.4.4 (Complete graph)**.** A graph $G = (V, E)$ such that $\delta_i = |V| - 1$ for all $i \in V$ is called a complete graph of size $|V| - 1$.

Definition 2.4.5 (Regular Graph)**.** A graph G is regular if all vertices have the same degree δ. We denote by $G(n, \delta)$ a regular graph with n vertices that all have degree δ.

An example of an empty graph and a complete with 4 vertices is given in figure 2.1a and figure 2.1b respectively. An example of a regular graph in which all vertices have degree 2 is given in figure 2.1c.

Note that there is no consistency in the literature when it comes to denoting graphs without edges; they are also called *null graphs* and *trivial graphs*. The concept is important for us though, since we know that the CJT does not allow



(a) Empty Graph (b) Complete Graph (c) Regular Graph

Figure 2.1: Examples of three types of graphs

for any links between agents in a network. This means that there are no edges and the CJT thus applies to empty graphs only. In Coughlan's framework, on the other hand, all agents are connected with each other. Therefore, we will say that Coughlan's framework applies to *complete graphs* only. We chose to define complete graphs as graphs in which $\delta_i = |V| - 1$ for all vertices, rather than $\delta_i = |V|$ for all vertices. It is true that each agent receives a signal from itself, but since this is always the case - also in Condorcet's model - we chose to leave out these reflexive arrows for simplicity in this thesis. It is important to keep in mind though that we assume that in every network - whether empty, complete or regular - agents always receive a signal from themselves.

We can see that empty graphs and complete graphs are a special case of regular graphs, one where the degree is 0 or $|V| - 1$ for all vertices respectively. Thus regardless of the differences between the CJT and Coughlan's model, they have in common that they both apply to regular graphs.

2.5 Partial communication on regular networks

So far we have seen Condorcet's model with no communication on the one hand, and Coughlan's model with full communication on the other hand. The main goal of this thesis will be to generalise both theorems to partial communication settings, that is, where some agents talk to some other agents, but not necessarily to everyone. In the remainder of this chapter, we will look at partial communication in regular networks. We will be interested in two main things: the expected accuracy of an individual agent after communication and the expected accuracy of the majority of the group. We will determine them in this respective order.

2.5.1 Partial communication and individual accuracy

In this section we will determine the probability that an agent is correct in a regular network of a given degree. Since in this chapter we focus on h-regular networks only, we know that each agent receives the same number h of signals from other agents. In section 2.1 we introduced the posterior π and the threshold t for belief such that an agent will believe a to be the correct alternative if and only if the posterior belief π in a is strictly bigger than the threshold t. We now are going to determine when the posterior π is bigger than our threshold t and express this in terms of the number of k a-signals that an agent has to receive. In other words, how many signals of a does one have to receive, in order to believe that a is the true state of the world? This question is answered by the following lemma. Note that throughout the thesis we assume the base of log to be 2.

Lemma 2.5.1. *Let* $\tau = \frac{h}{2} + \frac{1}{2}\frac{\log(\frac{t}{1-t}) - \log(\frac{r}{1-r})}{\log(\frac{p}{1-p})}$,

$$\pi > t \text{ if and only if } k > \tau$$

Proof. Using Observation 2.1.1 to fill in π:

$$\begin{aligned}
\pi &> t && \text{iff} \\
\frac{rp^k(1-p)^{(h-k)}}{rp^k(1-p)^{(h-k)} + (1-r)p^{(h-k)}(1-p)^k} &> t && \text{iff} \\
\frac{rp^k(1-p)^h}{(1-p)^k} &> \frac{trp^k(1-p)^h}{(1-p)^k} + \frac{t(1-r)p^h(1-p)^k}{p^k} && \text{iff} \\
rp^{2k}(1-p)^h &> trp^{2k}(1-p)^h + t(1-r)p^h(1-p)^{2k} && \text{iff} \\
p^{2k}(r(1-p)^h - tr(1-p)^h) &> t(1-r)p^h(1-p)^{2k} && \text{iff} \\
p^{2k}r(1-p)^h(1-t) &> t(1-r)p^h(1-p)^{2k} && \text{iff} \\
\frac{p^{2k}}{(1-p)^{2k}} &> \frac{t(1-r)p^h}{(1-t)r(1-p)^h} && \text{iff} \\
\left(\frac{p^2}{(1-p)^2}\right)^k &> \frac{t(1-r)p^h}{(1-t)r(1-p)^h} && \text{iff} \\
k &> \log_{\frac{p^2}{(1-p)^2}}\left(\frac{(t-tr)}{(r-tr)}\left(\frac{p}{1-p}\right)^h\right) \\
&> \frac{\log(\frac{t-tr}{r-tr}(\frac{p}{1-p})^h)}{\log\frac{p^2}{(1-p)^2})} \\
&> \frac{h}{2} + \frac{1}{2}\frac{\log(\frac{t}{1-t}) - \log(\frac{r}{1-r})}{\log(\frac{p}{1-p})}
\end{aligned}$$

□

Looking at τ, we make a few interesting observations of the influence of the parameters on the value of τ. Firstly, we can see that the prior r and the threshold t have a similar effect on the number of k signals an agent has to receive such that $\pi > t$. If the prior is low, meaning that the agent has a strong prior belief in b, or if the threshold is very high, then the agent needs relatively many signals in favour of a to believe a. Meanwhile, if the prior for a is high or the threshold is low, then the agent needs relatively few signals in favor of a

to believe a. Intuitively, this is what we would expect. After all, a high prior indicates that an agent already attaches a strong belief to a, and thus it makes sense that it needs less evidence in favour of a to believe a. A low threshold indicates that with a relatively low probability an agent will be convinced of a and thus it will need few signals to believe a.

Secondly, we can see that τ increases linearly in h. This means that if h becomes twice as big, then τ becomes twice as big. This intuitively makes sense: assuming that the remaining parameter values stay fixed, then if an agent receives double the number of h total signals, it also needs double the number of k a-type signals for her posterior to cross the threshold.

Thirdly, the higher the value of p, the closer the value of τ to $\frac{h}{2}$. Namely, as p increases, the value of $\log(\frac{p}{1-p})$ increases. Consequently, the value of $\frac{h}{2} + \frac{1}{2}\frac{\log(\frac{t}{1-t})-\log(\frac{r}{1-r})}{\log(\frac{p}{1-p})}$ decreases, bringing the value of τ closer to $\frac{h}{2}$. As a result, the prior r and the threshold t have less weight in determining the value k in the case p is high. This is an interesting observation, as it implies that competent agents attach more value to the signals they receive, while less competent agents attach relatively more weight to their prior and the threshold.

The intuition behind this observation is that as competence increases, the agent trusts the signals with higher confidence. During the Bayesian update, the agent puts a weight on the signals she receives and then derives a posterior. These weights depend on the informativeness of each signal. The more informative a signal, the more weight it receives. An increase in p is essentially an increase in the informativeness of the signal. In our model, a signal that is correct with probability 0.5 is not informative it all. In this case an agent will only trust its prior and discard the signal. A signal that is correct with probability 0.95 is very informative. Or more extremely, if $p = 1$ then no matter the agent's prior, it will trust the signal. Therefore, as p increases, the agent will attach more weight to the signals and thus τ moves to $\frac{h}{2}$.

Fourthly, we can observe that if $t = r$ then k reduces to $\frac{h}{2}$. This is stated in the following corollary.

Corollary 2.5.1.1. *If $t = r$ then $\pi > t$ if and only if $k > \frac{h}{2}$.*

Proof. If $t = r$ then $\log(\frac{t}{1-t}) - \log(\frac{r}{1-r}) = 0$ and thus $k > \frac{h}{2}$. □

This is a useful observation that we will use later on. In words it means that if the prior and the threshold are equal, then for the agent to believe that the state of the world is a, she must receive a majority of a-type signals. That is, strictly more than half of the signals an agent receives has to be of the a-type in order for her to believe that a is the state of the world. This is the set-up of so-called *maximum-likelihood estimation.* We will focus on this setting in section 3.3 in the next chapter.

It is worth mentioning already that although this result might come as a surprise at first, it does make sense. After all, if $r = t$ this means that the prior belief in a of an agent is equal to the threshold to belief a. Consequently, the agent is completely unbiased and therefore attaches equal weight to each incoming signal. An example might illustrate this. Suppose $r = t = 0.5$. In this case it is easy to see when it comes to the prior of the agent, it is completely unbiased as to whether a or b is the state of the world. The threshold value of 0.5 indicates the same thing: it makes the agent unbiased as to whether a or b is the true state of the world [3] Since the agent has no bias whatsoever, it makes sense that it lets its belief be determined by the observed signals only. It does so in an unbiased way by attaching equal weight to each signal by letting k be equal to $\frac{h}{2}$. Now in other cases that $r = t$, the agent is also unbiased. Of course, looking only at the value $r \neq 0.5$ the agent has a bias, but the point is that this biased is *compensated* or weighted out by letting $r = t$. What we can thus learn from this observation, is that the parameters r and t are two different ways of creating a bias towards one of the two alternatives. In the special case that $r = t$ the biases are balanced out such that the agent ends up being unbiased.

Lemma 2.5.1 gives us the probability that k is larger than a certain number. Now, we calculate that probability for a given h, based on the distribution of k given h and given that $\theta = a$. So, we want to know, given a specific total number of signals h and given that $\theta = a$, the probability that the number of a-type signals is big in enough to meet the threshold. This is given by the following lemma.

Lemma 2.5.2. *Let* $\tau = \frac{h}{2} + \frac{1}{2}\frac{\log(\frac{t}{1-t}) - \log(\frac{r}{1-r})}{\log(\frac{p}{1-p})}$. *The probability that the number of k a-type signals exceeds the threshold t, given h and $\theta = a$ is:*

$$P(k > \tau | \theta = a) = \sum_{k=\lfloor \tau \rfloor + 1}^{h+1} \binom{h+1}{k} p^k (1-p)^{h+1-k}$$

Proof. The probability mass function of a Binomial distribution expresses the probability of getting exactly k successes in n independent Bernoulli trials, where the chance of getting a correct signal is given by p and the chance of getting an incorrect signal is given by $(1 - p)$. Now the probability of getting k guilty signals follows a Binomial distribution, only we are interested in the probability of getting *more than* τ many k successes, as proven in Lemma 2.5.1. That is, we want to know the probability that an agent gets more than τ correct signals. We will express this by $k = \lfloor \tau \rfloor + 1$. The upper bound is of k is given by $h + 1$, as an agent receives h many signals from other agents and always 1 signal from itself. □

In sum, the formula expresses the probability of getting exactly k successes in $h + 1$ independent Bernoulli trials, where k ranges from $\lfloor \tau \rfloor + 1$ till $h + 1$.

Similarly, we calculate the probability of getting b right. That is, we want to express the probability that the number of a-type signals is small enough such that it does not exceed the threshold, given h and given that the b would be the true state of the world.

[3] This is not *completely* true. There is a slight bias towards b as the threshold is strict, but the effect is negligible for this intuitive explanation.

Lemma 2.5.3. *Let* $\tau = \frac{h}{2} + \frac{1}{2}\frac{\log(\frac{t}{1-t}) - \log(\frac{r}{1-r})}{\log(\frac{p}{1-p})}$. *The probability that the number of k a-type signals does not meet the threshold t, given h and $\theta = b$ is:*

$$P(k \leq \tau | \theta = b) = \sum_{k=0}^{\lfloor \tau \rfloor} \binom{h+1}{k} p^{h+1-k}(1-p)^k$$

Proof. As in the proof of Lemma 2.5.2, the probability of getting k a-type signals follows a Binomial distribution. In this case, building on Lemma 2.5.1, we are interested in the probability of getting *less than or equal to* τ many k a-type signals. We will express this by $k = \lfloor \tau \rfloor$, which is now our upper bound of k. The lower bound of k is given by 0: for an agent to believe that the $\theta = b$, she has to receive more than τ many signals. That means that for her to believe $\theta = b$, she can receive from 0 to τ many a-type signals. Since each agent will always receive 1 signal from herself, we can see this as $h + 1$ many trials. □

In sum, the formula expresses the probability of getting exactly k a-type in $h + 1$ independent Bernoulli trials, where k ranges from 0 till $\lfloor \tau \rfloor$.

If we put Lemma 2.5.2 and Lemma 2.5.3 together, we obtain the probability that an agent is correct in general after information exchange. We thus obtain the following theorem, that expresses the expected accuracy of an agent in any regular network of size h.

Theorem 2.5.4. *Let* $p_{(n,h)}$ *be the expected individual accuracy in a regular network* $G = (n, \delta)$ *where* $\delta = h$. *Then for a regular network with n nodes in which each* $\delta_i = h$, *where h is fixed,* $p_{(n,h)}$ *is given by:*

$$\begin{aligned} p_{(n,h)} =& \mathbb{P}(\pi > t | \theta = a) \cdot \mathbb{P}(\theta = a) + \mathbb{P}(\pi \leq t | \theta = b) \cdot \mathbb{P}(\theta = b) \\ =& \left(\sum_{k=\lfloor \tau \rfloor + 1}^{h+1} \binom{h+1}{k} p^k (1-p)^{h+1-k} \right) \cdot r \\ &+ \left(\sum_{k=0}^{\lfloor \tau \rfloor} \binom{h+1}{k} p^{h+1-k}(1-p)^k \right) \cdot (1-r) \end{aligned}$$

where $\tau = \frac{h}{2} + \frac{1}{2}\frac{\log(\frac{t}{1-t}) - \log(\frac{r}{1-r})}{\log(\frac{p}{1-p})}$.

Proof. This follows from combining Lemma 2.5.1, Lemma 2.5.2 and Lemma 2.5.3 and the fact that the probability of being correct consists of the probability of believing a in case $\theta = a$ and believing b in case $\theta = b$. □

Observe that since in a regular network the number of signals h an agent receives is the same for each agent, it follows that the expected accuracy $p_{(n,h)}$ will have the same value for each agent. We thus ended up with a way to calculate the expected accuracy of an individual agent in any regular network.

We will now show how Theorem 2.5.4 can capture both Coughlan's as well as Condorcet's model. In Condorcet's model, the network is empty, that is no one communicates with each other. As a result, the accuracy $p_{(n,h)}$ reduces to the accuracy p of an individual agent. This makes sense: since agents do not communicate with each other, their accuracy cannot improve due to communication and thus their accuracy remains unchanged. We prove this in the following corollary.

Corollary 2.5.4.1. *If $h = 0$ then $p_{(n,h)} = p$.*

Proof.

$$\begin{aligned} p_{(n,0)} =& \left(\sum_{k=1}^{1} \binom{1}{k} p^k (1-p)^{1-k} \right) \cdot r + \left(\sum_{k=0}^{0} \binom{1}{k} p^{1-k} (1-p)^k \right) \cdot (1-r) \\ =& pr + p(1-r) \\ =& p \end{aligned}$$

□

In Coughlan's model, the network is complete, that is, every agent communicates with every other agent. The resulting formula to calculate the accuracy of each individual agent is stated in the following corollary.

Corollary 2.5.4.2. *If $h = n - 1$ then*

$$p_{(n,n-1)} = \left(\sum_{k=\lfloor \tau \rfloor + 1}^{n} \binom{n}{k} p^k (1-p)^{n-k} \right) \cdot r + \left(\sum_{k=0}^{\lfloor \tau \rfloor} \binom{n}{k} p^{n-k} (1-p)^k \right) \cdot (1-r)$$

Proof. We simply substitute h with $n - 1$. □

We finally formulate a lemma that will be of importance in the next section.

Lemma 2.5.5. *If $p > 0.5$ then $p_{(n,h)} \geq p > 0.5$.*

Proof. This follows from the fact that $p_{(n,0)} = p$ and $p_{(n,h+1)} \geq p_{(n,h)}$. The former follows from Corollary 2.5.4.1. The latter follows from the observation that $p_{(n,h+1)} - p_{(n,h)}$ is positive. □

Corollary 2.5.5 states that if the competence of each agent is strictly bigger than a half, that is, an agent is more likely than not to get it right, then the expected accuracy of that agent will also be higher than a half. Moreover, observe that since $p_{(n,h+1)} - p_{(n,h)} \geq 0$, it follows that the larger h gets, the better the accuracy.

2.5.2 Partial communication and group accuracy

In this final section of the chapter we will determine the probability that the majority of agents is correct in a regular network for any degree. In particular, we will generalize the asymptotic result of the CJT to regular networks for any degree.

Theorem 2.5.6. *Let $k \in \mathbb{N}$. If $\delta_i = h$ for all $v \in V$ and $p > 0.5$ then* $\lim_{n \to \infty} p_{maj}(n) = 1$.

Proof. Each agent receives h signals that each have an accuracy $p > 0.5$. By equation 2.3 of the CJT (Theorem 2.2.1), the probability that the majority of these k signals is correct is bigger or equal than p. Thus, the probability that the majority of the k signals each agent receives is correct is strictly bigger than 0.5. So, in expectation, strictly more than half of the agents have access to a majority of correct signals. Then, by the Law of Large Numbers $\lim_{n \to \infty} p_{maj}(n) = 1$ □

We thus ended up with generalized version of the asymptotic result of the CJT for regular networks with any degree. In particular, this means that that for both Coughlan's as well as Condorcet's model, the probability that the majority of agents is correct will be 1 as the size of the networks approaches infinity.

2.6 Summary

In this chapter we set out the framework of Condorcet and Coughlan. Condorcet shows that without communication, groups are better at identifying the truth than individuals. Coughlan formulated a model for public communication. While the two models are at opposite ends when it comes to the form of communication, the structure of the communication networks shares one thing: it is a regular network, meaning that every agent is connected to the same number of other agents.

We set out to determine what is left of the group accuracy and the individual accuracy when partial communication takes place in a regular network. We formulated Theorem 2.5.4 to calculate the expected accuracy of an individual agent in any regular network. Furthermore, we showed in Theorem 2.5.6 that the asymptotic result of the Condorcet Jury theorem remains intact in any regular network.

Chapter 3

Individual Accuracy in Random Networks

In the previous chapter we became acquainted with Theorem 2.2.1 about group accuracy under no information exchange and with Observation 2.3.1 about group accuracy under full information exchange. We observed that both theorems are only valid on regular networks. We already extended both the expected individual's accuracy as well as the accuracy of the group under majority rule to regular networks with degrees of any number. The goal of the remainder of the thesis will be to determine the individual accuracy as well as the group accuracy for the class of random networks rather than the class of regular networks. In this chapter we will look at the expected accuracy of an individual agent in a random network.

We proceed as follows. In section 3.1 we introduce the formal definition of a random network. In section 3.2 we formulate a theorem that expresses the expected individual accuracy for any agent in a random network. Next, we will determine the influence of various parameters on the expected individual accuracy. Finally, in section 3.3 we show that the expected individual accuracy that is determined via Bayesian inference, becomes equal to the maximum likelihood estimation approach in case the threshold is equal to the prior.

3.1 Random communication networks

Following van Steen (2010), we introduce the basic concept of a random network. The basic idea of a random network is that we are given a graph with n vertices, and every two vertices are connected with a probability c. Paul Erdös and Alfred Rényi (1959) were the first to introduce random networks and hence the Erdös-Rényi networks are now known as the classical random networks. Our definition of a random network is what is called in the literature an *Erdös-Rényi Random Network*.

Definition 3.1.1 (Random Network)**.** Let $n \in \mathbb{N}$ and $0 \leq c \leq 1$. A random network $G(n,c)$ is graph $G = (V, E)$ such that each two vertices are connected by an edge with probability c.

A random graph induces a probability distribution over all possible graphs on n. Therefore, for a given n and c, a graph $G'(n,c)$ and another graph $G''(n,c)$ can be very different. Although both graphs will have n vertices, there may be edges in $G'(n,c)$ that are not in $G''(n,c)$ and the other way around. Note that in a random network $G(n,c)$ there are no reflexive arrows and there is at most one edge between two distinct vertices.

Among the graphs that that a random graph induces are all the regular graphs on n. In particular, for any graph G the probability of obtaining G from $G(n,c)$ is $\mathbb{P}(G) = c^m(1-c)^{\binom{n}{2}-m}$ where m is the number of edges in G.

We are interested in looking at random graphs rather than regular graphs only, because the class of random graphs extends the class of regular graphs. This implies that anything we say about random graphs, will also apply to regular graphs. All that we said in the previous chapter, can thus be seen as a kind of *warm-up* for a more general framework that will be laid out in the remainder of the thesis.

The agents in the random networks that we study in this thesis always have a certain competence and prior. Furthermore, there is always a given threshold t that determines when an agent believes an alternative to be the state of the world. We therefore introduce the following shorthand notation of a *uniformly competent random network*.

Definition 3.1.2 (Uniformly competent random network (UCRN))**.** A uniformly competent random network is a tuple $\langle N, c, p, r, t \rangle$ with $|N| = n$, consisting of a random network $G(n,c)$ where each n has competence p, prior r and threshold t.

UCRN's are the networks that we will be concerned with. They are the most general networks, as regular graphs as also among UCRN's. In the next section, we will define the expected accuracy of agents in an UCRN.

3.2 Individual accuracy in random networks

In the previous chapter we showed how to determine the expected accuracy of an agent in a regular network. In a regular network, we are given how many signals an agent receives. The only varying element in a regular network is thus the ratio between correct and incorrect signals an agent receives. Meanwhile, in a random network we add another varying element, namely the number of signals an agent receives. To generalise the expected accuracy of an agent from regular networks to random networks, we thus have to determine the probability that an agent receives a certain number of signals.

We are given a random network $G(n,c)$ and we are interested in the probability that an agent receives h signals from *other agents*. We say *other* agents, because we assume that each agent receives a signal from itself and thus the only random factor is the number of links it has in the network. Then h is a random variable that is determined by a binomial distribution. That is, we view the process of creating links as $n-1$ independent Bernoulli trials, where a link is created with probability c. The probability that an agent receives h signals from *other* agents is given by:

$$\mathbb{P}(h) = \binom{n-1}{h} c^h (1-c)^{(n-1)-h} \tag{3.1}$$

Now we can update Theorem 2.5.4 with equation 3.1, to obtain the expected accuracy of an agent in a random network.

Theorem 3.2.1. *Let $p_{(n,c)}$ be expected individual accuracy in an UCRN $\langle N, c, p, r, t\rangle$, then $p_{(n,c)}$ is given by:*

$$\begin{aligned}
p_{(n,c)} =& \mathbb{P}(\pi > t|\theta = a) \cdot \mathbb{P}(\theta = a) + \mathbb{P}(\pi \leq t|\theta = b) \cdot \mathbb{P}(\theta = b) \\
=& \sum_{h=0}^{n-1} \left(\sum_{k=\lfloor\tau\rfloor+1}^{h+1} \binom{h+1}{k} p^k (1-p)^{h+1-k} \cdot \binom{n-1}{h} c^h (1-c)^{(n-1)-h} \right) \cdot r \\
+& \sum_{h=0}^{n-1} \left(\sum_{k=0}^{\lfloor\tau\rfloor} \binom{h+1}{k} p^{h+1-k} (1-p)^k \cdot \binom{n-1}{h} c^h (1-c)^{(n-1)-h} \right) \cdot (1-r)
\end{aligned}$$

where $\tau = \frac{h}{2} + \frac{1}{2} \frac{\log(\frac{t}{1-t}) - \log(\frac{r}{1-r})}{\log(\frac{p}{1-p})}$

Proof. Let us focus on the left-hand side of the equation. We saw in Lemma 2.5.1 that the probability that the number of k a-type signals exceeds the threshold t, given h and $\theta = a$ is:

$$P(k > \tau|\theta = a) = \sum_{k=\lfloor\tau\rfloor+1}^{h+1} \binom{h+1}{k} p^k (1-p)^{h+1-k}$$

To get the probability of getting guilt correct for arbitrary h we add up all the probabilities of k meeting the threshold for each specific h, where h can range from 0 to $n-1$. This is because we let h denote the number of signals an agent gets from *other* agents, of which there are $n-1$ many. We then multiply these probabilities with the probability for each h to occur. This is given by equation 3.1:

$$\binom{n-1}{h} c^h (1-c)^{(n-1)-h}$$

Finally, we multiply this whole expression with the prior probability of a, which is r. As such, we obtain that the probability of getting a correct, denoted by p^a, for arbitrary h is:

$$\begin{aligned}p^a =&\mathbb{P}(\pi > t|\theta = a) \cdot \mathbb{P}(\theta = a)\\ =&\sum_{h=0}^{n-1}\left(\sum_{k=\lfloor\tau\rfloor+1}^{h+1}\binom{h+1}{k}p^k(1-p)^{h+1-k}\cdot\binom{n-1}{h}c^h(1-c)^{(n-1)-h}\right)\cdot r\end{aligned}$$

The right-hand side of the equation is obtained using similar reasoning. We saw in Lemma 2.5.3 that the probability that the number of k a-type signals does not exceed the threshold t, given h and $\theta = b$ is:

$$\mathbb{P}(k \leq \tau|\theta = b) = \sum_{k=0}^{\lfloor\tau\rfloor}\binom{h+1}{k}p^{h+1-k}(1-p)^k$$

To get the probability of getting b correct for arbitrary h we add up all the probabilities of k not meeting the threshold, where h ranges from 0 to $n-1$, and multiply these probabilities with the probability for each h to occur. Finally we multiply this whole expression with the prior probability of b, which is $(1-r)$. As such we obtain that the probability of getting b correct,denoted by p^b, for arbitrary h is given by:

$$\begin{aligned}p^b =&\mathbb{P}(\pi \leq t|\theta = b) \cdot \mathbb{P}(\theta = b)\\ =&\sum_{h=0}^{n-1}\left(\sum_{k=0}^{\lfloor\tau\rfloor}\binom{h+1}{k}p^{h+1-k}(1-p)^k\cdot\binom{n-1}{h}c^h(1-c)^{(n-1)-h}\right)\cdot(1-r)\end{aligned}$$

So, now we have the probability of an agent believing a when a is the true state of the world - $P(\pi > t|\theta = a) \cdot P(\theta = a)$ - and the probability of an agent believing b when b is the true state of the world - $P(\pi \leq t|\theta = b) \cdot P(\theta = b)$. Since these two events are disjoint, we simply add them together to get the probability of an agent to be correct. □

Notice that due to the way we set up the threshold t, it is impossible for an agent to be indecisive. That is, it will *always* believe that either a is the true alternative, or that b is the true alternative. In particular, we determined that an agent believes a if and only if its posterior belief in a is *strictly* bigger than the threshold t. Obviously, this means that our model has a bias towards alternative b. It is important to realise that this bias carries over to Theorem 3.2.1. Because there is a slight bias towards believing b, this means that agents will have a slightly higher probability to be correct in case b is the true alternative.

In the remainder of this section, we will determine the influence of each single parameter on the equation given in Theorem 3.2.1. We will use methods from comparative statics, which is a method used in economics to determine the change in the outcome of a model as a result from changes in the parameters of the model (see e.g. Kehoe (1989)). The method makes use of differential calculus to analyse the impact of an infinitesimal change in the parameter of the model to its outcome - in our case the expected accuracy. In particular, we will derive the first partial derivative of each parameter that appears in the function for expected accuracy. The nature of this derivative - either positive or negative - will then indicate whether an infinitesimal increase in the parameter will increase or decrease the expected accuracy.

3.2.1 Influence on $p_{(n,c)}$ of the size n of the network

We will investigate what happens when there is one agent and when the number of agents approaches infinity. Finally, we will determine what happens as the number of agents grows in general.

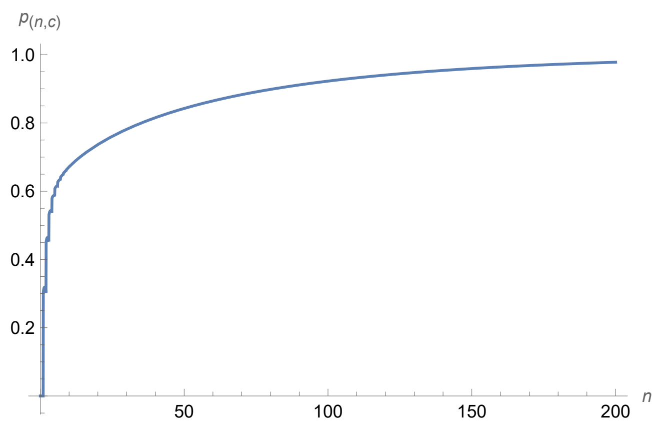


Figure 3.1: Influence of n on probability of correctness. Remaining parameter values are fixed to $c = 0.5, p = 0.6, r = 0.5, t = 0.5$.

As a start, we plotted the influence of the size of the network on the expected accuracy $p_{(n,c)}$ in figure 3.1. As can be seen, generally the accuracy increases as the size of the networks grows. We prove this formally in the following theorem.

Theorem 3.2.2. *As n increases, $p_{(n,c)}$ increases.*

Proof. Since n is a discrete variable, taking only integer values, there is no partial derivative of $p_{(n,c)}$ with respect to n. We therefore compute the value for $p_{((n+1),c)}$ and compare it with the value for $p_{(n,c)}$ as follows.

$$p_{((n+1),c)} - p_{(n,c)} =$$

$$(\sum_{h=0}^{n}\left(\sum_{k=\lfloor\tau\rfloor+1}^{h+1}\binom{h+1}{k}p^k(1-p)^{h+1-k}\cdot\binom{n}{h}c^h(1-c)^{n-h}\right)\cdot r$$

$$+\sum_{h=0}^{n}\left(\sum_{k=0}^{\lfloor\tau\rfloor}\binom{h+1}{k}p^{h+1-k}(1-p)^k\cdot\binom{n}{h}c^h(1-c)^{n-h}\right)\cdot(1-r))$$

$$-(\sum_{h=0}^{n-1}\left(\sum_{k=\lfloor\tau\rfloor+1}^{h+1}\binom{h+1}{k}p^k(1-p)^{h+1-k}\cdot\binom{n-1}{h}c^h(1-c)^{(n-1)-h}\right)\cdot r$$

$$+\sum_{h=0}^{n-1}\left(\sum_{k=0}^{\lfloor\tau\rfloor}\binom{h+1}{k}p^{h+1-k}(1-p)^k\cdot\binom{n-1}{h}c^h(1-c)^{(n-1)-h}\right)\cdot(1-r)))$$

There are two differences between $p_{((n+1),c)}$ and $p_{(n,c)}$. Firstly, for $p_{((n+1),c)}$ the parameter h is summed from 0 to n, whereas for $p_{(n,c)}$ parameter h is summed from 0 to $n-1$. Therefore, the first difference is that an extra value for h, namely $h = n$ is evaluated. The following formula expresses this:

$$\begin{aligned}&\left(\sum_{k=\lfloor\tau\rfloor+1}^{n+1}\binom{n+1}{k}p^k(1-p)^{n+1-k}\cdot\binom{n}{n}c^n(1-c)^{n-n}\right)\cdot r\\&+\left(\sum_{k=0}^{\lfloor\tau\rfloor}\binom{n+1}{k}p^{n+1-k}(1-p)^k\cdot\binom{n}{n}c^n(1-c)^{n-n}\right)\cdot(1-r)\end{aligned} \tag{3.2}$$

Equation 3.2 positive, as n, k, p, c, r are all positive values. Furthermore the domains of p and c are an open interval between 0 and 1, so $(1-p)$ and $(1-c)$ cannot become a negative value.

The second difference is that in $p_{((n+1),c)}$ the summation is being multiplied with

$$\binom{n}{h}c^h(1-c)^{n-h}$$

Meanwhile in $p_{(n,c)}$ the same summation is being multiplied with

$$\binom{n-1}{h}c^h(1-c)^{(n-1)-h}$$

Now if the numerator in binomial coefficient increases, this will increase the outcome.

$$
\begin{aligned}
&\binom{n}{h} c^h (1-c)^{n-h} - \binom{n-1}{h} c^h (1-c)^{(n-1)-h} \\
=&\frac{(1-c)^{n-h} c^h \left(\frac{(n-1)!}{(c-1)(n-h-1)!} + \frac{n!}{(n-h)!}\right)}{h!}
\end{aligned} \tag{3.3}
$$

Again since c, n, h are all positive and the domain of c is an open interval between 0 and 1, it follows that equation 3.3 is a positive expression.

Thus if we change the value from n to $n+1$ two parts of the expression of $p_{(n,c)}$ are affected. We have shown that these changes both result in $p_{((n+1),c)}$ being bigger than $p_{(n,c)}$. Hence, this implies that $p_{(n,c)}$ grows as n grows. □

In Theorem 3.2.3 we show that if there is only one agent, the probability that she will be correct is equal to her competence.

Theorem 3.2.3. *If* $n = 1$ *then* $p_{(n,c)} = p$

Proof.

$$
\begin{aligned}
p_{(n,c)} =&\sum_{h=0}^{n-1} \left(\sum_{k=\lfloor \tau \rfloor + 1}^{h+1} \binom{h+1}{k} (p^k (1-p)^{h+1-k} \cdot \binom{n-1}{h} c^h (1-c)^{(n-1)-h} \right) \cdot r \\
&+ \sum_{h=0}^{n-1} \left(\sum_{k=0}^{\lfloor \tau \rfloor} \binom{h+1}{k} p^{h+1-k} (1-p)^k \cdot \binom{n-1}{h} c^h (1-c)^{(n-1)-h} \right) \cdot (1-r) \\
=&\sum_{h=0}^{0} \left(\sum_{k=\lfloor \frac{0}{2} \rfloor + 1}^{1} \binom{1}{k} (p^k (1-p)^{1-k} \right) \cdot r \\
&+ \sum_{h=0}^{0} \left(\sum_{k=0}^{\lfloor \frac{0}{2} \rfloor} \binom{1}{k} p^{1-k} (1-p)^k \right) \cdot (1-r) \\
=&p \cdot r + p \cdot (1-r) \\
=&p
\end{aligned}
$$

□

In the last theorem of this subsection, we show that as the network size goes to infinity, the expected accuracy goes to 1.

Theorem 3.2.4. $\lim_{n \to \infty} p_{(n,c)} = 1$

Proof. As $n \to \infty$, by the weak law of large numbers the average number of correct signals converges to p. Since $p > 0.5$, it follows that $p_{(n,c)} \to 1$. □

3.2.2 Influence on $p_{(n,c)}$ of the connectednes c of the network

In this section we investigate the influence of the probability that any two nodes are connected on the expected accuracy. As a start, we plotted the influence of c on $p_{(n,c)}$. As can be seen in figure 3.2, the accuracy increases as the value of c increases. We prove this formally in the following theorem.

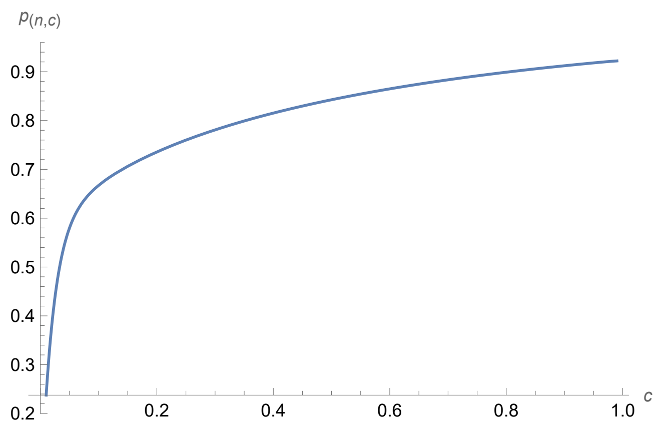


Figure 3.2: Influence of c on probability of correctness. Remaining parameter values are fixed to $n = 50, p = 0.6, r = 0.5, t = 0.5$.

Theorem 3.2.5. *If c increases, then $p_{(n,c)}$ increases.*

Proof. The partial derivative of $p_{(n,c)}$ with respect to parameter c is:

$$\frac{\partial p_{(n,c)}}{\partial c} =$$

$$\frac{\partial}{\partial c}\sum_{h=0}^{n-1}\left(\sum_{k=\lfloor\tau\rfloor+1}^{h+1}\binom{h+1}{k}p^k(1-p)^{h+1-k}\cdot\binom{n-1}{h}c^h(1-c)^{(n-1)-h}\right)\cdot r$$

$$+\frac{\partial}{\partial c}\sum_{h=0}^{n-1}\left(\sum_{k=0}^{\lfloor\tau\rfloor}\binom{h+1}{k}p^{h+1-k}(1-p)^k\cdot\binom{n-1}{h}c^h(1-c)^{(n-1)-h}\right)\cdot(1-r)$$

$$=\sum_{h=0}^{n-1}\left(\sum_{k=\lfloor\tau\rfloor+1}^{h+1}\frac{\partial}{\partial c}\left(\left(\binom{h+1}{k}p^k(1-p)^{h+1-k}\cdot\binom{n-1}{h}c^h(1-c)^{(n-1)-h}\right)\cdot r\right)\right.$$

$$\left.+\sum_{h=0}^{n-1}\left(\sum_{k=0}^{\lfloor\tau\rfloor}\frac{\partial}{\partial c}\left(\left(\binom{h+1}{k}p^{h+1-k}(1-p)^k\cdot\binom{n-1}{h}c^h(1-c)^{(n-1)-h}\right)\cdot(1-r)\right)\right)\right.$$

$$=\sum_{h=0}^{n-1}(\sum_{k=\lfloor\tau\rfloor+1}^{h+1}\binom{h+1}{k}p^k(1-p)^{h+1-k}$$

$$\cdot\binom{n-1}{h}c^h((n-1)-h)(1-c)^{(n-1)-h-1}+hc^{h-1}(1-c)^{(n-1)-h})\cdot r$$

$$+\sum_{h=0}^{n-1}(\sum_{k=0}^{\lfloor\tau\rfloor}\binom{h+1}{k}p^{h+1-k}(1-p)^k$$

$$\cdot\binom{n-1}{h}c^h((n-1)-h)(1-c)^{(n-1)-h-1}+hc^{h-1}(1-c)^{(n-1)-h})\cdot(1-r)$$

$$=\sum_{h=0}^{n-1}(\sum_{k=\lfloor\tau\rfloor+1}^{h+1}\binom{h+1}{k}p^k(1-p)^{h+1-k}$$

$$\cdot\binom{n-1}{h}(1-c)^{n-h-2}(c^h(n-1-h)+ch(1-c)))\cdot r$$

$$+\sum_{h=0}^{n-1}(\sum_{k=0}^{\lfloor\tau\rfloor}\binom{h+1}{k}p^{h+1-k}(1-p)^k$$

$$\cdot\binom{n-1}{h}(1-c)^{n-h-2}(c^h(n-1-h)+ch(1-c)))\cdot(1-r)$$

From the first to the second line we use the sum rule, from the second to the third line we use the summation rule and from the third to fourth line we use the product rule.

Note that h, k, p, n, τ, r, c can only take positive values. Furthermore $0 < p < 1$ and $0 < c < 1$ and thus $(1-p)$ and $(1-c)$ are positive too. Hence the partial derivative with respect to c - $\frac{\partial p_{(n,c)}}{\partial c}$ - is a positive formula. It thus follows that $p_{(n,c)}$ increases as c increases. □

We make two further interesting observations. Firstly, we show that as c goes to 0, then the accuracy goes to p. This makes sense, as when $c \to 0$, this means that the probability that a node has a link with another node is as good as 0. Therefore, the accuracy of agent reduces to its own competence.

Theorem 3.2.6. *If $c \to 0$ then $p_{(n,c)} \to p$*

Proof. First of all, observe that in case $h > 0$ and $c \to 0$ then $c^h = 0$. As a result,

$$\sum_{h>0}^{n-1} \left(\sum_{k=\lfloor\tau\rfloor+1}^{h+1} \binom{h+1}{k} p^k (1-p)^{h+1-k} \cdot \binom{n-1}{h} c^h (1-c)^{(n-1)-h} \right) = 0$$

and

$$\sum_{h>0}^{n-1} \left(\sum_{k=0}^{\lfloor\tau\rfloor} \binom{h+1}{k} p^{h+1-k} (1-p)^k \cdot \binom{n-1}{h} c^h (1-c)^{(n-1)-h} \right) = 0$$

Therefore, as $c \to 0$ the value of $p_{(n,c)}$ can be simplified for only evaluation for $h = 0$. Observe that if $h = 0$ and $c \to 0$ then $c^h = 0^0 = 1$. $p_{(n,c)}$ thus can be simplified to:

$$\begin{aligned}
p_{(n,c)} =& \left(\sum_{k=\lfloor\tau\rfloor+1}^{1} \binom{1}{k} p^k (1-p)^{1-k} \cdot \binom{n-1}{0} c^0 (1-c)^{(n-1)} \right) \cdot r \\
&+ \left(\sum_{k=0}^{\lfloor\tau\rfloor} \binom{1}{k} p^{1-k} (1-p)^k \cdot \binom{n-1}{0} c^0 (1-c)^{n-1} \right) \cdot (1-r) \\
=& \left(\sum_{k=\lfloor\tau\rfloor+1}^{1} \binom{1}{k} p^k (1-p)^{1-k} \cdot (1-c)^{(n-1)} \right) \cdot r \\
&+ \left(\sum_{k=0}^{\lfloor\tau\rfloor} \binom{1}{k} p^{1-k} (1-p)^k \cdot (1-c)^{n-1} \right) \cdot (1-r) \\
=& \left(\sum_{k=\lfloor\tau\rfloor+1}^{1} \binom{1}{k} p^k (1-p)^{1-k} \right) \cdot r \\
&+ \left(\sum_{k=0}^{\lfloor\tau\rfloor} \binom{1}{k} p^{1-k} (1-p)^k \right) \cdot (1-r) \\
=& rp + (1-r)p \\
=& p
\end{aligned}$$

□

The final interesting observation on the influence of c we make is that as $c \to 1$, then the accuracy approaches that of a complete network. Observe that this is a special case of Theorem 2.5.4, in which we use n instead of $h+1$. This is because the Theorem 2.5.4 is suitable for any regular network, and the case where $c \to 1$ we are looking at a regular, complete network in which every agent receives n signals.

Theorem 3.2.7. *If* $c \to 1$ *then*

$$p_{(n,c)} = \left(\sum_{k=\lfloor\tau\rfloor+1}^{n} \binom{n}{k} p^k(1-p)^{n-k}\right) \cdot r + \left(\sum_{k=0}^{\lfloor\tau\rfloor} \binom{n}{k} p^{n-k}(1-p)^k\right) \cdot (1-r)$$

Proof. First of all, observe that in case $(n-1)-h \neq 0$ and $c \to 1$ then $(1-c)^{(n-1)-h} = 0$. Therefore, as $c \to 1$ the value of $p_{(n,c)}$ can be simplified to only evaluating for $h = (n-1)$. Note that this intuitively makes sense, as it means that if the probability of being connected to any other agent goes to 1, we evaluate $p_{(n,c)}$ only for when an agent receives signals from all other agents $(n-1)$. Filling in $h = (n-1)$ we get:

$$\begin{aligned}
p_{(n,c)} &= \left(\sum_{k=\lfloor\tau\rfloor+1}^{n} \binom{n}{k} p^k(1-p)^{n-k} \cdot \binom{n-1}{n-1} c^{n-1}(1-c)^{(n-1)-(n-1)}\right) \cdot r \\
&+ \left(\sum_{k=0}^{\lfloor\tau\rfloor} \binom{n}{k} p^{n-k}(1-p)^k \cdot \binom{n-1}{n-1} c^{n-1}(1-c)^{(n-1)-(n-1)}\right) \cdot (1-r) \\
&= \left(\sum_{k=\lfloor\tau\rfloor+1}^{n} \binom{n}{k} p^k(1-p)^{n-k} \cdot c^{n-1}\right) \cdot r \\
&+ \left(\sum_{k=0}^{\lfloor\tau\rfloor} \binom{n}{k} p^{n-k}(1-p)^k \cdot c^{n-1}\right) \cdot (1-r)
\end{aligned}$$

As $c \to 1$ we can simplify to:

$$p_{(n,c)} = \left(\sum_{k=\lfloor\tau\rfloor+1}^{n} \binom{n}{k} p^k(1-p)^{n-k}\right) \cdot r + \left(\sum_{k=0}^{\lfloor\tau\rfloor} \binom{n}{k} p^{n-k}(1-p)^k\right) \cdot (1-r)$$

□

In sum, we can conclude that as the probability of connectedness increases, the expected accuracy of agents increases. Furthermore, if $c \to 0$ then the expected accuracy is equal to p and if $c \to 1$ then the expected accuracy is equal to the expected accuracy of an agent in a regular network.

3.2.3 Influence on $p_{(n,c)}$ of the competence p of agents

The influence of an increase in competence p is not what one might expect. Intuitively, we would expect that as the competence p increases, the expected accuracy increases as well. Generally, that is by looking at a rather large increase in p, this is indeed the case. However, a small increase in p turns out to be able to decrease the expected accuracy. Visually, this can be seen in figure 3.3. We see that the overall tendency is to grow, but there are drops in accuracy in between. In this section we will explain why this is the case.

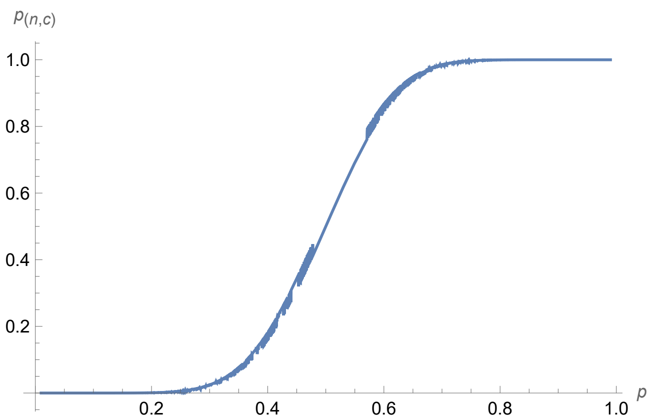


Figure 3.3: Influence of p on probability of correctness. Remaining parameter values are fixed to $n = 50, c = 0.5, r = 0.5, t = 0.5$

We cannot, as was the case for the derivative with respect to n and c before, simply take the partial derivative of $p_{(n,c)}$, as p occurs both in the summation as well as in τ. We will therefore first determine what happens to $\lfloor \tau \rfloor$ with an infinitesimal increase in p. Regarding the value of $\lfloor \tau \rfloor$ there are two different cases for an infinitesimal increase in p:

Case (1) the value of $\lfloor \tau \rfloor$ remains the same

Case (2) the value of $\lfloor \tau \rfloor$ becomes $\lfloor \tau \rfloor - 1$

To see that this is true, recall that

$$\tau = \frac{h}{2} + \frac{1}{2}\frac{\log(\frac{t}{1-t}) - \log(\frac{r}{1-r})}{\log(\frac{p}{1-p})}$$

We take the partial derivative of τ with respect to p:

$$\frac{\partial \tau}{\partial p} = \frac{h(1-p)\left(\frac{p}{(1-p)^2}+\frac{1}{1-p}\right)}{p\log\left(\frac{p^2}{(1-p)^2}\right)} - \frac{(1-p)^2\left(\frac{2p^2}{(1-p)^3}+\frac{2p}{(1-p)^2}\right)\log\left(\frac{(b-br)\left(\frac{p}{1-p}\right)^h}{r-br}\right)}{p^2\log^2\left(\frac{p^2}{(1-p)^2}\right)}$$

$$= -\frac{h\log\left(\frac{p^2}{(p-1)^2}\right)+2\log\left(\frac{b(r-1)\left(\frac{p}{1-p}\right)^h}{(b-1)r}\right)}{(p-1)p\log^2\left(\frac{p^2}{(p-1)^2}\right)}$$

We see that the partial derivative of τ with respect to p is negative, thus τ decreases as the value of p increases. As we take the floor of τ, there are two possible cases. Firstly, τ decreases but $\lfloor\tau\rfloor$ remains the same. Secondly, τ decreases such that $\lfloor\tau\rfloor$ becomes $\lfloor\tau\rfloor - 1$.

Now we will determine what happens to $p_{(n,c)}$ in both cases.

Case (1) Suppose the value of $\lfloor\tau\rfloor$ remains the same. In this case we can take the derivative $p_{(n,c)}$ and treat $\lfloor\tau\rfloor$ as a constant.

$$\frac{\partial p_{(n,c)}}{\partial p} =$$

$$\frac{\partial}{\partial p}\sum_{h=0}^{n-1}\left(\sum_{k=\lfloor\tau\rfloor+1}^{h+1}\binom{h+1}{k}p^k(1-p)^{h+1-k}\cdot\binom{n-1}{h}c^h(1-c)^{(n-1)-h}\right)\cdot r$$

$$+\frac{\partial}{\partial p}\sum_{h=0}^{n-1}\left(\sum_{k=0}^{\lfloor\tau\rfloor}\binom{h+1}{k}p^{h+1-k}(1-p)^k\cdot\binom{n-1}{h}c^h(1-c)^{(n-1)-h}\right)\cdot(1-r)$$

$$=\sum_{h=0}^{n-1}\left(\sum_{k=\lfloor\tau\rfloor+1}^{h+1}\frac{\partial}{\partial p}\left(\left(\binom{h+1}{k}p^k(1-p)^{h+1-k}\cdot\binom{n-1}{h}c^h(1-c)^{(n-1)-h}\right)\cdot r\right)\right.$$

$$\left.+\sum_{h=0}^{n-1}\left(\sum_{k=0}^{\lfloor\tau\rfloor}\frac{\partial}{\partial p}\left(\left(\binom{h+1}{k}p^{h+1-k}(1-p)^k\cdot\binom{n-1}{h}c^h(1-c)^{(n-1)-h}\right)\cdot(1-r)\right)\right)\right.$$

$$=\sum_{h=0}^{n-1}\left(\sum_{k=\lfloor\tau\rfloor+1}^{h+1}\binom{h+1}{k}p^k(h+1-k)(1-p)^{h+1-k-1}+kp^{k-1}(1-p)^{h+1-k}\cdot\right.$$

$$\binom{n-1}{h}c^h(1-c)^{(n-1)-h}\cdot r$$

$$+\sum_{h=0}^{n-1}\left(\sum_{k=0}^{\lfloor\tau\rfloor}\binom{h+1}{k}p^{h+1-k}k(1-p)^{k-1}+(h+1-k)p^{h+1-k}(1-p)^k\cdot\right.$$

$$\binom{n-1}{h}c^h(1-c)^{(n-1)-h}\cdot(1-r)$$

In the first line we use the sum rule, in the second line the summation rule and in the third we use the product rule and the power rule

Note that h, k, p, n, τ, r, c can only take positive values. Furthermore $0 < p < 1$ and $0 < c < 1$ and thus $(1-p)$ and $(1-c)$ are positive too. Hence the partial derivative with respect to p - $\frac{\partial p_{(n,c)}}{\partial p}$ - is a positive formula. It thus follows that $p_{(n,c)}$ increases as p increases and $\lfloor \tau \rfloor$ stays constant.

Case (2) We will determine what happens to $p_{(n,c)}$ in case $\lfloor \tau \rfloor$ becomes $\lfloor \tau \rfloor - 1$. The difference between $\lfloor \tau \rfloor$ and $\lfloor \tau \rfloor - 1$ is given by

$$\sum_{h=0}^{n-1} \left(\binom{h+1}{\lfloor \tau \rfloor} p^k (1-p)^{h+1-\lfloor \tau \rfloor} \cdot \binom{n-1}{h} c^h (1-c)^{(n-1)-h} \right) \cdot r$$
$$- \sum_{h=0}^{n-1} \left(\binom{h+1}{\lfloor \tau \rfloor} p^{h+1-\lfloor \tau \rfloor} (1-p)^{\lfloor \tau \rfloor} \cdot \binom{n-1}{h} c^h (1-c)^{(n-1)-h} \right) \cdot (1-r)$$

Looking at equation 3.2.3, it turns out that we cannot infer anything about the in- or decrease of the expected accuracy. The reason is that if $\lfloor \tau \rfloor$ becomes $\lfloor \tau \rfloor - 1$, the probability of getting a correct gets one more time evaluated, whereas probability of getting b corrects gets one less time evaluated. Now recall that we stipulated that an agent believes a is the state of the world if $k > \tau$, and we translated this to $k = \lfloor \tau \rfloor + 1$. Meanwhile we stipulated that an agent believes b if $k \leq \tau$, which translated to $k = \lfloor \tau \rfloor$. Now in the case that τ becomes $\tau - 1$ this means that the threshold for believing a is lowered. That is, with a lower number of a-type signals, an agent will believe a. This implies that up to a lower number of a-type, an agent will believe b. For example, suppose that previously $\lfloor \tau \rfloor = 50$, then with 50 a-type signals the agent believed b, whereas with the situation that $\lfloor \tau \rfloor = \lfloor \tau \rfloor - 1 = 49$ the agent will believe a with 50 signals.

Hence, it turns out that there is no general thing to say here: if the value of τ decreases, it can be that $p_{(n,c)}$ both increases as well as decreases, depending on the prior probability of a. This is also what explain the drops as seen in figure 3.3. However, we still see that the general tendency is that the accuracy increases as p increases. We infer that it must be that the cases in which the value of $\lfloor \tau \rfloor$ changes are rare, so therefore it *is* generally the case that $p_{(n,c)}$ increases as p increases. Proving this exactly is beyond the scope of this thesis.

3.2.4 Influence on $p_{(n,c)}$ of the prior r and the threshold t

In this section, we finally give an impression of the influence of the prior and the threshold. Plots of both are given in figure 3.4 and figure 3.5. Regarding the value of the prior, we see that the expected accuracy is lowest when $r = 0.5$. Regarding the value of the threshold, we see that it the expected accuracy is highest when $t = 0.5$. We do not give exact derivations for these results, as for the remainder of the thesis we will assume value 0.5 for r and t.

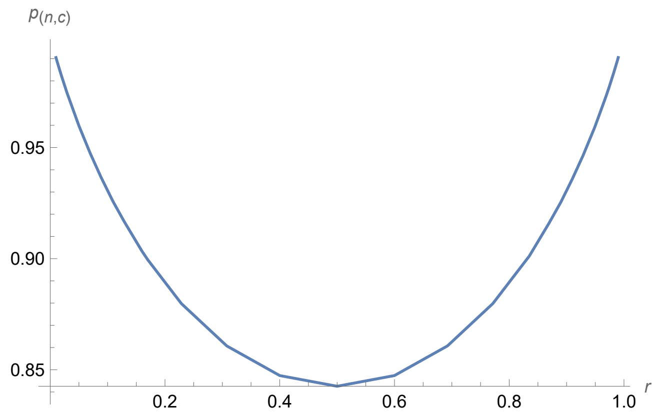


Figure 3.4: Influence of r on probability of correctness. Remaining parameter values are fixed to $n = 50, c = 0.5, p = 0.6, t = 0.5$

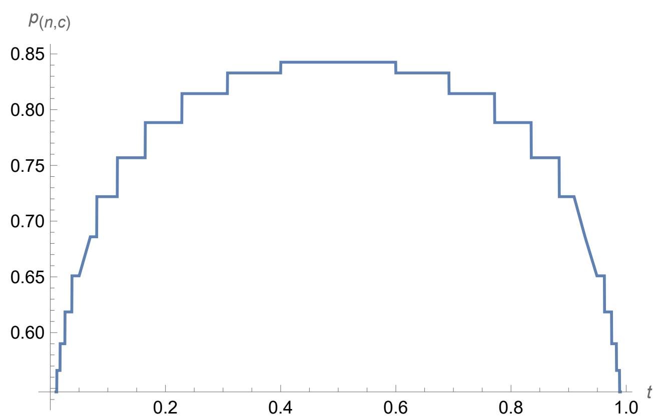


Figure 3.5: Influence of t on probability of correctness. Remaining parameter values are fixed to $n = 50, c = 0.5, p = 0.6, r = 0.5$

3.3 Maximum Likelihood Estimation

In the chapter 2, we saw in Corollary 2.5.1.1 that in case the threshold and prior are equal, an agent has to receive *more than half* a-type signals in order to actually believe a. In this section, we will show that Theorem 3.2.1, which represents the accuracy of an individual agent in a random network, in that case reduces to the maximum likelihood estimation approach.

Maximum likelihood estimation (henceforward denoted by *MLE*) is a method in statistics of estimating the parameters of a probability distribution, given observed data (Brandt et al., 2016, Chapter 8.3). In MLE, those parameters are chosen that maximize the likelihood that the assumed model results in the observed data. In our current setting, the observed data are the signals agents receive. The agents know that each signal is correct with a certain probability; this is the probability distribution. It is the goal of the agents to determine which parameter - alternative a or b - maximizes the likelihood of the observed set of signals.

In our more general Bayesian setting when τ is not necessarily equal to $\frac{h}{2}$, of course the agents are also trying to choose that alternative that maximises the likelihood of their observed signals. Yet, there is a difference between the Bayesian approach and the MLE approach. In the Bayesian approach, agents calculate a posterior probability distribution: the posterior probability of alternative a given the observed signals. So, we have a probability that the agents attaches to a and b. Meanwhile, the MLE approach returns a single fixed value, namely that alternative that maximises the likelihood of the observed signals given the state of the world. Unlike in the Bayesian approach, there is no probability attached to the value, it just returns whatever alternative is most likely. In this sense, the Bayesian approach is thus more *fine-grained*.

The fine-grained information comes with a cost: the Bayesian method also required more input parameters. In particular, it works with a prior probability and a threshold t. The MLE approach always implicitly assumes equal priors, and therefore in fact does not mention priors and posteriors. This can be seen both as a benefit as well as a loss. On the one hand, the MLE approach is simpler, on the other hand the Bayesian approach is more general and provides more information. Either way, we thought it interesting to show that in the case that the prior and the threshold value are equal, the Bayesian method boils down to MLE. In particular, in Theorem 3.3.1 we show what the calculation of the individual accuracy boils down to in case $r = t$.

Theorem 3.3.1. *Given an UCRN $\langle N, c, p, r, t \rangle$, if $t = r$ then the expected accuracy $p_{(n,c)}$ of an agent can be simplified to:*

$$p_{(n,c)} = \sum_{h=0}^{n-1} \left(\binom{n-1}{h} c^h (1-c)^{(n-1)-h} \sum_{k=\lfloor \frac{h+1}{2} \rfloor}^{h+1} \binom{h+1}{k} p^k (1-p)^{h+1-k} \right)$$

Proof. Recall that in case $t = r$ then $k > \frac{h}{2}$ as shown in Corollary 2.5.1.1. So $p_{(n,c)}$ then reduces to the following.

$$\begin{aligned} p_{(n,c)} = &\sum_{h=0}^{n-1} \left(\sum_{k=\lfloor \frac{h}{2} \rfloor + 1}^{h+1} \binom{h+1}{k} p^k (1-p)^{h+1-k} \cdot \binom{n-1}{h} c^h (1-c)^{(n-1)-h} \right) \cdot r \\ &+ \sum_{h=0}^{n-1} \left(\sum_{k=0}^{\lfloor \frac{h}{2} \rfloor} \binom{h+1}{k} p^{h+1-k} (1-p)^k \cdot \binom{n-1}{h} c^h (1-c)^{(n-1)-h} \right) \cdot (1-r) \end{aligned}$$

It then follows that:

$$\sum_{k=\lfloor \frac{h}{2} \rfloor+1}^{h+1} \binom{h+1}{k} p^k (1-p)^{h+1-k} = \sum_{k=0}^{\lfloor \frac{h}{2} \rfloor} \binom{h+1}{k} p^{h+1-k} (1-p)^k$$

And thus we can simplify the equation to 3.3.1. □

Like in Theorem 3.2.1, Theorem 3.3.1 also has a bias towards b. Suppose an agent receives an even number of signals, say 20. Then $\lfloor \frac{21}{2} \rfloor = 11$. As $\pi > t$ if and only if $k > \tau = \lfloor \frac{h+1}{2} \rfloor$ it follows that in case of an even number of signals, an agent believes a if strictly more than half of the signals is for a. In case of 20 signals, the agent will believe a if and only if it receives 11 a-type signals. If the agent receives less tan 11 a-type signals, it believes b. Now suppose an agent receives an odd number of signals, say 19. Then $\lfloor \frac{20}{2} \rfloor = 10$. This means again that the agent will believe a if and only if it receives a strict majority of a-type signals, namely 10. If it receives less than 10 a-type signals, it believes b.

In the current setting, agents get one noisy private signal and possibly some noisy signals from agents with which they communicate. The signals are noisy in the sense that they aim to say something about the true state of the world, but they don't do that perfectly. The goal of agents is to decide the true state of the world θ and they know there are two options. Either $\theta = a$ or $\theta = b$. We know that each agent i receives a private signal that tells with probability p_i correctly which alternative is correct, where $p_i = p_j$ for all $i, j \in N$. The signals agents receives from other agents are correct with this same probability p. We can think of each agent's signal s_i as a random variable generated by the probability distribution p. We assume that these signals are independent. Thus, during communication, an agent receives a sequence of identically distributed independent random variables that can take two values: a or b. The most likely state of the world then is the one that has the highest likelihood of generating the observed sequence of signals. Now it turns out that the way to determine this most likely state of the world is by using the majority rule: an agent should attach its belief to that alternative of which it receives a majority of signals. This is proved in the CJT by de Condorcet (1785). And this is indeed what Theorem 3.3.1 also tells us. Therefore, Theorem 3.3.1 shows that in case $r = t$, agents are using MLE.

This should not come as a surprise, as we basically already proved the CJT in Chapter 2. After all, the CJT states that a group under majority rule performs equally well or better than any individual that composes the group. This is because the majority rule is a voting rule that, given a voting profile, selects the alternative that corresponds to the most likely correct alternative. The only difference is that in the CJT setting, we are talking about a voting. In the current setting, we are not (yet) interested in voting, but about forming beliefs on an individual level. It turns out, however, that the exact same process takes place, only on a different level: that of the individual rather than the group.

As a result, we can also transfer the three postulates of the CJT to the individual level. Firstly, individuals become more accurate as they receive more signals. Secondly, communicating agents are more accurate than agents that do not communicate. Thirdly, agents that communicate with infinitely many other agents achieve perfect accuracy.

3.4 Summary

In this chapter we introduced random networks, in which agents are connected with each other according to a probability c. We formulated Theorem 3.2.1 which provides a method to determine the expected accuracy of an agent in any random network. We tested the influence of the size of the network, the connectedness and the competence on the expected accuracy and obtained intuitive results. Furthermore, we showed that in case the threshold value is equal to the prior, our Bayesian approach reduces to the maximum likelihood estimation approach.

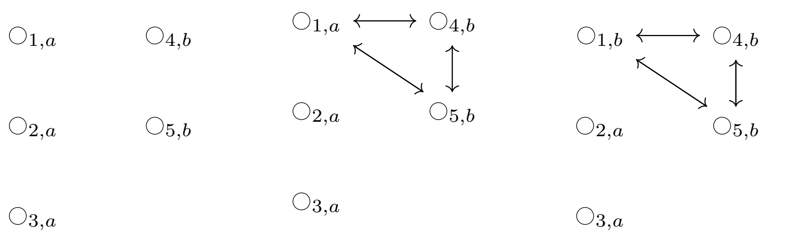
It is at this point interesting to look back on some of the related work we discussed in the introduction. There we discussed the work of Hahn et al. (2019) and Ladha (1992), who showed that with too much communication the competence of a group decreases, because of too much correlation. The reader might wonder how it can be that we showed individual accuracy increases as a result of communication. This is because our result applies to individual accuracy, where we assume that the signals agents receive are independent. The result of Hahn et al. (2019) and Ladha (1992) applies to group accuracy, and agents are not independent anymore after communication. In the next chapter, we will explain how we will deal with correlation.

Chapter 4

Group Accuracy in Random Networks

In the previous chapter we have established a method to measure the individual accuracy of agents that communicate with each other according to connections in a random network. We now shift our attention from individual accuracy after communication to group accuracy after communication in random networks. In this way, we aim to provide an intermediate result between Condorcet's model, for which we have group accuracy after no communication, and Coughlan's model, for which we have group accuracy after full communication. Our goal will be to determine the group accuracy via majority rule in random networks. Note that this is non-trivial, because after communication agents' beliefs and therefore also votes are not independent anymore.

There are several ways in which communication can affect group accuracy in random networks. We will focus on how the amount of communication connections an agent has in a network influences group accuracy. We will refer to this as the *spread of communication*. So, with spread of communication we mean the difference in how much agents communicate in the network. Spread of communication can influence the group accuracy. In particular, as we will show in Section 4.1, if all agents communicate equally as much, then if a majority of agents has a correct initial signal, it follows intuitively that the correct signal is more often communicated. Moreover, we will show that if each signal has the same weight, then asymptotically the group reaches a correct majority with probability 1. Therefore, in Section 4.1.2 we will set out to determine under which circumstances big differences in the amount of communication between agents do not occur. Our results are twofold. Firstly, we will show that asymptotically, that is, as the number of agents approaches infinity, agents will all communicate equally as much. Secondly, we will use Chernoff Bounds (Chernoff (1952)) to show that even with lower conditions on the number of agents, under certain conditions there is a high probability that the agents in a random network communicate equally as much.



(a) Initial signals (b) Before communication (c) After communication

Figure 4.1: Process of belief updates and communication in a network in which $\theta = a$

4.1 Evenly spread communication

In this section we will show that if all agents communicate equally as much, a correct majority of signals implies that the correct signals is communicated a majority of times. To see why it is important that every agent communicates equally as much, consider the following example.

In Figure 4.1 we are given three graphs that represent the process of communication in a network. Figure 4.1a represent the initial signals, Figure 4.1b is the network before communication and Figure 4.1c is the network after communication. Each graph represents a network with five nodes representing agents, numbered from $\{1, ..., 5\}$. Next to each node, the alternative - a or b - is written, which represents the private signal of the agent. Each arrow represents a communication relation between the corresponding nodes. So, in Figure 4.1b agent 1 is communicating with agent 4 and vice versa. If there is no arrow, then agents are not communicating.

Suppose a is the true state of the world and $r = t$ such that agents will update their belief according to majority rule. We can see in Figure 4.1a that a majority of the agents has a correct signal, namely agent 1,2 and 3. We assume agents are in a random network and thus relations between agents are established according to a probability c. Suppose Figure 4.1b is the resultant network. We can see in this network that communication is not evenly spread. All agents with a b signal communicate, whereas only one out of the three agents that has an a signal communicates. In Figure 4.1c we show the network after communication. Since all agent update their belief according to majority rule, agent 1,4 and 5 will update their belief to b, the remaining agents 2 and 3 will update their belief to a 4.1a. If the agents in Figure 4.1c would now vote, b would be in the majority, even though a majority of agents had a correct signal.

In this example, a belief that is relatively rare in a network is over-represented in the signals agents receive from other agents during communication. After all, b is the signal that initially is received by a minority yet during communication a majority of agents receives a majority of b-type signals. This phenomenon is in the literature called the *majority illusion effect* (see Lerman et al. (2016)).

As this example illustrates, the majority illusion effect can effect the accuracy of the majority of the group.

Moreover, intuitively we can see in this example that if all agents would communicate equally as much, then the signal that would have been most often communicated would have been the correct signal a. In this section we will make this intuition precise.

4.1.1 Weights and degrees as measures of communication

Recall that we have a random network $G(n, c)$. Each agent j sends a signal s_j into the network. The signal can be either a or b. We would like to express how much a certain type of signal is communicated in the network. For this we introduce *weights*, which are expressed in terms of degrees. Recall that the degree of a vertex j, denoted by $\delta(j)$, is the number of edges that are incident to the vertex. Now the weight of a signal j is defined as follows.

Definition 4.1.1 (Weight of a signal)**.** The weight w_j of an agent's j signal s_j is given by:

$$w_j = \frac{\delta(j)}{\sum_{i \in N}(\delta(i))}$$

Basically, the weight divides the number of times the signal of an agent j is seen by the number of times all signals in the network are observed. As such it expresses how often a signal relative to a network is observed.

When we defined the weight of a signal, we refrained from counting the fact that each agent receives a signal from herself. We can do this, as weight is a relative measure. Moreover, in random networks loops do not occur, so when we talk about the degree of a vertex, we will never be counting a loop. In contrast, when talking about expected accuracy, we do have to count the signal an agent receives from itself, because this measure is not relative.

We introduce two further notions. Firstly, we define the notion of the *degree of a signal type* that counts how often a certain type of signal is observed in the network. Secondly, we introduce the related notion of the *weight of a signal type* that expresses the total weight of a certain type of signal in the network. This will enable us to express how much weight signal type has in the network.

Definition 4.1.2 (Degree of a signal type)**.** Let $x \in \{a, b\}$. The number of times the signal type x is observed, denoted by $\delta(x)$ is defined as:

$$\delta(x) = \sum_{\{j \in N | s_j = x\}} \delta(j)$$

Definition 4.1.3 (Weight of a signal type)**.** Let $x \in \{a, b\}$. The total weight of all x-type signals, $w(x)$, is given by:

$$w(x) = \frac{\delta(x)}{\sum_{j \in N}(\delta(j))}$$

We will make a distinction between a weighted majority of signals, a majority of signals and a majority of votes. The distinction corresponds to three steps in the truth-tracking process. Firstly, agents receive signals. We say that we have a *majority of signals* of a signal $x \in \{a, b\}$ if $\sum_{\{j \in N | s_j = x\}}(s_j) > \sum_{\{j \in N | s_j = b\}}(s_j)$, where $y \in \{a, b\}$ and $x \neq y$. Secondly, agents communicate and receive signals from each other. We say that a signal $x \in \{a, b\}$ has a *weighted majority* over the other signal $y \in \{a, b\}$ in case $w(x) > w(y)$. In words, if a signal has a weighted majority this means that it is communicated a majority of times. Thirdly, agents update their beliefs according to the received signals and vote according to their belief. Recall that $majority(\boldsymbol{v}) = \{x \in A | v(x) \geq \lfloor \frac{n}{2} \rfloor\}$ where $v(x) = |\{v_i \in \boldsymbol{v} | v_i = x\}|$ and $\hat{k}$ is an integer between 0 and n. Recall furthermore that in case of a tie, an alternative is chosen by a random coin toss. In this case, we speak of a majority of votes for $x \in \{a, b\}$ in case $majority(\boldsymbol{v}) = \{x \in A | v(x) \geq \lfloor \frac{n}{2} \rfloor\}$. We have a majority of votes for a if for a majority of agents it is the case that $\pi > t$, while we have a majority of votes for b if for a majority of agents it is the case that $\pi \leq t$.

4.1.2 Correct majority and evenly spread communication

Using this new machinery, we will show in this section that if each signal has the same weight, then if we have a majority of x-type signals then we have a weighted majority for x. We will first show that if signals do not have the same weight, then the implication doesn't hold. That is, if we have a majority of signals for x, it does not necessarily imply that we have a weighted majority for x if communication is not evenly spread.

Observation 4.1.1. *Let $x, y \in A$ such that $x \neq y$. Given a random network $G(n, c)$, if $\sum_{\{j \in N | s_j = x\}}(s_j) > \sum_{\{j \in N | s_j = y\}}(s_j)$ then it is not necessarily the case that $w(x) > w(y)$.*

Proof. Consider the graph in Figure 4.2. The model should be read as explained in Section 3.1, yet here we only show the model during communication. In Figure 4.2, we have that $\sum_{\{j \in N | s_j = b\}}(s_j) > \sum_{\{j \in N | s_j = a\}}(s_j)$ yet $w(a) > w(b)$. □

If we impose that each signal has the same weight, however, then the implication does hold. Note that if each signal has the same weight, this means that each vertex has the same degree and thus we are dealing with a regular network.

Lemma 4.1.1. *Let $x, y \in A$ such that $x \neq y$. Given a regular network $G(n, \delta)$, if $\sum_{\{j \in N | s_j = x\}}(s_j) > \sum_{\{j \in N | s_j = y\}}(s_j)$ then $w(x) > w(y)$.*

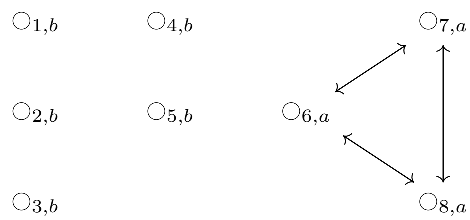


Figure 4.2: Model with weighted majority for a and majority of signals for b

Proof. Without loss of generality, given a regular network $G(n, \delta)$, let $\sum_{\{j \in N | s_j = a\}}(s_j) > \sum_{\{j \in N | s_j = b\}}(s_j)$. Assume by contradiction that $w(b) > w(a)$. This would imply that $\delta(b) > \delta(a)$, which means that $\sum_{\{j \in N | s_j = b\}}(\delta(j)) > \sum_{\{j \in N | s_j = a\}}(\delta(j))$. Since the network is regular, we know that $\delta(i) = \delta(j)$ for all $i, j \in N$. Therefore, we would have $\sum_{\{j \in N | s_j = b\}}(s_j) > \sum_{\{j \in N | s_j = a\}}(s_j)$, which is a contradiction. So we conclude $w(a) > w(b)$. □

Hence, Observation 4.1.1 implies that if all signals have the same weight, then if we have a majority of signals for x then we have a weighted majority for x. That is, if everyone communicates to equally many agents, then if the correct signal is more prevalent, then the correct signal is also communicated a majority of the time.

A correct weighted majority, does not necessarily imply a correct majority of votes. For this we would also need to know to which agents the correct signals are communicated. If the correct signals are communicated foremost to the same agent, we still cannot guarantee that a majority of agents receives a majority of correct signals. Figure 4.2 is a good illustration of this. In the figure, we have a weighted majority for a. Yet, we can see that it is only a minority of agents that receives a majority of a-type signals.

However, having a correct weighted majority does in a sense help reaching a correct majority of votes. More precisely, it rules out situations such as illustrated in Figure 4.1. In this figure, we had a correct majority of signals, but an incorrect weighted majority which led to an incorrect majority of votes. As a result, a minority of agents ended up believing a. By focusing on situations in which each signal has the same weight, we can reduce the chance of such a phenomenon happening. Moreover, as we will prove in the following theorem, we can guarantee a correct majority of votes asymptotically. That is, we show that if $t = r$ and all signals have the same weight, then a correct majority of signals implies a correct majority of votes.

Theorem 4.1.2. *If $t = r$ and $w_j = \frac{1}{n}$ for all s_j such that $j \in N$, and $p > 0.5$ then $\lim_{n \to \infty} p_{maj} \to 1$.*

Proof. Since $w_j = \frac{1}{n}$ for all s_j such that $j \in N$ this implies $\delta(j) = \delta(i) = k$ for all $i, j \in N$ and some integer k. This means that each agent receives k independent signals. We know each signal has the same accuracy $p > 0.5$. By equation 2.3 of the CJT (Theorem 2.2.1), the probability that the majority of these k signals is correct is strictly bigger than 0.5. By lemma 4.1.1 this means that in expectation we have a correct weighted majority. But we have even something stronger, namely, strictly more than half of the agents receive a majority of correct signals. Since $t = r$, this means that in expectation we have a correct majority of beliefs and thus votes. By the Law of Large Numbers $\lim_{n\to\infty} p_{maj}(n) = 1$. □

Theorem 4.1.2 should not come as a surprise. In Chapter 2 we stated that in regular networks, if $p > 0.5$ we will obtain a correct majority. Now if everyone communicates equally as much, we basically have a regular network. So, the reader might wonder why we added this extra step to measure communication. This is because we want to know the group accuracy in random networks, and in random networks communication is not necessarily evenly spread. In remainder of this chapter we will therefore discuss methods to measure how far a random network is from a regular network. The farther away a random network is from a regular network, the higher the chance that incorrect minorities turn into incorrect majorities and as a result the group accuracy decreases.

It is important to realize that the Theorem 4.1.2 depends on the assumption that $t = r$. We need this assumption, because in this case we know that an agent ends up believing the signal of which it receives a majority. As we discussed in Chapter 2 section 2.5 however, the value of p can also bring τ close to $\frac{h}{2}$. In particular, we observed that the higher the value of p, the closer the value of τ to $\frac{h}{2}$. Moreover, the higher the value of p, the less influence the prior r and the threshold t have in determining the value of τ. Thus, even if $t \neq r$, but we have a rather high competence p, it is likely that an agents ends up believing the signal of which it receives a majority.

4.2 Determining the spread of communication

So far we have results for the accuracy of the majority for the two extremes: when $c = 0$ and $c = 1$. Our aim in this chapter is to determine this accuracy for random networks where $0 < c < 1$. To achieve this, we study the spread of communication, since in the previous section we saw that if communication is evenly spread and $r = t$, then a correct majority of signals implies a correct majority of votes.

However, there is another important reason why we are interested in the spread of communication. After all, the aim of this chapter is to show that in random networks, despite correlation between agents, we can under some constraints still obtain accurate majorities. Recall that in Condorcet's model (1785) there is not communication and thus also no correlation. In this case majorities are better than non-communicating individuals. In Coughlan's model (2000), majorities are only just as good as fully-communication and therefore fully correlated individuals. In Chapter 3 we already saw that partially communicating individuals are better than non-communicating individuals, but we want to know that happens to the accuracy of the majority since beliefs are now partially correlated. Now the point is that when signals (tend to) have the same weight, this means that all pieces of information are most likely seen a same number of times. Therefore, in this case, correlation is also evenly spread. As a result, correlation does not disturb the process because it happens 'homogeneously' throughout the group. Hence, studying the spread of weights therefore also enables us to indirectly say something about the correlation. In particular, when signals (tend to) have the same weight, we can ignore correlation, which significantly simplifies obtaining results about the group accuracy.

In this section, we will discuss three cases in which communication is evenly spread: the extreme case, the asymptotic case and the approximate case.

4.2.1 Extreme case: Condorcet and Coughlan

There are two cases in which we are sure that all signals have the same weight. This is when $c = 0$ and when $c = 1$.

Observation 4.2.1. *If $c = 0$ or $c = 1$ then $w_j = \frac{1}{n}$ for all s_j where $j \in N$.*

Proof. If $c = 0$ then for all $v_j \in V$ it is the case that $\delta(v_j) = 0$. Therefore $w_j = \frac{0}{\sum_{j \in N}(0)} = \frac{1}{n}$ for all s_j such that $j \in N$.

Similarly, if $c = 1$ then for all $v_j \in V$ it holds that $\delta(v_j) = N$. Therefore $w_j = \frac{N}{\sum_{j \in N}(N)} = \frac{1}{n}$ for all s_j where $j \in N$. □

Intuitively, observation 4.2.1 says that if agents are connected to nobody or to everyone, then all signals have the same weight. This makes sense as in both cases all the signals are observed by the same number of people: by no one or by all agents respectively. It also makes sense, as we already knew that Condorcet and Coughlan are both dealing with regular networks. Obviously, in regular networks, all signals have the same weight since weight is determined via degrees and all vertices in a regular network have the same degree.

As we already stated in Theorem 2.5.6 for both Coughlan's as well as Condorcet's model, the probability that the majority of agents is correct will be 1 as the size of the networks approaches infinity. This now also follows from Theorem 4.1.2.

4.2.2 Asymptotic result: infinitely large networks

The following lemma shows that if the size of the network approaches infinity, then all vertices have the same weight.

Lemma 4.2.1. *As $n \to \infty$ and $0 \leq c \leq 1$ then $w_j = \frac{1}{n}$ for all s_j such that $j \in N$.*

Proof. As $n \to \infty$ it follows by the weak law of large numbers that average number of connections an agent has converges to the mean, which is $c(n-1)$. It then follows that every signal has the same weight, namely $\frac{c(n-1)}{\sum_1^n (c(n-1))} = \frac{1}{n}$. □

Now, intuitively, if every signal has the same weight, and every signal is correct with probability bigger than 0.5, then as n approaches infinity, the probability that a majority of agents will form a correct belief should approach 1. This is indeed what we showed in Theorem 4.1.2. So, combining Theorem 4.1.2 and lemma 4.2.1

Theorem 4.2.2. *Given an UCRN* $\langle N, c, p, r, t\rangle$ *with* $r = t$ *and* $p > 0.5$ *it follows that* $\lim_{n\to\infty} p_{maj}(n) = 1$.

Proof. From lemma 4.2.1 we know that as $n \to \infty$ then $w_j = \frac{1}{n}$ for all s_j such that $j \in N$. Since also $t = r$ and $p > 0.5$ it then follows from Theorem 4.1.2 that $\lim_{n\to\infty} p_{maj}(n) = 1$. □

Theorem 4.2.2 shows that asymptotically, a correct majority is guaranteed on any random network, independently of the specific network $G(n, c)$. So, after generalising the asymptotic result from $c = 0$ to the whole class of regular networks in Theorem 2.5.6, we now generalised the asymptotic result to the class of random networks.

4.2.3 Approximate result: Chernoff Bounds

The asymptotic result is nice as it shows that as the number of agents approaches infinity, the probability of a correct majority is certain. Yet, in real life we are never dealing with infinite number of agents. So, this makes us wonder, what can we say about the case when there is a finite number of agents? In this section we will spell out the relation between the size of the network n and the connectedness c of the network and the probability to reach a correct majority.

If $0 < c < 1$ and n is finite, then obviously the possibility arises that signals do not have the same weight. In fact, this is very likely. However, it is not the case that the weights of all the signals are all over the place. This is because we generate networks from a random process where agents are always connected with each other with probability c. In particular, the degree of a vertex follows a binomial distribution. This means there is a known mean degree and that the degrees will be spread around the mean degree, with the probability of being far away from the mean decreasing as the distance from the mean increases.

Before we make this precise, let's look back at our example in Figure 4.2. Although the network shows that a weighed majority does not necessarily imply a majority of votes, the network has a low probability of being generated from a random network with parameter c. After all, in the graph most nodes have a degree that is far from the average degree of $7 \cdot c$, no matter what value we attach to c. Assuming that the competence of each agent is higher than 0.5, it is most likely that b is the right signal. This means that a would be an incorrect weighted majority. If the correct b-nodes had some edges, then it would have been likely that we obtained a correct weighted majority, since there are many more b nodes. Intuitively, this graph is very unlikely to be generated from a random process, and the goal of this subsection is to make this unlikeness precise by making use of Chernoff Bounds (Chernoff (1952)) and studying properties of the variation of the degrees.

We start off with the latter: studying the properties of the variation of degrees. As we are working on random graphs, edges between vertices occur with probability c. We are thus given a distribution of degrees. The distribution of degrees is given by the binomial distribution. Given a random network $G(n, c)$ the vertex degree is a random variable X_δ that follows the binomial distribution:

$$P(X_\delta = k) = \binom{n-1}{h} c^h (1-c)^{n-1-h} \tag{4.1}$$

The Binomial Distribution can be considered as the distribution of the number of successes in a series of h trials with two possible outcomes in each trial: 1 (success) and 0 (failure). In our case, success can be seen as making a connection and exchanging a signal, and a failure as not exchanging a signal.

Following known properties of the binomial distribution, the mean vertex degree $\bar{\delta}$ is computed as as follows:

$$\bar{\delta} = c(n-1) \tag{4.2}$$

Furthermore, the variance $var(X_\delta)$ is given by:

$$var(X_\delta) = (n-1)c(1-c) \tag{4.3}$$

From the variance, we calculate the standard deviation σ as follows:

$$\sigma = \sqrt{Var(X_\delta)} = \sqrt{(n-1)c(1-c)} \tag{4.4}$$

Both the variance and the standard deviation are measures of the amount of variation of a set of degrees. The variance is the expectation of the squared difference of the random variable X_δ from the mean degree. The variance measures how far the set of degrees is spread out from the average mean degree. The standard deviation is the square root of the variance. A low standard deviation indicates that X_δ tends to be close to the mean degree, while a high standard deviation indicates that X_δ is farther away from the mean. We will use the variance and standard deviation to determine in which case the expected variation in degrees is maximal, as stated in the following observation.

Observation 4.2.2. *For the vertex degree X_δ that follows a binomial distribution, we make the following observations.*

(1) The variance and standard deviation are maximal when $c = 0.5$.

(2) The variance and standard deviation are minimal when $c = 0$ and $c = 1$.

(3) The variance and standard deviation decrease as they are farther away from $c = 0.5$.

Intuitively, Claim 1 in observation 4.2.2 tells us that we have the highest chance of differences in weight if $c = 0.5$. So, distortions in group accuracy due to weight differences are highest when $c = 0.5$. Claim 2 in observation 4.2.2 implies that distortions in group accuracy due to weight difference are impossible when $c = 0$ or $c = 1$. Finally, Claim 3 in observation 4.2.2 implies that that the farther c is away from 0.5, then with higher probability we get small differences in weights between signals, thus bringing us closer to group accuracy in the Condorcet and Coughlan setting.

The variation and standard deviation already give an indication of how the parameter c influences the spread of communication in a network and thus its influence on the group accuracy. However, we would like to say with more precision what the chance is for certain weight differences to occur. In particular, we would like to determine the probability of observing a degree within a given distance from the mean degree. For this, we introduce the use of the Multiplicate Chernoff Bound (Chernoff (1952)).

Definition 4.2.1 (Multiplicative Chernoff Bound)**.** Suppose $X_1, ..., X_n$ are independent random variables taking values in $\{0, 1\}$. Let X denote their sum and let $\mu = E[X]$ denote the sum's expected value. Then for any $d > 0$ we have the following.

$$P(X > (1 + d)\mu) < \left(\frac{e^d}{(1 + d)^{1+d}}\right)^{\mu} \tag{4.5}$$

$$P(X < (1 - d)\mu) < \left(\frac{e^{-d}}{(1 - d)^{1-d}}\right)^{\mu} \tag{4.6}$$

In words, the multiplicative Chernoff bound gives us an upper and lower bound on the probability that the random variable X is a certain distance d away from the mean.

We will use the Chernoff bound as follows. For every node we run $n - 1$ Bernoulli trials to determine the connections with other nodes. So we have $X_1, ..., X_{n-1}$ independent Bernoulli random variables. Each value either takes value 1 or 0, depending on whether there is a connection or not. Then $X = \sum_1^{n-1} X_1, ..., X_{n-1}$ denotes the total number of connections a node has. In other words, it denotes the degree. We already saw that the average degree is $\bar{\delta} = c(n - 1)$. Then applying the Multiplicative Chernoff bound, we obtain in the following theorem.

Theorem 4.2.3. *Let $G(n, c)$ be a random network. The probability that X_δ is $d > 0$ away from the mean degree is determined by the following two bounds.*

$$P(X_\delta > (1 + d)(c(n - 1))) < \left(\frac{e^d}{(1 + d)^{1+d}}\right)^{c(n-1)} \tag{4.7}$$

$$P(X_\delta < (1-d)(c(n-1))) < \left(\frac{e^{-d}}{(1-d)^{1-d}}\right)^{c(n-1)} \tag{4.8}$$

Proof. This follows from using the Multiplicative Chernoff bound and the fact that the mean degree $\bar{\delta}$ is equal to $c(n-1)$. □

In words, Theorem 4.2.3 bounds the probability that the degree of any node is d away from the mean degree. We introduce the following shorthand notation for the upper bound υ and the lower bound λ:

$$\upsilon = \left(\frac{e^{d}}{(1+d)^{1+d}}\right)^{c(n-1)}$$

$$\lambda = \left(\frac{e^{-d}}{(1-d)^{1-d}}\right)^{c(n-1)}$$

To get a feeling for Theorem 4.2.3, consider the following example. Suppose $n = 50$, $c = 0.3$, $d = 0.1$. Firstly observe that in this case $\bar{\delta} = c(n-1) = 15$ and $1.1\bar{\delta} = (1+d)\bar{\delta} = 16$ and $0.9\bar{\delta} = (1-d)\bar{\delta} = 13$. The Multiplicative Chernoff bound then tells us that the probability that the degree is bigger than 16 is at most 0.93 and the probability that the degree is smaller than 13 is also at most 0.93. This means that possibly with a high probability, there might be a difference of 3 degrees. This, on its turn, could change an incorrect minority into a weighted majority and as such lower the group accuracy.

In the following lemma, we spell out the influence of the size of the network n, the connectedness c and the distance d on the upper bound υ and the lower bound λ.

Observation 4.2.3. *υ and λ have the following properties:*

(1) As n increases, υ and λ decrease.

(2) As c increases, υ and λ decrease.

(3) As d increases, υ and λ decrease.

Proof. We prove the claims one by one using partial derivatives.

(1)

$$\frac{\partial \upsilon}{\partial n} = c \log\left((d+1)^{-d-1}e^{d}\right)\left((d+1)^{-d-1}e^{d}\right)^{c(n-1)}$$

$$\frac{\partial \lambda}{\partial n} = c \log\left((1-d)^{d-1}e^{-d}\right)\left((1-d)^{d-1}e^{-d}\right)^{c(n-1)}$$

The derivative with respect to n is negative and thus υ decreases as n increases. Similarly for λ.

(2)

$$\frac{\partial \upsilon}{\partial c} =(n-1)\log\left((d+1)^{-d-1}e^{d}\right)\left((d+1)^{-d-1}e^{d}\right)^{c(n-1)}$$

$$\frac{\partial \lambda}{\partial c} = \left((1-d)^{-1-d}e^{-d}\right)^{c(-1+n)p}(-1+n)p\log(e^{-d}(1-d)^{-1+d})$$

The derivative with respect to c is negative and thus υ decreases as c increases. Similarly for λ.

(3)

$$\frac{\partial \upsilon}{\partial d} =c(n-1)\left(e^{d}(d+1)^{-d-1}+e^{d}(d+1)^{-d-1}\left(\frac{-d-1}{d+1}-\log(d+1)\right)\right)\left((d+1)^{-d-1}e^{d}\right)^{c(n-1)-1}$$

$$\frac{\partial \lambda}{\partial d} =c(n-1)\left((1-d)^{d-1}e^{-d}\left(\log(1-d)-\frac{d-1}{1-d}\right)-(1-d)^{d-1}e^{-d}\right)\left((1-d)^{d-1}e^{-d}\right)^{c(n-1)-1}$$

The derivative with respect to d is negative and thus υ decreases as d increases. Similarly for λ.

□

From Observation 4.2.3 we can infer that in larger networks and more connected networks, the upper and lower bounds become tighter. The latter is in line with Observation 4.2.2, as we showed that the farther away c is from 0.5, the less variation in degrees there is. This intuitively makes sense. If any connection between two agents is basically established by a random coin toss, it makes sense that there occurs a lot of variation in the number of connections agents have. The more unfair our coin toss, the less variation in the number of connections between agents. The fact that in larger networks the upper and lower bound become tighter is also intuitively understandable. The larger the network, the more the degree can converge to the mean degree by following the law of large numbers This result is in line with Lemma 4.1.2, according to which in an infinitely large network, every signal has the same weight. Lemma 4.2.3 says that the bigger the network, the higher the probability that the weights are similar.

Observation 4.2.3 also shows that with bigger distances, that is a greater difference in degrees, the upper and lower bounds become tighter. This also makes sense intuitively. After all, the bigger the difference, the higher the probability that the degree of a vertex will fall within the range of the mean degree plus or minus the difference d.

What do these results imply for the group accuracy? From Theorem 4.1.2 we know that in a random network such that $t = r$ and $p > 0.5$, in the limit a correct majority is reached with probability 1. Theorem 4.2.3 gives a handle to measure the probability that a random network is a certain distance away from a regular network. The farther away a random network is from a regular network, the more chance on distortions in group accuracy due to unevenly spread communication. There is then a higher chance that incorrect minorities of signals turn into incorrect majorities of votes such as in Figure 4.1, with the result that the expected group accuracy decreases. Observation 4.2.3 in particular shows implies that a decreased expected group accuracy due to unevenly spread communication is is less likely in larger networks and in more connected networks.

In the following subsections, we will describe some applications of Theorem 4.2.3 for specific distances, to give an impression on the probabilities of a random network being close to a regular network.

Application Chernoff Bound: 1 degree distance

Suppose we allow the distance in number of signals to be at most 1. This means that we allow agents to receive one more or one less signal than the mean. This means that two agents can at most have a difference of two signals. This seems the tightest bound we can have, rather than just demanding every agent to have the exact same number of signals. The following lemma shows that for a difference of at most 1 signals away from the mean, d should be less than or equal to $\frac{1}{c(n-1)}$.

Lemma 4.2.4. *If* $(1+d)\bar{\delta} - \bar{\delta} \leq 1$ *then* $d \leq \frac{1}{c(n-1)}$.

Proof.

$$\begin{aligned}(1+d)((c(n-1)) - c(n-1) &= 1 \\ cd(n-1) &= 1 \\ d &= \frac{1}{c(n-1)}\end{aligned}$$

□

Observe that as we always allow for only one node of distance more or less than the mean, the allowed distance becomes smaller and smaller the bigger n and c get. This is because d is expressed in terms of percentage and not in terms of an absolute value. As a result, demanding a distance of at most 1 degree makes most sense for relatively small networks.

Example 1. Let $n = 10$ and $c = 0.5$. Then $d = 0.22$ and $P(X_\delta > 1+0.22)\bar{\delta} < 0.90$ and $P(X_\delta < 1-0.22)\bar{\delta} < 0.89$.

If we decrease c to 0.2 then the probability decreases too. Namely, $d = 0.56$ and $P(X_\delta > 1 + 0.56)\bar{\delta} < 0.79$ and $P(X_\delta < 1 - 0.56)\bar{\delta} < 0.70$.

The previous example indicates that if we allow for at most one degree deviation from the mean, the Chernoff Bounds are high. This means that the probability that the degree deviates more than one degree from the mean is high. As the network size n increases, the allowed distance $d \leq \frac{1}{c(n-1)n}$ will only become smaller. Obviously, a smaller allowed distance is harder to meet. In general, the requirement of a maximum of one degree deviation from the mean thus seems hard to meet.

Even with only a one degree deviation from the mean, an incorrect minority of signals could change into an incorrect majority of votes. However, given a specific network, Theorem 4.2.3 can express that this might be unlikely. For example, suppose we have a UCRN $\langle N, c, p, r, t \rangle$ with $n = 5$ and $c = 0.25$. In this UCRN the mean degree is 1. Applying the Chernoff Bounds in Theorem 4.2.3, the probability that a vertex has a degree of 2 is at most 0.68, and the probability that a vertex has a degree of 0 is at most 0.39. Based on this information, we can say that the network in Figure 4.3 has thus not a very high probability of occurring from our UCRN. After all, in the figure, agents 1 and 3 have a degree equal to the mean degree. Agent 2 has a degree that is one less than the mean degree, which has a probability of occurring of at most 0.39. Agent 4 and 5 have a degree that is one more than the mean, which, for each agent independently, has a probability of occurring of at most 0.68. In this network, a is the majority signal, while b has a weighted majority. If a would be the correct signal, this could potentially lead to an incorrect majority of votes.

What is more, we can see that there is also an asymmetry when it comes to the type of signals that are communicated: b is much more often communicated than a, even though each node has a degree close to the mean degree. This is a factor that we do not express wit Theorem 4.2.3, but intuitively we can say that an asymmetry in the distribution of types is less likely than a symmetry. All in all, the network in Figure 4.3 in total seems not to have a very high probably of occurring from our given UCRN, and Theorem 4.2.3 helps expressing this probability. So, even though it is possible to obtain an incorrect majority due to 1 degree deviations from the mean, given a specific UCRN, we can bound this probability using Theorem 4.2.3.

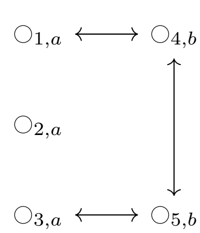
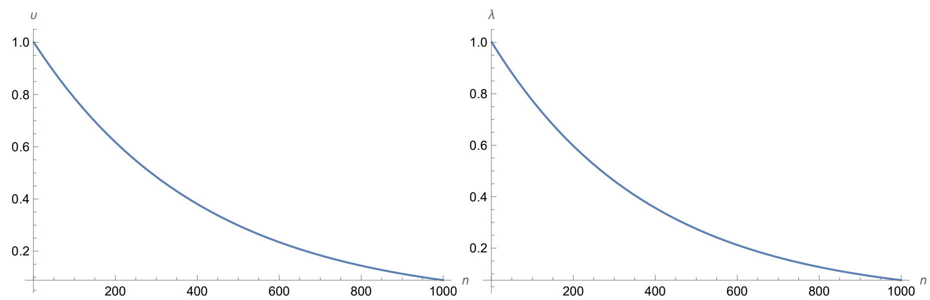


Figure 4.3: Network in which all degrees deviate at most 1 degree from the mean degree $\bar{\delta} = 2$



(a) Upper Chernoff bound υ (b) Lower Chernoff bound λ

Figure 4.4: Upper and Lower Chernoff bound while $d = 0.1$ and $c = 0.5$

Application Chernoff Bound: percentage

Instead of looking at a specific value that we allow the degrees to deviate from the mean, we can also look at a percentage that we allow degrees to deviate from the mean. The advantage of this method is that this requirement will differ per size of the network. If the network is small, the maximum number degrees are allowed to deviate from the mean is relatively small. Meanwhile, if the network is large, the maximum number of degrees a node is allowed to deviate from the mean is relatively large.

Example 2. Suppose $d = 0.1$. Then if $\bar{\delta} = 10$ we allow for $\delta = 9$ and $\delta = 11$. Meanwhile if $\bar{\delta} = 100$ then we allow for $\delta = 90$ and $\delta = 110$.

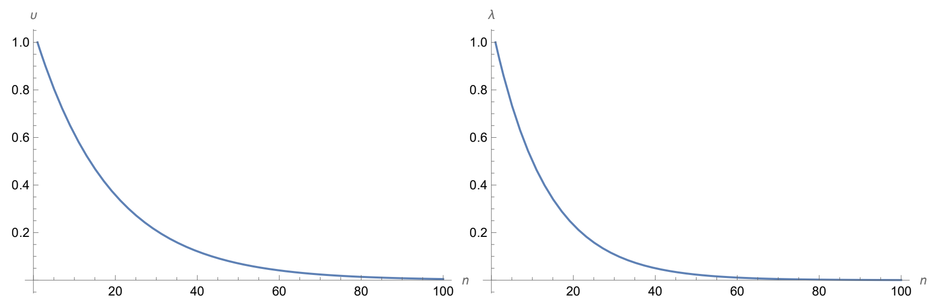
Observe furthermore that if we let $d \leq 0.1$ this implies that weight difference will also be less than or equal to 10%.

Observation 4.2.4. *If $d \leq 0.1$ then any $w(j)$ of any signal s_j of an agent j is 10% bigger or smaller than $\bar{\delta}$.*

Proof. Recall $w_j = \frac{\delta(j)}{\sum_{i \in N} \delta(i)}$. The denominator $\sum_{i \in N} \delta(i)$ stays the same in each calculation, yet if $d \leq 0.1$ then $\delta(j)$ can be at most $d\delta(j) \cdot 1.1$ and at least $\delta(j) \cdot 0.9$. In the first case $w_j = 1.1 \cdot \bar{\delta}$ while in the latter $w_j = 0.9 \cdot \bar{\delta}$. □

It turns out that a distance of 0.1 still has high bounds. The network has to be very large for there to be a low probability that degrees deviate a maximum of 10% from the mean. This is illustrated by Figure 4.4a and 4.4b below, showing the upper- and lowerbounds of the probability to exceed the maximum of 10% deviation. As we can see in both figures, a distance of 0.1 is bounded by a relatively low probability only for networks with about 500 nodes or more. In those cases we can say that there is only a 30% probability that nodes deviate more than 10% from the mean degree.

It makes sense though that for very small networks a distance of 0.1 is bounded by a high probability. Suppose we have a network of 10 nodes with $c = 0.3$ and $d = 0.1$. Then the mean degree $\bar{\delta} = 3$ and $d \cdot \bar{\delta} = 3.3$ while $d \cdot \bar{\delta} = 2.7$.



(a) Upper Chernoff bound υ (b) Lower Chernoff bound λ

Figure 4.5: Upper and Lower Chernoff bound while $d = 0.5$ and $c = 0.5$

So basically, in this case we are asking what the probability is that a node is not even one full degree away from the mean degree. It is sensible that this probability is low and thus the bounds are very high.

The smaller the network becomes, the higher the probability that the degree will be more than 0.1 distance away more the mean degree. We therefore also studied distances of 0.5. The results are shown in figure 4.5a and 4.5b.We see that networks of about 30 nodes or more can stay below this bound with a probability of more than 70%.

Like in the case where we investigated a distance of 1 degree from the mean, with a distance of 0.5, it is also possible that an incorrect minority of signals could change into an incorrect majority of votes. Again Theorem 4.2.3 can help in expressing the likeliness of this event for a specific network. For example, suppose we have a UCRN $\langle N, c, p, r, t\rangle$ with $n = 10$ and $c = 0.3$ and $d = 0.5$. The mean degree in this case is 3. Applying the Chernoff Bounds in Theorem 4.2.3, the probability that a vertex has a degree of $1.5 \cdot 3 = 4$ is at most 0.74, and the probability that a vertex has a degree of $0.5 \cdot 4 = 1$ is at most 0.66. Based on this information, we can say that the network in Figure 4.6 has thus a moderate probability of occurring from our UCRN. In the figure, all agents, except for agent 4 and 6, have a degree equal to the mean degree. Agent 4 has a degree of 2, which occurs with a probability of at most 0.66. Agent 6 has a degree of 4, which occurs with a probability of at most 0.74. In this network of Figure 4.6, both signals occur equally often. Due to the spread of communication, $\delta(b) = 16$ and $\delta(a) = 14$ and thus b has a weighted majority. If a would be the correct signal, this could potentially lead to an incorrect majority of votes. With the use of Theorem 4.2.3, we can express that this network has a probability of at most 0.78 of occurring. This is an upper bound, and thus most likely the probability of such a network occurring is actually much lower. So like before, even though it is possible to obtain an incorrect majority due to 50% percent deviations from the mean, given a specific UCRN, we can bound this probability using Theorem 4.2.3.

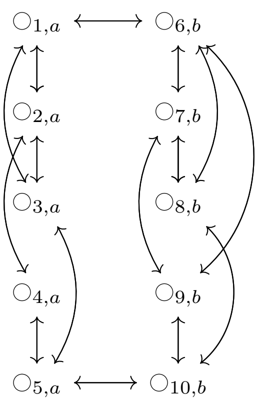


Figure 4.6: Network in which all degrees deviate at most 0.5 percent from the mean degree $\bar{\delta} = 3$

4.3 Summary

In this chapter, we focused on determining the group accuracy in random networks. One of the factors that can influence group accuracy in random networks it the spread of communication. As the majority illusion effect shows, an asymmetry in communication can lead the group astray. Therefore, we have introduced weights to measure the spread of communication that takes place in a random network. This weight represents the influence due to communication of an agent in a network. We set out three reasons to study networks in which signals (tend to) have the same weight. Firstly, evenly spread communication implies that correlation is also homogeneously spread in the network. Secondly, Lemma 4.1.1 shows that if each signal has the same weight, a correct majority of signals implies a correct weighted majority. Thirdly, Lemma 4.1.2 states that if every signal has the same weight, then asymptotically, the group reaches a correct majority with probability 1. This implies that on random networks, despite correlation between agents, we can still obtain accurate majorities under some constrains, even as accurate as in Condorcet's or Coughlan's model.

In second half of the chapter we set out to determine under what circumstances signals in a network have a high probability to have the same weight. We first showed that this is the case when there is either no communication at all or when everyone communicates with everyone. However, we are mostly interested in what happens between those two extreme cases. For this our results were twofold. We derived one asymptotic result, showing that as the network size approached infinity, every agent has the same weight. For an approximate, yet, more realistic result, using the Multiplicative Chernoff Bound, we formulated Theorem 4.2.3 that bounds the probability that the weight of a signal in the network is a given distance away from the mean weight. We gave a general impression on the bounds for certain distances. Furthermore, we showed how to use Theorem 4.2.3 to determine the probability for a given network to occur given an UCRN, thereby being able to tell the probability that asymmetries in spread of communication might lead the group accuracy astray. Finally, using properties of the variance and standard deviation, we showed that weight differences are most likely when $c = 0.5$ and decrease the farther away c is from 0.5. And the less differences in weight, the less likely group accuracy will be distorted due to unevenly spread communication.

In this chapter we only studied the effect of the amount of communication connections on the group accuracy. This is, however, only one factor that can influence the group accuracy. Other than the amount, it is also important to whom agents are connected. In particular, as Stewart et al. (2019) show, even when each person communications with equally many other agents and the number of agents supporting each alternative a or b is equally many, the structure of the network can be such that the voting outcome is swayed towards one alternative. To get a better grip on the group accuracy in random networks, future research should thus not only determine the probability of differences in amount of communication, but also study the specific structure of the communication networks.

Chapter 5

Conclusion

5.1 Summary of results

This thesis has provided a middle ground between to existing models. On the one hand, Condorcet's model (1785) allows for fairly inaccurate agents to vote on an epistemic issue that will be decided by majority rule. There is no communication and thus no correlation between votes. As a result, the majority rule ensures a high group accuracy in this case. On the other hand, Coughlan's model (2000) is set up for rather accurate agents that share all their information with each other. As a result there is full correlation between votes, and the individual accuracy will be equal to the group accuracy.

The middle ground is the situation in which communication is not fully public, nor completely absent. Our aim was to study both individual accuracy as well as group accuracy under majority rule for partial communication. And we did so for both regular as well as random networks. In Chapter 2 we studied regular communication networks, in which agents talk to a (strict) subset of the other agents, but the size of this subset is the same for all agents in the network. We derived Theorem 2.5.4 to determine the expected individual accuracy of an agent in a regular network. We observed that both Coughlan's as well as Condorcet's model are regular networks, and thus the expected accuracy of an individual agent can be derived from Theorem 2.5.4. Furthermore, we stated that asymptotically, if $p > 0.5$ the probability of a correct majority is 1 for any regular network. Again, since Coughlan's as well as Condorcet's models are regular networks, it follows from this theorem that in their models Theorem 2.5.4 applies too.

In Chapter 3 we studied individual accuracy in random networks. We introduced universally competent random networks (UCRN): a random network with n agents, all with the same competence $p > 0.5$, that are connected with each other with probability c. Each agent has a prior probability r about the state of the world and a threshold t for belief. We formulated Theorem 3.2.1 that expresses the expected individual accuracy in an UCRN. As the class of regular networks in contained in the class of random networks, it follows that Theorem 3.2.1 is a generalisation of Theorem 2.5.4. We spelled out the influence on the expected individual accuracy of the size of the network n, the connectedness c and the competence p. We finally observed that if the prior is equal to the threshold for belief, then the Bayesian approach coincides with the maximum likelihood estimation approach.

In Chapter 4 we studied group accuracy in random networks. More precisely, we studied how unevenly spread communication can influence the group accuracy. We introduced weights to measure the spread of communication. The first half of Chapter 4 consisted of arguing why it is important for the group accuracy that signals (tend to) have the same weight and the second half of Chapter 4 consisted of determining the probability that signals in a random network (tend to) have the same weight. Regarding the first half, we gave three reasons. Firstly, evenly spread communication implies that correlation is also homogeneously spread in the network. Secondly, Lemma 4.1.1 shows that if each signal has the same weight, a correct majority of signals implies a correct weighted majority. Having a correct weighted majority reduces the chance of scenarios such as the majority illusion effect (Lerman et al. (2016)), in which an asymmetry in communication can lead the group astray. Thirdly, Lemma 4.1.2 states that if every signal has the same weight, then asymptotically, the group reaches a correct majority with probability 1. This implies that on random networks, despite correlation between agents, we can still obtain accurate majorities under some constrains, even as accurate as in Condorcet's or Coughlan's model. Regarding the second half, we discussed three scenarios for determining the probability that signals in a random network have the same weight. Firstly, we showed that signals have the same weight if $c = 0$ or $c = 1$. Secondly, signals have the same weight if the network approached infinity. Thirdly, we formulated Theorem 4.2.3 that bounds the probability that the weight of a signal in the network is a given distance away from the mean weight.

An overview of the main results we discussed in this thesis is given in Table 5.1. Our main results concern both expressions for expected individual accuracy as well as group accuracy, as shown in the rows of Table 5.1. The type of networks for which we determined accuracy are divided in four. Firstly, there is Condorcet's network in which $c = 0$. Secondly, there is Coughlan's network in which $c = 1$. Thirdly, we discussed regular networks in which all agents are connected with the same number of other agents. Thirdly, there are random networks in which each two agents are connected with each other with a probability c.

5.2 Future work

We believe that this thesis provides a good framework to study the influence of communication on the accuracy of a group, as well as of the individual that compose the group. Yet, there are still many ways in which our framework can be expanded and has opened up a new area for various research questions.

	Condorcet	**Coughlan**	**Regular**	**Random**
Individual Accuracy	Corollary 2.5.4.1	Corollary 2.5.4.2	Theorem 2.5.4	Theorem 3.2.1
Group Accuracy	Theorem 2.2.1	Observation 2.3.1	Theorem 2.5.6	Lemma 4.1.2 Theorem 4.2.3

Table 5.1: Summary of results for expected individual accuracy and group accuracy, in four types of networks: Condorcet's and Coughlan's network, regular networks and random networks.

Firstly, we made various observations on the influence of the parameters on the individual accuracy as well as group accuracy of a communication network, but more work on the exact influence of the parameters would benefit a better understanding of the agent's accuracy in communication networks. For instance, we only shortly discussed the influence of the prior and the threshold on the expected individual accuracy in Chapter 3, but we do not have a precise understanding of these parameters yet. What is more, we only studied the influence of each parameter separately. It would be interesting to see how certain parameters together influence the expected accuracy. We already observed for example that in case the prior is equal to the threshold, then the Bayesian approach coincides with the maximum likelihood estimation approach.

Secondly, it is important to realise that in this thesis, we constantly have talked about *expected* accuracy. This is because we studied a given class of networks with some given parameters, but we never studied particular networks themselves. It is thus important to realise that the actual accuracy in a given network can be very different than our given expected accuracy. What is nice about our approach, is that it enables us to say something about a very big and general class of networks. The drawback, however, is that it is therefore not very sensitive to specific network structures. This was for example visible in Chapter 4. Theorem 4.2.3 enables us, given an UCRN, to bound the probability that group accuracy is distorted as a result of unevenly spread communication. However, this probability is merely a *bound* and could be made more precise if we had more information about the specific structure of the network.

In line with this remark, more work can still be done on the group accuracy in random networks. As we already pointed out, determining group accuracy after communication is non trivial, as after communication, beliefs are not independent anymore. What is more, since we are dealing with random networks, we do not know the exact structure of the network. This makes it in the current framework hard to derive a precise result on the expected group accuracy in a random network. Our path has been to focus on the influence of the number of connections agents have in a network. However, other than *number* of connections, also the *type* of signals an agent receives will influence the group accuracy. This plays a role not only in random networks, but also in regular networks. To fully study the influence of communication on the accuracy of agents, future research could determine the spread of types of signals throughout networks.

Thirdly, we only focused on group accuracy with infinitely many agents, stating that the likelihood of a correct majority tends to certainty as the group becomes infinitely big. One might argue that this assumption makes our group accuracy result not applicable to real-world applications. One way to go about this would be to generalise the growing-reliability claim of the CJT to regular and random networks, according to which larger groups are better at finding the truth than smaller groups of individuals, without assuming there are infinitely many agents. Indeed, as Dietrich and Spiekermann (2019) argue, this claim would be better applicable to real applications. Thus, generalising also this claim of the CJT to random networks would benefit the applicibality of our framework.

Fourthly, it would be interesting to extent the current synchronic process to a diachronic process. Right now, we study a setting in which agents receive signals, share them with each other, update their belief and vote. Now assume that this voting outcome is publicly announced, and then the same process starts over again. How will the accuracy be influenced if we iterate this process? A challenge for this approach will be to determine the posterior probability after the public voting outcome. Since communication took place before the voting round, the votes are not independent and thus one would need to determine the expected correlation.

Bibliography

Anderson, E. (2006). The Epistemology of Democracy. *Episteme*, 3(1-2):8–22.

Baltag, A., Rad, S. R., and Smets, S. (2021). Tracking probabilistic truths: a logic for statistical learning. *Synthese*, 199(3-4):9041–9087.

Baltag, A. and Smets, S. (2008). Probabilistic dynamic belief revision. *Synthese*, 165(2):179–202.

Brandt, F., Conitzer, V., Endriss, U., Lang, J., and Procaccia, A. D. (2016). Handbook of computational social choice. In *Handbook of Computational Social Choice*.

Chambers, S. (2003). Deliberative democratic theory. *Annual review of political science*, 6(1):307–326.

Chernoff, H. (1952). A Measure of Asymptotic Efficiency for Tests of a Hypothesis Based on the sum of Observations. *The Annals of Mathematical Statistics*, 23(4):493–507.

Chung, H. and Duggan, J. (2020). A Formal Theory of Democratic Deliberation. *American Political Science Review*, 114(1):14–35.

Coughlan, P. J. (2000). In Defense of Unanimous Jury Verdicts: Mistrials, Communication, and Strategic Voting. *American Political Science Review*, 94(2):375–393.

de Condorcet, M. (1785). Essay on the application of analysis to the probability of decisions: Returned to the plurality of votes. *From the Royal Printing Office*.

Dietrich, F. and Spiekermann, K. (2013). Epistemic democracy with defensible premises. *Economics and Philosophy*, 29(1).

Dietrich, F. and Spiekermann, K. (2019). Jury Theorems. In Fricker, M., Graham, P. J., Henderson, D., and Pedersen, N. J. L. L., editors, *The Routledge Handbook of Social Epistemology*, pages 386–396. Routledge, New York.

Ding, H. and Pivato, M. (2021). Deliberation and epistemic democracy. *Journal of Economic Behavior and Organization*, 185:138–167.

Erdös, P. and Rényi, A. (1959). On Random Graphs. *Publicationes Mathematicae*, 6:290–297.

Feddersen, T. and Pesendorfer, W. (1998). Convicting the Innocent: The Inferiority of Unanimous Jury Verdicts under Strategic Voting. *American Political Science Review*, 92(1):23–35.

Feld, S. L. (1991). Why Your Friends Have More Friends Than You Do. *American Journal of Sociology*, 96(6):1464–1477.

Foley, R. (1993). *Working Without a Net*. Oxford University Press, Oxford.

Goldbach, R. (2015). *Modeling Democratic Deliberation*. Master Thesis, Institute for Logic, Language and Computation, Amsterdam.

Grofman, B., Owen, G., and Feld, S. L. (1983). Thirteen theorems in search of the truth. *Theory and Decision*, 15(3):261–278.

Hahn, U., von Sydow, M., and Merdes, C. (2019). How communication can make voters choose less well. *Topics in cognitive science*, 11(1):194–206.

Kehoe, T. J. (1989). Comparative Statics. In Eatwell, J., Milgate, M., and Newman, P., editors, *General Equilibrium*, pages 76–83. The Macmillan Press Limited, London and Basingstoke.

Ladha, K. K. (1992). The condorcet jury theorem, free speech, and correlated votes. *American Journal of Political Science*, pages 617–634.

Leitgeb, H. (2014). The stability theory of belief. *Philosophical Review*, 123(2):131–171.

Lerman, K., Yan, X., and Wu, X. Z. (2016). The "majority illusion" in social networks. *PLoS ONE*, 11(2):e0147617.

Lumet, S. (1957). *Twelve Angry Men*. [Film] Orion-Nova Productions.

Mackay, C. (1962). *Extraordinary popular delusions and the madness of crowds*. New York, [new ed.], 12th pr. edition.

Michelini, M. (2021). *When Being the Fifth Wheel Pays Off: Wisdom of the Crowds with Costly Information*. Master Thesis, Institute for Logic, Language and Computation, Amsterdam.

Stewart, A. J., Mosleh, M., Diakonova, M., Arechar, A. A., Rand, D. G., and Plotkin, J. B. (2019). Information gerrymandering and undemocratic decisions. *Nature*, 573(7772):117–121.

Surowiecki, J. (2004). *The Wisdom of Crowds: Why the Many are Smarter Than the Few and How Collective Wisdom Shapes Business*, volume 75.

van Benthem, J., Gerbrandy, J., and Kooi, B. (2009). Dynamic update with probabilities. *Studia Logica*, 93(1):67–96.

van Ditmarsch, H., van der Hoek, W., and Kooi, B. (2008). *Dynamic Epistemic Logic*, volume 337. Springer, Dordrecht.

van Steen, M. (2010). *An Introduction to Graph Theory and Complex Networks.* Maarten van Steen, 1st edition.