On Dynamic Stubbornness in the Concatenated Friedkin-Johnsen Model

Johan Swartling Sennhed

LUND UNIVERSITY

Department of Automatic Control

BSc Thesis
TFRT-7669
ISSN 0280-5316

Department of Automatic Control
Lund University
Box 118
SE-221 00 LUND
Sweden

Printed in Sweden by Tryckeriet i E-huset
Lund 2024

Abstract

We study opinion dynamics in social networks using a version of the concatenated Friedkin-Johnsen model, where the opinion evolution of the participating agents is allowed to continue over sequences of related issues. These agents have a certain associated stubbornness that anchors them to their initial opinion, which we allow them to update between issues. The central purpose of the present work is to investigate the dynamics of this update, or in other words how different stubbornness-updating functions affect the behaviour of the model. For two convex and concave families of updating functions, we numerically explore which conditions guarantee that the agents will reach consensus, and conclude that the chosen functions can be well-approximated by linear ones with regard to consensus convergence criteria. We also study how dynamic stubbornness affects the depolarizing properties of the model, and show that agents becoming more stubborn does not always lead to more polarized opinion distributions. Lastly, we introduce a version of the model where dissatisfied agents entirely refuse to change their opinion until other agents approach their position enough. For this model, we provide some sufficient conditions for consensus, and give numerical examples that illustrate its oscillatory behaviour.

Acknowledgements

I would first like to thank my supervisors: Emma Tegling for many inspiring discussions and for introducing me to the fascinating subject of Opinion Dynamics, and David Ohlin for incredible amounts of insightful feedback as well as allowing me to develop his idea for the model discussed in Chapter 4. This has truly been a very fun and engaging project, much thanks to your great mentorship. I would also very much like to thank my close friend Sandra Bjarnetoft, who inspired me to pursue writing a Bachelor's thesis in the first place, for which I am very grateful.

Contents

1

Introduction

As a social species relying heavily on cooperation, humans have always had something to gain by understanding how people might influence one another, for example to be able to effectively agree on a common goal to work towards. This is all the more relevant in the globalized and highly connected modern world, where individuals might interact with, and be influenced by, more people than ever previously possible. A very concrete example of this is that of social media, which has become an ubiquitous feature of modernity, and at the same time is often attributed as being a cause of the increasing polarization of contemporary politics [Kubin and Sikorski, 2021]. This plainly highlights the need to increase our understanding of how opinions and influence spreads, and to develop strategies for effectively modeling it in a social network context. One of the efforts to this end that has garnered growing attention during recent years is an interdisciplinary field drawing from psychology, sociology, engineering and mathematics, known to us as opinion dynamics [Proskurnikov and Tempo, 2017; Noorazar et al., 2020; Peralta et al., 2022; Xia et al., 2011].

We adopt a control-based perspective of this field, building upon graph theory-oriented models that have initially been constructed on the basis of sociological and psychological work [French, 1956; DeGroot, 1974; Friedkin and Johnsen, 1990] to only then be given more meticulous mathematical attention [Proskurnikov and Tempo, 2017]. These purely mathematical models are often initially very simple to make rigorous analysis of their properties possible. However, in the construction of such models, a significant portion of the intricacies and nuances of reality is undoubtedly lost. The field of opinion dynamics is largely dependent on this balancing act - creating models that are simple enough that clear-cut mathematical results can be derived, while also being intricate enough to give meaningful insight into real issues [Noorazar et al., 2020]. In this pursuit, we believe that there can be great value in adding new layers of complexity to existing models, while retaining analytical tractability, in the hope that this can render the models more useful and able to depict increasingly sophisticated concepts. This is also how other models in the field have formed [Proskurnikov and Tempo, 2017].

In this thesis, we focus on one of the most widely used models in the field, the Friedkin-Johnsen (FJ) model, in which the agents that partake in a discussion update their opinions by weighted averaging of all opinions they interact with [Proskurnikov and Tempo, 2017]. More specifically, we consider a model that treats multiple discussions arranged in sequence, known as the concatenated FJ-model, that has been used to model, among other things, the 15 years of discussions leading up to the Paris agreement [Bernardo et al., 2021]. Inspired by the work [Ohlin et al., 2022], we model a scenario where agents update their stubbornness, or willingness to compromise, between each issue. This update is modeled as a reaction to the outcome of the discussion, in our case determined by a vote, and is carried out by a stubbornness-updating function with this explicit purpose. This is what we refer to as the concatenated FJ-model with dynamic stubbornness.

Aims and Contributions

This thesis aims to investigate how different stubbornness-updating functions affect the behaviour of the concatenated FJ-model. We numerically determine what conditions are needed to guarantee that the model reaches consensus for two different families of stubbornness-updating functions, and find that these are approximated well by linear ones in terms of the aforementioned conditions. Furthermore, we study how different stubbornness-updating functions affect the polarization of the opinion distribution, measured in multiple relevant metrics, and show that a strictly larger stubbornness-updating function does not always give more polarized results. Finally, we introduce an entirely novel version of the model, featuring a special stubbornness-maximizing threshold update function. We analyze this model numerically and theoretically, providing sufficient conditions that guarantee consensus, both generally and for some special cases.

Thesis Outline

In Chapter 2, we provide some theoretical background before introducing our model along with the two concepts (consensus convergence and depolarization) we utilize to compare the effects of different stubbornness-updating functions on it. In Chapter 3 we use these concepts to analyze how different update functions affect the behaviour of the model, both theoretically and numerically. Subsequently in Chapter 4, we present a version of the model with a stubbornness-maximizing threshold, which we find to have interesting properties, analyze it theoretically and provide some numerical examples of its behaviour. Following this, results and conclusions are summarized in Chapter 5, where we also discuss potential directions for future work. Finally, some auxiliary proofs are also provided in the appendix.

2

Preliminaries and Problem Setup

2.1 Graph Theory

Here, a short introduction to graph theory is provided, as many such these concepts are integral to opinion-dynamical modeling. All definitions in this chapter can be found in a standard textbook such as [Godsil and Royle, 2001], or for a slightly more extensive introduction than ours that is also centered on opinion dynamics, we would refer to [Proskurnikov and Tempo, 2017], which is the main inspiration for the way the topic is presented here.

DEFINITION 2.1 (Graph) A (directed) graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ consists of a pair of sets: The set $\mathcal{V} = \{1, 2, ..., n\}$ containing the *nodes* of the graph and the set $\mathcal{E}$ containing its *edges*. The elements in $\mathcal{E}$ are pairs of nodes, and the edge $(i, j) \in \mathcal{E}$ represents a link from the nodes i to j, $(i, j \in \mathcal{V})$. □

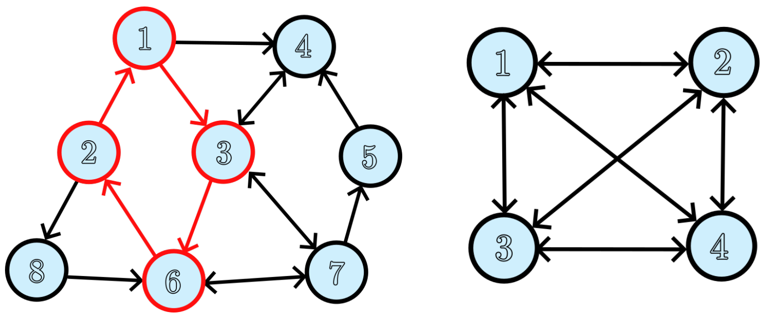
DEFINITION 2.2 (Subgraph) A graph $\mathcal{G}' = (\mathcal{V}', \mathcal{E}')$ is called a subgraph of a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ if $\mathcal{V}' \subset \mathcal{V}$ and $\mathcal{E}' \subset \mathcal{E}$ such that $i, j \in \mathcal{V}'$ for all $(i, j) \in \mathcal{E}'$. More plainly, a subgraph is the result obtained when taking a graph and removing some of its nodes and edges. □

DEFINITION 2.3 (Walk) A (directed) walk is a finite or infinite sequence of nodes $i_0, ..., i_k \in \mathcal{V}$ for which $(i_{n-1}, i_n) \in \mathcal{E}$. It can be thought to represent moving between the nodes along the edges that connect them, going from the starting node i_0 to the end node i_k. The *length* of the walk is equal to the number of nodes in the sequence minus one, or the number of edges traversed. A *path* is a walk where no node appears twice, and a *trail* is a walk were no edge appears twice. □

DEFINITION 2.4 (Cycle) A cycle is a trail with at least one node, where the first and last nodes are the same, but no other nodes appear twice. □

DEFINITION 2.5 (Connectivity) Two nodes are *connected* if there exists a walk between them. A node is called a *root node* if it is connected to all other nodes. A graph is called *quasi-strongly connected* if it has at least one root node. If every node is a root node, meaning that there exists a walk between any two nodes, it is said to be *strongly connected*. □

DEFINITION 2.6 (Complete Graph) In a complete graph, the edge (i, j) exists in $\mathcal{E}$ for any two nodes $i, j \in \mathcal{V}$. In other words, all nodes are connected to each other with walks of length 1 in both directions. A complete graph is also called *fully connected*. □



A stronly connected graph with 8 nodes. A cycle is marked in red.

A complete graph with 4 nodes.

Figure 2.1 Illustration of strongly connected a graph with 8 nodes (left), and a complete graph with 4 nodes (right), with an arrow between two nodes indicating that a link between them exists in the edge set.

2.2 Historical Background

In this section, we begin introducing our model with a brief historical background. The concatenated FJ-model has gradually evolved over many years, and introducing these evolutions chronologically is a convenient way to introduce the mathematical concepts step by step while also providing historical context. It should also be mentioned here that many other opinion dynamical models exist, focusing on different aspects of opinion formation. For a more complete overview of the field, see e.g. [Proskurnikov and Tempo, 2017; Noorazar et al., 2020; Peralta et al., 2022; Xia et al., 2011]. The models relevant to this thesis mainly pertain to situations where a group are actively seeking to establish consensus. An obvious example that is useful to keep in mind throughout is that of lawmakers trying to agree on policy or negotiating a political deal [Ohlin et al., 2022; Bernardo et al., 2021; DeGroot, 1974]. Naturally, the models may be applied much more generally than that, to practically any group of people, between some of which there is an exchange of opinions concerning some issue. However, the examples mentioned above appear both salient and illustrative to us, as the benefits of reaching consensus are clearly established.

French-DeGroot

Our model can be said to have its origins in the 50s, beginning with the work of the Psychologist John R.P. French in creating a formalized theory of the social power individuals are able to exert over each other [French, 1956]. Later, it was generalized and given a more formal mathematical shape by Morris DeGroot [DeGroot, 1974] and is thus known as the French-DeGroot model.

The basic principle is that during a discussion, participants will gain new knowledge, insights and perspectives by interacting with the other parties, that will in turn lead to them to update their stance on the issue(s) at hand [DeGroot, 1974]. The model also aims to capture some aspects of fairly intricate social phenomena, such as the power dynamics and relationships in the group, and to which extent different participants interact with each other [French, 1956]. When the opinion-updating process is iterated, all participants will after some time, under very slight conditions on the structure of the network, come to a consensus [Proskurnikov and Tempo, 2017].

Mathematically, this is defined as follows. Consider a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ where the nodes in $\mathcal{V}$ represent the n participants in the meeting, from now on referred to as *agents*. The edges can be thought of as representing interactions or connections between the agents, with the link (j,i) being present in $\mathcal{E}$ indicating that agent j is exerting some amount of influence on i. The extent of this influence is determined by the *influence weights* w_{ij}, which are collected for all agents in the matrix $W = (w_{ij})$. Typically the weights are normalized such that W is row-stochastic, meaning that $\sum_j w_{ij} = 1$. Each agent i is also associated with some opinion $y_i(t)$ defined on the interval $[0,1]$ where $t = 0,1,2,...$, and the change this opinion undergoes during the meeting is modeled as

$$y_i(t+1) = \sum_{j=1}^{n} w_{ij} y_j(t) \tag{2.1}$$

or for all agents at once as

$$y(t+1) = Wy(t) \tag{2.2}$$

where $y(t) = (y_1, y_2, ..., y_n)$ is the *opinion vector* or *opinion distribution*. At every time step, the new opinion of agent i becomes a weighted average of the opinions of all agents that i interacts with. We will only deal with the case where $y \in \mathbb{R}^{n\times 1}$, but it could just as easily be a matrix encoding multiple opinions on different subjects discussed in the same meeting [Proskurnikov and Tempo, 2017]. We will now briefly cover some convergence criteria.

DEFINITION 2.7 The model (2.2) is convergent if the limit

$$\lim_{t\to\infty} y(t) = \lim_{t\to\infty} W^t y(0) \tag{2.3}$$

exists for all possible initial opinion vectors $y(0)$. □

By convergence to *consensus*, we mean that all agents come to share the same opinion as the number of iterations goes to infinity. More formally, that $\lim_{t\to\infty} y_i(t) = y_\infty \forall\ i\ \in \mathcal{V}$ for some final opinion $0 \le y_\infty \le 1$.

The French-DeGroot model is always convergent if the weight $w_{ii} > 0\ \forall\ i \in V$, and always reaches consensus if $\mathcal{G}$ at least quasi-strongly connected [Proskurnikov and Tempo, 2017].

One of the main advantages of this model is that it is very simple, and thus easy to analyze. This is also an obvious point of critique. In reality, many social networks exhibit more lasting conflict, failing to reach a consensus or even forming polarized clusters [Proskurnikov and Tempo, 2017], which is clearly not captured by the French-DeGroot model. As long as there is at least one root, in this context referring to an agent whose opinion exerts some influence on all others, everyone will be able to agree, which is certainly not very realistic in many situations.

Friedkin-Johnsen

A model that is arguably better suited than French-DeGroot for capturing the dynamics of longer-lasting disagreement is the FJ model. It was introduced in 1990 in a joint work by Noah E. Friedkin (sociologist) and Eugene Johnsen (mathematician) [Friedkin and Johnsen, 1990], and has since come to be very influential, partly because it is one of relatively few models to be experimentally verified for smaller groups [Friedkin and Johnsen, 1999; Proskurnikov and Tempo, 2017]. It can be said to expand upon the work of French and DeGroot by proposing that the agents might be made differently susceptible to interpersonal influence [Friedkin and Johnsen, 1999], or simply more or less *stubborn*. Capturing this mathematically, each agent i is also associated with a *stubbornness coefficient* θ_i : $0 \le \theta_i \le 1$ and opinion is updated as

$$y_i(t+1) = (1-\theta_i)\sum_{j=1}^{n} w_{ij} y_j(t) + \theta_i y_i(0) \tag{2.4}$$

or for all agents at once as

$$y(t+1) = (I-\Theta)Wy(t) + \Theta y(0) \tag{2.5}$$

where $\Theta = diag(\theta_1, ..., \theta_n)$.

This essentially tethers the agents to their initial opinion, with the strength of the tethering determined by the stubbornness coefficient. It is assumed that the initial opinions have been formed by some exogenous conditions, possibly different for each agent, and that individual agents may be inclined to hold on to their beliefs more or less tightly [Proskurnikov and Tempo, 2017; Friedkin and Johnsen, 1990]. An agent with $\theta_i = 1$ is referred to as *fully stubborn* and will not alter their opinion at all, while the model reduces to French-DeGroot (2.4) for $\theta_i = 0$. Observe that the stubbornness coefficient has no intrinsic relation to the opinion value, and that if an agent has an initial opinion value that differs greatly from all other agents, it is not necessarily a stubborn agent. Stubbornness can instead be thought of as a measurement of how difficult it is to convince an agent to change their opinion. There is a game-theoretic interpretation of the FJ-model that gives useful intuition regarding this: Agents can at every time step be interpreted as striving to minimize their individual cost function

$$J_i(s) = (1-\theta_i)\sum_{j=1}^{n} w_{ij}(y_i(t)-y_j(t))^2 + \theta_i(y_i(t)-y_i(0))^2 \tag{2.6}$$

and the limit of y as $t \to \infty$ as a Nash equilibrium [Proskurnikov and Tempo, 2017].

Some literature instead employs a slight variation of the model, and uses the terms *expressed opinion* for y_i and *internal opinion* for $y_i(0)$, suggesting that an agent's internal opinion remains fixed [Matakos et al., 2017]. This is an illustrative terminology for some situations - during a political negotiation, an agent's deeply held beliefs might not change, but they can still accept an agreement that involves some compromise. Here, the stubbornness coefficient might instead be thought of as modelling the negotiation tactic of the agent and how much they are willing to compromise [author's comment].

Notably, the model is fully deterministic - e.g. when it is clear that the model will converge, the iteration process is not strictly necessary, and the final opinion distribution can be explicitly calculated. Letting $t \to \infty$, (2.5) gives that:

$$\lim_{t\to\infty} y(t+1) - (I-\Theta)Wy(t) = \lim_{t\to\infty}(I-(I-\Theta)W)y(t) = \Theta y(0) \Leftrightarrow$$

$$y(\infty) = (I-(I-\Theta)W)^{-1}\Theta y(0), \tag{2.7}$$

where $y(\infty)$ is the opinion distribution the model converges to. An important special case can be found when $\mathcal{G}$ is fully connected and has equal weights, $W = \frac{1}{n}\mathbf{1}_n\mathbf{1}_n^T$. Then the weighted mean opinion reduces to the mean opinion $\bar{y} = \frac{y_1+y_2+\ldots+y_n}{n}$ for every agent and each agent updates their opinion according to

$$y_i(t+1) = (1-\theta_i)\bar{y}(t) + \theta_i y_i(0). \tag{2.8}$$

In this case, there is a useful upper bound for how much the distance between the maximum and minimum opinions can decrease.

THEOREM 2.1
Let $\mathcal{G}$ be a graph of opinionated agents with the influence matrix $W = \frac{1}{n}\mathbf{1}_n\mathbf{1}_n^T$. If we denote the maximum and minimum opinion values $y_{max}(t)$ and $y_{min}(t)$ respectively,

and the maximum stubbornness value among all the agents $\hat{\theta}$, it holds that

$$\lim_{t\to\infty}(y_{max}(t)-y_{min}(t)) \leq \hat{\theta}(y_{max}(0)-y_{min}(0)) \qquad \square$$

Proof. By definition, it holds that $y_{max}(t) > \bar{y}(t) > y_{min}(t)$ if $y_{max}(t) \neq y_{min}(t)$, meaning that the distance between the maximum and minimum opinions will be reduced during every time step. If an agent with one of these opinions were to have a higher θ-value, it would move its opinion less during every time step, meaning that the distance between the maximum and minimum opinions would also reduce less. This means that the least possible reduction must be found if both the *max*- and *min*-agent has the stubbornness value $\hat{\theta}$. Letting $t \to \infty$ in (2.8) gives that

$$\lim_{t\to\infty}(y_i(t)-y_j(t)) = \bar{y}(t)(\theta_j-\theta_i)+\theta_i y_i(0)-\theta_j y_j(0)$$

and if we let $i = max$, $j = min$ and let $\theta_i = \theta_j = \hat{\theta}$, we get

$$\lim_{t\to\infty}(y_{max}(t)-y_{min}(t)) \leq \hat{\theta}(y_{max}(0)-y_{min}(0)). \qquad \square$$

This model is said to converge when the limit (2.7) exists, which it does if at least one root agent has a stubbornness-value greater than 0 [Proskurnikov and Tempo, 2017]. The distribution that the model converges to is guaranteed to be a point in the convex hull of the initial conditions, meaning that no agent can have a final opinion value that is greater than the maximum or less than the minimum opinion at $t = 0$. However, convergence to consensus is not generally achieved [Wang et al., 2021], in contrast to the French-DeGroot model.

Concatenated Friedkin-Johnsen

In recent years, a version of the FJ-model that depicts a series of issues being treated in sequence has been proposed, known as the concatenated FJ-model [Ohlin et al., 2022; Bernardo et al., 2021; Wang et al., 2021; Mirtabatabaei et al., 2014; Wang et al., 2022]. Notably, a version of this model was used to model the many years of negotiations that preceded the Paris climate agreement, and was able to identify some of the most influential agents in the negotiation process [Bernardo et al., 2021].

In a concatenated FJ-model, a secondary timescale of *issues* is introduced, and different aspects of the model allowed to depend on both the issue s and the time step t within the issue. Imagine first a set of people discussing some issue and trying to come to an agreement. This is modeled by the standard FJ-model. After this meeting, be it immediately or after some length of time, the group is once again assembled to discuss another issue. These issues can, but do not necessarily have to be related, for the model to be used, and it is intuitively clear that a model with this setup could give more insight into the situations initially described. A parliament discussing a wide range of different political issues in sequence is an intuitive example, or as in the case with the Paris agreement, the negotiations for a single agreement being too complex to resolve in a single meeting and instead handled as a collection of related issues [Bernardo et al., 2021].

Between any two such meetings, some conditions might change. The composition of the group can be altered, relationships developed, negotiation tactics shifted and so on. The effects of different such changes tend to be the focus of work done using the concatenated FJ-model [Ohlin et al., 2022; Bernardo et al., 2021; Mirtabatabaei et al., 2014]. In [Mirtabatabaei et al., 2014], the influence weights in W are updated between issues as a function of social power, while the topics of discussion are independent. Focusing on multilateral large-scale international negotiations, [Bernardo et al., 2021] uses both changing network structure to model the work done in many different committees towards the same goal, and changing stubbornness to reflect how active an agent is during a specific meeting. However, there is no discussion of how the stubbornness changes across issues. That is instead the focus of [Ohlin et al., 2022], that introduces a more formal mathematical framework for working with dynamic stubbornness. This framework will form the basis for the models investigated in this thesis.

2.3 Defining the Model

Here, we define our version of the model, which we call the *concatenated FJ-model with dynamic stubbornness*, using concepts from the previous section. Consider the following model: Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a graph with n opinionated agents and the influence matrix $W = (w_{ij})$ as before. During each issue s, the opinion vector $y(s,t)$ is updated as in the FJ-model (2.5). The issues in question are presumed to be closely related, such that the opinion evolution can be considered to continue over issues, meaning that $y(s,\infty) = y(s+1,0)$. A similar strategy is also employed in [Bernardo et al., 2021]. This means, using (2.7) that

$$y(s+1,0) = (I-(I-\Theta(s))W)^{-1}\Theta(s)y(s,0). \tag{2.9}$$

As can be seen in the above equation (2.9), we also have an issue-dependent stubbornness matrix $\Theta(s) = \mathrm{diag}(\theta_1(s),...,\theta_n(s))$ where $0 \leq \theta_i(0) < 1$.

We will focus on the same stubbornness-updating mechanism as in [Ohlin et al., 2022], considering the case when the issue s is settled with a vote at the end of discussion, the result of which becomes the outcome of the meeting. This result is taken to be the median opinion, which the well-established median voter theorem states is a good approximation for many voting systems [Black, 1948]. Let $\mu(y)$: $\mathbb{R}^n \to \mathbb{R}$ be the voting function that takes an opinion distribution and returns the median opinion value. Then we can define the distance $\delta_i(y) = |y_i - \mu(y)|$ from the opinion of agent i to the voting result, and update stubbornness based on the size of δ_i:

$$\Theta(s+1) = \mathrm{diag}[f(\theta_1(s), \delta_1(y(s+1,0))),...,f(\theta_n(s), \delta(y(s+1,0)))] \tag{2.10}$$

where $f(\theta_i, \delta_i) : \mathbb{R}, \mathbb{R} \to \mathbb{R}$ is the *stubbornness-updating function* for the system. Principally, f could be made to also depend on other aspects of the model such as the W-matrix, but some more complicated such dependence is beyond the scope for this thesis.

We say that the model reaches consensus if any one of the concatenated FJ-models does, which is equivalent to all opinions in y becoming equal as $s, t \to \infty$.

Since we will be interested in comparing the effects of different stubbornness-updating functions on a network, we more generally define a concatenated FJ-model with dynamic stubbornness as a map $\mathcal{F} : \mathcal{G}, f \to y(s,t)$ that takes a graph of opinionated agents (associated with some initial y and θ-values) as well as some stubbornness-updating function $f(\theta, \delta)$, and returns what the opinion distribution will be at time step t of issue s if the concatenated FJ-model (2.9) is applied to $\mathcal{G}$ and stubbornness is updated using for each agent using f between issues.

In [Ohlin et al., 2022], two choices of stubbornness-updating-functions are discussed, both with a linear dependence on the distance δ. One with increasing stubbornness

$$f(\theta_i, \delta_i) = \theta_i + c(1 - \theta_i)\delta_i \tag{2.11}$$

where $0 \leq c \leq 1$, which can render the system $\mathcal{F}(\mathcal{G}, f)$ unable to reach consensus if $\theta \to 1$ for some agents faster than the speed at which consensus is approached. The other has decreasing stubbornness for all δ_i, which guarantees convergence to consensus if $\mathcal{G}$ is strongly connected and at least one agent maintains non-zero stubbornness across all issues [Ohlin et al., 2022]. A conservative theorem that guarantees consensus for $\mathcal{F}$ using (2.11) is also provided: (For a proof, see [Ohlin et al., 2022].)

THEOREM 2.2
With f as in (2.11), the following conditions guarantee that $\mathcal{F}(\mathcal{G}, f)$ will reach consensus:

1. $\mathcal{G}$ is a complete graph
2. $\mathcal{G}$ has equal weights ($W = \frac{1}{n}\mathbf{1}_n\mathbf{1}_n^T$)
3. $cd(0) + \theta_{max}(0) < 1$

where $d(s) = \max_{i,j} |y_i(s,0) - y_j(s,0)|$ and $\theta_{max}(s) = \max\{\theta_i(s)\}$.

The work [Ohlin et al., 2022] is, to the best of our knowledge, the only previous paper to discuss dynamic stubbornness in terms of a stubbornness-updating function dependent on a voting result, and both θ-updating functions proposed are both are fairly straightforward linear dependencies. How using a wider range of functions to update stubbornness might affect the dynamics of the model has not been explored, which leads us to the purpose of this thesis.

2.4 Methods for Model Comparisons

The aim of this thesis is to study the effect of dynamic stubbornness on the properties of the concatenated FJ-model. Specifically how the model $\mathcal{F}$ is affected by different stubbornness-updating functions $f(\theta_i, \delta_i)$. In Sections 3 and 4, we present a selection of updating dynamics and compare them based on some areas of interest. More specifically we consider two problems, defined below. Firstly, an often discussed property of opinion dynamical models are under what conditions consensus can be guaranteed [Proskurnikov and Tempo, 2017], which becomes a natural point of comparison between our models.

Problem 1: Compare different stubbornness updating functions f in terms of the model's convergence to a consensus or non-consensus state.

This is mainly done numerically for a set of example cases, as the general model proves analytically challenging. Theorem 2.2 is also quite conservative, and how the conditions it requires for guaranteeing consensus change with the choice of f is thus not a very good indication of how the actual model behaviour changes. One special choice of f is however treated more theoretically in Section 4.

Another interesting point of comparison between the different versions of the model is the state of the opinion distribution in the cases where consensus is not attained. To be able to quantitatively compare the effects of different stubbornness-updating functions in this regard, we look to the work done regarding polarization in social networks, and study the following problem.

Problem 2: In the cases where the model does not converge to a consensus, compare different stubbornness updating functions by studying the polarization of the resulting opinion distribution.

Opinion polarization is a field of study that has received much attention in recent years, meaning that there are existing frameworks that we can use for such a comparison. The work [Biondi et al., 2023] provides a general framework for analysis of polarization in different FJ-models, that can very easily be extended to concatenated FJ for our purposes. It supports use of different polarization metrics, of which three relevant ones are listed below.

$$NDI(y) = \sum_{(i,j)\in\mathcal{E}} w_{ij}(y_i - y_j)^2 \tag{2.12}$$

$$GDI(y) = \sum_{i,j\in\mathcal{V}:i<j} (y_i - y_j)^2 \tag{2.13}$$

$$V(y) = \sum_{i\in\mathcal{V}} (y_i - \bar{y})^2 \tag{2.14}$$

NDI (2.12) and *GDI* (2.13) are referred to as the Network and Global *Disagreement Index*, the former considering the network structure and measuring the total disagreement between neighbours, while the latter measures the disagreement between all agents [Biondi et al., 2023]. V is simply the variance of the opinion distribution (with division by the network size omitted). From [Biondi et al., 2023], we also retrieve the following definitions:

DEFINITION 2.8 (Polarizing model) Let $\Phi(y)$ be a polarization metric for the opinion distribution y, and let $\mathcal{M} : y_0 \rightarrow y(t)$ be some opinion formation model. Then $\mathcal{M}$ is *polarizing* in Φ if there exists an initial opinion vector y_0 such that the final opinion distribution $\lim_{t\rightarrow\infty} y(t) = y_\infty$ fulfills $\Phi(y_\infty) > \Phi(y_0)$. If $\mathcal{M}$ is not polarizing, it is instead said to be *depolarizing* under Φ.

DEFINITION 2.9 (Polarization shift) When $\mathcal{M}$ is (de)polarizing, the induced (de)polarization for an initial opinion vector under a polarization metric Φ is measured by the *polarization shift* $\Delta_{\mathcal{M}}\Phi(y_0) = \Phi(y_\infty) - \Phi(y_0)$. □

The definition (2.8) is easily expandable to a concatenated FJ-model $\mathcal{F}$ by taking y_∞ to mean the final opinion vector of $\mathcal{F}$ rather than the final opinion vector of a single FJ-model, or in other words by letting $y_\infty = lim_{s\rightarrow\infty} y(s,t)$. It is proven in [Biondi et al., 2023] that the FJ-model is depolarizing under our metrics when the graph on which it acts is fully connected, so to get a metric that can be used to compare model variations even when this is the case we also introduce our own concept of *depolarizability*:

DEFINITION 2.10 (Depolarizability) Let $\mathcal{A}$ and $\mathcal{B}$ be two different (concatenated) FJ-models and y_0 some initial opinion vector. We then say that, if both $\Delta_{\mathcal{A}}\Phi(y_0)$ and $\Delta_{\mathcal{B}}\Phi(y_0) \leq 0$ and $|\Delta_{\mathcal{A}}\Phi(y_0)| > |\Delta_{\mathcal{B}}\Phi(y_0)|$, $\mathcal{A}$ has *higher depolarizability* than $\mathcal{B}$ for y_0 under Φ. Or, equivalently, $\mathcal{A}$ is *more depolarizing* than $\mathcal{B}$ for y_0. If $\mathcal{A}$ has higher depolarizability than $\mathcal{B}$ for all y_0, we simply say that $\mathcal{A}$ has higher depolarizability (is more depolarizing) than $\mathcal{B}$. □

3

Dynamic Stubbornness Updates in the Concatenated FJ-Model

3.1 Choosing the Stubbornness-Updating Function

Desired characteristics

As previously described, dynamic stubbornness in a concatenated FJ-model is intended to model the agents' willingness to compromise changing between meetings. Letting the change in an agent's stubbornness after an issue be a function of the distance from their final opinion to the result of the latest vote essentially ties it to how satisfied they are with the outcome of the meeting [Ohlin et al., 2022] - the larger $\delta_i(y(s,\infty))$ is, the less agent i can be said to have 'gotten their way'. We propose that stubbornness increasing for an agent could represent said agent being dissatisfied with the result of the meeting, and resolving to compromise less during the next issue. Stubbornness decreasing would then correspond to an agent feeling that they are getting their way, perhaps relaxing somewhat going into the next meeting and being more open to compromise. Capturing this mathematically could be done in a few different ways.

Suppose that an agent is not satisfied if the result of the meeting does not exactly agree with their opinion. This would correspond to some $f = g(\theta_i, \delta_i)$ with $g(\theta_i, 0) = \theta_i$ and $g(\theta_i, \delta_i) \geq \theta_i$ for $\delta_i > 0$. Another choice of function that captures both behaviours, let us call it $f = h(\theta_i, \delta_i)$, could be made in the following way. Suppose that agents with small values of δ_i could be considered more or less satisfied, and in response will decrease their stubbornness. Agents with large δ_i on the other hand are classified as more or less *dissatisfied*, increasing their stubbornness. This of course also implies the existence of a point, or series of points, for which stubbornness remains the same, that constitutes a threshold between satisfaction and dissatisfaction. Where should this threshold be placed? This could be modeled as being different for every agent, but every agent having their own stubbornness-updating function would greatly complicate the model, so we instead presume that all agents share the same distance to the border, and also choose to consider the case where the threshold is only one δ_i-value wide. For this threshold distance $\delta_i = a \in (0,1)$, we have that $h(\theta_i, a) = \theta_i$, while $h(\theta_i, \delta_i) < \theta_i$ for all $\delta_i < a$ and $h(\theta_i, \delta_i) > \theta_i$ if $\delta_i > a$.

We might also have considered another type of stubbornness-updating function where $f(\theta_i, \delta_i) \leq \theta_i$ for all $0 \leq \delta_i \leq 1$. This would correspond to the agents being very optimistic, staying equally or becoming more open to compromise regardless of the result of the meeting. However, concatenated FJ-models are guaranteed to reach consensus for fairly limited requirements on θ, as long as $\theta_i < 1 \ \forall \ i$ [Wang et al., 2021]. Stubbornness only being able to reduce or stay the same is thus not believed to be able to produce interesting non-consensus states to the same degree as g and h, so this type of function was therefore not examined further.

The types of types of stubbornness-updating functions g and h are those that we will study in this section. For both types, we will also require that f be monotonically increasing ($\frac{df(\theta_i, \delta_i)}{d\delta_i} \geq 0$) in the domain. If this was not the case, there could be cases where increased disagreement leads to lower resulting stubbornness, which is very hard to justify under the presumptions that we have made.

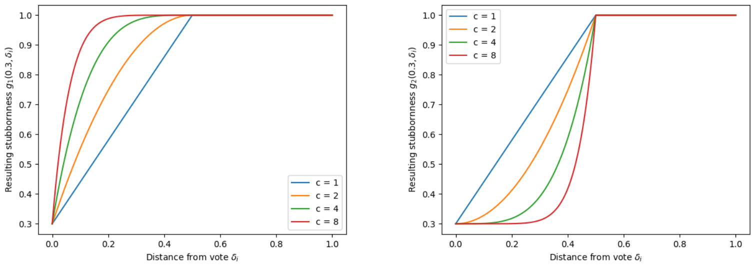
Non-linear functions

For different choices of f to be studied and compared numerically, some more explicit choices of functions fulfilling the requirements described above have to be made. For this purpose we consider the following choices of g:

$$g_1(\theta_i, \delta_i) = \begin{cases} 1 - (1-\theta_i)(1 - \frac{1}{b}\delta_i)^c & \text{if } \delta_i < b \\ 1 & \text{if } \delta_i \geq b \end{cases} \tag{3.1}$$

$$g_2(\theta_i, \delta_i) = \begin{cases} \theta_i + (1-\theta_i)(\frac{1}{b}\delta_i)^c & \text{if } \delta_i < b \\ 1 & \text{if } \delta_i \geq b \end{cases} \tag{3.2}$$

These functions both incorporate two scalar constants $c \geq 1$ and b ($0 \leq b \leq 1$), meaning that they are not really two distinct functions but rather two families of continual functions, one concave and one convex. The purpose of c is to scale the 'curvature' of g. A choice of $c = 1$ gives a linear function, and if also $b = 1$ it is equivalent to the function (2.11) with $c = 1$. If instead $c \to \infty$, the functions become step functions, g_1 going from 0 to 1 at $\delta_i = 0$ and g_2 doing the same at $\delta_i = b$. For all choices of c in between these extremes, g_1 and g_2 are concave and convex functions in δ_i, respectively. The other constant, b, determines at what δ_i the functions will reach their maximum value of 1. Some examples of these functions for $b = 0.5$ are displayed in Figure 3.1.



Concave functions g_1 Convex functions g_2

Figure 3.1 The stubbornness updating functions $g_1(\theta_i,\delta_i)$ and $g_2(\theta_i,\delta_i)$ for $\delta_i = 0.3$, $b = 0.5$ and varying c-values

Linear approximations

To help evaluate these non-linear stubbornness-updating functions, we define piecewise linear functions intended as approximations of g_1 and g_2. To approximate g_1, we use the function l_1.

$$l_1(\theta_i,\delta_i) = \begin{cases} \frac{c}{b}(1-\theta_i)\delta_i + \theta_i \text{ if } \frac{c}{b}(1-\theta_i)\delta_i + \theta_i < 1 \\ 1 \text{ otherwise} \end{cases} \tag{3.3}$$

This is simply the piecewise linear function for which $g_1(\theta_i,0) = l_1(\theta_i,0)$ and $g_1'(\theta_i,0) = l_1'(\theta_i,0)$ that levels out when $l_1 = 1$. These functions are displayed side by side in Figure 3.2

How should g_2 be approximated? The strategy used to get an approximation for g_1 is not likely to be effective, since that strategy would give that $g_2(\theta_i,\delta_i) \approx \theta_i$ for all δ_i. Instead, we suggest that an approximation may be made using a piecewise linear function such that each linear segment becomes one of the asymptotes of g_2 as $c \to \infty$, and consider the function

$$l_2(\theta_i,\delta_i) = \begin{cases} 1 \;\text{ if } (\delta_i > b) \\ \max[(5(1-\theta_i)(\frac{1}{5b})^c\delta_i + \theta_i),\ (\frac{1-k}{b}\delta_i + k)] \text{ otherwise} \end{cases} \tag{3.4}$$

for some constant k. The first term in the max-expression in (3.4) follows g_2 for small δ_i and becomes equal to θ_i for large c, while the second follows the steep rise that becomes the line $\theta_i = b$ when $c \to \infty$. For reference, l_2 and g_2 can be viewed together in Figure 3.3. In the numerical tests in section 3.2, k was chosen such that the absolute value of the integral

$$\int_0^1 \frac{1-k}{b}\delta_i + k - g_2(\theta_i,\delta_i)\,d\delta_i$$

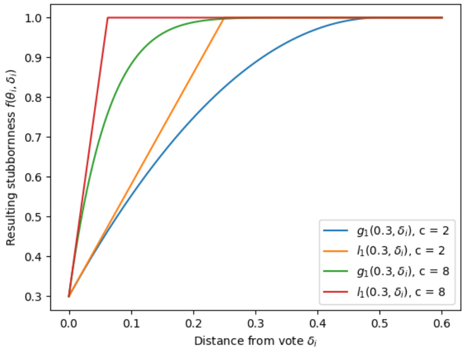


Figure 3.2 The stubbornness updating function $g_1(\theta_i, \delta_i)$ and its corresponding linear approximation l_1 for $\theta_i = 0.3$ with $c = 2$ and $c = 8$. The approximation is based on the derivative of g_1 in $\delta_i = 0$ and worsens as δ_i increases.

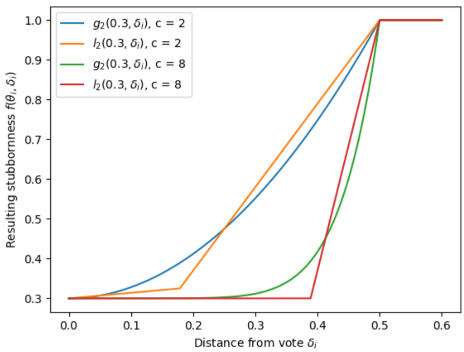


Figure 3.3 The stubbornness updating function $g_2(\theta_i, \delta_i)$ and its corresponding linear approximation l_2 for $\theta_i = 0.3$ with $c = 2$ and $c = 8$.

is minimized, giving $k = \frac{1}{2}(\theta_i - c + \theta_i c + 1)$, where c is the curvature constant in (3.2).

Note that the functions l_1 and l_2 are by no means chosen to be the best possible approximations of g in terms of having similar conditions that guarantee convergence to consensus, but simply because they superficially appear to approximate its main features fairly well.

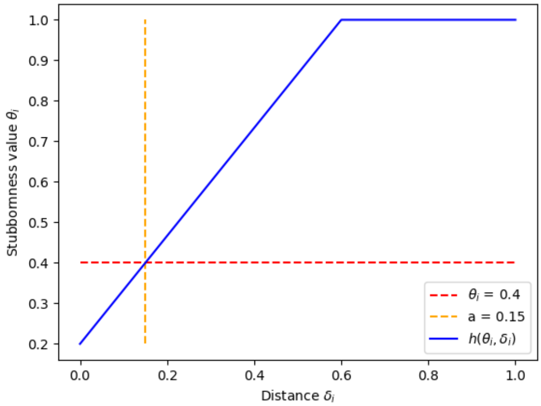


Figure 3.4 Illustration of the h updating function (3.5) with $\theta_i = 0.4$, $a = 0.15$ and $r = 0.5$. For $\delta_i < 0.15$, stubbornness will decrease, and for $\delta_i > 0.15$ it will increase.

Decreasing Stubbornness

Stubbornness-updating functions that fulfill the criteria of h, for which decreasing stubbornness is possible, can be obtained by replacing θ_i in the right hand side of (3.1) and (3.2) with some quantity in the interval $[0, \theta_i]$. However, this does not give direct control of the distance a at which stubbornness starts to increase. For this purpose, we provide a piecewise linear function.

$$h(\theta_i, \delta_i) = \begin{cases} \theta_i \text{ if } r = 1 \\ \theta_i r + \frac{\theta_i(1-r)}{a}\delta_i \text{ if } r < 1 \text{ and } \delta_i < \frac{a(1-\theta_i r)}{\theta_i(1-r)} \\ 1 \text{ otherwise} \end{cases} \tag{3.5}$$

In this setup, it is ensured that stubbornness is reduced for all $\delta_i < a$, increased for all $\delta_i > a$ and unchanged for $\delta_i = a$. The parameter $r \in [0, 1]$ defines the maximal relative stubbornness reduction possible. For example, $r = 0.5$ ensures that θ_i will be halved for an agent that has $\delta_i = 0$.

3.2 Convergence to Consensus

Non-Linear Stubbornness Updating Dynamics

To investigate the behaviour of the different variations of the model numerically, a large amount of parameter sweeps of $\mathcal{F}(\mathcal{G}, g)$ were performed, where $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ was fully connected and agents assigned equal weights. These were carried out for different values of b and c in $g(\theta_i, \delta_i)$ over the parameters that govern convergence to consensus in Theorem 2.2; maximal initial stubbornness $\theta_{max}(0)$ and maximal initial opinion distance $d(0)$. For each combination of $\theta_{max}(0)$ and $d(0)$, $\mathcal{F}(\mathcal{G}, g)$ was simulated 100-250 times, each time with a randomly generated initial opinion distribution. As our interests lie in guaranteeing consensus, if any one of these simulations did not reach consensus, the parameter combination was labeled as such.

This is also why the simulation was computed so many times. For parameter combinations that are close to the limit below which consensus is guaranteed, convergence to a non-consensus state seems to become increasingly rare, so to make sure the results were accurate, we chose to do a large number of tests. It would have been ideal to do even more, but we were limited by computational time, as one single parameter sweep took up to 8 hours. That this compromise had to be made is visible in the results. For example, if a result shows that initial opinion distributions for a given combination of $\theta_{max}(0)$ and $d(0)$ exist such that the model sometimes converges to a non-consensus state, this should also be the case for all higher $d(0)$. Simply take the exact same opinion distribution and scale it up such that $d(0)$ increases, but all relative distances between the opinions remain the same. The model dynamics should be exactly the same for this distribution, meaning that it also should not reach consensus. However, the results do in some cases seem to indicate that this actually were the case, see for example Figure 3.6. This is obviously not true, and could have been mitigated by increasing the number of tests further.

Not surprisingly, the results showed that for a stubbornness-updating function that is strictly larger than another, fewer combinations of $\theta_{max}(0)$ and $d(0)$ lead to $\mathcal{F}$ reaching consensus. Figures 3.5 and 3.6 show the outcomes of parts of a series of parameter sweeps for different c-values, $|\mathcal{V}| = 8$, $b = 0.5$. These convergence criteria and how they change with c does however not appear to be decided by the nonlinear features of g. In fact, the conditions for consensus for the functions l_1 and l_2 proved to be very similar to those for g_1 and g_2. This is illustrated plainly in Figures 3.7 and 3.8 that show some examples from a series of parameter sweeps where all parameter combinations were tested for both $\mathcal{F}(\mathcal{G}, l)$ and $\mathcal{F}(\mathcal{G}, g)$ using the same set of initial opinion distributions. The similarity is striking, especially given that the functions l were not chosen as the best approximations of g with regards to consensus convergence criteria. It thus seems likely that for every choice of g, there is a linear function that approximates it well enough that the difference in conditions that guarantee convergence is minimal.

Based on this, we also reason that an attempt to extend the Theorem 2.2 to non-linear stubbornness-updating functions would be somewhat superfluous. It is based on worst case scenarios and is very conservative [Ohlin et al., 2022], and we believe that a similarly conservative hypothetical theorem for non-linear functions could not meaningfully capture the nuanced differences in the consensus convergence criteria.

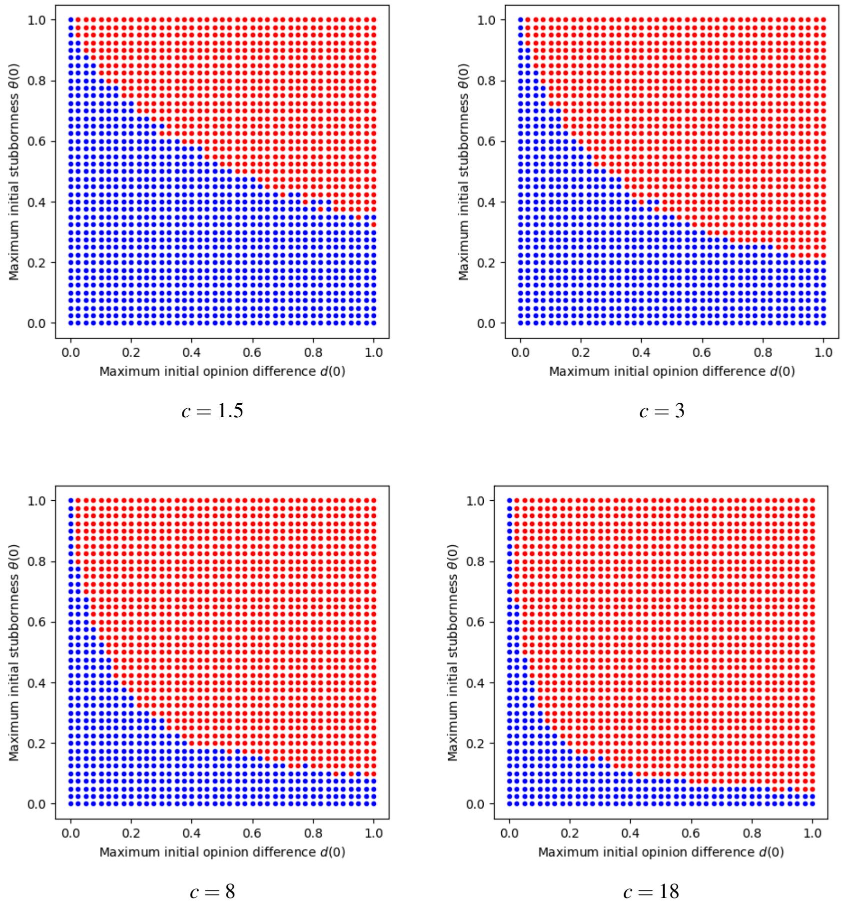
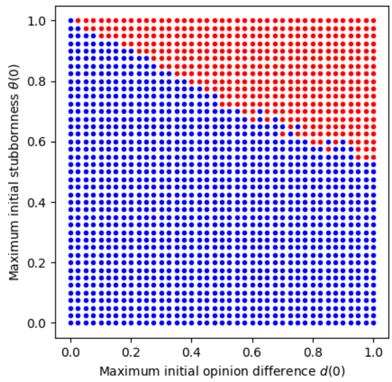
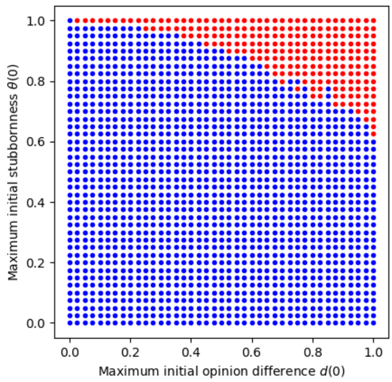


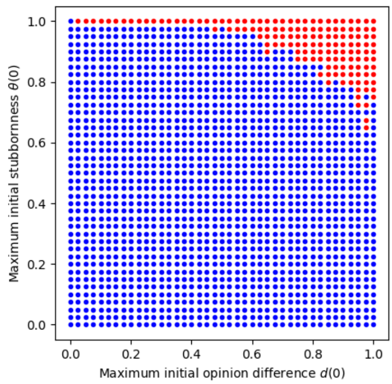
Figure 3.5 The results of a series of parameter sweeps of $\mathcal{F}(\mathcal{G}, g_1)$ over $d(0)$ and $\theta_{max}(0)$, with $\mathcal{G}$ containing 8 agents, where the model was simulated 100 times for each combination of parameters. A blue dot indicates that all simulations reached consensus, while a red dot signifies that at least one simulation converged to a non-consensus state.



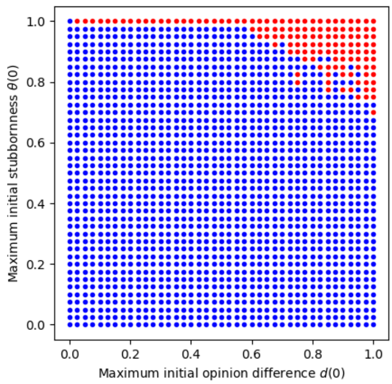
$c = 1.5$



$c = 3$



$c = 8$



$c = 18$

Figure 3.6 The results of a series of parameter sweeps of $\mathcal{F}(\mathcal{G}, g_2)$ over $d(0)$ and $\theta_{max}(0)$, with $\mathcal{G}$ containing 8 agents, where the model was simulated 250 times for each parameter combination. A blue dot indicates that all simulations reached consensus, while a red dot signifies that at least one simulation converged to a non-consensus state.

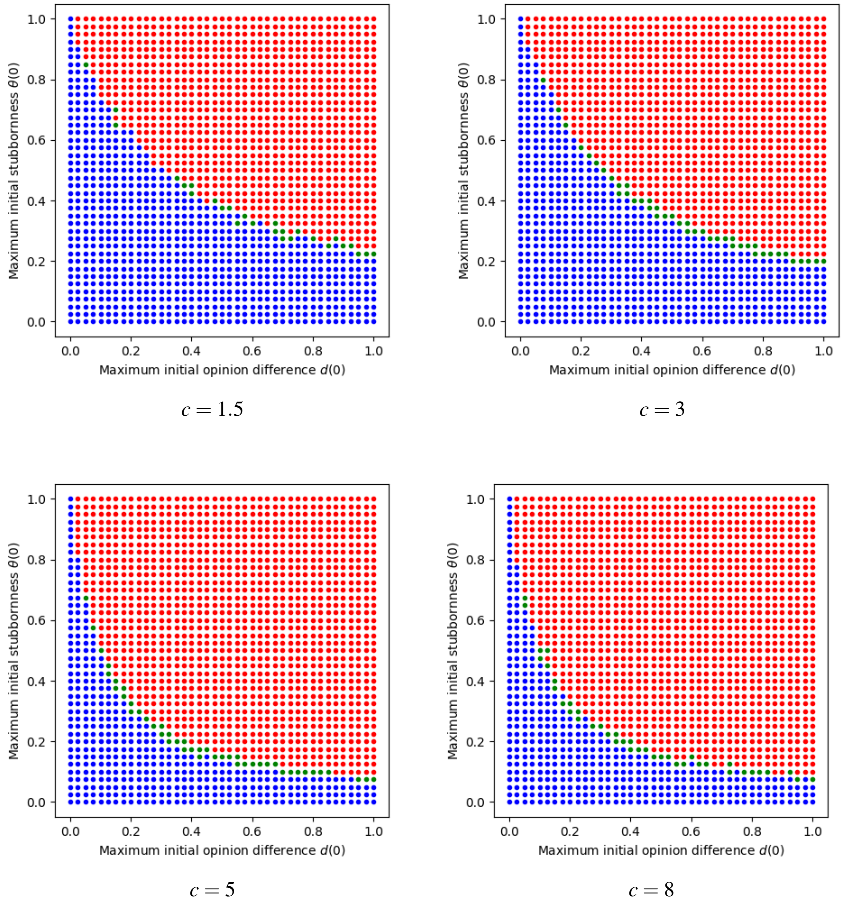


Figure 3.7 A series of parameter sweeps comparing the consensus convergence conditions of $\mathcal{F}(\mathcal{G}, g_1)$ and $\mathcal{F}(\mathcal{G}, l_1)$. The network contained 8 agents and the model was simulated 100 times for each combination of parameters. A blue dot indicates that all simulations reached consensus, and red one signifies that both g_1 and l_1 did not reach consensus at least once. Green dots represent all simulations reaching consensus only for g_1.

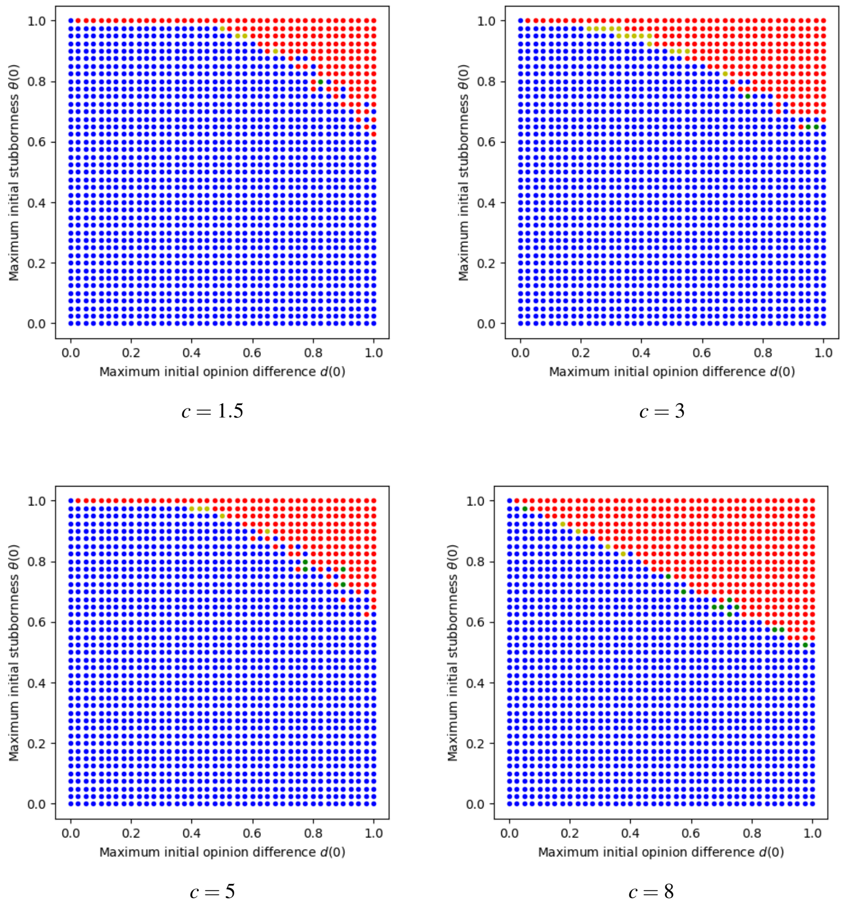
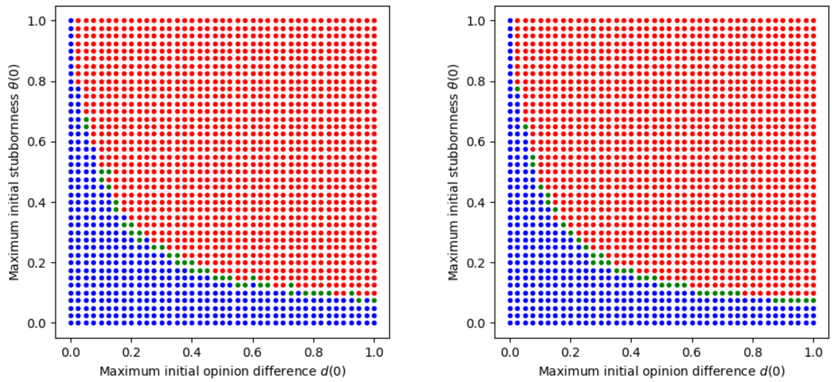


Figure 3.8 A series of parameter sweeps comparing the consensus convergence conditions of $\mathcal{F}(\mathcal{G}, g_2)$ and $\mathcal{F}(\mathcal{G}, l_2)$. The network contained 8 agents and the model was simulated 250 times for each combination of parameters. A blue dot indicates that all simulations reached consensus, and red one signifies that both g_2 and l_2 did not reach consensus at least once. Green dots represent all simulations reaching consensus only for g_2, and yellow ones the same for l_2.

Let us now try to develop an intuitive understanding of why the non-linear functions g can be so well approximated by linear functions with regard to consensus convergence criteria. Firstly, observe that large distances δ_i are more rare. When the initial opinion distribution is randomized, the initial mean opinion will tend to lie close to $y = 0.5$ rendering $\delta_i > 0.5$ impossible, which is why $b = 0.5$ was chosen for the tests shown in 3.5. Also, since the model has been defined in such a way that the first vote takes place only after the first issue, the maximal value of δ_i experienced by any agent will be reduced further. This means that the behaviour of g for small δ_i will tend to have a disproportionately large influence over how θ_i changes over time for a given agent. In simpler terms, the shape of g for larger δ_i tends to matter less than for smaller δ_i. While this can give some intuition for the behaviour of the model, it does not at all render the results uninteresting. Rather, when consensus is and is not reached is governed by the most extreme cases, and not by what tends to be common behaviour, so it is not obvious that intuition for standard behaviour carries over to intuition for consensus criteria.

We also remark that when consensus is not attained, this can not always be attributed to some agent with a high δ_i quickly jumping from a much lower stubbornness value to $\theta_i = 1$. Rather, it can be a question of whether the speed of convergence to consensus exceeds the rate at which stubbornness increases, as also observed in [Ohlin et al., 2022]. This means that for cases where consensus is almost attained, the opinion span may be very small as $s \to \infty$, meaning that a linear approximation of g_1 based on the derivative in $\delta_i = 0$ will become better and better.

Another observation is that the amount of agents in $\mathcal{G}$ does not seem to have a large impact on which simulations reach consensus. This is not the main issue being investigated, but still is worth briefly mentioning since all examples shown so far have 8 agents, which could be problematic if the number of agents were an important factor. Figure 3.9 is a comparison between networks of 8 and 50 agents for one of the parameter sweeps for g_1 and l_1, and the conditions are clearly very similar. This not very surprising - the conditions that guarantee convergence are as previously mentioned mainly governed by the most extreme agents. If two agents at either end of the opinion spectrum can become fully stubborn and stop $\mathcal{F}$ from reaching consensus, the amount of agents with opinions in between the extremes should be inconsequential.



8 agents

50 agents

Figure 3.9 Parameter sweeps for networks of 8 and 50 agents using the g_1 and l_1 update functions with $c = 5$ to update stubbornness between issues. As before, a blue dot indicates that all simulations reached consensus, and a red one signifies that a non-consensus state was reached at least once for both g_1 and l_1, while a green one means that the model reached consensus in every test only when g_1 was used.

While not exhaustive, our results point to that the difference that the stubbornness-updating functions g as compared with linear functions makes on the likelihood of the model converging to consensus is fairly small, under the restrictions that we have placed on them. The influence of the stubbornness-updating function, while naturally playing an important role in determining for which opinion distributions the model is able to reach consensus, appears to be mainly governed by its rough overall shape. While allowing for a broader range of stubbornness-updating functions does to some extent allow for greater control of model behaviour, the results suggest that there is little reason to choose a set of less analytically tractable functions over low-order piecewise linear approximations of them, when only considering requirements for consensus. We would also argue that there is little relevance to investigating the consensus convergence properties of more irregular stubbornness-updating functions than those discussed here, as there is no clear argument for, or empirical evidence known to the author, that suggests that they would constitute a realistic model.

Decreasing Stubbornness

Regarding stubbornness-updating functions with some possibility for decreasing stubbornness, specifically the function (3.5), some more clear-cut theoretical results could be determined, presented in the following theorem.

THEOREM 3.1
Let $\mathcal{F}(\mathcal{G},h)$ be a concatenated FJ-model where $\mathcal{G}$ is strongly connected and let $h(\theta_i,\delta_i)$ be the stubbornness-updating function (3.5) with $r<1$. Then, consensus is guaranteed for all $\theta_{max}(0)d(0)<2a$. □

Proof. Let y_α and y_β denote the maximum and minimum opinion values present in y respectively, with the stubbornness values of the corresponding agents denoted by $\theta_\alpha(s)$ and $\theta_\beta(s)$. Increasing the stubbornness of either agent is bound to slow the averaging dynamics, and decrease the rate of convergence to consensus. We therefore consider $\theta_\alpha(0)=\theta_\beta(0)=\theta_{max}(0)$, which may be equal to 1, and investigate how the quantity $\theta_\gamma(s)=\theta_\alpha(s)+\theta_\beta(s)$ changes. To stop the model from reaching consensus, both θ_α and θ_β must approach values that are arbitrarily close or equal to 1 faster than the speed at which the model approaches consensus, meaning that θ_γ must become arbitrarily close or equal to 2.

First, we note that if the median $\mu=\frac{y_\alpha-y_\beta}{2}$ at the end of an issue, both agents will decrease their stubbornness since $\delta_\alpha=\delta_\beta$ and $\delta_a+\delta_\beta<2a$ gives that $\delta_\alpha,\delta_\beta<a$. Regard now the case where $\mu>\frac{y_\alpha-y_\beta}{2}$. If $\theta_\alpha\geq\theta_\beta$, α will decrease their stubbornness more than β increases theirs. Suppose that $\theta_\alpha(s+1)=\theta_\alpha(s)-c$. There is then some distance $\delta_\alpha=a-p$ with $p>0$ that corresponds to this decrease. We must then also have that $\delta_\beta<a+p$. The decrease of θ_α is given by

$$\Delta\theta_\alpha=c=\theta_\alpha(s)-\theta_\alpha(s+1)=\theta_\alpha(s)r+\theta_\alpha(s)\frac{1-r}{a}(a-p)-\theta_\alpha(s)=\theta_\alpha(s)\frac{a(1-r)}{p}$$

and the increase of θ_β is found to be

$$\Delta\theta_\beta=\theta_\beta(s+1)-\theta_\beta(s)<\theta_\beta(s)r+\theta_\beta(s)\frac{1-r}{a}(a+p)-\theta_\beta(s)=\theta_\beta(s)\frac{a(1-r)}{p}$$

giving that $\Delta\theta_\beta<\Delta\theta_\alpha$. However, if instead $\theta_\alpha<\theta_\beta$, the opposite is true and α might decrease their stubbornness less than β increases theirs. However, if θ_β is large enough, this no longer holds. The way h is constructed, θ_β will at some point actually become equal to 1, meaning that it can no longer increase. If θ_β is to be arbitrarily close to 1, we might say that $\theta_\beta\gg r$ such that $\frac{1-\theta_\beta}{1-r}\to 0$. In this case $h(\theta_\beta,\delta_\beta)=1$ for all $\delta_\beta>a$. For a visual aid, see Figure 3.10.

This means that $\theta_\gamma(s)$ being equal to or arbitrarily close to 2 is not a stable state. If the model finds itself in this state, stubbornness will update such that one of the stubbornness values $\theta_\alpha(s+1)$ or $\theta_\beta(s+1)$ will have decreased more than the other has increased. Suppose that it is θ_β that increases, giving $\theta_\beta(s+1)>\theta_\alpha(s+1)$. Then, by the end of the next issue, if $\mu>\frac{y_\alpha-y_\beta}{2}$, this process will continue, bringing $\theta_\alpha(s+2)$ further from 1 and $\theta_\gamma(s+2)$ further from 2. If on the other hand $\mu<\frac{y_\alpha-y_\beta}{2}$, $\theta_\beta(s+2)$ will have decreased more than $\theta_\alpha(s+2)$ has increased, which also brings $\theta_\gamma(s+2)$ further from 2. This reduction must continue until it is no longer

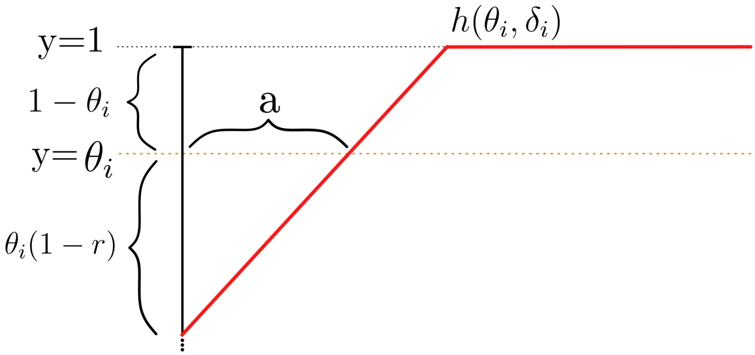


Figure 3.10 Illustration comparing the size of the quantities $1-\theta_i$ and $1-r$. When θ_i becomes arbitrarily close to 1, it will be set to 1 by h in the next stubbornness update.

true that $\frac{1-\theta_i}{1-r} \to 0$ for at least one of θ_α and θ_β, at which point it can no longer be considered to be arbitrarily close to 1. This means that for every issue where both θ_α and θ_β are arbitrarily close to 1, there exists a later issue where this is not the case. By this logic, there must be infinitely many issues during which the distance $y_\alpha - y_\beta$ decreases in a meaningful way, which guarantees consensus. $\square$

That we need $h(\theta_i, \delta_i) > 0$ stems from the convergence requirements laid out in [Wang et al., 2021] that stubbornness must be larger than 0 for at least one agent. If multiple agents have zero stubbornness and $\mathcal{G}$ has a cycle, the agents can become trapped in a loop. If for example $W = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, and $\theta_1 = \theta_2 = 0$, the agents will just keep swapping opinions with each other. Figure 3.11 below illustrates for which conditions the theorem guarantees consensus as compared to which conditions led to consensus for a numerical example.

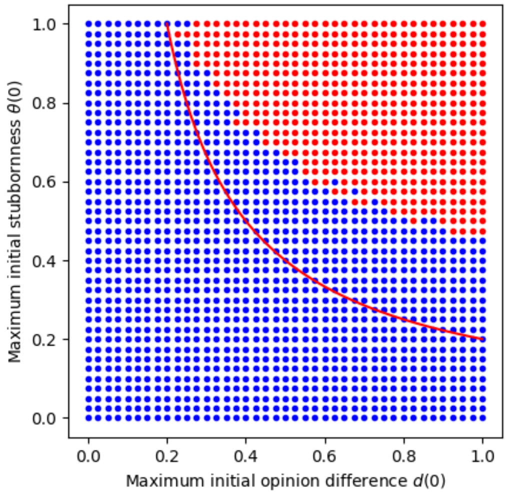


Figure 3.11 Results of a parameter sweep of $\mathcal{F}(\mathcal{G}, h)$ with $a = 0.1$ with 8 agents, where $\mathcal{G}$ was fully connected and the model was simulated 50 times for each parameter combination. The red line is $d(0)\theta_{max}(0) = 2a$, below which Theorem 3.1 guarantees convergence to consensus.

3.3 Characterization of the Non-Consensus State

In the previous section, we were only concerned with whether the models reached consensus, and all parameter combinations that made it possible for $\mathcal{F}$ to converge to a non-consensus state were treated as equal. In this section, we instead focus on this state. Why is this something that warrants further investigation? Describing an n-dimensional opinion distribution using only the binary concept of consensus, while often useful, results in a lot of information about the state of the system being lost. In fact, for every opinion distribution that is classified as having reached consensus, there are infinitely many unique distributions that are not. Some of these would undoubtedly be described as being closer to consensus than others, some more scattered and some more grouped. In our opinion, an analysis of the effect different stubbornness-updating functions have on the final opinion distribution of $\mathcal{F}$ is not complete without some analysis of these states.

The results from the numerical experiments described in Section 3.2 indicated that a more aggressive stubbornness-updating mechanism tends to lead to a more scattered final opinion distribution. This can be mathematically captured using the polarization concept as previously described in Section 2.4. For example, the g_1

functions tended to produce final opinion distributions that were more polarized for higher c-values. This is quite intuitive. A more aggressive stubbornness-updating mechanism means that agents will tend to have higher stubbornness values, rendering them less willing to change their opinion, which would in turn lead to more scattered final opinion distributions. It seems even more likely when considering the game-theoretic interpretation of the FJ-model dynamics. Recall that agents can at every time step be interpreted as striving to minimize the cost function

$$J_i(t) = (1-\theta_i)\sum_{j=1}^{n} w_{ij}(y_i(s,t)-y_j(s,t))^2 + \theta_i(y_i(s,t)-y_i(s,0))^2. \tag{3.6}$$

The first term in J_i is proportional to a partial sum of the NDI polarization index, so increasing stubbornness for the agent i essentially means that the agent prioritizes reducing NDI less over keeping the distance to its original opinion small, which one would presume should also lead to more polarized opinion distributions. A natural question to ask at this point is whether this pattern is more generally true?

The hypothesis that we are building towards is that increasing stubbornness more for any given agent will result in a more polarized final opinion distribution, or more formally that if $\mathcal{A}(\mathcal{G}, f_{\mathcal{A}})$ and $\mathcal{B}(\mathcal{G}, f_{\mathcal{B}})$ are two concatenated FJ-models such that $f_{\mathcal{B}}(\theta_i, \delta_i) \geq f_{\mathcal{A}}(\theta_i, \delta_i)$ for all y, $\mathcal{A}$ is more depolarizing than $\mathcal{B}$. However, our results show that this is in fact not true in the general case, as summarized in the proposition below.

One possible explanation for this could be that, when an agent uses (3.6) to determine the optimal position it should move its opinion to, it does this under the assumption that the other agents' opinions stay fixed [Proskurnikov and Tempo, 2017]. This means that for some opinion distributions, an agent can overshoot the opinion value that *would have* minimized J_i, and in such a circumstance, a higher θ_i value would result in a larger reduction of NDI for that time step. When considering this, the possibility that a strictly larger stubbornness-updating function can lead to increasing polarization seems slightly less unintuitive.

PROPOSITION 3.2
Let $\mathcal{G}$ be a graph of $N > 4$ opinionated agents and let $\mathcal{A}(\mathcal{G}, f_{\mathcal{A}})$, $\mathcal{B}(\mathcal{G}, f_{\mathcal{B}})$ be two concatenated FJ-models such that the stubbornness-updating function $f_{\mathcal{B}}(\theta_i, \delta_i) \geq f_{\mathcal{A}}(\theta_i, \delta_i)$ for all (θ_i, δ_i), with $f_{\mathcal{B}}(\theta_i, \delta_i) > f_{\mathcal{A}}(\theta_i, \delta_i)$ for at least some combination (θ_i, δ_i). Under these conditions it does not generally hold that $\mathcal{A}$ is more depolarizing than $\mathcal{B}$ under any of the polarization metrics GDI, NDI and V. $\square$

Proof. We provide a proof by counterexample. Assume that $\mathcal{G}$ is fully connected and has equal weights, which renders NDI and GDI proportional. It is thus sufficient to only consider the effect on the variance and GDI-metrics. The initial opinion distribution will be as follows: place $n > 2$ agents with opinion $y_n = 1$, one agent with opinion $y_1 = 0$ and finally place one agent with opinion $y_2 = d$ such that the mean opinion $\bar{y}(s,t) > d$ at $s = t = 0$. Note that it is not essential that the opinions span the whole interval $[0,1]$, but what is important is the relative distances between the agent's opinions. The setup could be scaled down arbitrarily, and the argument would still hold. Let us now assume that all agents are assigned the same stubbornness $\theta_1 = \theta_2 = \theta_n$ at the beginning of the first issue. Under that condition, (2.8) gives us that the relative distances between the agents will be preserved at the end of issue $s = 0$. The median will be placed at $\mu = y_n$, guaranteeing that for issue $s = 1$ we can have $\theta_1 > \theta_2 > \theta_n$. This means that during this issue, the relative distance between y_1 and y_2 will increase, and the distance between y_2 and y_n will decrease. Suppose now that after some issue, $\theta_1 \to 1$, while $1 > \theta_2 \geq \theta_n$. If d is chosen sufficiently small, and the starting stubbornness is sufficiently large, the mean opinion can at this point still be larger than y_2. As long as this is the case, the distance between y_2 and y_1 will continue to increase, with θ_2 being able to increase more than θ_n in between each issue. We can thus reach a point where also $\theta_2 \to 1$ while $\theta_n < 1$. Now, as long as the n-agents do not also become fully stubborn, which is easily ensured - the median will always be with this group since is consists of more than half the agents - y_n will converge to a point exactly in between the stationary values of y_1 and y_2, since this minimizes the cost function (3.6). If we denote $y_2 - y_n = y_n - y_1 = p$, we can calculate the polarization:

$$\Phi_V = \sum_{i \in \mathcal{V}} (y_i - \bar{y})^2 = 2p^2$$

$$\Phi_{GDI} = \sum_{i,j \in \mathcal{V}, i<j} (y_i - y_j)^2 = 2np^2 + 4p^2$$

which is proportional to p for both metrics. We also observe that the distance p is directly determined by the final y_2-value. If θ_2 had been given a higher value at even a single issue, θ_2 could have reached 1 earlier and the distance p would have been smaller. Observe also that this could have been achieved without changing the opinion at which θ_1 became arbitrarily close to 1, since y_2 must at some point come closer to y_n than y_1 ever did during the scenario. It would thus be sufficient to replace the stubbornness-updating function with one that returns higher values only for distances δ_i lower than those reached by the agent 1 to achieve this effect. $\square$

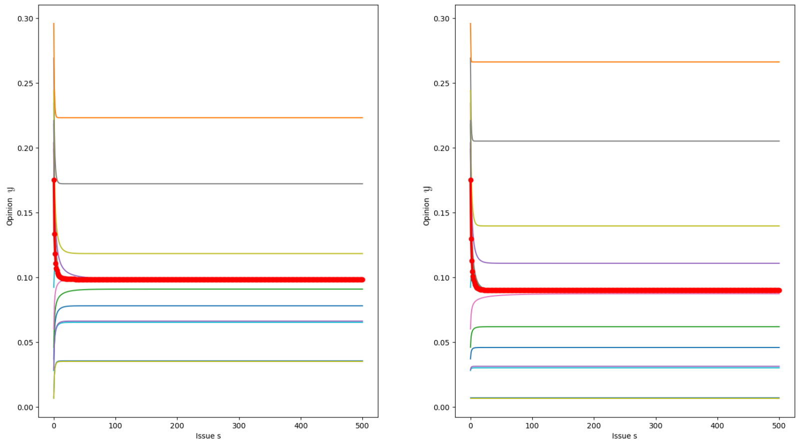
A similar argument to this could possibly be made for $N = 4$, but it is more complicated since the median is not fixed at y_n. Another possible route of investigation feasible for small graph sizes is to simply explicitly calculate the final opinion distribution for an FJ-model with a general initial distribution using (2.7), and search for a distribution for which depolarization is minimized for some $1 < \theta_i < 0$. Numerical experiments to this end show that such initial opinion distributions do exist for $N = 3$ and $N = 4$. However, this does not necessarily prove that the proposition holds also in these cases. Since the first vote is not taken until after the first issue, it is also necessary that some such distribution might first be attained through the model dynamics from an initial distribution in which all opinions are in the interval

$[0,1]$. We have not been able to find any such distribution, so for these network sizes the problem remains unsolved.

For the two-agent case it is however possible to prove that lower stubbornness is guaranteed to lead to a less polarized final opinion distribution. This is fairly easily concluded for a fully connected, equally weighed graph. In this case, the opinions will be approaching each other but are unable to surpass each other, meaning that increased stubbornness for any agent for any time step results in the distance between the two opinions reducing less for that time step. This also turns out to be true more generally, as long as W is not the identity matrix. A proof by explicit calculation using (2.7) can be found in Appendix A.1.

Summarizing the above discussion, we can conclude that for any concatenated FJ-model $\mathcal{F}(\mathcal{G}, f)$ with more than 4 agents, and any two stubbornness-updating functions f_1 and f_2, $f_1 \leq f_2$ does not guarantee that $\mathcal{F}(\mathcal{G}, f_1)$ is more depolarizing than $\mathcal{F}(\mathcal{G}, f_2)$ under the studied polarization metrics. For the two-agent case, the opposite is true, while the situation is not clear for other network sizes. Still, this is a fairly significant result, as it disproves the intuitive theory that higher stubbornness always leads to lower depolarization.

It should however be noted that even for networks with more than 4 agents, numerical results indicate that higher stubbornness resulting in lower polarization is a rare occurrence. For example, when comparing the depolarization of $\mathcal{F}(\mathcal{G}, g_1)$ and $\mathcal{F}(\mathcal{G}, l_1)$ where g_1 and l_1 are the functions with the same name from Sections 3.1, $\mathcal{F}(\mathcal{G}, l_1)$ was found to be more depolarizing for all randomly generated initial opinion distributions tested. An example of how these stubbornness-updating functions tended to lead to different final opinion distributions can be found in Figure 3.12. In other words, the intuitive hypothesis of larger stubbornness-updating functions yielding more polarized results is often true, albeit not generally so.



$\mathcal{F}(\mathcal{G}, g_1)$ $\mathcal{F}(\mathcal{G}, l_1)$

Figure 3.12 An example of one of the most commonly observed cases when comparing the depolarization of $\mathcal{F}(\mathcal{G}, g_1)$ and $\mathcal{F}(\mathcal{G}, l_1)$, the polarization of the final opinion vector is clearly greater when l_1 is used to update stubbornness. The median opinion is marked in red.

4

The Concatenated FJ-model with a Stubbornness-Maximizing Threshold

While the restrictions that are placed on the stubbornness-updating functions in Chapter 3 are justified, our choice to only study a few families of functions, while chosen fairly generally, led to some interesting model variants being left out. In this section, we highlight one such variant and present a version of the concatenated FJ-model with a special stubbornness-updating function that relies on instantly maximizing stubbornness, which has proved to have interesting properties. We begin by defining the problem and the model in section 4.1 using the same concatenated FJ-framework that we have used throughout. Then we outline the behaviour of the model in section 4.2 and present some theorems to that same effect. Furthermore, some illustrative numerical examples of the model are presented in section 4.3, and finally its practical applicability is discussed in section 4.4.

4.1 Problem Setup

Previously, when introducing stubbornness-updating functions with some possibility of stubbornness reduction in section 3.1, we introduced the notion of a *threshold* for δ_i above which agents are classified as *dissatisfied*. The model now to be presented is based on a similar principle, but but lets dissatisfied agents act differently in response to their dissatisfaction. Instead of agents above the dissatisfaction threshold becoming gradually more stubborn, and agents below it less so, we now consider the effects of letting all dissatisfied agents immediately become *fully* stubborn. This could be seen as the threshold corresponding to a ’red line’ in the negotiation process, that renders an agent completely unwilling to cooperate in any way if their δ_i-value exceeds it. Such a model would obviously reach consensus very rarely if this is the sole stubbornness-updating mechanic. However, we have yet to discuss what should happen when an agents opinion is below the threshold. If being above the threshold corresponds to the agent becoming unwilling to cooperate, an agent below the threshold should stay invested in the meeting and keep being willing to compromise. This could be modeled as a stubbornness-reduction for all or some of these agents as before, but we will be considering the case when agents below the threshold simply retain their original stubbornness. What makes this model different from the models with stubbornness-reducing functions previously discussed is that we propose that if the opinion of an agent that has previously become fully stubborn ends up below the threshold for some later issue, they should immediately *revert to their original stubbornness*. Consider the following chain of events: Suppose that some agent becomes dissatisfied and refuses to cooperate. The other agents still want to reach a consensus and are willing to compromise, so they will begin to approach the opinion of the dissatisfied agent. If the other agents are willing to make enough compromises, the stubborn agent will recognize their efforts and once again start to cooperate. This is in simple terms the intended behaviour of the setup.

In our terminology, the model can be defined as $\mathcal{F}(\mathcal{G}, f)$ where f is the function (4.1), being passed the initial stubbornness value θ_i^0 instead of $\theta_i(s)$ after every issue. We also reiterate that initial stubbornness values are chosen such that $0 \leq \theta_i^0 < 1$. No agent starts out as fully stubborn, and can thus only reach this state in response to disagreeing with a vote.

$$f(\theta_i, \delta_i) = \begin{cases} \theta_i \text{ if } \delta_i \leq \tau \\ 1 \text{ if } \delta_i > \tau \end{cases} \tag{4.1}$$

This is what we refer to as the model having a *stubbornness-maximizing threshold*, with the *threshold distance* denoted by τ. Also note that when discussing this model, we generally presume that the first vote is taken *before* the first issue ($s = 0$), meaning that $\theta_i(0) = f(\theta_i^0, \delta_i(y(0,0))$ which is not necessarily equal to θ_i^0. Since the only values that $\theta_i(s)$ can take are θ_i^0 and 1, the only difference this makes is that the dynamics of the model are permitted to play out over the whole interval $y_i \in [0,1]$, rather than an interval that has been shrunk over an issue with no fully stubborn agents. For every distribution that can be attained after a first issue of averaging with no fully stubborn agents, there is undoubtedly a 'scaled up' distribution that covers the whole interval $[0,1]$ which will exhibit exactly the same dynamics. In other words, there are no cases that are omitted by adopting this convention.

4.2 Model Properties

The behaviour of the model will arguably be most interesting when the initial opinion distribution is such that agents on both sides of the median find themselves dissatisfied. How can the agents in between then act to make the system reach consensus? One effective strategy could be to first compromise with one side, persuading the stubborn agents there to return to their original stubbornness values, to then move collectively to persuade the other side. However, this will only be guaranteed to happen for a select few network structures, like if the agents in the middle form a strongly connected subgraph whose only connections to the rest of the graph go through agents on one side of the median. Whether consensus remains possible is governed by where the Nash equilibrium of the median agent for the current issue is located. If it is located such that any one fully stubborn agent reverts to their original stubbornness before the next issue, the model will continue to evolve and consensus to remain within reach.

If the graph is instead fully connected and has equal weights, all agents that are not fully stubborn will move their opinions towards a common equilibrium point equal to the mean of all the fully stubborn agents' opinions, as if trying to make all of them happy at once. The question then becomes whether this point happens to lie within τ of any of the fully stubborn agents or not. This can have a sort of 'tipping point' effect. A previously fully stubborn agent that rejoins the averaging dynamics can radically change which way agents want to move, which paves the way for some very interesting model behaviour.

Suppose for instance that there is more fully stubborn agents with opinions 'above' the median than 'below' it. Then, the equilibrium point of the median agent is more likely to lie at a higher opinion value. However, moving to this opinion value might cause some of the fully stubborn agents on the top of the spectrum to once again become mobile before the next issue, potentially altering the distribution such that there are now more fully stubborn agents *below* the median. If this is possible, the model could potentially oscillate between these two states. Can such oscillations occur? For how long could such an oscillation persist? Can agents keep moving back and forth between the satisfied and dissatisfied states, such that this pattern could be sustained indefinitely? These are some of the questions we will explore in this section. As it turns out, the model does indeed have some oscillatory properties, explored numerically in section 4.3.

We begin however by presenting some theoretical results, starting with some conditions for consensus. This is a very intricate problem, and we are not able to determine all initial opinion distributions that will lead to consensus for a given network structure. However, we can provide some sufficient conditions that guarantee consensus will be reached.

THEOREM 4.1
Let $\mathcal{G}$ be a strongly connected graph with $n \geq 2$ opinionated agents. For a concatenated FJ-model $\mathcal{F}(\mathcal{G}, f)$ with threshold stubbornness such that f is the function (4.1), consensus is guaranteed:

- For any initial opinion distribution $y(0)$ if and only if the threshold $\tau \geq 0.5$.

- For any combination of y and τ that fulfill $\tau \geq \frac{d(0)}{2}$ where $d(s)$ is the maximal opinion difference $y_{max} - y_{min}$ in $y(s,0)$. □

Proof. We begin by proving that $\tau \geq \frac{d(0)}{2}$ guarantees consensus. A compelling argument for that consensus will be reached when there are no fully stubborn agents is that during any issue, the distance between the maximum and minimum opinions $y_{max} - y_{min}$ must decrease. We can employ the following reasoning: Since $\mathcal{G}$ is strongly connected, there exists at least one agent with opinion $y_i(s,0) < y_{max}(s,0)$ that exerts an influence over the y_{max} agent. Furthermore, there are by definition no agents with opinion $y_i(s,0) > y_{max}(s,0)$, making it impossible for any other opinion to reach $y_i(s,t) > y_{max}(s,0)$ during s. Also, $y_i(s,t) = y_{max}(s,0)$ for some t is only possible if i only is influenced by the y_{max}-agent and $\theta_i = 0$, in which case $y_i(s,1) = y_{max}(s,0)$. But then, $y_i(s,2) < y_{max}(s,0)$ since the y_{max}-agent has decreased its opinion by $t = 2$. This gives that $y_{max}(s,0) < y_{max}(s-1,0)$. Note that $y_{max}(s-1,0)$ and $y_{max}(s,0)$ do not necessarily belong to the same agent, rather $y_{max}(s,t)$ and $y_{min}(s,t)$ are simply to be interpreted as the maximum and minimum opinions present at the specified issue s and time t.

We can also make a corresponding argument for y_{min}, giving $y_{min}(s,0) > y_{max}(s-1,0)$, and it clearly holds that $y_{max}(s,0) - y_{min}(s,0) < y_{max}(s-1,0) - y_{min}(s-1,0)$. This argument also holds if either the agent with y_{max} or y_{min} have $\theta = 1$, but not both. For example, if the agent with the opinion y_{min} has $\theta = 1$ and the agent with y_{max} has $\theta < 1$, we have that $y_{max}(s,0) < y_{max}(s-1,0)$ and $y_{min}(s,0) = y_{min}(s-1,0)$, which still means that $y_{max}(s,0) - y_{min}(s,0) < y_{max}(s-1,0) - y_{min}(s-1,0)$.

Now, if we consider the stubbornness-updating mechanism and reintroduce the possibility of fully stubborn agents, a key observation is that if $\tau \geq \frac{d(0)}{2}$, fully stubborn agents can only exist on one side of the opinion spectrum at a time. For example, if $d(0) = 1$ and the median at the beginning of issue s is initialized at $\mu > 0.5$, the agent with y_{min} can be fully stubborn, but the agent with y_{max} cannot, and $y_{max} - y_{min}$ will decrease over the issue. If the opinions were to change in such a way that $\mu < 0.5$ at the beginning of issue $s+1$, the opposite is true and $y_{max} - y_{min}$ will once again decrease. This guarantees convergence to consensus.

To prove that situations exist where $\tau < \frac{d(0)}{2}$ and consensus is impossible, consider the following example: Let $\tau = \frac{d(0)}{2} - \varepsilon$ for some small $\varepsilon > 0$ and let the graph contain n nodes. If n is even, initialize $n/2$ agents with opinion $y_i = 0$ and the other half with $y_i = d(0)$. If n is odd, first place one agent with $y_i = \frac{d(0)}{2}$ and then distribute the remaining $n-1$ agents as before. This guarantees that both the initial mean opinion and median $\mu_0 = \frac{d(0)}{2}$, while it holds for all agents at either end of the opinion spectrum that $|y_i - \mu_0| = \frac{d(0)}{2} > \tau = \frac{d(0)}{2} - \varepsilon$, even as $\varepsilon \to 0$. This completes the proof. □

COROLLARY 4.2

If the first vote does not happen until after the first issue and $\mathcal{G}$ is a complete graph with equal weights, a lower τ that guarantees consensus is $\tau \geq \frac{d(1)}{2} = \frac{d(0)\theta^0_{max}}{2}$, where θ^0_{max} is the maximum stubbornness value any agent $\mathcal{V}$ is initialized with. □

Proof. Delaying the vote until after the first issue effectively shrinks the range of opinions that an agent can possibly hold. Theorem 2.1 from Section 2.2 gives that

$$d(1) = \theta^0_{max} d(0) \qquad \square$$

as long as there are no full stubborn agents during the first issue, which is exactly the case if we delay the vote until after issue $s = 0$. Now, exactly the same reasoning as for the main theorem can be applied on this reduced interval, giving $\tau \geq \frac{\theta^0_{max} d(0)}{2}$.

Now, we will present two results pertaining to when it can be guaranteed that an agent with an opinion that was within the threshold immediately after the most recent vote will retain their opinion within the threshold also when the next vote is taken. This problem proves surprisingly complicated, even under the condition that the network be fully connected and equal-weighed. The obvious reason that could make agents increase their distance to the median opinion over an issue is simply that of varying stubbornness. If the agent with the median opinion has lower stubbornness than an agent with an opinion just on the verge of the threshold, the median agent could move their opinion such that the other agent is unable to 'catch up', leaving their opinion over the threshold during the next vote. This is however not the only reason we must concern ourselves with. We must also consider that the mean opinion might *change agent* (or agent pair, if there are an even number of agents). For example, consider a situation where the median opinion $\mu(s,0)$ is close to 0, the mean opinion $\bar{y}(s,0) \approx \mu(s,0) + \tau$ and there are two agents α and β with opinions close to the threshold, $y_\alpha(s,0) < \mu(s,0) + \tau < y_\beta(s,0)$ such that β is fully stubborn while α is not. During the upcoming issue, there is likely to be a net movement of opinions towards higher opinion values. Since $\bar{y}(s,0) \approx \mu + \tau$, there must be fully stubborn agents with opinions greater than $\mu(s,0) + \tau$ at the start of the issue. These will not move their opinions at all, so the mean value $\bar{y}(s,t)$ will have to adjust in a way that keeps the other agents moving their opinion upwards at every time step, until an equilibrium is found. In this scenario, it is possible that both α and the agent that held the median at the beginning of the issue surpass the opinion of β, which would make y_β the new median opinion, see figure 4.1. That means than when considering whether $y_\alpha(s)$ stays within the interval $[\mu(s+1,0) - \tau, \mu(s+1,0) + \tau]$, $\mu(s+1,0)$ might be anywhere between $\mu(s,0) + \tau$ and $\bar{y}(s+1,0)$.

This internal order of the opinion vector, and in turn the agent with the median opinion, can change as an effect of two things:

1. An agent that is not fully stubborn having their opinion surpass the opinion of some fully stubborn agent. This is the case that is described above.

2. Two agents α and β with $\theta_\alpha \neq \theta_\beta$ that are both not fully stubborn moving their opinions in the same direction. This could happen if for example $\bar{y}(s,0) > y_\alpha(s,0) > y_\beta(s,0)$ while $\theta_\alpha > \theta_\beta$.

Note however that it is not possible for two agents that are moving their opinions towards each other to cross them, since none of them can pass the mean opinion when $\mathcal{G}$ has equal weights.

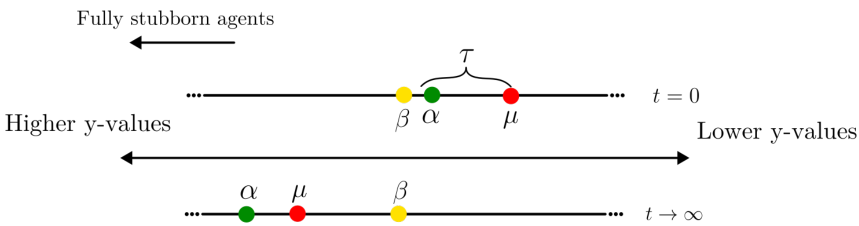


Figure 4.1 Illustration of how the internal order of opinions might change over the course of an issue in the prescence of fully stubborn agents.

To handle these issues, we approach the problem from two distinct directions, in each making assumptions that address one of the ways that the agents might increase their distance to the median opinion. In our first result, we only consider choices of $\tau \geq \frac{d(0)}{2}$, which makes the issue of the median opinion moving between agents much less complicated to handle, and let the maximum and minimum stubbornness values act as the main variable. In the second, we instead assume that all agents that are not fully stubborn share the same stubbornness value, which effectively eliminates the problem of agents outrunning one another. The results are provided below in the form of two theorems.

THEOREM 4.3
Let $\mathcal{G}$ be a complete graph with equal weights containing an odd number $n > 2$ of opinionated agents that are initialized with stubbornness values $\theta_i^0 \in [\theta_{min}, \theta_{max}]$. For a concatenated FJ-model using (4.1) with $\tau \geq \frac{d(0)}{2}$ to update stubbornness between issues, applied to $\mathcal{G}$, it holds for any agent i that $\delta_i(s) \leq \tau \Rightarrow \delta_i(s+l) \leq \tau$, $l \in \mathbb{Z}^+$ if $\theta_{max} - \frac{1}{2} \leq \frac{\theta_{min}}{2}$. In other words, under this condition, an agent who has at one point been within τ of the voting result will remain in that range for all subsequent issues. □

Proof. Initially we observe that since there is no incentive for any agent to decrease their opinion value to a point below the initial minimum or above the initial maximum, all opinions will remain in an interval with width $d(0)$ for all (s,t). This means that $d(0)$ essentially only acts as a scale factor for the whole problem, and we will from now on argue as if $d(0) = 1$.

Let us now show that the median changing agents is not an issue under our conditions. First, for fully stubborn agents, observe that for an agent γ to have been made fully stubborn, $|y_\gamma(s,0) - \mu(s,0)| > \tau$. This means that γ and μ must lie in different halves of the opinion spectrum, $y_\gamma(s,0) < 0.5 \Leftrightarrow \mu(s,0) > 0.5$ and vice versa. That in turn means that if the median agent α moves their opinion past y_γ, making γ the new median agent, all opinion values that are within τ of y_α are also within τ of y_γ. Say for example that the initial opinion placement $y_\alpha(s,0) = \mu(s,0) < \frac{1}{2}$ and $y_\gamma(s,0) > \frac{1}{2}$ shifts such that $y_\alpha(s+1,0) > y_\gamma(s+1,0) = \mu(s+1,0) > \frac{1}{2}$. Then, all potential opinion values larger than μ are within τ regardless of whether the median is with α or γ, and more potential values smaller than μ are within τ of it if $\mu = y_\gamma$ than if $\mu = y_\alpha$. We do thus not need to consider the presence of a potential agent γ since it makes conditions less restrictive. Now, considering the problem of an agent 'overtaking' the median, which is easily handled. There is no way for such an agent to overtake the median agent, when considering the case where the median agent has the minimal stubbornness value, so simply considering this case becomes a straightforward solution to the problem.

Now, consider two agents, α and β with $y_\alpha(s,0) > y_\beta(s,0)$, either of which could be the median at the start of issue s. It will undoubtedly be the easiest for them to move in a way that ensures that $|y_\alpha(s+1,0) - y_\beta(s+1,0)| > \tau$ if we let $y_\alpha(s,0) - y_\beta(s,0) = \tau$, so we assume that $y_\alpha(s,0) = r + \tau$, $y_\beta(s,0) = r$ for some $r \in [0, 0.5]$. The situation when $r > 0.5$ is already covered by symmetry. The opinion values at the end of the issue can be explicitly calculated using (2.8) from section 2.2:

$$y_\alpha(s+1,0) = (1-\theta_{min})\bar{y}(s+1,0) + \theta_{min}(r+\tau)$$

$$y_\beta(s+1,0) = (1-\theta_{max})\bar{y}(s+1,0) + \theta_{max} r$$

Since $y_\alpha(s,0) > y_\beta(s,0)$ and $\theta_\alpha < \theta_\beta$, we know that $y_\alpha(s+1,0) > y_\beta(s+1,0)$ and can calculate the opinion distance between the agents at issue $s+1$ as

$$y_\alpha(s+1,0) - y_\beta(s+1,0) = (\theta_{max} - \theta_{min})\bar{y}(s+1,0) + r(\theta_{min} - \theta_{max}) + \theta_{min}\tau \geq \theta_{max} - \theta_{min}(1-\tau).$$

We want to guarantee that this distance does not exceed $\tau \geq \frac{1}{2}$, giving that

$$\theta_{max} - \theta_{min}(1-\tau) \leq \tau \Leftrightarrow \theta_{max} - \tau \leq \theta_{min}(1-\tau)$$

which gives the most restrictive conditions on stubbornness when $\tau = \frac{1}{2}$, giving the final result that

$$\theta_{max} - \frac{1}{2} \leq \frac{\theta_{min}}{2}. \qquad \square$$

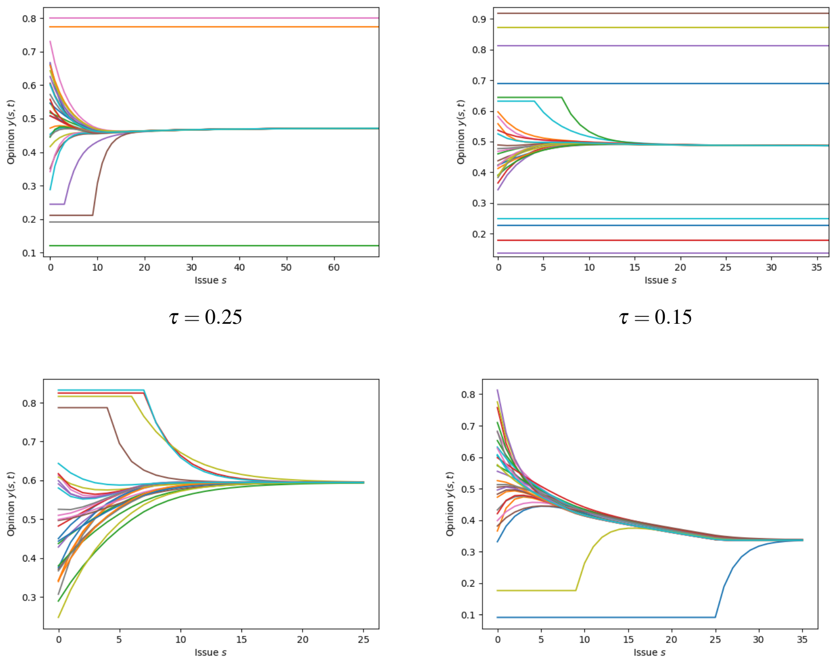
THEOREM 4.4
Let $\mathcal{G}$ be a complete graph with equal weights containing $n \geq 2$ opinionated agents who all share the same default stubbornness $\theta_i = a < 1$. For a concatenated FJ-model using (4.1) to update stubbornness between issues, applied to $\mathcal{G}$, it holds for any agent i that $\delta_i(s) \leq \tau \Rightarrow \delta_i(s+l) \leq \tau$, $l \in \mathbb{Z}^+$ if $\tau > \frac{d(0)}{6}$. □

The full proof of this theorem is quite extensive, and has for that reason been relegated to the appendix. In short, the principle is to first prove that two agents that are not fully stubborn can never increase their distance between their opinions. This is fairly straightforward. Then, the majority of the proof is dedicated to showing that the $\tau > \frac{d(0)}{6}$ is a sufficient bound in all potential situations where the median can switch agents.

4.3 Numerical Examples

The model has been examined numerically, focusing on the case where $\mathcal{G}$ is fully connected and has equal weights. Different values of τ and θ_i^0 were used, along with opinions sampled from different combinations of normal distributions as well as from uniform distributions. Figure 4.2 below shows some examples of the cases that were most commonly observed and easily achieved. The wider the threshold τ, the more often consensus tends to be observed, but non-consensus states are still commonly observed throughout.
This can be explained by that there are many ways to place a selection of opinions close to the extremes such that the mean of that selection is placed close to the middle. The mean agent will, in broad terms, sweep up fully stubborn agents until the point of equilibrium is longer than a threshold's length from the nearest fully stubborn agent's opinion. How many of the agents that are made fully stubborn after the first vote that will end up being 'swept up' by the median of course depends on the initial opinion distribution.

Occasionally, some oscillation-like behaviour is observed, examples of which are provided in Figure 4.3. These oscillations became more pronounced when destabilizing the initial opinion distribution, see Figure 4.4 for two examples. The movement of the cluster of opinions around the median switching direction in this way was in our numerical experiments always associated with a new previously fully stubborn agent joining their opinion to the median. That these two events might be related is understandable when considering that accession of a fully stubborn agent to the median can cause the mean of the remaining fully stubborn agents to move to the other end of the opinion spectrum. However, this association also makes it appear unlikely that oscillation might continue indefinitely. It is completely clear that indefinite oscillation would be impossible if both of these statements were to



$\tau = 0.25$ $\tau = 0.15$

$\tau = 0.25$ $\tau = 0.25$

Figure 4.2 Examples of common simulation outcomes. The graph contains 25 agents in all cases, and initial opinions were sampled from a normal distribution with $\sigma = \frac{1}{6}$.

be true.

1. An oscillation is always associated with an agent ceasing to be fully stubborn.
2. The conditions for any of the theorems 4.3 or 4.4 are fulfilled.

The model can then only oscillate until all agents have joined their opinion to the median, whereupon consensus must be reached, and oscillation stop. That it is impossible under the conditions of Theorem 4.3 is of course obvious since $\tau \geq \frac{d(0)}{2}$ also guarantees consensus. However, we can not guarantee the truth of the first statement and reiterate that it is only a hypothesis based in non-extensive numerical experiments.

Another point of note that reduces the likelihood of indefinite oscillations is that the theorems regarding the possibility of leaving the median seem to be quite con-

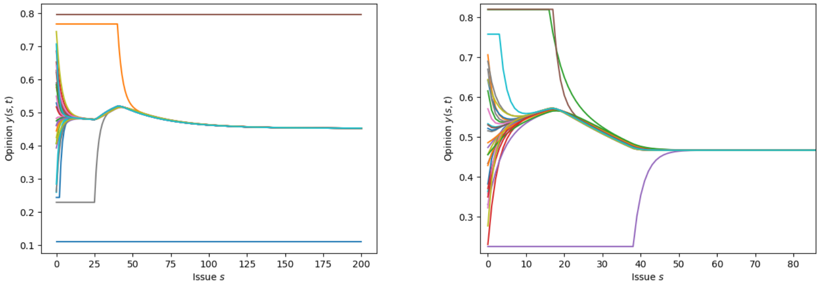


Figure 4.3 Examples of simulations were the median opinion showed small signs of oscillation. Both examples use 25 agents with initial θ-values sampled from a uniform distribution between 0 and 0.6. Initial opinions were sampled from a normal distribution with $\sigma = \frac{1}{6}$, and τ is equal to 0.25 in both examples.

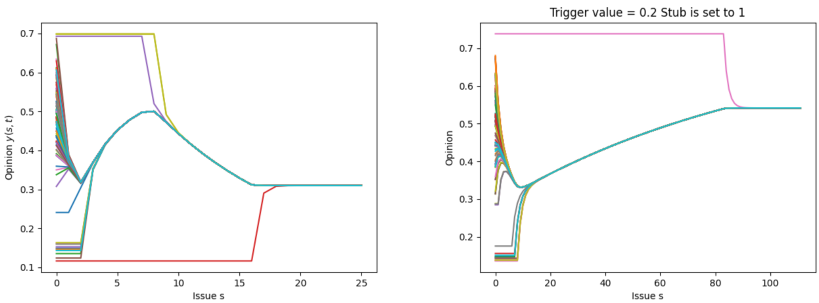


Figure 4.4 Examples of simulations were the median opinion showed some signs of oscillation. These examples use 100 agents, all with $\theta_i^0 = 0.1$. 90% of opinions were sampled from a normal distribution with $\sigma = 0.1$ centered on $y = 0.5$, while the remaining 10% were sampled from a distribution with $\sigma = 0.0167$ centered on $y = 0.1$. τ is equal to 0.2 in both examples.

servative. Agents that are not fully stubborn tend to converge quickly and 'move as one', rather than staying spread out across the available interval, which can be seen in both Figures 4.2 and 4.3, making it all the more difficult for any agent to leave the median. This is however not that surprising when considering that the worst case opinion distributions analyzed in the proofs of both theorems 4.3 and 4.4 consist of two tightly bound opinion clusters far apart, with only one or a few opinions in

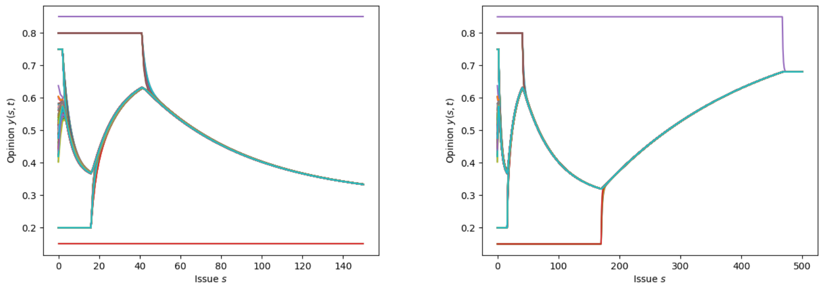


Figure 4.5 Simulation of an opinion distribution crafted especially to create oscillations, with initial stubbornness randomized between 0.3 and 0.6 and $\tau = 0.17$. Since the oscillations become slower and slower, the leftmost figure focuses on the first 150 issues such that all oscillations are visible.

between. These are some of the most high-cost situations possible under the game-theoretic interpretation of the model, since the first term of (3.6) is maximized for a given opinion interval when opinions are evenly spread out over the extremes. This makes these scenarios quickly dispersed, and very unlikely to occur spontaneously under the dynamics of the model, even if they are possible starting configurations that must be taken into account in a general proof.

Some brief experimentation was conducted with the aim of provoking more long-lasting oscillations than those that occurred spontaneously with randomly generated opinion distributions. To ensure that the distribution remained unbalanced for as long as possible, we deliberately placed agents in multiple tight clusters of decreasing size on alternating sides of the median, meaning that each oscillation becomes slower than the previous one, since the weight per agent is constant. The result of this experiment can be seen in Figure 4.5. In total, this network has 300 agents, 115 of which were sampled from a normal distribution with $\sigma = 0.05$ centered on $y = 0.5$. As for the remaining ones, 120 agents had $y_i(0) = 0.75$, 44 had 0.2, 16 had 0.8, 4 had 0.15 and 1 agent had $y_i(0) = 0.85$. This pattern could likely be repeated if the number of agents is expanded, but due to time constraints we were unfortunately unable to explore this further. We instead remark that how the initial opinion distribution and threshold size should be chosen to maximize the amount of oscillations for a given number of agents, and whether infinite oscillations for a finite number of agents are possible, are both intriguing problems for future research.

4.4 Realism

A point of critique against this model might be a lack of realism. Since the 'anchoring point' at the initial opinion is relocated at the beginning of every new issue, it does not constitute a very firm anchor in the presence of fully stubborn agents. Because of this mechanic, fully stubborn agents on one side of the opinion distribution will always attract all other agents towards them, making those agents move their opinions very far from their initial values at the start of the first issue. This raises the question whether such a mechanic makes for a realistic model. Is it reasonable that an agent with an initial opinion close to $y = 1$ can consider moving their opinion to the polar opposite of the opinion spectrum, just to move a fully stubborn agent with an opinion close to $y = 0$ out of the dissatisfied state? This is a valid point to make, but we argue that this is at least somewhat mitigated under the right interpretation of the model. Here, the terminology of *internal* and *expressed* opinion could be useful. An agent that changes their opinion greatly to achieve consensus must not be interpreted as them actually being convinced to change their fundamental beliefs, but prioritizing consensus over their internal opinion, and changing their expressed opinion to achieve that goal. Also, the result of the concatenated FJ-model might not always have to be measured in terms of the final opinion distribution. Consider for example a number of political actors with different ideologies trying to reach an agreement of cooperation to be able to work together towards a joint goal. Such an agreement might cover a wide array of issues, and for each agent to be satisfied with the overall agreement, they will need to be satisfied with the result of some portion of the issues, while willing to compromise on others. In other words, they require that the decision made on some issues is sufficiently close to their political position, and are then willing to compromise on other issues such that an agreement can be reached. This is precisely what the stubbornness-maximizing threshold achieves by letting agents remain fully stubborn until the result of the vote comes close enough. However, we would like to clarify this is an argument largely based on intuition that is yet to be verified. Still, this example highlights areas to which the model could be of potential relevance, even if the author also believes that it is fascinating in its own right as a purely mathematical construct.

5

Summary and Future Work

In this thesis, we have investigated how dynamic stubbornness affects the properties of the concatenated FJ-model, with a focus on what conditions guarantee convergence to consensus, as well as on depolarization. Consensus convergence criteria were investigated numerically for a fairly wide family of functions. The results were not general due to the of the numerical nature of the experiments, but nonetheless indicated that which conditions guarantee convergence to consensus is mainly decided by the overall outline of the stubbornness-updating functions, and not by their finer details. This was concluded as the consensus convergence properties of the investigated functions were approximated well by piecewise linear functions of no more than 3 pieces. For a stubbornness-updating function where stubbornness can decrease for some agents while increasing for others, we also presented some conditions that guarantee convergence to consensus. Regarding polarization, we observed numerically that when comparing two versions of the model with stubbornness dynamics such that a new stubbornness value for any agent will always become higher (or equal) in one than in the other for the same opinion distribution, the former tended to be less depolarizing than the latter. However, we also proved that this is not generally true, and that there are initial opinion distributions for which a strictly larger stubbornness-updating function will produce less polarized results.

Furthermore, we presented the concatenated FJ-model with a stubbornness maximizing threshold as an example of a stubbornness dynamic with interesting properties. These properties were investigated both numerically and theoretically, resulting in multiple theorems focusing on different model features, namely conditions that guarantee convergence to consensus as well as conditions under which the opinion of an arbitrary agent cannot come to exceed the threshold distance to the median opinion after once having been sufficiently close to it. It was also found that it is possible for the median opinion to oscillate to some degree under the right conditions, although the maximum possible extent of these oscillations, along with precisely defined conditions under which they can occur, remain undetermined. This renders this particular version of the model a possible and somewhat intriguing avenue of future research, even if some criticism can be directed at it regarding its practical applicability.

We would also like to point out that this thesis takes a primarily theoretical and mathematical perspective on all models discussed and that investigating the empirical validity of them is a clear and important direction of future work regarding dynamic stubbornness in concatenated FJ. While the models do stem from a set of fairly simple and intuitive assumptions, it is not obvious without validation that these assumptions are correct, or that they can meaningfully capture the complex web of interactions that constitute real-world social networks. Inquiring into this by attempting to empirically validate the models, or potentially by integrating them with research from other relevant fields such as Psychology or Sociology would be highly interesting. While the models may be considered fascinating in their own right, the ambition of opinion dynamical modeling is of course to do more than deliver curious mathematical concepts. Also, while our results do point to that the detailed features of the stubbornness-updating function might not be paramount in determining some parts of the model behaviour, results from less mathematically oriented research in this area could still help narrow down which types of functions ought to be used.

Bibliography

Bernardo, C., L. Wang, F. Vasca, Y. Hong, G. Shi, and C. Altafini (2021). "Achieving consensus in multilateral international negotiations: the case study of the 2015 paris agreement on climate change". *Science Advances* **7**:51, eabg8068. DOI: 10.1126/sciadv.abg8068.

Biondi, E., C. Boldrini, A. Passarella, and M. Conti (2023). "Dynamics of opinion polarization". *IEEE Transactions on Systems, Man, and Cybernetics: Systems* **PP**, pp. 1–12. DOI: 10.1109/TSMC.2023.3268758.

Black, D. (1948). "On the rationale of group decision-making". *Journal of Political Economy* **56**:1, pp. 23–34. ISSN: 00223808, 1537534X. (Visited on 2024-05-15).

DeGroot, M. H. (1974). "Reaching a consensus". *Journal of the American Statistical Association* **69**:345, pp. 118–121. ISSN: 01621459. URL: http://www.jstor.org/stable/2285509 (visited on 2024-04-23).

French, J. R. P. J. (1956). "A formal theory of social power." *Psychological Review* **63**:3, pp. 181–194. ISSN: 0033-295X.

Friedkin, N. and E. Johnsen (1999). "Social influence networks and opinion change". *Advances in Group Processes* **16**.

Friedkin, N. E. and E. C. Johnsen (1990). "Social influence and opinions". *The Journal of Mathematical Sociology* **15**:3-4, pp. 193–206. DOI: 10.1080/0022250X.1990.9990069.

Godsil, C. D. and G. Royle (2001). *Algebraic graph theory.* Graduate texts in mathematics: 207. Springer. ISBN: 0387952209.

Kubin, E. and C. von Sikorski (2021). "The role of (social) media in political polarization: a systematic review". *Annals of the International Communication Association* **45**:3, pp. 188–206. DOI: 10.1080/23808985.2021.1976070.

Matakos, A., E. Terzi, and P. Tsaparas (2017). "Measuring and moderating opinion polarization in social networks". *Data Mining and Knowledge Discovery* **31**:5, pp. 1480–1505. ISSN: 1573-756X. DOI: 10.1007/s10618-017-0527-9.

Mirtabatabaei, A., P. Jia, N. E. Friedkin, and F. Bullo (2014). "On the reflected appraisals dynamics of influence networks with stubborn agents". In: *2014 American Control Conference*, pp. 3978–3983. DOI: 10.1109/ACC.2014.6859256.

Noorazar, H., K. R. Vixie, A. Talebanpour, and Y. Hu (2020). "From classical to modern opinion dynamics". *International Journal of Modern Physics C* **31**:07, p. 2050101. DOI: 10.1142/S0129183120501016.

Ohlin, D., F. Bencherki, and E. Tegling (2022). "Achieving consensus in networks of increasingly stubborn voters". In: *2022 IEEE 61st Conference on Decision and Control (CDC)*, pp. 3531–3537. DOI: 10.1109/CDC51059.2022.9992899.

Peralta, A. F., J. Kert'esz, and G. Iñiguez (2022). "Opinion dynamics in social networks: from models to data". *ArXiv* **abs/2201.01322**.

Proskurnikov, A. V. and R. Tempo (2017). "A tutorial on modeling and analysis of dynamic social networks. part i". *Annual Reviews in Control* **43**, pp. 65–79. ISSN: 1367-5788.

Wang, L., C. Bernardo, Y. Hong, F. Vasca, G. Shi, and C. Altafini (2021). "Achieving consensus in spite of stubbornness: time-varying concatenated friedkin-johnsen models". In: *2021 60th IEEE Conference on Decision and Control (CDC)*, pp. 4964–4969. DOI: 10.1109/CDC45484.2021.9683466.

Wang, L., G. Chen, Y. Hong, G. Shi, and C. Altafini (2022). "A social power game for the concatenated friedkin-johnsen model". In: *2022 IEEE 61st Conference on Decision and Control (CDC)*, pp. 3513–3518. DOI: 10.1109/CDC51059.2022.9992374.

Xia, H., H. Wang, and Z. Xuan (2011). "Opinion dynamics: a multidisciplinary review and perspective on future research". *IJKSS* **2**, pp. 72–91. DOI: 10.4018/jkss.2011100106.

A

Appendix

A.1 Depolarization in the Two-Agent Case

Here, we show that proposition 3.2 does not hold for a network of two agents.

PROPOSITION A.1
Let $\mathcal{G}$ be a graph of 2 opinionated agents and let $\mathcal{A}(\mathcal{G}, f_{\mathcal{A}})$, $\mathcal{B}(\mathcal{G}, f_{\mathcal{B}})$ be two concatenated FJ-models such that $f_{\mathcal{B}}(\theta_i, \delta_i) \geq f_{\mathcal{A}}(\theta_i, \delta_i)$ for all (θ_i, δ_i), with $f_{\mathcal{B}}(\theta_i, \delta_i) > f_{\mathcal{A}}(\theta_i, \delta_i)$ for at least some combination (θ_i, δ_i). It then holds that $\mathcal{A}$ is more depolarizing than $\mathcal{B}$ under any of the polarization metrics GDI, NDI and V. □

Proof. We prove this by explicitly calculating the final opinion vector of a non-concatenated FJ-model with a general initial opinion distribution through the equation

$$y(\infty) = V^{-1}\Theta y(0). \tag{A.1}$$

where $V = (I - (I - \Theta)W)$, previously introduced as equation 2.7. We define the following variables:

$$y(0) = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

$$\Theta = \begin{pmatrix} \theta_1 & 0 \\ 0 & \theta_2 \end{pmatrix}$$

$$y(\infty) = \begin{pmatrix} y_1^\infty \\ y_2^\infty \end{pmatrix}$$

$$W = \begin{pmatrix} a & 1-a \\ 1-b & b \end{pmatrix}$$

Then we can express V as

$$V = \begin{pmatrix} a(\theta_1 - 1) + 1 & (1-a)(\theta_1 - 1) \\ (1-b)(\theta_2 - 1) & b(\theta_2 - 1) + 1 \end{pmatrix} \tag{A.2}$$

and its inverse as

$$V^{-1} = \frac{1}{\det V}\begin{pmatrix} b(\theta_2-1)+1 & (1-a)(1-\theta_1) \\ (1-b)(1-\theta_2) & a(\theta_1-1)+1 \end{pmatrix} \tag{A.3}$$

where

$$\det V = \theta_1(1-b)+\theta_2(1-a)+\theta_1\theta_2(a+b-1) \tag{A.4}$$

Combining (A.1) and (A.3) we get the following expression for $y(\infty)$:

$$y(\infty) = \frac{1}{\det V}\begin{pmatrix} \theta_1 y_1(b(\theta_2-1)+1)+y_2\theta_2(\theta_1-1)(a-1) \\ \theta_2 y_2(a(\theta_1-1)+1)+y_1\theta_1(\theta_2-1)(b-1) \end{pmatrix} \tag{A.5}$$

The strategy now will be to calculate the polarization $\Phi(y(\infty))$ and show that any increase to stubbornness also will increase the polarization. Since the agents are completely arbitrary, it is enough to prove that the derivative

$$\frac{d\Phi(y(\infty))}{d\theta_1}$$

is strictly positive for all $\theta_1,\theta_2,a,b \in (0,1), y_1 \neq y_2$ for all 3 polarization metrics. However, since the network has to be fully connected (otherwise opinions would be stationary in the 2-agent case), GDI and NDI are proportional, so it is enough to consider GDI and variance.

For GDI we get

$$\Phi_{GDI}(y(\infty)) = (y_1^\infty - y_2^\infty)^2 = \frac{1}{(\det V)^2}(\theta_1 y_1(2-\theta_2)-\theta_2 y_2(2-\theta_1))^2$$

which differentiates to

$$\frac{d\Phi_{GDI}(y(\infty))}{d\theta_1} = \frac{2\,(1-a)\,\theta_2^3\cdot(y_2-y_1)^2\,\theta_1}{(((b+a-1)\,\theta_2+1-b)\,\theta_1+(1-a)\,\theta_2)^3}$$

The numerator of this expression is clearly always positive. To evaluate the denominator, we first observe that $(1-a)\theta_2$ and $1-b$ always are positive quantities. To maximize the contribution from the potentially negative quantity $b+a-1$, let $\theta_2 = 1$. Then, the denominator evaluates to $(a\theta_1+(1-a)\theta_2)^3$ which is also positive. This proves the proposition for GDI.

For variance, we see that

$$\Phi_V(y(\infty)) = (y_1^\infty - \frac{y_1^\infty+y_2^\infty}{2})^2 + (y_2^\infty - \frac{y_1^\infty+y_2^\infty}{2})^2 = \frac{1}{2}((y_1^\infty)^2 - 2y_1^\infty y_2^\infty + (y_2^\infty)^2) =$$

$$= \frac{1}{2}(y_1^\infty - y_2^\infty)^2 = \frac{1}{2}\Phi_{GDI}(y(\infty))$$

which also proves the proposition for the variance metric. □

A.2 Proof of Theorem 4.4

Proof. We begin by proving that there is no way for any two agents that are not fully stubborn to increase their distance between their opinions during an issue. In every time step during issue s, each y_i for which $\delta_i(s) \leq \tau$ will be updated as in (2.8):

$$y_i(s,t+1) = (1-a)\bar{y}(s,t) + a y_i(s,0). \tag{A.6}$$

If we let $t \to \infty$, the model will converge and the final opinion of each agent can be explicitly calculated as

$$y_i(s,\infty) = (1-a)\bar{y}(s,\infty) + a \cdot y_i(s,0). \tag{A.7}$$

Now regard two arbitrary agents α and β with $y_\alpha(s,0) > y_\beta(s,0)$ that both fulfil $\delta_i(s) \leq \tau$. Applying A.7 directly yields that

$$y_\alpha(s,\infty) - y_\beta(s,\infty) = a(y_\alpha(s,0) - y_\beta(s,0)).$$

This proves that α and β will reduce the distance between each other's opinions, while preserving the sign of the opinion distance, e.g. $y_\alpha(s,0) > y_\beta(s,0) \Rightarrow y_\alpha(s,\infty) > y_\beta(s,\infty)$. This is important since which agent (or agent pair, if n is even) holds the median opinion could change if the internal order of opinions in the opinion vector y does, which in turn could make $\delta_i(s+1) > \delta_i(s)$ for some agent.

However, we have not yet considered fully stubborn agents. For example, if $y(s,0)$ is initially ordered such that $y_i < y_{i+1}$ and n is odd, the median $\mu(s,0) = y_{(n+1)/2}$. Let $y_{(n+1)/2} = c$ for some arbitrary constant c, $y_{(n+1)/2+1} = c+\tau$ and $y_{(n+1)/2+2} = c+\tau+\varepsilon$ where ε is an arbitrarily small constant. This means that $y_{(n+1)/2}$ and $y_{(n+1)/2+1}$ will have stubbornness $\theta_i = a$ when issue s starts, while $y_{(n+1)/2+2}$ will be fully stubborn and thus not change at all during s. If $\bar{y}(s,\infty) > c+\tau+\varepsilon$, $y_{(n+1)/2}$ and $y_{(n+1)/2+1}$ might update in such a way that $y_{(n+1)/2+2}(s,\infty) < y_{(n+1)/2}(s,\infty) < y_{(n+1)/2+1}(s,\infty)$. Then we have that $\mu(s,\infty) = \mu(s+1,0) = y_{(n+1)/2+2}$, and we have no guarantees about the distance between $y_{(n+1)/2+2}$ and $y_{(n+1)/2+1}$.

We must now derive sufficient conditions on τ and a such that $\delta_i(s+1) \not> \delta_i(s)$ even if the above situation occurs. The strategy will be as follows. We define the most extreme case where:

1. Opinions with $\delta(s) = \tau$ are present on both sides of the median.
2. The median agent moves their opinion as far as possible during s.

These conditions are however not necessarily independent. To achieve the furthest possible movement, placing the median at one end of the opinion spectrum would be ideal, but in such a situation it is impossible to have opinions $\delta(s) = \tau$ on both sides of the median. If we can show that $\tau > 1/6$ ensures that $\delta_i(s+1) \not> \delta_i(s)$ for all such cases, we will have proven the theorem. The considerations are slightly different depending on whether n is odd or even, so we will consider them separately. If n is odd, consider the following setup. Let $n = 2N+1$. Place $N-2$ agents with $y(s,0) = 1$, place $m \leq N$ agents with $y(s,0) = 0$ and $N-m$ agents with $y_\alpha(s,0) = \varepsilon$ for some small constant ε. Then place one agent μ (the initial median) with $y_\mu = pd + \varepsilon$ for some $p : 0 \leq p \leq 1$, one agent β with opinion $y_\beta = \tau(1+p)+\varepsilon$ and one agent γ with $y_\gamma = 2d + 2\varepsilon$. This ensures that y_α and y_β are as far as possible from the median while within d of it, and that y_γ is as close to y_μ as possible without being within d of it. This means that α, β and μ will be the only agents that are not fully stubborn, and that it will be as easy as possible for y_μ to surpass y_γ.

The value p is used to weigh the two extreme cases against each other, $p = 1$ guarantees that opinions with the maximum possible $\delta(s)$ are present on both sides of y_μ, while if $p = 0$ there are no opinions smaller than the median, but the greatest possible movement of the y_μ during s is ensured. We are now interested in finding the lowest possible value of τ such that $\delta_\beta(s+1) \not> \delta_\beta(s)$ for all combinations of a, p and m.

First, let us consider the special case when $p = 1$ and $m = N$. In this configuration, it is clear that $\bar{y}(s,\infty) < 0.5$, since N fully stubborn agents have $y(s,0) = 0$, while only $N-2$ fully stubborn agents have $y(s,0) = 1$. This means that $y_\beta(s,\infty) < 0.5$. We also know that $y_\gamma > 2d$, and we want to bound d such that $\tau > y_\beta(s,\infty) - y_\gamma$, which in turn gives $\tau > y_\beta(s,\infty) - 2\tau \Leftrightarrow \tau > \frac{y_\beta(s,\infty)}{3} > \frac{1}{6}$.

Moving on to the other cases, an important observation is that for $m < N$ and $a \leq 0.5$ it holds that $y_\beta(s,\infty) - y_\alpha(s,\infty) \leq 0.5\tau(1+p) \leq \tau$, which directly gives that $\delta_\alpha(s+1) \not> \delta_\alpha(s)$ since $y_\beta(s,\infty) \leq y_\mu(s,\infty) \leq y_\alpha(s,\infty)$. We thus only need to consider $a > 0.5$ for the following arguments.

The next step will be to find a bound for the movement of y_μ during s. Here we observe that for all $m < N$, the only difference between $m = k$ and $m = k-1$ for a given p is that $y_\mu(s,\infty)$ will be larger when $m = k-1$, as one less opinion is then anchored to $y = 0$. Thus, we will assume $m = 0$ to ensure that the bound will be valid for all $m < N$. Since we know all initial opinions exactly, any final opinion can be explicitly calculated. If we let $\varepsilon \to 0$ and denote $y_\mu(s,\infty)$ by y_μ^∞, we get

$$\bar{y}(s,\infty) = \frac{N-2+Ny_\alpha(s,\infty)+y_\mu^\infty+y_\gamma(s,\infty)+y_\beta(s,\infty)}{2N+1} <$$

$$\frac{N+N(y_\mu^\infty - ap\tau)+y_\mu^\infty}{2N+1} = \frac{y_\mu^\infty(N+1)+N(1-ap\tau)}{2N+1}$$

Using this, we can construct an upper bound for y_μ^∞ as follows:

$$y_\mu^\infty = (1-a)\bar{y}(s,\infty) + ap\tau < (1-a)\frac{y_\mu^\infty(N+1) + N(1-ap\tau)}{2N+1} + ap\tau \Leftrightarrow$$

$$y_\mu^\infty(N(1+a)+a) < (1-a)(N(1-ap\tau)) + (2N+1)ap\tau \Leftrightarrow$$

$$y_\mu^\infty < \frac{Nap\tau(1+a) + N(1-a) + ap\tau}{N(1+a)+a}$$

Since this expression increases with N, the inequality will be valid for all N if we let $N \to \infty$:

$$y_\mu^\infty < \lim_{N\to\infty} \frac{Nap\tau(1+a)+N(1-a)+ap\tau}{N(1+a)+a} = \frac{ap\tau(1+a)+1-a}{1+a}$$

This can now in turn be used to find sufficient bounds for τ.

Let us first consider $p = 1$. Figure A.1 shows the setup for $t = 0$ and $t \to \infty$ in a critical case. We see that for $\delta_\beta(s+1) \le \tau$ to hold, we need to ensure that $y_\beta(s,\infty) - y_\gamma < \tau$. This is enough even if the α-agents also surpass y_γ and become the median - the distance from y_β to y_γ will be larger.

$$y_\beta(s,\infty) - y_\gamma < \tau \Leftrightarrow y_\mu^\infty + a\tau < 3\tau \Leftrightarrow \frac{ad + a^2\tau + 1 - a}{1+a} + a\tau < 3\tau$$

$$\Leftrightarrow 2a^2\tau - a\tau - 3\tau < a - 1 \Leftrightarrow \tau > \frac{a-1}{2a^2-a-3}$$

This quantity decreases with increasing a in the interval, which is easily verifiable.

$$a > 0.5 \Rightarrow \frac{a-1}{2a^2-a-3} < \frac{1}{2(3+\frac{1-1}{2})} = \frac{1}{6} \Rightarrow \tau > \frac{1}{6}$$

Now, for the final case where $p < 1$: Here, m is rendered irrelevant as $y_\mu(s,0) < \tau$. This case is also illustrated in figure A.1. We identify that for $\delta_\beta(s+1) > \delta_\beta(s)$, the following inequalities must hold simultaneously:

$$\begin{cases} y_\beta(s,\infty) - y_\alpha(s,\infty) > \tau \\ y_\beta(s,\infty) - y_\gamma > \tau \end{cases} \iff \begin{cases} ad(1+p) > \tau \\ \frac{apd(1+a)+1-a}{1+a} - \tau(1+p) + ad > \tau \end{cases}$$

If we apply the condition that $d > \frac{1}{6}$ we get the following restrictions on p:

$$\begin{cases} p < \frac{1}{a} - 1 \\ p > \frac{a^2-7a+4}{1+a^2} \end{cases}$$

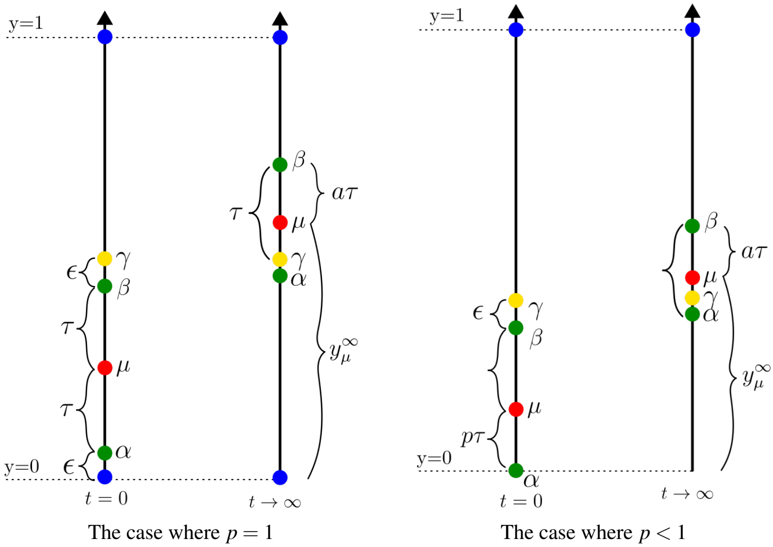


Figure A.1 Illustration of two potential edge cases for odd n

$a = 0.5$ gives $\frac{1}{a} - 1 = 1 = \frac{a^2-7a+4}{1+a^2}$ and for all $a > 0.5$, $\frac{a^2-7a+4}{1+a^2} > \frac{1}{a} - 1$, which means that $\tau < \frac{1}{6}$ is a sufficient bound also in this case. This proves the theorem for all odd n.

Now, for even values of n we must consider a slightly different setup. Let $n = 2N$. and place $N - j$ agents with $y(s,0) = 0$ and j agents with $y_\alpha(s,0) = \varepsilon$. Furthermore, place $N - k - 1$ agents with $y(s,0) = 1$ and k agents with $y_\gamma(s,0) + 2\tau + 2\varepsilon$ and finally place one β-agent, $y_\beta(s,0) = 2\tau + \varepsilon$. This ensures that the median opinion $y_\mu(s,0) = \frac{y_\alpha(s,0)+y_\beta(s,0)}{2}$. Note that we are still denoting the median opinion by y_μ even though there is no agent μ. As before, the α and β-agents are the only ones that will not be fully stubborn during s.

We will now consider two main scenarios: $j > k$ and $j \leq k$. Beginning with $j > k$, we note that if $y_\alpha(s,\infty) > y_\gamma$, we have that $y_\mu(s,\infty) = y_\alpha(s,\infty)$. Since $y_\beta(s,\infty) - y_\alpha(s,\infty) = 2a\tau$, it is immediately clear that $\delta_\beta(s+1) \not> \delta_\beta(s)\ \forall\ a \leq 0.5$.

For $a > 0.5$, we seek an upper bound for $y_\alpha(s,\infty)$ using (2.7).

$$y_\alpha(s,\infty) = (1-a)\frac{n-k-1+y_\alpha(s,\infty)(j+1)+2a\tau+2\tau k}{2n} < (1-a)\frac{n+y_\alpha(s,\infty)n}{2n} \iff$$

$$\iff y_\alpha(s,\infty)(2n - n(1-a)) < n(1-a) \iff y_\alpha(s,\infty) < \frac{1-a}{1+a}$$

We then regard the case when $y_\alpha(s,\infty) < y_\gamma < y_\beta(s,\infty)$. Then, the median will move to $y_\mu = \frac{y_\alpha(s,\infty)+y_\gamma}{2}$ and we receive the following condition on τ:

$$y_\beta(s,\infty) - y_\mu(s,\infty) < \tau \Leftrightarrow y_\alpha + 2ad - \frac{y_\alpha(s,\infty)+y_\gamma}{2} < \tau$$

$$\Leftrightarrow 2y_\alpha(s,\infty) + 4a\tau - y_\alpha(s,\infty) - 2\tau < 2\tau \Leftrightarrow \tau > \frac{-y_\alpha(s,\infty)}{4(1-a)}$$

This is clearly fulfilled as τ cannot be negative. Since this inequality only holds for $y_\alpha(s,\infty) < y_\gamma$, let us however also restrict τ such that it is always true.

$$y_\alpha(s,\infty) < y_\gamma \Rightarrow \frac{1-a}{1+a} < 2\tau \Leftrightarrow \tau > \frac{1-a}{2(1+a)}$$

We know that $\frac{1-a}{2(1+a)} \leq \frac{1}{6}$ for all $a > 0.5$, so we once again find that $\tau > \frac{1}{6}$ is sufficient.

Now for the final case of $j \leq k$. Here, $y_\mu(s,\infty) = y_\gamma$ when $y_\alpha(s,\infty) > y_\gamma$ which gives a condition $y_\beta - y_\gamma < \tau$. When $y_\alpha(s,\infty) < y_\gamma < y_\beta(s,\infty)$, we will have the same condition as before, $y_\beta - \frac{y_\alpha(s,\infty)+y_\gamma}{2}$. We shall adopt a slightly different strategy and derive an upper bound for y_β. To ensure that it is valid for all $j \leq k$, let $k = j$, which will allow for the furthest possible opinion movement.

$$y_\beta = (1-a)\bar{y}(s,\infty) + ay_\beta(s,0) \leq \frac{1-a}{2n}(n-1-k+y_\beta+k(y_\beta-2a\tau)+2\tau k)+2ad$$

We also note that the maximum possible y_β must be found when $a = 0$ which yields

$$y_\beta(s,\infty)(2n-1-k) \leq n-1-k+2\tau k \Leftrightarrow y_\beta(s,\infty) \leq \frac{n-1-k+2}{2n-1-k}$$

Applying our condition on $\tau > \frac{1}{6}$ gives

$$y_\beta(s,\infty) \leq \frac{n-1-\frac{2k}{3}}{2n-1-k}$$

Let us now expand on this by showing that $y_\beta(s,\infty) \leq \frac{1}{2}$. This can be done by contradiction:

$$\frac{n-1-\frac{2k}{3}}{2n-1-k} > \frac{1}{2} \Leftrightarrow 2(n-1-\frac{2k}{3}) > 2n-1-k \Leftrightarrow -2-\frac{4k}{3} > -1-k \Leftrightarrow \frac{k}{3} < -1$$

But $k > 0$, so we must have that $y_\beta(s,\infty) \leq \frac{1}{2}$. Applying this to our condition $y_\beta - y_\gamma < \tau$, we see that

$$y_\beta - y_\gamma \leq \frac{1}{2} - y_\gamma < \tau \Leftrightarrow \frac{1}{2} - 2\tau < \tau \Leftrightarrow \tau > \frac{1}{6}$$

Now for the very final condition $y_\beta - \frac{y_\alpha(s,\infty)+y_\gamma}{2} < \tau$:

$$y_\beta - \frac{y_\alpha(s,\infty)+y_\gamma}{2} \leq \frac{1}{2} - \frac{\frac{1}{2}+y_\gamma}{2} = \frac{\frac{1}{2}+2a\tau-2\tau}{2} < \tau$$

$$\Leftrightarrow 1+4a\tau-4\tau < 4\tau \Leftrightarrow \tau > \frac{1}{4(2-a)} \Rightarrow \tau > \frac{1}{8} \qquad \square$$

This, at last, proves the theorem.

Lund University
Department of Automatic Control
Box 118
SE-221 00 Lund Sweden

Document name
BACHELO'S THESIS

Date of issue
May 2024

Document Number
TFRT-7669

Author(s)
Johan Swartling Sennhed

Supervisor
David Ohlin, Dept. of Automatic Control, Lund University, Sweden
Emma Tegling, Dept. of Automatic Control, Lund University, Sweden
Giacomo Como, Dept. of Automatic Control, Lund University, Sweden (examiner)

Title and subtitle
On Dynamic Stubbornness in the Concatenated Friedkin-Johnsen Model

Abstract

We study opinion dynamics in social networks using a version of the concatenated Friedkin-Johnsen model, where the opinion evolution of the participating agents is allowed to continue over sequences of related issues. These agents have a certain associated stubbornness that anchors them to their initial opinion, which we allow them to update between issues. The central purpose of the present work is to investigate the dynamics of this update, or in other words how different stubbornnessupdating functions affect the behaviour of the model. For two convex and concave families of updating functions, we numerically explore which conditions guarantee that the agents will reach consensus, and conclude that the chosen functions can be well-approximated by linear ones with regard to consensus convergence criteria. We also study how dynamic stubbornness affects the depolarizing properties of the model, and show that agents becoming more stubborn does not always lead to more polarized opinion distributions. Lastly, we introduce a version of the model where dissatisfied agents entirely refuse to change their opinion until other agents approach their position enough. For this model, we provide some sufficient conditions for consensus, and give numerical examples that illustrate its oscillatory behaviour.

Keywords

Classification system and/or index terms (if any)

Supplementary bibliographical information

ISSN and key title
0280-5316

ISBN

Language
English

Number of pages
1-63

Recipient's notes

Security classification

http://www.control.lth.se/publications/