2022

# Smoothed Bounded-Confidence Opinion Dynamics on the Complete Graph

Solomon Valore-Caplan

# Smoothed Bounded-Confidence Opinion Dynamics on the Complete Graph

**Solomon Valore-Caplan**

Heather Zinn-Brooks, Advisor

Vin de Silva, Reader

Harvey Mudd College

**Department of Mathematics**

May, 2022

# Abstract

We present and analyze a model for how opinions might spread throughout a network of people sharing information. Our model is called the smoothed bounded-confidence model and is inspired by the bounded-confidence model of opinion dynamics proposed by Hegselmann and Krause. In the Hegselmann–Krause model, agents move towards the average opinion of their neighbors. However, an agent only factors a neighbor into the average if their opinions are sufficiently similar. In our model, we replace this binary threshold with a logarithmic weighting function that rewards neighbors with similar opinions and minimizes the effect of dissimilar ones. This weighting function can be tuned with parameters $\gamma$ and $\delta$ and recovers the Hegselmann–Krause model as $\gamma$ approaches infinity.

We analyze the effect of $\gamma$ and $\delta$ on some of the stationary states of the smoothed bounded-confidence model on the complete graph. In particular, we analyze the stationary states with consensus and those with two distinct opinions.

# Contents

# List of Figures

# Acknowledgments

Many thanks to Prof. Heather for all of her mentorship on this research project, the thesis writing process, and much life advice! I would also like to thank Phil Chodrow for welcoming me onto this project and all the work that he and Prof. Heather had already done.

Thanks to Prof. Jakes, DruAnn Thomas, and Melissa Hernandez-Alvarez for everything they did to create our senior thesis experience this year.

Thanks to Sydney for keeping me company during many late nights spent writing this thesis in Shan classrooms.

# Chapter 1

# Introducing Some Classical Models of Opinion Dynamics

## 1.1 Towards a Mathematical Model of Opinions

We are all in a constant state of forming new opinions and adapting old ones in response to new pieces of information. Our opinions about local restaurants, political platforms, and favorite movies all live in our mind with some degree of flux. Sometimes these opinions change in direct response to a new experience (maybe one of the restaurants you frequent has just changed their menu and it is not to your liking). However, our opinions can also change by simply learning the current opinions of other people. The restaurant's menu may not change at all, but if you start to interact with a number of people who have a negative opinion of the restaurant's food, that can affect your opinion as well. It is this latter phenomenon that we will be investigating.

In this thesis, we analyze a model of opinion dynamics based on the notion that our opinions tend to move towards the opinions of others that we share information with. In particular, we will be analyzing the smoothed bounded-confidence model, which has been recently proposed by Brooks and Chodrow (2022). It is inspired by some of the foundational models of opinion dynamics, most notably the models of Taylor (1968) and Hegselmann and Krause (2002). In particular, the smoothed bounded-confidence model is a tunable model which recovers either the Taylor or Hegselmann–Krause model for particular parameter choices.

Like both the Taylor and Hegselmann–Krause models, the smoothed bounded-confidence model is defined on a network of people. That is, we consider how opinions propagate on a particular graph structure. In this thesis, we will be focused on characterizing the behavior of the smoothed bounded-confidence model on complete graphs. We hope that an understanding of the dynamics on this graph with simple structure can inform an understanding of the dynamics on more complicated graphs. (Homs-Dones et al. (2021) has shown that there exist relationships between dynamics on a network and those dynamics on its subgraphs.)

## 1.2 Introduction to Nonlinear Systems

There are many ways that one could choose to mathematically represent opinions. Some modelers choose to assign agents to opinion categories and have them update which opinion category they belong to (see, for instance, Yildiz et al. (2013) where each agent has one of two possible opinions).

In the smoothed bounded-confidence model, as in the Taylor and Hegselmann–Krause models, we instead use an interval of the real numbers to represent the opinions of our agents. That is, each agent $X_i$ in the graph has an opinion $x_i \in \mathbb{R}$. (To keep consistent throughout this thesis, we will be using the opinion interval $x_i \in [-1, 1]$.) Then, the opinion of $X_i$ changes continuously in time as a nonlinear function of the opinions of its neighbors.

Because the opinion profile $\mathbf{x}$ evolves according to a nonlinear function, we will be leveraging theory from the study of nonlinear systems to study its behavior. We introduce some key ideas from nonlinear systems here, but for a more complete description of this theory see Alligood et al. (1996), Guckenheimer and Holmes (2013), Strogatz (1994), or other textbooks on the subject).

**Definition 1.2.1.** *Stationary States* (Alligood et al. (1996)). A constant solution of the autonomous differential equation $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$ is called a stationary state of the equation. A stationary state necessarily satisfies $\mathbf{f}(\mathbf{x}) = \mathbf{0}$.

Stationary states go by other names as well (such as equilibria and fixed points) but we will call them stationary states throughout this thesis for consistency. In our model, stationary states are distributions of opinions across our network such that none of the opinions are changing with time.

**Definition 1.2.2.** *Asymptotic Stability* (Alligood et al. (1996)). A stationary state $\mathbf{x}^*$ is asymptotically stable if it is both stable (every initial point $\mathbf{x_0}$ chosen very close to $\mathbf{x}^*$ has the property that the solution $\mathbf{f}(t, \mathbf{x_0})$ stays close to $\mathbf{x}^*$ for $t \geq 0$) and attracting (the trajectories of nearby initial conditions converge to it). We say that $\mathbf{x}^*$ is unstable if it is not stable.

Characterizing the stability of the stationary states of our system is important for understanding the dynamics. Knowing which stationary states are actually being approached by solution trajectories can help us understand the behavior of the solution trajectories in general. The following theorem will help us determine the stability of our stationary states.

**Theorem 1.1.** (*Alligood et al.* (1996)). *Let $\mathbf{x}^*$ be a stationary state of $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$. If the real part of each eigenvalue of $\mathbf{J_f}(\mathbf{x}^*)$ is strictly negative, then $\mathbf{x}^*$ is asymptotically stable. If the real part of at least one eigenvalue is strictly positive, then $\mathbf{x}^*$ is unstable.*

Here, $\mathbf{J_f}(\mathbf{x}^*)$ refers to the Jacobian of $\mathbf{f}$ at $\mathbf{x}^*$. The Jacobian of $\mathbf{f}$ at $\mathbf{x}^*$ is a matrix of partial derivatives of $\mathbf{f}$ evaluated at $\mathbf{x}^*$ and is the best linear approximation of $\mathbf{f}$ at $\mathbf{x}^*$ (see: Alligood et al. (1996)).

With this nonlinear systems vocabulary under our belt, we move on to defining our model.

## 1.3 The Smoothed Bounded-Confidence Model

Before introducing the Taylor and Hegselmann–Krause models, we define the smoothed bounded-confidence model. Having this definition in mind will be useful to us as we motivate the special cases of the Taylor and Hegselmann–Krause models that the smoothed bounded-confidence model can recover.

We consider a graph defined by a vertex set of agents $\mathcal{N} = \{X_1 \ldots, X_N\}$ and edges $\mathcal{E} \subset \mathcal{N} \times \mathcal{N}$. (In summation bounds, we often refer to an agent $X_i$ simply as $i$ for visual clarity.) The agents are partitioned into two subsets: the persuadable nodes $\mathcal{P}$ whose opinions can update with time and the zealots $\mathcal{Z}$ whose opinions are fixed. Each agent $X_i$ has an opinion $x_i$ which is governed by the dynamics

$$\frac{dx_i}{dt} = f_i(x) \triangleq \begin{cases} \frac{\sum_{j \in \mathcal{N}} w(x_i, x_j)(x_j - x_i)}{\sum_{j \in \mathcal{N}} w(x_i, x_j)} & i \in \mathcal{P} \\ 0 & i \in \mathcal{Z} \end{cases} \tag{1.1}$$

where $w(x_i, x_j)$ is the weighting function

$$w(x_i, x_j) = \begin{cases} \frac{1}{1+e^{-\gamma(x_i - x_j)^2 - \gamma\delta}} & (i, j) \in \mathcal{E} \\ 0 & \text{otherwise} \end{cases}$$

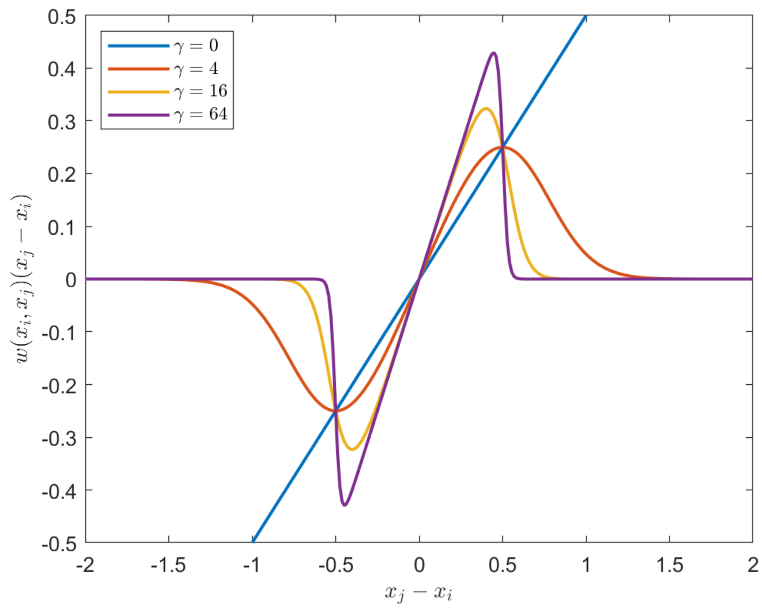


**Figure 1.1** The $x_j$ contribution to the weighted average that governs the dynamics of $x_i$ for $\delta = 0.25$ (note that the step occurs at $\sqrt{\delta} = 0.5$)

and $\gamma, \delta \geq 0$ are model parameters.

In this model, the agents move towards the weighted average of their neighbors' opinions, where the weighting is dictated by $w(x_i, x_j)$. To get a sense of how this weighted average operates for various values of the model parameter $\gamma$, see Figure 1.1.

## 1.4 Taylor's Model of Continuous Averaging

In this section, we describe in further detail the model proposed in Taylor (1968). Taylor introduced this model as an extension of the model proposed in Abelson (1967), in an attempt to explain how a lack of consensus could emerge from an averaging-based model of opinion dynamics. In Abelson's original model, the system always converged to consensus (that is, to a point where all of the agents had the same opinion value). He suggested the development of a variation on his model that would include agents with fixed opinions to drive faction formation. Taylor accepted this challenge and proposed his model soon afterwards.

Taylor called these agents with fixed opinions "constant sources." To keep with current convention in the literature (seen as early as Mobilia (2003)), we will refer to them as "zealots." We will refer to the other agents as "persuadable nodes." From a modelling perspective, the interpretation of zealots is incredibly flexible. In Brooks and Porter (2020) zealots are used to represent media outlets whose content is consumed by agents in the network but whose platforms are not changed by the state of the network. The zealots could also represent actual people in the network who are just too stubborn to change their minds on the opinion at hand. They could represent the opinions of people who matter to the persuadable nodes but are not present for the exchange of information that is occurring in the network. The point here is that zealots can have meaningful and reasonable real-world interpretations—they are not just an artificial constraint introduced in pursuit of non-consensus stationary states.

In Taylor's model, we begin with a network of agents. To keep notational conventions consistent throughout this thesis we will denote the set of graph vertices $\mathcal{N}$ and the set of graph edges $\mathcal{E}$. Each agent $X_i \in \mathcal{N}$ has an opinion $x_i \in [-1, 1]$. In Taylor's model, the edges are weighted and directed. Each directed edge $(i, j)$ has associated with it a weight $a_{ij} \geq 0$ that represents the rate at which agent $j$ influences agent $i$. The dynamics of $\mathbf{x}$ are defined as

$$\frac{dx_i}{dt} = \sum_{j=1}^{n} a_{ij}(x_j - x_i).$$

Taylor observes that if we define

$$a_{ii} = - \sum_{(j \neq i)=1}^{n} a_{ij},$$

then we can define a matrix $A = (a_{ij})$, and the dynamics can be rewritten as

$$\frac{d\mathbf{x}}{dt} = A\mathbf{x}.$$

In Taylor's model, each agent's opinion is always moving towards the weighted average of its neighbors' current opinions.

### 1.4.1 A Special Case of the Taylor Model

The smoothed bounded-confidence model can recover a particular case of the Taylor model. In particular, we will define here the case of the Taylor model that can be recovered by the smoothed bounded-confidence model on the complete graph. First, we partition our agents $\mathcal{N}$ into two subsets: the zealots $\mathcal{Z}$ and the persuadable nodes $\mathcal{P}$. Now, we define

$$a_{ij} = \begin{cases} (-|\mathcal{N} - 1|) & i = j, i \in \mathcal{P} \\ 1 & i \neq j, i \in \mathcal{P} \\ 0 & i \in \mathcal{Z} \end{cases}$$

This formulation of the model corresponds to a complete graph, where every persuadable node is being influenced equally by every other node on the graph (including the zealots).

**Lemma 1.1.** *Let $A = (a_{ij})$ be defined with components*

$$a_{ij} = \begin{cases} (-|\mathcal{N} - 1|) & i = j, i \in \mathcal{P} \\ 1 & i \neq j, i \in \mathcal{P} \\ 0 & i \in \mathcal{Z} \end{cases}$$

*for a graph with at least one zealot. Then, the only stationary state of the Taylor model is*

$$x_i = \begin{cases} \frac{z_m + \cdots + z_n}{|\mathcal{Z}|} & X_i \in \mathcal{P} \\ z_i & X_i \in \mathcal{Z} \end{cases}$$

*where $z_m, \ldots, z_n$ are the opinions of the zealots.*

*Proof.* Let $A_{\mathcal{P}}$ and $\mathbf{x}_{\mathcal{P}}$ be the restrictions of $A$ and $\mathbf{x}$ to the persuadable dimensions. The zealots have fixed opinions $z_m \ldots z_n$ and are stationary for any value of the persuadable nodes. So, the dynamics of the whole system are really just the dynamics of the persuadable nodes, which we can rewrite as

$$\frac{d\mathbf{x}_{\mathcal{P}}}{dt} = A_{\mathcal{P}}\mathbf{x}_{\mathcal{P}} + (z_m + \cdots + z_n)\mathbf{1}.$$

Thus, a persuadable opinion profile $\mathbf{x}_{\mathcal{P}}$ is stationary when

$$A_{\mathcal{P}}\mathbf{x}_{\mathcal{P}} = -(z_m + \cdots + z_n)\mathbf{1}. \tag{1.2}$$

It is clear from the definition of $A_{\mathcal{P}}$ that it is a rank $|\mathcal{P}|$ matrix. This means that Equation 1.2 has a unique solution. This solution must then be the only stationary state of the system.

We now simply check that

$$\mathbf{x}_{\mathcal{P}} = \frac{(z_m + \cdots + z_n)}{|\mathcal{Z}|}\mathbf{1}$$

is a stationary state of the system by taking $A_{\mathcal{P}}\mathbf{x}_{\mathcal{P}}$ and finding that

$$A_{\mathcal{P}}\mathbf{x}_{\mathcal{P}} = -(z_m + \cdots + z_n)\mathbf{1}$$

as desired. □

This result is important because it means that the equally-weighted Taylor model on a complete graph still converges to consensus amongst the persuadable nodes, even with the presence of zealots.

## 1.5 The Hegselmann–Krause Model of Conditional Receptivity

In this section, we introduce the Hegselmann–Krause model of opinion dynamics in further detail. This model was first proposed in Hegselmann and Krause (2002) and has been extensively studied ever since. Hegselmann and Krause call their model a "bounded confidence" model. This is in reference to their introduction of an opinion difference threshold (the *confidence bound*) that each agent uses to filter the inclusion of other agents' opinions into their own opinion update. That is, each agent $X_i$ has a confidence bound $\epsilon_i$ and the dynamics of its opinion $x_i$ only depends on another opinion $x_j$ if $|x_j - x_i| \leq \epsilon_i$.

The original formulation of the Hegselmann–Krause model was defined as a map in discrete time. That is, the dynamics are defined as

$$x_i(t+1) = |I(i, x(t))|^{-1} \sum_{j \in I(i, \mathbf{x}(t))} x_j(t)$$

where $I(i, \mathbf{x}) = \left\{1 \leq j \leq n \mid d(x_i, x_j) \leq \epsilon_i\right\}$ and $d(x_i, x_j)$ is a metric on the opinion space. In Hegselmann and Krause (2002) they use $d(x_i, x_j) = |x_i - x_j|$.

A continuous-time analog of the Hegselmann–Krause model is proposed in Blondel et al. (2010). There, the dynamics are essentially defined as

$$\frac{dx_i}{dt} = \sum_{j \in I(i, \mathbf{x})} (x_j(t) - x_i(t)).$$

The move to continuous time requires additional results about existence and uniqueness of solutions (provided in Blondel et al. (2010)) and analysis of dynamics of functions rather than dynamics of mappings.

Neither the original Hegselmann–Krause model nor the Blondel et al. extension are defined on a graph structure. Rather, they are posed for a collection of $n$ agents who exchange information with each other if they are sufficiently close in opinion space. That being said, the Hegselmann–Krause model has been studied on graph networks before (see Fortunato (2005), for instance) and it is a natural extension of the model.

In the Hegselmann–Krause model (and its continuous-time analog), it is a well-known feature that the stationary states consist of "clusters" of agents with the same opinion that are each spaced further than $\epsilon$ apart from the others. In this thesis, we will refer to these clusters as "factions" and will analyze the conditions under which distinct factions form on the complete graph.

### 1.5.1 A Special Case of the Hegselmann–Krause Model

The smoothed bounded-confidence model is defined so that it can both recover the behavior of direct averaging models (like the Taylor model) and of bounded-confidence models. In this section, we will describe the exact version of the Hegselmann–Krause model that can be recovered by the smoothed bounded-confidence model on the complete graph.

The continuous Hegselmann–Krause model on the complete graph with a squared distance threshold $\delta$ would look like

$$\frac{dx_i}{dt} = \sum_{j \in I(i,\mathbf{x})} (x_j(t) - x_i(t))$$

where $I(i,\mathbf{x}) = \left\{1 \leq j \leq n \mid d(x_i, x_j) \leq \epsilon_i\right\}$ for $d(x_i, x_j) = (x_i - x_j)^2$.

We observe that we could equivalently replace the sum condition with an indicator function

$$\mathbb{1}(x_i, x_j) = \begin{cases} 1 & (x_i - x_j)^2 \leq \delta \\ 0 & (x_i - x_j)^2 > \delta, \end{cases}$$

and let

$$\frac{dx_i}{dt} = \sum_{j=1}^{n} \mathbb{1}(x_i, x_j) \cdot (x_j(t) - x_i(t)).$$

The sign of the dynamics is preserved after renormalizing, which means we can renormalize this expression to

$$\frac{dx_i}{dt} = \frac{\sum_{j=1}^{n} \mathbb{1}(x_i, x_j) \cdot (x_j(t) - x_i(t))}{\sum_{j=1}^{n} \mathbb{1}(x_i, x_j)}.$$

Now, we note that in the limit as $\gamma \to \infty$, the smoothed bounded-confidence model is

$$\frac{dx_i}{dt} = \frac{\sum_{j=1}^{n} w(x_i, x_j) \cdot (x_j(t) - x_i(t))}{\sum_{j=1}^{n} w(x_i, x_j)}$$

where the weighting function is converging pointwise to

$$w(x_i, x_j) = \begin{cases} 1 & (x_i - x_j)^2 < \delta \\ 1/2 & (x_i - x_j)^2 = \delta \\ 0 & (x_i - x_j)^2 > \delta. \end{cases}$$

Thus, in the limit as $\gamma \to \infty$, the smoothed bounded-confidence model is very nearly the same as the continuous Hegselmann–Krause model on squared distance. The fact that $w(x_i, x_j) = 1/2$ when $x_i$ and $x_j$ are exactly $\sqrt{\delta}$ apart in opinion space is an unfortunate artifact of the pointwise convergence of the model, but we will address it more carefully when we analyze the $\gamma \to \infty$ case in later chapters.

# Chapter 2

# The Smoothed Bounded-Confidence Model

In this chapter, we introduce general results about the smoothed bounded-confidence model. This involves reproducing some results from Brooks and Chodrow (2022). These results and terminology apply to graph structures other than the complete graph, but we will apply them to the complete graph for our analysis in Chapters 3 and 4.

## 2.1 The Smoothed Bounded-Confidence Model on a Network Without Zealots

As we discussed in Section 1.4, the Abelson model of opinion dynamics on a connected graph always converges to consensus and this is what prompted the introduction of zealot nodes in the Taylor model. In this section, we briefly show that the smoothed bounded-confidence model also converges to consensus on a connected graph when there are no zealots (for any finite choice of $\gamma$).

**Lemma 2.1.** *Let $G = \{\mathcal{N}, \mathcal{E}\}$ be a finite connected graph with no zealots. Then, all stationary states of the smoothed bounded-confidence model with finite $\gamma$ on $G$ have consensus.*

*Proof.* We prove this by contradiction. Assume that there exists a stationary state $\mathbf{x}^*$ which contains at least two distinct opinions.

Let $X_j$ be an agent in $\mathcal{N}$ (not necessarily unique) such that $x_j$ is the maximum opinion attained on $G$.

Recall the dynamics of the smoothed bounded-confidence model provided in Equation 1.1. We consider the dynamics of $x_j$. Denoting the set of $X_j$'s neighbors as $\mathcal{N}_{X_j}$, we observe that

$$\frac{dx_j}{dt} = C_j \sum_{i \in \mathcal{N}_{X_j}} \frac{(x_i - x_j)}{1 + e^{-\gamma(x_i - x_j)^2 - \gamma\delta}}$$

(where the denominator of the dynamics has been absorbed into a positive constant $C_j$).

Now, since $x_j$ is the maximum opinion value attained on $G$, $(x_i - x_j)$ is non-positive. Since the denominator of the summand is strictly positive, the entire summand is always non-positive. Since $\mathbf{x}^*$ is a stationary state by assumption, it must be that every summand is equal to 0. That is, every $i \in \mathcal{N}_{x_j}$ must have $x_i = x_j$. More colloquially, if $X_j$ is an agent whose opinion attains the maximum opinion value in a stationary state, then all of the neighbors of $X_j$ must also attain the maximum.

Now, recall that our graph $G$ is connected and contains two distinct opinions. Let $X_i \neq X_j \in \mathcal{N}$ such that $x_i$ is the minimum opinion attained on $G$. Since the graph is connected, there exists a path from $X_i$ to $X_j$. That is, there is a set of agents $\{A_1, A_2, \dots, A_n\} \in \mathcal{N}$ such that $\{(X_i, A_1), (A_1, A_2), \dots, (A_n, X_j)\} \subset \mathcal{E}$.

We just showed, however, that in a stationary state of the dynamics all neighbors of a node that attains the maximum must also attain the maximum. This propagates down the path from $X_j$ to $X_i$ and implies that $X_i$ attains the maximum opinion value as well. This contradicts the definition of $X_i$ and therefore implies that there cannot exist a stationary state with at least two opinions on a graph with no zealots. □

We can show similarly that in a graph with exactly one zealot $Z_1$ with opinion $z_1$, the only stationary state is $\mathbf{x} = z_1\mathbf{1}$.

**Lemma 2.2.** *Let $G = \{\mathcal{N}, \mathcal{E}\}$ be a finite connected graph with exactly one zealot $Z_1$, with opinion $z_1$. Then, the only stationary state of the smoothed bounded-confidence model with finite $\gamma$ on $G$ is $\mathbf{x}^* = z_1\mathbf{1}$.*

*Proof.* First, let $G_1, G_2, \dots, G_m$ be the connected components of $G - Z_1$. We will prove that in $\mathbf{x}^*$ each connected component contains only the opinion $z_1$.

Consider an arbitrary $G_k$. If $G_k$ contains only one vertex then it's clear that that vertex must have opinion $z_1$ to be stationary. So, let us assume that $G_k$ contains more than one vertex.

As in Lemma 2.1, we let $X_j$ be an agent which obtains the maximum opinion value attained on $G_k$. As in that proof, we note that the dynamics of $x_j$ are

$$\frac{dx_j}{dt} = C_j \sum_{i \in \mathcal{N}_{X_j}} \frac{(x_i - x_j)}{1 + e^{-\gamma(x_i - x_j)^2 - \gamma\delta}}$$

where $\mathcal{N}_{X_j}$ is the set of neighbors of $X_j$ (in the uncut graph $G$) and the denominator of the dynamics has been absorbed into a positive constant $C_j$. Now, we observe that for $x_j$ to be stationary (as is necessary for $\mathbf{x}^*$ to be a stationary state) requires that either

1. All of the neighbors of $X_j$ have opinion $x_j$, or
2. $X_j$ is connected to $Z_1$ and $z_1 > x_j$.

Since $G$ is connected, there exists a path from $X_j$ to $Z_1$. If condition 2 is never met on that path, then $Z_1$ must have opinion $x_j$. Otherwise, $z_1$ must be strictly greater than $x_j$. Either way, we conclude that $z_1 \geq x_j$.

We repeat a similar analysis with an agent $X_i$ that attains the minimum opinion value on $G_k$, and find that for $X_i$ to be stationary requires either that

1. All of the neighbors of $X_i$ have opinion $x_i$, or
2. $X_i$ is connected to $Z_1$ and $z_1 < x_i$,

which similarly implies that $z_1 \leq x_i$.

We know that $x_i \leq x_j$, so the only way that we can have both $z_1 \geq x_j$ and $z_1 \leq x_i$ is if $x_i = x_j = z_1$. Thus, the minimum and maximum opinion values attained on $G_k$ are both $z_1$, which is what we wanted to show. □

Throughout Chapters 3 and 4, we will be analyzing the smoothed bounded-confidence model on a complete graph with two zealots because (as we have just shown) any fewer than two zealots leads to rather straightforward dynamics.

## 2.2 The Smoothed Bounded-Confidence Jacobian

In this section, we reproduce analysis from Brooks and Chodrow (2022) of the Jacobian of the smoothed bounded-confidence model. In particular, we will describe a matrix $\mathbf{M}_{\mathcal{P}}$ that is similar to the restriction of the Jacobian to the persuadable subsystem. Since similar matrices have the same eigenvalue spectra, we can perform stability analysis with Theorem 1.1 by evaluating the eigenvalues of $\mathbf{M}_{\mathcal{P}}$ instead of the Jacobian.

We begin by putting the Jacobian in block structure based on the persuadable and zealot nodes. We let

$$\mathbf{J} = \begin{bmatrix} \mathbf{J}_{\mathcal{P}} & \mathbf{J}_{\mathcal{P}\mathcal{Z}} \\ \mathbf{J}_{\mathcal{Z}\mathcal{P}} & \mathbf{J}_{\mathcal{Z}} \end{bmatrix}$$

where the upper block rows correspond to functions $f_i(\mathbf{x})$ with $i \in \mathcal{P}$ and the lower rows correspond to functions $f_i(\mathbf{x})$ with $i \in \mathcal{Z}$. Similarly, the columns in the left blocks correspond to partial derivatives with respect to $\mathbf{x}_j$ where $j \in \mathcal{P}$ and the columns in the right blocks correspond to $j \in \mathcal{Z}$.

Since the zealot nodes do not change their opinions, we observe that the lower two blocks vanish and we are left with

$$\mathbf{J} = \begin{bmatrix} \mathbf{J}_{\mathcal{P}} & \mathbf{J}_{\mathcal{P}\mathcal{Z}} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}.$$

As long as we have at least one zealot, then we have some zero rows in our matrix. At first it seems that this might be a problem for our eigenvalue analysis because each zero row corresponds to a standard basis vector with eigenvalue 0 (and recall that Theorem 1.1 only guarantees asymptotic stability for a system when all of its eigenvalues are strictly negative).

However, upon further reflection we observe that these eigenvectors with eigenvalue 0 correspond to perturbations of the zealot nodes. We are only concerned with whether a stationary state is stable to perturbations of its *persuadable* nodes, so we can ignore these eigenvalues in our stability analysis. Really, the stability of our system under perturbations of the persuadable nodes is determined entirely by the eigenvalue spectrum of $\mathbf{J}_{\mathcal{P}}$.

Now, we begin to evaluate what this matrix looks like for the smoothed bounded-confidence model. It will be convenient for us to first define a couple of new objects. We use $\mathbf{W}(\mathbf{x})$ to represent a weight matrix, defined as $w_{ij}(\mathbf{x}) = w(x_i, x_j)$. We also define a vector $\mathbf{s}$ as

$$s_i = \sum_{j \in \mathcal{N}} w_{ij} \tag{2.1}$$

for all $i \in \mathcal{N}$. (This vector is equivalently defined in Brooks and Chodrow (2022) as $\mathbf{s} = \mathbf{W}\mathbf{1}$). Note that we can now redefine our update operator for a persuadable node $i$ as

$$f_i(\mathbf{x}) = \frac{1}{s_i} \sum_{j \in \mathcal{N}} w_{ij}(x_j - x_i).$$

Now, we begin computing the necessary derivatives for $\mathbf{J}_{\mathcal{P}}$. The $(i, j)$ entry of $\mathbf{J}_{\mathcal{P}}$ is $\partial f_i(x)/\partial x_j$ where $i, j \in \mathcal{P}$. We assume first that $i$ and $j$ are distinct and not adjacent. Then, they do not directly affect each other's opinions and so that entry of $\mathbf{J}_{\mathcal{P}}$ is zero.

To compute the other components of $\mathbf{J}_{\mathcal{P}}$, we begin by assuming that $i$ and $j$ are adjacent (and distinct). As we take the partial derivatives of $f_i$ we consider the definition of $f_i$ from the equation above. (Note that $s_i$ is a function of $x_j$, so we use the product rule.) We find

$$\frac{\partial f_i(\mathbf{x}^*)}{\partial x_j} = \left(\frac{\partial}{\partial x_j}\frac{1}{s_i}\right)\sum_{k\in\mathcal{N}} w_{ik}(x_k - x_i) + \frac{1}{s_i}\left(\frac{\partial}{\partial x_j}\sum_{k\in\mathcal{N}} w_{ik}(x_k - x_i)\right).$$

We observe that the summation on the left is actually equivalent to $s_i f_i(\mathbf{x})$. Since we are computing $\mathbf{J}_{\mathcal{P}}$ at a stationary state, this must vanish. We turn then to the term on the right. This partial derivative is only non-zero when $k = j$, so we can replace the summation with that particular summand. Altogether, we have reduced our equation to

$$\begin{aligned}\frac{\partial f_i(\mathbf{x}^*)}{\partial x_j} &= \frac{1}{s_i}\left(\frac{\partial}{\partial x_j}\left[w_{ij}(x_j - x_i)\right]\right)\\ &= \frac{1}{s_i}\left(\frac{\partial w_{ij}}{\partial x_j}(x_j - x_i) + w_{ij}\right).\end{aligned}$$

At this point, we plug in the explicit weight function. We find that

$$\begin{aligned}\frac{\partial w_{ij}}{\partial x_j} &= \frac{\partial}{\partial x_j}\left(\frac{1}{1+e^{-\gamma(x_i-x_j)^2-\gamma\delta}}\right)\\ &= -\left(\frac{1}{1+e^{-\gamma(x_i-x_j)^2-\gamma\delta}}\right)\left(\frac{e^{-\gamma(x_i-x_j)^2-\gamma\delta}}{1+e^{-\gamma(x_i-x_j)^2-\gamma\delta}}\right)(2\gamma(x_j - x_i))\\ &= -2\gamma w_{ij}(1-w_{ij})(x_j - x_i).\end{aligned}$$

Now, our expression for the $i, j$ entry of $\mathbf{J}_{\mathcal{P}}$ (given $i$ and $j$ are adjacent) is

$$\frac{\partial f_i(\mathbf{x}^*)}{\partial x_j} = \frac{w_{ij}}{s_i}(1 - 2\gamma(1-w_{ij})(x_j - x_i)^2) \tag{2.2}$$

Finally, we compute $\mathbf{J}_{\mathcal{P}}$ when $i = j$. We observe that $\partial w_{ik}/\partial x_i = -\partial w_{ik}/\partial x_k$. Then, one can show via algebraic manipulation that

$$\frac{\partial f_i(\mathbf{x})}{\partial x_i} = -\sum_{j\neq i\in\mathcal{N}}\frac{\partial f_i(\mathbf{x})}{\partial x_j}. \tag{2.3}$$

With this in mind, we define some new matrices which we can sum to obtain the Jacobian. In particular, we define a matrix **S** as

$$S_{ij} = \begin{cases} s_i & i = j \\ 0 & \text{otherwise} \end{cases}$$

where **s** is defined as in Equation 2.1.

We also define a matrix **Q** with entries $Q_{ij} = w_{ij}(1 - w_{ij})(x_i - x_j)^2$.

Finally, we define a diagonal matrix **R** as

$$R_{ij} = \begin{cases} (\mathbf{Q1})_i & i = j \\ 0 & \text{otherwise} \end{cases}$$

where **Q1** is the vector obtained from multiplying the matrix **Q** by the all ones vector.

For each of these matrices (and the weight matrix **W**) we notate the restriction of the matrix to its persuadable entries as $\mathbf{S}_{\mathcal{P}}$, $\mathbf{Q}_{\mathcal{P}}$, $\mathbf{R}_{\mathcal{P}}$, and $\mathbf{W}_{\mathcal{P}}$, respectively.

These matrices are defined such that the restriction of the Jacobian matrix to the persuadable nodes can be written as

$$\mathbf{J}_{\mathcal{P}} = \mathbf{S}^{-1}[(\mathbf{W}_{\mathcal{P}} - \mathbf{S}_{\mathcal{P}}) - 2\gamma(\mathbf{Q}_{\mathcal{P}} - \mathbf{R}_{\mathcal{P}})].$$

By checking the on-diagonal and off-diagonal entries of this matrix, we can verify that it matches the Jacobian expressions we found in Equations 2.2 and 2.3. We now define $\mathbf{M}_{\mathcal{P}} = [(\mathbf{W}_{\mathcal{P}} - \mathbf{S}_{\mathcal{P}}) - 2\gamma(\mathbf{Q}_{\mathcal{P}} - \mathbf{R}_{\mathcal{P}})]$, so that

$$\mathbf{J}_{\mathcal{P}} = \mathbf{S}^{-1}\mathbf{M}_{\mathcal{P}}.$$

Now, since $\mathbf{S}^{-1}$ is a diagonal matrix with positive entries, it has a square root $\mathbf{S}^{-1/2}$. Then, since $\mathbf{S}^{-1/2}$ and $\mathbf{M}_{\mathcal{P}}$ are both symmetric matrices, they commute with each other. Altogether, this means we have

$$\mathbf{J}_{\mathcal{P}} = \mathbf{S}^{-1/2}\mathbf{M}_{\mathcal{P}}\mathbf{S}^{-1/2}.$$

Thus, $\mathbf{J}_{\mathcal{P}}$ is similar to $\mathbf{M}_{\mathcal{P}}$. Since similar matrices have the same eigenvalue spectra, we can look at the eigenvalues of $\mathbf{M}_{\mathcal{P}}$ to determine the stability of a stationary state. We will use this $\mathbf{M}_{\mathcal{P}}$ matrix to determine stability on the complete graph in Theorem 3.1.

## 2.3 Balanced Exposure and the Harmonic Solution

In Brooks and Chodrow (2022), a special class of graphs are defined that have a property they call "balanced exposure". We say that a graph meets the balanced exposure condition if it has exactly two zealots, and each persuadable node is either connected to neither zealot or both of them. This will be useful to us when we analyze the complete graph with two zealots in Chapter 3 because that graph meets the balanced exposure condition. As they do in their paper, we assume that the two zealots in a balanced exposure graph are located at the extremes of the opinion spectrum (that is, at 1 and $-1$).

The following result involves the existence of a stationary state in a balanced exposure graph and a characterization of its stability.

**Theorem 2.1** (Brooks and Chodrow)**.** *Let* $\{\mathcal{N}, \mathcal{E}\}$ *be a graph with balanced exposure. For any* $\gamma$, $\mathbf{x}^* = \mathbf{0}$ *is a stationary state of* $\mathbf{F}$. *This state is linearly stable if and only if*

$$\frac{2\gamma e^{\gamma(1-\delta)}}{1 + e^{\gamma(1-\delta)}} < 1.$$

This theorem proves that $\mathbf{x}^* = \mathbf{0}$ is always a stationary state of a balanced exposure graph and provides a condition for when that state is stable. In Sections 3.3 and 3.4, we will be analyzing the stability of stationary states for a special case of balanced exposure (the complete graph). There, we will take advantage of the structure of the complete graph and extend this theorem in such a way that we characterize the existence and stability of a whole class of stationary states that includes $\mathbf{x}^* = \mathbf{0}$ and extends beyond it.

# Chapter 3

# Consensus on the Complete Graph

In this chapter, we explore the conditions under which consensus forms on a complete graph with two zealots and when this consensus is stable to perturbations of the persuadable agents' opinions.

## 3.1 Why the Complete Graph?

The complete graph is a network structure that merits individual study. It is a natural network to investigate the smoothed bounded-confidence model on for many reasons. From an abstract mathematical perspective, its highly structured nature lends itself nicely to the eigenvalue analysis we will perform to characterize the stability of its stationary opinion distributions. Through an opinion dynamics lens, the complete graph represents a group of agents who are all sharing information with each other, which is a natural scheme for the sharing of information in a group of people. We could use a complete graph to represent the dynamics of a group of friends talking, of participants in a debate, or of a discussion on an online forum.

Recall from 2.1 that the smoothed bounded-confidence model of opinion dynamics can only converge to consensus on any connected graph unless the system includes at least two zealots. To that end, we incorporate two zealots into our complete graph, one at each extreme of the opinion spectrum. We will name the zealots $Z_1$ and $Z_2$, with $Z_1$ fixed at opinion $-1$ and $Z_2$ at 1.

As we will see later on in this chapter, the size of the complete graph does not affect the stationarity and stability of consensus opinion states. It has some effect on convergence time, but we will not be analyzing convergence time here. This is why we will often discuss our system in this chapter without specifying the size of the complete graph.

## 3.2 Why Study Consensus?

As we try to understand the patterns of opinions that form in real networks of people, an important (and natural) question arises: will consensus form? In our daily lives, in a variety of settings and scales, groups of people come to agreement and groups of people split into factions. We investigate the conditions under which consensus and fragmentation occur not only because it is interesting to study but because we are often in environments where we are either actively pursuing consensus (e.g., a group presentation) or actively pursuing faction formation (e.g., seeking a diversity of intellectual positions in a philosophical debate), and this provides insights into how the desired behavior might be cultivated. We now look to our model to understand these conditions for consensus and fragmentation.

**Definition 3.2.1.** *Consensus.* We say that an opinion profile $\mathbf{x}$ is at consensus if all persuadable nodes in the network have the same opinion (that is, $x_i = x_j$ for all $X_i, X_j \in \mathcal{P}$).

The zealots are excluded from our definition of consensus because otherwise consensus would be impossible achieve. This calls attention to the seemingly paradoxical fact that we are beginning our analysis of this system with opinion profiles at consensus, despite introducing zealots to prevent the persuadable agents from always converging to consensus. The difference now is that when agents manage to reach consensus it is despite the zealots' competing influences, which makes the analysis of consensus dynamics with zealots far more interesting.

## 3.3 Stationary Consensus in the Complete Graph

Now, we characterize the stationarity of opinion profiles with consensus.

Recall that at any given time, the opinion $x_i$ of a persuadable agent $i$ is subject to the dynamics

$$\frac{dx_i}{dt} = \frac{\sum_{j \in \mathcal{N}} w(x_i, x_j)(x_j - x_i)}{\sum_{j \in \mathcal{N}} w(x_i, x_j)},$$

where $w(x_i, x_j)$ is the weighting function

$$w(x_i, x_j) = \frac{1}{1 + e^{-\gamma(x_i - x_j)^2 - \gamma\delta}}.$$

(Note that while the general definition of the weighting function from Equation 1.1 is conditionally defined depending on whether $i$ and $j$ are adjacent, all nodes are adjacent in the complete graph so we condense the conditional definition here.)

Consider a consensus opinion distribution $\mathbf{x}^*$ with $x_i^* = k$ for all $i \in \mathcal{P}$. For such a solution to be stationary requires that

$$0 = \frac{\sum_{j \in \mathcal{N}} w(k, x_j)(x_j - k)}{\sum_{j \in \mathcal{N}} w(k, x_j)}.$$

Since all of the persuadable nodes have opinion $k$, the only non-zero terms in the numerator summation come from the zealots. After considering this and multiplying through by the denominator we end up with our final condition for a stationary consensus at $k$:

$$0 = w(k, 1)(1 - k) + w(k, -1)(-1 - k). \tag{3.1}$$

This stationary condition is essentially just checking whether the influence that $Z_1$ has over the consensus opinion (given by $w(k, 1)(1 - k)$) is canceled out by the influence that $Z_2$ has over the consensus opinion (given by $w(k, -1)(-1 - k)$.

As we discussed in Section 2.3, a consensus opinion at $\mathbf{x} = \mathbf{0}$ is always stationary under balanced exposure (every agent is connected to both zealots or neither of them). Since the complete graph certainly has balanced exposure, we should expect a stationary consensus distribution at $\mathbf{x} = \mathbf{0}$. This is verified by observing that when $k = 0$, we have $w(k, 1) = w(k, -1)$ regardless of the values of $\gamma$ and $\delta$, and thus the stationary condition is always met.

Now, we describe the non-zero opinion values at which consensus is stationary. At any opinion value, there are two competing factors that balance the influence of the zealots. On the one hand, if an agent's opinion value is close to a zealot's, it means that the weighting function between that agent and that zealot is relatively high. On the other hand, it means that the difference in opinion is low, which weakens the strength of the pull that comes from the $(x_i - x_j)$ term. Similarly, being far from a zealot will

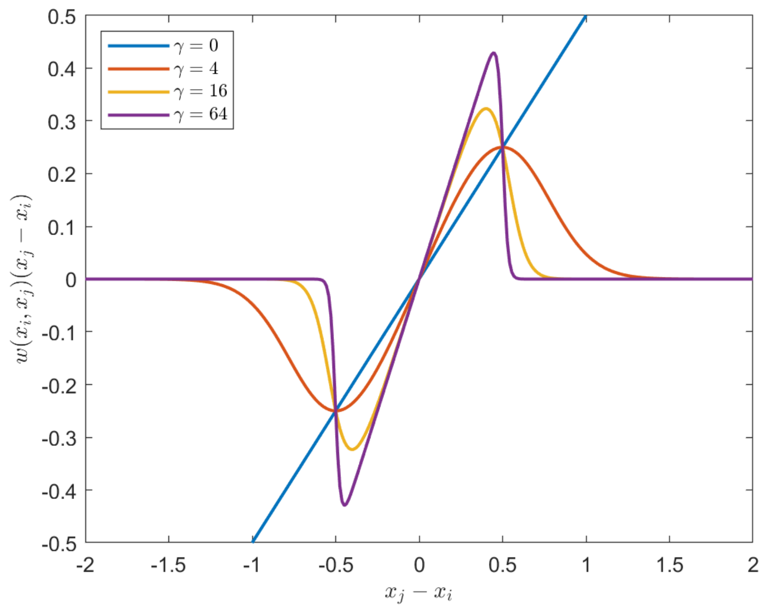


**Figure 3.1** The $x_j$ contribution to the weighted average that governs the dynamics of $x_i$ for $\delta = 0.25$ (note that the contribution of $x_j$ is the greatest near $|x_j - x_i| = \sqrt{\delta} = 0.5$)

lower the effect of the weighting function, but will increase the effect of the $(x_i - x_j)$ term. The change in this $(x_i - x_j)$ term is linear in $k$, but the change in the weighting function is not, which creates a pull strength profile that is strongest at mid-range distances (near where $(x_i - x_j)^2 = \sqrt{\delta}$) and weaker at short or long distance, as pictured in Figure 3.1. This non-monotonic profile means that the pulls of the two zealots can cancel each other out at multiple values of $k$.

To see how this non-monotonic influence profile can create multiple stationary states, we will look at an example. Figure 3.2 illustrates the dynamics of consensus when $\gamma = 2$ and $\delta = 2$ by plotting the numerator of Equation 2.2. For this parameter combination, we see that there are five opinion values at which consensus is stationary (the five values of $k$ for which $dx_i/dt = 0$). This version of the figure conveys where the stationary states are, but it doesn't give a clear picture of how the locations of the stationary states materialize from the zealots' competing influences. To gain that insight, we look to Figure 3.3, which decomposes the numerator of the

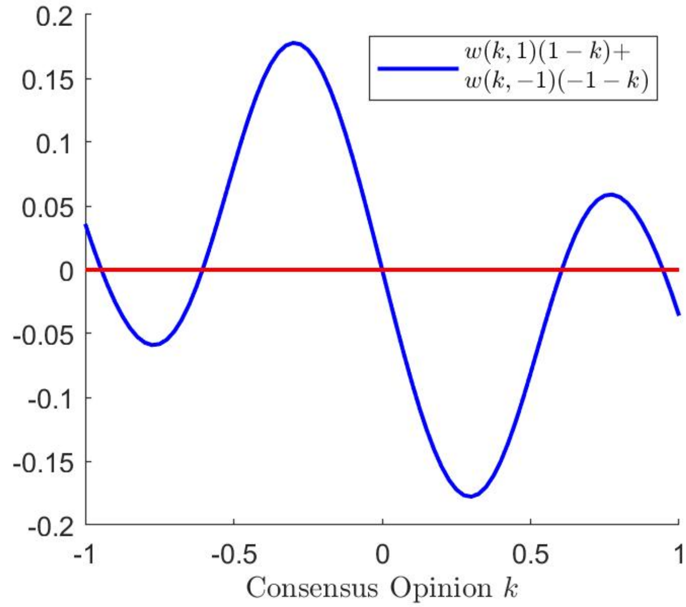


**Figure 3.2** Finding opinions with stationary consensus for $\gamma = 2$ and $\delta = 2$. The vertical axis plots the stationarity condition from 3.1.

dynamics term into the contribution from each zealot. Then, the stationary states are the values of $k$ where the influences of the two zealots are equal. We can see how the nonlinearity of the zealots' influences are what allows for multiple function intersections, and thus multiple stationary consensus points.

## 3.4 Stability of Consensus in the Complete Graph

Now that equation (3.1) provides an expression for where the stationary consensus values are, we will start to characterize their stability. In Figure 3.2, we can tell which of the stationary states are stable against perturbations in $k$ by looking at the sign of

$$\frac{d}{dk}\left[\frac{dx_i}{dt}\right].$$

For instance, consider the stationary state at $k = 0$. The $k-$derivative of $dx_i/dt$ there is negative, which means that a small perturbation of $k$ towards

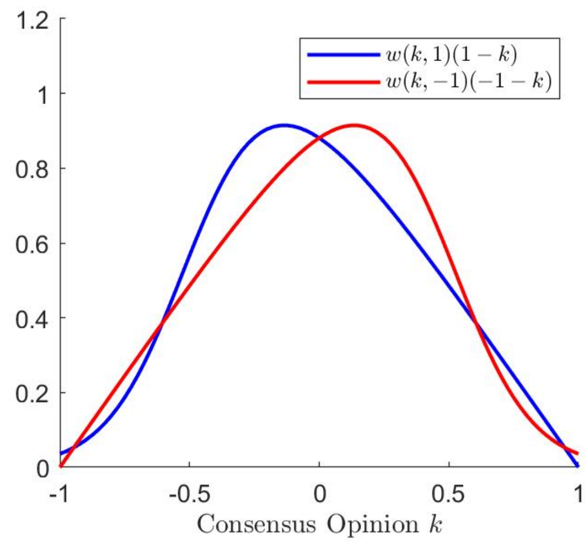


**Figure 3.3** Decomposing the stationarity condition from Equation 3.1 into the $Z_1$ and $Z_2$ contributions. The function intersections correspond to stationary consensus profiles. Here, $\gamma = 2$ and $\delta = 2$.

$-1$ would result in a consensus opinion that would increase. Similarly, a small perturbation towards 1 would result in a consensus opinion that wants to decrease. This means that the stationary state at $k = 0$ is stable for this combination of $\gamma$ and $\delta$. Figure 3.3, which decomposes the dynamics into the influence of each zealot, provides us another way to understand this stability. We notice that for values of $k$ slightly less than 0, the influence from the zealot at 1 is greater than the influence from the zealot at -1. Similarly, when $k$ is slightly greater than 0, the influence from the zealot at -1 is greater. Together, this means that small perturbations in $k$ will be corrected by the imbalances in the zealot influence.

Analyzing the $k$-derivative of $dx_i/dt$ at stationary states provides a good intuition for the stability analysis that we will now perform, but is not sufficient to prove whether a stationary state is stable against all perturbations. When we make a plot like the one in Figure 3.2, we are completely reducing our problem to a single dimension: the consensus opinion $k$. Therefore, we can only say that these stationary states are stable or unstable against perturbations that preserve the consensus. After any other perturbation, our one-dimensional simplification of the dynamics is no longer valid. We would like to be able to say whether stationary consensus opinions are stable when subjected to perturbations that disrupt their consensus.

To do so, we extend Theorem 2.1 to cover all stationary consensus states on the complete graph.

**Theorem 3.1.** *Let $\{\mathcal{N}, \mathcal{E}\}$ be a complete graph with zealots at* 1 *and* $-1$*. If the consensus distribution* $\mathbf{x} = k\mathbf{1}$ *is stationary, then it is stable if and only if*

$$-(v_1 + v_{-1}) + \gamma[v_1(1 - v_1)(1 - k)^2 + v_{-1}(1 - v_{-1})(1 + k)^2] < 0,$$

*where*

$$v_1 = \frac{1}{1 + e^{\gamma((1-k)^2-\delta)}}$$

*and*

$$v_{-1} = \frac{1}{1 + e^{\gamma((-1-k)^2-\delta)}}.$$

Before we prove this theorem, note that when $k = 0$, this condition reduces to $-v + \gamma(2v(1 - v) < 0$, where

$$v = \frac{1}{1 + e^{\gamma(1-\delta)}}.$$

Since $v$ is strictly positive, we divide through by it and the condition turns into $-1 + 2\gamma(1 - v) < 0$, which is the condition that was proven to govern stability at $k = 0$ in Theorem 2.1. Now that we've assured ourselves that the original theorem on the complete graph can be recovered from this new theorem, we will prove the new theorem.

*Proof.* Recall the definition of

$$\mathbf{M}_{\mathcal{P}} = (\mathbf{W}_{\mathcal{P}} - \mathbf{S}_{\mathcal{P}}) - 2\gamma(\mathbf{Q}_{\mathcal{P}} - \mathbf{R}_{\mathcal{P}})$$

from Section 2.2. As mentioned there, since $\mathbf{M}_{\mathcal{P}}$ is similar to $\mathbf{J}_{\mathcal{P}}$, we can characterize the stability of our system with the spectrum of $\mathbf{M}_{\mathcal{P}}$.

Let $u = \frac{1}{1+e^{-\gamma\delta}}$. On the complete graph with $n$ persuadable nodes, we have

$$\begin{aligned}
\mathbf{W}_{\mathcal{P}} &= uJ_n \\
\mathbf{S}_{\mathcal{P}} &= u(nI_n) + (v_1 + v_{-1})I_n \\
\mathbf{Q}_{\mathcal{P}} &= \mathbf{0} \\
\mathbf{R}_{\mathcal{P}} &= (v_1(1 - v_1)(1 - k)^2 + v_{-1}(1 - v_{-1})(-1 - k)^2) \cdot I_n,
\end{aligned}$$

where $I_n$ is the $n \times n$ identity matrix and $J_n$ is the $n \times n$ ones matrix. Then, we have

$$\mathbf{M}_{\mathcal{P}} = u(J_n - nI_n) + (-(v_1 + v_{-1}) + \gamma(v_1(1-v_1)(1-k)^2 + v_{-1}(1-v_{-1})(-1-k)^2))I_n.$$

Since $nI_n - J_n$ is the graph Laplacian of the complete graph, we know that $u(J_n - nI_n)$ is negative semi-definite. Then, if

$$-(v_1 + v_{-1}) + \gamma(v_1(1 - v_1)(1 - k)^2 + v_{-1}(1 - v_{-1})(-1 - k)^2)) \leq 0$$

the second term is negative semi-definite and thus so is $\mathbf{M}_{\mathcal{P}}$. In this case, the stationary consensus is stable.

Now, we assume that

$$-(v_1 + v_{-1}) + \gamma(v_1(1 - v_1)(1 - k)^2 + v_{-1}(1 - v_{-1})(-1 - k)^2)) > 0.$$

Since $(J_n - nI_n)1 = 0$, we observe that

$$\mathbf{M}_{\mathcal{P}}\mathbf{1} = -(v_1 + v_{-1}) + \gamma(v_1(1 - v_1)(1 - k)^2 + v_{-1}(1 - v_{-1})(-1 - k)^2))\mathbf{1}.$$

Thus, $\mathbf{1}$ is an eigenvector of $\mathbf{M}_{\mathcal{P}}$ with a positive eigenvalue and the stationary consensus is unstable. □

## 3.5 A Brief Introduction to Bifurcations

One of the important objectives in studying a nonlinear system is to classify how its dynamics depend on its parameters. Our system depends on the parameters $\gamma$ and $\delta$, and we observe that both our stationarity condition (see Equation 3.1) and our stability condition (see Theorem 3.1) depend on both parameters.

To analyze more meaningfully how our system's dynamics depends on its parameters, we look to bifurcation theory. A brief introduction to bifurcation theory should begin with a definition of what a bifurcation is, but that's easier said than done. There are many texts on nonlinear systems which go into depth on bifurcation theory, including but not limited to Alligood et al. (1996), Guckenheimer and Holmes (2013), and Strogatz (1994). They all give similarly loose definitions of what a bifurcation is, which I paraphrase here.

**Definition 3.5.1.** *Bifurcation point.* As a parameter of a nonlinear system is varied, there are critical values of the parameter at which the structure of

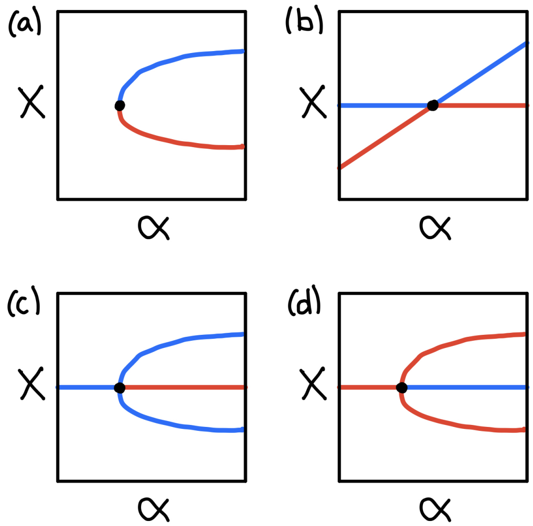


**Figure 3.4** Sketches of the four canonical one-dimensional bifurcations in a parameter $\alpha$: (a) saddle-node, (b) transcritical, (c) supercritical pitchfork, (d) subcritical pitchfork. The blue lines denote stable solutions and the red lines unstable ones.

stationary states in the system qualitatively change (e.g., stationary states appear, disappear, or change stability). One such critical value of the parameter is called a *bifurcation value*, and the point $x^*$ at which a stationary state is appearing, disappearing, or changing stability is called the *bifurcation point*.

Guckenheimer and Holmes (2013) describes why a more formal definition of a bifurcation is difficult. Essentially, it becomes quite complicated to formally define what one means for the flows of the system to have "qualitatively changed." They note that even once a more formal definition is established, it can obscure our understanding of the systems we are actually interested in analyzing.

With that being said, it is well-known that there are four canonical types of bifurcations in a one-dimensional system (displayed in Figure 3.4). There are saddle-node bifurcations, where a pair of stationary states (one stable and one unstable) appear from thin air past a critical value of the parameter. There are transcritical bifurcations, where a pair of stationary states exchange stabilities at a critical parameter value. There are supercritical pitchfork bifurcations, where a stable stationary state becomes unstable and becomes surrounded by a new pair of stable stationary states. Finally, there are subcritical pitchfork bifurcations, which are the same as supercritical pitchforks but with the stabilities reversed.

The diagrams in Figure 3.4 are known as "bifurcation diagrams" and are a helpful way to visualize these critical parameter values. While there are several ways one can present them (see Strogatz (1994) for a few), the structure of this figure is a common convention. In this convention, the system variable is plotted on the vertical axis against the parameter value on the horizontal axis. At every parameter value, we mark the stable and unstable stationary states of the system.

What results is a picture of how the stationary states depend on the parameter value. In these one-dimensional bifurcation diagrams, it's clear to see where the flows of the system lead and how the different canonical bifurcations are qualitatively changing the structure of the system's stationary states. As we analyze consensus in the complete graph, we will see saddle-node bifurcations and both varieties of pitchfork bifurcation.

## 3.6 Understanding Stability with Bifurcation Diagrams

Now that we are equipped with a stationarity condition and a stability condition (Equation 3.1 and Theorem 3.1), we can make bifurcation diagrams for consensus in the complete graph to understand how $\gamma$ and $\delta$ qualitatively affect the system.

Since we have two system parameters, we will be fixing one of them at a time and letting the other vary. For instance, let's fix $\delta = 2$ while letting $\gamma$ vary (as is depicted in Figure 3.5). As is the convention, we place $\gamma$ on the horizontal axis to emphasize that it is the variable being varied. On the vertical axis, then, we plot the $k$ values of our stationary states.

With such a diagram, we start to get a better idea of how the stationary states relate to each other. We observe some qualitative trends for when $\delta = 2$ in Figure 3.5. We see that at small values of $\gamma$, the only stationary consensus happens when $k = 0$, and this consensus remains stable for all $\gamma$. At about $\gamma = 1.75$, new stationary states appear in a pair of saddle-node bifurcations. Two of them are stable and quickly approach the zealot opinions. The other two are unstable and slowly make their way back towards the center. One important thing to note about this bifurcation diagram is that the outermost solution is always stable. This means that every opinion vector must converge—there's no way for it to explode to infinity. This is not true

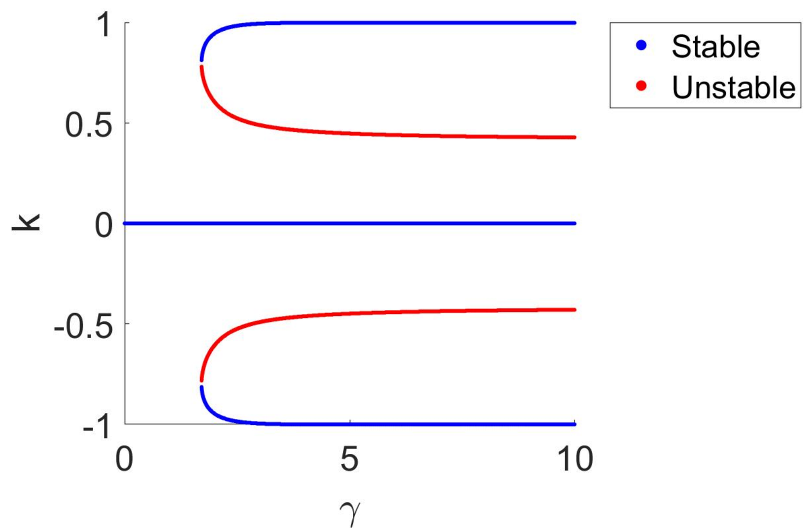


**Figure 3.5** Bifurcation diagram in $\gamma$ with fixed $\delta = 2$. Consensus at zero is always stable, and at $\gamma^* \approx 1.75$ a pair of saddle-node bifurcations introduce four more stationary consensus states.

for all nonlinear systems, but we expect it to be true for these smoothed bounded-confidence diagrams because the smoothed bounded-confidence model always converges (Brooks and Chodrow (2022) proves this using a standard fixed point theorem).

Figure 3.5 has provided us a better idea of how the stationary states behave as a function of $\gamma$ at a particular value of $\delta$ (that is, at $\delta = 2$). We can develop a picture of how our system evolves in the $\delta$ dimension of the parameter space if we start to vary $\delta$ and look at how the bifurcation diagrams themselves change. For instance, in Figure 3.6, we can see what the bifurcation diagram in $\gamma$ looks like for $\delta$ values of 1.6 and 1.5. As $\delta$ gets smaller, parts of the unstable "inner" solutions move towards 0 until they eventually meet the solution there, introducing a pair of subcritical pitchfork bifurcations and creating a region of $\gamma$ for which a 0 consensus is unstable.

In fact, the three bifurcation diagrams that we've seen (see Figures 3.5 and 3.6) are part of a two-dimensional surface that describes the stationary consensus states as a function of both $\delta$ and $\gamma$. These bifurcation diagrams are slices of this surface at particular values of $\delta$. Figure 3.7 shows a numerical approximation of what this surface looks like and shows how these three bifurcation diagrams fit into that surface. Figure 3.8 shows an overhead

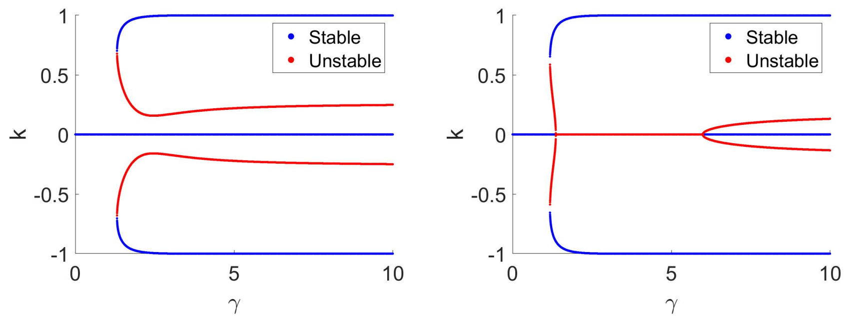


**Figure 3.6** A $\gamma$ bifurcation diagram with $\delta = 1.6$ (left) and one with $\delta = 1.4$ (right). At some critical value of $\delta$ between 1.4 and 1.6, the structure of the bifurcation diagram undergoes qualitative changes.

view of the bifurcation surface, marking how many stationary states exist in each region of the parameter space. On every boundary of this region plot, we find a bifurcation of our system. For instance, consider the boundary separating the region with one stationary state and the region with five stationary states. Crossing that boundary in parameter space corresponds to the location of a pair of saddle-node bifurcations. Crossing the 1 and 3 boundary corresponds to a supercritical pitchfork and crossing the 3 and 5 boundary corresponds to a subcritical pitchfork.

## 3.7 Conclusions about Consensus

Now that we've used bifurcation diagrams to visualize the stationarity points of our system and their stability, it's time to tie these results back to our model formulation.

From a modeling perspective, the stable stationary states are of more interest to us because they are what simulations of such a model could actually converge to. In terms of stable stationary consensus, we observe from Figure 3.7 that they are overwhelmingly located at either $k = 0$ or very near $k = \pm 1$. There do appear to be some stable stationary consensus points with other values of $k$ but another look at the bifurcation diagrams in $\gamma$ confirms that these points exist for a very narrow region of $\gamma$ and seem to be more a product of the stable stationary states emerging from a saddle-node bifurcation than a real structural phenomenon.

The three regions in Figure 3.8 then correspond to the following three conditions:

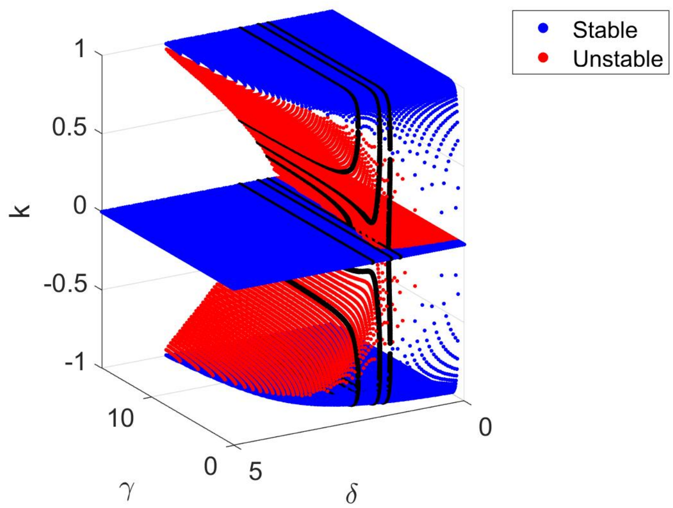


**Figure 3.7** The bifurcation diagram for our system over both $\gamma$ and $\delta$. The surfaces were produced with a mesh size of 1/20 in both $\gamma$ and $\delta$. The black lines mark the locations of the three one-dimensional bifurcation diagrams from Figures 3.5 and 3.6.

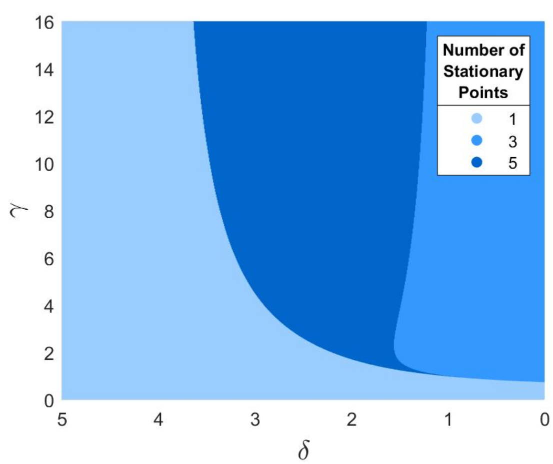


**Figure 3.8** This bifurcation region plot gives an overhead view of the number of stationary states in $(\gamma, \delta)$ space to help visualize the shape of the surface. The boundaries on the interior of this figure correspond to bifurcations of the system (either pairs of saddle-node bifurcations, subcritical pitchforks, or supercritical pitchforks).

1. The only stable consensus opinion is 0.

2. The only stable consensus opinions are $\approx \pm 1$.

3. Consensus opinions of 0 and $\approx \pm 1$ are all stable.

In the Abelson model, the only stationary solution is the harmonic solution $k = 0$. Thus, it makes sense that we see a single stationary state for small values of $\gamma$ no matter what the size of $\delta$ is.

The behavior of the Hegselmann–Krause model is slightly more complicated.

- When $\sqrt{\delta} > 2$, we recover the Abelson model and expect to see a lone stationary state at $k = 0$.

- When $1 < \sqrt{\delta} < 2$, we expect $k = 0$ to remain stable because even after a small perturbation off of $\mathbf{0}$, all of the persuadable nodes would still be receptive to both zealots. In this parameter range, though, there are stationary states at the zealots which remain stable after a perturbation off the zealot because the other zealot remains out of range.

- When $\sqrt{\delta} < 1$, we have the interesting phenomenon that the zealots are no longer in the receptivity range of 0. That is, if we have a consensus at 0 and then perturb the persuadable agents, we do not converge back to a stationary consensus at 0 (unless the perturbation happens to be in the direction of $\langle -1, -1, 1, 1 \rangle$ or something similar). This means that the 0 consensus would be unstable. The stationary consensus points near the zealots, though, would still be stable.

That is, as $\gamma$ approaches infinity in our bifurcation surface, we would expect to see that the interval $\sqrt{\delta} > 2$ lies in the region with one stationary state, the interval $1 < \sqrt{\delta} < 2$ lies in the region with five stationary states, and the interval $0 < \sqrt{\delta} < 1$ lies in the region with three stationary states. Indeed, this is supported by Figure 3.8.

Having verified that we recover the expected stationary states for small $\gamma$ and large $\gamma$, it's time to interpret what's happening at mid-range values of $\gamma$ using Figure 3.8.

- For $0 < \sqrt{\delta} < 1$, nothing much interesting happens at mid-range values of $\gamma$. As discussed before, very small $\gamma$ leads to a stable consensus at 0. Then, beyond some threshold value of $\gamma$, the 0 consensus loses its stability. This threshold value of $\gamma$ doesn't seem to depend on $\sqrt{\delta}$ very strongly.

- On $1 < \sqrt{\delta} < 2$, there's an interesting feature. While $\sqrt{\delta}$ is still near 1 and $\gamma$ is low, the 0 consensus remains unstable despite being within $\sqrt{\delta}$ of both zealots.

  Also on this interval of $\sqrt{\delta}$, the region corresponding to the lone stable stationary consensus at 0 goes from covering the entire interval to a narrow range near $\sqrt{\delta} = 2$. The transition is slow enough that there are values of $\delta$ (consider $\delta \approx 3.5$, for instance) where the zealot stationary states do not exist for relatively high values of gamma despite the fact that the two zealots are not within $\sqrt{\delta}$ of each other.

We have found that the behavior of the smoothed bounded-confidence Model mirrors the behavior of the Abelson model at small $\gamma$ and that of the Hegselmann–Krause model for large $\gamma$, as expected. In the intermediary, the transition between the behaviors is slow enough that there are substantial parts of parameter space where the structure of stationary states defies the intuitive analysis of agents being pulled together when their opinions are closer than $\sqrt{\delta}$.

# Chapter 4

# Dynamics of Two Factions

In the previous chapter, we looked at opinion profiles in consensus and described their stationarity and stability. In this chapter, we relax the consensus condition to analyze the dynamics of two opinion factions on the complete graph.

When dealing with consensus, we were free to omit the size of our complete graph because of the symmetry involved. In our investigation of faction dynamics, we will have to be a bit more careful. Now that we will have persuadable nodes at different opinion values exerting influence on each other, the size of our persuadable pool matters quite a lot. Consider the following–when there are only two persuadable agents, the strength with which one pulls on the other is roughly comparable to the pulls of the zealots. When there are a thousand persuadable agents divided into two opinion factions of five hundred agents each, the pulls of the zealots on any given agent would become essentially negligible relative to the pull of the hundreds of persuadable agents in the other opinion cohort.

## 4.1 Opinion Factions Under Hegselmann–Krause

Before we dive in to the dynamics of the Smoothed-Bounded Confidence model with two opinion factions, we will look at the Hegselmann–Krause model over squared distance. That is, we let the weighting function be

$$w(x_i, x_j) = \begin{cases} 1 & (x_i - x_j)^2 \leq \delta \\ 0 & \text{otherwise.} \end{cases}$$

We are studying the dynamics under this weighting function because it is remarkably similar to the limit of the smoothed bounded-confidence weighting function in the limit as $\gamma \to \infty$. In the limit as $\gamma \to \infty$, the smoothed bounded-confidence weight function becomes a heaviside function

$$w(x_i, x_j) = \begin{cases} 1 & (x_i - x_j)^2 < \delta \\ 1/2 & (x_i - x_j)^2 = \delta \\ 0 & \text{otherwise.} \end{cases}$$

We are studying the Hegselmann–Krause version of the system rather than the $\gamma \to \infty$ version of the system because the inclusion of a "half-weight" at $(x_i - x_j)^2 = \delta$ complicates the analysis and has a negligible impact on the dynamics of the system. For a further discussion of this, see Section 4.3.

The following terminology will be useful to us as we analyze the squared distance Hegselmann–Krause system.

**Definition 4.1.1.** *Receptivity.* For adjacent agents $X_i$ and $X_j$ with opinions $x_i$ and $x_j$ we say that $X_i$ and $X_j$ are receptive to each other if $(x_i - x_j)^2 \leq \delta$. We represent receptivity as $X_i \sim X_j$.

**Definition 4.1.2.** *Receptivity Set of an opinion profile.* For an opinion profile $\mathbf{x}$ on a network with agents $\{X_1, \ldots, X_n\}$, we define the receptivity set $R$ of $\mathbf{x}$ as

$$R(\mathbf{x}) = \{(X_i, X_j) \mid X_i \sim X_j\}.$$

These receptivity sets will be useful to us because they partition the space of opinion profiles and it is easy to define the dynamics of the Hegselmann–Krause system for an opinion profile given its receptivity set.

### 4.1.1 Receptivity Sets of Two Factions

It will help us to define some faction-related terminology that is specific to the complete graph.

We will continue our convention from Chapter 3 of referring to the zealot with opinion $-1$ as $Z_1$ and the zealot with opinion 1 as $Z_2$. We will refer to the two persuadable opinions present in our graph as $x_1$ and $x_2$. Without loss of generality, we will assume that $x_1 < x_2$. We let $\alpha$ be the number of persuadable nodes with opinion $x_1$ and $\beta$ be the number of persuadable nodes with opinion $x_2$.

Although there are many persuadable nodes with opinion $x_1$, the structure of the complete graph means that they are either all receptive to another given node or none of them are. So, we let $P_1$ be a representative agent from the $x_1$ faction and $P_2$ be a representative agent from the $x_2$ faction. Then, we will use the notation $P_1 \sim P_2$ to mean that all nodes from the $x_1$ faction and all nodes from the $x_2$ faction are receptive to each other. (This extends to the zealots, too. We use $P_1 \sim Z_1$ to mean that all of the $x_1$ agents are receptive to $Z_1$.)

This also allows us to use a condensed version of the receptivity set. Rather than list out every pair of receptive nodes, we can use $P_1$ and $P_2$ as representatives of the factions. That is, our receptivity sets will be subsets of

$$\{(Z_1 \sim P_1), (Z_1 \sim P_2), (Z_1 \sim Z_2), (P_1 \sim P_2), (P_2 \sim Z_1), (P_2 \sim Z_2)\}. \tag{4.1}$$

**Theorem 4.1.** *Let $x_1 \leq x_2$ on the complete graph with corresponding (not necessarily distinct) factions $P_1$ and $P_2$. Then, all opinion profiles belong to one of the following nine receptivity sets (defined up to sign-flip symmetry).*

$$\begin{aligned}
R_1 &= \{\emptyset\} \\
R_2 &= \{(Z_1, P_1)\} \text{ or} \\
&= \{(P_2, Z_2)\} \\
R_3 &= \{(Z_1, P_1), (P_2, Z_2)\} \\
R_4 &= \{(Z_1, P_1), (P_1, P_2), (P_2, Z_2)\} \\
R_5 &= \{(Z_1, P_1), (Z_1, P_2), (P_1, P_2), (P_2, Z_2)\} \text{ or} \\
&= \{(Z_1, P_2), (P_1, P_2), (P_1, Z_2), (P_2, Z_2)\} \\
R_6 &= \{(P_1, P_2)\} \\
R_7 &= \{(Z_1, P_1), (Z_1, P_2), (P_1, P_2)\} \text{ or} \\
&= \{(P_1, P_2), (P_1, Z_2), (P_2, Z_2)\} \\
R_8 &= \{(Z_1, P_1), (Z_1, P_2), (P_1, P_2), (P_1, Z_2), (P_2, Z_2)\} \\
R_9 &= \{(Z_1, P_1), (Z_1, P_2), (Z_1, Z_2), (P_1, P_2), (P_1, Z_2), (P_2, Z_2)\} \\
R_{10} &= \{(Z_1, P_1), (P_1, P_2)\} \text{ or} \\
&= \{(P_1, P_2), (P_2, Z_2)\}
\end{aligned}$$

*Of these receptivity sets, the stationary states with distinct opinion factions all belong to $R_1, R_2, R_3, R_4$, or $R_5$ and stationary states with consensus all belong to $R_6, R_7, R_8$, or $R_9$. No stationary states belong to $R_{10}$.*

*Proof.* There are $2^6 = 64$ candidate subsets of the set described in Equation 4.1. However, many of these are impossible for an opinion profile to achieve.

For instance, consider the set element $(Z_1, Z_2)$. Because the distance between the opinions of $Z_1$ and $Z_2$ is always greater than the distance between any other two agents, this element can only appear in the receptivity set if every other pair of agents appears as well. With this type of consideration in mind, we will systematically demonstrate that all stationary states with two factions on the complete graph belong to one of these nine receptivity sets and (along the way) demonstrate that each one corresponds uniquely to stationary states with either consensus or factions.

Throughout this proof, we will make reference to an opinion profile $\mathbf{x}$ and its dynamics $\mathbf{F}(\mathbf{x})$. While technically these vectors are of length $\alpha + \beta$, we are currently only interested in evaluating the stationary states of the dynamics, so we can safely reduce them to the two dimensions

$$\begin{bmatrix} x_1' \\ x_2' \end{bmatrix} = \mathbf{F}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

where $x_1$ and $x_2$ represent the opinions of the factions $P_1$ and $P_2$.

**1. Neither Faction Receptive to a Zealot.** First, we consider an opinion profile where neither faction is receptive to either zealot. Since the zealots must also not be receptive to each other, this must describe either $R_1 = \{\emptyset\}$ or $R_6 = \{(P_1, P_2)\}$.

Any opinion profile with receptivity set $R_1$ has $\mathbf{F}(\mathbf{x}) = \mathbf{0}$ and is therefore a stationary state. Since $P_1 \nsim P_2$, it follows that $x_1 \neq x_2$ and $R_1$ corresponds to stationary states with distinct factions.

Consider now a stationary opinion profile $\mathbf{x}^*$ with receptivity set $R_6$. The dynamics of the persuadable subsystem are

$$\mathbf{F}(\mathbf{x}^*) = \begin{bmatrix} \beta(x_2 - x_1) \\ \alpha(x_1 - x_2) \end{bmatrix}$$

which is equal to $\mathbf{0}$ if and only if $x_1 = x_2$. Thus, $R_6$ corresponds to stationary states at consensus.

**2. Exactly One Faction Receptive to a Zealot.** Let's first assume that $P_1$ is the faction that is receptive to a zealot. If either persuadable faction were receptive to the "opposite" faction, then both factions would have to be receptive to at least one zealot which falls outside of our current scope. So, we will assume that $(Z_1 \sim P_1)$.

To maintain that $P_2$ is receptive to neither zealot requires that $(Z_1 \nsim P_2), (Z_1 \nsim Z_2), (P_1 \nsim Z_2)$, and $(P_2 \nsim Z_2)$. Then, the only remaining receptivity sets are $R_{10} = \{(Z_1, P_1), (P_1, P_2)\}$ and $R_2 = \{(Z_1, P_1)\}$. While $R_{10}$ is a valid receptivity set, it is never achieved by a stationary state.

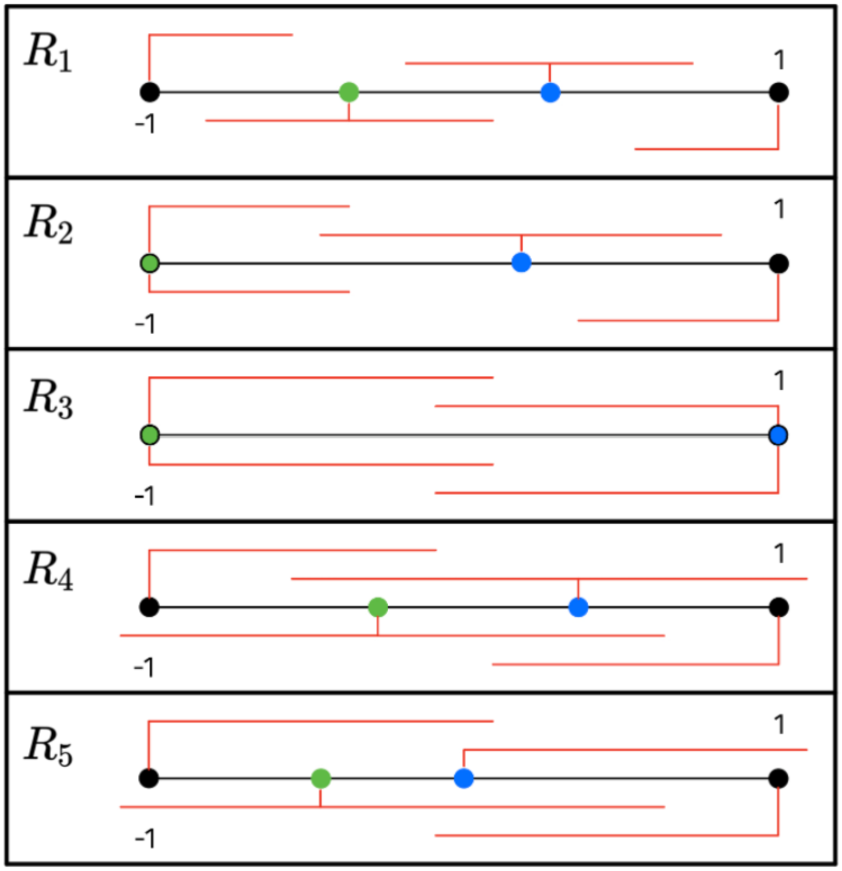


**Figure 4.1** Visualizations of the five receptivity sets which could be present in a stationary state with fragmentation. The examples from cases $R_1$, $R_2$, and $R_3$ are precisely stationary states. We will calculate later in this chapter what stationary states with $R_4$ or $R_5$ precisely look like—the diagrams here are simply to give a qualitative representation. In all five diagrams, $P_1$ is colored green, $P_2$ is colored blue, and intervals of width $2\sqrt{\delta}$ are marked in red and centered on each agent. If an agent lies within another agent's interval, the two are receptive to each other.

For proof of this, consider the dynamics of $\mathbf{F}(\mathbf{x})$ for an opinion profile $\mathbf{x}$ with $R(\mathbf{x}) = R_{10}$. We would have

$$\mathbf{F}(\mathbf{x}) = \begin{bmatrix} (-1 - x_1) + \beta(x_2 - x_1) \\ \alpha(x_1 - x_2) \end{bmatrix}.$$

For this to be stationary would require that $x_1 = x_2 = -1$. However, since $R_{10}$ demands that $(Z_1 \nsim P_2)$ this is impossible.

That leaves us to consider $R_2 = \{(Z_1, P_1)\}$. For an opinion profile $\mathbf{x}$ with $R(\mathbf{x}) = R_2$, the dynamics of the persuadable subsystem are

$$\mathbf{F}(\mathbf{x}) = \begin{bmatrix} (-1 - x_1) \\ 0 \end{bmatrix}.$$

which is stationary for any $x_2$ that satisfies the opinion profile so long as $x_1 = -1$. Since $R_2$ requires that $P_1 \nsim P_2$, it follows that $x_1 \neq x_2$ and thus a stationary opinion profile with $R_2$ has distinct factions.

At the beginning of this section, we assumed that $P_1$ was the only faction receptive to its neighboring zealot. We can repeat the same analysis with $P_2$, arriving at receptivity sets $\{(P_1, P_2), (P_2, Z_2)\}$ and $\{(P_2, Z_2)\}$. We classify these receptivity sets under $R_{10}$ and $R_2$, respectively, because they correspond to the same opinion profiles with the signs of both faction opinions flipped.

**3. Both Factions Receptive to Exactly One Zealot (the same one).** Let $\mathbf{x}$ be an opinion profile where the two factions are receptive to the same zealot. This implies that they are also receptive to each other and not the other zealot. Then, $R(\mathbf{x})$ can only be $R_7 = \{(Z_1, P_1), (Z_1, P_2), (P_1, P_2)\}$.

For an $R_7$ opinion profile to be at stationarity requires that

$$\mathbf{F}(\mathbf{x}) = \begin{bmatrix} (-1 - x_1) + \beta(x_2 - x_1) \\ (-1 - x_2) + \alpha(x_1 - x_2) \end{bmatrix} = \mathbf{0}.$$

This is satisfied only by $x_1 = x_2 = -1$. (Similarly, for the other receptivity set in $R_7$ the only stationary state is $x_1 = x_2 = 1$.) This is a stationary state with consensus.

**4. Both Factions Receptive to Exactly One Zealot (different ones).** Now, we let $\mathbf{x}$ be an opinion profile where each faction is receptive to different zealots. With the restriction that each faction is receptive to only one zealot, the receptivity set of this profile can only be $R_3 = \{(Z_1, P_1), (P_2, Z_2)\}$ or $R_4 = \{(Z_1, P_1), (P_1, P_2), (P_2, Z_2)\}$.

If $R(\mathbf{x}) = R_3$, then the dynamics of the persuadable subsystem are

$$\mathbf{F}(\mathbf{x}) = \begin{bmatrix} (-1 - x_1) \\ (1 - x_2) \end{bmatrix}.$$

This is stationary only when $x_1 = -1$ and $x_2 = 1$ (which makes this a stationary state with distinct factions).

If $R(\mathbf{x}) = R_4$, then the dynamics of the persuadable subsystem are

$$\mathbf{F}(\mathbf{x}) = \begin{bmatrix} (-1 - x_1) + \beta(x_2 - x_1) \\ (1 - x_2) + \alpha(x_1 - x_2) \end{bmatrix}.$$

This stationarity condition requires a bit more algebra but turns out to be

$$\mathbf{x}^* = \frac{1}{1 + \alpha + \beta} \begin{bmatrix} \beta - \alpha - 1 \\ \beta - \alpha + 1 \end{bmatrix},$$

which corresponds to distinct factions for any allowable choice of $\alpha$ and $\beta$.

**5. Exactly One Faction Receptive to Both Zealots.** Let $\mathbf{x}$ be an opinion profile where exactly one faction is receptive to both zealots. Without loss of generality, let $P_1$ be this faction. Note that $P_2$ must be receptive to $Z_2$ because the distance between $P_1$ and $Z_2$ is at least as great as the distance between $P_2$ and $Z_2$. Altogether, the only receptivity set that can represent $\mathbf{x}$ is $R_5 = \{(Z_1, P_2), (P_1, P_2), (P_1, Z_2), (P_2, Z_2)\}$.

The dynamics of the persuadable subsystem here are

$$\mathbf{F}(\mathbf{x}) = \begin{bmatrix} (-1 - x_1) + \beta(x_2 - x_1) + (1 - x_1) \\ (1 - x_2) + \alpha(x_1 - x_2) \end{bmatrix}.$$

For this to be stationary requires that

$$\mathbf{x}^* = \frac{1}{2\beta + \alpha + 2} \begin{bmatrix} -\alpha \\ -(2 + \alpha) \end{bmatrix},$$

which means $\mathbf{x}^*$ has two distinct factions for any choice of $\alpha$ and $\beta$.

If we let $P_2$ be the faction receptive to both zealots, the same analysis applies for a sign-flipped version of $\mathbf{x}^*$.

**6. Both Factions Receptive to Both Zealots.** Let $\mathbf{x}$ be an opinion profile where both factions are receptive to both zealots. It's given, then, that $R(\mathbf{x})$ must have the four $(Z_i \sim P_j)$ receptivity elements. Since $(Z_1 \sim P_2)$ and $(P_1 \sim Z_2)$, we must have $(P_1 \sim P_2)$. Then, $R(\mathbf{x})$ must be either $R_8$ or $R_9$. Either way, the dynamics of the persuadable subsystem are

$$\mathbf{F}(\mathbf{x}) = \begin{bmatrix} (-1 - x_1) + \beta(x_2 - x_1) + (1 - x_1) \\ (-1 - x_2) + \alpha(x_1 - x_2) + (1 - x_2) \end{bmatrix}.$$

Then, the only stationary state $\mathbf{x}^*$ with $R(\mathbf{x}^*) = R_8$ or $R(\mathbf{x}^*) = R_9$ is $x_1 = x_2 = 0$. This is a stationary state with consensus.

□

To visualize what the regions associated with these receptivity sets actually look like on the phase plane for various values of $\sqrt{\delta}$, look to Figure 4.2. Something important to notice is that not every receptivity set is present for any given value of $\sqrt{\delta}$. This means that although every region besides $R_{10}$ is *capable* of having a stationary state inside of it, that point is only actually stationary for values of $\sqrt{\delta}$ where it falls in the correct region. In the next subsection, we characterize when this occurs.

### 4.1.2 Stationary States of Two Factions

We will go through all of the receptivity sets and briefly describe for which values of $\sqrt{\delta}$ their stationary states exist. During this section, we will relax the condition that $x_1 \leq x_2$ to reflect the full reality of the phase plane. All the analysis we've done so far applies for $x_2 < x_1$ but with flipped faction labels.

#### Existence of Stationary States for $R(\mathbf{x}) \neq R_4, R_5$

We will come back to $R_4$ and $R_5$ because their conditions are more complicated.

First, consider $R_1 = \{\emptyset\}$. As we showed in section 1 of the proof of Theorem 4.1, the entire $R_1$ region consists of stationary states whenever it exists. This region exists whenever $\sqrt{\delta}$ is small enough that all four of our agents can avoid being receptive to each other (i.e., when $\sqrt{\delta} < 2/3$).

Now, $R_2 = \{(Z_1, P_1)\}$ or $\{(P_2, Z_2)\}$. We found in section 2 of Theorem 4.1 that any opinion profile with $R_2$ is stationary so long as $x_1 = -1$ or $x_2 = 1$, respectively. Now that we allow $x_2 < x_1$, there are two more symmetric sets of stationary states (one where $x_1 = 1$ and one where $x_2 = -1$). These four symmetric sets of stationary states exist so long as $\sqrt{\delta} < 1$. Otherwise, there's no way for one of the factions to avoid being receptive to both zealots.

For $R_3 = \{(Z_1, P_1), (P_2, Z_2)\}$ we found in section 4 of Theorem 4.1 that the only possible stationary state is $(x_1, x_2) = (-1, 1)$. On the full phase plane there also exists $(x_1, x_2) = (1, -1)$. Both of these stationary states exist for all $\sqrt{\delta} < 2$.

The $R_6$ region is present for all $\sqrt{\delta} < 1$ and always contains the origin $x_1 = x_2 = 0$, which is its corresponding stationary state.

The $R_7$ region is present for all $\sqrt{\delta} < 2$ and always contains the points $(-1, -1)$ and $(1, 1)$, which are its stationary states.

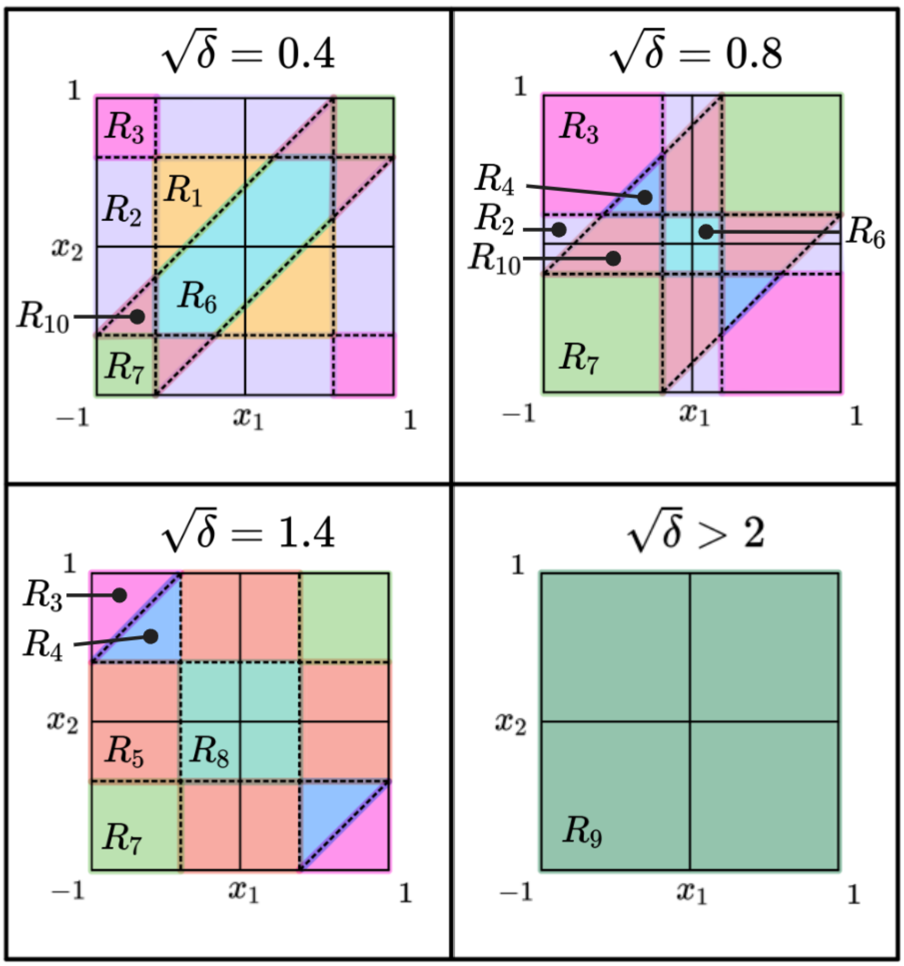


**Figure 4.2** Visualizing the receptivity sets as regions in phase space (with the relaxation that $x_1$ may be larger than $x_2$).

The $R_8$ region replaces the $R_6$ region for $1 \leq \sqrt{\delta} < 2$ and always contains the origin $x_1 = x_2 = 0$, which is its corresponding stationary state.

The $R_9$ region replaces the $R_8$ region for $\sqrt{\delta} \geq 2$ and always contains the origin, which is its corresponding stationary state.

**Existence of Stationary States for $R_4$ and $R_5$**

The existence of $R_4$ and $R_5$ stationary states is more complicated because the location of the stationary states depends on both $\alpha$ and $\beta$.

In section 4 of Theorem 4.1 we found that one of the four $R_4$ stationary states is at the coordinates

$$\mathbf{x}^* = \frac{1}{1+\alpha+\beta}\begin{bmatrix}\beta-\alpha-1\\ \beta-\alpha+1\end{bmatrix}.$$

For this point in phase space to actually fall into an $R_4$ region requires that $x_1 < -|1-\sqrt{\delta}|$ and $x_2 > |1-\sqrt{\delta}|$ and $x_2 - x_1 < \sqrt{\delta}$. We will further analyze this case in the $\alpha = \beta$ case in Section 4.2.

In section 5 of Theorem 4.1 we also established that one of the four stationary states in an $R_5$ region has coordinates

$$\mathbf{x}^* = \frac{1}{2\beta+\alpha+2}\begin{bmatrix}-\alpha\\ -(2+\alpha)\end{bmatrix}.$$

This point in phase space will actually be in a region corresponding to $R_5$ if $P_2$ is receptive to both zealots but $P_1$ is not receptive to $Z_2$.

The condition that $P_1$ is not receptive to $Z_2$ is given by

$$\sqrt{\delta} < 1 + \frac{2+\alpha}{2\beta+\alpha+2}.$$

The condition that $P_2$ is receptive to $Z_1$ is given by

$$\sqrt{\delta} \geq 1 + \frac{\alpha}{2\beta+\alpha+2}.$$

This means that $\mathbf{x}^*$ exists when

$$1 + \frac{\alpha}{2\beta+\alpha+2} \leq \sqrt{\delta} < 1 + \frac{2+\alpha}{2\beta+\alpha+2}. \tag{4.2}$$

Thus, the interval of $\sqrt{\delta}$ values for which this stationary state actually exists has a width of

$$\frac{2}{2\beta+\alpha+2}.$$

As $\beta$ and $\alpha$ increase, this very quickly approaches zero.

This makes sense, because when $\beta$ and $\alpha$ get large, the zealots become more and more negligible. The more negligible the pull of the zealots, the closer together the opinions of the factions must be to remain stationary. This requires a more and more precise value of $\sqrt{\delta}$ to keep $P_2$ receptive to both zealots without letting $P_1$ become receptive to $Z_2$.

## 4.2 Equally-Sized Factions in Hegselmann–Krause

In this section, we consider the dynamics of equally-sized factions (i.e., we let $\beta = \alpha$). This produces some simpler analysis and clearer visualizations while still allowing us to learn about how the system behavior changes as the factions grow large. In particular, this makes much simpler the description of the stationary states from regions $R_4$ and $R_5$.

### 4.2.1 Equal Faction Stationary States with $R_4$ Receptivity

For unbalanced factions, we found that a stationary state $\mathbf{x}^*$ with $R(\mathbf{x}^*) = R_4$ would have coordinates

$$\mathbf{x}^* = \frac{1}{1+\alpha+\beta}\begin{bmatrix}\beta-\alpha-1\\ \beta-\alpha+1\end{bmatrix}.$$

Now, we can simplify this to

$$\mathbf{x}^* = \frac{1}{1+2\alpha}\begin{bmatrix}-1\\ 1\end{bmatrix}.$$

Importantly, this means that the $R_4$ stationary states for equally-sized factions are sign-flip symmetric (that is, $x_1 = -x_2$). For such a stationary state, the conditions for falling within the $R_4$ region become

$$\frac{-1}{1+2\alpha} < -|1-\sqrt{\delta}|$$

and

$$\frac{2}{1+2\alpha} < \sqrt{\delta}.$$

By noting that $\alpha \geq 1$ and considering the cases $\sqrt{\delta} < 1$, $\sqrt{\delta} = 1$, and $\sqrt{\delta} > 1$ separately we are able to get rid of the absolute value and further simplify the conditions to

$$\frac{2}{1+2\alpha} < \sqrt{\delta} < \frac{2+2\alpha}{1+2\alpha}. \tag{4.3}$$

This is analogous to the interval condition we were able to establish for $R_5$ in Equation 4.2, which means we can now more easily compare the two.

Note that $\sqrt{\delta} = 1$ satisfies the $R_4$ stationary state condition for any $\alpha$. This is because the stationary state is approaching $(0, 0)$ as $\alpha \to \infty$, but $x_1$ is always negative and $x_2$ is always positive. That means that no matter how close to 0 the opinions get, it will never be close enough that $P_1$ becomes receptive to $Z_2$ (or $P_2$ to $Z_1$) when $\sqrt{\delta} = 1$.

### 4.2.2 Equal Faction Stationary States with $R_5$ Receptivity

For unbalanced factions, we found that a stationary state $\mathbf{x}^*$ with $R(\mathbf{x}^*) = R_5$ would have coordinates

$$\mathbf{x}^* = \frac{1}{2\beta + \alpha + 2} \begin{bmatrix} -\alpha \\ -(2+\alpha) \end{bmatrix}.$$

Now, we can simplify this to

$$\mathbf{x}^* = \frac{1}{3\alpha + 2} \begin{bmatrix} -\alpha \\ -(2+\alpha) \end{bmatrix}.$$

We already established an interval of $\sqrt{\delta}$ for which the $R_5$ stationary state exists in Subsection 4.1.2 (see Equation 4.2). We let $\beta = \alpha$ and the $\sqrt{\delta}$ interval becomes

$$\frac{4\alpha + 2}{3\alpha + 2} \leq \sqrt{\delta} < \frac{4\alpha + 4}{3\alpha + 2}. \tag{4.4}$$

Note that now that our factions are equally-sized, $\sqrt{\delta} = 4/3$ statisfies the $R_5$ condition for any selection of $\alpha$. This is because this $R_5$ stationary state approaches $(-1/3, -1/3)$ as $\alpha \to \infty$. Since $x_1$ never exceeds $-1/3$ and $x_2$ never falls below $-1/3$, the $R_5$ stationary state coordinates do fall in $R_5$ for any $\alpha$ when $\sqrt{\delta} = 4/3$.

### 4.2.3 Visualizing Existence of Stationary States

Now that Equations 4.3 and 4.4 give us clear $\sqrt{\delta}$ intervals in terms of $\alpha$ where $R_4$ and $R_5$ stationary states exist, we visualize how these intervals evolve with $\alpha$ using Figure 4.3.

We observe that the $\sqrt{\delta}$ intervals do briefly overlap for $\alpha = 1$, but not for any larger value of $\alpha$. This is because (as we noted in the previous section) the two families of stationary states are approaching different consensus

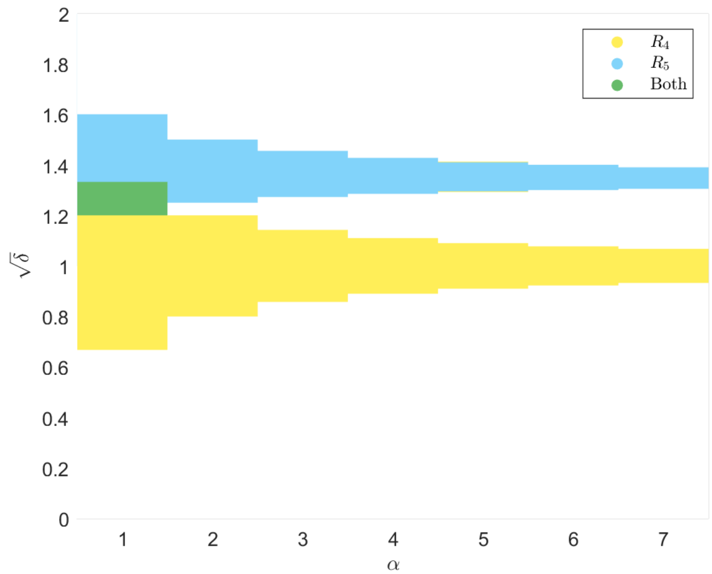


**Figure 4.3** The $\sqrt{\delta}$ intervals for which $R_4$ and $R_5$ stationary states exist (as a function of $\alpha$).

points as $\alpha \to \infty$. So, even though both of their intervals are shrinking, the $R_4$ interval is shrinking towards 1 and the $R_5$ interval is shrinking towards $4/3$.

To visualize what it looks like for a phase plane to have both $R_4$ and $R_5$ stationary states, we look to Figure 4.4. There, we can see how selecting a $\sqrt{\delta}$ value between $6/5$ and $4/3$ results in the $x_1$ and $x_2$ nullclines intersecting in both the $R_4$ and $R_5$ regions. Note that they also intersect in the $R_3$, $R_7$, and $R_8$ regions, which are the other stationary states we expect to find based on the $\sqrt{\delta}$ conditions we found at the beginning of Subsection 4.1.2.

### 4.2.4 Conclusions about Factions in Hegselmann–Krause

We have completely characterized the possible stationary states with two factions for the Hegselmann-Krause model. After removing those with consensus (that is, the ones where the faction opinions are not actually distinct), there are only five types of stationary states (associated with receptivity sets $R_1, \ldots, R_5$).

Of these five types of stationary states, the first three are simple to

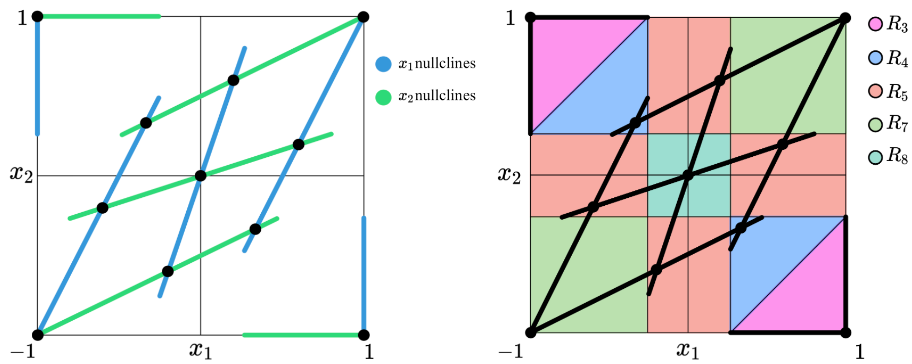


**Figure 4.4** Nullclines for the $\alpha = \beta = 1$ system with $\sqrt{\delta} = 1.25$. The figure on the right overlays the nullclines on the receptivity sets of the phase plane.

understand and do not depend on $\alpha$ or $\beta$ (which means that they behave the same for balanced or unbalanced factions).

In an $R_1$ stationary state, the factions and zealots are simply all too far from each other to have any influence. The $R_1$ region has the unique property that any point inside of it is a stationary state.

In an $R_2$ stationary state, one of the factions is at a zealot opinion and the other is too far from anything to be influenced. A given $R_2$ region does contain an infinite set of stationary states, but they only take up a sliver of the whole region.

In an $R_3$ stationary state, both factions are at opposite zealots.

The $R_4$ and $R_5$ stationary states are a bit more complicated because they exist in the middle of the opinion space and their location is a function of $\alpha$ (and $\beta$, in the unbalanced case). We were able to specify their location and in the $\alpha = \beta$ case we characterized the $\sqrt{\delta}$ intervals for which each of them exists. We found that these intervals become small very quickly as $\alpha$ increases, which means that these stationary states are rather fragile except at small $\alpha$. That being said, for $\alpha = 1$ the $\sqrt{\delta}$ intervals are substantial and they even overlap, leading to a narrow $\sqrt{\delta}$ range where both $R_4$ and $R_5$ stationary states exist (see Figures 4.3 and 4.4).

Without zealots, the Hegselmann-Krause model on the complete graph can only converge to something akin to $R_1$, where every opinion faction is simply too far from the others to have any influence. While the $R_4$ and $R_5$ stationary states don't exist for as wide of a $\delta$ range as the other stationary states, they are proof that the introduction of zealots allows for a type of faction formation where the factions are still being influenced by multiple other opinions, which is qualitatively different from the factions that form in the Hegselmann–Krause model without zealots.

## 4.3 How is Hegselmann–Krause different from $\gamma \to \infty$ smoothed bounded-confidence?

Before returning to the smoothed bounded-confidence model, we will carefully account for how the Hegselmann–Krause model that we have analyzed in Sections 4.1 and 4.2 is different from the limit of the smoothed bounded-confidence model as $\gamma \to \infty$.

As we already noted in Section 4.1, the only difference in dynamics is at points in the phase plane where $(x_i - x_j)^2 = \sqrt{\delta}$ (where $x_i$ and $x_j$ are either faction opinions or zealot opinions). For all other points in the phase plane, the dynamics are exactly the same under the two models. Thus, the only points that need special consideration beyond what we've done already are those who lie on the lines $x_1 = \pm(1 - \sqrt{\delta})$, $x_2 = \pm(1 - \sqrt{\delta})$, and the two lines defined by $\sqrt{\delta} = |x_1 - x_2|$. These are the boundaries that separate the $R_i$ regions, as can be seen in Figure 4.2.

We find that at some particular values of $\sqrt{\delta}$, there are stationary states in the $\gamma \to \infty$ smoothed bounded-confidence model that do not appear in the Hegselmann-Krause model. However, these stationary states are unstable and structurally fragile (only appearing at one precise value of $\delta$)

**Theorem 4.2.** *When $\sqrt{\delta} = 16/13$, there are four stationary states in the $\gamma \to \infty$ smoothed bounded-confidence model that do not appear in the HK model. When $\sqrt{\delta} = 7/4$, there are four. When $\sqrt{\delta} = 4/3$, there are two. For any other $\sqrt{\delta} > 1$, the stationary states of the two models are identical.*

*Proof.* The dynamics of the two models is certainly the same on the interior of any $R_i$ region, which means we must only check the points on the boundaries. We begin by considering the boundary at $x_2 = -1 + \sqrt{\delta}$ (though our analysis will apply to the other three boundaries by symmetry).

According to the smoothed bounded-confidence model with $\gamma \to \infty$, the dynamics of $x_2$ on the line $x_2 = -1 + \sqrt{\delta}$ are

$$x_2' = x_1 + 3 - \frac{5}{2}\sqrt{\delta}.$$

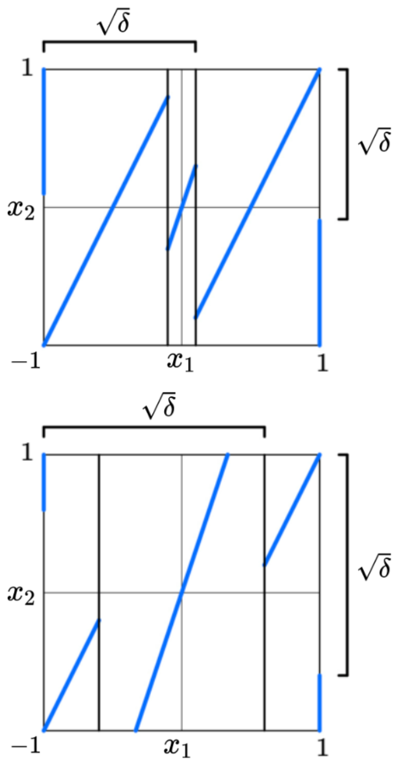


**Figure 4.5** The $x_1$ nullclines for various values of $\sqrt{\delta}$.

For a point to be stationary requires that $x_2' = 0$, so we find that a stationary state on the $x_2 = -1 + \sqrt{\delta}$ boundary would require

$$x_1 = \frac{5}{2}\sqrt{\delta} - 3.$$

Thus, the line

$$\mathbf{x} = \begin{bmatrix} -1 \\ -3 \end{bmatrix} + \sqrt{\delta} \begin{bmatrix} 1 \\ 5/2 \end{bmatrix} \tag{4.5}$$

(given $\sqrt{\delta} > 1$) describes the location of the point on the $-1 + \sqrt{\delta}$ boundary where $x_2' = 0$, as a function of $\sqrt{\delta}$.

Now, we need to find where on this line $x_1' = 0$. To do so, we first check whether it intersects any of the $x_1$ nullclines. The $x_1$ nullclines do change with $\sqrt{\delta}$, but they always consist of vertical line segments at $x_1 = \pm 1$, lines of slope 2 passing through $(-1, -1)$ and $(1, 1)$ wherever $P_1$ is receptive to

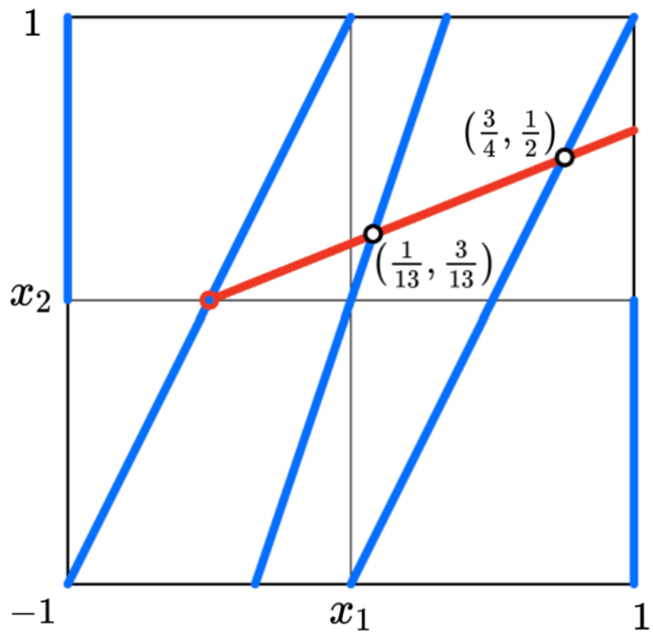


**Figure 4.6** The line of nullpoints on the $x_2 = -1+\sqrt{\delta}$ boundary (parametrized in $\sqrt{\delta}$). The blue lines represent all points that are part of an $x_1$ nullcline for some value of $\sqrt{\delta} > 1$.

only one zealot, and a line of slope 3 passing through $(0,0)$ wherever $P_1$ is receptive to both zealots (see: Figure 4.5).

For the moment, we sketch all the possible points that are part of an $x_1$ nullcline for some $\sqrt{\delta} > 1$ and check to see if they cross the line from Equation 4.5 (see Figure 4.6). We find that there are indeed two intersections, one at $(1/13, 3/13)$ and one at $(3/4, 1/2)$. Now, for these to actually be stationary states requires that the $x_1$ nullcline involved in the intersection actually exists at those coordinates, for the proper $\sqrt{\delta}$ value.

We confirm that when $\sqrt{\delta} = 16/13$, $P_1$ is receptive to both zealots at $(1/13, 3/13)$. Similarly, when $\sqrt{\delta} = 7/4$, $P_1$ is receptive only to $Z_1$ at $(3/4, 1/2)$. Thus, $(1/13, 3/13)$ is a stationary state for $\sqrt{\delta} = 16/13$ and $(3/4, 1/2)$ is a stationary state for $\sqrt{\delta} = 7/4$.

The same analysis applies on the boundaries $x_2 = 1 - \sqrt{\delta}$ and $x_1 = \pm(-1 + \sqrt{\delta})$. This leads to three other stationary states at $\sqrt{\delta} = 16/13$ (those being $(3/13, 1/13)$, $(-1/13, -3/13)$, and $(-3/13, -1/13)$) and three

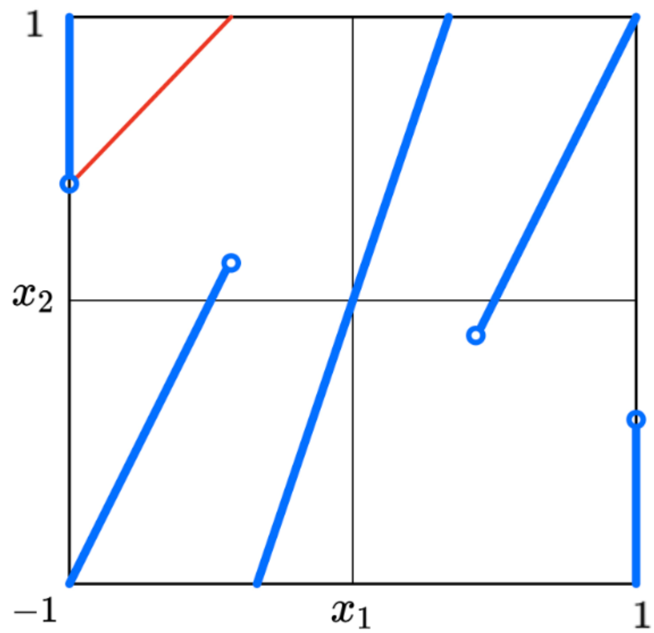


**Figure 4.7** The $x_1$ nullclines of the $\gamma \to \infty$ smoothed bounded-confidence model for a particular value of $\sqrt{\delta}$. The $x_1$ nullclines of the $\gamma \to \infty$ smoothed bounded-confidence model (pictured in blue) never intersect the line $x_2 = x_1 + \sqrt{\delta}$ (pictured in red).

other stationary states at $\sqrt{\delta} = 7/4$ (those being $(1/2, 3/4)$, $(-3/4, -1/2)$, and $(-1/2, -3/4)$).

A stationary state could also appear if an $x_1$ nullpoint intersects with an $x_2$ nullpoint. By the symmetry of the system, we can tell that this will happen when the line from Equation 4.5 intersects with its inverse

$$\mathbf{x} = \begin{bmatrix} -3 \\ -1 \end{bmatrix} + \sqrt{\delta} \begin{bmatrix} 5/2 \\ 1 \end{bmatrix}.$$

This occurs at $(1/3, 1/3)$, when $\sqrt{\delta} = 4/3$. Therefore, when $\sqrt{\delta} = 4/3$ there is a stationary state at $(1/3, 1/3)$ where both the $x_1$ and $x_2$ coordinates fall on an $R_i$ boundary. Note that since this point falls on the line $x_1 = x_2$, it has only one symmetric partner, at $(-1/3, -1/3)$.

Now all that remains is to consider stationary states that lie on the boundaries at $|x_2 - x_1| = \sqrt{\delta}$. Consider first the boundary at $x_2 = x_1 + \sqrt{\delta}$. We observe from Figure 4.7 that when $\sqrt{\delta} > 1$, this boundary never intersects any of the $x_1$ nullclines. (In Figure 4.7, it looks like $x_1$ might be stationary at $(x_1, x_2) = (-1, -1+\sqrt{\delta})$, but that $x_1$ nullcline disappears before the boundary.)

Since this boundary never intersects a nullcline, its only hope for containing a stationary state is if an $x_1$ and $x_2$ nullpoint that are both on the boundary intersect.

We find that on the boundary $x_2 = x_1 + \sqrt{\delta}$, the condition $x_2' = 0$ implies that $x_1 = 1-(3/2)\sqrt{\delta}$ and the condition $x_1' = 0$ implies that $x_2 = -1+(3/2)\sqrt{\delta}$. The only $\sqrt{\delta}$ for which these coordinates lie on the given boundary is $\sqrt{\delta} = 1$, which falls outside of our current scope.

We have found that for $\sqrt{\delta} > 1$ there are only ten stationary states in the $\gamma \to \infty$ limit of the smoothed bounded-confidence model that do not appear in the Hegselmann-Krause model. They appear in two sets of four symmetric states and one set of two states, and each of the sets are only present for one particular value of $\sqrt{\delta}$. □

A similar analysis could be performed for $\sqrt{\delta} = 1$ and $\sqrt{\delta} < 1$. The lesson here, though, is that the additional stationary states that are generated by the inclusion of the 1/2 weight at a distance of $\sqrt{\delta}$ are not significant because there are only a handful of them and they are very structurally unstable (each only present for an exact value of $\sqrt{\delta}$).

## 4.4 Dynamics of Two Factions for Large $\gamma$

Now that we have described the behavior of the smoothed bounded-confidence model in the $\gamma \to \infty$ limit, we will connect it back to the finite $\gamma$ model.

Throughout this section, we will refer to the dynamics of the smoothed bounded-confidence model for a particular selection of $\gamma$ as $\mathbf{F}_\gamma$. We refer to the $\gamma \to \infty$ limit of the dynamics as $\mathbf{F}_\infty$. (When we compare $\mathbf{F}_\gamma$ and $\mathbf{F}_\infty$ it is implied that both systems have otherwise identical parameters.)

### 4.4.1 The $\mathbf{F}_\gamma$ Phase Plane for Large $\gamma$

In this section, we compare the stationary states of $\mathbf{F}_\gamma$ for large $\gamma$ with the stationary states of $\mathbf{F}_\infty$. We find that the phase planes look remarkably similar, though the $\mathbf{F}_\gamma$ phase planes have some additional unstable stationary states not present in the $\mathbf{F}_\infty$ planes.

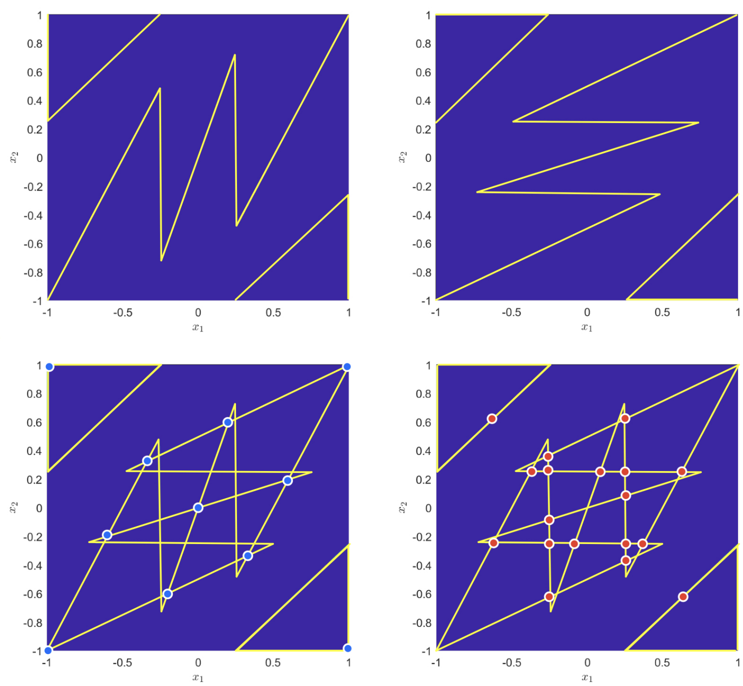


**Figure 4.8** The $x_1$ and $x_2$ nullclines of the system for $\gamma = 300$, $\sqrt{\delta} = 1.25$, $\alpha = \beta = 1$. The top row shows the $x_1$ (left) and $x_2$ (right) nullclines separately. The bottom row shows the nullclines simultaneously with stable stationary states marked in blue (on the left) and unstable stationary states marked in red (on the right).

We begin by comparing $\mathbf{F}_{300}$ and $\mathbf{F}_{\infty}$ with $\sqrt{\delta} = 1.25$, $\alpha = 1$, and $\beta = 1$. Recall that we have already plotted the nullclines of the $\mathbf{F}_{\infty}$ case in Figure 4.4. After plotting the $x_1$ and $x_2$ nullclines of $\mathbf{F}_{300}$ in Figure 4.8, we find that the $x_1$ and $x_2$ nullclines of $\mathbf{F}_{300}$ look very similar to those of the $\mathbf{F}_{\infty}$ case. For instance, the stable stationary states present for $\mathbf{F}_{\infty}$ are all present for $\mathbf{F}_{300}$ (comparing the lower left component of Figure 4.8 with Figure 4.4). The substantial difference is that the discontinuities present in the $\mathbf{F}_{\infty}$ nullclines in Figure 4.4 have been patched up in the $\mathbf{F}_{300}$ phase planes.

The discontinuities are not present in the $\mathbf{F}_{300}$ case because the dynamics $\mathbf{F}_{\gamma}$ remain continuous for any finite $\gamma$ (even though they are converging pointwise to the discontinuous expression of $\mathbf{F}_{\infty}$.) The continuity of the dynamics means that the nullclines must form boundaries between regions of the phase plane where opinions are decreasing and regions of the phase

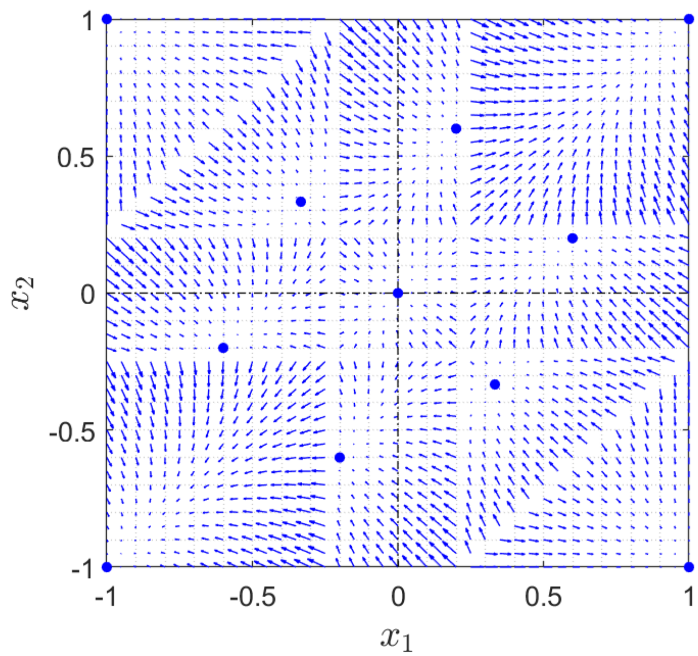


**Figure 4.9** The $(x_1, x_2)$ slope field of the smoothed bounded-confidence model for $\alpha = \beta = 1$, $\gamma = 300$, $\sqrt{\delta} = 1.25$. The stable stationary states are marked in blue.

plane where opinions are increasing (by the Intermediate Value Theorem). In the $\mathbf{F}_\infty$ model this is not the case. In fact, the boundaries between the $R_i$ regions in the $\mathbf{F}_\infty$ case are separating positive and negative dynamics for $x_1$ or $x_2$ without the dynamics ever passing through 0.

These additional components in the nullclines may not seem substantial, but they can actually introduce a lot of additional stationary states. In Figure 4.8, we see that for this particular parameter combination there are 18 additional stationary states in the $\mathbf{F}_{300}$ case that are not present in the $\mathbf{F}_\infty$ case, all of which are unstable. Further, we note that all of these additional stationary states occur near the $R_i$ boundaries of the $\mathbf{F}_\infty$ case.

This example helps motivate the following theorem describing the relationship between the stationary states of $\mathbf{F}_\gamma$ and $\mathbf{F}_\infty$ at large $\gamma$. The following theorem is paraphrased from results from Brooks and Chodrow (2022) and requires definition of some new notation.

Let $\mathcal{X}(\gamma)$ be the set of stationary states of $\mathbf{F}_\gamma$ and let

$$\mathcal{X}_a = \left\{ \mathbf{x} : \left| |x_i - x_j| - \sqrt{\delta} \right| \geq a \quad \forall (i,j) \in \mathcal{E} \right\}.$$

That is, $\mathcal{X}_a$ is the set of all opinion vectors such that no two adjacent agents have opinions that are "nearly exactly $\sqrt{\delta}$ apart" (where "nearly" means closer than $a$).

Finally, let $\mathcal{X}_a(\gamma)$ be $\mathcal{X}(\gamma) \cup \mathcal{X}_a$.

**Theorem 4.3.** *Brooks and Chodrow* (2022) *Let* $\{\mathcal{X}_a(\gamma_1), \mathcal{X}_a(\gamma_2), \ldots, \mathcal{X}_a(\gamma_n), \ldots\}$ *be an infinite sequence such that* $\gamma_1 < \gamma_2 < \cdots < \gamma_n$*. Then, all limit points of this sequence belong to* $\mathcal{X}_a(\infty)$*.*

Colloquially, this means that if we just ignore the stationary states of $\mathbf{F}_\gamma$ that occur very near the $R_i$ boundaries in $\mathbf{F}_\infty$, then all $\mathbf{F}_\gamma$ stationary states belong to some sequence of stationary states that is approaching a stationary state of $\mathbf{F}_\infty$ as $\gamma$ grows large.

Note that this applies to the smoothed bounded-confidence model on any graph (not just the complete graph) and applies to any stationary state (not just consensus or dual factions). That being said, our investigation in Figures 4.4 and 4.8 certainly supports this theorem.

### 4.4.2 A Stronger Relationship Between $\mathbf{F}_\gamma$ and $\mathbf{F}_\infty$?

On the one hand, our comparison of Figures 4.4 and 4.8 does support the claim of Theorem 4.3. However, it also hints at a potentially stronger version of the theorem.

Note that the theorem makes no claim about the reverse inclusion. That is, is every point in $\mathcal{X}_a(\infty)$ realizable as the limit of a sequence of points built from a sequence $\{\mathcal{X}_a(\gamma_1), \mathcal{X}_a(\gamma_2), \ldots, \mathcal{X}_a(\gamma_n), \ldots\}$?

First, it appears that none of the $\mathbf{F}_\infty$ stationary states lie on the $R_i$ regions (and from 4.3 we know this to be true for nearly all selections of $\sqrt{\delta}$). This means that (for almost any $\sqrt{\delta}$) we have $\mathcal{X}_a(\infty) = \mathcal{X}(\infty)$.

Then, we note that all eleven of the stationary states that appear in Figure 4.4 are being approximated by stable stationary states in Figure 4.8. That is, it does seem that every $\mathbf{F}_\infty$ stationary state is being approximated by sequences of $\mathbf{F}_\gamma$ stationary states as $\gamma$ increases. Taking these observations together gives the intuition for the following conjecture.

**Conjecture 4.1.** *Let* $\mathbf{x}^*$ *be a stationary state of* $\mathbf{F}_\infty$ *such that*

1. *Every persuadable agent is receptive to a zealot*

2. *No two adjacent agents have opinions that differ by exactly* $\sqrt{\delta}$.

*Then,* $\mathbf{x}^*$ *is the limit point of a sequence* $\{\mathbf{x}_1, \mathbf{x}_2 \dots\}$ *such that* $\mathbf{x}_i \in \mathcal{X}(\gamma_i)$ *for a sequence of* $\gamma$ *such that* $\gamma_1 < \gamma_2 < \cdots < \gamma_n$.

*Proof.* (Sketch). Condition 2 of the conjecture ensures us that there exists a neighborhood $N(\mathbf{x}^*)$ such that every point in $N(\mathbf{x}^*)$ has the same receptivity set as $\mathbf{x}^*$. Note that the "nullclines" of the $\mathbf{F}_\infty$ system are discontinuous collections of hyperplanes. Since every point in $N(\mathbf{x}^*)$ has the same receptivity set, the nullcline hyperplanes have no discontinuities in $N(\mathbf{x}^*)$.

The fact that every agent in $\mathbf{x}^*$ is receptive to at least one other agent is enough to demonstrate that all of these hyperplanes have dimension $m - 1$, where $m$ is the number of agents in the system.

We think that the condition that every persuadable agent is receptive to a zealot is sufficient to prove that these $m$ hyperplanes are linearly independent. However, we are not sure that this is sufficient (and, if so, it may not be necessary). Another potential (less stringent) condition is that every agent is "connected" to a zealot through a receptivity "chain." (That is, if the receptivity set of $\mathbf{x}^*$ defined the edges of a graph, every connected component of that graph would have a zealot in it.)

Now, we can prove that, in $N(\mathbf{x}^*)$, the nullclines of $\mathbf{F}_\gamma$ are manifolds that are converging absolutely to the $m$ linearly independent $m-1$-dimensional hyperplanes as $\gamma \to \infty$.

It remains to show that the intersection of manifolds approaches the intersection of hyperplanes that they are converging to absolutely. This would complete the proof.

$\square$

This result would be stronger than Theorem 4.3 because it doesn't require us to ignore the behavior of $\mathbf{F}_\gamma$ on the $R_i$ boundaries. (And, if true, it better captures the fact that $\mathcal{X}(\infty)$ is a subset of the limit points of $\mathcal{X}(\gamma)$, and not the other way around.)

# Chapter 5

# Conclusions and Future Work

## 5.1 Conclusions

In this thesis, we have analyzed the smoothed bounded-confidence model of opinion dynamics introduced in Brooks and Chodrow (2022) as applied to the complete graph.

In Chapter 1, we showed how the smoothed bounded-confidence model can be tuned to recover the definition of two classical opinion models: Taylor's model and a Hegselmann-Krause inspired model (originally presented in Taylor (1968) and Hegselmann and Krause (2002), respectively).

In Chapter 2, we presented some general properties of the model, including both original proofs of some results and reproductions from Brooks and Chodrow (2022) of others. Among these results, we proved that at least two zealots must be incorporated into the smoothed bounded-confidence model to obtain stationary states other than consensus.

In Chapter 3, we narrowed our focus to the complete graph with two zealots. There, we created a bifurcation surface which let us visualize the stationary consensus values throughout our $(\gamma, \delta)$ parameter space. We saw that consensus on the smoothed bounded-confidence model recovers the Taylor and Hegselmann–Krause behaviors for small and large $\gamma$, respectively, and that mid-range values of $\gamma$ are a transition zone with behavior not quite like that of either of the classical models.

In Chapter 4, we began to analyze two-dimensional stationary states by considering opinion vectors that consisted of two opinion factions. We began by analyzing the $\gamma \to \infty$ version of the smoothed bounded-confidence model, completely characterizing its stationary states with two factions. Then, using a mixture of numerical results and analytical proof, we established similarities and differences between the stationary states of the finite $\gamma$ model and those of the $\gamma \to \infty$ model.

## 5.2 Future Work

In this section, we detail some possible directions for future work related to this project.

### 5.2.1 Investigating Small and Medium $\gamma$ on Two Factions

In Section 4.4 we establish a relationship between the behavior of our model for large $\gamma$ and the Hegselmann–Krause model. However, we do not consider in much detail what happens with two factions at small or medium values of $\gamma$. In particular, does the model reproduce Taylor-like behavior at low $\gamma$? Are there mid-range values of $\gamma$ for which the model behavior is different than that of Taylor's model or the Hegselmann–Krause model (as we saw in the consensus case)?

### 5.2.2 A Stricter Relationship Between SBC Stationary States and HK Stationary States.

In Section 4.4 we provided the sketch of a proof that nearly all Hegselmann–Krause stationary states are approached by limit sets of finite $\gamma$ stationary states as $\gamma$ approaches $\infty$.

Fleshing out the proof of this conjecture would be valuable both in clarifying whether the theorem statement is true and in providing further insight into the kinds of Hegselmann–Krause stationary states that it applies to. Right now we have conjectures about which stationary states the theorem applies to but it is not clear that these conditions are necessary, let alone sufficient.

### 5.2.3 Exploring Higher Dimensions (More Factions!)

In this thesis, we explored the behavior of stationary states with consensus (which mostly reduced to a one-dimensional system) and those with two opinion factions (which mostly reduced to a two-dimensional system). In these low dimensions, we were able to gain valuable insights to the system through bifurcation diagrams and phase planes.

It would make sense to extend the exploration of the smoothed bounded-confidence model on the complete graph to higher dimensional space (i.e., what do stationary states look like when we allow more than two opinion factions to form?) It would be difficult to do the analysis in higher dimensions, but it would be a valuable insight into the dynamics of the system.

### 5.2.4 Exploring Other Graph Structures

In this thesis, we concentrated on the dynamics of the complete graph. These dynamics were highly non-trivial and we only characterized the behavior of stationary states with up to two opinions in them. Although there is still more work to be done on understanding the smoothed bounded-confidence model on the complete graph, it would also be valuable to work on applying the lessons that we have learned on the complete graph to other graph structures.

This could take many different forms. Perhaps it could involve looking at other graphs that have a relatively high degree of structure to them and comparing the conditions for consensus and faction formation. It could also mean investigating the behavior of networks that consist of a few complete graphs that are sparsely connected and comparing and contrasting their behavior with that of the isolated complete graph. Or, maybe one could take a more simulation-oriented angle and see how consensus and faction formation patterns differ on a graph with less structure to it. These are merely a couple of ideas for how the lessons that we have learned about the smoothed bounded-confidence model could be exported from the complete graph to other types of graph structures.

# Bibliography


Abelson, Robert P. 1967. Mathematical models in social psychology. *Advances in Experimental Social Psychology*, vol. 3, 1–54. Academic Press.

Alligood, Kathleen, Tim Sauer, and James Yorke. 1996. *Chaos: an introduction to dynamical systems*. Springer.

Blondel, Vincent D, Julien M Hendrickx, and John N Tsitsiklis. 2010. Continuous-time average-preserving opinion dynamics with opinion-dependent communications. *SIAM Journal on Control and Optimization* 48(8):5214–5240.

Brooks, Heather Z, and Phil Chodrow. 2022. Bifurcations in a tunably nonlinear model of opinion dynamics. *Manuscript in preparation* .

Brooks, Heather Z, and Mason A Porter. 2020. A model for the influence of media on the ideology of content in online social networks. *Physical Review Research* 2(2):023,041.

Fortunato, Santo. 2005. On the consensus threshold for the opinion dynamics of krause–hegselmann. *International Journal of Modern Physics C* 16(02):259–270.

Guckenheimer, John, and Philip Holmes. 2013. *Nonlinear oscillations, dynamical systems, and bifurcations of vector fields*. Springer Science & Business Media.

Hegselmann, Rainer, and Ulrich Krause. 2002. Opinion dynamics and bounded confidence models, analysis, and simulation. *Journal of artificial societies and social simulation* 5(3).

Homs-Dones, Marc, Karel Devriendt, and Renaud Lambiotte. 2021. Nonlinear consensus on networks: Equilibria, effective resistance, and trees of motifs. *SIAM Journal on Applied Dynamical Systems* 20(3):1544–1570.

Mobilia, Mauro. 2003. Does a single zealot affect an infinite group of voters? *Physical review letters* 91(2):028,701.

Strogatz, Steven H. 1994. *Nonlinear dynamics and chaos: with applications to physics, biology, chemistry, and engineering*. CRC press.

Taylor, Michael. 1968. Towards a mathematical theory of influence and attitude change. *Human Relations* 21(2):121–139.

Yildiz, Ercan, Asuman Ozdaglar, Daron Acemoglu, Amin Saberi, and Anna Scaglione. 2013. Binary opinion dynamics with stubborn agents. *ACM Transactions on Economics and Computation (TEAC)* 1(4):1–30.