

NONLINEAR DYNAMICS OF MULTI-AGENT MULTI-OPTION
BELIEF AND OPINION FORMATION

ANASTASIA SERGEYEVNA BIZYAEVA

A DISSERTATION
PRESENTED TO THE FACULTY
OF PRINCETON UNIVERSITY
IN CANDIDACY FOR THE DEGREE
OF DOCTOR OF PHILOSOPHY

RECOMMENDED FOR ACCEPTANCE BY
THE DEPARTMENT OF
MECHANICAL AND AEROSPACE ENGINEERING
ADVISER: NAOMI EHRICH LEONARD

SEPTEMBER 2022

© Copyright by Anastasia Sergeyevna Bizyaeva, 2022.

All Rights Reserved

Abstract

Many self-organized collective behaviors in natural, social, and technological groups require coordinated decision-making on multiple alternatives. For example, groups must reach consensus on a movement direction to navigate through space in unison or allocate members across different tasks to forage and explore their environment. In this dissertation we present a new modeling framework for the study of multi-alternative collective decision-making in social systems in nature and society, and for the design of such decision-making in technological teams.

First, we describe a new model of social belief and opinion formation as a dynamic and nonlinear process in a multi-agent system. When the agents are identical, the model has a small number of interpretable parameters that characterize intrinsic properties and biases of the networked agents. Belief formation is synthesized through a communication graph that describes a structured set of cooperative and antagonistic relationships between agents, and a belief system graph that describes the logical alignment between options or topics.

We present this model alongside analysis grounded in the theory of nonlinear dynamical systems. We establish that the network generically exhibits a sharp transition from a state of indecision among agents to their commitment to strong beliefs or opinions as the amount of attention to social interactions is increased. We investigate how the model parameters, the communication graph, and the belief system graph inform the allocation of agents across options in this transition and the how relative influence of different agents' biases informs the network decision. We prove conditions under which the belief formation dynamics yield agreement, disagreement, multi-stability of equilibria, and oscillations.

Finally, we illustrate how this modeling framework can be utilized for the design of sensitive and adaptable collective behaviors when model parameters are allowed to be dynamic. First, we introduce tunable dynamic feedback laws for agents' attention to social interactions which provably trigger cascades of strong opinion formation that spread across the entire network in response to a local input. Next, we prove that individual nodes can adapt their decision state locally without affecting the state of the rest of the network through altering the sign of their social interactions.

Acknowledgements

I am overwhelmed with gratitude for the support and mentorship I received during my time in Princeton. First and foremost I want to thank my advisor Professor Naomi Ehrich Leonard for her encouragement and kindness. I am continuously inspired by Naomi's enthusiasm and out-of-the-box creativity. I feel privileged to have had the opportunity to learn from her these past five years. I have grown immensely as a researcher and as a communicator thanks to Naomi's feedback and guidance. Naomi always went out of her way to highlight my contributions to others and to include me in all relevant conversations and collaborations. No matter how new I was to the subject matter of a discussion, I always felt like an equal and respected collaborator. I walk away from my Ph.D. with a newly gained confidence in my abilities and in my ideas thanks to Naomi.

The work in this dissertation would not be possible without the help and guidance of Professor Alessio Franci. Alessio is a co-author on most of the papers I produced during my Ph.D. From our first meeting in the fall semester of my first year at Princeton, Alessio had unwavering faith in my ability and in the validity of my ideas. Whenever I was stuck I could always turn to him for a helpful reference and whenever I made a mistake he helped me correct it without judgment. I came in to my Ph.D. wishing to gain a deep understanding of dynamical systems theory, and I am a step closer to this goal thanks to Alessio.

I am thankful for my interactions with all of the members of the Leonard Lab with whom I had the privilege to cross paths: Liz Davison, Renato Pagliara, Peter Landgren, Desmong Zhong, Bec Gray, Yunxiu Zhou, Mari Kawakatsu, Anthony Savas, Justice Mason, Udari Madhushani, Justin Lidard, Charlotte Cathcart, Isla Xi Han, Giovanna Amorim, Daniel Shen, Zahra Aminzare, Biswadip Dey, Shinkyu Park, Kenza Hamidouche, Christine Allen-Blanchette, and María Santos. Thank you for all of the lockdown social hours on zoom, the impromptu coffee runs and lunch outings, and the inspiring conversations. I am going to miss being part of this fantastic community.

I am grateful for all of the multi-disciplinary research efforts I had an opportunity to partake in during my time at Princeton. In particular I am thankful to Professor Jonathan Cohen, Sebastian Musslick, Lena Rosendahl, and all other members of the Project 6 community I have met throughout the years. The opportunity to learn about cognitive science and neuroscience helped me see my own work in a new light. I am also thankful to Professor Keena Lipsitz and Professor Yphtach Lelkes for their fascinating insights on political polarization and the forces that drive it. It was a privilege to explore with them the interplay between mathematical models and theories in social science.

I want to thank Professor Clarence Rowley and Professor Anirudha Majumdar for serving on

my dissertation committee these past five years and providing helpful advice. Professor Clarence Rowley and Professor Alessio Franci also provided feedback as readers of this dissertation, which greatly improved its contents. I am also grateful to Professor Simon Levin for his advice throughout my time at Princeton, and for his service as an examiner alongside Professor Anirudha Majumdar at my Final Public Oral examination. I also want to thank all of the administrators in the Mechanical and Aerospace Engineering department who helped me navigate the bureaucracy throughout the years, especially Jill Ray, Tom Wing, and Katerina Zara.

I would not be where I am today without my friends and family. I am immensely grateful to my husband Sam Otto, a.k.a. Red, for being my teammate and partner in crime throughout these years. My parents Olga and Serguei endured many hardships throughout the years that enabled me to have this opportunity, for which I cannot thank them enough. My sister Katya is the strongest person I have ever met, and she inspires me to be my best self every day. Sam's parents Craig and Lynn and all of their extended family embraced me into their family with unbounded warmth.

I want to thank William, Aneiss, Alex, Mayia, Travis, Jamie, Szilard, Liya, Mac, Brian, Chris, and all other members of spam chat, a.k.a. "Book" "Club", for the strange and robust community we have formed over the years. Thank you all for the long conversations, food photos, boba nights, and pandemic voice chat "coffee shop" work hours. Thank you Soha for forcing me to be a little more social, Jessica for our walks and conversations, and Alby for our coffee runs. I am grateful to the Princeton Prison Teaching Initiative for providing a valuable opportunity to give back to the community. Teaching PTI courses was one of the most meaningful experiences of my Ph.D.

The last five years were shaped by many ups and downs - a cross-country move, a family death, a pandemic, a wedding, and plenty of worldwide turbulence. This dissertation is a testament to the strength of my support system which empowered me to see light in times of darkness, stress, and uncertainty. I can only hope to return this favor to my future friends, colleagues, and mentees.

The following funding sources supported the research in this dissertation: National Science Foundation Grant CMMI-1635056, Office of Naval Research Grant #N00014-19-1-2556, Army Research Office Grant #W911NF-18-1-0325, Princeton University first year fellowship, Gordon Y.S. Wu Fellowship in Engineering, and Howard Crathorne Phillips Fellowship in Mechanical Engineering. This material is based upon work supported by the National Science Foundation Graduate Research Fellowship Program under Grant #DGE-2039656. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. This dissertation carries T#3436 in the records of the Department of Mechanical and Aerospace Engineering.

To Sam (a.k.a. Red), whose love and support got me through these rough years.
To my parents Olga and Serguei, whose leap of faith across continents gave me this opportunity.
To Katya, my wonderful sister who always cheers me on from afar 🍷.

Contents

Abstract	iii
Acknowledgements	iv
I Dynamic formation of opinions, values, or beliefs	1
1 Introduction	2
1.1 Overview and motivation	2
1.2 Outline of contributions	4
2 Background: mathematical preliminaries	7
2.1 Basic notation	7
2.2 Networked multi-agent systems and graph theory	8
2.3 Dynamical systems theory	11
2.3.1 Equilibria and their stability	12
2.3.2 Invariant manifolds	12
2.3.3 Reduction principle and parameter dependence	13
2.3.4 Bifurcation problems and Lyapunov-Schmidt reduction	15
2.3.5 Elementary bifurcations	19
2.3.6 Unfolding theory and universal unfolding of a pitchfork bifurcation	25
3 Social formation of beliefs as a nonlinear dynamical system	28
3.1 Motivation	28
3.1.1 Formation of opinions and beliefs is a nonlinear process	28
3.1.2 Classic models of social opinion formation	29
3.1.3 Defining properties of nonlinear opinion formation model	31
3.2 New model for dynamics of beliefs and opinions	32

3.3	Indecision-breaking bifurcations	35
3.4	Homogeneous parameters	39
4	Belief-forming bifurcations: single topic or two mutually exclusive options	42
4.1	Model specialization to scalar state variable	42
4.2	Networks of cooperative and competitive agents	46
4.3	Mixed-sign networks	52
4.4	Higher-order bifurcations and mode interaction	57
5	Belief-forming bifurcations: multiple options	68
5.1	Linear analysis	69
5.2	Pitchfork bifurcation and its unfolding	73
5.3	Hopf bifurcation	80
5.4	Mode interaction example: invariant 2-torus	86
5.5	Consensus, dissensus, and multistability on symmetric belief systems	90
6	Dynamic parameters	92
6.1	Opinion cascades: tunable sensitivity and robustness	93
6.2	Dynamic switching	102
7	Final remarks: applications and future extensions	107
7.1	Applications	107
7.1.1	Game theory and social dilemmas	108
7.1.2	Dynamic task allocation	110
7.1.3	Cooperative navigation and collaboration with humans	111
7.1.4	Emergent properties of sociopolitical systems	113
7.2	Future extensions	113
II	Papers	115
8	Overview	116
8.1	Outline	116
8.2	Author contributions	117

9	Nonlinear Opinion Dynamics with Tunable Sensitivity	119
9.1	Introduction	119
9.2	Notation	122
9.3	Nonlinear multi-option opinion dynamics	122
9.3.1	Linear Averaging Models: Drawbacks and Extensions	123
9.3.2	Nonlinear Multi-option Extension of Weighted-average Models: Defining Properties	124
9.3.3	A General Nonlinear Opinion Dynamics Model	126
9.3.4	Generality and Connection to Existing Models	128
9.3.5	Clustering and Model Reduction	131
9.3.6	A Minimal Opinion Network Model	132
9.4	Agreement and disagreement opinion formation	133
9.4.1	Agreement and Disagreement States	134
9.4.2	Opinions Form through a Bifurcation	134
9.4.3	Patterns of Opinion Formation for Two Options	137
9.4.4	Consensus and Dissensus Generic for Transitive Symmetry	140
9.5	Attention dynamics and tunable sensitivity	142
9.5.1	Dynamic State Feedback Law for Attention	143
9.5.2	Tunable Sensitivity and Robustness for a Single Agent	144
9.5.3	Opinion Cascades with Tunable Distributed Sensitivity	145
9.5.4	Dynamics on weights: agreement-disagreement transitions	148
9.6	Final Remarks	150
9.7	Appendix	150
9.7.1	Extension to Heterogeneous Inter-option Coupling	150
9.7.2	Well-definedness of Model	151
9.7.3	Proof of Theorem 9.3.5	152
9.7.4	Proof of Theorem 9.4.2	153
9.7.5	Proof of Proposition 9.4.9	153
10	Patterns of Nonlinear Opinion Formation on Networks	154
10.1	Introduction	154
10.2	Opinion Dynamics Model	156
10.3	Agreement and Disagreement	158

10.4 Agent Centrality	161
10.5 Graph Symmetry	162
10.6 Agent Sensitivity and Opinion Cascades	164
10.7 Discussion	165
11 Control of Agreement and Disagreement Cascades with Distributed Inputs	167
11.1 Introduction	167
11.2 Opinion Dynamics Model	169
11.3 Lyapunov-Schmidt Reduction	171
11.4 Constant Attention: Sensitivity to Input	172
11.5 Dynamic Attention: Cascades and Tunable Sensitivity to Input	174
11.6 Final Remarks	180
12 Switching transformations for decentralized control of opinion patterns in signed networks: application to dynamic task allocation	182
12.1 Introduction	183
12.2 Notation and Mathematical Preliminaries	184
12.3 Nonlinear Opinion Dynamics Model	184
12.3.1 Network opinion formation through bifurcation	185
12.4 Switching Transformation as a Design Tool for Synthesis of Opinion Patterns	187
12.4.1 Signed graphs and switching	187
12.4.2 Nonlinear opinion patterns on switch equivalent graphs	188
12.4.3 Synthesis of nonlinear opinion patterns	190
12.5 Dynamic Switching	191
12.5.1 Monotonicity and structural balance	191
12.5.2 Instantaneous switching	191
12.6 Applications to Multi-Robot Task Allocation	193
12.7 Final Remarks	195
A Supporting material for Chapter 4	196
A.1 Proof of Proposition 4.4.1	196
B Supporting material for Chapter 5	200
B.1 Directional derivatives	200
B.2 Proof of Theorem 5.2.2	201

B.3 Proof of Theorem 5.3.1	203
--------------------------------------	-----

Part I

Dynamic formation of opinions, values, or beliefs

Chapter 1

Introduction

1.1 Overview and motivation

In this dissertation we explore how a group of social decision-makers can form beliefs and reach a decision in evaluating multiple alternatives, accounting for locally available information and social relationships. Such capacity for dynamic belief formation and autonomous decision-making is a key component of *swarm intelligence*. The term “swarm intelligence” is often used to describe the ability of collectives in nature and technology to self-organize and perform complex tasks without a centralized coordination mechanism [26]. Self-organizing systems have the following key properties, identified in [76]:

1. They are *dynamic* entities, evolving in time through a variety of complex positive and negative feedback mechanisms;
2. They have *localized nonlinear interactions* which lead to emergence of macroscopic properties in the group which are more sophisticated than the properties of any given individual;
3. Emergent properties in these systems arise from *bifurcations*, or parametrized transitions characteristic to a nonlinear system;
4. The characteristic emergent properties of these systems are often *multi-stable*, which means that given the same set of parameters, the system can converge to one of several different behaviors.

Common examples of emergent behaviors in self-organizing systems in nature include flocking in starling murmurations [17, 85, 222], honeybee house-hunting [179–181], ant foraging [46, 155, 172],

and spatial decision-making in fish schools [40, 121, 189, 204] among many others. Principles of self-organization are also widespread features of human societies [42]. People organize into communities [91], move through space in complex traffic patterns [56], vote in elections to select their laws and leaders [59], and make financial transactions to participate in the global economy [7]. In online social networks interactions between people can yield unwanted effects - emergence of echo chambers [37], political polarization [141], and descent of individuals into extremism [3]. Understanding the fundamental mechanisms that lead to these emergent features of social systems is an important step towards their mitigation, and therefore towards ensuring the healthy functioning of modern society.

Self-organization is also an important *design* principle for many technologies such as robotic swarms [16, 29, 164], smart electric power grids [129, 203], and mobile sensor networks [53, 176]. These technologies share many similarities with human and animal social groups, in that they are made up of many individual nodes or agents which can interact through communication or sensing and make autonomous decisions. Design of decentralized networked technologies can be a tricky task, as it requires engineers to develop simple rules that individuals can follow to achieve a desired behavioral outcome at the level of the group. In other words it requires design of emergent properties in a self-organized system.

While the literature on flocking, formation control, consensus formation, task allocation, and other collective behaviors in distributed technological systems is exceptionally broad – see for example the surveys [29, 55, 150, 151, 219] – autonomous swarm technology is nevertheless in a nascent stage of development. In designing swarm behavior algorithms researchers often derive behavioral rules for technological units from studies of social animals and human social systems. Such design techniques have seen great success in generating self-organized behaviors in laboratory settings for decades. However despite this success, deployment of swarm technologies in real-world industrial applications is minimal at best [50, 51, 173, 178]. Among the many reasons which contribute to this technological lag are the following:

1. *Lack of theoretical guarantees for the emergent behavior of the collective* [178]. When designing a self-organizing system it is important to ensure not only that a desired property will emerge, but also that unwanted properties will not emerge. Ideally, collective behavior algorithms need to be developed alongside rigorous theory. This is a challenging task because self-organizing systems by their nature are highly nonlinear.
2. *Inability of generated collective behaviors to adapt to meaningful environmental cues* [51]. Many examples of collective behavior in the literature are built with a particular objective in mind

- for example, formation of a specific shape in space or allocation of agents across tasks in a pre-determined distribution. However in real world settings objectives may change, and autonomous swarms need the capacity to adapt their behavior in response to changes in their environment or to input from human collaborators.

These shortfalls of existing design methods impede widespread adoption of decentralized autonomous swarm technologies especially because such systems are often safety-critical, with potentially catastrophic failure consequences [213]. For example, unpredictable or inflexible behavior in a group of autonomous vehicles on the road can lead to a collision or a traffic jam. It is imperative to develop new design techniques grounded in rigorous theory which enable technological teams to self-organize in a robust and reliable manner, yet allow for sensitivity to meaningful cues and adaptation of the group behavior. In this dissertation, we take a first step towards addressing this challenge by presenting a new mechanistic modeling framework which describes how a group reaches a decision on multiple interconnected topics or options. This model is presented alongside with analysis which describes the group’s emergent properties in terms of its social network, belief system, and the intrinsic properties of the agents.

We are interested in modeling belief formation in particular because evaluating and making decisions on context-dependent options is a fundamental building block of almost any collective behavior. Formal taxonomies of swarm behaviors typically distinguish between collective decision-making such as consensus formation [206] and task allocation [77, 114], and other swarm behaviors such as collective navigation, foraging, and flocking [16, 29]. However at their core, many of these behaviors can be viewed abstractly as a sequence of collective decisions on spatially embedded options. This view is consistent with findings in the literature on collective behavior in animals. For example, spatial movement patterns in various animal groups have been linked to outcomes of social decisions [40, 189, 194]. Theory developed for collective belief formation and decision-making therefore has potential to be used in a wide variety of contexts. We envision the modeling and analysis presented in this dissertation as a tool both for understanding how various emergent properties arise in groups of biological and socio-political decision-makers, and for development of new collective behavior algorithms for coordination of robotic teams and other autonomous swarm technologies.

1.2 Outline of contributions

Motivated by the above discussion, the following contributions are presented in Part I of this dissertation:

1. In Chapter 2 we review some mathematical ideas which are utilized throughout the dissertation. The topics we review include fundamental definitions from the theory of signed graphs and an overview of nonlinear dynamical systems with a focus on local bifurcation theory.
2. In Chapter 3 we motivate and present a new model which describes how a network of communicating agents forms beliefs on several interconnected topics. The belief formation mechanism captured by this model combines nonlinear integration of social evidence with the capacity of agents to have intrinsic biases or access to task-relevant information. These features in our model capture the abstract features of a broad class of biological and technological decision-makers. A fundamental mechanism behind formation of strong beliefs on a network according to this model is a sharp transition from indecision to decision, which is parametrized by the amount of attention individuals allocate towards information obtained from their social interactions. We refer to this transition as an indecision-breaking or belief-forming bifurcation.
3. In Chapters 4 and 5 we present a body of rigorous analysis of this model, grounded in the theory of nonlinear dynamical systems. This analysis characterizes the properties of beliefs formed on the network at the onset of the indecision-breaking bifurcation. We show that this bifurcation can result in a wide variety of outcomes, including appearance of multi-stable opinionated equilibria and onset of periodic oscillations of beliefs. We establish how the communication graph between agents and a belief system graph, which describes the alignment between different options, inform patterns of equilibrium beliefs that emerge on the network, as well as relationships between agents' beliefs in the periodic oscillations. Specifically we identify several spectral properties of these graphs that determine the amount of attention necessary for onset of bifurcation, how the agents distribute themselves across the options post-bifurcation, and which agents are the most influential in swaying the group belief formation outcome. We present a simple synthesis procedure that can be used to realize a desired pattern of beliefs on any network through manipulating signs of local interactions between agents. Importantly, we show that formation of beliefs according to this model can lead both to agreement and disagreement solutions on the network. This provides a unified framework for consensus formation and task allocation, which are traditionally treated as separate design objectives [206].
4. In Chapter 6 we illustrate how this modeling framework can be used for design of flexible and adaptable behaviors by allowing model parameters to be dynamic, through instantaneous updates or through dynamic feedback of network states. First we show how dynamic attention

to social interactions results in sensitive cascades of opinion formation in which the entire group state transitions from indecision to decision in response to an input or a sensor measurement introduced locally at one or more nodes. Next we show how individual agents can change their belief state dynamically without affecting the behavior of the rest of the network through updating the sign of interactions with their neighbors on the communication network.

5. In Chapter 7 we conclude with an overview of applications and future extensions of the modeling framework presented in this dissertation.

Chapter 2

Background: mathematical preliminaries

In this chapter we review fundamental concepts from the theory of nonlinear dynamical systems and from the theory of graphs. Both of these subjects are vast, and the overview in this chapter is not meant to be a comprehensive summary. For a thorough development of dynamical systems theory, see [82–84, 90, 109, 117]. For a detailed treatment of graphs and signed graphs, see [41, 137, 207, 223, 224].

2.1 Basic notation

In this dissertation we use the symbol \mathbb{R} to denote the real numbers and the symbol \mathbb{C} to denote the complex numbers. For a complex number $x = a + ib = re^{i\phi}$, $\bar{x} = a - ib = re^{-i\phi}$ signifies the complex conjugate, $|x| = \sqrt{x\bar{x}} = r$ signifies the modulus, and $\arg(x)$ signifies its argument ϕ . Matrices are represented by capital letters, e.g. A , and vectors are represented by boldface symbols, e.g. \mathbf{v} . The inner product of two vectors \mathbf{v}, \mathbf{w} is $\langle \mathbf{w}, \mathbf{v} \rangle = \bar{\mathbf{w}}^T \mathbf{v}$. We define $\mathbf{0} \in \mathbb{R}^N$ as the zero vector and $\text{diag}(\mathbf{v})$ as a diagonal matrix whose diagonal entries are components of \mathbf{v} . The standard $N - 1$ simplex Δ_{N-1} is the set of all vectors $\mathbf{x} = (x_1, \dots, x_N) \in \mathbb{R}^N$ which satisfy $\sum_{i=1}^N x_i = 1$, $x_i \geq 0$ for all $i = 1, \dots, N$.

We denote the spectrum of a square matrix A as $\sigma(A) = \{\lambda_1, \dots, \lambda_N\}$ and its spectral radius $\rho(A) = \max\{|\lambda_i|, \lambda_i \in \sigma(A)\}$. We say an eigenvalue $\lambda \in \sigma(A)$ is a leading eigenvalue of A if $\text{Re}(\lambda) \geq \text{Re}(\mu)$ for all $\mu \in \sigma(A)$. We say a leading eigenvalue λ of A is a dominant eigenvalue if $\lambda = \rho(A)$. A right eigenvector \mathbf{v}_i of A corresponding to λ_i satisfies $A\mathbf{v}_i = \lambda_i\mathbf{v}_i$, and a left eigenvector

\mathbf{w}_i of A corresponding to λ_i satisfies $\mathbf{w}_i A = \lambda_i \mathbf{w}_i$. Given two vectors \mathbf{v}, \mathbf{w} or two matrices M, N , we say $\mathbf{v} \succeq \mathbf{w}$ whenever $v_i \geq w_i$ for all i (similarly $M \succeq N$ whenever $M_{ij} \geq N_{ij}$ for all i, j). Analogously, $\mathbf{v} \succ \mathbf{w}$ whenever $v_i > w_i$ for all i (similarly $M \succ N$ whenever $M_{ij} > N_{ij}$ for all i, j). For two matrices $M, N \in \mathbb{R}^{m \times n}$ we define the element-wise Hadamard product $M \odot N \in \mathbb{R}^{m \times n}$ as $(M \odot N)_{ij} = M_{ij} N_{ij}$. For two matrices $M \in \mathbb{R}^{m \times n}, N \in \mathbb{R}^{k \times l}$ we define the Kronecker product, or direct product, $M \otimes N \in \mathbb{R}^{mk \times nl}$ as matrix

$$M \otimes N = \begin{pmatrix} M_{11}N & M_{12}N & \dots & M_{1n}N \\ M_{21}N & M_{22}N & \dots & M_{2n}N \\ \vdots & \vdots & \ddots & \vdots \\ M_{m1}N & M_{m2}N & \dots & M_{mn}N \end{pmatrix}.$$

We say a real square matrix A has the Perron-Frobenius property if it has a dominant eigenvalue $\lambda = \rho(A)$ and a corresponding eigenvector satisfying $\mathbf{v} \succeq \mathbf{0}$; A has the strong Perron-Frobenius property if the dominant eigenvalue is unique and satisfies $\lambda \geq |\lambda_i|$ for all $\lambda_i \neq \lambda$ in the spectrum $\sigma(A)$ and its corresponding eigenvector satisfies $\mathbf{v} \succ \mathbf{0}$. A real square matrix is irreducible if it cannot be transformed into an upper triangular matrix through similarity transformations. A real square matrix A is eventually positive (eventually nonnegative) if there exists a positive integer k_0 such that $A^k \succ 0_{N \times N}$ ($A^k \succeq 0_{N \times N}$) for all integers $k > k_0$. Several results in this dissertation take advantage of the following proposition.

Proposition 2.1.1. [149, Theorem 2.2] *The following statements are equivalent for a real square matrix A : (1) A and A^T have the strong Perron-Frobenius property; (2) A is eventually positive; (3) A^T is eventually positive.*

2.2 Networked multi-agent systems and graph theory

The mathematical framework of *multi-agent systems* captures the common features shared by many natural and technological collectives. An *agent* in a multi-agent system is a dynamic unit which takes autonomous actions and exchanges information with other agents through local interactions [137, 216]. Interactions between agents can be grouped into two broad categories. The first of these is direct communication, or sending and receiving of signals and messages. The second category is sensing, which captures information exchange through vision, sound, pheromone signals, and other environment-mediated interactions. Prominent examples of multi-agent systems include robotic

teams, mobile sensor networks, power grids, ensembles of neurons, groups of social animals, and human social networks. The actions of individual agents in such systems can add up to complex emergent behaviors at the level of the group. Some collective phenomena which have been extensively studied through the lens of multi-agent systems include consensus formation [5, 45, 153, 182], task allocation [6, 110, 122, 130], synchronization [2, 140, 196], and collective movement [146, 152].

Local interactions between agents together make up an interaction *network*. This network is a fundamental feature of a multi-agent system and often plays a key role in defining its global emergent behavior. Mathematically, interaction geometry of networked agents is captured by a *graph*. A graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is a mathematical object made up of a set of N vertices or nodes $\mathcal{V} = \{1, \dots, N\}$ and a set of edges \mathcal{E} . Each node in \mathcal{V} represents an agent, and whenever an edge e_{ik} belongs to the edge set \mathcal{E} we say agent i receives information about agent k - i.e. there is an arrow from node i to node k . In this dissertation we assume that all graphs are *simple*, meaning that for any two nodes $i, k \in \mathcal{V}$ there is at most one edge e_{ik} that starts at i and ends at k . The structure of a simple graph \mathcal{G} is captured by its *adjacency matrix* $A = (a_{ik})$ where $a_{ik} = 0$ whenever $e_{ik} \notin \mathcal{E}$ and $a_{ik} \neq 0$ otherwise. Most generally, an adjacency matrix can be *weighted*, meaning the nonzero entries a_{ik} can take on any real value that represents the sign and strength of influence from agent k on agent i . In this dissertation we will typically consider graphs with *unweighted* adjacency matrices for which $a_{ik} \in \{0, 1\}$ or *signed unweighted* adjacency matrices for which $a_{ik} \in \{0, 1, -1\}$ for all $i, k \in \mathcal{V}$. The graph is *undirected* whenever $a_{ik} = a_{ki}$ for all $i, k \in \mathcal{V}$, and *directed* otherwise. Both directed and undirected networks naturally arise in various multi-agent systems. For example, in human social networks face-to-face interactions are typically undirected, whereas online interactions through social media are typically directed.

Consider an unsigned graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ that consists of sets of nodes or vertices \mathcal{V} and edges \mathcal{E} . The in-degree of vertex i on a graph \mathcal{G} is the number of edges that end in node i , i.e. $\sum_k a_{ik}$. A walk or path on a graph is a finite or infinite sequence of edges that joins a sequence of vertices, for example e_{12}, e_{25}, e_{53} is a walk of length 3 that visits nodes 1, 2, 5, 3 in that order. A graph \mathcal{G} is said to be strongly connected if in its topology there exists a path from any vertex to any other vertex. A graph is strongly connected if and only if its adjacency matrix is an irreducible matrix with the strong Perron-Frobenius property. A closed path is a walk of positive length that starts and ends at the same vertex. Finally, \mathcal{G} is *bipartite* if its vertex set \mathcal{V} can be partitioned into two disjoint subsets $\mathcal{V}_1, \mathcal{V}_2$ such that for any edge $e_{ik} \in \mathcal{E}$, either $i \in \mathcal{V}_1$ and $j \in \mathcal{V}_2$ or $i \in \mathcal{V}_2$ and $j \in \mathcal{V}_1$.

In this dissertation we frequently consider signed graphs $\mathcal{G} = (\mathcal{V}, \mathcal{E}, s)$ where $s : \mathcal{E} \rightarrow \{-1, 1\}$ is the signature function for the graph. A signed adjacency matrix A has positive and negative entries,

defined as $a_{ik} = s(e_{ik})$ if $e_{ik} \in \mathcal{E}$ and $a_{ik} = 0$ otherwise. The sign of a walk on a signed graph is the product of the signatures of all of the edges passed on the walk, e.g. $s(e_{12})s(e_{25})s(e_{53})$. A signed graph has a property called *structural balance* when the sign of every closed path on the graph is positive.

The remaining paragraphs this section are taken verbatim from [18]. Let $\mathcal{W} \subset \mathcal{V}$ be a subset of nodes on a signed graph \mathcal{G} . *Switching* a set \mathcal{W} on the graph \mathcal{G} refers to a mapping of the graph \mathcal{G} to $\mathcal{G}^{\mathcal{W}} = (\mathcal{V}, \mathcal{E}, s_{\mathcal{W}})$ where the signature of all the edges in \mathcal{E} between nodes in \mathcal{W} and nodes in its complement $\mathcal{V} \setminus \mathcal{W}$ reverses sign. We introduce the *switching function* $\theta : \mathcal{V} \rightarrow \{1, -1\}$, where for any $i \in \mathcal{V}$, $\theta(i) = -1$ if $i \in \mathcal{W}$ and $\theta(i) = 1$ otherwise. Then the signature of the switched graph $\mathcal{G}^{\mathcal{W}}$ is generated as

$$s_{\mathcal{W}}(e_{ik}) = \theta(i)s(e_{ik})\theta(k) \quad (2.1)$$

for all $e_{ik} \in \mathcal{E}$. From (2.1) we see that the signature update for an edge between agents i and k depends only on their membership in the switching set \mathcal{W} . Thus, the edges between i and k flip sign if and only if exactly one of i, k is in the switching set \mathcal{W} , and does not change sign if i, k are both in \mathcal{W} or in $\mathcal{V} \setminus \mathcal{W}$. Importantly, switching a set \mathcal{W} all at once generates the same graph $\mathcal{G}^{\mathcal{W}}$ as sequentially switching individual vertices in \mathcal{W} . If \mathcal{G} can be transformed into $\mathcal{G}^{\mathcal{W}}$ by switching, \mathcal{G} and $\mathcal{G}^{\mathcal{W}}$ are *switching equivalent graphs*. A switching transformation preserves the sign of all closed paths, and therefore preserves structural balance.

Let θ be the function for switching from graph \mathcal{G} to $\mathcal{G}^{\mathcal{W}}$, with adjacency matrices A and $A^{\mathcal{W}}$, respectively. Define the *switching matrix* $\Theta = \text{diag}(\theta(1), \theta(2), \dots, \theta(N))$. The adjacency matrices of \mathcal{G} and its switching $\mathcal{G}^{\mathcal{W}}$ are related as

$$A^{\mathcal{W}} = \Theta^{-1}A\Theta. \quad (2.2)$$

Since Θ is diagonal and $\theta(i) = \pm 1$, $\Theta^{-1} = \Theta$; we write Θ^{-1} in (2.2) to emphasize that the matrices A and $A^{\mathcal{W}}$ are similar. We refer to (2.2) as a *switching transformation* of the adjacency matrix A , and A and $A^{\mathcal{W}}$ as *switching equivalent adjacency matrices*. The following proposition is adapted from [224, Proposition II.5].

Proposition 2.2.1. [18] *Suppose $\mathcal{G}, \mathcal{G}^{\mathcal{W}}$ are switching equivalent with adjacency matrices A and $A^{\mathcal{W}}$ and associated switching matrix Θ . Then 1) A and $A^{\mathcal{W}}$ are isospectral, i.e. have the same set of eigenvalues; 2) \mathbf{v} (\mathbf{w}) is a right (left) eigenvector of A corresponding to eigenvalue λ if and only if $\Theta\mathbf{v}$ ($\Theta\mathbf{w}$) is a right (left) eigenvector of $A^{\mathcal{W}}$ with the same eigenvalue.*

2.3 Dynamical systems theory

A *dynamical system* is a mathematical object which describes the deterministic evolution of one or more variables in discrete or continuous time. In continuous time a dynamical system is typically represented by a system of *ordinary differential equations (ODEs)*. In this dissertation we study autonomous, continuous time dynamical systems of the form

$$\frac{d\mathbf{x}}{dt} =: \dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{p}) \quad (2.3)$$

where $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$ is the system *state*, $\mathbf{p} = (p_1, \dots, p_{n_p}) \in \mathbb{R}^{n_p}$ is the set of system *parameters*, and $\mathbf{f} : \mathbb{R}^n \times \mathbb{R}^{n_p} \rightarrow \mathbb{R}^n$ is a smooth vector field $\mathbf{f}(\mathbf{x}, \mathbf{p}) = (f_1(\mathbf{x}, \mathbf{p}), \dots, f_n(\mathbf{x}, \mathbf{p}))$. By smoothness we mean that the vector field \mathbf{f} is continuously differentiable with respect to the variables and the parameters as many times as necessary for the methods we will discuss. When convenient, we may omit the explicit mention of the vector field dependence on the parameters \mathbf{p} and simply state $\dot{\mathbf{x}} = f(\mathbf{x})$. The *flow* generated by (2.3) is the operator $\varphi^t : \mathbb{R}^n \rightarrow \mathbb{R}^n$ which satisfies (2.3), meaning $\left. \frac{d}{dt}(\varphi^t(\mathbf{x})) \right|_{t=\tau} = f(\varphi^\tau(\mathbf{x}))$. The flow operator φ^τ thereby maps any initial condition $\mathbf{x}(0)$ to some point $\mathbf{x}(\tau)$ at time τ . The set $\varphi^t(\mathbf{x}(0)) := \mathbf{x}(t)$ parametrized by time t is often referred to as a *trajectory* or an *orbit* of the dynamical system (2.3) based at $\mathbf{x}(0)$. For the remainder of this section, we assume that (2.3) is defined on a compact forward-invariant *trapping region* $W \subset \mathbb{R}^n$, meaning that for any initial condition $\mathbf{x}(0) = \mathbf{x}_0 \in W$ the associated solution trajectory $\mathbf{x}(t)$ remains in W for all $t \geq 0$. Existence and uniqueness of solutions to (2.3) is guaranteed for all $t \geq 0$ as long as the system is defined on W [109, Theorem 3.3]. With this established, one of the primary aims of dynamical systems theory is to rigorously describe features of (2.3) qualitatively, without finding explicit expressions for the general solutions $\mathbf{x}(t)$. At the core of such qualitative descriptions lies the notion of *topological equivalence*, which we now formally define. Recall that a *homeomorphism* is a continuous and invertible map with a continuous inverse.

Definition 2.3.1 (Topological equivalence). [117, Definition 2.1] *Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, $\mathbf{f} : \mathbb{R}^n \rightarrow \mathbb{R}^n$, $\mathbf{g} : \mathbb{R}^n \rightarrow \mathbb{R}^n$. The dynamical system $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$ is topologically equivalent to the dynamical system $\dot{\mathbf{y}} = \mathbf{g}(\mathbf{y})$ if there exists a homeomorphism $h : \mathbb{R}^n \rightarrow \mathbb{R}^n$ which maps the trajectories $\mathbf{x}(t)$ onto the trajectories $\mathbf{y}(t)$ and which preserves the direction of time.*

2.3.1 Equilibria and their stability

Given a choice of parameters \mathbf{p} , an important class of solutions of (2.3) are the *equilibria* or *fixed points* defined as points $\mathbf{x}^* \in W$ at which the vector field vanishes, $\mathbf{f}(\mathbf{x}^*, \mathbf{p}) = \mathbf{0}$. We say an equilibrium \mathbf{x}^* is

- *stable* if for every neighborhood $U \subset W$ of \mathbf{x}^* , there exists a neighborhood $U_1 \subset U$ of \mathbf{x}^* so that for all initial conditions $\mathbf{x}_0 \in U_1$, the associated solution $\mathbf{x}(t)$ is in U for all $t \geq 0$;
- *locally asymptotically stable* if it is stable and we can find U_1 for which $\mathbf{x}(t) \rightarrow \mathbf{x}^*$ as $t \rightarrow \infty$;
- *locally exponentially stable* if it is locally asymptotically stable, and there exist positive constants K, λ for which $\|\mathbf{x}(t)\| \leq Ke^{-\lambda t} \|\mathbf{x}_0\|$ for all $\mathbf{x}_0 \in U_1$;
- *unstable* if it is not stable.

If the asymptotic or exponential stability definition is satisfied for all neighborhoods in W , we say the equilibrium is *globally* asymptotically or exponentially stable.

At an equilibrium \mathbf{x}^* , the *linearization* of (2.3) is the set of linear ODEs

$$\dot{\mathbf{y}} = D\mathbf{f}(\mathbf{x}^*, \mathbf{p})\mathbf{y} \tag{2.4}$$

where $\mathbf{y} = \mathbf{x} - \mathbf{x}^*$ and $D\mathbf{f}(\mathbf{x}^*, \mathbf{p}) \in \mathbb{R}^{n \times n}$ is the Jacobian matrix of the linearization, with $(D\mathbf{f})_{ij} = \frac{\partial f_i}{\partial x_j}$. An equilibrium \mathbf{x}^* is called *hyperbolic* if none of the eigenvalues of its Jacobian matrix $D\mathbf{f}(\mathbf{x}^*, \mathbf{p})$ are zero or purely imaginary. By the Hartman-Grobman theorem [90, Theorem 1.3.1], near a hyperbolic equilibrium \mathbf{x}^* , the nonlinear system (2.3) is topologically equivalent to its linearization (2.4). As a consequence, the stability of a hyperbolic equilibrium \mathbf{x}^* can be directly inferred from the stability of the origin in the linear system (2.4). Specifically, whenever all eigenvalues of $D\mathbf{f}(\mathbf{x}^*, \mathbf{p})$ have a negative real part the equilibrium is locally exponentially stable, and it is unstable if at least one of the eigenvalues has a positive real part [95, Theorems 8.5 and 8.6]. When all eigenvalues of $D\mathbf{f}(\mathbf{x}^*, \mathbf{p})$ have a negative real part we say the matrix is *Hurwitz*.

2.3.2 Invariant manifolds

Existence and local geometry of invariant manifolds near an equilibrium \mathbf{x}^* for the flow of (2.3) can also be inferred from the linearization (2.4). In general, an invariant set for a flow φ^t is a subset $S \subset W$ which satisfies $\varphi^t(\mathbf{x}(0)) \in S$ whenever $\mathbf{x}(0) \in S$, for all $t \in \mathbb{R}$. Eigenspaces of the Jacobian matrix $D\mathbf{f}(\mathbf{x}^*, \mathbf{p})$ are invariant for the flow of the linear system (2.4). The *stable eigenspace* E^s , the

unstable eigenspace E^u , and the *center eigenspace* E^c are direct products of all of the (generalized) eigenspaces of $D\mathbf{f}(\mathbf{x}^*, \mathbf{p})$ corresponding to eigenvalues with negative, positive, and zero real parts respectively. As implied by the names of these sets, whenever $\mathbf{y}(0) \in E^s$, $\mathbf{y}(t) \rightarrow \mathbf{0}$ at an exponential rate as $t \rightarrow \infty$ and whenever $\mathbf{y}(0) \in E^u$, $\mathbf{y}(t) \rightarrow \mathbf{0}$ at an exponential rate as $t \rightarrow -\infty$. The behavior of trajectories of (2.4) along E^c is characterized neither by exponential growth nor by exponential decay. In the simplest case when E^c corresponds to a simple zero eigenvalue, all points in E^c are fixed points for the dynamics (2.4), and when E^c corresponds to a complex-conjugate eigenvalue pair the trajectories oscillate at a constant amplitude [90, Chapter 1.2].

Suppose \mathbf{x}^* is a hyperbolic equilibrium of the nonlinear system (2.3) and its linearization about \mathbf{x}^* (2.4) has $\dim E^s = n_s$ and $\dim E^u = n_u$. By the (un)stable manifold theorem [90, Theorem 1.3.2] there exist a local *stable manifold* $W_{loc}^s(\mathbf{x}^*)$ and a local *unstable manifold* $W_{loc}^u(\mathbf{x}^*)$ of dimensions n_s and n_u respectively which are invariant under the flow of (2.3) and tangent to E^s, E^u at $\mathbf{x} = \mathbf{x}^*$. Global stable and unstable manifolds $W^s(\mathbf{x}^*), W^u(\mathbf{x}^*)$ can be obtained by following trajectories in $W_{loc}^s(\mathbf{x}^*), W_{loc}^u(\mathbf{x}^*)$ backwards and forwards in time. Furthermore these manifolds are unique and as smooth as the vector field \mathbf{f} . Whenever $\mathbf{x}(0) \in W^s(\mathbf{x}^*)$, $\mathbf{x}(t) \rightarrow \mathbf{x}^*$ as $t \rightarrow \infty$ and whenever $\mathbf{x}(0) \in W^u(\mathbf{x}^*)$, $\mathbf{x}(t) \rightarrow \mathbf{x}^*$ as $t \rightarrow -\infty$. Trajectories starting near \mathbf{x}^* decay towards \mathbf{x}^* along $W^s(\mathbf{x}^*)$ and evolve away from \mathbf{x}^* along $W^u(\mathbf{x}^*)$.

Furthermore we cannot infer the behavior of the trajectories along $W_c(\mathbf{x}^*)$ from the flow of the linearized system along E^c . When \mathbf{x}^* is a non-hyperbolic equilibrium, the Hartman-Grobman theorem does not apply and the flow of the nonlinear system (2.3) is not topologically equivalent to the flow of its linearization. However we can still use the linearized system (2.4) to infer information about the behavior of its trajectories near \mathbf{x}^* . Suppose \mathbf{x}^* is a non-hyperbolic equilibrium of the nonlinear system (2.3) and its linearization about \mathbf{x}^* (2.4) has $\dim E^s = n_s$, $\dim E^u = n_u$, and $\dim E^c = n_c$. By the center manifold theorem [90, Theorem 3.2.1] there exist manifolds $W^s(\mathbf{x}^*), W^u(\mathbf{x}^*), W^c(\mathbf{x}^*)$ which are invariant under the flow of (2.3), with the stable and unstable manifolds $W^s(\mathbf{x}^*), W^u(\mathbf{x}^*)$ defined as before and a *center manifold* $W^c(\mathbf{x}^*)$ which is tangent to E^c at $\mathbf{x} = \mathbf{x}^*$. Analogously to the hyperbolic case, $W^s(\mathbf{x}^*)$ and $W^u(\mathbf{x}^*)$ are unique; however $W^c(\mathbf{x}^*)$ may not be unique.

2.3.3 Reduction principle and parameter dependence

An important consequence of the center manifold theorem is sometimes referred to as the *reduction principle* [117, Theorem 5.2], [90, p.130]. By this principle, in a neighborhood of a non-hyperbolic

equilibrium \mathbf{x}^* , the nonlinear system (2.3) is topologically equivalent to the system

$$\dot{\mathbf{y}}_s = -\mathbf{y}_s \tag{2.5a}$$

$$\dot{\mathbf{y}}_u = \mathbf{y}_u \tag{2.5b}$$

$$\dot{\mathbf{y}}_c = \mathbf{g}(\mathbf{y}_c, \mathbf{p}) \tag{2.5c}$$

with $(\mathbf{y}_s, \mathbf{y}_u, \mathbf{y}_c) \in W^s(\mathbf{x}^*) \times W^u(\mathbf{x}^*) \times W^c(\mathbf{x}^*)$. While the trajectories along stable and unstable manifolds are characterized by decay and growth, respectively, more complex dynamics can happen along a center manifold $W^c(\mathbf{x}^*)$. Furthermore in many systems of interest $n_u = 0$ and all nearby trajectories decay to $W^c(\mathbf{x}^*)$ as $t \rightarrow \infty$. Characterizing the qualitative properties of the flow of the n -dimensional system (2.3) near \mathbf{x}^* boils down to studying the reduced n_c -dimensional system (2.5c), the restriction of (2.3) to a center manifold at \mathbf{x}^* .

Since $W^c(\mathbf{x}^*)$ is typically nonlinear and not necessarily unique, deriving a reduced system of the form (2.5c) poses a challenge. For many systems, the task of finding a closed-form expression for $\mathbf{g}(\mathbf{y}_c, \mathbf{p})$ is not analytically tractable. A technique referred to as *center manifold reduction* is often used to find a truncated series approximation for the dynamics along a center manifold. Suppose the system (2.3) is written in the block diagonal form

$$\dot{\mathbf{x}}_s = A_s \mathbf{x}_s + \mathbf{f}_s(\mathbf{x}_s, \mathbf{x}_u, \mathbf{x}_c) \tag{2.6a}$$

$$\dot{\mathbf{x}}_u = A_u \mathbf{x}_u + \mathbf{f}_u(\mathbf{x}_s, \mathbf{x}_u, \mathbf{x}_c) \tag{2.6b}$$

$$\dot{\mathbf{x}}_c = A_c \mathbf{x}_c + \mathbf{f}_c(\mathbf{x}_s, \mathbf{x}_u, \mathbf{x}_c) \tag{2.6c}$$

where $(\mathbf{x}_s, \mathbf{x}_u, \mathbf{x}_c) \in \mathbb{R}^{n_s} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_c}$, $A_s \in \mathbb{R}^{n_s \times n_s}$ has negative real-part eigenvalues, $A_u \in \mathbb{R}^{n_u \times n_u}$ has positive real-part eigenvalues, and $A_c \in \mathbb{R}^{n_c \times n_c}$ has eigenvalues on the imaginary axis. We can always express a system in the form (2.6) using a suitable change of variables, typically by choosing a coordinate system that transforms the linear part of (2.3) into its Jordan normal form. Without loss of generality assume that $\mathbf{x}^* = \mathbf{0}$ and the functions $\mathbf{f}_s, \mathbf{f}_u, \mathbf{f}_c$ vanish at the origin along with their partial derivatives (meaning all of the linear terms in the dynamics are captured by A_s, A_u, A_c).

Near $\mathbf{x} = \mathbf{0}$, a center manifold can be represented as a local graph over the coordinates \mathbf{x}_c due to its tangency to E_c at the origin,

$$W^c(\mathbf{0}) = \{(\mathbf{x}_s, \mathbf{x}_u, \mathbf{x}_c) \mid \mathbf{x}_s = \mathbf{h}_s(\mathbf{x}_c), \mathbf{x}_u = \mathbf{h}_u(\mathbf{x}_c)\}. \tag{2.7}$$

Then the restriction of (2.3) to a center manifold near $\mathbf{x} = \mathbf{0}$ is captured by the reduced system

$$\dot{\mathbf{x}}_c = A_c \mathbf{x}_c + \mathbf{f}_c(\mathbf{h}_s(\mathbf{x}_c), \mathbf{h}_u(\mathbf{x}_c), \mathbf{x}_c). \quad (2.8)$$

To approximate the graphs \mathbf{h}_s and \mathbf{h}_u we observe that on a center manifold (2.7) $\dot{\mathbf{x}}_s = D\mathbf{h}_s(\mathbf{x}_c)\dot{\mathbf{x}}_c$, and $\dot{\mathbf{x}}_u = D\mathbf{h}_u(\mathbf{x}_c)\dot{\mathbf{x}}_c$. Since a center manifold is invariant under the dynamics (2.6) yields a system of $n_s + n_u$ equations which implicitly define the graphs $\mathbf{h}_s, \mathbf{h}_u$:

$$D\mathbf{h}_s(\mathbf{x}_c) \left(A_c \mathbf{x}_c + \mathbf{f}_c(\mathbf{h}_s(\mathbf{x}_c), \mathbf{h}_u(\mathbf{x}_c), \mathbf{x}_c) \right) = A_s \mathbf{h}_s(\mathbf{x}_c) + \mathbf{f}_s(\mathbf{h}_s(\mathbf{x}_c), \mathbf{h}_u(\mathbf{x}_c), \mathbf{x}_c), \quad (2.9a)$$

$$D\mathbf{h}_u(\mathbf{x}_c) \left(A_c \mathbf{x}_c + \mathbf{f}_c(\mathbf{h}_s(\mathbf{x}_c), \mathbf{h}_u(\mathbf{x}_c), \mathbf{x}_c) \right) = A_u \mathbf{h}_u(\mathbf{x}_c) + \mathbf{f}_u(\mathbf{h}_s(\mathbf{x}_c), \mathbf{h}_u(\mathbf{x}_c), \mathbf{x}_c). \quad (2.9b)$$

Then by taking out Taylor series expansions for $\mathbf{h}_s, \mathbf{h}_u, \mathbf{f}_s, \mathbf{f}_u, \mathbf{f}_c$ and matching up coefficients in (2.9) we can recover approximate expressions that define a center manifold (2.7) and the reduced dynamics along it (2.8). For a detailed treatment of center manifold reduction with several worked out examples of computing these approximations, see [90, Chapter 3.2].

In studying the flow of (2.3) we are often interested in how its qualitative features depend on the set of model parameters \mathbf{p} . Suppose \mathbf{x}^* is an equilibrium (2.3) for some choice of parameters $\mathbf{p} = \mathbf{p}^*$. To account for the parameter vector, we append to (2.3) the trivial system of ODEs $\dot{\mathbf{p}} = \mathbf{0}$. With this amendment, a center manifold $W^c(\mathbf{x}^*, \mathbf{p}^*)$ has dimension $n_c + n_p$, and the restriction of the extended system to its center manifold is of the form

$$\dot{\mathbf{x}}_c = \mathbf{g}(\mathbf{x}_c, \mathbf{p}) \quad (2.10a)$$

$$\dot{\mathbf{p}} = \mathbf{0}. \quad (2.10b)$$

Following the center manifold reduction procedure for this extended system yields a reduced system of equations of the form (2.10) in which the evolution of state variables explicitly depends on the parameters \mathbf{p} . Furthermore, this reduced system inherits symmetries from the original dynamical system [36].

2.3.4 Bifurcation problems and Lyapunov-Schmidt reduction

Consider a parametrized system (2.3). It is often the case that varying one or more of the system parameters changes the topological features of its flow. For example equilibria and invariant sets may appear, disappear, change in structure, or change in stability. These parametrized topological

changes in the flow are called *bifurcations*. Characterizing bifurcations in a nonlinear system is a major focus of dynamical systems theory. In this dissertation we are interested in particular in *local* bifurcations which arise when the stability of an equilibrium point, periodic orbit, or some other invariant set of (2.3) changes. *Global* bifurcations are another category of topological changes in the flow which cannot be fully described through local analysis. Global bifurcations are outside of the scope of this dissertation, and we refer the reader to [90, Chapter 6] for a treatment of this topic.

Equilibria of (2.3) correspond to choices of (\mathbf{x}, \mathbf{p}) which form the level sets

$$\mathbf{f}(\mathbf{x}, \mathbf{p}) = \mathbf{0}. \tag{2.11}$$

Suppose $(\mathbf{x}^*, \mathbf{p}^*)$ is a solution to (2.11) which corresponds to a hyperbolic equilibrium of the dynamics (2.3). By the implicit function theorem [82, Appendix 1] there exist neighborhoods $U \subset W$ of \mathbf{x}^* and $V \subset \mathbb{R}^{n_p}$ of \mathbf{p}^* as well as a smooth function $X : V \rightarrow U$ such that for every $\mathbf{p} \in V$, there is a unique $\mathbf{x} = X(\mathbf{p}) \in U$ which satisfies $\mathbf{f}(X(\mathbf{p}), \mathbf{p}) = \mathbf{0}$. However if $(\mathbf{x}^*, \mathbf{p}^*)$ has a zero eigenvalue, the implicit function theorem does not apply. Intuitively this means that a necessary condition for the number of equilibria of (2.3) to change as one or more of the parameters \mathbf{p} is varied is the existence of a solution $(\mathbf{x}^*, \mathbf{p}^*)$ to (2.11) at which the Jacobian matrix $D\mathbf{f}(\mathbf{x}^*, \mathbf{p}^*)$ is not invertible, i.e. has at least one zero eigenvalue. When such a point exists, it is referred to as a *singular point* or a *singularity*. The field of *singularity theory* is concerned with the study of properties of differentiable manifolds near their singular points [9]. In a dynamical systems context these manifolds are the sets of equilibria defined by (2.11), their singular points $(\mathbf{x}^*, \mathbf{p}^*)$ are called *bifurcation points*, and the parameters which are varied are the *bifurcation parameters* [82]. At a bifurcation point the system (2.3) is not *structurally stable*, which means that under arbitrarily small perturbation it can generate flow that is not topologically equivalent to the flow of the original unperturbed system. The set of all values (\mathbf{x}, \mathbf{p}) which satisfy (2.11) is referred to as the *solution set* or the *bifurcation diagram* of \mathbf{f} .

Suppose the system (2.3) has a bifurcation point $(\mathbf{x}^*, \mathbf{p}^*)$. As established in Section 2.3.3, any nontrivial dynamics that emerge near this bifurcation point will happen along a n_c -dimensional center manifold. When the system dimension n is greater than n_c , to study its bifurcations we must first derive an $(n_c + n_p)$ -dimensional reduced representation of the system that captures the behavior we are interested in describing. Center manifold reduction is a popular method that can be used for this purpose. However for high-dimensional systems it is easy to see how this approach becomes analytically inconvenient. For example, consider a system with one zero eigenvalue, $n_s = n - 1$ stable eigenvalues, and one parameter p . The local graph approximation of its center manifold (2.7)

reads

$$\mathbf{h}_s(x_c, p) = \begin{pmatrix} a_{11}p + b_{11}p^2 + b_{12}px_c + b_{13}x_c^2 + c_{11}p^3 + c_{12}p^2x_c + c_{13}px_c^2 + c_{14}x_c^3 + \mathcal{O}(4) \\ \vdots \\ a_{n_s1}p + b_{n_s1}p^2 + b_{n_s2}px_c + b_{n_s3}x_c^2 + c_{n_s1}p^3 + c_{n_s2}p^2x_c + c_{n_s3}px_c^2 + c_{n_s4}x_c^3 + \mathcal{O}(4) \end{pmatrix}$$

where a_{ij}, b_{ij}, c_{ij} are first, second, and third order coefficients which must be solved using the relationship defined in (2.9). Approximating a center manifold through third order thus involves solving for $8(n-1)$ coefficients from several nested series expansions, which is a cumbersome task for systems with large n . This task is complicated even further when the function $\mathbf{f}(\mathbf{x}, \mathbf{p})$ depends on arbitrary network structure, as will be the case for the models we consider in later chapters of this dissertation. For these reasons an alternative reduction method, the so-called *Lyapunov-Schmidt reduction*, often proves useful [208].

A center manifold reduction finds a subsystem of the original dynamics (2.3) which captures the local features of the flow we are interested in studying. In contrast, a Lyapunov-Schmidt reduction is a static reduction method which applies to the the solution sets of (2.11). Let $P \in \mathbb{R}^{n \times n}$ be the projection matrix $P : \mathbb{R}^n \rightarrow \text{range}(D\mathbf{f}(\mathbf{x}^*, \mathbf{p}^*))$ and $\mathcal{I} - P$ the projection onto its complement in \mathbb{R}^n . In a Lyapunov-Schmidt reduction we split the system (2.11) into the equivalent pair

$$P\mathbf{f}(\mathbf{x}, \mathbf{p}) = \mathbf{0}, \quad (2.12a)$$

$$(\mathcal{I} - P)\mathbf{f}(\mathbf{x}, \mathbf{p}) = \mathbf{0}. \quad (2.12b)$$

Restricted to the range of $D\mathbf{f}(\mathbf{x}^*, \mathbf{p}^*)$, (2.12a) is invertible and the implicit function theorem applies. We can solve for $n - n_c$ of the variables using a restriction (2.12a), and substitute for these variables in (2.12b). This eventually results in an n_c -dimensional parametrized system $\mathbf{g}(\mathbf{y}, \mathbf{p}) = 0$ whose bifurcation diagram near its bifurcation point $(\mathbf{y}^*, \mathbf{p}^*)$ is in one-to-one correspondence with the bifurcation diagram of (2.11) near its bifurcation point $(\mathbf{x}^*, \mathbf{p}^*)$. For a general treatment of Lyapunov-Schmidt reduction including its development for infinite-dimensional systems and for systems with symmetries see [82, Chapter VII].

In the simplest possible case, consider a dynamical system with one parameter $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, p)$ with bifurcation diagram $\mathbf{f}(\mathbf{x}, p) = 0$ and suppose the system has a bifurcation point $(\mathbf{x}, p) = (\mathbf{0}, 0)$ at which the Jacobian matrix $D\mathbf{f}(\mathbf{0}, 0)$ has a simple zero eigenvalue and no other eigenvalues on the imaginary axis. Let $\mathbf{v}, \mathbf{w} \in \mathbb{R}$ be the right and left null eigenvectors of $D\mathbf{f}(\mathbf{0}, 0)$, respectively and

define the k -th order directional derivative of \mathbf{f} along the ordered set of vectors $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ as

$$\begin{aligned} (d^k \mathbf{f})_{\mathbf{x}, p}(\mathbf{v}_1, \dots, \mathbf{v}_k) &= \frac{\partial}{\partial t_1} \dots \frac{\partial}{\partial t_k} \mathbf{f} \left(\mathbf{x} + \sum_{i=1}^k t_i \mathbf{v}_i, p \right) \Big|_{t_1 = \dots = t_k = 0} \\ &= \sum_{i, \dots, j=1}^n \frac{\partial^k \mathbf{f}}{\partial x_i \dots \partial x_j}(\mathbf{x}, p) (\mathbf{v}_1)_i \dots (\mathbf{v}_k)_j. \end{aligned} \quad (2.13)$$

For compactness of notation, define $J = D\mathbf{f}(\mathbf{0}, 0)$, let J^{-1} be the inverse of the restriction of J to its range, and let $P = \mathcal{I} - \frac{1}{\|\mathbf{w}\|} \mathbf{w} \mathbf{w}^T$ be the projection onto the range of J . Typically we cannot derive a closed-form expression for a Lyapunov-Schmidt reduction, and instead we approximate a Lyapunov-Schmidt reduction of $\mathbf{f}(\mathbf{x}, p)$ with a truncated series expansion

$$f(x, p) = ap + bpx + cx^2 + dx^3 + h.o.t. \quad (2.14)$$

with the coefficients of (2.14) defined as

$$a = \left\langle \mathbf{w}, \frac{\partial \mathbf{f}}{\partial p}(\mathbf{0}, 0) \right\rangle, \quad b = \left\langle \mathbf{w}, \left(d \frac{\partial \mathbf{f}}{\partial p} \right)_{\mathbf{0}, 0}(\mathbf{v}) - (d^2 \mathbf{f})_{\mathbf{0}, 0} \left(\mathbf{v}, J^{-1} P \frac{\partial \mathbf{f}}{\partial p}(\mathbf{0}, 0) \right) \right\rangle, \quad (2.15a)$$

$$c = \langle \mathbf{w}, (d^2 \mathbf{f})_{\mathbf{0}, 0}(\mathbf{v}, \mathbf{v}) \rangle, \quad d = \left\langle \mathbf{w}, (d^3 \mathbf{f})_{\mathbf{0}, 0}(\mathbf{v}, \mathbf{v}, \mathbf{v}) - 3(d^2 \mathbf{f})_{\mathbf{0}, 0}(\mathbf{v}, J^{-1} P (d^2 \mathbf{f})_{\mathbf{0}, 0}(\mathbf{v}, \mathbf{v})) \right\rangle. \quad (2.15b)$$

Computing a Lyapunov-Schmidt reduction approximation through third order for a particular system with a simple eigenvalue crossing therefore involves evaluating the four coefficients (2.15). This computation is often easier to carry out than solving for the $8(n-1)$ coefficients involved in a third-order center manifold approximation discussed in the previous section. The series (2.14) can be expanded to account for more higher-order terms following the procedure outlined in [82, Chapter I.3]. However for many problems, the terms described above are sufficient to capture the local topological features of the bifurcation diagram. The exact number of terms required in the expansion depends on the codimension of the bifurcation, see [82, Chapter II] for a detailed development. The computation of the coefficients (2.15) is further simplified when $\mathbf{f}(\mathbf{x}, p)$ has certain commonly encountered properties. For example, if $\mathbf{x} = \mathbf{0}$ is an equilibrium for all values of the parameter p , then $a = 0$ and the second-order dependence in d disappears. If \mathbf{f} has an odd symmetry in the state variable, i.e. $\mathbf{f}(-\mathbf{x}, p) = -\mathbf{f}(\mathbf{x}, p)$ then $a = 0, c = 0$, and the quadratic dependence in d similarly disappears. Computing a Lyapunov-Schmidt reduction in these cases mostly boils down to projecting several directional derivatives of \mathbf{f} onto the left null eigenspace of J . Furthermore if we choose \mathbf{v}, \mathbf{w} satisfying $\langle \mathbf{v}, \mathbf{w} \rangle > 0$, then stability of equilibria of the original system $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, p)$

is the same as the stability of the corresponding equilibria of $\dot{x} = f(x, p)$ [82, Theorem I.4.1]. A Lyapunov-Schmidt reduction therefore captures much of the same information about the original system as a center manifold reduction.

One disadvantage of a Lyapunov-Schmidt reduction is that we do not learn anything about the shape of a center manifold on which the bifurcating branches will appear, or about the dynamics of the system away from the equilibria. However, we can still infer the local structure of a center manifold close to the bifurcation point since it will be arbitrarily close to $\text{span}(\mathbf{v})$ by the center manifold theorem [90, Theorem 3.2.1]. More generally, the Lyapunov-Schmidt reduction can always be performed in a way which ensures that the reduced system of equations inherits the symmetries from the full system [82, Chapter VII, §3].

2.3.5 Elementary bifurcations

Once a reduced system is obtained, either through center manifold reduction or through Lyapunov-Schmidt reduction, we are ready to analyze the local bifurcation. Local bifurcations are typically classified by their *codimension*. In the singularity theory literature, the codimension of a bifurcation problem is defined as the number of parameters in its universal unfolding [82, Definition III.1.3] - see Section 2.3.6 for a discussion of unfoldings.¹ By this definition, the codimension of a bifurcation problem also coincides with the number of defining conditions for identifying the bifurcation, minus two [82, Corollary III.2.6]. We will elaborate on this connection further in this chapter. Bifurcations with a low codimension are most likely to be observed in models of realistic systems [82, Chapter IV]. In this section we describe several of the simplest possible bifurcations of codimension zero, one, and two, which are sometimes referred to as *elementary bifurcations*. The first three elementary bifurcations we will consider are bifurcations of equilibria, for which we will assume without loss of generality that the system $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, p)$ has a bifurcation point at $(\mathbf{x}, p) = (\mathbf{0}, 0)$, $n_c = 1$, and a one-dimensional reduced system has been obtained whose bifurcation diagram is described by

$$f(x, p) = 0 \tag{2.16}$$

with a smooth function $f : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ for which $f(0, 0) = 0$. Recall that a *diffeomorphism* is a homeomorphism where the map and its inverse are both differentiable.

¹In the classic presentation of dynamical systems theory in [90], the codimension of a bifurcation is defined to be the smallest number of parameters which must be varied to capture the bifurcation. This definition is different from the one presented in the text. By the definition of [90], all of the elementary bifurcations discussed in this chapter have codimension one; however by the definition in the text a saddle-node bifurcation has codimension zero, a transcritical bifurcation has codimension one, and a pitchfork bifurcation has codimension two.

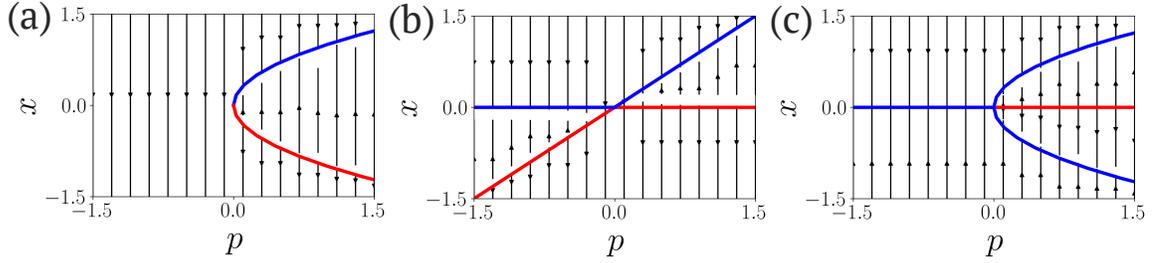


Figure 2.1: Bifurcation diagrams for the system $\dot{x} = g(x, p)$ defined by a normal form of (a) saddle-node bifurcation with $\varepsilon_1 = 1$, $\varepsilon_2 = -1$; (b) transcritical bifurcation with $\varepsilon = -1$; (c) pitchfork bifurcation with $\varepsilon_1 = 1$, $\varepsilon_2 = -1$. Blue lines are curves of stable equilibria, red lines are curves of unstable equilibria, and gray arrows indicate direction of the flow of the dynamics $\dot{x} = g(x, p)$ forward in time.

Definition 2.3.2 (Equivalence of bifurcation diagrams). *Bifurcation diagrams $f(x, p) = 0$ and $g(x, p) = 0$ with $f(0, 0) = g(0, 0) = 0$ are equivalent if they can be related locally near the origin through*

$$G(x, p)f(X(x, p), P(p)) = g(x, p)$$

where $G : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ is a nonzero and positive function, and $(X, P) : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R} \times \mathbb{R}$ is a local diffeomorphism which preserves the orientation of x and p . They are strongly equivalent if in addition, $P(p) = p$ [82, p.5].

This definition of equivalence is slightly more restrictive than the topological equivalence of Definition 2.3.1 since the bifurcation parameter cannot be transformed. Next we will state for each of the three elementary steady-state bifurcations a *normal form* $g(x, p)$, i.e. a simple equation that captures all of the “interesting” topological features of the bifurcation diagram. Furthermore we will describe conditions for the bifurcation diagram of the reduced system $f(x, p) = 0$, and therefore for the original system $\mathbf{f}(\mathbf{x}, p) = \mathbf{0}$, to be strongly equivalent to the bifurcation diagrams $g(x, p) = 0$ of each of these normal forms in the sense of Definition 2.3.2. For equivalence conditions for several other, more degenerate bifurcation problems see [82, Chapter IV.2] where the various equivalence conditions are discussed in the text and summarized in Table 2.3.² In general, a local bifurcation is referred to as *supercritical* if new topological features (e.g. new equilibria) appear in the flow of the dynamics for values of the bifurcation parameter p above its bifurcation value, it is *subcritical* if these features appear for p below its bifurcation value. When new features appear on both sides of the bifurcation point, the phenomenon is *transcritical*.

²An alternative analysis approach involves transforming a reduced system on a center manifold into its normal form through application of nonlinear coordinate transformations. For details of normal form theory, see [90, Chapter 3.3].

Saddle-node bifurcation

A *saddle-node bifurcation*, sometimes also called a *fold* or a *limit point* bifurcation, has a normal form

$$g(x, p) = \varepsilon_1 p + \varepsilon_2 x^2 \quad (2.17)$$

where $\varepsilon_1, \varepsilon_2 \in \{1, -1\}$. For example when $\varepsilon_1 = 1$ and $\varepsilon_2 = -1$ for $p < 0$, (2.17) has no solutions, and two solutions appear for $p > 0$; see Figure 2.1(a) for a bifurcation diagram. A bifurcation diagram of $f(x, p)$ is equivalent to the bifurcation diagram of (2.17) if and only if at $x = p = 0$ it satisfies the following conditions [82, Chapter II, Proposition 9.1]:

$$f(0, 0) = \left. \frac{\partial f}{\partial x} \right|_{(0,0)} = 0, \quad \varepsilon_1 = \text{sign} \left(\left. \frac{\partial f}{\partial p} \right|_{(0,0)} \right) \neq 0, \quad \varepsilon_2 = \text{sign} \left(\left. \frac{\partial^2 f}{\partial x^2} \right|_{(0,0)} \right) \neq 0. \quad (2.18)$$

In terms of the Lyapunov-Schmidt reduction coefficients (2.15), this means a saddle-node bifurcation is observed whenever $a \neq 0$ and $b \neq 0$, with criticality and stability of the solution branches determined by the signs of a, b . Recall that $f(0, 0) = \left. \frac{\partial f}{\partial x} \right|_{(0,0)} = 0$ are general conditions that establish existence of a bifurcation point. The rest of the conditions in (2.18) are the defining conditions for a Saddle-Node bifurcation. There are two defining conditions, $\varepsilon_1 \neq 0$ and $\varepsilon_2 \neq 0$, which means that the saddle-node bifurcation is codimension zero.

Transcritical bifurcation

A *transcritical bifurcation*, sometimes also called a *simple bifurcation* is characterized by the exchange of stability properties between two intersecting branches of equilibria. It is called transcritical because it is neither supercritical nor subcritical, as the solution branches persist for p above and below the bifurcation point. A normal form for the transcritical bifurcation is

$$g(x, p) = px + \varepsilon x^2 \quad (2.19)$$

where $\varepsilon \in \{1, -1\}$. For example when $\varepsilon_1 = 1$ and $\varepsilon_2 = -1$, the solution $x = 0$ is stable for $p < 0$ and unstable for $p > 0$, and the solution $x = p$ is unstable for $p < 0$ and stable for $p > 0$; see Figure 2.1(b) for a bifurcation diagram. A bifurcation diagram of $f(x, p)$ is equivalent to the bifurcation diagram of (2.17) if and only if at $x = p = 0$ it satisfies the following conditions [82, Chapter II,

Proposition 9.3]:³

$$f(0,0) = \frac{\partial f}{\partial x} \Big|_{(0,0)} = \frac{\partial f}{\partial p} \Big|_{(0,0)} = 0, \quad \det \begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial p} \\ \frac{\partial^2 f}{\partial x \partial p} & \frac{\partial^2 f}{\partial p^2} \end{pmatrix} \Big|_{(0,0)} < 0, \quad \varepsilon = \text{sign} \left(\frac{\partial^2 f}{\partial x^2} \Big|_{(0,0)} \right) \neq 0. \quad (2.20)$$

Assuming no quadratic dependence of $f(x, p)$ on the bifurcation parameter p the Hessian determinant condition simplifies to require $\frac{\partial^2 f}{\partial x \partial p} \Big|_{(0,0)} \neq 0$, and thus in terms of the Lyapunov-Schmidt reduction coefficients (2.15), a transcritical bifurcation is observed whenever $a = 0$, $b \neq 0$, and $c \neq 0$ with criticality and stability of the solution branches determined by the signs of b, c .⁴ Identifying a transcritical bifurcation requires three defining conditions in (2.20), in addition to the two general conditions $f(0,0) = \frac{\partial f}{\partial x} \Big|_{(0,0)} = 0$. Therefore a transcritical bifurcation has codimension one.

Pitchfork bifurcation

A *pitchfork bifurcation* gets its name from the distinctive shape of its solution branches. A normal form for the pitchfork bifurcation is

$$g(x, p) = \varepsilon_1 p x + \varepsilon_2 x^3 \quad (2.21)$$

where $\varepsilon_1, \varepsilon_2 \in \{1, -1\}$. For example, when $\varepsilon_1 = 1$ and $\varepsilon_2 = -1$, for $p < 0$ there is a single stable equilibrium at $x = 0$. For $p > 0$, the equilibrium at the origin is unstable and two stable solution branches appear which are reflection-symmetric about zero - see Figure 2.1(c). A bifurcation diagram of $f(x, p)$ is equivalent to the bifurcation diagram of (2.17) if and only if at $x = p = 0$ it satisfies the following conditions [82, Chapter II, Proposition 9.2]:

$$f(0,0) = \frac{\partial f}{\partial x} \Big|_{(0,0)} = \frac{\partial f}{\partial p} \Big|_{(0,0)} = \frac{\partial^2 f}{\partial x^2} \Big|_{(0,0)} = 0, \quad (2.22)$$

$$\varepsilon_1 = \text{sign} \left(\frac{\partial^2 f}{\partial p \partial x} \Big|_{(0,0)} \right) \neq 0, \quad \varepsilon_2 = \text{sign} \left(\frac{\partial^3 f}{\partial x^3} \Big|_{(0,0)} \right) \neq 0.$$

In terms of the Lyapunov-Schmidt reduction coefficients (2.15), this means a pitchfork bifurcation is observed whenever $a = c = 0$, $b \neq 0$, and $d \neq 0$, with criticality and stability of the solution

³In [82] the equivalence conditions are derived for a normal form $\varepsilon(x^2 - p^2)$ which is strongly equivalent to (2.19) with the same choice of ε . We state (2.19) in the text as it is a more commonly encountered normal form for a transcritical bifurcation.

⁴When there is a possible quadratic dependence on the bifurcation parameter in the general system, to compute the Hessian determinant in the equivalence conditions one must compute the coefficient of p^2 in the Lyapunov-Schmidt reduction. We did not include this coefficient in the terms we elected to show in (2.15), however the formula for this coefficient can be easily computed following the reduction procedure outlined in [82, Chapter I.3].

branches determined by the signs of b, d . Identifying a pitchfork bifurcation requires four defining conditions in (2.22), in addition to the two general conditions $f(0, 0) = \left. \frac{\partial f}{\partial x} \right|_{(0,0)} = 0$. Therefore a pitchfork bifurcation has codimension two.

Hopf bifurcation

The final elementary bifurcation we will consider is a *Hopf bifurcation* in which a limit cycle, i.e. a periodic solution, appears as a parameter is varied. Since this bifurcation is not a bifurcation of equilibria, it doesn't correspond to a zero eigenvalue in the Jacobian matrix of $\mathbf{f}(\mathbf{x}, p)$. Assume without loss of generality that $(\mathbf{x}, p) = (\mathbf{0}, 0)$ is an equilibrium of the system $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, p)$. Then $(\mathbf{0}, 0)$ is a *Hopf bifurcation point* if it satisfies the following properties:

- The Jacobian matrix $D\mathbf{f}(\mathbf{0}, 0)$ has a complex conjugate pair of eigenvalues $\pm i\omega(0)$;
- No other eigenvalues of $D\mathbf{f}(\mathbf{0}, 0)$ lie on the imaginary axis;
- Let $\lambda(p) = r(p) + i\omega(p)$, $\bar{\lambda}(p) = r(p) - i\omega(p)$ be the eigenvalues of $D\mathbf{f}(\mathbf{x}, p)$ which are smoothly parametrized by p for which $r(0) = 0$; then $\left. \frac{\partial r}{\partial p} \right|_{(\mathbf{0}, 0)} \neq 0$.

These conditions establish that at $(\mathbf{0}, 0)$, a pair of complex conjugate eigenvalues of $D\mathbf{f}(\mathbf{x}, p)$ crosses the imaginary axis with nonzero speed as the bifurcation parameter p is varied. Showing that a system satisfies these properties is sufficient to conclude the emergence of a family of periodic orbits of $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, p)$ in a Hopf bifurcation [90, Theorem 3.4.2]. A condition for criticality and stability of the resulting branches of periodic orbits is presented in [90]. This condition relies on first computing a reduction of the system to a center manifold, and then transforming it into normal form. Computing the necessary coefficient can be cumbersome for many systems, since it involves approximating the dynamics on a center manifold to third order. Here we present an alternative approach to classify the criticality and stability of the Hopf bifurcation which relies on Lyapunov-Schmidt reduction and singularity results developed in [82, Chapter VIII]. Before continuing this discussion, we establish a modified definition of equivalence for bifurcation diagrams that is restricted to bifurcation diagrams which have an odd symmetry.

Definition 2.3.3 (Z_2 -equivalence of bifurcation diagrams). *Bifurcation diagrams $f(x, p) = 0$ and $g(x, p) = 0$ with $f(0, 0) = g(0, 0) = 0$ are Z_2 -equivalent if they can be related locally near the origin through*

$$G(x, p)f(X(x, p), P(p)) = g(x, p)$$

where $G : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ is a nonzero and positive function, and $(X, P) : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R} \times \mathbb{R}$ is a local diffeomorphism which preserves the orientation of x and p . Additionally, the functions X is odd in x and the function G is even in x [82, Chapter VI, Definition 2.5].

As it turns out, whenever the eigenvalues of $D\mathbf{f}(\mathbf{0}, 0)$ satisfy the stated Hopf hypotheses, for values of (\mathbf{x}, p) near the bifurcation point there exists a one-to-one correspondence between orbits of small amplitude periodic solutions of $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, p)$ with period near $2\pi/\omega(0)$ and solutions on a reduced bifurcation diagram $f(x, p) = 0$ where $x \in \mathbb{R}$ and the function $f(x, p)$ has an odd symmetry in x [82, Chapter VIII, Theorem 2.1]. Furthermore, this reduced bifurcation equation $f(x, p)$ is Z_2 -equivalent to a pitchfork normal form $g(x, p) = \varepsilon_1 px + \varepsilon_2 x^3$ where $\varepsilon_1 = \text{sign} \left(\frac{\partial^2 f}{\partial x \partial p} \Big|_{(0,0)} \right) = \text{sign}(r(0))$ and $\varepsilon_2 = \text{sign} \left(\frac{\partial^3 f}{\partial x^3} \Big|_{(0,0)} \right)$ [82, Chapter VIII, Theorem 3.2].

We can use Lyapunov-Schmidt reduction to compute the necessary coefficients in the approximation of $\mathbf{f}(\mathbf{x}, p)$. Specifically, let $f(x, p) = apx + bx^3 + h.o.t..$ Then by [82, Chapter VIII, Proposition 3.3] the coefficients a, b are defined as

$$a = \frac{1}{2} \text{Re} \left(\left\langle \mathbf{w}, \left(d \frac{\partial \mathbf{f}}{\partial p} \right)_{\mathbf{0},0} (\mathbf{v}) \right\rangle \right) \quad (2.23a)$$

$$b = \frac{1}{4} \text{Re} \left(\left\langle \mathbf{w}, (d^2 \mathbf{f})_{\mathbf{0},0}(\mathbf{v}, \mathbf{b}_1) + (d^2 \mathbf{f})_{\mathbf{0},0}(\bar{\mathbf{v}}, \mathbf{b}_2) + \frac{1}{4} (d^3 \mathbf{f})_{\mathbf{0},0}(\mathbf{v}, \mathbf{v}, \bar{\mathbf{v}}) \right\rangle \right) \quad (2.23b)$$

where $\mathbf{b}_1, \mathbf{b}_2$ are defined through the relationships

$$D\mathbf{f}(\mathbf{0}, 0)\mathbf{b}_1 = -\frac{1}{2}(d^2 \mathbf{f})_{\mathbf{0},0}(\mathbf{v}, \bar{\mathbf{v}}), \quad (D\mathbf{f}(\mathbf{0}, 0) + 2i\mathcal{I})\mathbf{b}_2 = -\frac{1}{4}(d^2 \mathbf{f})_{\mathbf{0},0}(\mathbf{v}, \bar{\mathbf{v}}) \quad (2.24)$$

and the complex-valued eigenvectors \mathbf{v}, \mathbf{w} are defined as

$$D\mathbf{f}(\mathbf{0}, 0)\mathbf{v} = -i\mathbf{v}, \quad \mathbf{w}D\mathbf{f}(\mathbf{0}, 0) = i\mathbf{w} \quad (2.25)$$

and normalized to satisfy $\bar{\mathbf{w}}^T \mathbf{v} = 2, \mathbf{w}^T \mathbf{v} = 0$ which is always possible due to biorthogonality of right and left eigenvectors [82, Lemma 2.4]. Whenever the second derivatives in $\mathbf{f}(\mathbf{x}, p)$ vanish, such as when it is odd in the state variable, the coefficient b in (2.23) simplifies to

$$b = \frac{1}{16} \text{Re} \left(\langle \mathbf{w}, (d^3 \mathbf{f})_{\mathbf{0},0}(\mathbf{v}, \mathbf{v}, \bar{\mathbf{v}}) \rangle \right)$$

and the cumbersome computation of $\mathbf{b}_1, \mathbf{b}_2$ is not necessary.

Once we have computed coefficients a and b , we can easily establish stability of periodic orbits

of $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, p)$ due to the one-to-one correspondence between these periodic orbits and the zeros of $f(x, p)$. Specifically, suppose the origin $\mathbf{x} = 0$ is stable for $p < 0$ and unstable for $p > 0$, which means $a > 0$. If the nontrivial branches of the pitchfork bifurcation in the reduced bifurcation equation $f(x, p)$ appear supercritically ($b < 0$), then the periodic orbits in the original system also bifurcate supercritically and are stable. If instead the nontrivial bifurcation branches of the pitchfork bifurcation appear subcritically, then the periodic orbits in the original system also bifurcate subcritically and are unstable [82, Chapter VIII, Theorem 4.1]. In summary, to establish the criticality and the stability of a Hopf bifurcation, it is sufficient to compute the coefficient b in (2.23).

2.3.6 Unfolding theory and universal unfolding of a pitchfork bifurcation

The discussion in this section is a summary of some ideas developed in Chapters I and III of [82]. Unfolding theory aims at classifying the various ways in which different bifurcation diagrams change in response to perturbations or changes in parameters. Consider a bifurcation diagram $g(x, p) = 0$, and suppose there exists a function $G(x, p, \mathbf{r})$ with $\mathbf{r} \in \mathbb{R}^k$, $G : \mathbb{R} \times \mathbb{R} \times \mathbb{R}^k \rightarrow \mathbb{R}$ which satisfies

$$G(x, p, \mathbf{0}) = g(x, p). \quad (2.26)$$

Then we refer to G as a *k-parameter unfolding* of g , and to \mathbf{r} as *unfolding parameters*. An unfolding is thereby a parametrized perturbation of the original system. Suppose $G(x, p, \mathbf{r}), H(x, p, \mathbf{s})$ are unfoldings of g , with $\mathbf{r} \in \mathbb{R}^k$ and $\mathbf{s} \in \mathbb{R}^l$ where k and l can be distinct, and suppose that for every choice of $\mathbf{s} \in \mathbb{R}^l$ the bifurcation diagram of H (with bifurcation parameter p) is equivalent to a bifurcation diagram of G for some choice of parameters $\mathbf{r} \in \mathbb{R}^k$. Then we say that H *factors through* G . For a more mathematically precise definition of this idea, see [82, Chapter III, Definition 1.1].

An unfolding $G(x, p, \mathbf{r})$ of g is said to be *versal* if every other unfolding of g factors through it, i.e. if for various choices of the parameters \mathbf{r} , the bifurcation diagram of $G(x, p, \mathbf{r})$ reproduces all of the topologically distinct bifurcation diagrams which are possible to obtain through a perturbation of $g(x, p)$. An unfolding is then said to be *universal* if it is versal and the set of perturbation parameters \mathbf{r} has the smallest possible cardinality. Although not every bifurcation problem admits a universal unfolding, most bifurcation problems do. A key result of unfolding theory is the Universal Unfolding Theorem which establishes a criterion to identify the number of parameters required for a versal unfolding of a bifurcation problem to be universal [82, Chapter III, Theorem 2.3]. Universal unfoldings of the various elementary bifurcations of equilibria are summarized in [82, Chapter IV.3].

We conclude this chapter with a discussion of an unfolding problem which is particularly relevant

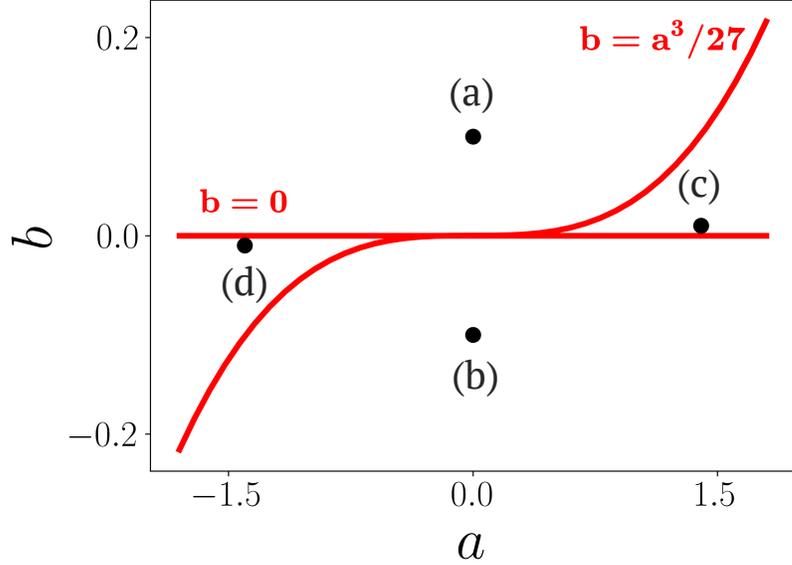


Figure 2.2: Four regions in $a-b$ parameter plane which generate topologically distinct persistent bifurcation diagrams in the pitchfork bifurcation universal unfolding (2.28); points labeled (a)-(d) correspond to the parameter combinations that generate the corresponding bifurcation diagrams shown in Figure 2.3.

to the analysis contained in this dissertation, unfolding of a pitchfork bifurcation. A universal unfolding of a pitchfork bifurcation contains two parameters. First we state a necessary and sufficient condition for an unfolding of a pitchfork bifurcation to be universal, found in [82, Chapter III, Proposition 4.4]. Suppose $f(x, p) = 0$ is a bifurcation diagram which is equivalent to a pitchfork bifurcation $g(x, p) = \pm px \pm x^3$. Suppose that $F(x, p, a, b)$ is a two-parameter unfolding of $f(x, p)$ with unfolding parameters a, b . Then F is a universal unfolding of f if and only if the following condition is satisfied at $x = p = a = b = 0$:

$$\det \begin{pmatrix} 0 & 0 & f_{xp} & f_{xxx} \\ 0 & f_{xp} & f_{pp} & f_{pxx} \\ F_a & F_{ax} & F_{ap} & F_{axx} \\ F_b & F_{bx} & F_{bp} & F_{bxx} \end{pmatrix} \neq 0 \quad (2.27)$$

where subscripts indicate partial derivatives.

Finally we describe the topologically distinct persistent bifurcation diagrams that can be recovered in a universal unfolding of a pitchfork bifurcation. To do this, we consider a supercritical pitchfork bifurcation normal form $g(x, p) = px - x^3$ and a universal unfolding of this normal form

$$G(x, p, a, b) = px + ax^2 - x^3 + b. \quad (2.28)$$

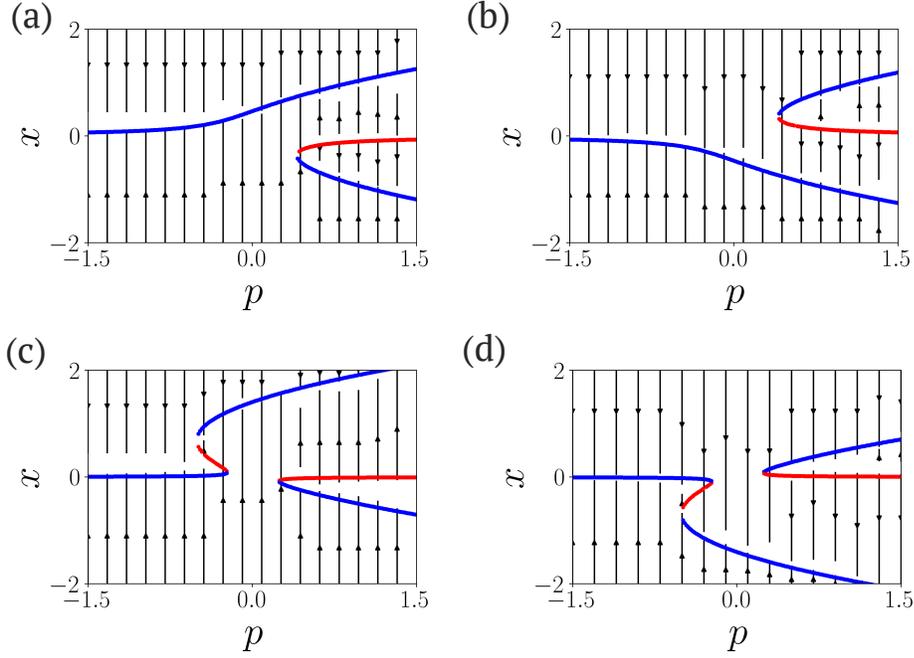


Figure 2.3: Four representative persistent bifurcation diagrams in a universal unfolding of a pitchfork normal form (2.28). (a) $a = 0, b = 0.1$; (b) $a = 0, b = -0.1$; (c) $a = 1.4, b = 0.01$; (d) $a = -1.4, b = -0.01$. Blue points are stable equilibria, red points are unstable equilibria, and gray arrows indicate direction of the flow of the dynamics $\dot{x} = G(x, p, a, b)$ forward in time.

There are four possible persistent bifurcation diagrams. Each of the topologically distinct bifurcation diagrams corresponds to parameter combinations in one of four regions of the $a - b$ parameter plane separated by the curves $b = 0, b = a^3/27$ - see Figure 2.2. In all four bifurcation diagrams, the symmetric pitchfork of Figure 2.1(c) breaks up into two continuous branches of equilibria, see Figure 2.3. One of these branches exists for all values of p , which we refer to as the primary branch; the second branch of equilibria appears in a saddle-node bifurcation at some bifurcation value $p > 0$. In a small neighborhood of the bifurcation point $p = 0$ of the unperturbed system, there is a unique nonzero stable equilibrium which belongs to the primary solution branch. When $b > 0$ the primary branch of solutions is positive, and when $b < 0$ it is negative. Furthermore when $-|a^3/27| < b < |a^3/27|$ and $\text{sign}(a) = \text{sign}(b)$, the primary solution branch exhibits two saddle-node bifurcations. In the first bifurcation, the equilibria lose stability and the branch folds back onto itself. At a second bifurcation point the branch regains stability and continues forward in p - see Figure 2.1(c) and (d).

Chapter 3

Social formation of beliefs as a nonlinear dynamical system

In this chapter we introduce the nonlinear model of belief and opinion formation whose formulation and analysis are the primary contributions of this dissertation. This chapter summarizes and expands on the work presented in [19], also contained in Chapter 9. The following parts of this chapter are modified directly from [19], with some sentences stated verbatim: section 3.1.2 and section 3.1.3.

3.1 Motivation

3.1.1 Formation of opinions and beliefs is a nonlinear process

Collective decision-making and belief formation in social and biological systems is the emergent outcome of many simultaneous distributed computations. In order to reach a decision about one or more options, topics, or candidates, individuals accumulate social and perceptual evidence about the value of these alternatives over time [79, 107]. This evidence accumulation process is typically understood to be nonlinear. For example, biologically plausible models of perceptual decision-making feature nonlinearities that mimic layers of processing in the brain [4, 24, 174, 205]. Nonlinearity is also necessary to capture critical phase transitions in behavior of animals on the move which make decisions among spatially embedded alternatives [39, 40, 189]. In human social systems, emergence of political polarization has recently been linked to irreversible tipping points in the dynamics of political opinions through a number of independently developed nonlinear models [11, 131, 201]. Adoption of new social norms is also characterized by nonlinear tipping points. Consensus in fa-

vor of a new social norm emerges spontaneously when the number of individuals who conform to this norm crosses a critical threshold [34, 89]. In similar vain, nonlinear models have been shown to capture the qualitative dynamics of echo chamber formation in Twitter debates [14], of opinion polarization on politically charged topics in survey data [15], of post-election government negotiations in parliamentary democracies [64], and of transitions between peace and war in international relationships between governments [142]. Informed by these observations, we argue that formation of opinions and beliefs in a group is best understood as a dynamic self-organizing process which is necessarily nonlinear.

In this dissertation we develop and analyze a mechanistic agent-based model of the process through which a group evaluates multiple alternatives to form opinions and make decisions. Motivated by our understanding of natural and sociopolitical decision-makers, we model this process as a nonlinear multi-agent dynamical system in which individuals accumulate evidence about options or topics over time through a social network. In our modeling we are also guided by recently developed theory which describes general properties of multi-agent, multi-option decisions in networked dynamical systems [68],[69]. According to this theory, the transition from indecision to decision in a group corresponds to a synchrony-breaking bifurcation in a nonlinear, multi-agent model. Agreement and disagreement decisions are both possible and general outcomes of social opinion formation, which are expected even when the communication graph is all-to-all. Furthermore, many properties of these emergent decisions are informed by the symmetries and the architecture of the influence network among the agents. The model we present allows for the study and design of emergent self-organizing features of social systems in nature and technology.

3.1.2 Classic models of social opinion formation

In developing a new agent-based model, we build on a multidisciplinary literature of *opinion dynamics* models. Opinion formation in a social network is classically modeled as a weighted-averaging process, as originally introduced by DeGroot [45]. In this framework an agent’s opinion $x_i \in \mathbb{R}$ reflects how strongly the agent supports an issue or topic of interest. The real-valued opinion is updated in discrete time as a weighted average of the agent’s own and other agents’ opinions, i.e.,

$$x_i(T + 1) = a_{i1}x_1(T) + \cdots + a_{iN_a}x_{N_a}(T) \tag{3.1}$$

where $a_{i1} + \cdots + a_{iN_a} = 1$ and $a_{ik} \geq 0$. The weight a_i describes the influence of the opinion of agent k on the opinion of agent i .

A key drawback of linear weighted-average models is that consensus among the agents is the only possible macroscopic outcome on a strongly connected network. As observed in [135], this necessarily happens because the attraction strength of agent i 's opinion toward agent k 's opinion increases linearly with the difference of opinions between the two agents. In other words, the more divergent the two agents' opinions are, the more strongly they are attracted to each other, which is paradoxical from an opinion formation perspective. Classical weighted-average opinion dynamic models therefore cannot explain emergence of persistent disagreement in social networks. In fact, a recent review notes that agreement is the inevitable outcome according to most formal models of social opinion formation [12]. However in real-world social systems disagreement is commonly encountered. In some cases disagreement can be an unwanted feature, such as partisan polarization on political issues. Disagreement can also be a desired outcome for the group, for example when individuals need to distribute themselves among different alternatives to forage or perform separate tasks.

To overcome these limitations of averaging models, a number of prominent variations on these models have been proposed. For example in “bounded confidence” models, agents average network opinions but delete communication links to any neighbors whose opinions are sufficiently divergent from their own [22,44,93,94]. In a similar spirit, “biased assimilation” models instead incorporate a self-feedback into the interaction weights of an averaging model [43,218]. This self-feedback accounts for an individual's bias towards evidence that conforms with its existing beliefs. The linear model and its variations have also been extended to the case of signed networks, where the weights a_{ik} can be negative [5,127,184]. Linear averaging dynamics on signed networks can support persistent disagreement, however these disagreement solutions require fine-tuned network architecture and distribution of weights. Small perturbations to the network weights in a signed network typically collapse disagreement solutions to a trivial consensus with $x_i = 0$ for all i , i.e. disagreement is not a structurally robust feature of these systems. In this chapter our goal is to derive a novel nonlinear extension of classical opinion dynamics models that captures consensus formation and persistent disagreement as general and robust outcomes of the opinion formation process, without the need to fine-tune network structure or to delete communication links.

3.1.3 Defining properties of nonlinear opinion formation model

We model opinion formation as a continuous-time dynamical system. The linear weighted-average discrete-time opinion dynamics (3.1) can equivalently be written as

$$x_i(T+1) = x_i(T) + \left(-x_i(T) + a_{i1}x_1(T) + \dots + a_{iN_a}x_{N_a}(T) \right).$$

This discrete-time update rule is the unit time-step Euler discretization of the continuous time linear dynamics

$$\dot{x}_i = -x_i + a_{i1}x_1 + \dots + a_{iN_a}x_{N_a}. \quad (3.2)$$

Observe that (3.1) and (3.2) have exactly the same steady states with the same (neutral) stability. These linear consensus dynamics (3.2) are determined by two terms: a weighted-average opinion-exchange term, modeling the pull felt by agent i toward the weighted group opinion, and a linear damping term, which can be interpreted as the agent's *resistance* to changing its opinion.

Our nonlinear extension of (3.2) incorporates the following features.

1. Opinion exchanges are saturated. We assume that there is a limit on the amount of influence that can be exerted on each agent by its social network. A natural way to incorporate this assumption mathematically is by allowing agents to apply a bounded saturating nonlinearity to the social information they accumulate about each option. This modeling assumption is also supported by the general features which appear in models of complex behavior in nature and technology. Saturating nonlinearities appear in virtually every natural and artificial signaling network due to bounds on action and sensing. For example, dynamics that evolve according to saturating interactions appear in spatially localized and extended neuronal population models of thalamo-cortical dynamics [211, 212], in Hopfield neural network models [98, 99, 145], in models of perceptual decision making [23, 205], and in control systems with sensor and actuator saturations [103, 126]. Saturated interactions between decision-makers also effectively bound the attraction between opinions, thus overcoming the linear weighted-average model paradox discussed in section 3.1.2.

2. Evaluation of multiple interdependent options. Classic models of opinion formation typically capture dynamics of opinions on a single topic or two options. Allowing for an arbitrary number of options in our model makes it relevant to a wide range of applications, for example, in task allocation problems where options represent tasks or in strategic settings where options represent strategies. We extend the model to multiple options by suitably generalizing the agent's opinion state space, analogous to existing multi-option averaging models such as [65, 72, 125, 147, 157, 160, 221].

3. Agents have allocable attention. Because an agent’s attention to exchanged opinions may be variable, we introduce, for each agent i , two parameters, $d_i > 0$ and $u_i \geq 0$, that weight the relative influence of the linear resistance term and the nonlinear opinion-exchange term, respectively. When the *resistance parameter* d_i dominates the *attention parameter* u_i , the agent is weakly attentive to other agents’ opinions. When u_i dominates d_i , the agent is strongly attentive to other agents’ opinions. A shift from a weakly attentive to a strongly attentive state can be induced, for instance, by a time-urgency (election day approaching) or a spatial-urgency (target getting closer) to form an informed collective opinion. The attention parameter u_i can also be used to model *social effort*, *excitability*, or *susceptibility* of agent i to social influence.

4. Agents have exogenous inputs. For each agent, we introduce an input parameter b_{ij} , which represents an input signal from the environment or a bias or predisposition that directly affects agent i ’s opinion of option j . For example, the input b_{ij} can be used to model the exogenous influence of agent i ’s initial opinions, as in [71], where agents hold on to their initial opinions (sometimes called “stubborn” agents as in [78]).

3.2 New model for dynamics of beliefs and opinions

In this section we formally introduce the nonlinear extension of classic opinion formation models which incorporates the features described in Section 3.1.3.

We extend the state space of each agent to account for multiple options in two complementary ways. First, consider a network of N_a agents. In scalar averaging models, the opinion state $x_i \in \mathbb{R}$ of agent i can represent favoring ($x_i > 0$) or disfavoring ($x_i < 0$) a single topic or option. Whenever $x_i = 0$, the agent is neutral on the topic, and the strength of favoring or disfavoring is proportional to the magnitude $|x_i|$. In the same spirit, in order to account for N_o topics or options we define the *belief state* or *value state* of agent i to be $\mathbf{Z}_i = (z_{i1}, \dots, z_{iN_o}) \in \mathbb{R}^{N_o}$, where z_{ij} is the *belief* or *valuation* of agent i on option j . As in the scalar case, $z_{ij} > 0$ (< 0) corresponds to favoring (disfavoring) an option, with $z_{ij} = 0$ representing neutrality. The total system state is $\mathbf{Z} = (\mathbf{Z}_1, \dots, \mathbf{Z}_{N_a}) \in \mathbb{R}^{N_a N_o}$.

In addition to attention u_i , resistance d_i , and bias b_{ij} parameters introduced in Section 3.1.3, we consider four classes of coupling weights between options:

1. Intra-agent, same-option coupling α_j^i ;
2. Intra-agent, inter-option coupling β_{jl}^i , with $j \neq l$;
3. Inter-agent, same-option coupling γ_j^{ik} , with $i \neq k$;

4. Inter-agent, inter-option coupling δ_{jl}^{ik} , with $i \neq k, j \neq l$.

In this notation, upper indices in the gains $\alpha_j^i, \beta_{jl}^i, \gamma_j^{ik}, \delta_{jl}^{ik}$ correspond to agents, and lower indices correspond to options. These four classes of interactions are illustrated in Figure 3.1, with dashed lines used for intra-agent couplings and solid lines used for inter-agent couplings.

The function $F_{ij}(\mathbf{Z})$ captures the net information about option j gathered by agent i , accounting for social interactions, external information and biases captured in b_{ij} , and its linear resistance to forming strong opinions:

$$F_{ij}(\mathbf{Z}) = -d_i z_{ij} + u_i \left(S_1 \left(\alpha_j^i z_{ij} + \sum_{k=1}^{N_a} \gamma_j^{ik} z_{kj} \right) + \sum_{l=1}^{N_o} S_2 \left(\beta_{jl}^i z_{il} + \sum_{k=1}^{N_a} \delta_{jl}^{ik} z_{kl} \right) \right) + b_{ij}. \quad (3.3)$$

Agents form beliefs about options by integrating information about each option according to

$$\dot{z}_{ij} = F_{ij}(\mathbf{Z}). \quad (3.4)$$

The functions S_1, S_2 in (3.3) are generic bounded sigmoidal saturating functions satisfying $S_q(0) = 0, S_q'(0) = 1, q \in \{1, 2\}$.¹ S_1 is applied to information along the same option dimension, and S_2 saturates cross-option influence. In principle, S_1 and S_2 can be the same function; however distinguishing between the two allows for the agents to have a different level of response to these two sources of influence. In all of the simulations shown in this dissertation we use saturating functions of the form

$$S_q(y) = A_q \tanh \left(\frac{1}{A_q} \left(y + \varepsilon_q \tanh(y^2) \right) \right) \quad (3.5)$$

with $A_q > 0, \varepsilon_q \geq 0$; however all of the results and discussion hold for generic saturating functions.

In the dynamics (3.4), the coupling gains $\alpha_j^i, \beta_{jl}^i, \gamma_j^{ik}, \delta_{jl}^{ik}$ characterize the qualitative properties of the interactions. To avoid redundancy between the self-coupling weight α_j^i and the resistance d_i , we assume that $\alpha_j^i \geq 0$. This means that whenever α_j^i is nonzero, it represents self-reinforcement of agent i about option j . The same-option coupling weight γ_j^{ik} captures the strength of the drive for social imitation of agent i towards agent k on option j ; $\gamma_j^{ik} > 0$ represents a cooperative interaction in which agent i seeks to mimic agent k , and $\gamma_j^{ik} < 0$ represents a competitive or antagonistic interaction in which agent i seeks to take on the opposite opinion of agent k . The inter-option coupling weights $\beta_{jl}^i, \delta_{jl}^{ik}$ capture interdependence between options that can arise, for example, from

¹The stated conditions are generic, as they can be imposed through rescaling and translation of variables. Additional nondegeneracy conditions (in the sense of [82]), such as $S_q''(0) \neq 0, S_q'''(0) \neq 0$, can be imposed when necessary to allow for the most general outcomes of opinion-formation dynamics, since symmetries in the saturation function impose additional restrictions on solutions of the dynamics (3.4),(3.7); the condition $S_q''(0) \neq 0$ is not necessary when $N_o = 1$ for the value dynamics (3.4) or when $N_o = 2$ for the constrained opinion dynamics (3.7).

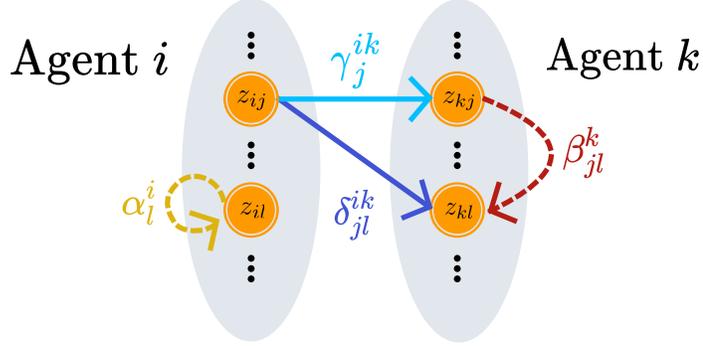


Figure 3.1: Illustration of the four classes of coupling weights in the model (3.3). The arrow directions use the sensing convention: a connection from z_{ij} to z_{kl} means $F_{ij}(\mathbf{Z})$ depends on z_{kl} , modulo the labeled gain.

a structured belief system, or from some logical or environmental constraints on the options. The role and interpretation of $\beta_{jl}^i, \delta_{jl}^{ik}$ will be explored in more detail in Chapter 5.

Next we present a second way to extend the state space of scalar opinion dynamics to multiple options. In scalar opinion dynamics, the state x_i can be alternatively interpreted as the decision variable for choosing between two “mutually-exclusive” options - for example in [43, 87]. Then when $x_i = 0$, agent i is undecided. When $x_i > 0 (< 0)$ agent i favors option 1 (2) and disfavors option 2 (1). This formalism is equivalent to the agent’s state space consisting of two scalar variables \tilde{z}_{i1} and \tilde{z}_{i2} constrained by a condition $\tilde{z}_{i1} + \tilde{z}_{i2} = 0$. Extending this convention to N_o options, we define each agent’s *opinion state* $\tilde{\mathbf{Z}}_i = (\tilde{z}_{i1}, \dots, \tilde{z}_{iN_o})$, which satisfy the constraint

$$\tilde{z}_{i1} + \dots + \tilde{z}_{iN_o} = 0 \quad (3.6)$$

for all $i = 1, \dots, N_a$. The constraint (3.6) allows us to interpret each agent’s state space as the $(N_o - 1)$ -dimensional simplex. This is because bounded dynamics on a trapping region in the affine subspace defined by (3.6) that contains the origin can be mapped to the standard simplex through a simple translation of the midpoint and a rescaling of the trajectories. For details on this interpretation and a formal statement mapping the dynamics constrained by (3.6) to the simplex, see [19, Appendix A]. Define $P_0 = \mathcal{I}_{N_o} - \frac{1}{N_o} \mathbf{1}_{N_o} \mathbf{1}_{N_o}^T$ to be the projection onto this affine subspace $\mathbf{1}_{N_o}^\perp \subset \mathbb{R}^{N_o}$. The total opinion state of the system is $\tilde{\mathbf{Z}} = (\tilde{\mathbf{Z}}_1, \dots, \tilde{\mathbf{Z}}_{N_a}) \in V$ where the state space $V = \underbrace{\mathbf{1}_{N_o}^\perp \times \dots \times \mathbf{1}_{N_o}^\perp}_{N_a \text{ times}} \subset \mathbb{R}^{N_a N_o}$. Then the dynamics of the opinion state of each agent are given by the projection of the evidence in favor of each option onto the opinion space V :

$$\dot{\tilde{\mathbf{Z}}}_i = P_0 \mathbf{F}_i(\tilde{\mathbf{Z}}) \quad (3.7)$$

where $\mathbf{F}_i(\tilde{\mathbf{Z}}) = (F_{i1}(\tilde{\mathbf{Z}}), \dots, F_{iN_o}(\tilde{\mathbf{Z}}))$. The system (3.7) is well-defined in the sense that it V is forward-invariant for the dynamics, its trajectories are bounded, and there exists a simple linear rescaling that reinterprets the dynamics of each agent's opinion state on a classical simplex Δ_{N_o-1} where $\mathbf{y} \in \Delta_{N_o-1}$ satisfies $y_j \geq 0$, $\sum_{j=1}^{N_o} y_j = K > 0$. For a detailed development of this well-definedness, see [19, Appendix A]. We show in [19, Section 3.4] that several prominent linear and nonlinear models of opinion formation are recovered from (3.7) with various constraints on the model parameters.

The model (3.3), (3.4), (3.7) can also be used more coarsely, with each belief state z_{ij} modeling average belief on option j in a subset of the agents on the network. Formally this is a valid generalization whenever the structure of the coupling weights partitions the agent vertices into approximately synchronized clusters. We do not develop this theory here and instead refer the reader to [2, 75, 154, 177, 188, 191] for a detailed treatment of cluster synchronization on dynamic networks and to [19, Section III-E] for a model reduction of (3.7) to a clustered manifold. These observations illustrate the versatility in the potential applications of our model to capture dynamics of opinion formation at different scales. For example, in [120] we use the belief states in this model to represent average ideological positions among the elite members of the two political parties in the United States Congress.

Finally, we state the following observation which connects the dynamics of beliefs (3.4) to the dynamics of constrained opinions (3.7).

Proposition 3.2.1. *If $\mathbf{Z}(t)$ is a solution of (3.4), then $\tilde{\mathbf{Z}}(t) = P_0\mathbf{Z}(t)$ is a solution of (3.7).*

Proposition 3.2.1 follows trivially from comparing (3.4) and (3.7). This observation allows us to focus our analysis on the unconstrained value dynamics (3.4), and interpret the results on the opinion simplex when necessary via a simple projection.

3.3 Indecision-breaking bifurcations

According to the model-independent theory [68],[69], one of the key features of nonlinear value-formation and decision-making models is the bifurcation of opinionated solution branches from a fully synchronous state of network indecision. In the models (3.4),(3.7) this fully synchronous state corresponds to the trivial neutral state $\mathbf{Z} = \mathbf{0}$ ($\tilde{\mathbf{Z}} = \mathbf{0}$, respectively) which is an equilibrium of the dynamics whenever $\mathbf{b} = \mathbf{0}$ ($P_0\mathbf{b} = \mathbf{0}$). In this section we illustrate that such indecision-breaking bifurcations are indeed a feature of our proposed model which can be observed robustly

across network architectures and specific parameter selections, and for any number of options under evaluation. Without loss of generality we focus our discussion on belief or value formation according to (3.4).

First, we present two motivating simulations that illustrate the network indecision-breaking behavior for a $N_a = 5$ agents evaluating several options according to (3.4). In Figures 3.2 and 3.3 we show the trajectories over time of the value states $z_{ij}(t)$ for agents evaluating 3 and 8 options, respectively, for 200 time steps of a simulation. Let $\mathcal{N}(a, b)$ denote the normal distribution centered at a , with standard deviation b . The following list summarizes the choice of parameters for these simulations:

- The communication weights between agents were generated randomly, with each α_j^i drawn from $\mathcal{N}(1, 0.1)$, and each $\beta_{jl}^i, \gamma_j^{ik}, \delta_{ji}^k$ was randomly drawn from $\mathcal{N}(0, 0.1)$;
- The resistance d_i for each agent was randomly drawn from $\mathcal{N}(1, 0.05)$;
- Agents have randomly generated small biases, with each b_{ij} drawn from $\mathcal{N}(0, 0.01)$;
- Initial conditions are randomized, with each $z_{ij}(0)$ drawn from $\mathcal{N}(0, 0.01)$;
- Each attention parameter u_i was drawn randomly from $\mathcal{N}(u_0, 0.01)$, with $u_0 = 0.57$ for part (a), and $u_0 = 0.6$ for part (b) in both figures;
- For the two saturation functions (3.5), $A_1 = 1, A_2 = 0.5, \varepsilon_1 = 0.5, \varepsilon_2 = 0.5$.

The same initial conditions and parameter selections were used to generate plots in part (a) and (b) of the two figures, with the exception of the value of u_0 as described.

In Figures 3.2 and 3.3 we observe a similar phenomenon. When the attention to social interactions u_0 is at its lower value in part (a), network beliefs z_{ij} remain small over time, slightly offset from zero due to the small additive biases. When u_0 is increased by a small amount in part (b) the network opinions evolve away from zero and settle on an equilibrium state in which many of the agents form strong beliefs. At this equilibrium, the final belief states of most agents are much greater in magnitude than their small biases b_{ij} . Relatively weak coupling of agents' belief states is sufficient to generate this transition. These simulations illustrate that whenever the average attention on the network is below some threshold, the linear resistance in (3.3) dominates and agents do not form strong opinions. Once this threshold is crossed, the influence of the social interactions dominates. In the following theorem we prove that these observations hold in general for the dynamics (3.3),(3.4) by analyzing stability of the indecision equilibrium $\mathbf{Z} = \mathbf{0}$. This theorem generalizes [19, Theorem IV.1] and motivates the rest of the analysis presented in this dissertation.

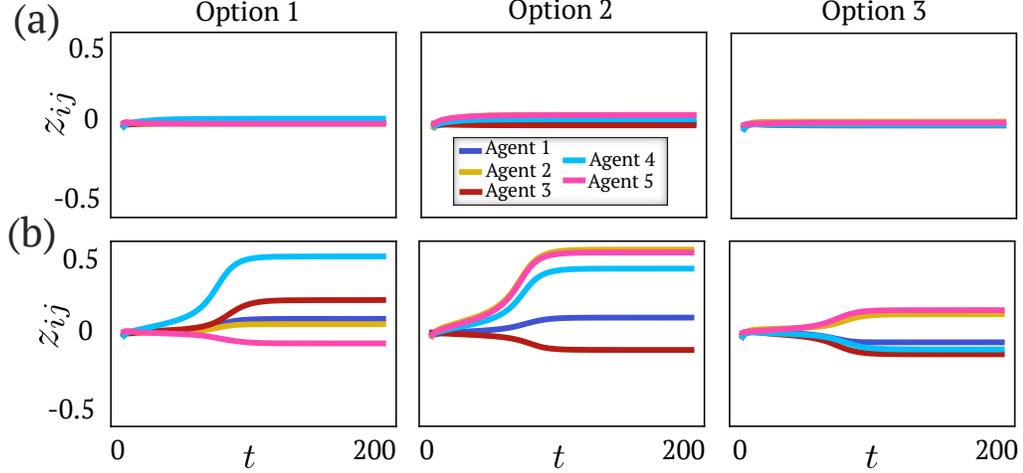


Figure 3.2: Trajectories of belief states $z_{ij}(t)$ of 5 agents evaluating 3 options according to (3.4) with randomized parameters; (a) average network attention level is below bifurcation threshold and beliefs of all agents remain small; (b) average network attention level is above bifurcation threshold and agents form strong beliefs.

Theorem 3.3.1 (Stability of Indecision Equilibrium). *Consider the model (3.3),(3.4) with $b_{ij} = 0$, let $u_i = u \cdot \tilde{u}_i$ for all $i = 1, \dots, N_a$, where $u \geq 0$, $u_i > 0$, and \otimes is the Kronecker product of matrices. Let $J(u) = -\text{diag}(d_1, \dots, d_{N_a}) \otimes \mathcal{I}_{N_o} + u\tilde{J}$ be the Jacobian matrix of the dynamics evaluated at the neutral equilibrium $\mathbf{Z} = \mathbf{0}$ for some u , where entries of the matrix \tilde{J} are of the form $\tilde{u}_i \alpha_j^i, \tilde{u}_i \beta_{jl}^i, \tilde{u}_i \gamma_j^{ik}, \tilde{u}_i \delta_{jl}^{ik}$. Assume \tilde{J} has at least one eigenvalue with positive real part. Then there exists a critical value $u^* > 0$; whenever $u < u^*$, the neutral equilibrium $\mathbf{Z} = \mathbf{0}$ is locally exponentially stable, and whenever $u > u^*$ it is unstable.*

Proof. Let $\Delta = \text{diag}(d_1, \dots, d_{N_a}) \otimes \mathcal{I}_{N_o}$. When $u = 0$, the system of ODEs (3.3),(3.4) reduces to the linear system $\dot{\mathbf{Z}} = -\Delta \mathbf{Z}$ and the origin is globally exponentially stable. Eigenvalues of $J(u)$ are continuous functions of u [124] and by this continuity $J(u)$ is Hurwitz for sufficiently small $u > 0$, which means the origin is locally exponentially stable. Define $\hat{J}(\epsilon) = \epsilon \Delta + \tilde{J}$. When $\epsilon = 1/u$, the eigenvalues of $\hat{J}(\epsilon)$ and of $J(u)$ are related by a factor of $\epsilon > 0$. Note that $\hat{J}(0)$ has at least one eigenvalue with positive real part, and by eigenvalue continuity the sign of the real part of the eigenvalues is preserved for sufficiently small $\epsilon > 0$. Therefore for some sufficiently large $u > 0$, $J(u)$ has at least one eigenvalue with positive real part and $\mathbf{Z} = \mathbf{0}$ is unstable. Existence of critical value u^* at which the origin loses stability follows once again from continuity of eigenvalues and from the intermediate value theorem. \square

Theorem 3.3.1 establishes existence of a critical level of attention $u = u^*$, at which the system (3.3),(3.4) undergoes a local bifurcation as the origin loses stability. This bifurcation can be observed

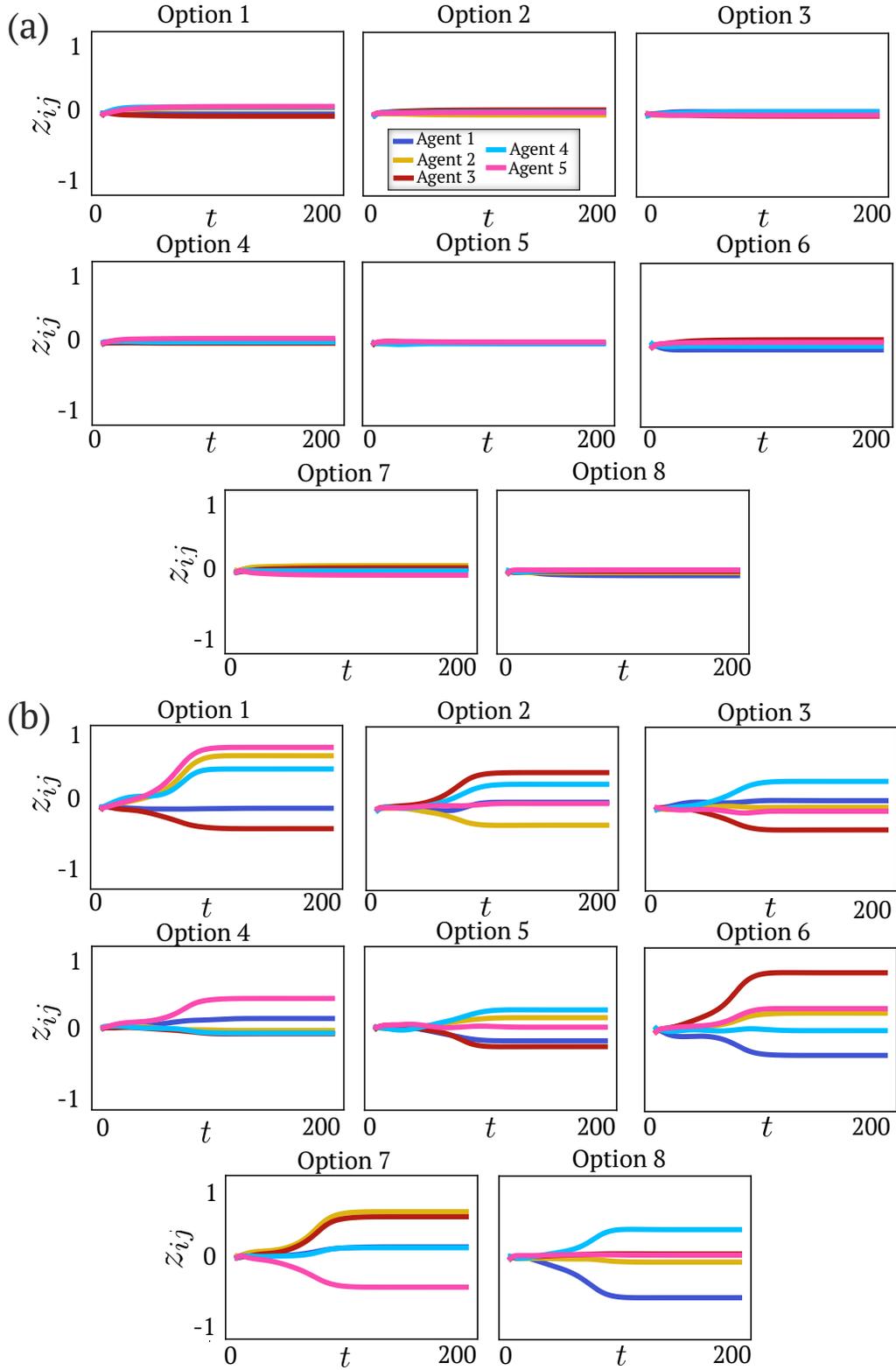


Figure 3.3: Trajectories of belief states $z_{ij}(t)$ of 5 agents evaluating 8 options according to (3.4) with randomized parameters; (a) average network attention level is below bifurcation threshold and beliefs of all agents remain small; (b) average network attention level is above bifurcation threshold and agents form strong beliefs.

whenever \tilde{J} , a matrix containing the linearization about the origin of the nonlinear terms in (3.4), has at least one eigenvalue with positive real part. This necessary condition is easily satisfied, which is most easily seen for networks of agents with no self-reinforcement, $\alpha_j^i = 0$ for all i, j . In this case, $\text{Tr}(\tilde{J}) = \sum_{i=1}^{N_a} \sum_{j=1}^{N_o} \tilde{u}_i \alpha_j^i = 0$ which means either that all of the eigenvalues of \tilde{J} lie on the imaginary axis, or more likely that at least one of the eigenvalues has a positive real part. More generally, \tilde{J} has an eigenvalue in the right half-plane if and only if its symmetrization $\frac{1}{2}(\tilde{J} + \tilde{J}^T)$ has a positive real eigenvalue. In turn this property is generic for symmetric square matrices as a consequence of Wigner's semicircle law [198], and more specifically for sparse symmetric square matrices by the analysis in [116].

The indecision-breaking bifurcation of Theorem 3.3.1 typically results in the emergence of new equilibria, such as the configurations the network settles on in part (b) of the Figures 3.2 and 3.3. Other possible outcomes include emergence of periodic oscillations of opinions, and other more complex dynamics on attracting invariant sets. In the next couple of chapters of this dissertation we investigate how the choice of network structure, communication weights, and biases in the model informs the properties these indecision-breaking bifurcations.

3.4 Homogeneous parameters

In its full generality, the model (3.3),(3.4) involves many parameters. However, it is often convenient to consider this model in its *homogeneous regime*, which we now describe. First, we assume that agents on the network are identical in their resistance to forming beliefs and in the amount of attention they allocate towards social interactions, i.e.

$$d_i := d > 0, \quad u_i := u \geq 0, \quad i = 1, \dots, N_a.$$

Next, we decompose the structure of the network coupling weights into the influence of two signed directed graphs that are fundamental to describing the belief formation process. The first of these is the *communication graph* among agents, $\mathcal{G}_a = (\mathcal{V}_a, \mathcal{E}_a, s_a)$ where $\mathcal{V}_a = \{1, \dots, N_a\}$ is the vertex set corresponding to the N_a agents, \mathcal{E}_a is the edge set, and $s_a : \mathcal{E}_a \rightarrow \{-1, 1\}$ is the *signature* of the communication graph \mathcal{G}_a . When $e_{ik} \in \mathcal{E}_a$, agent k is a neighbor of agent i and the belief state of agent k influences the belief formation of agent i . When $s_a(e_{ik}) = 1$, agent i is *cooperative* towards agent k , and whenever $s_a(e_{ik}) = -1$ it is *competitive* or *antagonistic* towards agent k . We assume that \mathcal{G}_a is *simple*, i.e. contains no self-loops $e_{ii} \notin \mathcal{E}_a$ for all $i \in \mathcal{V}_a$, and there is at most one edge

e_{ik} in \mathcal{E}_a that begins at vertex i and ends at vertex k for all $i, k \in \mathcal{V}_a$. The signed adjacency matrix of the communication graph is the matrix $A_a \in \mathbb{R}^{N_a \times N_a}$ whose entries are defined as $(A_a)_{ik} = 0$ if $e_{ik} \notin \mathcal{E}_a$ and $s_a(e_{ik})$ otherwise.

The second fundamental graph $\mathcal{G}_o = (\mathcal{V}_o, \mathcal{E}_o, s_o)$ encodes the interdependence of the various options. We refer to \mathcal{G}_o as the *belief system graph*, analogous to the terminology utilized in [72, 221]. Analogously we can consider \mathcal{G}_o to represent some context-dependent *social norms* [92]. The nodes in the vertex set $\mathcal{V}_o = \{1, \dots, N_o\}$ correspond to distinct options, or topics, which are evaluated by the agents. An edge $e_{jl} \in \mathcal{E}_o$ thereby signifies that formation of beliefs about option j is affected by the beliefs about option l . The signature function $\sigma_o : \mathcal{E}_o \mapsto \{-1, 1\}$ describes whether two options are positively or negatively aligned according to the belief system. We also assume that \mathcal{G}_o is simple, and we define the signed adjacency matrix $A_o \in \mathbb{R}^{N_o \times N_o}$ whose entries are defined as $(A_o)_{jl} = 0$ if $e_{jl} \notin \mathcal{E}_o$ and $s_o(e_{jl})$ otherwise. We assume that the belief system graph is inherent to the options that are being evaluated, and thereby shared by all agents in the group.

Using these definitions, we specialize the coupling weights in the model as

$$\alpha_j^i = \alpha, \quad \beta_{jl}^i = \beta(A_o)_{jl}, \quad \gamma_j^{ik} = \gamma(A_a)_{ik}, \quad \delta_{jl}^{ik} = \delta(A_a)_{ik}(A_o)_{jl}.$$

In this specialization we assume homogeneity in the magnitude of the coupling gains $\alpha, \beta, \gamma, \delta \geq 0$ among the agents. These gains correspond to the four distinct classes of interactions, as illustrated by the four differently colored arrows in Figure 3.1. Heterogeneity on the network therefore arises from the architecture of the communication and belief system graphs.

With these assumptions in place, the model of belief formation (3.3),(3.4) becomes

$$\begin{aligned} \dot{z}_{ij} = & -dz_{ij} + u \left(S_1 \left(\alpha z_{ij} + \gamma \sum_{\substack{k=1 \\ k \neq i}}^{N_a} (A_a)_{ik} z_{kj} \right) \right. \\ & \left. + \sum_{\substack{l \neq j \\ l=1}}^{N_o} S_2 \left(\beta(A_o)_{jl} z_{il} + \delta(A_o)_{jl} \sum_{\substack{k=1 \\ k \neq i}}^{N_a} (A_a)_{ik} z_{kl} \right) \right) + b_{ij} := F_{ij}^h(\mathbf{Z}). \end{aligned} \quad (3.8)$$

Importantly, (3.8) still supports potential heterogeneity in the external information about options that is captured by the distributed biases b_{ij} . For completeness of presentation we also state (3.8) in vector-matrix form:

$$\dot{\mathbf{Z}} = -d\mathbf{Z} + u\mathbf{S}_1(((\alpha\mathcal{I}_{N_a} + \gamma A_a) \otimes \mathcal{I}_{N_o})\mathbf{Z}) + \sum_{l=1}^{N_o} \mathbf{S}_2(((\beta\mathcal{I}_{N_a} + \gamma A_a) \otimes M_l)\mathbf{Z}) + \mathbf{b} \quad (3.9)$$

where $\mathbf{Z} \in \mathbb{R}^{N_a N_o}$ is the state vector, $\mathbf{S}_m(\mathbf{y}) = (S_m(y_1), \dots, S_m(y_n))$ for $\mathbf{y} \in \mathbb{R}^n$, and $M_l \in \mathbb{R}^{N_o \times N_o}$ is the matrix whose column l coincides with column l of A_o , with zero entries in all other columns. Analogously, the dynamics of constrained opinions in the homogeneous regime become

$$\dot{\tilde{\mathbf{Z}}}_i = P_0 \mathbf{F}_i^h(\tilde{\mathbf{Z}}). \quad (3.10)$$

There are several reasons that make the homogeneous model (3.8) compelling. First of all, in this form the model contains a small number of parameters which carry a clear interpretation in the context of belief formation. This makes the homogeneous model highly amenable to rigorous and interpretable analysis, as we will illustrate over the next several chapters of this dissertation. Second of all, homogeneity of agents in a group is often a valid assumption for the study and design of collective behavior. For instance, groups in nature are often composed of individuals that are indistinguishable from one another in their role [25, 26]. Analogously, robotic teams and other engineered collectives are commonly made up of interchangeable components. In fact homogeneity of system components is frequently used as motivation for the deployment of robotic swarms, in which many simple and easily replaceable robotic units work together to perform complex tasks [29, 111, 175]. Finally, rigorous understanding of the model behavior in its homogeneous regime paves the way for systematic study of the effects of heterogeneity, since small heterogeneity can be considered as a perturbation of the homogeneous system. We take advantage of this observation later on in this dissertation when we relax the homogeneity assumption to allow agents to dynamically modify some of their parameters.

Chapter 4

Belief-forming bifurcations: single topic or two mutually exclusive options

Most of the material presented in sections 4.1,4.2,4.3 is a summary of analysis that appears in print in various parts of [18–21] that are also included as Chapters 9,10,11,12 in Part II of this dissertation. Several of the propositions, theorems, figures, and figure captions in these sections are taken verbatim from these sources as cited, with minor edits for consistency. The surrounding discussion of the contributions is original, with exceptions noted explicitly in the text. The analysis in section 4.4 does not yet appear in print, however an early version of it was presented as part of my general examination in January 2019 and in the 2019 SIAM Conference on Applications of Dynamical Systems. I presented parts of analysis in this chapter in the 2020 SIAM Conference on the Life Sciences; in the 2020 UCLA Institute for Pure and Applied Mathematics, Mathematical Challenges and Opportunities for Autonomous Vehicles Program, Workshop IV: Social Dynamics beyond Vehicle Autonomy;; in the 2021 SIAM Conference on Applications of Dynamical Systems; in the 2021 American Control Conference; and in the 2021 IEEE Conference on Decision and Control.

4.1 Model specialization to scalar state variable

In this chapter we study local indecision-breaking bifurcations of equilibria in the homogeneous model (3.8) specialized to one topic, or two mutually exclusive options. We present two ways to

derive this specialization.

First, consider the homogeneous value dynamics (3.8) for a network of agents evaluating a single topic. In this case, the belief state \mathbf{Z}_i of each agent is simply the scalar state $z_{i1} \in \mathbb{R}$. To simplify notation, we relabel $z_{i1} = x_i$ and $b_{i1} = b_i$, with the network state $\mathbf{x} = (x_1, \dots, x_{N_a})$. Furthermore, we observe that for a single topic, the belief system graph \mathcal{G}_o is trivial and inter-option dependencies captured by saturations S_2 do not play a role in the dynamics (3.8). To further simplify notation, we drop the subscript from the saturation function S_1 . With these modifications, we arrive at the model

$$\dot{x}_i = -dx_i + uS \left(\alpha x_i + \gamma \sum_{\substack{k=1 \\ k \neq i}}^{N_a} (A_a)_{ik} x_k \right) + b_i. \quad (4.1)$$

The internal gain $\alpha \geq 0$ describes the strength of each agent's self-reinforcement of its own beliefs. The social gain $\gamma \geq 0$ describes the strength of each agent's drive to imitate opinions of the neighbors with which it has a cooperative relationship, and to reject the opinions of the neighbors it has an antagonistic relationship with. A version of (4.1) with no self-reinforcement and with inhomogeneous resistance weights d_i that reflect the in-degree of each agent on the communication graph \mathcal{G}_a has been studied in the literature as a biologically inspired model for robust consensus formation on two options [61, 63, 87]. A version of this model was also recently independently developed in the sociophysics literature as a model of echo chamber formation in online social networks [14, 74]. As mentioned in chapter 3, the nondegeneracy condition $S''(0) \neq 0$ is not necessary in the scalar opinion setting and it is natural to assume the saturation function S has an odd symmetry, $S(-y) = -S(y)$.

Alternatively consider the homogeneous model of opinion formation on the simplex for two mutually exclusive options (3.10). When $N_o = 2$, each agent's state space $\tilde{\mathbf{Z}}_i = (\tilde{z}_{i1}, \tilde{z}_{i2})$ is constrained by $\tilde{z}_{i1} + \tilde{z}_{i2} = 0$ and can be represented by a single state variable $x_i = \tilde{z}_{i1} = -\tilde{z}_{i2}$. Observe that $\dot{\tilde{z}}_{i1} = \frac{1}{2} \left(F_{i1}^h(\tilde{\mathbf{Z}}) - F_{i2}^h(\tilde{\mathbf{Z}}) \right)$, and the dynamics of the scalar opinion variable x_i reduce to

$$\dot{x}_i = -dx_i + u \left(\hat{S}_1 \left(\alpha x_i + \gamma \sum_{\substack{k=1 \\ k \neq i}}^{N_a} (A_a)_{ik} x_k \right) - \hat{S}_2 \left(\beta x_i + \delta \sum_{\substack{k=1 \\ k \neq i}}^{N_a} (A_a)_{ik} x_k \right) \right) + \hat{b}_i. \quad (4.2)$$

where $\hat{S}_1(y) = \frac{1}{2}(S_1(y) - S_1(-y))$, $\hat{S}_2(y) = \frac{1}{2}(S_2(y) - S_2(-y))$ are saturating functions with an odd symmetry and $\hat{b}_i = \frac{1}{2}(b_{i1} - b_{i2})$. For (4.2) we allow the inter-option gains $\beta, \delta \in \mathbb{R}$ to be negative as well as positive. Whenever $(A_o)_{12} = (A_o)_{21} = 0$, $\beta = \delta = 0$ and (4.2) reduces to (4.1).

When $\mathbf{b} = \mathbf{0}$, the Jacobian of (4.2) evaluated at the origin $\mathbf{x} = \mathbf{0}$ is

$$J = (-d + u(\alpha - \beta))\mathcal{I} + u(\gamma - \delta)A_a. \quad (4.3)$$

By inspection of (4.3) we see that as long as $\gamma \neq \delta$, the spectral properties of J are inherited from the communication graph \mathcal{G}_a . Specifically, whenever \mathbf{v} is a right eigenvector of A_a with eigenvalue λ , \mathbf{v} is also a right eigenvector of J with eigenvalue $\xi = -d + u(\alpha - \beta) + u\lambda(\gamma - \delta)$. The eigenvalues of A_a determine the critical value of attention u^* at which the origin loses stability according to Theorem 3.3.1. Furthermore, the local geometry of the center manifold along which the resulting bifurcation appears near $(\mathbf{x}, u) = (\mathbf{0}, u^*)$ is informed by an eigenspace of A_a .

In the following theorem we establish that a typical bifurcation of the origin in the model (4.1) is a pitchfork bifurcation that results in appearance of two new opinionated equilibria on the network.

Theorem 4.1.1 (Pitchfork Bifurcation). *[20, Theorem IV.1] Consider (4.1) and define*

$$u^*(\lambda) = \frac{d}{\alpha + \lambda\gamma}, \quad (4.4)$$

where λ is a simple real eigenvalue of adjacency matrix A_a . Let $\mathbf{v} = (v_1, \dots, v_{N_a})$ and $\mathbf{w} = (w_1, \dots, w_{N_a})$ be right and left unit eigenvectors, respectively, corresponding to λ . Assume that (i) for all eigenvalues $\xi \neq \lambda$ of A , $\text{Re}[\xi] \neq \lambda$; (ii) $\alpha + \lambda\gamma \neq 0$; (iii) $\langle \mathbf{w}, \mathbf{v}^3 \rangle \neq 0$. Let $f(z, u, \mathbf{b})$ be a Lyapunov-Schmidt reduction of (4.1) at $(\mathbf{x}, u, \mathbf{b}) = (\mathbf{0}, u^*, \mathbf{0})$.

A. Bifurcation problem $f(z, u, \mathbf{0}) = 0$ has a symmetric pitchfork singularity at $(z, u, \mathbf{b}) = (0, u^*, \mathbf{0})$. For values of $u > u^*$ and sufficiently small $|u - u^*|$, two branches of equilibria branch off from $\mathbf{x} = \mathbf{0}$ in a pitchfork bifurcation along a manifold tangent at $\mathbf{x} = \mathbf{0}$ to $\text{span}\{\mathbf{v}\}$.

When $\text{sign}\{\langle \mathbf{w}, \mathbf{v}^3 \rangle / \langle \mathbf{w}, \mathbf{v} \rangle\}(\alpha + \lambda\gamma) > 0$ (< 0) the bifurcation happens supercritically (subcritically) with respect to u . Whenever \mathbf{x}^* is an equilibrium of (4.2), $-\mathbf{x}^*$ is also an equilibrium.

B. Bifurcation problem $f(z, u, \mathbf{b}) = 0$ is an N_a -parameter unfolding of the symmetric pitchfork, and $\frac{\partial f}{\partial b_i}(z, u, \mathbf{b}) = w_i$.

Using Theorem 4.1.1 we can describe the properties of indecision-breaking bifurcations of (4.1) on a communication graph \mathcal{G}_a whenever its signed adjacency matrix A_a has a dominant eigenvalue λ^* . Suppose $u = u^*(\lambda^*)$ and \mathbf{w}, \mathbf{v} are the left and right unit eigenvectors of A_a corresponding to λ^* . We always assume that the choice of \mathbf{w}, \mathbf{v} satisfies $\langle \mathbf{w}, \mathbf{v} \rangle > 0$.

1. When agents are unbiased, at the bifurcation point $u = u^*(\lambda^*)$ the indecision equilibrium $\mathbf{x} = \mathbf{0}$ loses stability and two stable branches of opinionated equilibria appear supercritically.

For values of u near $u^*(\lambda^*)$, the two equilibria $\pm \mathbf{x}^*$ can be approximated to first order as multiples of the eigenvector \mathbf{v} . The eigenvector \mathbf{v} thus informs the pattern of opinions the network takes on at bifurcation. For example, if $\text{sign}(v_i) = \text{sign}(v_k)$ then agents i and k will agree on their choice of option. Furthermore, if $|v_i| > |v_k|$ then agent i will form a stronger belief at equilibrium than agent k .

2. When agents have nonzero small biases b_i the direction of unfolding of the pitchfork bifurcation near $u = u^*(\lambda^*)$ is determined by the quantity $\langle \mathbf{b}, \mathbf{w} \rangle$, the projection of the network biases onto the left null eigenspace of the Jacobian (4.3). Whenever $\langle \mathbf{b}, \mathbf{w} \rangle > 0 (< 0)$ the equilibria \mathbf{x}^* in the continuous branch of the pitchfork unfolding satisfy $\langle \mathbf{x}^*, \mathbf{v} \rangle > 0 (< 0)$. We can interpret the relative magnitude of the entries of \mathbf{w} as a measure of *influence* or *centrality* of the agent. For example if $|w_i| > |w_k|$ and agents i, k receive small inputs of equal magnitude $|b_i| = |b_k|$, in the projection $\langle \mathbf{b}, \mathbf{w} \rangle = \sum_{l=1}^{N_a} b_l w_l$ the quantity $b_i w_i$ dominates the quantity $b_k w_k$. Then the absence of other biases of the network $\text{sign}(\langle \mathbf{b}, \mathbf{w} \rangle) = \text{sign}(b_i w_i)$ and the bias of agent i determines the direction of unfolding due to its higher centrality ranking.

Figure 4.1 shows an example of a numerically generated bifurcation diagram illustrating a pitchfork bifurcation and its unfolding established by Theorem 4.1.1 on a small network. Note that the right eigenvector \mathbf{v} informs the opinion pattern at equilibrium because its span is a linear approximation of the center manifold on which new equilibria appear as the state of indecision loses stability. The linear term in the series expansion of this approximation dominates whenever $|u - u^*(\lambda^*)|$ is small, which means that any of the qualitative descriptions of the equilibria that are based on this linear analysis apply only in a small neighborhood of the bifurcation point. In practice, we observe the characterizations of equilibria in the linear regime typically persist for values of u far away from the bifurcation point. Usually the pattern of signs on the network predicted by \mathbf{v} remains unchanged, and the magnitudes of opinions of different agents get closer to one another as the nonlinearities in (4.4) saturate to their maximum value. However we do not derive a center manifold approximation beyond first order to justify this observation, and in general it would be difficult to carry out such a computation for an arbitrary network architecture. In the next theorem we establish that there is a range of values for the attention parameter u near the bifurcation point $u^*(\lambda^*)$ for which the three described equilibria are the only equilibria admitted by the system.

Theorem 4.1.2 (Uniqueness of equilibria). [18] Consider (4.1) on a signed graph \mathcal{G}_a ; suppose the signed adjacency matrix A_a has a simple leading eigenvalue $\lambda^* > 0$ and let λ_2 be an eigenvalue of A satisfying $\text{Re}(\lambda_2) \geq \text{Re}(\lambda_i)$ for all eigenvalues $\lambda_i \neq \lambda^*$ of A_a . 1) $\mathbf{x} = \mathbf{0}_N$ is globally asymptotically

stable on a forward-invariant compact set $\Omega \subset \mathbb{R}^n$ containing the origin $\mathbf{x} = \mathbf{0}_N$, for all $u \in [0, u^*(\lambda^*)$); 2) if $\text{Re}(\lambda_2) \geq -\alpha/\gamma$, $u \in (u^*(\lambda^*), u^*(\lambda_2))$, the only equilibria the system admits are $\mathbf{0}_N$, \mathbf{x}_1^* , and \mathbf{x}_2^* ; 3) if $\text{Re}(\lambda_2) < -\alpha/\gamma$, the only equilibria the system admits in Ω for all $u > u^*$ are $\mathbf{0}_N$, \mathbf{x}_1^* , and \mathbf{x}_2^* .

An analogous result to Theorems 4.1.1 and 4.1.2 can be established for (4.2) whenever $\gamma \neq \delta$, with the critical value of attention instead being

$$u^*(\lambda) = \frac{d}{\alpha - \beta + \lambda(\gamma - \delta)}. \quad (4.5)$$

In the next two sections we study properties of the equilibria that result from the pitchfork bifurcation in Theorem 4.1.1. We already established that the bias vector \mathbf{b} determines the unfolding of the bifurcation, therefore we mainly study properties of the equilibria of the symmetric pitchfork bifurcation for a network of unbiased agents. We focus our attention on the simpler model (4.1), but note that analogous results hold more generally for (4.2) with $\gamma \neq \delta$.

4.2 Networks of cooperative and competitive agents

First, we consider (4.1) on networks that have purely cooperative or purely competitive (i.e. antagonistic) relationships among the agents. Since all of the interactions have the same sign, in this section we constrain the communication graph \mathcal{G}_a to be unsigned with $(A_a)_{ik} \in \{0, 1\}$ for all $i, k = 1, \dots, N_a$, and instead absorb the sign of the social interactions into the social gain $\gamma \in \mathbb{R}$. With this modification, $\gamma > 0$ describes a network cooperative agents, and $\gamma < 0$ describes a network of competitive agents. For the results stated in this section, we establish the following assumptions and definitions:

- Assume the communication graph \mathcal{G}_o is strongly connected;
- $\lambda_{max} > 0$ is the Perron-Frobenius eigenvalue of A_a , with right and left unit eigenvectors \mathbf{v}_{max} , \mathbf{w}_{max} ; recall from the Perron-Frobenius theorem that we can choose $\mathbf{v}_{max} \succ \mathbf{0}$, $\mathbf{w}_{max} \succ \mathbf{0}$;
- Define the set of smallest real-part eigenvalues of A_a as $\Lambda_{min} = \{\lambda = \text{argmin}_{\mu \in \sigma(A_a)} \text{Re}(\mu)\}$;
- Whenever $|\Lambda_{min}| = 1$, let $\Lambda_{min} = \{\lambda_{min}\}$, where $\lambda_{min} \in \mathbb{R}$ with right and left unit eigenvectors \mathbf{v}_{min} , \mathbf{w}_{min} ; since $\langle \mathbf{v}_{max}, \mathbf{w}_{min} \rangle = \langle \mathbf{v}_{min}, \mathbf{w}_{max} \rangle = 0$, both \mathbf{v}_{min} , \mathbf{w}_{min} each have at least one positive and at least one negative component.

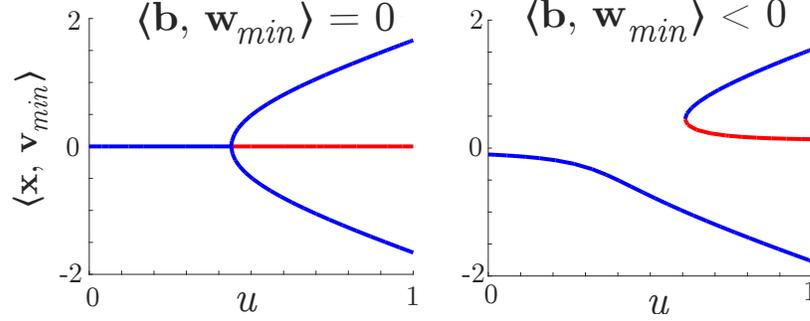


Figure 4.1: [19] Bifurcation diagrams showing the symmetric pitchfork bifurcation (left) and its unfolding (right) for two-option opinion dynamics (9.14) with $d = \alpha = 1$ in the disagreement regime ($\gamma = -1$) for three agents communicating over an undirected line graph. Blue (red) curves represent stable (unstable) equilibria. The vertical axis is the projection of the system equilibria \mathbf{x} onto \mathbf{w}_{min} ($\mathbf{w}_{min} = \mathbf{v}_{min}$ since the graph is undirected). Left: $\mathbf{b} = (0.2, 0, -0.2)$; right: $\mathbf{b} = -0.1\mathbf{w}_{min} + (0.2, 0, -0.2)$. Bifurcation diagrams generated with MatCont [47]. In the agreement regime, the diagrams look qualitatively the same with $\mathbf{w}_{min}/\mathbf{v}_{min}$ replaced with $\mathbf{w}_{max}/\mathbf{v}_{max}$ and \mathbf{b} modified appropriately (see Figure 1 in [21], included in his dissertation as Figure 11.1).

In the following theorem we establish an intuitive result: cooperation drives networked agents to agree on a choice of option, and competition or antagonism among the agents results in disagreement.

Theorem 4.2.1. [19],[20] *The following hold true for (4.1) with $b_i = 0$ for all $i = 1, \dots, N_a$:*

A. Cooperation leads to agreement: *If $\gamma > 0$, the neutral state $\mathbf{x} = 0$ is a locally exponentially stable equilibrium for $0 < u < u_a$ and unstable for $u > u_a$, with*

$$u_a = \frac{d}{\alpha + \gamma\lambda_{max}}. \quad (4.6)$$

At $u = u_a$, branches of agreement equilibria, $x_i \neq 0$, $\text{sign}(x_i) = \text{sign}(x_k)$ for all $i, k \in \mathcal{V}_a$, emerge in a pitchfork bifurcation along a center manifold tangent to $\text{span}(\mathbf{v}_{max})$;

B. Competition leads to disagreement: *Suppose $|\Lambda_{min}| = 1$. If $\gamma < 0$ the neutral state $\mathbf{x} = 0$ is a locally exponentially stable equilibrium for $0 < u < u_d$ and unstable for $u > u_d$, with*

$$u_d = \frac{d}{\alpha + \gamma\lambda_{min}}. \quad (4.7)$$

At $u = u_d$, branches of disagreement equilibria, $\text{sign}(x_i) = -\text{sign}(x_k)$ for at least one pair $i, k \in V$, $i \neq k$, emerge in a pitchfork bifurcation along a center manifold tangent to $\text{span}(\mathbf{v}_{min})$.¹

According to Theorem 4.2.1, whether a network reaches agreement or disagreement post bifurcation is determined by the signs of the interactions between agents rather than by the architecture of the network interconnections. In fact, agreement and disagreement can happen on the same network

¹For (4.5), the agreement regime is $\gamma - \delta > 0$ and the disagreement regime is $\gamma - \delta < 0$.

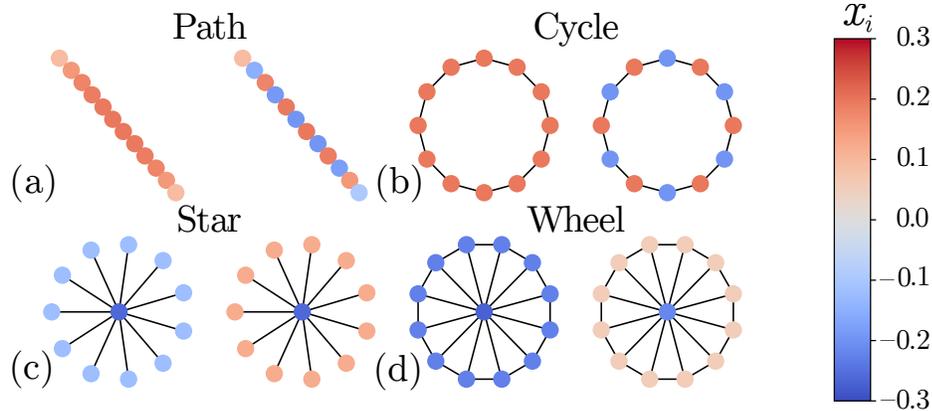


Figure 4.2: [20] Post-bifurcation patterns of agreement for purely cooperative agents with $\gamma > 0$ (left) and of disagreement for purely antagonistic agents with $\gamma < 0$ (right) on four small undirected networks.

- see Figure 4.2 for examples of agreement and disagreement equilibria that are reached by cooperative and antagonistic agents on several representative undirected networks that satisfy $|\Lambda_{min}| = 1$. In Figure 4.2 the color of each node represents the value of its opinion state x_i at equilibrium, with red nodes choosing option 1 and blue nodes choosing option 2. Darker nodes settled on stronger opinions than lighter nodes. Figure 4.2 illustrates that the network architecture clearly plays a role in selecting the pattern of agreement and disagreement that appears in this bifurcation, i.e. how the nodes are distributed across the two options in disagreement, and the relative strength of beliefs of individual nodes in relation to their neighbors. This happens because, as we observed from Theorem 4.1.1, these patterns of opinions at equilibrium follow the sign pattern of an eigenvector of the adjacency matrix of the graph. In the remainder of this section we formalize this connection between graph structure and emergent beliefs. In particular, we explore how the centrality of individual nodes on the communication graph \mathcal{G}_a and its graph symmetries inform the opinion patterns the network settles on.

Undirected graphs and agent centrality

By Theorem 4.1.1, the pitchfork bifurcation that gives rise to opinionated states on the network happens along a manifold tangent to $\text{span}(\mathbf{v})$, with \mathbf{v} the right eigenvector corresponding to the dominant eigenvalue of A_a . For networks of cooperative and competitive agents, we further identified in Theorem 4.2.1 eigenvectors \mathbf{v}_{min} and \mathbf{v}_{max} which govern the properties of the bifurcation. In this section, we show that for undirected networks these eigenvectors often reflect a classic measure of centrality on the network.

A *centrality measure* for a graph is an assignment of a numerical score to each node that ranks

its relative importance. In the social sciences, network centrality measures are often used to identify influential individuals in various social networks [66,118]. The *eigenvector centrality* is a popular centrality measure for unsigned networks that effectively ranks the nodes on a graph by the number of infinite walks that pass through them [27,28]. The eigenvector centrality score of node i on the graph \mathcal{G}_a is determined by entry i in the left Perron-Frobenius eigenvector \mathbf{w}_{max} of its adjacency matrix A_a , assuming \mathbf{w}_{max} is normalized to have nonnegative entries. In this section we illustrate how for all cooperative undirected graphs and for some competitive undirected graphs, the strength of commitment to options at equilibrium reflects the standard eigenvector centrality score of each agent. First, we establish this result for cooperative networks.

Proposition 4.2.2 (Agreement patterns and eigenvector centrality). *[20] Consider opinion dynamics (4.1) with undirected \mathcal{G}_a and $b_i = 0$ for all $i = 1, \dots, N_a$. When $\gamma > 0$, agreement equilibria $\mathbf{x} = (x_1, \dots, x_{N_a})$ described in Theorem 4.2.1.A satisfy $|x_i| < |x_k|$ if $(\mathbf{w}_{max})_i < (\mathbf{w}_{max})_k$ and $|x_i| = |x_k|$ if $(\mathbf{w}_{max})_i = (\mathbf{w}_{max})_k$ for all $i, k = 1, \dots, N_a$.*

A similar result holds for disagreement patterns on bipartite undirected graphs of antagonistic agents, i.e. graphs all of whose edges connect two disjoint subsets of vertices. Recall that the adjacency spectrum of bipartite graphs is symmetric about zero, which means $|\Lambda_{min}| = 1$ for all bipartite graphs. Furthermore $\lambda_{min} = -\lambda_{max}$ and $|(\mathbf{v}_{min})_i| = |(\mathbf{v}_{max})_i|$ for all $i = 1, \dots, N_a$ as long as \mathbf{v}_{min} and \mathbf{v}_{max} are normalized to the same length [190, Theorem 1.2]. This means that the disagreement bifurcation at $u = u_d$ on a bipartite undirected graph of antagonistic agents is always a pitchfork bifurcation whose properties are characterize in Theorem 4.1.1.

Proposition 4.2.3 (Bipartite disagreement patterns and eigenvector centrality). *[20] Consider (4.4) with undirected, connected, and bipartite \mathcal{G}_a , $b_i = 0$ and $\gamma < 0$. Let $\mathcal{V}_1, \mathcal{V}_2 \subset \mathcal{V}_a$ be the bipartition of the vertex set \mathcal{V}_a . Disagreement equilibria $\mathbf{x} = (x_1, \dots, x_{N_a})$ described in Theorem 4.2.1.B satisfy $|x_i| < |x_k|$ if $(\mathbf{w}_{max})_i < (\mathbf{w}_{max})_k$ and $|x_i| = |x_k|$ if $(\mathbf{w}_{max})_i = (\mathbf{w}_{max})_k$ for all $i, k = 1, \dots, N_a$. Moreover, $\text{sign}(x_i) = -\text{sign}(x_k)$ for all $i \in \mathcal{V}_1, k \in \mathcal{V}_2$.*

By Proposition 4.2.3, on undirected bipartite graphs the distribution of agents across the two options reflects the bipartition of the agents enforced by the graph architecture, and the magnitude of their beliefs relative to other agents reflects their centrality. Figure 4.2 illustrates the findings of Propositions 4.2.2 and 4.2.3. For example on the path, star, and wheel graphs the outer agents are less central than the inner agents according to the eigenvector centrality measure. This is reflected in the strength of the agents' beliefs at equilibrium, with the inner agents settling to a stronger

opinion about their chosen option in both agreement and disagreement regimes. Additionally in Figure 4.2 the path, even cycle, and star graphs are bipartite. As expected, the agents' pattern of beliefs at equilibrium reflects the underlying graph bipartition. Interestingly, the wheel graph is not bipartite; however at disagreement it assumes a similar color pattern to the star graph which has a similar geometry. The final beliefs of the central agent on the wheel graph is stronger than the beliefs of the outer agents. This observation suggests that at least for some graph structures, the magnitude of entries in the eigenvector \mathbf{w}_{min} reflects agent centrality in a similar manner to the standard eigenvector centrality vector \mathbf{w}_{max} even though the two do not coincide. As established in Theorem 4.1.1, \mathbf{w}_{min} and \mathbf{w}_{max} also rank the centrality of the agents by determining the relative influence of their biases b_i in determining the unfolding of the bifurcation diagram, as in Figure 4.1.

Notice that on the cycle graph in Figure 4.2 all agents are equally central and the magnitude of their beliefs at equilibrium is equal both in agreement and in disagreement, with disagreement opinions summing to zero over all agents. This property holds more generally on graphs that are K -regular, which we formalize in the following proposition.

Proposition 4.2.4 (Consensus and dissensus on regular graphs). *[20] If \mathcal{G}_a is undirected, connected, and K -regular, the agreement bifurcations at $u = u_a$ with $\gamma > 0$ give rise to **consensus** equilibria that satisfy $x_i = x_k$ for all $i, k \in \mathcal{V}_a$, and the disagreement bifurcations at $u = u_d$ with $\gamma < 0$ give rise to **dissensus** equilibria characterized by $\sum_{i=1}^{N_a} x_i = 0$.*

Directed graphs and agent centrality

In the previous section we discussed undirected graphs for which the adjacency matrix A_a is symmetric. As a consequence of this symmetry for undirected graphs the left and right adjacency eigenvectors coincide, $\mathbf{w}_{max/min} = \mathbf{v}_{max/min}$. The most strongly opinionated agents on undirected graphs also have the strongest influence in determining which equilibrium is selected by the network biases in the pitchfork bifurcation unfolding. This is an important design feature if we consider the biases of agents as distributed control inputs or sensor inputs that are meant to select for a desired behavior on the network. In order to most effectively influence the decision of agents following (4.1) on an undirected network, one can assign a bias to the most strongly opinionated agent.

For directed graphs, distribution of agents' beliefs decouples from their centrality with respect to the pitchfork unfolding. For example consider a purely cooperative network of agents with $\gamma > 0$. As before, the left eigenvector \mathbf{w}_{max} is the standard centrality eigenvector for their communication graph \mathcal{G}_a . On the other hand, the right eigenvector \mathbf{v}_{max} is the eigenvector centrality vector for the

graph \mathcal{G}'_a with adjacency matrix A_a^T . \mathcal{G}'_a is generated by reversing the direction of all of the arrows in \mathcal{G}_a . The distribution of beliefs on the network at an agreement equilibrium therefore reflects their centrality in the information flow network \mathcal{G}'_a , whereas their influence in determining the unfolding direction reflects their centrality in the sensing network \mathcal{G}_a . Similarly, for $\gamma < 0$ the eigenvectors \mathbf{w}_{min} and \mathbf{v}_{min} need not coincide, and the most opinionated agents may not be the most influential in determining the unfolding direction of the networked pitchfork bifurcation.

Graph symmetry

Recall that a graph automorphism for an unsigned graph $\mathcal{G}_a = (\mathcal{V}_a, \mathcal{E}_a)$ is a permutation of vertices \mathcal{V}_a that preserves adjacency between agents. The automorphism group of \mathcal{G}_a is the set of all of its automorphisms. A graph automorphism is also referred to as a *symmetry* of the graph. Symmetries of a graph generate a partition of its vertices:

Definition 4.2.5. [20] *Let Γ be the automorphism group of $\mathcal{G}_a = (\mathcal{V}_a, \mathcal{E}_a)$, and $i \in \mathcal{V}_a$ a vertex. An **orbit** of i is $O_i = \{k \in \mathcal{V}_a \mid k = \rho i \text{ for some } \rho \in \Gamma\}$. The orbits are equivalence classes that partition \mathcal{V}_a through the equivalence relation*

$$i \sim k \text{ if } k = \rho i \text{ for some } \rho \in \Gamma. \quad (4.8)$$

Connected to graph symmetry is the notion of *equivariance* of a dynamical system with respect to a symmetry group. Consider a dynamical system $\dot{\mathbf{x}} = \mathbf{h}(\mathbf{x})$, where the map $\mathbf{h} : \mathbb{R}^{N_a} \rightarrow \mathbb{R}^{N_a}$ is $\mathbf{h}(\mathbf{x}) = (h_1(\mathbf{x}), \dots, h_{N_a}(\mathbf{x}))$. Let Σ be a compact Lie group with elements σ that act on \mathbb{R}^{N_a} . Then \mathbf{h} is *σ -equivariant* for some $\sigma \in \Sigma$ if $\sigma \mathbf{h}(\mathbf{x}) = \mathbf{h}(\sigma \mathbf{x})$, and \mathbf{h} is *Σ -equivariant* if this holds true for all $\sigma \in \Sigma$ [83, Definition 1.7].² In the following propositions we establish that symmetries of the communication graph \mathcal{G}_a are symmetries of the belief formation dynamics (4.4).

Proposition 4.2.6 (Γ_a -equivariance). [20] *Consider (4.4) with $b_i = 0$ for $i = 1, \dots, N_a$. Let Γ_a be the automorphism group of \mathcal{G}_a . The dynamical system (4.4) is Γ_a -equivariant.*

Proposition 4.2.6 is a consequence of generic properties of dynamical systems with graph structure, see [8, 191].

Proposition 4.2.7 (Symmetry and equilibrium patterns). [20] *Consider opinion dynamics (10.1) with $b_i = 0$ for all $i \in \mathcal{V}_a$. Let Γ_a be the automorphism group of the undirected graph \mathcal{G}_a , and for any two vertices $i, k \in \mathcal{V}_a$ define the equivalence relation $i \sim k$ as in (4.8).*

²This paragraph is taken verbatim from [20]

A. Suppose $\gamma > 0$. For the agreement equilibria $\mathbf{x} = (x_1, \dots, x_{N_a})$ from Theorem 4.2.1.A, if $i \sim k$, then $x_i = x_k$.

B. Suppose $\gamma < 0$ and $|\Lambda_{min}| = 1$. For the disagreement equilibria $\mathbf{x} = (x_1, \dots, x_{N_a})$ from Theorem 4.2.1.B, if $i \sim k$, then $|x_i| = |x_k|$.

Using Proposition 4.2.7 we gain further insight into the agreement and disagreement patterns shown in Figure 4.2. For example for the star, wheel, and cycle graphs the outer agents are all within the same equivalence class of Definition 4.2.5 generated by permutation symmetries. This equivalence is reflected in the belief pattern at equilibrium, as the outer agents all form beliefs of equal magnitude. Similarly, the line graph has a reflection symmetry about its midpoint. This means $i \sim k$ whenever agent i and agent k are the same number of nodes away from one of the two endpoints of the line graph. Such agents form beliefs of equal magnitude in agreement and in disagreement.

4.3 Mixed-sign networks

In the previous section we described patterns of agreement and disagreement opinions on networks of purely cooperative or purely antagonistic agents, and we showed that cooperation among agents always leads to network agreement. On sign-homogeneous networks, cooperation between agents is therefore both necessary and sufficient for agreement. However more generally, we consider the model (4.4) on a communication graph with a mix of cooperative and antagonistic relationships among neighbors. It turns out that an all-positive signature of edges in \mathcal{G}_a is sufficient but not necessary for network agreement, and a negative signature on one or more edges of \mathcal{G}_a is necessary but not sufficient for disagreement. We illustrate this in the following proposition, in which we provide a sufficient condition for network agreement that allows for some negative interactions between agents. Recall that a matrix $A \in \mathbb{R}^{N \times N}$ is called eventually positive if there exists some integer $k_0 \geq 0$ such that $A^k \succ 0_{N \times N}$ for all $k \geq k_0$.

Proposition 4.3.1 (Agreement on mixed-sign graphs). *Consider (4.1) with $\gamma > 0$ and $b_i = 0$ for $i = 1, \dots, N_a$ on a signed directed communication graph \mathcal{G}_a . Suppose A_a is eventually positive. Then A_a has a unique dominant eigenvalue λ^* with corresponding eigenvector \mathbf{v} . The neutral state $\mathbf{x} = \mathbf{0}$ is a locally exponentially stable equilibrium for $0 < u < u^*(\lambda^*)$ where $u^*(\lambda^*)$ is defined as in (4.4). At $u = u^*(\lambda^*)$ the origin loses stability and two new equilibria appear in a pitchfork bifurcation along a manifold tangent to $\text{span}(\mathbf{v})$, where $\mathbf{v} \succ \mathbf{0}$. If A_a is eventually positive, $\mathbf{v} \succ \mathbf{0}$ and at each*

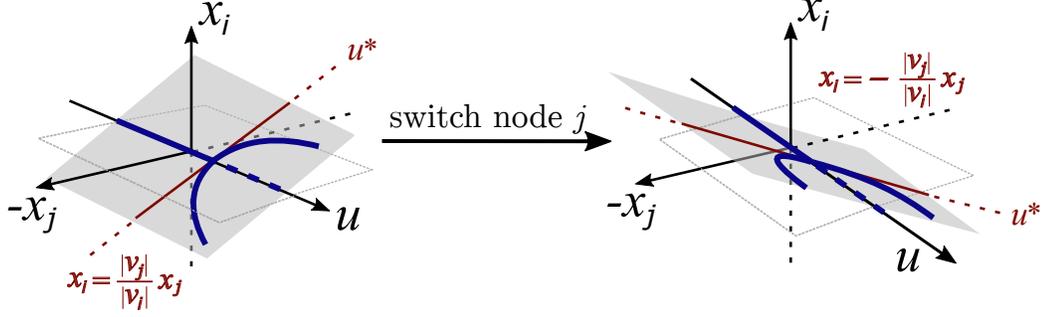


Figure 4.3: [18] Illustration of Corollary 4.3.3. The bifurcation diagram of the switched system is a “rotated” version of the original diagram because the sign of v_j flips.

nontrivial equilibrium $\mathbf{x} = (x_1, \dots, x_{N_a})$, $x_i \neq 0$ and $\text{sign}(x_i) = \text{sign}(x_k)$ for all $i, k \in \mathcal{V}_a$.

Proof. A square matrix A is eventually positive if and only if A and A^T possess the strong Perron-Frobenius property - see Proposition 2.1.1. This establishes the existence of a simple dominant eigenvalue $\lambda^* = \rho(A_a)$ and a corresponding positive eigenvector \mathbf{v} . The rest of the proposition is a direct consequence of Theorem 4.1.1. \square

Recall that two signed graphs $\mathcal{G} = (\mathcal{V}, \mathcal{E}, s)$, $\mathcal{G}^{\mathcal{W}} = (\mathcal{V}, \mathcal{E}, s_{\mathcal{W}})$ are switching equivalent if there exists a switching function $\theta : \mathcal{V} \rightarrow \{-1, 1\}$ that relates the signatures of the two graphs through $s_{\mathcal{W}}(e_{ik}) = \theta(i)s(e_{ik})\theta(k)$ for all $i, k \in \mathcal{V}$, $e_{ik} \in \mathcal{E}$. In this notation, $\mathcal{W} \subseteq \mathcal{V}$ is the set of nodes that is being switched, i.e. $\theta(i) = -1$ if $i \in \mathcal{W}$ and $\theta(i) = 1$ otherwise. The switching matrix $\Theta = \text{diag}(\theta(1), \dots, \theta(N))$ relates the adjacency matrices of the two switching equivalent graphs as $A^{\mathcal{W}} = \Theta^{-1}A\Theta$. In the following proposition we establish topological equivalence of the flow generated by (4.1) on two switching equivalent graphs.

Proposition 4.3.2 (Diffeomorphism between trajectories of switching equivalent systems). [18] Consider switching equivalent graphs $\mathcal{G}_a, \mathcal{G}_a^{\mathcal{W}}$ with adjacency matrices A_a and $A_a^{\mathcal{W}}$ and with switching matrix Θ . The trajectory $\mathbf{x}(t)$ is a solution to (4.4) on \mathcal{G} if and only if $\Theta\mathbf{x}(t)$ is a solution of (4.4) on $\mathcal{G}^{\mathcal{W}}$.

Equivalence of bifurcation diagrams of (4.1) on two switching equivalent graphs is a direct consequence of Proposition 4.3.2.

Corollary 4.3.3 (Switching a graph “rotates” a bifurcation diagram). [18] Consider (4.4) on switching equivalent graphs $\mathcal{G}_a, \mathcal{G}_a^{\mathcal{W}}$ with adjacency matrices A_a and $A_a^{\mathcal{W}}$ and with switching matrix Θ . The state $\mathbf{x} = \mathbf{x}^*$ is an equilibrium on the bifurcation diagram of (4.4) on \mathcal{G}_a at some value of u if and only if $\Theta\mathbf{x}^*$ is an equilibrium on the bifurcation diagram of (4.4) on $\mathcal{G}^{\mathcal{W}}$ for the same u .

By Corollary 4.3.3 whenever (4.1) exhibits a pitchfork bifurcation on some graph \mathcal{G} , the bifurcation diagram of (4.1) on a switching equivalent graph $\mathcal{G}^{\mathcal{W}}$ also exhibits a pitchfork bifurcation at the same critical attention value u^* . The mapping from \mathcal{G} to $\mathcal{G}^{\mathcal{W}}$ through a switching transformation effectively “rotates” the pitchfork bifurcation, mapping its nontrivial branches into different orthants in \mathbb{R}^{N_a} . We illustrate this intuition in Figure 4.3.

Next, we establish that switching a set of vertices is equivalent to switching its complement.

Proposition 4.3.4 (Switching complementary vertex sets generates the same flow). *[18] Consider two switching equivalent graphs $\mathcal{G}^{\mathcal{W}}$, $\mathcal{G}^{\mathcal{V}\setminus\mathcal{W}}$, generated by switching a set of vertices \mathcal{W} or its complement $\mathcal{V}\setminus\mathcal{W}$ on graph \mathcal{G} . The trajectory $\mathbf{x}(t)$ is a solution of (10.1) on $\mathcal{G}^{\mathcal{W}}$ if and only if it is also a solution of (10.1) on $\mathcal{G}^{\mathcal{V}\setminus\mathcal{W}}$.*

These results motivate a simple design procedure to build a signed adjacency matrix that ensures a desired allocation of agents across the two options, i.e. partitions the agents into two subgroups of opposite-sign opinion.

Step 1. Start with a strongly connected \mathcal{G}_a with an all-positive signature, i.e. $(A_a)_{ik} \in \{0, 1\}$ for all $i, k \in \mathcal{V}$. By Theorem 4.2.1, (4.1) on \mathcal{G}_a has an all-positive stable equilibrium \mathbf{x}_1^* and all-negative stable equilibrium \mathbf{x}_2^* .

Step 2. Define the switching set \mathcal{W} . In this step, the designer chooses which nodes are grouped together. The two partitions $\mathcal{W}, \mathcal{V}\setminus\mathcal{W}$ correspond to the two options or tasks.

Step 3. Update edge signatures of \mathcal{G}_a locally as $a_{ik}^{\mathcal{W}} = \theta(i)(A_a)_{ik}\theta(k)$. This edge signature update generates the switch-equivalent graph $\mathcal{G}_a^{\mathcal{W}}$ and groups all nodes in \mathcal{W} and all nodes in $\mathcal{V}\setminus\mathcal{W}$ together by sign. The dynamics (10.1) on $\mathcal{G}^{\mathcal{W}}$ is bistable with stable equilibria $\Theta\mathbf{x}_1^*$, $\Theta\mathbf{x}_2^*$. If $|\mathcal{W}| = M$, the equilibrium $\Theta\mathbf{x}_1^*$ corresponds to M negative nodes, and $\Theta\mathbf{x}_2^*$ to $N - M$ negative nodes.³

In the simulation shown in Figure 4.4 we illustrate this design procedure in action. We design a signed adjacency matrix for 10 agents which partitions the agents into a group of 3 and a group of 7 by sign. To generate the signed graph we first generate a random positive-signature graph for a network of ten agents. We then apply a switching transformation for nodes 1,2, and 3. Note that as a consequence of Proposition 4.3.4, we could equivalently choose to switch nodes 3-10 to generate the same bipartition of agents on the network. The figure shows the time trajectories of the networked opinions converging to one of the two bistable equilibria for a value of attention u slightly above the bifurcation point, as well as the final distribution of opinions on the network at the final time step of

³Paragraphs outlining steps 1-3 are taken mostly verbatim from [18]

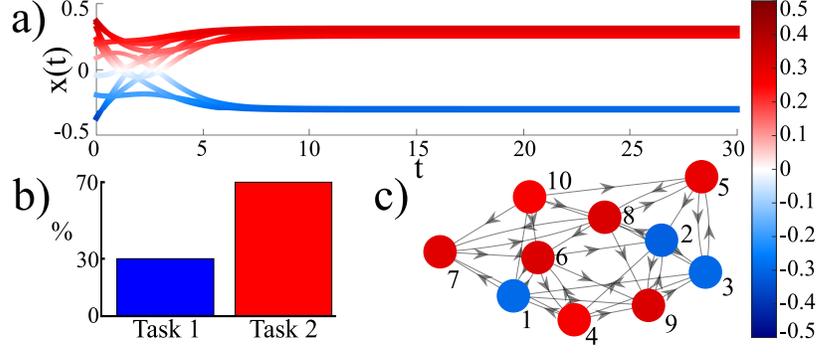


Figure 4.4: [18] Assigning 10 agents to a 30-70% distribution between two tasks by switching agents 1, 2 and 3. Edges that connect red nodes and blue nodes have negative signature, all other edges have positive signature. (a) Time trajectory of the opinion dynamics (4.1). (b) Final agent distribution. (c) Network diagram with the opinion of each agent at $t = 30$. Parameters: $S = \tanh$, $d = 1$, $\alpha = 1.2$, $\gamma = 1.3$, $u = 0.324$.

the simulation. From a different set of initial conditions, the group could settle to the second stable equilibrium at which nodes 1,2,and 3 would take on positively opinions and the rest of the group would take on negative opinions.

More generally, any pattern of disagreement can be generated using the outlined design procedure starting from a signed graph \mathcal{G}_a with an eventually positive adjacency matrix A_a . This follows straightforwardly from Proposition 4.3.1, since all possible disagreement patterns can be generated by applying switching transformation to a graph that generates agreement solutions. This means that there can be several different graph signatures for the same graph architecture $(\mathcal{V}_a, \mathcal{E}_a)$ that generate the same partition of nodes across the two options.

In the simulations we have seen so far, all agents' opinions tend away from zero post-bifurcation. This does not have to be the case in general. There are many graphs on which (4.1) exhibits a pitchfork bifurcation for which some of the agents' beliefs remain close to neutral. This happens whenever the right leading eigenvector \mathbf{v} of the adjacency matrix has some zero components. To formalize this observation, provided a graph \mathcal{G}_a with a simple leading eigenvalue and a corresponding right eigenvector \mathbf{v} , we will say agent i forms *strong beliefs* in a pitchfork bifurcation whenever $v_i \neq 0$, i.e. whenever the center manifold approximation has a linear dependence in the direction of x_i .

Proposition 4.3.5 (Strong beliefs). *Consider (4.1) with $b_i = 0$ for all $i = 1, \dots, N_a$ on a graph \mathcal{G}_a with signed adjacency matrix A_a . If A_a is switching equivalent to an eventually positive matrix, the indecision state $\mathbf{x} = \mathbf{0}$ loses stability in a supercritical pitchfork bifurcation along a center manifold on which all agents form strong beliefs.*

Proof. Suppose A_a has a simple leading eigenvalue λ^* with a right eigenvector $\mathbf{v} = (v_1, \dots, v_{N_a})$

for which $v_i \neq 0$ for all $i \in \mathcal{V}_a$, and a left eigenvector \mathbf{w} . Let $\mathcal{W} \subseteq \mathcal{V}$ be the set of nodes for which $v_i < 0$. Then we can define a switching function $\theta(i) = -1$ whenever $i \in \mathcal{W}$ and $\theta(i) = 1$ otherwise, with switching matrix $\Theta = \text{diag}(\theta(1), \dots, \theta(N_a))$. The switching equivalent adjacency matrix $A'_a = \Theta A_a \Theta$ has the strong Perron-Frobenius property by Proposition 2.2.1. Recall that A'_a has the strong Perron-Frobenius property if and only if it is eventually positive [149, Theorem 2.2]. The supercritical pitchfork bifurcation in the dynamics is established by Theorem 4.1.1 since $\langle \mathbf{w}, \mathbf{v}^3 \rangle = \sum_{i=1}^{N_a} w_i v_i^3 > 0$. \square

Proposition 4.3.5 establishes that in order for all agents to form strong beliefs at bifurcation, it is sufficient for the signed adjacency matrix A_a of the communication graph to be in a switching equivalence class with an eventually positive matrix. A special class of graphs whose adjacency matrices are always eventually positive are strongly connected graphs with an all-positive signature. Recall from Chapter 2 that a signed graph for which the sign of every closed path on the graph is positive is called structurally balanced. A structurally balanced signed graph is switching equivalent to an all-positive graph.

Let \mathcal{K} be an orthant of \mathbb{R}^N , $\mathcal{K} = \{\mathbf{x} \in \mathbb{R}^N \text{ s.t. } (-1)^{m_i} x_i \geq 0, i = 1, \dots, N\}$ with each $m_i \in \{0, 1\}$. The orthant \mathcal{K} generates a partial ordering “ $\leq_{\mathcal{K}}$ ” on \mathbb{R}^N where if $\mathbf{x}, \mathbf{y} \in \mathbb{R}^N$, $\mathbf{y} \leq_{\mathcal{K}} \mathbf{x}$ if and only if $\mathbf{x} - \mathbf{y} \in \mathcal{K}$. We say a system $\dot{\mathbf{x}} = f(\mathbf{x})$ on $\mathcal{U} \subseteq \mathbb{R}^N$ is *type \mathcal{K} monotone* if its flow preserves the partial ordering $\leq_{\mathcal{K}}$, i.e. if $\mathbf{x}_1(0) \leq_{\mathcal{K}} \mathbf{x}_2(0)$ implies $\mathbf{x}_1(t) \leq_{\mathcal{K}} \mathbf{x}_2(t)$ for all $t > 0$.⁴

Proposition 4.3.6 (Monotonicity and structural balance). *[18] Consider (4.1) on a signed graph \mathcal{G}_a with $b_i = 0$ for all $i = 1, \dots, N_a$. It is a type \mathcal{K} monotone system if and only if \mathcal{G}_a is switching equivalent to \mathcal{G}_a^+ , for which $\sigma(e_{ik}) = 1$ for all $e_{ik} \in \mathcal{E}$, i.e. \mathcal{G}_a is structurally balanced.*

Proposition 4.3.6 gives a necessary and sufficient condition for monotonicity of the dynamics (4.1). This is useful because monotone systems are well-behaved. For example, a locally stable Hopf bifurcation cannot happen in a monotone system, which rules out the possibility of networked beliefs settling on an oscillation. Furthermore, generically an equilibrium in a monotone system is either asymptotically stable, or there are two heteroclinic orbits connecting it to two other asymptotically stable equilibria [187]. This means that belief formation according to (4.1) on a graph satisfying proposition 4.3.6 is generically expected to result in convergence on a network equilibrium. For a detailed treatment of monotone dynamical systems see [97].

⁴This paragraph is taken verbatim from [18].

4.4 Higher-order bifurcations and mode interaction

So far in this chapter we described properties of equilibria that arise from indecision-breaking pitchfork bifurcations in the belief formation networks (4.1), (4.2). The pitchfork bifurcation described by Theorem 4.1.1 requires that the largest real-part eigenvalue of the signed adjacency matrix A_a is algebraically simple. Although this is frequently the case, there are many graphs for which this property does not hold. More generally, let $\lambda^* = \max_{\lambda \in \sigma(A_a)} \operatorname{Re}(\lambda)$ and define the set of leading eigenvalues of A_a as $\Lambda_{max} = \{\lambda \in \sigma(A_a) \text{ s.t. } \operatorname{Re}(\lambda) = \lambda^*\}$. Whenever $\gamma \neq \delta$ and $|\Lambda_{max}| = n$, there are exactly n eigenvalues of the Jacobian (4.3) that lie on the imaginary axis at the bifurcation point $u^*(\lambda^*)$ (4.5). This means the center manifold along which at $u^*(\lambda^*)$ is $(n + 1)$ -dimensional, and when $n > 1$ the resulting bifurcation is different from the pitchfork bifurcation of Theorem 4.1.1. Some of the possibilities of higher-order phenomena for (4.2) include the following:

- $\gamma \neq \delta$ and the leading eigenvalues are a complex-conjugate pair $\Lambda_{max} = \{\lambda, \bar{\lambda}\}$. In this case a Hopf bifurcation of periodic orbits is observed. This type of bifurcation occurs frequently on graphs which do not satisfy the conditions for forming strong beliefs in Proposition 4.3.5. We defer the analysis of this case to the following chapter, where we will examine oscillations in the more general multi-option model.
- $\gamma \neq \delta$ and $|\Lambda_{max}| = n > 1$, where the set Λ_{max} consists of copies of a real-valued eigenvalue with multiplicity n . Several branches of new equilibria will emerge in a higher-order bifurcation of equilibria on the network. Typically this happens when the graph \mathcal{G}_a has some nontrivial symmetry, see Figure 4.5 A and B for examples of equilibria generated by a graph with a leading eigenvalue of multiplicity 2. In fact for every undirected cycle graph with an all-negative signature and an odd N_a , the leading eigenvalue is $\lambda^* = 2 \cos\left(\frac{N_a-1}{N_a}\right)$ with multiplicity 2 and two eigenvalues cross zero simultaneously at the indecision-breaking bifurcation. Techniques from equivariant bifurcation theory and network bifurcation theory can be used to classify these higher-order phenomena for classes of networks with specific symmetries [81, 83, 84]. Such analysis lies outside of the scope of this dissertation.
- $\gamma = \delta$ and $\alpha > \beta$. In this case, the influence of the adjacency matrix in the Jacobian matrix (4.3) disappears, and all N_a eigenvalues of the Jacobian are simultaneously zero when $u = d/(\alpha - \beta)$.

For the remainder of this section we will focus our attention on the last case, which is particularly interesting due to its interpretation. Recall that in the model for opinion formation on two mutually

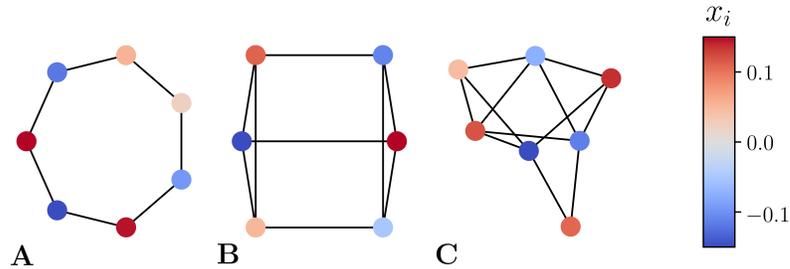


Figure 4.5: [20] Disagreement patterns of (4.1) on odd cycle (A), 3-regular (B) and randomly generated (C) graphs of purely antagonistic agents. All pictured undirected edges have a negative signature. The color of the pictured nodes represents the opinion x_i the node converged to at $t = 500$ of a simulation. All nodes have $b_i = 0$ and start from randomized initial conditions drawn from a uniform distribution between -1 and 1. System parameters are $d = 1$, $\alpha = 0.5$, $\gamma = 0.5$, $u = u_d + 0.01$.

exclusive options (4.2) the parameter γ reflects the strength of agents' drive for social imitation or reciprocity, and δ reflects the strength of alignment between the two options. When $\gamma > 0$, each agent is pulled to form an opinion choosing the option favored by its social network. When $\delta > 0$ agents are simultaneously pulled to choose the opposite option, since the two are positively correlated. When $\gamma = \delta$, these two conflicting social forces balance each other out. Effectively this situation means that agents desire both options, but can only have one. Due to the simplex constraint imposed by this mutual exclusivity, the Jacobian matrix (4.3) becomes maximally singular at the bifurcation point, independently of the architecture of the communication graph. When $\gamma = \delta$ we will refer to the agents as *conflicted*.

As a consequence of the high multiplicity of the crossing eigenvalue, indecision-breaking bifurcations in a network of conflicted agents can give rise to an exceptionally rich set of possible network equilibria. However for the same reason, this parameter regime is tricky to analyze. The task of characterizing bifurcations in arbitrarily large general networks of conflicted agents is potentially intractable. When multiple eigenvalues cross the imaginary axis simultaneously, their associated eigenspaces can have nonlinear coupling that gives rise to behaviors that are more complicated than would be expected if each eigenvalue was crossing the imaginary axis separately. This phenomenon is typically referred to as *mode interaction* [31, 84], and the complexity of the structure of resulting bifurcation branches is expected to grow significantly with the number of crossing eigenvalues. Some exploratory simulations of opinion formation two options on small networks of conflicted agents can be found in the undergraduate senior thesis [133], and indeed there are many outcomes the network can settle on as a single parameter of the model (4.1) is varied. To rigorously illustrate this complexity we will examine in detail a deceptively simple system of two conflicted agents.

Consider the dynamics (4.2) for two communicating unbiased agents and assume $(A_a)_{12} = (A_a)_{21}$.

Assume that the agents are conflicted, with $\gamma = \delta$, and define the parameter $c =: \gamma(A_a)_{12} = \delta(A_a)_{12} = \gamma(A_a)_{21} = \delta(A_a)_{21}$ with $c \in \mathbb{R}$. When $c > 0$ the two agents are conflicted because they are pulled to cooperate and choose both options, whereas when $c < 0$ the two agents are conflicted because they are pulled to reject their neighbor's opinion on both options. With these simplifications, the equations governing the opinion formation for the two agents become

$$\dot{x}_i = -dx_i + u \left(\hat{S}_1(\alpha x_i + cx_j) - \hat{S}_2(\beta x_i + cx_j) \right) \quad i, j \in \{1, 2\}, \quad i \neq j. \quad (4.9)$$

To simplify the calculations we choose specific saturating functions, the odd sigmoid functions $\hat{S}_1(x) = \frac{1}{a_1} \tanh(a_1 x)$ and $\hat{S}_2(x) = \frac{1}{a_2} \tanh(a_2 x)$ where $a_1, a_2 > 0$ and $a_2 > a_1$.

Proposition 4.4.1. *Consider (4.9) with $\alpha > \beta$ and define the quantities*

$$\begin{aligned} k &=: a_1^2(c + \alpha)^3 - a_2^2(c + \beta)^3, & \ell &=: -a_1^2(c - \alpha)^3 + a_2^2(c - \beta)^3, \\ m &=: \frac{3(a_1^2(c - \alpha)^2(c + \alpha) - a_2^2(c - \beta)^2(c + \beta))}{|-a_1^2(c - \alpha)^3 + a_2^2(c - \beta)^3|}, & n &=: \frac{3(-a_1^2(c + \alpha)^2(c - \alpha) + a_2^2(c + \beta)^2(c - \beta))}{|a_1^2(c + \alpha)^3 - a_2^2(c + \beta)^3|} \end{aligned} \quad (4.10)$$

The neutral state $\mathbf{x} = \mathbf{0}$ is locally exponentially stable for $0 \leq u < d/(\alpha - \beta)$ and unstable for $u > d/(\alpha - \beta)$. Assuming the nondegeneracy conditions

$$k \neq 0, \quad \ell \neq 0, \quad m \neq \text{sign}(\ell), \quad n \neq \text{sign}(k), \quad mn \neq \text{sign}(k) \text{sign}(\ell) \quad (4.11)$$

are satisfied for some choice of parameters $a_1, a_2, \alpha, \beta, c$, the bifurcation diagram of (4.9) is equivalent⁵ to a bifurcation diagram of a normal form for $(Z_2 \oplus Z_2)$ -symmetric nondegenerate bifurcation problems ,

$$\begin{aligned} \dot{x}_c &= -rx_c - \varepsilon_c x_c^3 - mx_c x_d^2, \\ \dot{x}_d &= -rx_d - \varepsilon_d x_d^3 - nx_c^2 x_d. \end{aligned} \quad (4.12)$$

where $r = u - d/(\alpha - \beta)$, $\varepsilon_c = \text{sign}(k)$, $\varepsilon_d = \text{sign}(\ell)$ and the parameters m, n are defined in (4.10).

A proof of Proposition 4.4.1 is developed Appendix A. The equations (4.12) define a family of normal forms parametrized by $\varepsilon_c, \varepsilon_d \in \{1, -1\}$ and $m, n \in \mathbb{R}$. The parameters m, n are referred to as *modal parameters*. The bifurcations in these normal forms are classified using tools from singularity theory in [82, Chapter X]. From this analysis we can deduce that there are several topologically distinct bifurcation diagrams that can be obtained for different choices of model

⁵Here we mean equivalent in the sense of Definition 2.3.2, with additional constraint that the $Z_2 \oplus Z_2$ symmetry is preserved.

parameters $\alpha, \beta, c, a_1, a_2$. There are two types of solution branches that can appear in this bifurcation. The first of these are the *pure mode* branches, that correspond to consensus between the agents $x_1 = x_2$ and dissensus between the agents $x_1 = -x_2$. Consensus and dissensus equilibria appear at the bifurcation for all choices of model parameters. Additional equilibria can sometimes appear in *mixed mode* branches that are a consequence of mode interaction. We will now describe in more detail the various possible bifurcation diagrams realized by the system (4.9), identifying each distinct bifurcation diagram by its parameter regime in the normal form (4.12). In the normal form, the coefficient ε_c determines whether consensus equilibria bifurcate supercritically ($\varepsilon_c > 0$) or subcritically ($\varepsilon_c < 0$). Similarly, $\varepsilon_d > 0 (< 0)$ indicates whether dissensus equilibria bifurcate supercritically or subcritically.

Case 1: $\varepsilon_c = \varepsilon_d = 1$. In this case, the consensus and dissensus solution branches both bifurcate supercritically, for values of $u > d/(\alpha - \beta)$. This happens whenever the parameter c satisfies

$$-\left| \frac{a_2^{2/3}\beta - a_1^{2/3}\alpha}{a_2^{2/3} - a_1^{2/3}} \right| < c < \left| \frac{a_2^{2/3}\beta - a_1^{2/3}\alpha}{a_2^{2/3} - a_1^{2/3}} \right|. \quad (4.13)$$

When this is satisfied, the system (4.9) realizes one of five topologically distinct bifurcation diagrams.

- I. $m > 1, n > 1$. In this parameter regime, consensus and dissensus solution branches are simultaneously stable at bifurcation. In addition to these pure mode solutions, four unstable mixed mode equilibria bifurcate from the origin supercritically.
- II. $m < 1, n > 1$. In this parameter regime, the consensus solutions are stable and the dissensus solutions are unstable. No mixed-mode branches bifurcate from the origin.
- III. $m < 1, n < 1, mn < 1$. In this parameter regime, the dissensus solutions are stable and the consensus solutions are unstable. No mixed-mode branches bifurcate from the origin.
- IV. $m > 1, n < 1$. In this parameter regime, consensus and dissensus solution branches are simultaneously unstable at bifurcation. In addition to these pure mode solutions, four stable mixed mode equilibria bifurcate from the origin supercritically. Depending on the parameters of the model (4.9), these mixed mode solutions can represent states of agreement, in which the two agents commit to the same option with different magnitudes of opinion, or mixed disagreement, in which the two agents commit to the opposite options with different magnitudes of opinion.

V. $m < 1$, $n < 1$, $mn > 1$. This parameter regime appears in the normal form, however we have not yet found a parameter combination in the model (4.9) that generates it. In this parameter regime, consensus and dissensus branches would both appear supercritically, and four mixed mode solutions would bifurcate subcritically. All of these solution branches would be unstable at bifurcation.

For an illustration of a representative bifurcation diagram of (4.9) in each of the first four described parameter regimes, see Figure 4.6. In this figure, dashed lines represent branches of unstable equilibria and solid lines represent stable branches of equilibria. Pink curves are the indecision equilibrium $\mathbf{x} = \mathbf{0}$, blue curves are the consensus equilibria, black curves are the dissensus equilibria, and red curves are the mixed mode branches. Additionally in Figure 4.7 we show a representative phase portrait in each of these four parameter regimes, for a value of u above the bifurcation point $d/(\alpha - \beta)$. The color of each of the equilibria in this figure is the same as the color of the solution branch they belong to in the bifurcation diagrams in Figure 4.6.

Case 2: $\varepsilon_c \neq \varepsilon_d$. In this case one of the pure mode branches bifurcates supercritically, and the other one subcritically.

(a) $\varepsilon_c = -1$, $\varepsilon_d = 1$. This parameter regime is realized whenever

$$c > \left| \frac{a_2^{2/3} \beta - a_1^{2/3} \alpha}{a_2^{2/3} - a_1^{2/3}} \right|. \quad (4.14)$$

In this parameter regime the consensus bifurcation happens subcritically and the dissensus bifurcation happens supercritically.

(b) $\varepsilon_c = 1$, $\varepsilon_d = -1$. This parameter regime is realized whenever

$$c < - \left| \frac{a_2^{2/3} \beta - a_1^{2/3} \alpha}{a_2^{2/3} - a_1^{2/3}} \right|. \quad (4.15)$$

In this parameter regime the dissensus bifurcation happens subcritically and the consensus bifurcation happens supercritically.

For each of these two scenarios, there are five possible topologically distinct bifurcation diagrams, separated in the $m - n$ parameter plane by the nondegeneracy conditions $m \neq -1$, $n \neq 1$, $mn \neq -1$. The descriptions of these bifurcation diagrams are analogous to descriptions of diagrams I-V provided for Case 1, excepting the criticality of the consensus or dissensus

solution branch. As before, in some of the parameter regimes mixed mode branches appear alongside the consensus and dissensus equilibria. To stay concise we do not enumerate the possibilities in this case and refer the reader to the analysis in [82, Chapter X]. We show a representative bifurcation diagram for each of the two categories in Figure 4.9.

For a concrete example, consider (4.9) with parameters $d = 1$, $\alpha = 0$, $\beta = -1$, $a_1 = 1$, and $a_2 = 2$. For these choices of parameters, the indecision-breaking bifurcation happens at $u = 1$. Varying a single parameter c generates six distinct bifurcation diagrams. In the following discussion we provide descriptions of the six parameter regimes; all of the numerical values have been rounded to four decimal points.

Case 1. When $-2.7024 < c < 2.7024$, $\varepsilon_c = \varepsilon_d = 1$ and all bifurcations happen supercritically.

- I. $-0.5149 < c < 0.5149$. Consensus and dissensus solution branches are simultaneously stable in this regime. When $c > 0$, the agents' opinions at consensus are stronger than the opinions at dissensus; when $c < 0$ dissensus opinions are stronger than consensus opinions.
- II. $0.5149 < c < 0.7224$ and $1.7952 < c < 2.7024$. Consensus equilibria are stable and dissensus equilibria are unstable in this parameter regime.
- III. $-1.6519 < c < -0.7224$ and $0.7224 < c < 1.6519$. Consensus and dissensus equilibria are both unstable in this parameter regime. Instead, four mixed mode equilibria are stable. When $c > 0$ the mixed mode equilibria represent agreement with $\text{sign}(x_1) = \text{sign}(x_2)$, with two agreement equilibria available for each option. Similarly, when $c < 0$ the mixed mode equilibria represent agreement with $\text{sign}(x_1) = -\text{sign}(x_2)$.
- IV. $-2.7024 < c < -1.7952$ and $-0.7224 < c < -0.5149$. Dissensus equilibria are stable and consensus equilibria are unstable in this parameter regime.

See Figure 4.6 for four bifurcation diagrams of (4.9) in each of these parameter regimes and Figure 4.7 for phase portraits at a supercritical value $u = 1.4$. In Figure 4.8 we show a graphical representation of a curve in the m, n parameter plane that is traced out as c is varied between -2.7024 and 2.7024 .

Case 2. (a) $c > 2.7024$: dissensus solutions are unstable and bifurcate supercritically; consensus solutions bifurcate subcritically and are also unstable at bifurcation. However the consensus solution branch undergoes a saddle-node bifurcation at some value of $u = u^* < 1$ and gains stability - see Figure 4.9 (a).

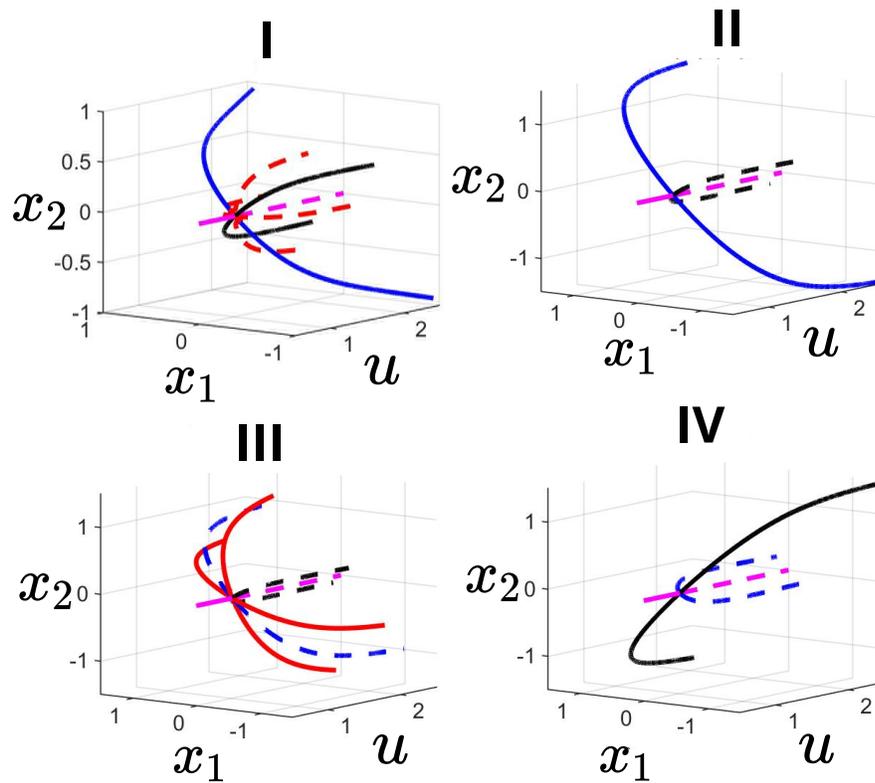


Figure 4.6: Representative bifurcation diagrams of (4.9) with coupling parameter c in each of the four distinct regions of the plane of modal parameters of normal form (4.12) pictured in Figure 4.8. Model parameters: $d = 1$, $\alpha = 0$, $\beta = -1$, $u = 1.4$, $a_1 = 1$, $a_2 = 2$. I) $c = 0.2$; II) $c = 0.6$; III) $c = 1$; IV) $c = -0.6$.

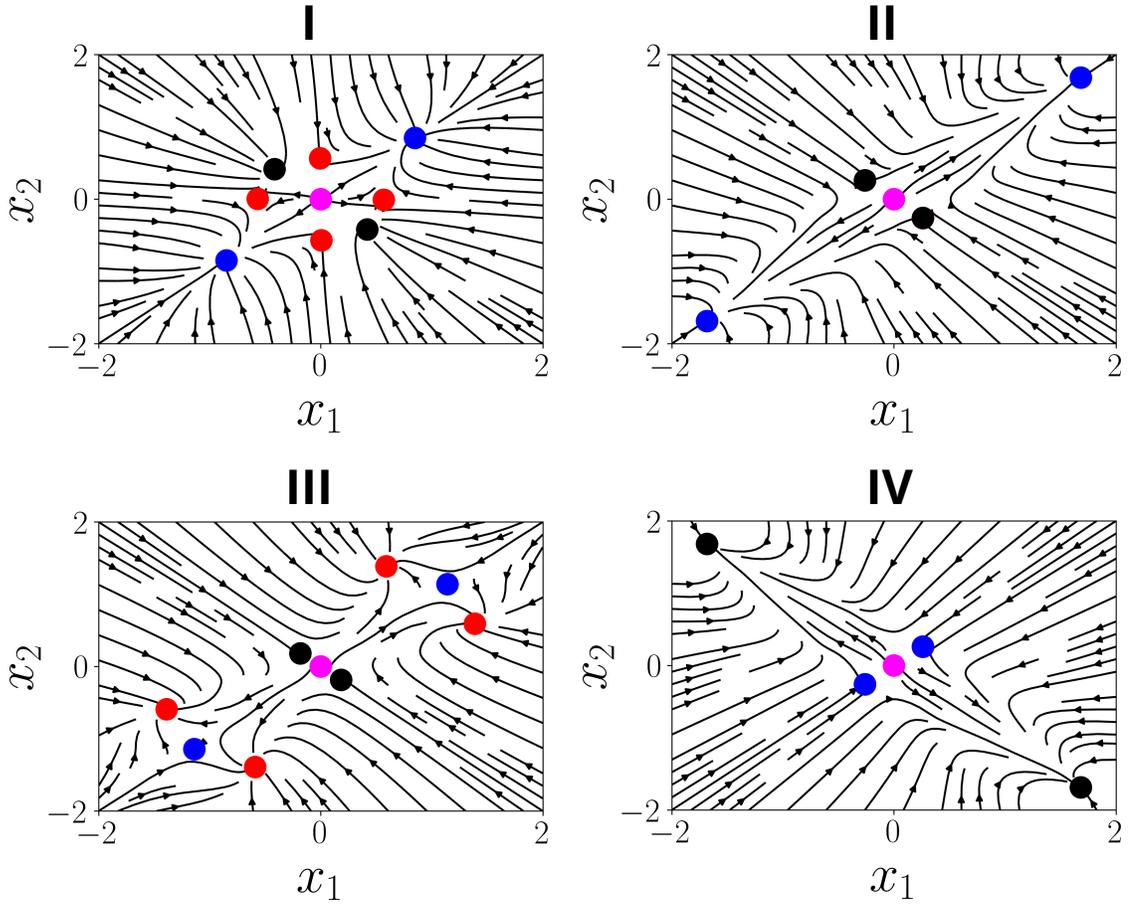


Figure 4.7: Representative phase portraits of (4.9) with coupling parameter c in each of the four distinct regions of the plane of modal parameters of normal form (4.12) pictured in Figure 4.8. Parameters: $d = 1$, $\alpha = 0$, $\beta = -1$, $u = 1.4$, $a_1 = 1$, $a_2 = 2$. I) $c = 0.2$; II) $c = 0.6$; III) $c = 1$; IV) $c = -0.6$. Colors of equilibria correspond to colors of the branches in the bifurcation diagrams in Figure 4.6.

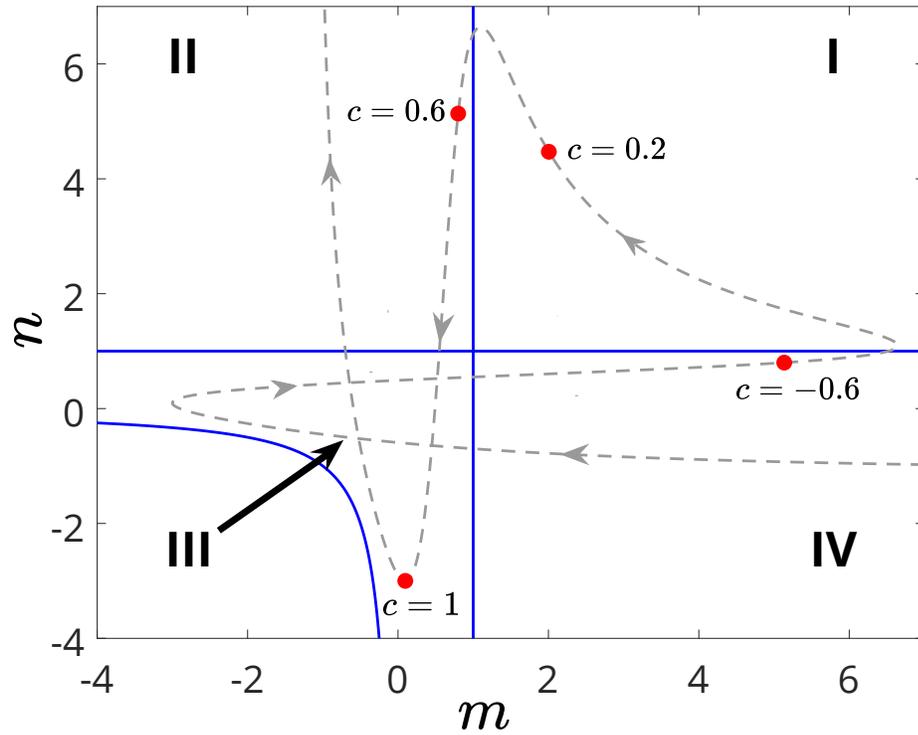


Figure 4.8: Parameter plane of $m-n$ modal parameters of the normal form (4.12), with blue lines indicating nondegeneracy conditions $m \neq 1$, $n \neq 1$, $mn \neq 1$ (the latter only shown for negative values of m, n). The four labeled regions correspond to regions I-IV discussed in the text for the normal form family with $\varepsilon_c = \varepsilon_d = 1$. Gray curve corresponds to values of m, n reached as the parameter c of (4.9) is varied between $-2.7024 < c < 2.7024$, with gray arrows indicating direction of increasing c . Other parameters: $d = 1$, $\alpha = 0$, $\beta = -1$, $u = 1.4$, $a_1 = 1$, $a_2 = 2$.

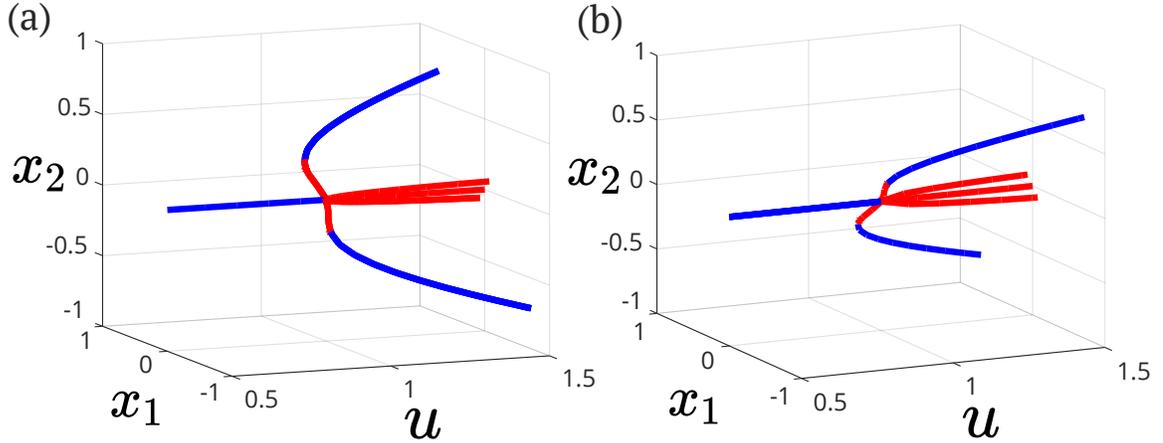


Figure 4.9: Bifurcation diagrams of (4.9) with $\varepsilon_c \neq \varepsilon_d$. Parameters: $d = 1, \alpha = 0, \beta = -1, a_1 = 1, a_2 = 2$; (a) $c = 3, \varepsilon_c = -1$, consensus solution branch bifurcates subcritically and regains stability in a saddle-node bifurcation; (b) $c = -3, \varepsilon_d = -1$, dissensus solution branch bifurcates subcritically and regains stability in a saddle-node bifurcation. Figure generated with MatCont numerical continuation package [47]. Red lines are unstable equilibria and blue lines are stable equilibria.

- (b) $c < -2.7024$: consensus solutions are unstable and bifurcate supercritically; dissensus solutions bifurcate subcritically and are also unstable at bifurcation. However the dissensus solution branch undergoes a saddle-node bifurcation at some value of $u = u^* < 1$ and gains stability - see Figure 4.9 (b).

Altogether we develop the following intuition about the opinion formation of two conflicted agents with the stated choice of system parameters. When the agents are weakly conflicted, i.e. when c is close to zero, it is possible for the agents' opinions to converge to any of the four possible combinations of choices between the two options (both choose option 1, both choose option 2, agent 1 chooses option 1 and agent 2 chooses option 2, agent 1 chooses option 2 and agent 2 chooses option 1). When $c > 0$ the agents are cooperative, and in this bistable regime consensus solutions reflect a stronger commitment to the choice of option than the dissensus state. Analogously, when $c < 0$ the agents are antagonistic and form stronger opinions when they disagree. Whether the group arrives at agreement or at disagreement depends entirely on the initial conditions, which represent initial beliefs or priors of the agents.

There exists a threshold magnitude of the c parameter beyond which the bistability between agreement and disagreement disappears. When $c > 0$ and above this threshold, the two agents' opinions will always converge to agreement. Agreement can either be consensus agreement, or a mixed agreement state in which one of the agents forms a stronger belief about the option, depending on the exact value of c . When $c < 0$ and its magnitude is above this threshold, the two agents' opinions will always converge to disagreement, which can either be a dissensus disagreement or a

mixed disagreement depending on choice of c .

Finally, there is a second threshold magnitude of the c parameter which corresponds to the switch of criticality for the consensus solution branch when c is positive or the dissensus solution branch when c is negative. Thus when the conflicted agents are strongly cooperative, it becomes possible for them to reach consensus at a lower urgency, i.e. for values of u below its indecision-breaking bifurcation point. An analogous conclusion holds for strongly antagonistic agents and dissensus.

Chapter 5

Belief-forming bifurcations: multiple options

In Chapter 4 we studied indecision-breaking bifurcations in the model (3.8) when the agents' opinions are captured by a single scalar opinion variable. Such opinion variable captures agents' beliefs on a single topic, or a choice on one of two mutually exclusive options. In this chapter we return to thinking about multiple options, with agent i 's beliefs on N_o options captured by the vector $\mathbf{Z}_i = (z_{i1}, \dots, z_{iN_o})$. In the scalar opinion setting we explored the role of the communication graph \mathcal{G}_a and discovered that the spectral properties of its signed adjacency matrix A_a play a defining role in determining how agents distribute themselves across the two options. With N_o options in addition to the communication graph \mathcal{G}_a , the belief system graph \mathcal{G}_o will play a role in determining the properties of the indecision-breaking bifurcation on the network. For easier reference, we restate the homogeneous model (3.8)

$$\dot{z}_{ij} = -dz_{ij} + u \left(S_1 \left(\alpha z_{ij} + \gamma \sum_{\substack{k=1 \\ k \neq i}}^{N_a} (A_a)_{ik} z_{kj} \right) + \sum_{\substack{l \neq j \\ l=1}}^{N_o} S_2 \left(\beta (A_o)_{jl} z_{il} + \delta (A_o)_{jl} \sum_{\substack{k=1 \\ k \neq i}}^{N_a} (A_a)_{ik} z_{kl} \right) \right) + b_{ij} := F_{ij}^h(\mathbf{Z}) \quad (5.1)$$

and its matrix-vector form (3.9)

$$\dot{\mathbf{Z}} = -d\mathbf{Z} + u\mathbf{S}_1(((\alpha\mathcal{I}_{N_a} + \gamma A_a) \otimes \mathcal{I}_{N_o})\mathbf{Z}) + \sum_{l=1}^{N_o} \mathbf{S}_2(((\beta\mathcal{I}_{N_a} + \gamma A_a) \otimes M_l)\mathbf{Z}) + \mathbf{b} \quad (5.2)$$

where $\mathbf{S}_m(\mathbf{y}) = (S_m(y_1), \dots, S_m(y_n))$ for $\mathbf{y} \in \mathbb{R}^N$, and $M_l \in \mathbb{R}^{N_o \times N_o}$ is the matrix whose column l coincides with column l of A_o , with zero entries in all other columns. For maximal generality, in this chapter we define the saturation functions S_1, S_2 as perturbations of odd saturation functions. Let $S_m(y) = \hat{S}_m(y) + g_m(y)$ with $\hat{S}_m(-y) = -\hat{S}_m(y)$, $\hat{S}'_m(0) = 1$, $g_m(0) = 0$, $g'_m(0) = 0$ for $m = 1, 2$ and suppose $g''_m(0)$ is small. When $g_1(y) = g_2(y) = 0$, S_1 and S_2 have an odd symmetry, The bounded perturbations $g_m(y) \neq 0$ capture a less idealized case with imperfect symmetry.

Recall that $\alpha, \gamma, \beta, \delta \geq 0$ are the gains that regulate the relative strengths of influence of an agent's own beliefs and the social network influence along each option dimension. We can interpret these gains in the following manner. The gain α is the strength of the agents' self-reinforcement, or their commitment to already-held beliefs; β is the strength of the agents' internal adherence to the belief system prescribed by \mathcal{G}_o . The gain γ captures the strength of social imitation, i.e. the agents' desire to mimic the beliefs of the neighbors towards whom they are cooperative and oppose the beliefs of those towards whom they are antagonistic. Finally, δ describes the agents' ideological commitment; when δ is large, the agents evaluate their neighbors' influence more holistically according to the value system \mathcal{G}_o rather than through pure imitation along each topic. A mixture of these four effects drives the formation of beliefs according to (5.1).

The results in this chapter do not yet appear in print and have not yet been submitted for publication; all of the work presented in this chapter is thereby original.

5.1 Linear analysis

As before, the neutral state $\mathbf{Z} = \mathbf{0}$ is an equilibrium of the homogeneous multi-option model (5.1) whenever the agents are unbiased, $\mathbf{b} = \mathbf{0}$. In this section we study the properties of the linearization of the system the about the neutral state. The Jacobian matrix of the linearization is

$$J(\mathbf{0}, u) = (-d + u\alpha)\mathcal{I}_{N_a} \otimes \mathcal{I}_{N_o} + u\gamma A_a \otimes \mathcal{I}_{N_o} + u\beta \mathcal{I}_{N_a} \otimes A_o + u\delta A_a \otimes A_o \quad (5.3)$$

where \otimes is the Kronecker product of matrices. The following proposition establishes that the eigenvalues and eigenvectors of (11.4) inherit their properties from the eigenvalues and eigenvectors of the signed adjacency matrices A_a and A_o of the communication and belief system graphs.

Proposition 5.1.1 (Eigenvalues and eigenvectors). *The following statements hold for (5.3), with some selection of parameters $d, u, \alpha, \gamma, \beta, \delta$:*

- 1) Recall that for a square matrix A , $\sigma(A)$ refers to the set of its eigenvalues. For each $\eta \in$

$\sigma(J(\mathbf{0}, u))$, there exists $\lambda \in \sigma(A_a)$ and $\mu \in \sigma(A_o)$ so that

$$\eta = -d + u(\alpha + \gamma\lambda + \beta\mu + \delta\lambda\mu) := \eta(u, \lambda, \mu); \quad (5.4)$$

2) Suppose λ_i is an eigenvalue of A_a with a right (left) eigenvector \mathbf{v}_i and μ_j is an eigenvalue of A_o with a right (left) eigenvector $\tilde{\mathbf{v}}_j$, then the vector $\mathbf{v}_i \otimes \tilde{\mathbf{v}}_j$ is a right (left) eigenvector of (5.3) with corresponding eigenvalue $\eta(u, \lambda_i, \mu_j)$.

Proof. 1) By Schur's unitary triangularization theorem [101, Theorem 2.3.1] there exist unitary matrices $U \in \mathbb{C}_{N_a \times N_a}$, $V \in \mathbb{C}_{N_o \times N_o}$ such that $U^* A_a U = \Delta_a$, $V^* A_o V = \Delta_o$ where Δ_a, Δ_o are upper triangular complex matrices with eigenvalues of A_a, A_o on the diagonal. Then, using the mixed-product property of the Kronecker product,

$$\begin{aligned} \Delta_J &= (U \otimes V)^* J(\mathbf{0}, u) (U \otimes V) = (-d + u\alpha)(U^* \mathcal{I}_{N_a} U) \otimes (V^* \mathcal{I}_{N_o} V) \\ &\quad + u\gamma(U^* A_a U) \otimes (V^* \mathcal{I}_{N_o} V) + u\beta(U^* \mathcal{I}_{N_a} U) \otimes (V^* A_o V) + u\delta(U^* A_a U) \otimes (V^* A_o V) \\ &= (-d + u\alpha)\mathcal{I}_{N_a} \otimes \mathcal{I}_{N_o} + u\gamma\Delta_a \otimes \mathcal{I}_{N_o} + u\beta\mathcal{I}_{N_a} \otimes \Delta_o + u\delta\Delta_a \otimes \Delta_o. \end{aligned}$$

The matrix Δ_J is upper triangular, with its diagonal entries corresponding to the eigenvalues of $J(\mathbf{0}_N, u)$. By inspection we see that all diagonal entries of Δ_J have the form $\eta(u, \lambda, \mu)$ for some $\lambda \in \sigma(A_a)$, $\mu \in \sigma(A_o)$.

2) The vector $\mathbf{v}_i \otimes \tilde{\mathbf{v}}_j$ is an eigenvector of $A_a \otimes A_o$ with eigenvalue $\lambda_i \mu_j$ by [100, Theorem 4.2.12]; it is also an eigenvector of $\mathcal{I}_{N_a} \otimes \mathcal{I}_{N_o}$, $A_a \otimes \mathcal{I}_{N_o}$, and $\mathcal{I}_{N_a} \otimes A_o$ with corresponding eigenvalues 1, λ_i, μ_j , respectively, and the proposition statement follows from multiplying (5.3) by $\mathbf{v}_i \otimes \tilde{\mathbf{v}}_j$. \square

Given some parameters $d, u, \alpha, \gamma, \beta, \delta$ we say the eigenvalue pair $\lambda \in \sigma(A_a), \mu \in \sigma(A_o)$ generate the Jacobian eigenvalue $\eta(u, \lambda, \mu)$ through the relationship (5.4). Recall that a leading eigenvalue ξ of a matrix A satisfies $\text{Re}(\xi) \geq \text{Re}(\xi_i)$ for all $\xi_i \in \sigma(A)$. For the remainder of this chapter we define $\mathcal{H} \subset \sigma(J(\mathbf{0}, u))$ to be the set of the leading eigenvalues of the Jacobian matrix (5.3) (accounting for algebraic multiplicity), and Λ to be the set of eigenvalue pairs (λ, μ) that generate each $\eta \in \mathcal{H}$. We are now ready to establish a result on the stability of the neutral indecision equilibrium $\mathbf{Z} = \mathbf{0}$.

Theorem 5.1.2 (Stability of network indecision equilibrium). *Consider (5.1) with $b_{ij} = 0$ for all $i \in \mathcal{V}_a, j \in \mathcal{V}_o$. Suppose for any $(\lambda, \mu) \in \Lambda$, $\alpha + \gamma \text{Re}(\lambda) + \beta \text{Re}(\mu) + \delta \text{Re}(\lambda\mu) > 0$ and note that this quantity is constant over Λ . Then $\mathbf{Z} = \mathbf{0}$ is locally exponentially stable whenever $u < u^*$ and*

unstable for $u > u^*$,

$$u^* = \frac{d}{\alpha + \gamma \operatorname{Re}(\lambda) + \beta \operatorname{Re}(\mu) + \delta \operatorname{Re}(\lambda\mu)}. \quad (5.5)$$

Let $\mathcal{N}(J)$ be the generalized eigenspace of $J(\mathbf{0}, u^*)$ corresponding to its leading eigenvalue set \mathcal{H} , and let $k = |\mathcal{H}|$ be the dimension of $\mathcal{N}(J)$. There exists a $(k+1)$ -dimensional invariant center manifold $W^c \subset \mathbb{R}^{N_a N_o + 1}$ passing through $(\mathbf{Z}, u) = (\mathbf{0}, u^*)$, tangent to $\mathcal{N}(J)$ at $u = u^*$. All trajectories of (5.1) starting at (\mathbf{Z}, u) near $(\mathbf{0}, u^*)$ converge to W^c exponentially as $t \rightarrow \infty$.

Proof. By Proposition 5.1.1, a leading eigenvalue of (5.3) is of the form $\eta(u, \lambda_s, \mu_p) = -d + u(\alpha + \gamma\lambda_s + \beta\mu_p + \delta\lambda_s\mu_p) =: \eta_{sp}$. When $u < u^*$, $\operatorname{Re}(\eta_{sp}) < 0$ and therefore $\operatorname{Re}(\eta_{ij}) < 0$ for all $\eta_{ij} \in \sigma(J(\mathbf{0}, u))$; for $u > u^*$, $\operatorname{Re}(\eta_{sp}) > 0$. Stability conclusions follow by Lyapunov's indirect method [109, Theorem 4.7], and the existence of an attracting center manifold follows by the Center Manifold Theorem [90, Theorem 3.2.1] since $\operatorname{Re}(\eta_{ij}) < 0$ for all eigenvalues $\eta_{ij} \in \sigma(J(\mathbf{0}, u))$, $\eta_{ij} \notin \mathcal{H}$ at $u = u^*$. \square

The instability of the origin in Theorem 5.1.2 corresponds to one or more eigenvalues of $J(\mathbf{0}, u)$ crossing the imaginary axis as u passes through its critical value u^* (5.5). By Proposition 5.1.1, the eigenvalues and eigenvectors of $J(\mathbf{0}, u)$ are generated by the eigenvalues and eigenvectors of the adjacency matrices A_a, A_o of the communication and belief system graphs $\mathcal{G}_a, \mathcal{G}_o$. In fact, for each $(\lambda_i, \mu_j) \in \Lambda$ with corresponding eigenvectors $\mathbf{v}_{ai}, \mathbf{v}_{oj}$, we know $\mathbf{v}_{ai} \otimes \mathbf{v}_{oj} \in \mathcal{N}(J)$. Since the center manifold W^c is tangent to $\mathcal{N}(J)$ at $u = u^*$, we can conclude that the communication and belief system graphs \mathcal{G}_a and \mathcal{G}_o play a defining role in determining the local structure of the new solutions that emerge along W^c . Next, we seek to understand which pairs of eigenvalues (λ, μ) generate the leading eigenvalue set \mathcal{H} . In the following proposition, we establish how the set of eigenvalue pairs Λ is selected by the communication gains γ, β, δ in the networked value dynamics (5.1).

Proposition 5.1.3 (Model parameters select the crossing eigenvalues).

Let $\lambda_{max} = \max_{\lambda_i \in \sigma(A_a)} \{\operatorname{Re}(\lambda_i)\}$, $\mu_{max} = \max_{\mu_j \in \sigma(A_o)} \{\operatorname{Re}(\mu_j)\}$.

Let $(\lambda\mu)_{max} = \max_{\lambda_k \in \sigma(A_a), \mu_l \in \sigma(A_o)} \{\operatorname{Re}(\lambda_k \mu_l)\}$.

The following statements hold:

- 1) Suppose $\lambda_{max} > 0$. Then with all other model parameters held constant, there exists a critical parameter value γ_c such that for every pair $(\lambda, \mu) \in \Lambda$, $\operatorname{Re}(\lambda) = \lambda_{max}$ whenever $\gamma > \gamma_c$;
- 2) Suppose $\mu_{max} > 0$. Then with all other model parameters held constant, there exists a critical value β_c such that for every pair $(\lambda, \mu) \in \Lambda$, $\operatorname{Re}(\mu) = \mu_{max}$ whenever $\beta > \beta_c$;
- 3) Suppose $(\lambda\mu)_{max} > 0$. Then with all other model parameters held constant, there exists a critical value δ_c such that for every pair $(\lambda, \mu) \in \Lambda$, $\operatorname{Re}(\lambda\mu) = (\lambda\mu)_{max}$ whenever $\delta > \delta_c$;

4) Suppose $\lambda_{max} \geq 0$ and $\mu_{max} \geq 0$ with $\gamma, \beta > 0$, and at least one of $\lambda_{max}, \mu_{max} \neq 0$. Then there exists a critical value K_c such that for every pair $(\lambda, \mu) \in \Lambda$, $\text{Re}(\lambda) = \lambda_{max}$ and $\text{Re}(\mu) = \mu_{max}$ whenever $(\gamma\lambda_{max} + \beta\mu_{max})/\delta > K_c$.

Proof. 1) Suppose $\eta(u, \lambda, \mu), \eta(u, \tilde{\lambda}, \tilde{\mu}) \in \sigma(J(\mathbf{0}_N, u))$, $\text{Re}(\lambda) = \lambda_{max}$, $\text{Re}(\tilde{\lambda}) < \lambda_{max}$.

$$\begin{aligned} & \text{Re}(\eta(u, \lambda, \mu)) - \text{Re}(\eta(u, \tilde{\lambda}, \tilde{\mu})) \\ &= u\gamma(\lambda_{max} - \text{Re}(\tilde{\lambda})) + u\beta(\text{Re}(\mu) - \text{Re}(\tilde{\mu})) + u\delta(\text{Re}(\lambda\mu) - \text{Re}(\tilde{\lambda}\tilde{\mu})) > 0 \end{aligned}$$

whenever

$$\gamma > \frac{1}{\lambda_{max} - \text{Re}(\tilde{\lambda})} \left(-\beta(\text{Re}(\mu) - \text{Re}(\tilde{\mu})) - \delta(\text{Re}(\lambda\mu) - \text{Re}(\tilde{\lambda}\tilde{\mu})) \right). \quad (5.6)$$

We can always find a sufficiently large value of γ so that (5.6) is satisfied for all possible combinations of $\tilde{\lambda}, \tilde{\mu}$. Then $\text{Re}(\eta(u, \lambda, \mu)) > \text{Re}(\eta(u, \tilde{\lambda}, \tilde{\mu}))$ whenever $\text{Re}(\lambda) = \lambda_{max}$ and $\text{Re}(\tilde{\lambda}) < \lambda_{max}$, regardless of the choice of $\mu, \tilde{\mu}$. From this we can conclude that $\mathcal{H} \subset \{\eta(u, \lambda, \mu) \text{ s.t. } \text{Re}(\lambda) = \lambda_{max}, \mu \in \sigma(A_o)\}$. Statements 2),3), and 4) follow by analogous arguments. □

Proposition 5.1.3 carries the following interpretation. For a network of decision-makers, the set \mathcal{H} is often generated by one of two sets of eigenvalue pairs of A_a, A_o . These are the following:

1. Let Λ_1 be a set of ordered pairs (λ, μ) with $\lambda \in \sigma(A_a)$, $\mu \in \sigma(A_o)$ defined as $(\lambda, \mu) \in \Lambda_1$ whenever $\text{Re}(\lambda) = \max_{\lambda_i \in \sigma(A_a)} \{\text{Re}(\lambda_i)\}$, $\text{Re}(\mu) = \max_{\mu_j \in \sigma(A_o)} \{\text{Re}(\mu_j)\}$.
2. Let Λ_2 be a set of ordered pairs $(\tilde{\lambda}, \tilde{\mu})$ with $\tilde{\lambda} \in \sigma(A_a)$, $\tilde{\mu} \in \sigma(A_o)$ defined as $(\tilde{\lambda}, \tilde{\mu}) \in \Lambda_2$ whenever $\text{Re}(\tilde{\lambda}\tilde{\mu}) = \max_{\lambda_k \in \sigma(A_a), \mu_l \in \sigma(A_o)} \{\text{Re}(\lambda_k \mu_l)\}$.

Based on the observations in Proposition 5.1.3, we conjecture that a typical belief-forming bifurcation at $u = u^*$ will happen along a manifold tangent to an eigenspace associated with either Λ_1 or Λ_2 . The two sets Λ_1, Λ_2 are not necessarily disjoint, and in some cases $\Lambda_2 \subseteq \Lambda_1$ or $\Lambda_1 \subseteq \Lambda_2$. For example $\Lambda_1 = \Lambda_2$ whenever both A_a and A_o have the strong Perron-Frobenius property. Whenever $\Lambda_1 \cap \Lambda_2 = \emptyset$, if Λ_1 generates \mathcal{H} we will say the resulting bifurcation is driven by social imitation, since it typically happens when the social imitation gain γ is strong. Similarly if Λ_2 generates \mathcal{H} we will say the bifurcation is driven by the belief system or ideology, since it typically corresponds to a strong ideological commitment gain δ .

5.2 Pitchfork bifurcation and its unfolding

The cardinality of the leading eigenvalue set \mathcal{H} corresponds to the the number of eigenvalues of $J(\mathbf{0}, u)$ that cross the imaginary axis with nonzero speed as u is varied through u^* when the origin $\mathbf{Z} = \mathbf{0}$ loses stability in an indecision-breaking bifurcation. In this section we examine the simplest case, when (5.3) has exactly one leading eigenvalue. In the following lemma we establish that this necessarily corresponds to an interaction of simple, real eigenvalues of A_a and A_o .

Lemma 5.2.1. *Suppose $|\mathcal{H}| = 1$. Then there exist simple and real eigenvalues $\lambda \in \sigma(A_a)$ and $\mu \in \sigma(A_o)$ for which $\mathcal{H} = \{\eta(u, \lambda, \mu)\}$; furthermore, it must hold that $(\lambda, \mu) \in \Lambda_1$ or $(\lambda, \mu) \in \Lambda_2$.*

Proof. Following (5.4), define $\eta(u, \lambda_i, \mu_j) = -d + u(\alpha + \gamma\lambda_i + \beta\mu_j + \delta\lambda_i\mu_j)$ and observe that $\eta(u, \lambda, \mu) \in \mathcal{H}$ implies $\eta(u, \bar{\lambda}, \bar{\mu}) \in \mathcal{H}$ since $\text{Re}(\eta(u, \lambda, \mu)) = \text{Re}(\eta(u, \bar{\lambda}, \bar{\mu}))$. Therefore if $\text{Im}(\lambda)$ and/or $\text{Im}(\mu)$ are nonzero, then $|\mathcal{H}| > 1$ and the assumption in the lemma is violated. Furthermore, if either λ, μ have multiplicity greater than 1, then $|\mathcal{H}| > 1$ as a consequence of Proposition 5.1.1. Therefore λ, μ must both be real and simple. The last statement then follows trivially from the definitions of \mathcal{H} , Λ_1 , and Λ_2 . \square

As a consequence of Lemma (5.2.1), whenever the leading eigenvalue of (5.3) is simple there are at most two manifolds along which the indecision-breaking bifurcation of the origin can appear, which greatly constraints the structure of the possible beliefs that can arise in this bifurcation. Next, we establish in the following theorem that whenever $J(\mathbf{0}, u)$ has a simple leading eigenvalue the indecision-breaking behavior of the system (5.1) is organized by a symmetric pitchfork singularity.

Theorem 5.2.2 (Multi-option pitchfork bifurcation). *Consider (5.1) with communication graph \mathcal{G}_a and belief system graph \mathcal{G}_o , and suppose $|\mathcal{H}| = 1$, i.e. the leading eigenvalue of (5.3) $\eta_{max} = -d + u(\alpha + \gamma\lambda_a + \beta\mu_o + \delta\lambda_a\mu_o) \in \mathbb{R}$ is simple for some $\lambda_a \in \sigma(A_a)$, $\mu_o \in \sigma(A_o)$. Assume $\alpha + \gamma\lambda_a + \beta\mu_o + \delta\lambda_a\mu_o > 0$. Let $\mathbf{v}_a, \mathbf{w}_a \in \mathbb{R}^{N_a}$ and $\mathbf{v}_o, \mathbf{w}_o \in \mathbb{R}^{N_o}$ be the right and left eigenvectors of A_a and A_o corresponding to λ_a and μ_o , respectively, normalized to satisfy $\langle \mathbf{w}_a, \mathbf{v}_a \rangle = 1, \langle \mathbf{w}_o, \mathbf{v}_o \rangle = 1$. Finally, let $f(y, u, \mathbf{b})$ be a Lyapunov-Schmidt reduction of (5.1) at $(\mathbf{Z}, u, \mathbf{b}) = (\mathbf{0}, u^*, \mathbf{0})$.*

1) *Suppose g_1 and g_2 are zero, i.e. the saturating functions S_1, S_2 have an odd symmetry. At $(\mathbf{Z}, u, \mathbf{b}) = (\mathbf{0}, u^*, \mathbf{0})$, the system undergoes a symmetric pitchfork bifurcation. For values of $u \neq u^*$, $|u - u^*| < \epsilon$ for small ϵ two branches of equilibria of (5.1) bifurcate from $\mathbf{Z} = \mathbf{0}$ along a manifold tangent at $u = u^*$ to $\text{span}\{\mathbf{v}_a \otimes \mathbf{v}_o\}$. If*

$$(S_1'''(0)(\alpha + \gamma\lambda_a)^3 + S_2'''(0)(\beta\mu_o + \delta\lambda_a\mu_o)^3) \langle \mathbf{w}_a, \mathbf{v}_a^3 \rangle \langle \mathbf{w}_o, \mathbf{v}_o^3 \rangle < 0 (> 0)$$

where $\mathbf{x}^3 = \mathbf{x} \odot \mathbf{x} \odot \mathbf{x}$, the bifurcation happens supercritically (subcritically). The two bifurcating fixed points are locally exponentially stable (unstable).

2) When one or more of g_1 , g_2 , and b_{ij} are nonzero, the bifurcation problem $f(y, u, \mathbf{b}) = 0$ is an unfolding of the pitchfork singularity near $(z, u, \mathbf{b}) = (0, u^*, \mathbf{0})$. If $\mathbf{b} = \mathbf{0}$ and

$$(g_1''(0)(\alpha + \gamma\lambda_a)^2 + g_2''(0)(\beta\mu_o + \delta\lambda_a\mu_o)^2) \langle \mathbf{w}_a, \mathbf{v}_a^2 \rangle \langle \mathbf{w}_o, \mathbf{v}_o^2 \rangle > 0 (< 0),$$

then the bifurcation at $u = u^*$ is transcritical, with the branch of equilibria satisfying $\langle \mathbf{v}_a \otimes \mathbf{v}_o, \mathbf{Z} \rangle > 0 (< 0)$ bifurcating subcritically. Furthermore, whenever

$$\langle \mathbf{w}_a \otimes \mathbf{w}_o, \mathbf{b} \rangle > 0 (< 0) \tag{5.7}$$

on a small neighborhood of u near u^* the local bifurcation diagram of (5.1) has a unique equilibrium which satisfies $\langle \mathbf{v}_a \otimes \mathbf{v}_o, \mathbf{Z} \rangle > 0 (< 0)$.

For a proof of Theorem 5.2.2, see Appendix B.2. We illustrate the intuition of Theorem 5.2.2 with the following example.

Example 5.2.3 (5 Agents, 3 Options). Consider a network of 5 agents evaluating 3 options according to (5.1), with communication graph \mathcal{G}_a and belief system graph \mathcal{G}_o characterized by the adjacency matrices

$$A_a = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 \\ -1 & 0 & -1 & 1 & -1 \\ 1 & 1 & 0 & 0 & 1 \\ 1 & -1 & 0 & 0 & -1 \\ 1 & 0 & 1 & -1 & 0 \end{pmatrix}, \quad A_o = \begin{pmatrix} 0 & -1 & 1 \\ 1 & 0 & -1 \\ 1 & 0 & 0 \end{pmatrix}. \tag{5.8}$$

Network diagrams of these communication and belief system graphs is shown in Figure 5.1 (a) and (b), with positive edges represented by blue lines and negative edges represented by red lines. Let the parameters of (5.1) be $d = 1$, $\alpha = \gamma = \beta = \delta = 0.1$. The eigenvalues of A_a are

$$\lambda_1 \approx 0.823, \quad \lambda_{2,3} \approx 0.745 \pm 1.106i, \quad \lambda_{4,5} \approx -1.157 \pm 0.327i$$

and the eigenvalues of A_o are

$$\mu_1 = 1, \quad \mu_{2,3} = -\frac{1}{2} \pm \frac{\sqrt{3}}{2}i.$$

The set of leading eigenvalues of A_a, A_o is $\Lambda_1 = \{(\lambda_1, \mu_1)\}$ and the set of eigenvalue pairs that

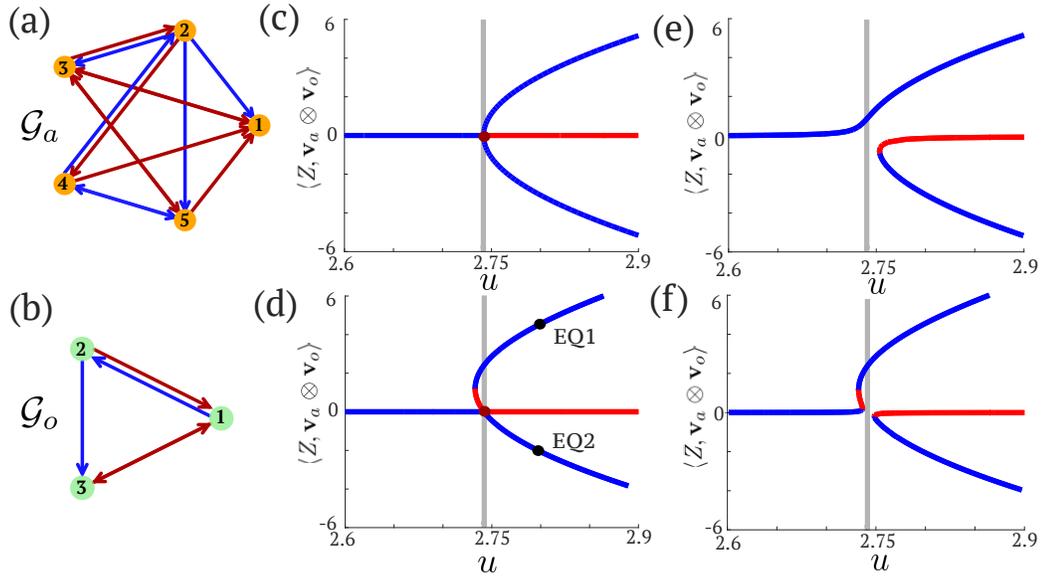


Figure 5.1: (a) and (b): communication and belief system graphs for Example 5.2.3. Red edges have a positive signature, and blue edges have a negative signature; (c) symmetric pitchfork bifurcation with $\varepsilon_1 = \varepsilon_2 = 0$, $\langle \mathbf{w}_a \otimes \mathbf{w}_o, \mathbf{b} \rangle = 0$; (d) pitchfork unfolding with $\varepsilon_1, \varepsilon_2 \neq 0$, $\langle \mathbf{w}_a \otimes \mathbf{w}_o, \mathbf{b} \rangle = 0$; (e) pitchfork unfolding with $\varepsilon_1 = \varepsilon_2 = 0$, $\langle \mathbf{w}_a \otimes \mathbf{w}_o, \mathbf{b} \rangle > 0$; (f) $\varepsilon_1, \varepsilon_2 \neq 0$, $\langle \mathbf{w}_a \otimes \mathbf{w}_o, \mathbf{b} \rangle > 0$. In (c)-(f) vertical gray line indicates the bifurcation point $u = u^* \approx 0.2742$, red lines represent unstable equilibria, and blue lines represent stable equilibria. Bifurcation diagrams in (c)-(f) generated with MatCont numerical continuation package [47].

maximize the real part of the product of eigenvalues of A_a, A_o is $\Lambda_2 = \{(\lambda_4, \mu_3), (\lambda_5, \mu_2)\}$. For the stated choice of model parameters, the leading eigenvalue of the Jacobian (5.3) is simple and generated by the set Λ_1 , with corresponding right eigenvectors

$$\mathbf{v}_a \approx (0.463, -0.272, 0.381, 0.743, 0.123)^T, \quad \mathbf{v}_o = (1, 0, 1)^T, \quad (5.9)$$

and left eigenvectors

$$\mathbf{w}_a \approx (0.999, 0.322, 0.739, 0.474, -0.069)^T, \quad \mathbf{w}_o = \left(\frac{1}{3}, -\frac{1}{3}, \frac{2}{3}\right)^T.$$

where values were rounded to three digits past the decimal point. The critical value of attention at which the indecision-breaking bifurcation occurs is

$$u^* = \frac{d}{\alpha + \gamma\lambda_1 + \beta\mu_1 + \delta\mu_1\lambda_1} \approx 2.742.$$

Let the two saturating functions in (5.1) be

$$S_1(y) = \tanh(y + \varepsilon_1 \tanh y^2), \quad S_2(y) = \frac{1}{2} \tanh(2(y + \varepsilon_2 \tanh y^2)) \quad (5.10)$$

where $\varepsilon_1, \varepsilon_2 \in \mathbb{R}$. When $\varepsilon_1 = \varepsilon_2 = 0$ and therefore $g_1 = g_2 = 0$, meaning S_1 and S_2 have an odd symmetry. In Figure 5.1 (c)-(f) we show four numerically generated bifurcation diagrams for (5.1). For diagrams (c) and (e), $\varepsilon_1 = \varepsilon_2 = 0$; for diagrams (d) and (f), $\varepsilon_1 = \varepsilon_2 = 0.1$. For diagrams (c) and (d), all agents are unbiased; in diagrams (e) and (f), the nonzero biases are $b_{11} = 0.001$, $b_{22} = 0.003$, $b_{53} = -0.002$ and

$$\langle \mathbf{w}_a \otimes \mathbf{w}_o, \mathbf{b} \rangle = (\mathbf{w}_a)_1(\mathbf{w}_o)_1 b_{53} + (\mathbf{w}_a)_2(\mathbf{w}_o)_2 b_{22} + (\mathbf{w}_a)_5(\mathbf{w}_o)_3 b_{11} \approx 0.0001 > 0.$$

Together these four diagrams illustrate how allowing for nonzero biases and for asymmetric saturation functions, the value dynamics (5.1) on these graphs realize all of the topologically distinct bifurcation diagrams in a universal unfolding of a pitchfork bifurcation [82, Chapter I]. When the agents are unbiased and both saturation functions have an odd symmetry, two nontrivial equilibria $\mathbf{Z}^*, -\mathbf{Z}^*$ emerge on the network in a supercritical pitchfork bifurcation shown in Figure 5.1(c); both of these equilibria are stable at bifurcation. When agents are unbiased and the saturation functions are asymmetric, the bifurcation at $u = u^*$ is transcritical which means one of the solution branches

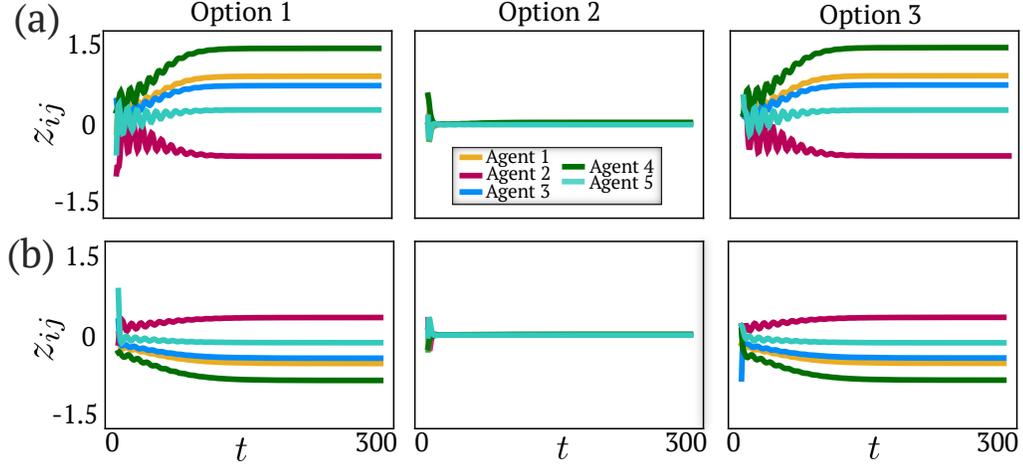


Figure 5.2: Trajectories $z_{ij}(t)$ for value dynamics (5.1) for 5 agents evaluating 3 options, with setup and parameters as described in Example 5.2.3. Initial conditions for each $z_{ij}(0)$ were generated randomly. The figure illustrates the distribution of agents' states z_{ij} along each option dimension as the network settles to (a) EQ1 from Figure 5.1(d), (b) EQ2 from Figure 5.1(d).

is unstable at bifurcation and gains stability in a saddle-node bifurcation at some value of $u < u^*$ - see Figure 5.1(d). Finally, when the agents are unbiased and $\langle \mathbf{w}_a \otimes \mathbf{w}_o, \mathbf{b} \rangle > 0 (< 0)$, in both cases the bifurcation diagram breaks up into two continuous curves. For values of u near u^* the system admits a single stable equilibrium that satisfies $\langle \mathbf{v}_a \otimes \mathbf{v}_o, \mathbf{Z} \rangle > 0 (< 0)$ - see Figure 5.1(e),(f).

Finally we can describe the relationship between the right eigenvectors $\mathbf{v}_a, \mathbf{v}_o$ (5.9) and the distribution of agents' beliefs at equilibrium post-bifurcation. The center manifold along which the bifurcation occurs is tangent at $u = u^*$ to $\text{span}(\mathbf{v}_a \otimes \mathbf{v}_o)$. This means that near the critical value u^* , the opinions formed on the network are approximately a multiple of the vector $\mathbf{v}_a \otimes \mathbf{v}_o$. To illustrate this point, in Figures 5.2 and 5.3 we show representative trajectories of the system with $\varepsilon_1 = \varepsilon_2 = 0$ and $\mathbf{b} = 0$ at a supercritical value of the attention parameter, $u = u^* + 0.05$. In part (a) of Figure 5.2 and in Figure 5.3, the trajectories of the system settle to the point labeled as EQ1 in the bifurcation diagram in Figure 5.1. Similarly, in part (b) in Figure 5.2 the trajectories settle to EQ2.

In Figure 5.2 the trajectories of the system are grouped by option and it can easily be seen that the relative strength of commitment of each agent along each option dimension is determined by the eigenvector of the communication graph \mathbf{v}_a . Specifically, if $|(\mathbf{v}_a)_i| > |(\mathbf{v}_a)_k|$ then $|z_{ij}| > |z_{kj}|$ for any option j on which agents i and k form strong opinions (i.e. $(\mathbf{v}_a)_i(\mathbf{v}_o)_j \neq 0$ and $(\mathbf{v}_a)_k(\mathbf{v}_o)_j \neq 0$). Similarly, if $\text{sign}((\mathbf{v}_a)_i) \neq \text{sign}((\mathbf{v}_a)_k)$ then $\text{sign}(z_{ij}) \neq \text{sign}(z_{kj})$, i.e. agents i and k disagree in their belief about any option j on which both form strong opinions. Specifically, in the eigenvector \mathbf{v}_a we see that the second entry has opposite sign from the rest. This means that agent 2 will agree with all other agents on any option on which the network forms strong opinions, and this is indeed

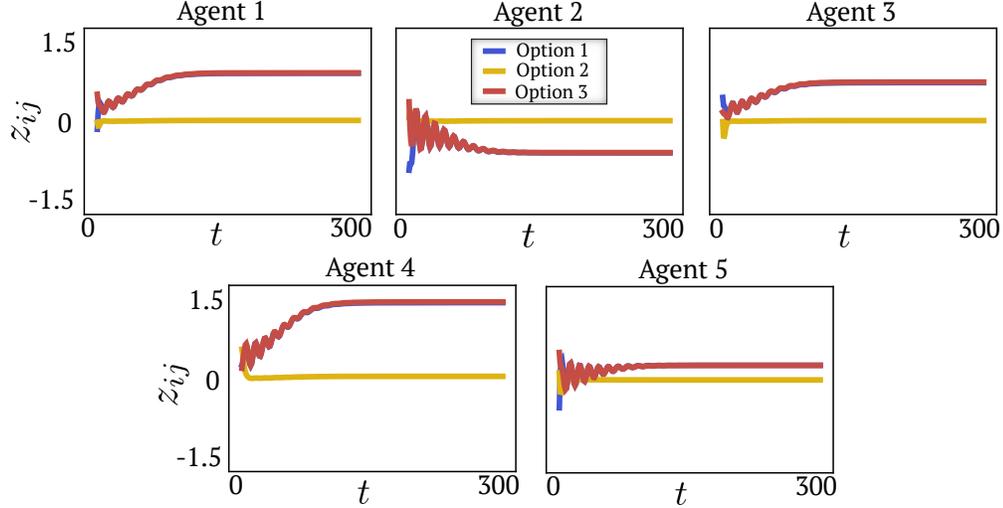


Figure 5.3: Same trajectories as Figure 5.2(a), grouped by agent instead of option. The Option 1 blue curve overlaps with the Option 3 red curve for most of the simulation.

the observed behavior for both equilibria to which the network converges in Figure 5.2.

Notice that the belief trajectories on option 2 remain near neutral at equilibrium. This observation can be inferred from the observation that the second entry of \mathbf{v}_o is zero. The role of \mathbf{v}_o can easily be deduced when we group the network belief trajectories by agent, instead of by option, in Figure 5.3. From this figure we can observe that \mathbf{v}_o defines the distribution of each agent's internal beliefs at equilibrium. Specifically, if $|(\mathbf{v}_o)_j| > |(\mathbf{v}_o)_l|$ then $|z_{ij}| > |z_{il}|$ for any agent i and options j, l on which it forms strong opinions. Similarly, if $\text{sign}((\mathbf{v}_o)_j) \neq \text{sign}((\mathbf{v}_o)_l)$ then $\text{sign}(z_{ij}) \neq \text{sign}(z_{il})$, i.e. agent i always either chooses option j and rejects option l , or chooses option l and rejects option j . For the presented example, $(\mathbf{v}_o)_1 = (\mathbf{v}_o)_3$ and $(\mathbf{v}_o)_2 = 0$; and indeed we see in Figure 5.3 that all agents remain close to neutral on option 2, and form comparatively strong beliefs of close magnitude about options 1 and 3.

We further illustrate the relationship between the eigenvectors \mathbf{v}_a , \mathbf{v}_o and the system trajectories in Figure 5.4. Let $\mathbf{Z}_j^\dagger = (z_{1j}, z_{2j}, z_{3j}, z_{4j}, z_{5j})$ be the vector of network beliefs on option j . In Figure 5.4(a) for each option j we normalize the vector \mathbf{Z}_j^\dagger to unit magnitude at each time step of the simulation of Figure 5.2(a), and project its absolute value onto the unit eigenvector \mathbf{v}_a . For both of the options on which the agents form strong opinions, options 1 and 3, this projection converges to a value close to 1 over time, which tells us that the network opinions along each option are settling to a multiple of \mathbf{v}_a . Analogously, in Figure 5.4(b) for each agent i we normalize its opinion vector \mathbf{Z}_i to unit magnitude at each time step of the simulation in Figure 5.3, and project its absolute value onto the unit eigenvector \mathbf{v}_a . For all agents, this projection converges to a value close to 1 over

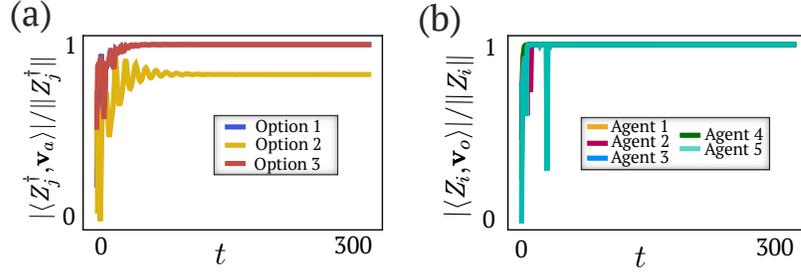


Figure 5.4: (a) Projections of the normalized trajectories in each panel of Figure 5.2(a) onto \mathbf{v}_a ; (b) projections of the normalized trajectories in each panel of Figure 5.3 onto \mathbf{v}_o .

time, which tells us that network opinions of each agents are settling to a multiple of \mathbf{v}_o .

In summary, in the pitchfork bifurcation and its unfolding established in Theorem 5.2.2, the eigenvector \mathbf{v}_a of the communication graph \mathcal{G}_a informs the relative beliefs of agents along each option dimension post-bifurcation, whereas the eigenvector \mathbf{v}_o of the belief system graph \mathcal{G}_o informs the internal distribution of beliefs of each agent. We saw in Example 5.2.3 how whenever $(\mathbf{v}_a)_i(\mathbf{v}_o)_j = 0$, the beliefs of agent i about option j remain near neutral at equilibrium post bifurcation, even if the agent forms strong beliefs on other options. We will say that agent i forms *strong beliefs* about option j in a pitchfork bifurcation whenever $(\mathbf{v}_a)_i(\mathbf{v}_o)_j \neq 0$. To conclude our discussion of bifurcations of equilibria, we establish a sufficient condition for all agents to form strong beliefs on all options, analogous to Proposition 4.3.5 for the scalar-variable system.

Proposition 5.2.4 (Strong beliefs on multiple options). *Consider (5.1) on a communication graph \mathcal{G}_a with a belief system graph \mathcal{G}_o , with corresponding signed adjacency matrices A_a, A_o . Suppose $b_{ij} = 0$ for all $i = 1, \dots, N_a, j = 1, \dots, N_o$ and $g_1 = g_2 = 0$. Whenever A_a and A_o are both switching equivalent to eventually positive matrices, the indecision state $\mathbf{Z} = \mathbf{0}$ loses stability in a pitchfork bifurcation along a center manifold on which all agents form strong beliefs on all options.*

Proof. The adjacency matrices A_a and A_o are each similar to a matrix that has the strong Perron Frobenius property by the argument presented in the proof of Proposition 4.3.5. Let $\lambda_{max} > 0$ and $\mu_{max} > 0$ be their dominant eigenvalues with corresponding right eigenvectors $\mathbf{v}_a, \mathbf{v}_o$. For any pair of eigenvalues $\lambda_i \in \sigma(A_a), \mu_j \in \sigma(A_o)$ that are not equal to λ_{max}, μ_{max} , $|\lambda_i \mu_j| < \lambda_{max} \mu_{max}$ and therefore $\text{Re}(\lambda_i \mu_j) < \lambda_{max} \mu_{max}$ which means $\Lambda_1 = \Lambda_2 = \{(\lambda_{max}, \mu_{max})\}$. The leading eigenvalue of the Jacobian (5.3) is therefore $\eta_{max} = -d + u(\alpha \lambda_{max} + \beta \lambda_{max} + \delta \lambda_{max} \mu_{max}) > 0$, and it is real and simple which means $|\mathcal{H}| = 1$ and a pitchfork bifurcation is established by Theorem 5.2.2. The vector $\mathbf{v}_a \otimes \mathbf{v}_o$ has no zero entries, which means all agents form strong beliefs on all options in this bifurcation. \square

5.3 Hopf bifurcation

So far in this dissertation we have focused our attention on describing belief patterns that arise in the networked dynamics (5.1) as a consequence of indecision-breaking pitchfork bifurcations of equilibria. Another common type of bifurcation encountered in this system is a Hopf bifurcation in which the origin $\mathbf{Z} = \mathbf{0}$ loses stability and gives rise to a family of periodic orbits. In general, oscillations are a meaningful phenomenon in the context of opinion and belief formation. Oscillations of beliefs and attitudes of have been hypothesized and observed in individuals [60] and in computational models of social opinion formation [30, 74, 210, 215]. Famously, public opinion surveys in the United States that track voters' policy positions over time show alternating swings towards more conservative and more liberal voter preferences [192, 193]. For a more informal example, oscillations in values are commonly encountered in social communities on topics such as fashion trends in clothing, hair styles, and musical tastes that tend to re-emerge in popularity skipping a few years or generations. In the context of decision-making, an oscillatory network state may enable a system to systematically explore the various available options by visiting different configurations over time without committing to a set of beliefs. In the following theorem we formalize the emergence of persistent periodic oscillations in (5.1) and describe the role of the communication graph \mathcal{G}_a and belief system graph \mathcal{G}_o in determining the relative amplitudes and phases of agents' opinion states in these oscillations.

Theorem 5.3.1 (Hopf bifurcation). *Consider (5.1) with communication graph \mathcal{G}_a and belief system graph \mathcal{G}_o , with $\mathbf{b} = \mathbf{0}$. Assume that i) the leading eigenvalues of the Jacobian (5.3) are generated by some combinations of $\lambda_{\pm} = \lambda_a \pm i\lambda_c \in \sigma(A_a)$ and $\mu_{\pm} = \mu_o \pm i\mu_c \in \sigma(A_o)$, with at least one of $\lambda_c, \mu_c \neq 0$; i.e. $\mathcal{H} = \{\eta(u), \overline{\eta(u)}\}$ with $\eta(u) = -d + u(\alpha + \gamma\lambda_a + \beta\mu_o + \delta(\lambda_a\mu_o + \lambda_c\mu_c)) + iu(\gamma\lambda_c - \beta\mu_c + \delta(-\lambda_a\mu_c + \lambda_c\mu_o))$; ii) $\alpha + \gamma\lambda_a + \beta\mu_o + \delta(\lambda_a\mu_o + \lambda_c\mu_c) > 0$.*

Suppose λ, μ generate $\eta(u)$. Let $\mathbf{v}_a, \mathbf{w}_a \in \mathbb{C}^{N_a}$ be the right and left eigenvectors of A_a corresponding to λ and $\bar{\lambda}$, respectively; let $\mathbf{v}_o, \mathbf{w}_o \in \mathbb{C}^{N_o}$ be the right and left eigenvectors of A_o corresponding to μ and $\bar{\mu}$, respectively. Choose the eigenvectors to satisfy the normalization

$$\langle \mathbf{w}_a \otimes \mathbf{w}_o, \mathbf{v}_a \otimes \mathbf{v}_o \rangle = 2, \quad \langle \overline{\mathbf{w}_a \otimes \mathbf{w}_o}, \mathbf{v}_a \otimes \mathbf{v}_o \rangle = 0. \quad (5.11)$$

1) *There is a unique 3-dimensional center manifold $W^c \subset \mathbb{R}^{N_a N_o} \times \mathbb{R}$ passing through $(\mathbf{Z}, u) = (\mathbf{0}, u^*)$, tangent to $\text{span}\{\text{Re}(\mathbf{v}_a \otimes \mathbf{v}_o), \text{Im}(\mathbf{v}_a \otimes \mathbf{v}_o)\}$ at $u = u^*$. There is a family of periodic orbits of (5.1) that bifurcates from the neutral equilibrium $\mathbf{Z} = \mathbf{0}$ along W_c at $u = u^*$;*

2) *When $|u - u^*|$ is small, the period of the solutions is near $2\pi/|\gamma\lambda_c + \beta\mu_c + \delta(\lambda_a\mu_c + \lambda_c\mu_o)|$, the*

difference in phase between $z_{ij}(t)$ and $z_{kl}(t)$ is near $\varphi = \arg((\mathbf{v}_a)_i(\mathbf{v}_o)_j) - \arg((\mathbf{v}_a)_k(\mathbf{v}_o)_l)$, and the amplitude of $z_{ij}(t)$ is greater than the amplitude of $z_{kl}(t)$ if and only if $|(\mathbf{v}_a)_i| |(\mathbf{v}_o)_j| > |(\mathbf{v}_a)_k| |(\mathbf{v}_o)_l|$;

3) Whenever

$$b_2 = \text{Re} \left(\left(S_1'''(0) ((\alpha + \gamma\lambda_a)^2 + (\gamma\lambda_c)^2) (\alpha + \gamma(\lambda_a + i\lambda_c)) \right. \right. \\ \left. \left. + S_2'''(0) (\mu_o^2 + \mu_c^2) ((\beta + \delta\lambda_a)^2 + (\delta\lambda_c)^2) (\mu_o(\beta + \delta\lambda_a) - \delta\mu_c\lambda_c \right. \right. \\ \left. \left. + i(\mu_c(\beta + \delta\lambda_a) + \delta\lambda_c\mu_o) \right) \langle \mathbf{w}_a, |\mathbf{v}_a|^2 \odot \mathbf{v}_a \rangle \langle \mathbf{w}_o, |\mathbf{v}_o|^2 \odot \mathbf{v}_o \rangle \right) < 0, \quad (5.12)$$

where $|\mathbf{x}|^2 = \bar{\mathbf{x}} \odot \mathbf{x}$, the bifurcating periodic solutions appear supercritically (for $u > u^*$) and are locally asymptotically stable; whenever (5.12) is instead positive, the solutions appear subcritically and are unstable.

For a proof of Theorem 5.3.1, see Appendix B.3. Observe that the period of oscillations established by Theorem 5.3.1 is inversely proportional to the social gains β, γ, δ . The beliefs of agents with weak social coupling and weak internal adherence to social beliefs will change slowly over time, whereas strong social coupling or strong adherence to a belief system results in high-frequency oscillations among all agents on the network. To illustrate the findings of Theorem 5.3.1 we consider the following examples. In both of these examples we define the saturation functions S_1, S_2 as (5.10) with $\varepsilon_1 = \varepsilon_2 = 0$.

Example 5.3.2 (Oscillations). *a) First we return to the dynamics of five agents evaluating three options on communication graph \mathcal{G}_a and belief system graph \mathcal{G}_o that we examined in Example 5.2.3. The adjacency matrices of the two graphs are stated in (5.8) and their network diagram representations can be found in Figure 5.1(a) and (b). We consider the dynamics (5.1) with parameters $d = 10$, $\alpha = \gamma = \beta = 0.1$, $\delta = 12$, $\varepsilon_1 = \varepsilon_2 = 0$. Compared to Example 5.2.3, the ideological coupling δ was increased substantially from 0.1 to 12; the damping coefficient d was also increased to keep the bifurcation value of u further away from zero. In proposition 5.1.3 we established that when the ideological coupling gain δ is large, then the network bifurcation is generated by the eigenvalues of A_a and A_o in the set Λ_2 , that maximize the real part of the eigenvalue product. This is indeed what happens with the stated choice of model parameters. The leading eigenvalue set \mathcal{H} of the Jacobian matrix (5.3) is generated by two pairs of the complex conjugate eigenvalues $\lambda_{4,5} \approx -1.157 \pm 0.327i$ and $\mu_{2,3} = -\frac{1}{2} \pm \frac{\sqrt{3}}{2}i$. The indecision-breaking bifurcation happens when two complex conjugate eigenvalues cross the imaginary axis, which suggests an onset of oscillations in a Hopf bifurcation.*

A choice of eigenvectors defined in Theorem 5.3.1 that satisfy the normalization condition (5.11) are

$$\begin{aligned}\mathbf{v}_a &= (0.468 + 0.132i, 0.073 + 0.119i, -0.584, -0.105 - 0.423i, 0.135 - 0.442i)^T, \\ \mathbf{v}_o &= (0.447i, -0.775, -0.387 - 0.224i)^T,\end{aligned}$$

$$\begin{aligned}\mathbf{w}_a &= (0.990 - 1.181i, -1.123 - 0.476i, 0.309 + 1.043i, -0.834 + 0.126i, -2.129 - 0.602i)^T, \\ \mathbf{w}_o &= (-0.577, -0.289 - 0.5i, 0.577)^T.\end{aligned}$$

These eigenvectors were generated numerically and all values have been rounded to three decimal points when necessary. The critical value of attention at which the bifurcation occurs is

$$u^* = \frac{d}{\alpha + \gamma \operatorname{Re}(\lambda_4) + \beta \operatorname{Re}(\mu_3) + \delta \operatorname{Re}(\mu_3 \lambda_4)} = \frac{d}{\alpha + \gamma \operatorname{Re}(\lambda_5) + \beta \operatorname{Re}(\mu_2) + \delta \operatorname{Re}(\mu_2 \lambda_5)} \approx 0.973.$$

Furthermore we compute a numerical estimate of the quantity (5.12) which yields $b_2 \approx -264 < 0$ (rounded to the nearest integer) which means the periodic orbits should appear supercritically, and will be stable at bifurcation according to the third part of Theorem 5.3.1.

Figure 5.5 shows simulated trajectories of the small oscillations the system settles on from randomized initial conditions at an attention value slightly above the bifurcation point, $u = 1$. The same trajectories are showed in parts (a) and (b) of the figure, grouped by beliefs about each option in part (a) and by each agent's internal beliefs in part (b). The following properties of the observed oscillations can be inferred from Theorem 5.3.1:

- The oscillation period is close to $\frac{2\pi}{|\gamma \operatorname{Im}(\lambda_5) + \beta \operatorname{Im}(\mu_2) + \delta \operatorname{Im}(\mu_2 \lambda_5)|} \approx 0.6279$
- Relationships between amplitudes of oscillations of beliefs on any given option can be inferred by the magnitude of the entries of the eigenvector \mathbf{v}_o (5.3.2); for example, we expect agent 3 to have the largest amplitude belief oscillations on all options, and agent 2 to have the lowest amplitude belief oscillations, which is consistent with the simulation - see Figure 5.5(a).
- Relationships between amplitudes of oscillations of beliefs internal to any agent can be inferred by the magnitude of the entries of eigenvector \mathbf{v}_a (5.3.2); for example, we expect belief oscillations about option 2 to have the largest amplitude, and belief oscillations about options 1 and 3 to be close in amplitude. Again, this is consistent with the simulation - see Figure 5.5(b).
- For all $i, k = 1, \dots, N_a$, $j, l = 1, \dots, N_o$, it holds that $\arg((\mathbf{v}_a)_i (\mathbf{v}_o)_j) \neq \arg((\mathbf{v}_a)_k (\mathbf{v}_o)_l)$, which

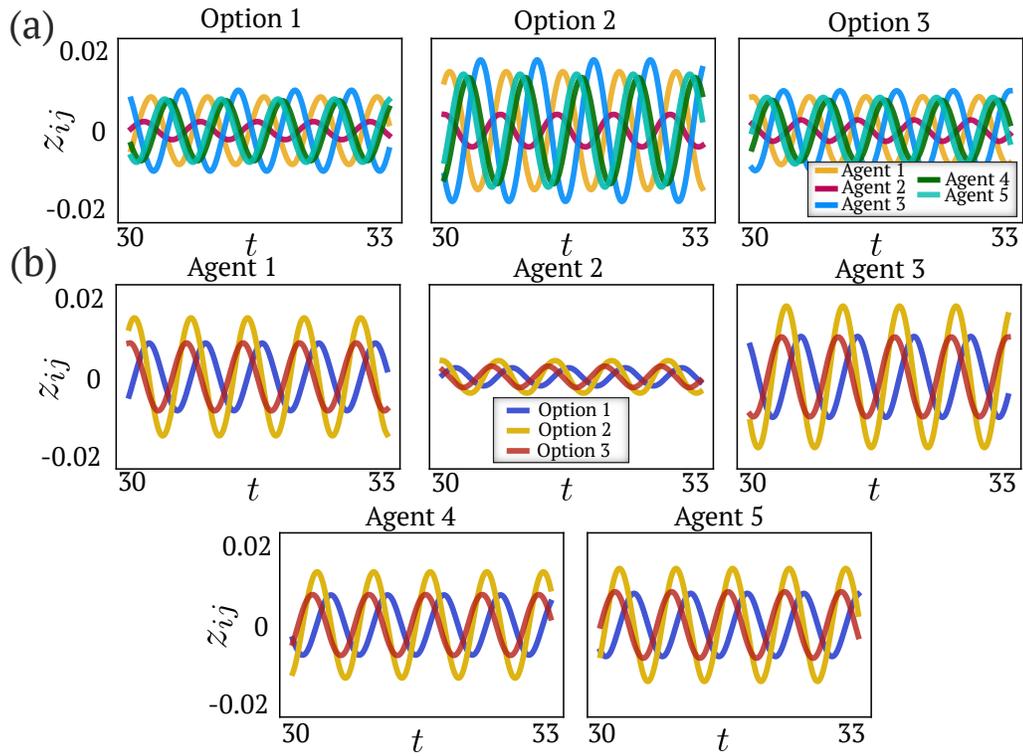


Figure 5.5: Trajectories $z_{ij}(t)$ of stable oscillation described in Example 5.3.2(a) on the same communication and belief system graphs shown in Figure 5.1(a),(b); parameters $d = 10$, $u = 1$, $\alpha = \gamma = \beta = 0.1$, $\delta = 12$, $\varepsilon_1 = \varepsilon_2 = 0$; (a) grouped by option; (b) grouped by agent.

means that there is a phase offset between any pair of periodic belief trajectories $z_{ij}(t)$, $z_{kl}(t)$. And indeed in Figure 5.5 we see that all trajectories are out-of-phase with one another, with the phase offset constant over time as all of the oscillations have the same period.

Interestingly, when the beliefs on these exact same graphs \mathcal{G}_a and \mathcal{G}_o converged to an equilibrium in Example 5.2.3, all agents remained neutral on option 2. In contrast, for a different choice of model parameter δ on the same exact network, oscillations of beliefs on option 2 have the strongest amplitude for all agents. This example illustrates clearly how a difference in the strength of the gains β, γ, δ can result in drastically different belief-forming behaviors of the agents. For the choice of \mathcal{G}_a and \mathcal{G}_o studied in Example 5.2.3 and in the present discussion, the agents' beliefs converge to equilibrium when social imitation dominates, and oscillate when ideology dominates in the dynamics (5.1).

b) In the previous example, we saw an oscillation of opinions in which all of the belief trajectories were out of phase with one another which means that at a random snapshot in time, there can be many potential configurations of beliefs in the network. This is typically what happens any time the leading eigenvalue set \mathcal{H} is generated by interactions of complex conjugate eigenvalues of both A_a and A_o . For completeness of discussion, we present another example to illustrate how certain networked opinions can synchronize in-phase or out of phase when \mathcal{H} is generated by an interaction of a simple eigenvalue of one graph with a complex-conjugate pair of the second.

We consider a network of four agents evaluating nine options on communication graph \mathcal{G}_a and belief system graph \mathcal{G}_o whose adjacency matrices are

$$A_a = \begin{pmatrix} 0 & 0 & -1 & 1 \\ 0 & 0 & 1 & 0 \\ -1 & 1 & 0 & -1 \\ 1 & -1 & 0 & 0 \end{pmatrix}, \quad A_o = \begin{pmatrix} 0 & 0 & -1 & -1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ -1 & 1 & 0 & -1 & 0 & 0 & 0 & -1 & 1 \\ 0 & -1 & 0 & 0 & 0 & -1 & -1 & 0 & 0 \\ -1 & 0 & 0 & -1 & 0 & 0 & -1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 & -1 & 0 & 0 \\ 0 & -1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \end{pmatrix}. \quad (5.13)$$

Network diagrams of these communication and belief system graphs are shown in Figure 5.6(a) and

(b). The eigenvalues of A_a are

$$\lambda_1 = 2, \quad \lambda_2 = 0, \quad \lambda_3 = \lambda_4 = -1$$

and the eigenvalues of A_o are

$$\mu_{1,2} \approx 1.467 \pm 0.599i, \quad \mu_{3,4} \approx 0.088 \pm 0.661i, \quad \mu_5 = \mu_6 = 0, \quad \lambda_7 \approx -0.447, \quad \mu_{8,9} = -1.332 \pm 0.486i.$$

where the values were rounded to three digits past the decimal point. The set of pairs of leading eigenvalues of A_a and A_o is $\Lambda_1 = \{(\lambda_1, \mu_1), (\lambda_1, \mu_2)\}$. Let the parameters of the model (5.1) be $d = 1$, $\alpha = \beta = \gamma = \delta = 0.1$. For this choice of parameters, Λ_1 generates the leading eigenvalues of the Jacobian (5.3) which are a complex conjugate pair, i.e. the indecision-breaking bifurcation is a Hopf bifurcation of the origin as established in Theorem 5.3.1. The critical value of attention at which the bifurcation occurs is

$$u^* = \frac{d}{\alpha + \gamma\lambda_1 + \beta \operatorname{Re}(\mu_1) + \delta\lambda_1 \operatorname{Re}(\mu_1)} = \frac{d}{\alpha + \gamma\lambda_1 + \beta \operatorname{Re}(\mu_2) + \delta\lambda_2 \operatorname{Re}(\mu_2)} \approx 1.351.$$

A choice of eigenvectors defined in Theorem 5.3.1 which satisfy the normalization (5.11) are

$$\mathbf{v}_a = (-0.539, 0.324, 0.647, -0.431)^T,$$

$$\mathbf{v}_o = (0.243 + 0.164i, 0, -0.630, 0.176 - 0.095i, 0.048 - 0.310i, 0, \\ -0.202 + 0.244i, 0.404 - 0.277i, -0.102 + 0.170i)^T,$$

$$\mathbf{w}_a = (-0.5, 0.5, 0.5, -0.5)^T,$$

$$\mathbf{w}_o = (0.979 + 0.796i, -1.184, -0.352i, -0.762 - 0.232i, 0.498 + 0.418i, -0.1970 - 1.523i, \\ 1.147 + 0.030i, -0.194 + 0.864i, 0.501 - 0.047i, 0.071 + 1.441i)^T$$

where values were rounded to three digits past the decimal point. With these eigenvectors we compute a numerical estimate for the quantity (5.12) to get $b_2 \approx -20708 < 0$ (rounded to the nearest integer), which means the periodic orbits should appear supercritically, and will be stable at bifurcation

according to the second part of Theorem 5.3.1.

Figure 5.6(c) and (d) shows simulated trajectories of oscillations on which the system settles from randomized initial conditions at an attention value slightly above the bifurcation point, $u = 1.4$. The same trajectories are shown in parts (c) and (d) of the figure, grouped by beliefs about each option in part (c) and by each agent's internal beliefs in part (d). In contrast to the first example we considered, we observe that the oscillations in beliefs of different agents about each option are either exactly in phase or exactly out of phase. This happens because \mathbf{v}_a is real, and $\arg((\mathbf{v}_a)_i(\mathbf{v}_a)_j)$ is the same across all j modulo an offset by π . Then if $\text{sign}((\mathbf{v}_a)_i) = \text{sign}((\mathbf{v}_a)_k)$, agents i and k agree on all options, i.e. their beliefs oscillate in phase. Analogously, if $\text{sign}((\mathbf{v}_a)_i) \neq \text{sign}((\mathbf{v}_a)_k)$ then agents i and k disagree on all options, i.e. their beliefs oscillate out of phase. However, all of the trajectories in the agents' internal dynamics are out of phase and the ordering of relative preference towards options changes over time.

Analogous conclusions would hold if \mathcal{G}_a and \mathcal{G}_o swapped roles, and we instead considered a network of nine agents evaluating four options. In this scenario, the internal belief oscillations of each agent would synchronize, meaning each agent would be rigid in its order of preference towards options; however, the ordering of preferences of various agents along each option would vary over time, as all of the trajectories would be out of phase.

To conclude our discussion of oscillations of networked beliefs, we establish that oscillations cannot happen as a primary bifurcation of the origin in (5.1) if the communication and belief system graphs are both undirected, i.e. if the adjacency matrices A_a and A_o are both symmetric.

Proposition 5.3.3 (No primary Hopf bifurcations on undirected graphs). *If communication graph \mathcal{G}_a and belief system graph \mathcal{G}_o are both undirected, the dynamics (5.1) cannot support the onset of oscillations in a local bifurcation of the indecision equilibrium $\mathbf{Z} = \mathbf{0}$.*

Proof. For undirected $\mathcal{G}_a, \mathcal{G}_o$, the adjacency matrices A_a and A_o are symmetric. This means that all eigenvalues $\lambda \in \sigma(A_a)$, $\mu \in \sigma(A_o)$ are real, and as a consequence all of eigenvalues of the Jacobian (11.4) are $\eta = -d + u(\alpha + \gamma\lambda + \beta\mu + \delta\lambda\mu) \in \mathbb{R}$. It is thereby impossible for the leading eigenvalues of (11.4) to be a complex-conjugate pair, which violates a necessary condition of Theorem 5.3.1. \square

5.4 Mode interaction example: invariant 2-torus

So far we studied in detail how equilibria and sustained oscillations appear in a network of agents forming beliefs according to (5.1). These are the most common and general expected outcomes of

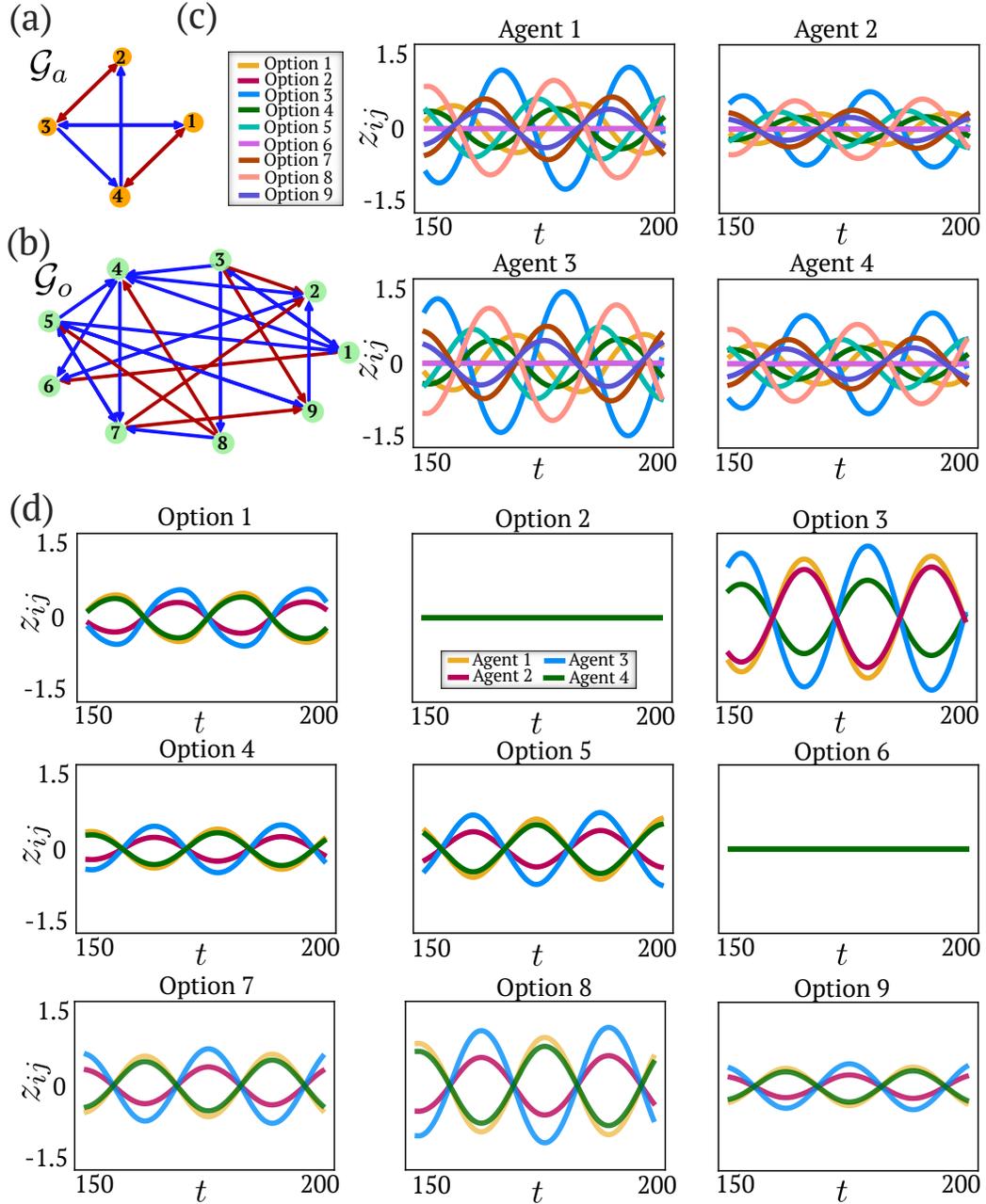


Figure 5.6: Trajectories $z_{ij}(t)$ of stable oscillation described in Example 5.3.2(b), $d = 1$, $u = 1.4$, $\alpha = \gamma = \beta = \delta = 0.1$, $\varepsilon_1 = \varepsilon_2 = 0$. (a) and (b): communication and belief system graphs with red edges representing negative connections and blue edges representing positive connections; (c) belief trajectories grouped by option; (d) belief trajectories grouped by agent.

belief formation. Here we do not prove the genericity of these bifurcations, but rather state it as a conjecture. However, genericity of pitchfork and transcritical bifurcations on networked dynamical systems has been established for some classes of graphs, for example for networks with a regular in-degree [80]. Although these belief-forming outcomes are generic, they are not all-exhaustive. For example when the communication graph \mathcal{G}_a and/or the belief system graph \mathcal{G}_o have some symmetry, the set of leading eigenvalues of the Jacobian \mathcal{H} can be made up of multiple real eigenvalues, multiple complex conjugate pairs of eigenvalues, or combinations of both. As discussed in Chapter 4 the dynamics of belief formation at bifurcation can become much more complicated in such scenarios due to mode interaction between the eigenspaces of the simultaneously crossing eigenvalues, which can give rise to much more complex dynamics.

To illustrate the possible complexity of the belief formation dynamics, we present a numerical example. Consider four unbiased agents forming beliefs on four options according to (5.1) with model parameters $d = 1$, $\alpha = \gamma = \beta = \delta = 0.1$. Let the communication graph \mathcal{G}_a and the belief system graph \mathcal{G}_o be defined by their adjacency matrices

$$A_a = \begin{pmatrix} 0 & -1 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix}, \quad A_o = \begin{pmatrix} 0 & -1 & 1 & -1 \\ -1 & 0 & -1 & 1 \\ -1 & -1 & 0 & 0 \\ -1 & 1 & 1 & 0 \end{pmatrix}.$$

Network diagrams of \mathcal{G}_a and \mathcal{G}_o are shown in Figure 5.7(a) and (b). The eigenvalues of A_o are $\mu_1 = 2$, $\mu_2 = 0$, $\mu_3 = \mu_4 = -1$. At the same time, all of the eigenvalues of A_a lie on the imaginary axis with $\lambda_{1,2} \approx \pm 1.618i$ and $\lambda_{3,4} \approx \pm 0.618i$. Then there are four eigenvalues of the Jacobian (5.3) which cross the imaginary axis simultaneously at the bifurcation point $u^* = d/(\alpha + \beta\mu_1) \approx 3.33$, which are two distinct complex-conjugate pairs. In Figure 5.7(c) we show simulated trajectories of the beliefs of the agents over time at a value of attention u slightly above the bifurcation point, $u = 3.34$. Agents' belief trajectories may seem random and irregular at first glance, however when they are visualized in phase space it becomes immediately clear that these trajectories are in fact ergodic and settle to an invariant torus in phase space - see Figure 5.7(d). This is an expected outcome of such bifurcation, as mode interaction of two complex conjugate pairs of eigenvalues typically results in the bifurcation of an invariant 2-torus in phase space [84]. This means that network beliefs settle on a low-dimensional manifold, despite their apparent irregularity. Furthermore since the torus bifurcation arises from an interaction of a simple eigenvalue of A_o and two complex conjugate

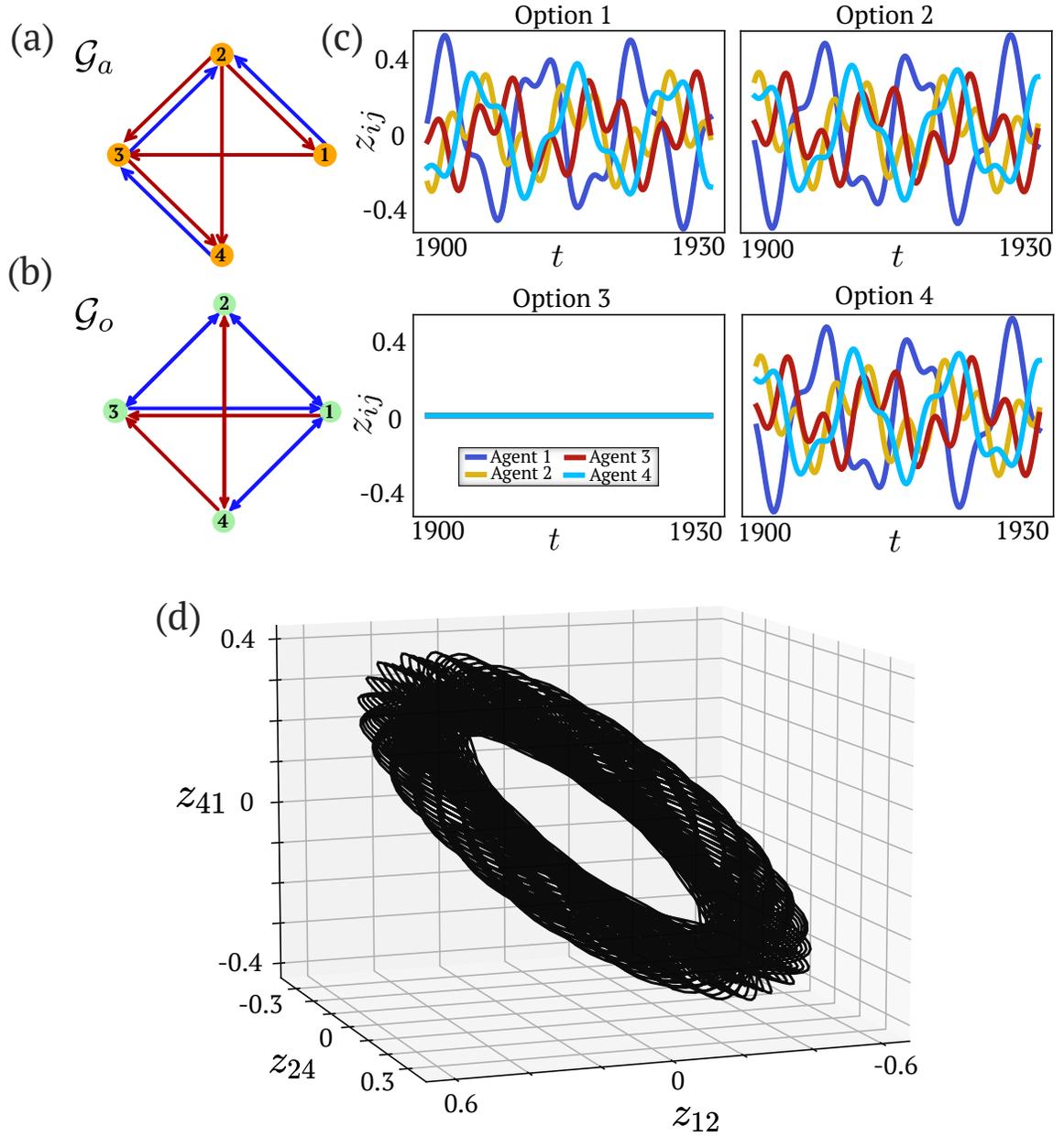


Figure 5.7: (a) and (b): communication and belief system graphs with red edges representing negative connections and blue edges representing positive connections; (c) belief trajectories grouped by option; (d) belief trajectories in phase space along three representative coordinates. Parameters: $d = 1$, $u = 3.34$, $\alpha = \gamma = \beta = \delta = 0.1$.

eigenvalue pairs of A_a , the internal dynamics of each agent are predictably structured. The right eigenvector of A_o corresponding to μ_1 is $\mathbf{v}_1 = (-1, 1, 0, 1)$. This tells us that all agents form strong beliefs on options 1, 2, and 4, and remain close to neutral on option 3 which is consistent with the simulation in Figure 5.7. Furthermore, we see that beliefs on options 2 and 4 are synchronized within each agent's belief trajectory, and beliefs on option 1 are anti-synchronized with those on options 2 and 4.

5.5 Consensus, dissensus, and multistability on symmetric belief systems

When a group evaluates options that are highly interchangeable, the adjacency matrix A_o of the belief system graph reflects a high degree of symmetry. When A_o has a transitive symmetry or most generally a homogeneously signed all-to-all connection topology, belief formation according to (3.8) and (3.10) has particularly interesting properties. For example, consider the case of simplex-constrained opinion dynamics (3.10) with an all-positive undirected communication matrix A_a with $A_o = \frac{1}{N_o} \mathbf{1}_{N_o} \mathbf{1}_{N_o}^T - \mathcal{I}_{N_o}$. The opinions of the group in this case will form in a bifurcation along either the *multi-option consensus space* defined as

$$W_c = \{(\tilde{\mathbf{Z}}_1, \dots, \tilde{\mathbf{Z}}_{N_a}) \mid \tilde{\mathbf{Z}}_i = \tilde{\mathbf{Z}}_k, \forall i, k\}, \quad (5.14)$$

or the *multi-option dissensus space* defined as

$$W_d = \{(\tilde{\mathbf{Z}}_1, \dots, \tilde{\mathbf{Z}}_{N_a}) \mid \tilde{\mathbf{Z}}_1 + \dots + \tilde{\mathbf{Z}}_{N_a} = 0\}. \quad (5.15)$$

On the consensus space W_c , agents have identical opinions. On the dissensus space W_d , agent opinions are balanced over the options such that the average opinion of the group is neutral. Consensus bifurcations happen when social imitation dominates, $\gamma - \delta > 0$, and dissensus bifurcations happen when ideology dominates, $\gamma - \delta < 0$. For example trajectories of group opinions settling on consensus and dissensus equilibria, see Figure 5.8. A similar splitting happens for unconstrained belief dynamics (3.8), with bifurcations expected along one of four distinct subspaces that correspond to full synchrony ($z_{ij} = z_{kl}$ for all i, j, k, l), consensus ($\mathbf{Z}_i = \mathbf{Z}_k$ and $\sum_{j=1}^{N_o} z_{ij} = 0$), deadlock ($z_{ij} = z_{ik}$ and $\sum_{i=1}^{N_a} z_{ij} = 0$ for all i, j, k), and dissensus ($\sum_{j=1}^{N_o} z_{ij} = 0$ and $\sum_{i=1}^{N_a} z_{il} = 0$ for all i, j, k, l). These ideas are developed thoroughly by Alessio Franci and collaborators in [68] and [69], and we mention

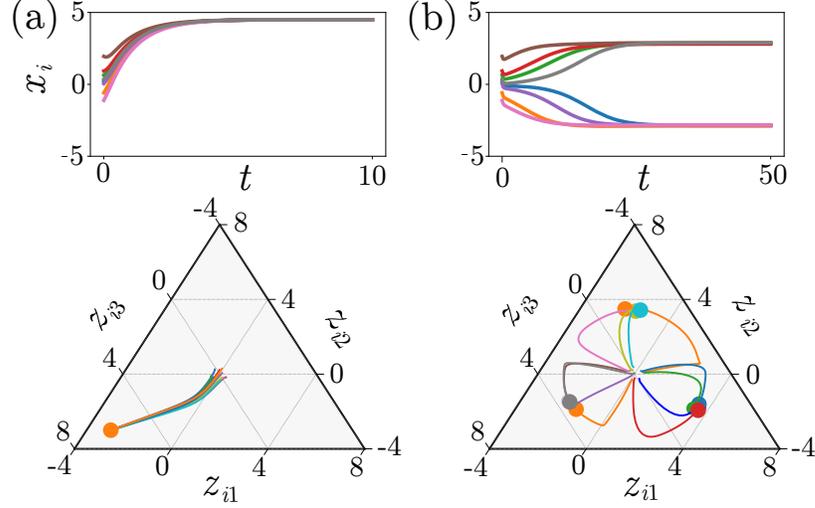


Figure 5.8: [19] Simulations of (3.10) for $N_o = 2$ options and $N_a = 8$ agents (top) and $N_o = 3$ options and $N_a = 12$ agents (bottom) with $A = \mathbf{11}^T - \mathcal{I}$. Opinions form (a) consensus when agents are cooperative: $\gamma = 0.2$, $\delta = -0.1$; (b) dissensus when agents are competitive: $\gamma = -0.1$, $\delta = 0.2$. In each plot, $\alpha = 0.2$, $\beta = 0.1$, $d = 1$, $u = 3$, $\hat{b} = 0$, and random initial conditions are the same. Communication weights $\alpha, \beta, \gamma, \delta$ were perturbed with small random additive perturbations drawn from a normal distribution with variance (a) 0.01, (b) 0.001 to illustrate the robustness of these trajectories. Ternary plots for three options generated with the help of [54].

them here for completeness of discussion. It is also worthy of note that in the presence of symmetry, bifurcations of equilibria will exhibit a high degree of multistability with branches of equilibria that emerge being related to one another by the symmetries of the network as a consequence of the Equivariant Branching Lemma [83].

These ideas are particularly important because analysis of the behavior of this model with maximal symmetry informs its possible behaviors when symmetry is broken through parameter heterogeneity or network structure. For a parametrized dynamical system, its instance with maximal symmetry in the equations often serves as an *organizing center* for the broader dynamics [82]. This means that many of the possible bifurcation diagrams of the heterogeneous model are recovered in a universal unfolding of the maximally symmetric case. In other words, predictions made for a symmetric system are remarkably robust and constrain the behavior of the dynamical system with broken symmetry. The analysis of the maximally symmetric system in [68] and [69] predicts the qualitative features of the solutions supported by the model (3.8). The analysis presented in this dissertation is complementary to this work, and can be understood as analysis of some of the possible unfoldings of the maximally symmetric case.

Chapter 6

Dynamic parameters

So far in this dissertation, in Chapter 3 we formulated a new model for distributed formation of beliefs on multiple options on a social network and showed that beliefs form through a bifurcation in which a network state of indecision loses stability. In Chapter 4 we characterized patterns of scalar beliefs or opinions at equilibrium using spectral properties of the signed communication graph \mathcal{G}_a . In Chapter 5 we studied belief-forming bifurcations in multi-option networks and established how the communication graph \mathcal{G}_a and the belief system graph \mathcal{G}_o play a role in determining the dynamic outcome, along with the ideological and social gains $\alpha, \gamma, \beta, \delta$ in the homogeneous model (3.8). All of the analysis we presented so far assumes that the model parameters $d, u, \gamma, \alpha, \delta$ are homogeneous and static. In this chapter, we relax this assumption. First in section 6.1, we will show how allowing each agent to dynamically modulate their level of attention to social interactions can enable cascades of opinion formation on the network which are tunably sensitive to the information contained in the distributed inputs \mathbf{b} . Next, in section 6.2 we will show how allowing agents to dynamically change the sign of their local interactions allows for decentralized control of the opinion patterns on the network. Both of these features are especially relevant for the use of this modeling framework to coordinate behavior in distributed technological teams, such as task allocation in robotic swarms. The analysis summarized in this chapter is contained in various parts of [18–21] which are also contained as Chapters 9,10,11,and 12 in Part II of this dissertation. All of the figures in this chapter also originate from these works. The ideas in section 6.1 are originally inspired by the adaptive controller design for nonlinear consensus formation in [87, 88] and cascade dynamics studied in [226]. I presented parts of analysis in this chapter in the 2020 SIAM Conference on the Life Sciences; in the 2020 UCLA Institute for Pure and Applied Mathematics, Mathematical

Challenges and Opportunities for Autonomous Vehicles Program, Workshop IV: Social Dynamics beyond Vehicle Autonomy; in the 2021 SIAM Conference on Applications of Dynamical Systems; in the 2021 American Control Conference; and in the 2021 IEEE Conference on Decision and Control.

6.1 Opinion cascades: tunable sensitivity and robustness

When a nonlinear system operates in a parameter regime near its bifurcation point, it is highly sensitive to perturbations and external signals. This is evident in our discussion of unfolding theory in Chapter 2.3 where we saw that a perturbed bifurcation diagram for a pitchfork bifurcation deforms significantly near its bifurcation point, but still resembles the unperturbed system away from the bifurcation point. Perturbations in such systems can lead to sudden and drastic changes in behavior. When bifurcations are encountered in control engineering applications, a common design goal is therefore to move the system away from its bifurcation point in order to improve its robustness to disturbance [35]. We instead argue that heightened sensitivity near bifurcation points can be a desired property for some systems, and we can take advantage of it through incorporating its features in controller design.

To illustrate the advantage of sensitivity, consider a robotic team operating in a safety-critical environment. In such environments individual robots may come across information that warrants an urgent response from the entire team - for example a fire that requires an evacuation, an individual in need of rescue, or an action cue from a human collaborator. For such purposes it is imperative to design distributed teams that are able to quickly and reliably respond to locally available information by committing to a decision or a task. A desired network response may be a consensus response in which all agents commit to the same choice of option or a disagreement response in which individuals allocate themselves across available tasks, depending on the specific context. In this section we illustrate how the model of opinion and belief formation (3.8) proposed and analyzed in this dissertation very naturally lends itself to the design of distributed control laws which enable teams to be highly sensitive and responsive to information, allowing for the network to reach both agreement and disagreement in response to a local cue. We develop and illustrate these ideas for decision-making on two options; for an extension to multiple options we refer the reader to the analysis in the work [67] on which I am a co-author.

Recall from our analysis in Chapter 4 that on many signed communication graphs \mathcal{G}_a , opinion formation on two options exhibits a pitchfork bifurcation. Let A_a be the signed adjacency matrix of \mathcal{G}_a with a simple leading eigenvalue $\lambda > 0$ and associated right and left eigenvectors \mathbf{v} , \mathbf{w} satisfying

$\langle \mathbf{w}, \mathbf{v} \rangle > 0$. Then the entries of \mathbf{v} inform the pattern of opinions which form at equilibrium in the pitchfork bifurcation, and the quantity $\langle \mathbf{w}, \mathbf{b} \rangle$ determines the primary branch of the unfolding, where $\mathbf{b} = (b_1, \dots, b_{N_a})$ is a vector of biases of individual agents about the two options. The biases of individuals represent distributed information about the two options on the network and they can be pre-determined or be obtained through sensor measurement or control input. Near its bifurcation point of network attention $u = u^*$, a group of agents forming opinions according to (4.1) is highly sensitive to the information contained in \mathbf{b} and commits to a configuration of opinions determined by the sign distribution of \mathbf{v} when $\langle \mathbf{w}, \mathbf{b} \rangle > 0$ and $-\mathbf{v}$ when $\langle \mathbf{w}, \mathbf{b} \rangle < 0$.

In order to take advantage of this high sensitivity to distributed information in the opinion formation process, we seek to design a dynamic feedback law that allows the level of attention on the network to increase over time from some pre-bifurcation value $u < u^*$ to some post-bifurcation value $u > u^*$. This transition will enable the system to pass through a highly sensitive region of the pitchfork bifurcation diagram and become informed by the distributed biases. In order for this process to be distributed, we relax the homogeneous attention assumption in (4.1) and define $u_i \geq 0$ to be the amount of attention agent i allocates towards the information from its social network. Furthermore, we allow this attention to be a dynamic state which evolves over time in response to dynamic feedback of opinion states of its neighbors. With these extensions, each agent is equipped with an opinion state x_i and an attention state u_i which evolve according to the coupled dynamics

$$\dot{x}_i = -dx_i + u_i S \left(\alpha x_i + \gamma \sum_{k \neq i}^{N_a} (A_a)_{ik} x_k \right) + b_i \quad (6.1a)$$

$$\tau_u \dot{u}_i = -u_i + S_u \left(x_i^2 + \sum_{k=1}^{N_a} \left((A_a)_{ik} x_k \right)^2 \right) \quad (6.1b)$$

where parameters d, γ, b_i are defined as before, S_u is a function that saturates inputs between a lower bound \underline{u} and an upper bound \bar{u} such that $\bar{u} > u^* \geq \underline{u} > 0$, and $\tau_u > 0$ is a time scale. The intuition behind the dynamic feedback law for the agents' attention (6.1b) is as follows. The function S_u saturates the squared magnitude of the opinions which agent i observes from its social network. It becomes more urgent for agent i to commit to an option if many of its neighbors are forming strong opinions for either option, and it ramps up its attention accordingly. We can think of (6.1b) as a form of social pressure between agents.

Before presenting our main result we examine a motivating simulation in Figure 6.1. In this figure, we show two simulations of the dynamics (6.1) on a balanced tree graph with all-negative interactions between agents. In both simulations, all agents on the network are unbiased except for

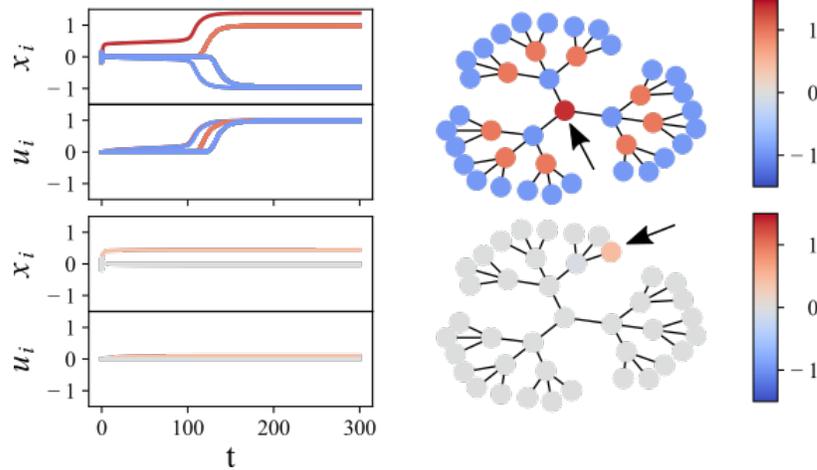


Figure 6.1: [20] Triggering a cascade on the balanced tree. Input $b_i = 0.4$ for agent marked with arrow; $b_i = 0$ for all other agents. Simulations of (6.1) start from small random initial conditions. Left: time trajectories of the dynamics. Right: final opinion, indicated by color, of simulation at $t = 300$. Parameters: $d = 1$, $\alpha = 1$, $\gamma = -1$, $u_i(0) = 0$ for all $i \in \mathcal{V}$; $S_u(y)$ is a rescaled logistic function.

the agent that is pointed out by a black arrow. In both simulations, a bias of equal magnitude is provided to the selected agent with all other parameters equal and the agents' initial opinions and attention are close to zero. On the left we show opinion and attention trajectories of the agents over time for both simulations, and on the right we show the final configuration of opinions on the network represented by color at $t = 300$ of the simulation. In the top simulation, the input is provided to a more central agent and we observe a network transition in which all agents take on strong opinions. In the bottom simulation, the input is provided to a less central agent and the network opinions remain close to neutral, with the attention level across all agents remaining low. Motivated by these observations, we define an *opinion cascade* as a network transition from a weakly to a strongly attentive state, where in a *weakly (strongly) attentive* state, the agents' attention is close to its lower (upper) saturation bound, i.e. $u \cong \underline{u}$ ($u \cong \bar{u}$). With this definition, in the top simulation of Figure 6.1 the network exhibits an opinion cascade, and the bottom simulation does not. In the rest of the simulations and in the stated theorem we utilize a modified Hill activation function:

$$S_u(y) = \underline{u} + (\bar{u} - \underline{u}) \frac{y^n}{(y_{th})^n + y^n}, \quad (6.2)$$

where threshold parameter $y_{th} > 0$ tunes the midpoint input value of the saturation. Similar analysis would hold with different forms of saturation function. In the following theorem we establish the properties of input-triggered opinion cascades.

Theorem 6.1.1 (Inputs to the network trigger opinion cascades). [19, 21] Consider the coupled

system (6.1) on a communication graph \mathcal{G}_a whose adjacency matrix has a simple leading eigenvalue $\lambda > 0$ with corresponding right and left eigenvectors \mathbf{v}, \mathbf{w} satisfying $\langle \mathbf{w}, \mathbf{v} \rangle > 0$. Let $u^* = \frac{d}{\alpha + \lambda \gamma}$. There exists $\varepsilon > 0$ such that for $u^* > \underline{u}$, $y_{th} < \varepsilon$, and n sufficiently large, the following generically hold:

1) There exists $p = p(y_{th}) > 0$ satisfying $\frac{\partial p}{\partial y_{th}} > 0$ such that, for $|\langle \mathbf{w}, \mathbf{b} \rangle| < p$, model (6.1) possesses a weakly attentive locally exponentially stable equilibrium;

2) The weakly attentive equilibrium loses stability in a saddle-node bifurcation for $|\langle \mathbf{w}, \mathbf{b} \rangle| = p$. No weakly attentive equilibria exist for $|\langle \mathbf{w}, \mathbf{b} \rangle| > p$ and all trajectories converge to a strongly attentive equilibrium $(\mathbf{x}^*, \mathbf{u}^*)$. At this equilibrium if $\langle \mathbf{w}, \mathbf{b} \rangle > 0 (< 0)$ then $\text{sign}(x_i) = \text{sign}(v_i) (-\text{sign}(v_i))$;

3) For $\gamma = 0$, with $\alpha > 0$, the strongly opinionated equilibrium $(\mathbf{x}^*, \mathbf{u}^*)$ satisfies $\text{sign}(x_i^*) = \text{sign}(b_i)$.¹

Theorem 6.1.1 explains why the same input triggered an opinion cascade on the graph shown in Figure 6.1 when it was provided to the innermost agent, and did not trigger a cascade when it was provided to the outermost agent. Cascades are triggered when the projection of the network biases onto the left leading eigenvector of the dsjacency matrix A_o is sufficiently large. The magnitude of the entries in \mathbf{w} therefore determine the influence, or centrality, of agents on the network in their ability to trigger network cascades. For the graph in Figure 6.1, the innermost agent is the most central and the outermost agent is one of the least central, according to \mathbf{w} . For the selected parameters in these simulations providing an input to the inner agent pushed the quantity $\langle \mathbf{w}, \mathbf{b} \rangle$ over its cascade threshold, and providing it to the outermost agent did not.

We further illustrate the properties of opinion cascades with three representative simulations on undirected graphs with the same topology in Figure 6.2. In these simulations the five agents all have initial small biases which do not trigger a cascade. At $t = 20$, the agent pointed out by the arrow increases its bias and a network cascade is triggered. This figure illustrates how different opinion-forming behaviors can be triggered on a network with the same biases and interconnection topology. In part (a), all of the edges in the graph \mathcal{G}_a are positive and the cascade results in an agreement decision on the network. In part (b) all of the edges are negative and the network distributes itself across the two options at disagreement. In part (c), $\gamma = 0$ which means the dynamics of opinions of the agents are decoupled in (6.1). However, the agents are still coupled through their attention dynamics, and a cascade is triggered. The configuration of opinions on the network post-cascade is not determined by an eigenvector of the adjacency matrix A_a ; instead the agents strongly commit

¹The statement of the theorem here has been modified to describe a broader class of signed graphs than in its original presentation; a proof of this version of the theorem is a minor modification of the proof in the cited works.

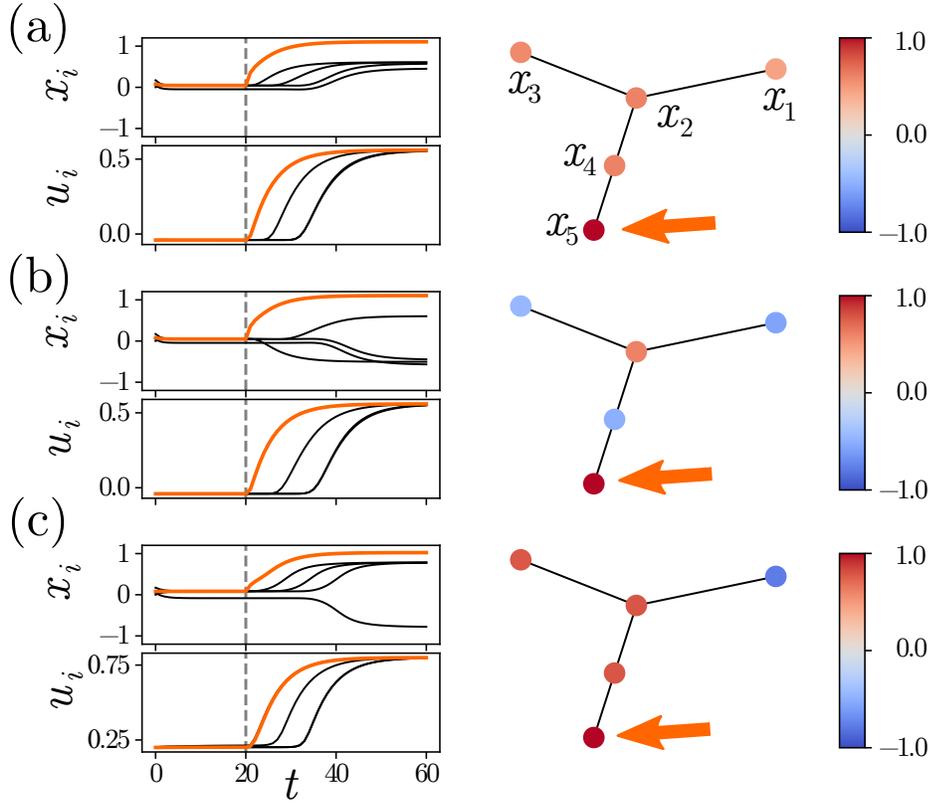


Figure 6.2: [19] Opinion cascades with opinion and attention dynamics defined in Theorem 9.5.1. For $t < 20$, $\mathbf{b} = (-0.05, 0.05, 0.05, 0.05, 0.05)$ for all three simulations. At $t = 20$ the input to agent 5 (indicated by the arrow) increases to $b_5 = 0.25$, which triggers an opinion cascade on the network. Plots show opinion and attention trajectories of the agents with agent 5 in orange. Network diagrams on the right show the opinion strength of each agent at $t = 60$ of the simulation. (a) Agreement cascade with $\gamma = 1$; the network chooses the positive option following the informed agent. (b) Disagreement cascade with $\gamma = -1$; agents' opinions on the network disperse following the sign structure of \mathbf{v}_{min} . (c) Agents are coupled through the attention dynamics only (i.e. $\gamma = 0$); opinion cascade causes each agent to amplify its small input and commit to a strong opinion. Other parameters: $\alpha = 2$, $n = 3$, $y_{th} = 0.1$, $\tau_u = 5$, $d = 1$, $\bar{u} = u^* + 0.3$, $\underline{u} = u^* - 0.3$, $u_i(0) = \underline{u}$ for all $i = 1, \dots, N_a$. $\mathbf{x}(0)$ generated randomly from a uniform distribution between -0.2 and 0.2 ; the same initial condition was used for all three simulations.

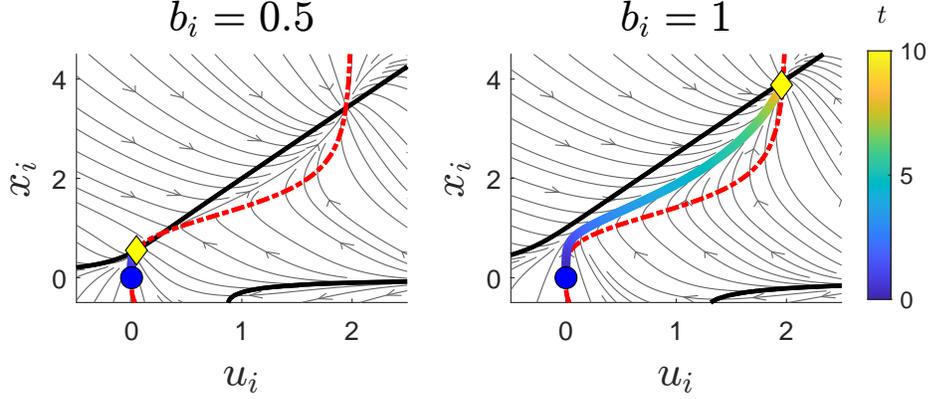


Figure 6.3: [19] Sensitivity of opinion formation to input magnitude. (u_i, x_i) -phase plane and trajectories of (6.1) for a single agent. Initial state $(u_i(0), x_i(0)) = (0, 0)$ is a blue circle, and final state a yellow diamond. Nullclines of (6.1a) are black solid and (6.1b) are red dashed. Gray arrows show flow streamlines. Color scale is time.

to the option which is favored by their small biases b_i .

Next, we emphasize that the parameters of the saturation function (6.2) allow a designer to tune two important aspects of the opinion cascade: the agents' sensitivity, i.e. how strong of a network input is required to trigger the onset of a cascade; and the agents' robustness, i.e. how strongly committed the agents are to their choice of option post-cascade. Specifically, the parameter y_{th} regulates the sensitivity of the network. If y_{th} is small, the saturation midpoint is small, and a small bias can trigger a network cascade. Increasing y_{th} increases the range of magnitudes of network biases which is ignored by the agents. This means that agents can be made sensitive to arbitrarily small inputs when necessary, but they can also be designed to reject small inputs as disturbance when necessary - for example when there is a high level of sensor noise. On the other hand, the upper bound \bar{u} of the saturation tunes the robustness of the system. The larger \bar{u} , the further away the final operating point of the system lies from the bifurcation point u^* of the underlying opinion dynamics model. This means that even if the inputs b_i change, the agents will remain committed to their opinion configuration.

For a simple illustration of these ideas we consider a single agent with a nonzero self-reinforcement $\alpha > 0$ whose opinion and attention dynamics evolve according to (6.1). Although there is no social network, due to the self-reinforcement mechanism built into both the opinion and the attention dynamics the agent's opinion dynamics reproduce the basic properties underlying the network opinion cascades. An added bonus of examining this simple system lies in the fact that its equilibria can be visualized as intersections of nullclines of the opinion dynamics and attention dynamics in the $x - u$ plane - see Figures 6.3 and 6.4.

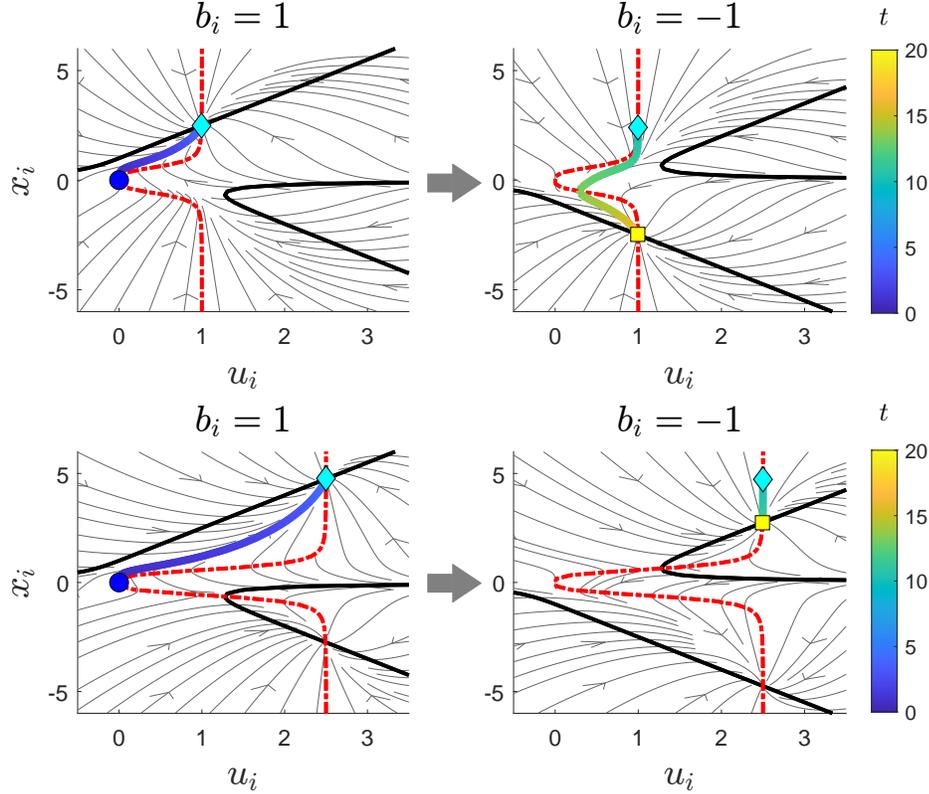


Figure 6.4: [19] Robustness of opinion formation to changes in input. (u_i, x_i) -phase plane and trajectories of (6.1); (Left) Input is $b_i = 1$, initial state $(u_i(0), x_i(0)) = (0, 0)$ is a blue circle, and final state is a cyan diamond. (Right) Input changes to $b_i = -1$, initial state is final state on left and final state is yellow square. Top: $\bar{u} = 1$, and agent changes opinion in direction of new input. Bottom: $\bar{u} = 2.5$, and agent retains opinion in original direction. Nullclines, streamlines, and time are drawn as in Figure 6.3.

First, consider Figure 6.3. This figure shows the same phenomenon as was shown for the network in Figure 6.1. When the agent receives a small input b_i , its attention nullcline (red) and the primary branch in the bifurcation unfolding diagram captured by its opinion nullcline (black) intersect in three places, including at a value of u_i close to zero. The agent remains weakly attentive and does not commit to a strong opinion. However as the input is increased, the continuous branch moves up and eventually the intersections near $u_i = 0$ disappear. This forces the agent's attention to increase and it commits to a strong opinion in the direction of the input. A similar mechanism in higher dimension governs the network cascades such as the one shown in Figure 6.1 - the nullcline of the attention dynamics on a network defines a bent hypersurface which intersects with a pitchfork bifurcation nullcline of the opinion dynamics. If the saturation function S_u chosen to generate Figure 6.3 was chosen with a higher parameter y_{th} , the red curve in the picture would move up and it would take a bigger input to remove the nullcline intersections near $u_i = 0$.

Next, consider Figure 6.4 which shows two simulations. In both simulations, for the first time

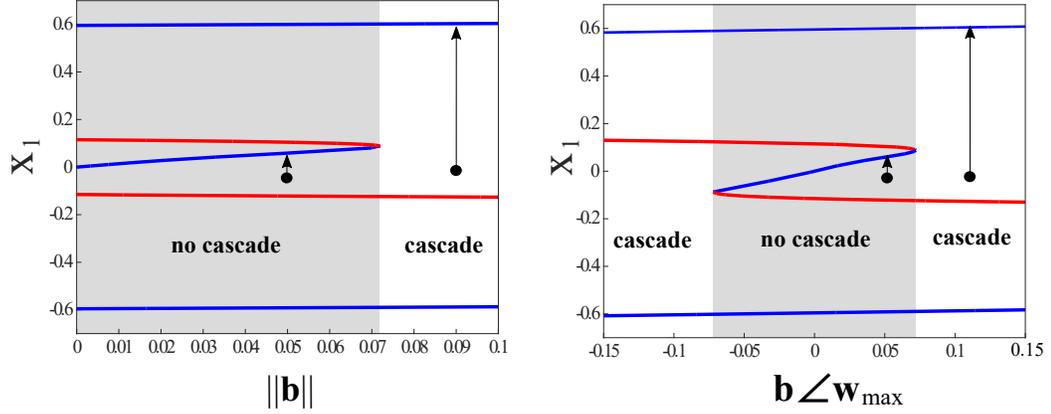


Figure 6.5: [21] Blue (red) lines track the first coordinate of the stable (unstable) equilibrium solutions of the coupled dynamics (11.9) on a 3-agent undirected line graph. Parameters: $u_{th} = 0.1$, $\gamma = 1$, $n = 3$, $\mathbf{b} = \|\mathbf{b}\| \cdot |\mathbf{b} \angle \mathbf{w}_{max}| \cdot \mathbf{w}_{max}$; left: $|\mathbf{b} \angle \mathbf{w}_{max}| = 0.1$, right: $\|\mathbf{b}\| = 0.1$ Bifurcation diagrams generated using MatCont [47].

steps the agent receives an input $b_i = 1$ and at $t = 10$ the input instantaneously changes sign. In the top simulation, $\bar{u} = 1$ and when the input switches, the agent’s opinion also switches. In the bottom simulation, $\bar{u} = 2.5$ and in the first ten time steps, it settles at a point further along the primary branch of the pitchfork unfolding in the opinion nullcline. It is therefore more robust, and its opinion does not switch when the input changes sign (although it lowers in magnitude). This illustrates on a low dimensional example the sensitivity-robustness tradeoff that’s built into the opinio cascade dynamics (6.1) - a similar (but trickier to visualize) mechanism happens for the networked dynamics.

Finally, let $\|\mathbf{b}\|$ designate the magnitude of the input or bias vector, and let $\mathbf{b} \angle \mathbf{w} := \langle \mathbf{w}, \mathbf{b} \rangle / \|\mathbf{b}\|$ its orientation relative to the centrality eigenvector \mathbf{w} . Increasing $\|\mathbf{b}\|$ or increasing $\mathbf{b} \angle \mathbf{w}$ has an analogous effect in triggering network cascades, which we illustrate in Figures 6.5 and 6.6; both can be used as controls for triggering network cascades.

In Figure 6.5 we show a bifurcation diagram which illustrates the saddle-node bifurcation established in Theorem 6.1.1 on a small network of agents with an all-positive signature, for which the cascade centrality eigenvector is \mathbf{w}_{max} . When $\|\mathbf{b}\|$ and/or $\mathbf{b} \angle \mathbf{w}$ are small, there is a stable equilibrium that is linearly offset from zero, and a network starting from a small initial condition will be trapped at this equilibrium. Past the threshold determined by the saddle-node bifurcation, all initial conditions instead evolve to a branch with a strong final opinion. When the design parameter y_{th} is increased, the saddle-node bifurcation point will move further away from zero and the system response will reject inputs of greater magnitude. We further illustrate the existence of this cascade threshold by heatmaps in Figure 6.6 which were generated by running 1.5×10^5 simulations of (6.1) on the three pictured small graph architectures with an all-positive signature and cascade central-

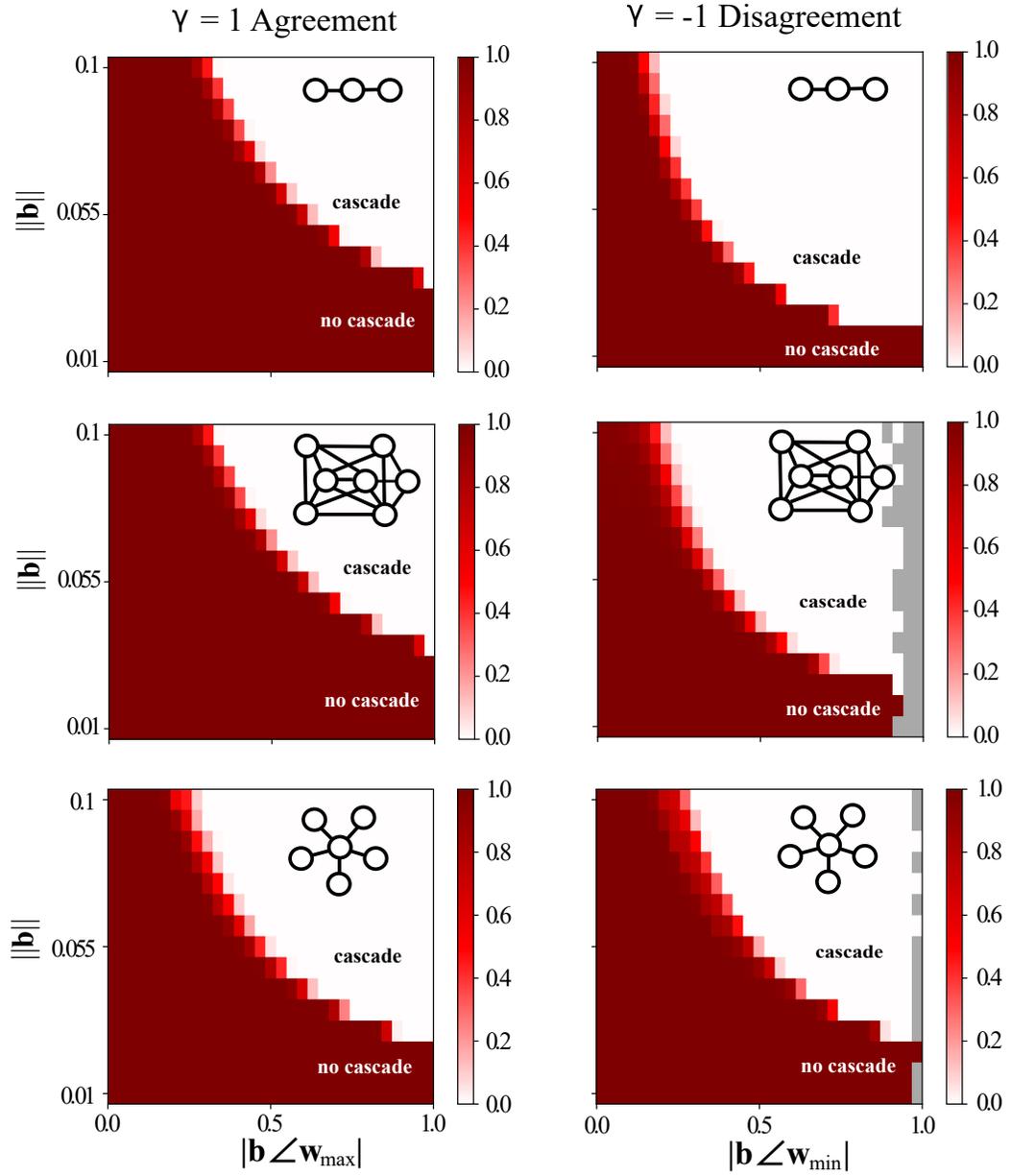


Figure 6.6: [21]Heatmaps with color corresponding to proportion of simulations in the given parameter range that did not result in a network cascade by $t = 500$. Dark red corresponds to no cascades, white to there always being a cascade. Grey squares are bins with no datapoints. Each plot corresponds to 1.5×10^5 distinct simulations on an undirected graph shown in the diagram. Simulation parameters: $\tau_u = 10$, $u_{th} = 0.2$, $\underline{u} = u_a - 0.01$ for $\gamma = 1$ (left plots) and $\underline{u} = u_d - 0.01$ for $\gamma = -1$ (right plots). For each simulation, inputs b_i were drawn from $\mathcal{N}(0, 1)$ and the input vector \mathbf{b} was normalized to a desired magnitude. There were 10000 simulations performed at each constant input magnitude, with 15 magnitudes sampled uniformly spaced between 0 and 0.1. The initial conditions for each simulation were $x_i = 0$, $u_i = 0$ for all $i = 1, \dots, N_a$.

ity eigenvector \mathbf{w}_{max} (left) and an all-negative signature and cascade centrality eigenvector \mathbf{w}_{min} (right). In these plots, the red region indicates combinations of input magnitude and orientation which did not trigger an opinion cascade, and the white region corresponds to a parameter regime in which cascades were always triggered. There is a clear and abrupt transition between these two dynamical regimes, which corresponds to the crossing of the saddle-node bifurcation threshold. If the design parameter y_{th} was increased, this threshold would also increase and red no-cascade region in these plots would expand outwards.

6.2 Dynamic switching

In the previous section we established a mechanism for a cascade of opinions to spread on a network in response to a local signal. This type of drastic transition in the entire network is one way in which a team of decentralized decision-makers can adapt to environmental cues, however in many situations it may be desirable for only some of the team members to alter their behavior when they encounter a relevant cue. For example if an autonomous robotic team is exploring an environment and some of its members run low on battery, these individuals need to have the capability to switch away from their current behavior and dock to a charging station. However it is advantageous for the group that these individuals remain plugged in to the overall communication network in case of a safety-critical concern that requires their participation. Any framework for the design and coordination of robust, responsive, and adaptable autonomous teams must enable individual members to alter their behavior in response to local cues without affecting the behavior of the rest of the team. In this section we explore how such a local change in commitment to an option or task can be implemented dynamically and in a distributed manner for a group of decision-makers following the belief formation dynamics (3.8).

To illustrate this point we examine once again the dynamics of opinion formation on two options for a network of unbiased agents with communication graph \mathcal{G}_a

$$\dot{x}_i = -dx_i + uS \left(\alpha x_i + \gamma \sum_{\substack{k=1 \\ k \neq i}}^{N_a} (A_a)_{ik} x_k \right). \quad (6.3)$$

So far we have shown that whenever the signed adjacency matrix A_a has a simple positive leading eigenvalue, opinions form according to (6.3) through a pitchfork bifurcation. In section 6.1 we also showed that this pitchfork bifurcation informs the outcome of opinion cascade when agents' attention

is dynamic. Whether the agents reach their final opinion configuration solely through the opinion dynamics or through opinion cascades, we know that the sign pattern of opinions on the network follows the sign pattern of the right leading eigenvector of the adjacency matrix, \mathbf{v} . On the surface this outcome may appear rigid since the final distribution of opinions is effectively pre-determined by the architecture and signature of \mathcal{G}_a . However, recall from Chapter 4.3 that the bifurcation digrams of (6.3) on two switching-equivalent graphs are related in a straightforward manner. For example if agent i is switched to obtain graph \mathcal{G}'_a from \mathcal{G}_a (i.e. the signature of all the edges that point in and out of agent i on the graph is flipped), then the opinion of agent i relative to its neighbors at equilibrium on graph \mathcal{G}'_a is the opposite of what it would be on \mathcal{G}_a . We take advantage of this observation and show that such a switching transformation can be implemented locally on a network, and it will result in a predictable network transition in which the agents which are being switched flip their opinion, and the opinions of the rest of their neighbors are unaffected (excepting some minor transient dynamics).

Consider a structurally balanced network of agents whose attention is above the pitchfork bifurcation point, $u > u^* = d/(\alpha + \lambda\gamma)$, and whose opinion states have settled to an equilibrium. This system is operating at an one of the stable opinionated equilibria established by the pitchfork bifurcation described in Theorem 4.1.2. First, we conjecture in the following assumption that the stable manifold of the origin is bounded for these graphs when $u > u^*$. Let \mathbf{w}, \mathbf{v} be the left and right eigenvectors of A_a corresponding to the simple leading eigenvalue of the adjacency matrix A_a .²

Assumption 6.2.1 (Stable manifold of origin is bounded; Fig. 6.7). *[18] Consider (10.1) on some structurally balanced graph \mathcal{G}_a with $u > u^*$ and $u < u_2$ when appropriate, as defined in Theorem 4.1.2. Let $\mathcal{U}' \subset \mathbb{R}^N$ be an open neighborhood containing the origin, and let $\mathbf{x} \in W^s(\mathbf{0}) \cap \mathcal{U}'$. 1) $|\langle \mathbf{w}, \mathbf{x} \rangle| < \varepsilon \|\mathbf{x}\|^2$ for some $0 < \varepsilon < 1$; 2) for equilibria $\mathbf{x}_k^* \neq \mathbf{0}$ of Proposition 12.3.1 with $k \in \{1, 2\}$, $|\langle \mathbf{w}, \mathbf{x}_k^* \rangle| > \varepsilon \|\mathbf{x}_k^*\|^2$.*

The assumption 6.2.1 has been verified through simulation on various network structures, however a proof of this assumption would require computing an approximation of the stable manifold of the origin, allowing for general graph architectures and parameters. Such a calculation is at best cumbersome and at worst intractable, and we leave it as an assumption in present work. A geometric illustration is presented in Figure 6.7. In the following lemma we characterize the region of attraction of the equilibria $\mathbf{x}_1^*, \mathbf{x}_2^*$.

Lemma 6.2.2 (Regions of attraction). *[18] Consider (6.3) on some structurally balanced graph*

²As discussed perviously, a structurally balanced graph always has a simple leading eigenvalue.

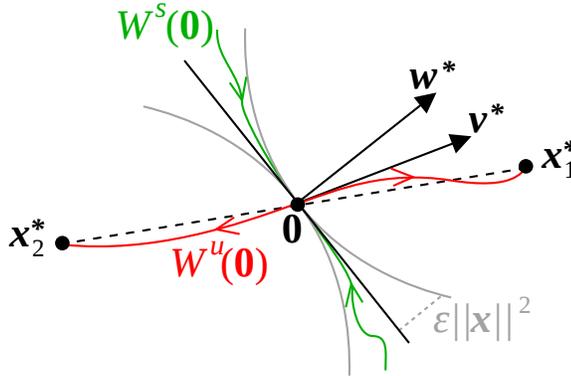


Figure 6.7: [18] Geometric intuition behind Assumption 12.5.2. The one-dimensional unstable manifold $W^u(\mathbf{0})$ of the origin (shown in red) forms heteroclinic orbits with the stable equilibria $\mathbf{x}_1^*, \mathbf{x}_2^*$, as is generically the case for monotone systems - see [187, Theorem 2.8].

\mathcal{G}_a with $u > u^*$ for all $i = 1, \dots, N$, on an open and bounded neighborhood Ω_r . Let $\mathbf{x}_1^*, \mathbf{x}_2^*$ be the nonzero equilibria on the pitchfork bifurcation of Theorem 4.1.2 with $\langle \mathbf{w}, \mathbf{x}_1^* \rangle > 0$. Consider an initial condition $\mathbf{x}(0)$ at $t = 0$. If $\langle \mathbf{w}, \mathbf{x}(0) \rangle > \varepsilon \|\mathbf{x}(0)\|^2 (< -\varepsilon \|\mathbf{x}(0)\|^2)$ then as $t \rightarrow \infty$, $\mathbf{x}(t) \rightarrow \mathbf{x}_1^* (\mathbf{x}_2^*)$.

Finally, we can establish our main result.

Theorem 6.2.3 (Instantaneous switching). [18] Consider (10.1) on some \mathcal{G}_a and let $\mathbf{x}_1^*, \mathbf{x}_2^*$ be the two nonzero equilibria on the pitchfork bifurcation of Theorem 4.1.2, with $\langle \mathbf{w}, \mathbf{x}_1^* \rangle > 0$. Let $\mathcal{G}_a^{\mathcal{W}}$ be switch equivalent to \mathcal{G} with the associated switching matrix Θ . Suppose at $t = 0$, $\mathbf{x}(0)$ is close to \mathbf{x}_i^* where $i = 1$ or 2 . If $|\langle \Theta \mathbf{w}, \mathbf{x}_i^* \rangle| > \varepsilon \|\mathbf{x}_i^*\|^2$ and $\langle \Theta \mathbf{w}, \mathbf{x}_i \rangle > 0 (< 0)$ then for (10.1) on $\mathcal{G}_a^{\mathcal{W}}$ as $t \rightarrow \infty$, $\mathbf{x}(t) \rightarrow \Theta \mathbf{x}_i$ ($\rightarrow -\Theta \mathbf{x}_i$).

The intuition behind Theorem 6.2.3 is the following. Instantaneously changing a structurally balanced graph \mathcal{G}_a to its switching equivalent $\mathcal{G}_a^{\mathcal{W}}$ results in a predictable transition of the system state. Namely, if the number of nodes in \mathcal{W} is small in comparison with the cardinality of \mathcal{V} , we expect that all nodes in \mathcal{W} will change sign, and all of the nodes in $\mathcal{V} \setminus \mathcal{W}$ will not. A simulation example of this behavior is shown in Figure 6.8. In this figure, opinions form on a graph that is initially all-positive. At the halfway point of the simulation, agent 1 on the graph is switched, which results in its opinion evolving away from the group, with all other opinions remaining unchanged. The precise number of nodes that can be switched simultaneously to generate this behavior depends on the eigenvector \mathbf{w}^* of the graph adjacency matrix, the value of the equilibrium \mathbf{x}_1^* , and the bound ε . In practice, it is often sufficient that $|\mathcal{W}| < \frac{1}{2}|\mathcal{V}|$.³

Recall that a switching of a node modifies only the sign of the pairwise interactions of that node with its neighbors. A switching transformation can therefore easily be implemented in a way that is

³This paragraph is modified directly from [18].

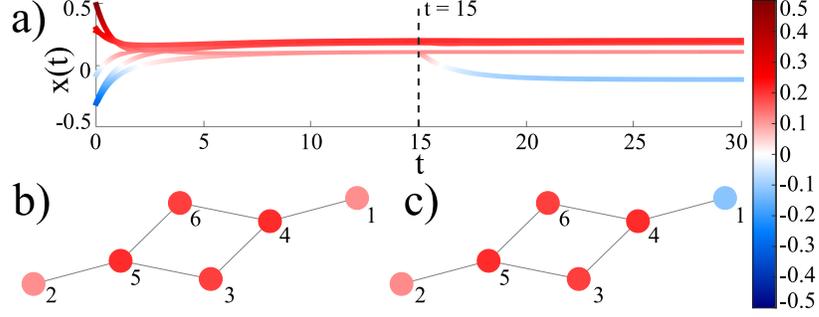


Figure 6.8: [18] Applying a switching transformation to agent 1 at $t = 15$. (a) Time trajectory of the opinion dynamics (b)-(c) Network diagram with the opinion of each agent at $t = 10$, $t = 30$. Parameters: $S = \tanh$, $d = 1$, $\alpha = 1.2$, $\gamma = 1.3$, $u = 0.294$.

agnostic to the global network topology. This means that a transition of network opinions of the type shown in Figure 6.8 can easily be triggered with a local cue. Concretely, let $\theta_i \in \{1, -1\}$ be an agent's switching state and suppose the agents share both their opinion x_i and their switching state θ_i with their neighbors on the communication graph \mathcal{G}_a . Furthermore let $(A_a)_{ik} = (A_a)_{ik}(0)\theta_i\theta_k$ where $(A_a)_{ik}(0) \in \{1, -1\}$ is determined by the initial graph signature. If agent i receives information about agent k on the communication graph \mathcal{G}_a , then it is aware of both switching states. If in response to some cue or control signal the switching state of agent i or of agent k changes, then the sign of $(A_a)_{ik}$ flips. This transition implements the instantaneous switching transformation we show in Figure 6.8.

Finally we show with a simulation example that a switching transformation can also be implemented through a dynamic law and still results in well-behaved predictable transitions in the opinion dynamics. In this example we allow each agent to update its edge weights with its neighbors through a dynamic feedback law which smooths out the previously instantaneous transition in the edge weights by integrating the quantity $(A_a)_{ik}(0)\theta_i\theta_k$:

$$\tau_a \frac{d}{dt} (A_a)_{ik} = -(A_a)_{ik} + (A_a)_{ik}(0)\theta_i\theta_k. \quad (6.4)$$

In Figure 6.9 we show a simulation of a sequence of switches facilitated by (6.4), implemented on a network which starts with an all-positive signature. Initially the agents converge to an agreement state. Then at $t = 10$ agents 1 and 2 change their switching state and evolve away from the group, followed by agent 3 at $t = 15$. Figure 6.9 illustrates how network opinion transitions with dynamic edge weights are well-behaved and similar to the transitions for an instantaneously switched graph. This observation motivates future explorations of various feedback laws through which a network can adapt and self-regulate its behavior via coupled dynamics of opinions and edge signatures.

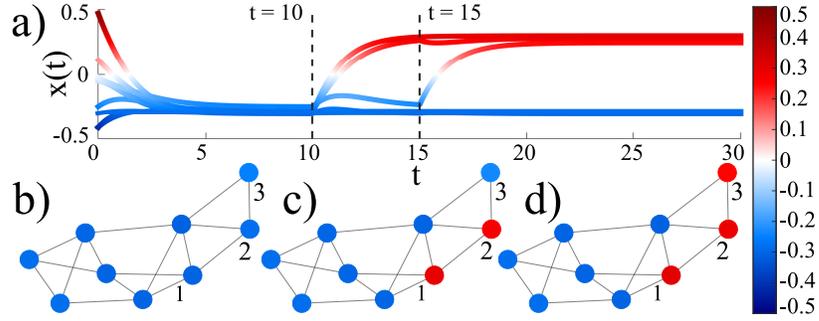


Figure 6.9: [18] Locally switching agents 1, 2 and 3 using (12.6). Agents 1, 2 switch at $t = 10$ and agent 3 at $t = 15$. (b)-(d) Network diagram with the opinion of each agent at $t = 10$, $t = 15$, $t = 30$. Parameters: $S = \tanh$, $d = 1$, $\alpha = 1.2$, $\gamma = 1.3$, $u = 0.315$, $\tau_a = 0.01$

In this chapter we established that a group of decentralized decision-makers can respond to local cues in a variety of meaningful ways. At one extreme, we have shown that a small localized signal picked up by a node in the network can be amplified, leading to a cascade of opinionation and a global change in the network state. On the other extreme, we have shown that individual nodes in the system can respond to local cues by changing their state without disrupting the behavior in the rest of the group. A key ingredient in the design of these flexible features of the decision process is the endowment of the parameters in the belief formation model (3.8) with their own dynamics. Although these ideas were explored separately, they can easily coexist in a system which makes the modeling framework developed in this dissertation an appealing tool for guiding future design of networked technologies. Together the ideas presented in this dissertation make up a new approach for the design of adaptable behaviors in a decentralized group of agents.

Chapter 7

Final remarks: applications and future extensions

7.1 Applications

The material presented in this dissertation constitutes primarily a theoretical contribution with the development and formal analysis of the model (3.8) and its extensions. However one of the driving motivations for developing this modeling framework is the broad range of potential applications of this analysis. These applications include using the model as a tool for generation of mechanistic hypotheses that provide insight into emergent phenomena in online social networks, sociopolitical systems, groups of social animals, and other systems driven by social interactions. On the other hand, the model can serve as a powerful tool that informs the design of distributed biologically inspired algorithms which enable technological teams such as robotic swarms to coordinate, make decisions, and allocate tasks dynamically in uncertain and time-varying environments. In this section we provide a brief summary of several applications of this modeling framework which are either actively under development or appear in print. I am a co-author on several of the discussed works which are cited where appropriate. However, none of the papers discussed in this section with the exception of [18] appear in Part II of this dissertation, as Part II only contains work for which I was the lead author.

7.1.1 Game theory and social dilemmas

The field of *game theory* develops mathematical models of decision-making outcomes of strategic interactions between individuals, typically referred to as players. Game theoretic models are widely utilized by economists as models of human behavior [73], and in engineering as tools for design [13]. These models are built with a fundamental underlying assumption that the players' decision-making strategies follow rational rules. Rationality in this context is captured in a system of *rewards* or *payoffs* which each individual seeks to maximize for themselves. Different games are classified by the number of players, the number of available strategies, and the structure of their payoffs for different strategy combinations. For example in a two-player two-strategy game the payoff structure is summarized in the following chart:

$$\begin{array}{c}
 \text{Player 2} \\
 \begin{array}{cc}
 & C & D \\
 \text{Player 1} & \begin{array}{|c|c|}
 \hline
 C & (P_{CC})_1, (P_{CC})_2 & (P_{CD})_1, (P_{DC})_2 \\
 \hline
 D & (P_{DC})_1, (P_{CD})_2 & (P_{DD})_1, (P_{DD})_2 \\
 \hline
 \end{array} &
 \end{array}
 \end{array} \tag{7.1}$$

In (7.1) there are two strategies, C and D (e.g. “cooperate” and “defect”). The entries in the chart represent the payoffs received when each player chooses the indicated strategy. When Player 1 chooses strategy C and player 2 chooses strategy D , Player 1 receives the payoff $(P_{CD})_1$ and Player 2 receives the payoff $(P_{DC})_2$. A *Nash equilibrium* in a game is a combination of strategies such that no single player can gain a greater payoff by changing its strategy, assuming all other players' strategies remain unchanged.

Social dilemmas are games whose Nash equilibria differ from the optimal configuration for the group [113]. A famous social dilemma is the two-player *prisoners' dilemma* in which the Nash equilibrium leads to both players defecting (choosing strategy D) whereas the best outcome for the group is both players cooperating (choosing strategy C). Standard game-theoretic models predict convergence of a prisoners' dilemma game to this mutual-defection Nash equilibrium. However experimental evidence supports emergence of cooperation in prisoners' dilemma games played by people in laboratory settings for some of the game trials [96, 132].

In political science and economics it has long been argued that cooperation emerges from *reciprocity* between individuals, i.e. players mimicking the observed strategies of their neighbors [10, 58, 86]. In the opinion dynamics model (3.8) reciprocity is modeled by the same-option interactions with gain γ whenever agents are cooperative, i.e. with $(A_a)_{ik} \in \{0, 1\}$. With this motivation

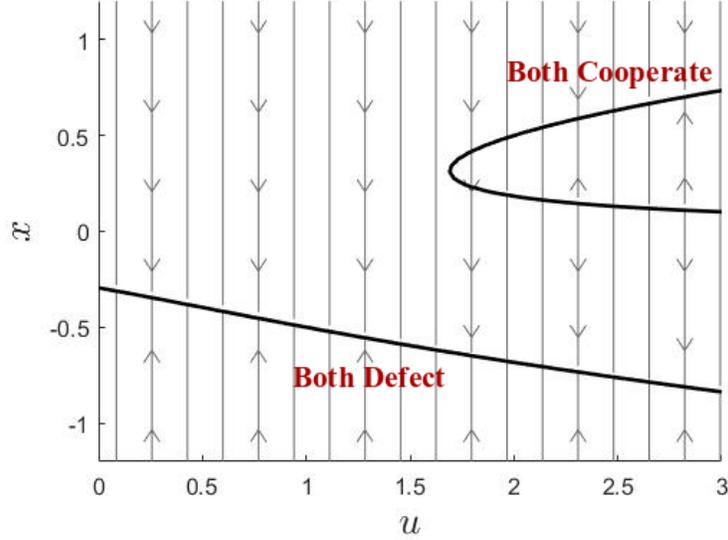


Figure 7.1: Pitchfork bifurcation unfolding explains emergence of cooperation in a two-player prisoners dilemma game with social reciprocity ($\gamma > 0$); $x = \frac{1}{2}(z_{11} + z_{21})$. Branch of equilibria with $x > 0$ corresponds to both players choosing the cooperation strategy, and $x < 0$ corresponds to both players choosing to defect.

in [158] we extend the opinion dynamics model (3.10) as a framework for the study of strategic interactions in multi-player games. To do this we allow the agents' biases b_{ij} in the model to be driven by the game-theoretic payoff weights (7.1); for a full mathematical development we refer the reader to [158]. With this modification the model captures a balance between payoff-driven rationality and social interaction-driven reciprocity in each agent's decision-making. Furthermore in this work we analyze the bifurcations in this modified model for two-strategy games with symmetric payoffs between agents $(P_{CC})_i = P_{CC}$, $(P_{CD})_i = P_{CD}$, $(P_{DC})_i = P_{DC}$, $(P_{DD})_i = P_{DD}$ for all i . Analogously to the analysis presented in Chapter 4 we prove that a pitchfork bifurcation organizes the dynamics of decision-making in two-strategy games on many graphs. Furthermore, we can make concrete predictions about different games based on the structure of their payoff weights. For example we find that the quantity

$$P^\perp = P_{CC} + P_{CD} - P_{DC} - P_{DD} \quad (7.2)$$

determines the direction of unfolding of the pitchfork bifurcation. Specifically if all interactions between agents are positive in sign then $P^\perp > 0$ (< 0) implies that the primary branch of solutions in the pitchfork unfolding corresponds to all agents choosing strategy C (strategy D).

For a concrete example, consider a two-player symmetric prisoners' dilemma game with strategy C corresponding to cooperation and strategy D to defection. This type of game is characterized by the payoff weight structure $P_{DC} > P_{CC} > P_{DD} > P_{CD}$. With this constraint imposed, $P^\perp < 0$

for all prisoners' dilemma games and the primary branch in the pitchfork bifurcation unfolding corresponds to the mutual-defection Nash equilibrium. When the social coupling u between agents is weak, the Nash equilibrium is the only solution for the two-player game. However as u is increased, reciprocity between the players plays a stronger role and a simultaneously stable mutual cooperation equilibrium appears - see Figure 7.1. This modeling framework therefore rigorously predicts emergence of cooperation in social dilemmas, which is consistent with experimental evidence. In future work we plan to generalize the multi-option analysis in Chapter 5 of this dissertation to games involving more than two strategies in the model [158].

7.1.2 Dynamic task allocation

The modeling framework introduced in this dissertation can serve as a new tool for the design of dynamic task allocation algorithms in distributed technological teams. Allocating and managing tasks in a team through self-organization is an especially important design problem for coordination of robotic swarms and multi-robot teams [29, 77, 110, 114]. Algorithms for distributed task allocation often rely on biological inspiration, for example taking inspiration from honeybees [104], fish [197, 200], and ants [115]. Other common algorithms include market-based approaches in which robots communicate and place bids to compete for tasks [48]. Since the form of the model (3.3) takes inspiration from social decision-making in biological and human groups, its use for design of task allocation algorithms is similar in spirit to these popular approaches in the broader literature. Furthermore the model and its analysis scales to arbitrarily many options, which makes it especially compelling for use in multi-robot task allocation applications.

As we have shown with the analysis in this dissertation, the model (3.8) enables a group of decision-makers to self-organize into a variety of potentially multistable preference configurations. This self-organization has many properties which are desirable in a task allocation algorithm, which we discuss in [18]. For example in Chapter 4 we established a simple procedure that allows a designer to pre-assign each node on the network into one of two teams which will choose opposite tasks in a two-task environment. When agents are unbiased, the resulting task allocation is maximally flexible in that there are two bistable equilibria available to the network, and the network can select either equilibrium based on its initial conditions. Initial conditions can reflect relevant environmental factors, for example the spatial proximity of the individual robots to the locations of the tasks. This enables the group decision to adapt in real time to the environment. Furthermore, using the dynamic switching mechanism explored in Chapter 6 agents can change their allocation in real-time

which allows for the design of adaptable robotic teams. An expansion of these ideas to more than two tasks is part of future work.

In [67] we introduced a different mechanism which extends the opinion dynamics model (3.10) to allow a group of social agents to self-organize and allocate themselves across tasks in a pre-determined distribution. In contrast to the approach we discuss in [18], in [67] agents are not pre-assigned into groups. Instead, they are made aware of the desired proportion of the group required at each task. The group then self-organizes into the desired distribution following the modified opinion dynamics vector field. This task allocation application very clearly illustrates the benefit of introducing attention dynamics and opinion cascades, developed for two options in Chapter 6 of this dissertation. In [67] the proposed task allocation approach was tested with and without feedback dynamics in the attention parameter, analogous to (6.1b). One of the key insights of these numerical experiments is that dynamic feedback to agents' attention and the resulting opinion cascades greatly improve the accuracy with which the group is able to track the desired distribution across tasks, compared to the static attention case.

7.1.3 Cooperative navigation and collaboration with humans

The model (3.8) is being explored as a promising design tool for enabling robotic teams to cooperatively navigate through space and interact with people. The ability of a robotic team to cooperate and self-organize is a feature that is important for coordinating navigation of robotic units towards spatially embedded tasks which may be discovered by members of the distributed team as they are exploring their environment [52,171]. Recently dynamical systems approaches to composing and coordinating navigational tasks in robotic units and teams have proved successful in generating robust, legible, and adaptive behaviors that accomplish design goals [169,170,199]. These observations motivate the use of our model (3.3),(3.4) for coordinating robotic navigation in various applications. Furthermore since the model takes inspiration from social decisions in human and other biological social systems, it has potential for generating motion that is highly legible to a human observer.

The indecision-breaking bifurcations that organize the behavior of a swarm following (3.8) are an important feature that can be leveraged for design. This mechanism can enable a group of robotic units to reliably avoid any deadlocks which can arise from pathological symmetries in the physical space or in the task urgency. This is analogous to how groups in nature avoid deadlocks - for example, symmetry breaking bifurcations have been shown to organize the spatial movement of animals as they approach spatial targets which are arranged in an equidistant fashion from the

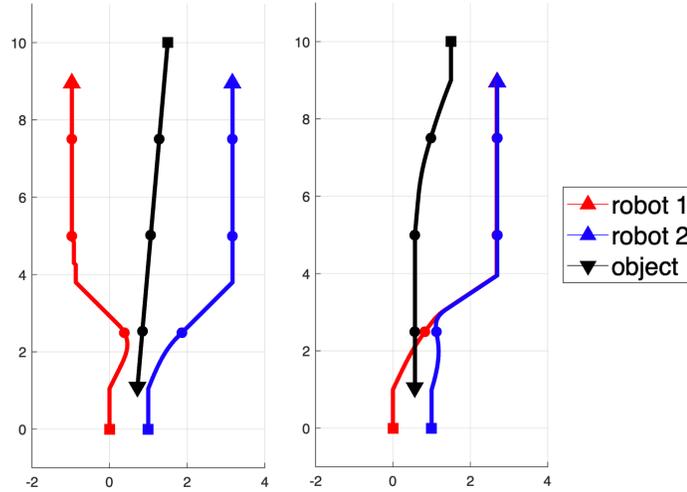


Figure 7.2: Figure from [32], simulation by Charlotte Cathcart. Birds eye view of a simulation of two robots (red and blue) avoiding collision with an oncoming mover (black) via the opinion dynamics model. Left: robots disagree on the direction of movement and avoid the oncoming mover by splitting up; right: robots agree on direction of movement and pass the oncoming mover on the same side.

animals [162, 189]. Symmetry breaking bifurcations are also a key mechanism that allows house-hunting honeybees to avoid deadlock when making a collective decision between nest sites which have equal value [167, 181]. In future work, we are exploring how a similar symmetry-breaking mechanism can be implemented in a robotic unit and in a team of robotic units whose decision-making is governed by the model (3.8), with the attention bifurcation parameter u being governed by proximity to the spatial targets. This design approach will enable a team to provably avoid deadlock in highly symmetric and pathological settings.

This modeling framework is also a promising tool for the design of collaborative decision-making by robotic teams which co-exist in environments with people. The ability for robotic units to navigate crowded environments and collaborate with people in their decision-making are important design challenges for intelligent robotic swarms [112, 202]. The following steps have been taken to explore the potential of this modeling framework to address these challenges:

1. An undergraduate senior thesis explored through simulation how the model (4.1) can be implemented on a robotic unit moving through space for reactive collision avoidance of an oncoming mover, such as a human navigating through a hallway [214].
2. Extended abstract [32] expands on these ideas and explores how a group of robotic units can interact with an oncoming mover using (4.1). We show in simulation that a group can collaboratively avoid an oncoming mover through a variety of strategies, for example agreeing to pass the unit on the same side, or splitting up - see Figure 7.2. This enables a group

of robots navigating space to react to their surroundings adapt their behavior flexibly. In upcoming work, interaction of a moving robot with a human using this framework is being explored through analysis and robotic experiments [33].

3. An undergraduate senior thesis explored through simulation, analysis, and robotic implementation how the game theoretic framework discussed earlier in this chapter can be used for collaborative lane merging between two moving autonomous vehicles [166].

7.1.4 Emergent properties of sociopolitical systems

The model we propose and analyze in this dissertation has the potential to aid social scientists in exploring hypotheses about mechanisms that give rise to emergent properties of social systems, such as online social networks and sociopolitical collectives. For example in [120] we used the model (4.1) to investigate possible mechanisms behind the asymmetric political polarization trends which have been observed over time in the voting records of partisan elites in the United States Congress. With this approach we were able to deduce that self-reinforcement of political opinions within each party is a more likely mechanism for explaining asymmetric trends in polarization than reflexive partisanship, i.e. parties opposing policies because they are supported by their opponent. This insight is an important step towards proposing evidence-based mitigation strategies to counteract the negative effects of polarization. A model analogous in its structure to the two-option model (4.1) was arrived at independently by researchers in sociophysics literature, in parallel to the modeling work presented in this dissertation. These researchers have shown that this model successfully reproduces qualitative features of echo chamber formation and polarization of opinions in debates in online social networks [14,15]. These works together suggest that the model and analysis we present in this dissertation can serve as a valuable tool in the future for gaining insight into the emergent properties of complex sociopolitical networks, and as a next step for exploring possible mitigation strategies to counteract polarization in societies.

7.2 Future extensions

The analysis of the model (3.8) presented in this dissertation is limited in a number of ways, which leaves opportunity for many future analyses and extensions of the model. We list a few of these below.

- Throughout this dissertation we have elected to use the attention parameter u as a bifurcation

parameter. Equivalently, any of the other model parameters $d, \alpha, \beta, \gamma, \delta$ can be used to establish an indecision-breaking bifurcation threshold. In fact we have already seen an example of this in our discussion of political polarization, wherein the increase in self-reinforcement α drove the partisan polarization. In future applications of this model, dynamics in any of these parameters can be linked to indecision-breaking. This observation paves the way for systematic comparisons of many different mechanistic hypotheses behind emergent features of social systems.

- The focus of our analysis so far has been the classification and description of primary indecision-breaking bifurcations. However, other local and global bifurcations can be found in (3.8), even in its homogeneous parameter setting. Equilibria and invariant sets which appear in these secondary bifurcations may be of equal interest for our understanding of emergent properties of social processes. This leaves room for many theoretical investigations of the rich dynamic behavior of (3.8).
- Analysis of belief formation on multiple topics in Chapter 5 is constrained to a single homogeneous belief system shared by all agents. However in real-world settings, different communities can subscribe to different sets of beliefs. Investigation of the properties of indecision-breaking bifurcations of (3.4) with two or more belief systems which coexist in the population of agents is an exciting open question. Mapping out these properties would also allow for a richer set of possibilities in the design of self organizing behaviors.
- So far we have always assumed that the communication topology between agents does not change over time and communication is instantaneous. However realistic social networks are time-varying, with agents forming and deleting communication links due to proximity, shared social groups, and other factors [123]. Furthermore realistic networked communication is not instantaneous, and possesses time delays [225]. These observations pose many important questions which must be addressed in future work. For example, are the multistable equilibria established in the analysis in this dissertation robust to changes in the communication topology and to time delays? Furthermore, can any pathological behavior arise on the network due to dynamic communication topology? These questions are important to address in order to establish safety and reliability of self-organized behaviors which rely on the model (3.4).

Part II

Papers

Chapter 8

Overview

Part II of this dissertation contains four peer-reviewed papers, with minor formatting modifications constituting the only differences between the original articles and the presented chapters. Only my first-author works are included in this dissertation, with other peer-reviewed publications I contributed to as a co-author cited throughout the text.

8.1 Outline

Chapter 9 presents a new dynamical system model of opinion formation on a social network. We show how this model generalizes several existing models of opinion dynamics in the literature. We prove that agreement and disagreement opinions can form on the network in a bifurcation of the neutral state. Furthermore we investigate how through parameter dynamics, a network can take advantage of its high sensitivity to input near a bifurcation point, and how this sensitivity can be tuned through controller design. This chapter appears in the IEEE Transactions on Automatic Control journal [19].

Chapter 10 studies a 2-option specialization of the model proposed in Chapter 9 on networks with homogeneous signs. Specifically, we show how graph symmetries and centrality of agents on the network informs the outcome of agreement and disagreement opinion formation. We show through simulation how centrality determines whether a signal to an agent will trigger a cascade of opinionation on a network of decision-makers with dynamic attention which is governed through a feedback law. This chapter appears in the Proceedings of the 2021 American Control Conference [20]

Chapter 11 formally expands the observations made in Chapter 10 about the role of centrality in

opinion cascades. First, this chapter establishes a formal proof that any bifurcation of the origin in the scalar opinion dynamics model of Chapter 9 must be a pitchfork bifurcation. Next, the chapter introduces a dynamic feedback law which governs each agent’s attention and formally connects opinion cascades on the network with a saddle-node bifurcation in the dynamics of opinions coupled with attention. This bifurcation establishes a tunable threshold which separates inputs which trigger network cascades from inputs which do not. We illustrate this threshold in a series of simulations. This chapter appears in the Proceedings of the 2021 Conference on Decision and Control [21]

Chapter 12 studies the scalar opinion dynamics model on networks with mixed-sign interactions. First the notion of switching equivalence of two signed graphs is defined, and a relationship between patterns of opinions on switching equivalent graphs is established. Next, we prove that agents can change the opinion pattern on the network dynamically, through a local switching transformation. Finally, applications of switching dynamics to multi-robot teams are explored through motivating scenarios. This chapter appears in the IEEE Control Systems Letters journal [18].

8.2 Author contributions

I am the lead author and lead contributor to the material presented in all four included papers. My advisor, Naomi Ehrich Leonard, and co-author Alessio Franci were closely involved with the development of all of the presented works throughout the various stages of planning, writing, and response to peer review. Alessio Franci was particularly involved in checking the mathematical results in all of these manuscripts and provided many helpful suggestions and references. Specific author contributions are listed below.

- In Chapter 9, Alessio Franci had the original idea to generalize nonlinear consensus models to include disagreement as a potential solution. With his guidance I developed a simple low-dimensional version of the presented model, which I presented for my general examination in 2019. The final version of the model which appears in the chapter was arrived at through several months of meticulous collaborative efforts between myself, Alessio Franci, and Naomi Ehrich Leonard. The original version of the proof for the boundedness result in this chapter was outlined by Alessio Franci, which I then iterated on to arrive at what is presented in the text. Similarly, an early version of Theorem 10.4.2 was originally written down by Alessio Franci. Alessio Franci also developed the original discussion of the comparison between our model with the linear averaging models, which was then iterated on by myself and Naomi Ehrich Leonard. Any discussion of axial subgroups of symmetry groups and how they inform solution

branches of bifurcations that is present in the text was the contribution of Alessio Franci. Figure 10.10 is adapted from the conference paper [21] and includes data from simulations run by Timothy Sorochkin. Figure 10.4 was adapted from a figure which was originally created by myself and Ayanna Matthews for the conference publication [20]. This paper went through many collaborative iterations and revisions, and the writing throughout the manuscript is a well-mixed combination of contributions by all three of the co-authors - myself, Alessio Franci, and Naomi Ehrich Leonard.

- In Chapter 10, Ayanna Matthews contributed to the development of Figures 11.1, 11.3, and 11.4. Ayanna Matthews also ran many helpful simulations which helped confirm the connection between standard network centrality measures and the strength of opinions formed in this model, as well as contributed to some of the writing introducing the model and describing its various terms. The idea to use a saturation function in the dynamic feedback law for attention was developed collaboratively with Naomi Ehrich Leonard, inspired by active feedback in neuronal systems.
- In Chapter 11, Timothy Sorochkin contributed to the development of Figures 12.2 and 12.4. He was involved with this paper from its inception, and ran many greatly helpful simulations of opinion cascades with various attention dynamics which helped narrow down the right parameter regimes and visualization techniques for the presented results.
- In Chapter 12, Giovanna Amorim created Figures 13.2, 13.4, and 13.5 and Alessio Franci created Figure 13.3. María Santos wrote a part of the introduction and Giovanna Amorim wrote part of the section introducing the model. María Santos and Giovanna Amorim collaboratively developed the applications of our analysis to multi-robot systems and together wrote section 13.6 of the paper.

Chapter 9

Nonlinear Opinion Dynamics with Tunable Sensitivity

ANASTASIA BIZYAEVA, ALESSIO FRANCI, AND NAOMI EHRICH LEONARD

We propose a continuous-time multi-option nonlinear generalization of classical linear weighted-average opinion dynamics. Nonlinearity is introduced by saturating opinion exchanges, and this is enough to enable a significantly greater range of opinion-forming behaviors with our model as compared to existing linear and nonlinear models. For a group of agents that communicate opinions over a network, these behaviors include multistable agreement and disagreement, tunable sensitivity to input, robustness to disturbance, flexible transition between patterns of opinions, and opinion cascades. We derive network-dependent tuning rules to robustly control the system behavior and we design state-feedback dynamics for the model parameters to make the behavior adaptive to changing external conditions. The model provides new means for systematic study of dynamics on natural and engineered networks, from information spread and political polarization to collective decision making and dynamic task allocation.

9.1 Introduction.

Opinion dynamics of networked agents are the subject of long-standing interdisciplinary interest, and there is a large and growing literature on agent-based models created to study mechanisms that drive the formation of consensus and opinion clustering in groups. These models appear, for example, in studies of collective animal behavior and voting patterns in human social networks. In

engineering, they are fundamental to designing distributed coordination of autonomous agents and dynamic allocation of tasks across a network.

Agent-based models are typically used to investigate parameter regimes and network structures for which opinions in a group converge over time to a desired configuration. However, natural groups exhibit much more flexibility than captured with existing models. Remarkably, groups in nature can rapidly switch between different opinion configurations in response to changes in their environment, and they can break deadlock, i.e., choose among options with little, if any, evidence that one option is better than another. Understanding the mechanisms that explain the temporal dynamics of opinion formation in groups and the ultra-sensitivity and robustness needed for groups to pick out meaningful information and to break deadlock in uncertain and changing environments is important in its own right. It is also pivotal to developing the means to design provably adaptable yet robust control laws for robotic teams and other networked multi-agent systems.

Motivated by these observations, we explore the following questions in this paper. How can a network of decision makers come rapidly and reliably to coherent configurations of opinions, including both agreement and disagreement, on multiple options in response to, or in the absence of, internal biases or external inputs? How can a network reliably transition from one configuration of opinions to another in response to change? How can the sensitivity of the opinion formation process be tuned so that meaningful signals are distinguished from spurious signals? To investigate these questions, we present an agent-based dynamic model of the opinion formation process that generalizes linear and existing nonlinear models. The model is rich in the behaviors it exhibits yet tractable to analysis by virtue of the small number of parameters needed to generate the full range of behaviors.

We emphasize that our modeling approach is distinct from existing models in the literature in the following way. Models of opinion formation are typically built on the fundamental assumption that individuals update their opinions through a linear averaging process [5, 38, 45, 71, 153]. Additional feedback dynamics are then often imposed on the coupling weights between agents, for example in bounded confidence models [22, 44, 93, 94], biased assimilation models [43, 218], and models of evolution of social power [105, 220]. Nonlinearity thereby arises through the superposition of linear opinion dynamics and nonlinear coupling-weight dynamics. When persistent disagreement is observed, it is necessarily the consequence of the dynamic updating of the coupling weights. However, state-dependent interactions are not the only way for a network to achieve structurally stable disagreement. We are instead proposing that the opinion update process itself is fundamentally nonlinear due to saturation of information. We introduce a new multi-option nonlinear model of opinion formation with saturated interactions in Section 9.3 and in Section 9.4 we prove that this

modeling assumption supports persistent disagreement with a completely static interaction network.

As is done for linear models, dynamic feedback can also be introduced to the nonlinear model parameters. We explore the effects of several dynamic parameter update laws in detail in Section 9.5. The feedback laws we consider are simple, yet they make our model adaptive to changing external conditions with tunable sensitivity and they allow robust and tunable transitions between distinctly different patterns of opinions.

Our model generalizes recent literature on opinion formation with input saturation [1, 49, 61, 62, 70, 87, 102]. Closely related to these are nonlinear models that leverage coupled oscillator dynamics [121, 139, 144], biologically inspired mean-field models [156], and the Ising model [162, 185].

Our primary contributions are as follows.

1) We introduce a new nonlinear model for the study of multi-agent, multi-option opinion dynamics. The model has a social term weighted by an attention parameter, which can also represent social effort or susceptibility to social influence, and an input term, which can represent, e.g., external stimuli, bias, or persistent opinions.

2) We show that the model exhibits a rich variety of opinion-formation behaviors governed by bifurcations. This includes rapid and reliable opinion formation and multistable agreement and disagreement, with flexible transitions between them. It also includes ultra-sensitivity to inputs near the opinion forming bifurcation, and robustness to disturbances and uncertainties, away from the bifurcation. Moreover, the behaviors are governed by a small number of key parameters, rendering the model analytically tractable. We prove the central role of the spectral properties of the network graph adjacency matrix in informing the model behavior.

3) We show how the model recovers a range of models in the literature for suitable parameter combinations and/or when linearized, and how the reliance on structurally unstable network conditions in linear models breaks down in the nonlinear setting. The central role of the network graph adjacency matrix in our nonlinear model generalizes the central role of the network graph Laplacian in opinion dynamics in the literature. We show that the right and left adjacency matrix eigenstructures determine patterns of opinion and sensitivity to inputs, respectively.

4) We introduce distributed adaptive feedback dynamics to the agent parameters. We show how design parameters in the attention feedback allow tunable sensitivity of opinion formation to inputs and robustness to changes in inputs, as well as tunable opinion cascades even in response to a single agent receiving an input.

5) We examine tunable transitions between consensus and dissensus using feedback dynamics also on network weights.

We define notation in Section 9.2. We present the new nonlinear opinion dynamics model in Section 9.3. In Section 9.4 we prove results on agreement and disagreement opinion formation for the new model. We introduce attention feedback dynamics and prove results on tunable sensitivity in Section 9.5 in the special case of two options. In Section 9.5.4, we illustrate feedback controlled transitions between agreement and disagreement. We conclude in Section 9.6.

9.2 Notation

Given $\mathbf{y} \in \mathbb{R}^N$, the norm $\|\mathbf{y}\|$ is the standard Euclidean 2-norm and $\text{diag}\{\mathbf{y}\} \in \mathbb{R}^{N \times N}$ is a diagonal matrix with y_i in row i , column i . Let $\mathcal{I}_N \in \mathbb{R}^{N \times N}$ be the identity matrix, $\mathbf{1}_N \in \mathbb{R}^N$ the vector of ones, and $P_0 = (\mathcal{I}_{N_o} - \frac{1}{N_o} \mathbf{1}_{N_o} \mathbf{1}_{N_o}^T)$ the projection onto $\mathbf{1}_{N_o}^\perp$. Let $\mathbb{R}\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ be the span of vectors $\mathbf{v}_1, \dots, \mathbf{v}_k \in \mathbb{R}^N$. We define $\mathbf{v}_i \in \mathbb{R}^N$ component-wise as (v_{i1}, \dots, v_{iN}) . Let U, V and W be vector spaces. U is the *direct sum* of V and W , i.e., $U = V \oplus W$, if and only if $U = V + W$ and $V \cap W = \{0\}$. Given matrices $B = (b_{ij}) \in \mathbb{R}^{m \times n}$ and $C = (c_{ij}) \in \mathbb{R}^{p \times q}$, the *Kronecker product* $B \otimes C \in \mathbb{R}^{mp \times nq}$ has entries $(B \otimes C)_{pr+q, qs+w} = b_{rs} c_{vw}$.

Let the set of vertices $V = \{1, \dots, N_a\}$ index a group of N_a agents, and let edges $E \subseteq V \times V$ represent interactions between agents. If edge $e_{ik} \in E$, then agent k is a *neighbor* of agent i . The communication topology between agents is captured by the *directed* graph $G = (V, E)$ and its associated adjacency matrix $A \in \mathbb{R}^{N_a \times N_a}$. A is made up of elements a_{ik} , and $a_{ik} \neq 0$ if and only if agent k is a neighbor of agent i . When A is symmetric (i.e., communication between agents is bidirectional), the graph is *undirected*.

9.3 Nonlinear multi-option opinion dynamics

In this section we present our nonlinear model of opinion dynamics for a network of interacting agents that form opinions about an arbitrary number of options. In Section 9.3.1 we recall the classical consensus model of DeGroot [45] and several of the extensions that have been proposed and studied in the literature. All of the cited models (with one exception noted) use an opinion update rule that depends on a linear weighted-average of exchanged opinions. In our model, as discussed in Section 9.3.2 and formalized in Section 9.3.3, we apply a saturation function to opinion exchanges, which makes the update rule fundamentally nonlinear, even before introducing extensions. The fundamentally nonlinear update rule makes all the difference with respect to generality and flexibility of the model as we show here and in the rest of the paper.

9.3.1 Linear Averaging Models: Drawbacks and Extensions

Opinion formation is classically modeled as a weighted-averaging process, as originally introduced by DeGroot [45]. In this framework an agent’s opinion $x_i \in \mathbb{R}$ reflects how strongly the agent supports an issue or topic of interest. The real-valued opinion is updated in discrete time as a weighted average of the agent’s own and other agents’ opinions, i.e.,

$$x_i(T + 1) = a_{i1}x_1(T) + \cdots + a_{iN_a}x_{N_a}(T) \quad (9.1)$$

where $a_{i1} + \cdots + a_{iN_a} = 1$ and $a_{ik} \geq 0$. The weights a_{ik} describe the influence of the opinion of agent k on the opinion of agent i and the matrix $A \in \mathbb{R}^{N_a \times N_a}$ with entries a_{ik} represents the structure of the influence network.

A key drawback of linear weighted-average models is that consensus among the agents is the only possible outcome. As observed in [134], this necessarily happens because the attraction strength of agent i ’s opinion toward agent k ’s opinion increases linearly with the difference of opinions between the two agents. In other words, the more divergent the two agents’ opinions are, the more strongly they are attracted to each other, which is paradoxical from an opinion formation perspective.

To overcome these limitations, a number of prominent variations on averaging models have been proposed. For example in “bounded confidence” models, agents average network opinions but delete communication links to any neighbors whose opinions are sufficiently divergent from their own [22, 44, 93, 94]. In a similar spirit, “biased assimilation” models instead incorporate a self-feedback into the interaction weights of an averaging model [43, 218]. This self-feedback accounts for an individual’s bias towards evidence that conforms with its existing beliefs. The linear model and its variations have also been extended to the case of signed networks, where the linear weights a_{ik} can be negative [5, 127, 184]. In [134] the authors do away with averaging altogether and instead propose that opinions form through a weighted-median mechanism.

In the present paper we propose an alternative perspective to this literature: driven by the above motivation and the model-independent theory developed in [68], we introduce a parsimonious non-linear extension of linear weighted-average opinion dynamics that leverages the saturation function.

The linear weighted-average discrete-time opinion dynamics (9.1) can equivalently be written as

$$x_i(T + 1) = x_i(T) + \left(-x_i(T) + a_{i1}x_1(T) + \cdots + a_{iN_a}x_{N_a}(T) \right).$$

This discrete-time update rule is the unit time-step Euler discretization of the continuous time linear

dynamics

$$\dot{x}_i = -x_i + a_{i1}x_1 + \cdots + a_{iN_a}x_{N_a}. \quad (9.2)$$

Observe that (9.1) and (9.2) have exactly the same steady states with the same (neutral) stability.

The linear consensus dynamics (9.2) are determined by two terms: a weighted-average opinion-exchange term, modeling the pull felt by agent i toward the weighted group opinion, and a linear damping term, which can be interpreted as the agent’s *resistance* to changing its opinion.

9.3.2 Nonlinear Multi-option Extension of Weighted-average Models: Defining Properties

Our goal is to derive a novel nonlinear extension of (9.2) satisfying the following defining properties.

1. Opinion exchanges are saturated. Saturated nonlinearities appear in virtually every natural and artificial signaling network due to bounds on action and sensing. For example, dynamics that evolve according to saturating interactions appear in spatially localized and extended neuronal population models of thalamo-cortical dynamics [211,212], in Hopfield neural network models [98,99,145], in models of perceptual decision making [23,205], and in control systems with sensor and actuator saturations [103,126]. Saturated interactions between decision-makers also effectively bound the attraction between opinions, thus overcoming the linear weighted-average model paradox mentioned above.

2. Multi-option opinion formation. Allowing for an arbitrary number of options makes the model relevant to a wide range of applications, for example, in task allocation problems where options represent tasks or in strategic settings where options represent strategies. We extend the model to multiple options by suitably generalizing the agent’s opinion state space, analogous to existing multi-option extensions of averaging models such as [65,72,125,147,157,160,221].

To construct this extension formally, observe that in the scalar opinion setting, $x_i > 0$ ($x_i < 0$) is usually interpreted as favoring (disfavoring) an option A and disfavoring (favoring) an option B. The strength of favoring or disfavoring is represented by the magnitude $|x_i|$ and $x_i = 0$ is interpreted as being neutral. This formalism is equivalent to one in which each agent is characterized by two scalar variables z_{iA} (modeling the preference of agent i for option A) and z_{iB} (modeling the preference of agent i for option B) that are “mutually-exclusive”, i.e., that satisfy $z_{iA} + z_{iB} = 0$. The scalar opinion is then obtained simply by defining $x_i = z_{iA}$. This observation leads to the following multi-option generalization of the state space of model (9.2). Given N_o options, we model each agent’s opinion state space as the subspace $\mathbf{1}_{N_o}^\perp \subset \mathbb{R}^{N_o}$. Thus, in our model, the opinion state of agent

$i, i = 1, \dots, N_a$, is described by the state variable $\mathbf{Z}_i \in \mathbf{1}_{N_o}^\perp$, with components $z_{ij}, j = 1, \dots, N_o$. When $\mathbf{Z}_i = 0$, we say that agent i is *neutral* or *unopinionated*. When $\mathbf{Z}_i \neq 0$ we say that the agent is *opinionated*. The full model state space is $V = \underbrace{\mathbf{1}_{N_o}^\perp \times \dots \times \mathbf{1}_{N_o}^\perp}_{N_a \text{ times}}$, and $\mathbf{Z} = (\mathbf{Z}_1, \dots, \mathbf{Z}_{N_a}) \in V$ is the system state. The origin $\mathbf{Z} = 0$ is the *neutral point*. Another way of interpreting our choice of $\mathbf{1}_{N_o}^\perp$ as an agent's state space comes from observing that $\mathbf{1}_{N_o}^\perp$ is the tangent space to the $(N_o - 1)$ -dimensional simplex in \mathbb{R}^{N_o} . Because $\mathbf{1}_{N_o}^\perp$ and the simplex are isomorphic, our modeling approach naturally applies to multi-option decision-making problems in which an agent's state space is the $(N_o - 1)$ -dimensional simplex. This is useful when the agents' opinions are interpreted as probabilities of choosing options, for example, in the case of mixed strategies in games where an option refers to a strategy [158]. For more details on the connection to simplex dynamics see Corollary 9.7.3 in the Appendix.

3. Agents have allocable attention. Because an agent's attention to exchanged opinions may be variable, we introduce, for each agent i , two parameters, $d_i > 0$ and $u_i \geq 0$, that weight the relative influence of the linear resistance term and the nonlinear opinion-exchange term, respectively. When the *resistance parameter* d_i dominates the *attention parameter* u_i , the agent is weakly attentive to other agents' opinions. When u_i dominates d_i , the agent is strongly attentive to other agents' opinions. A shift from a weakly attentive to a strongly attentive state can be induced, for instance, by a time-urgency (election day approaching) or a spatial-urgency (target getting closer) to form an informed collective opinion. The attention parameter u_i can also be used to model *social effort*, *excitability*, or *susceptibility* of agent i to social influence.

4. Agents have exogenous inputs. For each agent, we introduce an input parameter b_{ij} , which represents an input signal from the environment or a bias or predisposition that directly affects agent i 's opinion of option j . For example, the input b_{ij} can be used to model the exogenous influence of agent i 's initial opinions, as in [71], where agents hold on to their initial opinions (sometimes called "stubborn" agents as in [78]).

If the attention and/or bias parameters are hard or impossible to measure or control, which may be the case in sociopolitical applications, we can use standard homogeneity assumptions, e.g., $d_i = 1, u_i = u, b_{ij} = 0$ for all agents, and include random perturbations to capture modeling uncertainties. In technological applications (e.g. robotic swarms), however, tunable parameters of the model provide novel, analytically tractable means to design complex collective behaviors – see for example [67].

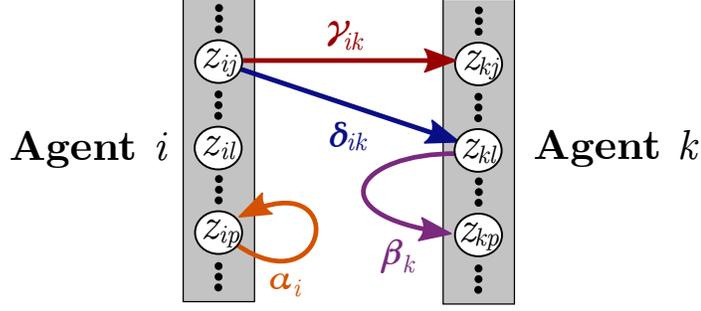


Figure 9.1: Illustration of the four classes of interactions. An arrow from z_{ij} to z_{kl} means the opinion of agent i about option j is influenced by the opinion of agent k about option l , modulo the labeled gain.

9.3.3 A General Nonlinear Opinion Dynamics Model

In the multi-option setting, there are four possible types of coupling in the resulting opinion network (see Figure 9.1):

- 1) *Intra-agent, same-option coupling*, with gain α_i ;
- 2) *Intra-agent, inter-option coupling*, with gain β_i ;
- 3) *Inter-agent, same-option coupling*, with gains γ_{ik} , $i \neq k$;
- 4) *Inter-agent, inter-option coupling*, with gains δ_{ik} , $i \neq k$.

Parameters α_i , β_i , γ_{ik} , δ_{ik} determine qualitative properties of opinion interactions. Parameter α_i determines sign and magnitude of opinion self-interaction for agent i . To avoid redundancy with resistance d_i , we assume $\alpha_i \geq 0$, i.e., either no self-coupling ($\alpha_i = 0$) or *self-reinforcing* coupling ($\alpha_i > 0$). Parameter β_i determines how different intra-agent opinions interact. Parameters γ_{ik} and δ_{ik} determine if i 's response to k is to *cooperate* ($\gamma_{ik} - \delta_{ik} > 0$) or *compete* ($\gamma_{ik} - \delta_{ik} < 0$). When different option dimensions have no interdependence, we can set $\beta_i = \delta_{ik} = 0$ for all $i, k = 1, \dots, N_a$.

The proposed general nonlinear opinion dynamics are

$$\dot{\mathbf{Z}}_i = P_0 \mathbf{F}_i(\mathbf{Z}) \quad (9.3a)$$

$$F_{ij}(\mathbf{Z}) = -d_i z_{ij} + b_{ij} + u_i \left(S_1 \left(\alpha_i z_{ij} + \sum_{k=1}^{N_a} \gamma_{ik} z_{kj} \right) + \sum_{l=1}^{N_o} S_2 \left(\beta_i z_{il} + \sum_{k=1}^{N_a} \delta_{ik} z_{kl} \right) \right) \quad (9.3b)$$

where $z_{ij}(t)$ is agent i 's opinion of option j at time t , $\mathbf{Z}_i(t) = (z_{i1}, \dots, z_{iN_o})(t) \in \mathbf{1}_{N_o}^\perp$ is agent i 's opinion state at time t as introduced in Section 9.3.2, and $\dot{\mathbf{Z}}_i = \frac{d\mathbf{Z}_i}{dt}$. $S_q : \mathbb{R} \rightarrow [-k_{q1}, k_{q2}]$ with $k_{q1}, k_{q2} \in \mathbb{R}^{>0}$ for $q \in \{1, 2\}$ is a generic sigmoidal saturating function satisfying constraints $S_q(0) = 0$, $S_q'(0) = 1$, $S_q''(0) \neq 0$, $S_q'''(0) \neq 0^1$. S_1 saturates same-option interactions, and S_2

¹ $S_q''(0) \neq 0$ is a nondegeneracy condition, in the sense of [82], only for $N_o > 2$. For $N_o = 2$, the simplex projection (9.3a) imposes odd symmetry of the opinion dynamics and makes this assumption unnecessary. See (9.5) and below.

saturates inter-option interactions. S_1 and S_2 could be the same but are distinguished in (9.3) for a more general statement of the model. We provide an even more general formulation of the model in Appendix 9.7.1 that makes use of an adjacency tensor and allows for the possibility of heterogeneous interactions between options. Let

$$\Gamma = [\gamma_{ik}] \in \mathbb{R}^{N_a \times N_a}, \quad \Delta = [\delta_{ik}] \in \mathbb{R}^{N_a \times N_a}. \quad (9.4)$$

In (9.3) the sum over the agents could be brought outside of the two sigmoids without altering the qualitative behavior of the model. Our choice in (9.3) corresponds to an opinion network with saturated inputs. Bringing the sum over the agents outside the sigmoids corresponds to an opinion network with saturated outputs. Either choice could be useful depending on the application. On the other hand, the sum over the options cannot be brought inside S_2 as the mutual exclusivity condition $\mathbf{Z}_i \in \mathbf{1}_{N_o}^\perp$ would lead to spurious term cancellations for some parameter choices. Intuitively, this means that opinions about different options are processed through different input channels. Dynamics (9.3) are well defined on the system state space V , as we rigorously prove in Appendix 9.7.2.

Let $\hat{b}_i = \frac{1}{N_o} \sum_{l=1}^{N_o} b_{il}$ be the *average input to agent i* and let $b_{ij}^\perp = b_{ij} - \hat{b}_i$ be the *relative input to agent i for option j* .

Lemma 9.3.1. *The dynamics (9.3) are independent of the average input \hat{b}_i in the sense that $\frac{\partial z_{ij}}{\partial \hat{b}_i} = 0$.*

Proof. Recall that P_0 is the projection onto $\mathbf{1}_{N_o}^\perp$ as defined in Section 9.2. Then $P_0 \mathbf{b}_i = \mathbf{b}_i^\perp$, and the conclusion follows trivially from the form of (9.3). \square

Lemma 9.3.1 implies that only relative inputs affect the location of the equilibria of the opinion dynamics (9.3).

Assumption 9.3.2. *In light of Lemma 9.3.1, for the remainder of the paper we assume without loss of generality that the average input $\hat{b}_i = 0$ for all $i = 1, \dots, N_a$. Thus, $b_{ij} = b_{ij}^\perp$.*

When relative inputs are absent, the system (9.3) always has the neutral point as an equilibrium.

Lemma 9.3.3. *$\mathbf{Z} = 0$ is an equilibrium for (9.3) if and only if there are no relative inputs, i.e., $b_{ij}^\perp = 0$ for all i and all j .*

When relative inputs are small, i.e., they do not dominate the dynamics, the formation of opinions in the general model (9.3) is governed by the balance between the resistance term, which inhibits

opinion formation, and the social term, which promotes opinion formation. For illustrative purposes, consider the case in which $u_i = u \geq 0$ for all i . Then for u small, resistance dominates and the system behaves linearly. The opinions z_{ij} remain small and their relative magnitudes are determined by the small inputs b_{ij} . For u large, the social term dominates and the system behaves nonlinearly.

Importantly, in this nonlinear regime, where u is large enough that opinion exchanges dominate resistance, opinions z_{ij} form that are much larger than, and potentially unrelated to, inputs b_{ij} , even for very small initial conditions. Opinion exchanges govern opinion formation through bifurcations at which the neutral equilibrium loses stability as discussed in the next section and formalized and investigated in the remainder of the paper.

9.3.4 Generality and Connection to Existing Models

The model (9.3) is general in the sense that it recovers a number of published opinion-formation, decision-making, and consensus models for specific sets of parameters and/or when linearized. In order to illustrate this we consider the model specialized to $N_o = 2$, as most of the models in the literature consider two-option scenarios. The opinion state of agent i is one-dimensional: following the notation introduced in Section 9.3.2 we define $x_i = z_{i1} = -z_{i2}$ as agent i 's opinion. Then, opinion dynamics (9.3) reduce to

$$\dot{x}_i = -d_i x_i + u_i \left(\hat{S}_1 \left(\alpha_i x_i + \sum_{k \neq i}^{N_a} \gamma_{ik} x_k \right) - \hat{S}_2 \left(\beta_i x_i + \sum_{k \neq i}^{N_a} \delta_{ik} x_k \right) \right) + b_i \quad (9.5)$$

where $\hat{S}_l(y) = \frac{1}{2}(S_l(y) - S_l(-y))$ are odd saturating functions for $l = 1, 2$, $b_i := b_{i1} = -b_{i2}$, and $d_i = \frac{1}{2}(d_{i1} + d_{i2})$. Let the network opinion state be $\mathbf{x} = (x_1, \dots, x_{N_a}) \in \mathbb{R}^{N_a}$ and vector of inputs be $\mathbf{b} = (b_1, \dots, b_{N_a}) \in \mathbb{R}^{N_a}$. When interactions between option dimensions are disregarded, i.e. with $\beta_i = \delta_{ik} = 0$ for all $i, k = 1, \dots, N_a$, the two-option model (9.5) further reduces to

$$\dot{x}_i = -d_i x_i + u_i \hat{S}_1 \left(\alpha_i x_i + \sum_{k \neq i}^{N_a} \gamma_{ik} x_k \right) + b_i \quad (9.6)$$

which, with appropriate restrictions on the model coefficients, recovers a number of nonlinear consensus models studied in recent literature. We illustrate this in the following example.

Example 9.3.4 (Specialization to nonlinear consensus protocols in the literature). *A. When $\alpha_i = 0$, $\gamma_{ik} \in \{0, 1\}$ (or more generally, $\gamma_{ik} \geq 0$), $u_i := u \geq 0$, and the resistance parameter d_i is defined as $d_i := \sum_{k=1}^{N_a} \gamma_{ik}$ with $k \neq i$ (the network in-degree for node i), (9.6) reduces to the nonlinear consensus dynamics of [1, 61, 70, 87].*

B. When $\alpha_i = 0$, $\gamma_{ik} \in \{0, 1, -1\}$ (or more generally, $\gamma_{ik} \in \mathbb{R}$), $u_i := u \geq 0$, and the resistance parameter d_i is defined as $d_i := \sum_{k=1}^{N_a} |\gamma_{ik}|$ with $k \neq i$, (9.6) reduces to the nonlinear consensus dynamics with antagonistic interactions studied in [62, 63].

In the nonlinear consensus models of Example 9.3.4, the formation of consensus opinions on the network is a bifurcation phenomenon. Namely when $b_i = 0$ for all $i = 1, \dots, N_a$ and $0 \leq u < u^*$, the neutral point $\mathbf{x} = 0$ is an asymptotically stable equilibrium. At a critical value $u = u^* > 0$ a pitchfork bifurcation is observed in both models, at which point $\mathbf{x} = 0$ loses stability and two nonzero asymptotically stable equilibria appear [87, Theorem 1], [63, Theorem 1]. For nonzero inputs, the pitchfork unfolds.

Importantly, the linearization of these models about the origin $\mathbf{x} = 0$ at $u = 1$ yields $\dot{\mathbf{x}} = -(D - \Gamma)\mathbf{x}$, where $D = \text{diag}(d_i) \in \mathbb{R}^{N_a \times N_a}$ is the degree matrix for the network. For the positive weights of Example 9.3.4.A this corresponds to the standard Laplacian consensus protocol [153], a continuous-time analogue of the weighted-average models discussed in Section 9.3.1. For the signed weights of Example 9.3.4.B this linearization is exactly the model of linear consensus with antagonistic, i.e., signed, interconnections [5, 127, 184].

In linear models, nonzero agreement (consensus) and disagreement (e.g., bipartite consensus and its generalizations) equilibria are never exponentially asymptotically stable because the model Jacobian has a zero eigenvalue. The eigenspace of the zero eigenvalue is $\mathbb{R}\{\mathbf{1}\}$ in the case of agreement, whereas it is spanned by a mixed-sign vector determined by the coupling topology in the case of disagreement [5, 136, 157, 165]. In other words, linear agreement and disagreement models are not structurally stable and arbitrary small unmodelled (nonlinear) dynamics will in general destroy the predicted behavior. Adding saturated opinion exchanges has a two-fold advantage: i) it makes the model generically structurally stable and, therefore, the agreement and disagreement equilibria hyperbolic (i.e., with no eigenvalues on the imaginary axis); ii) it weakens the necessary conditions for the existence of stable disagreement states.

In linear models, the existence of neutrally stable agreement or disagreement states is always linked to restrictive and non-generic assumptions on the coupling topology, for example, balanced coupling for consensus [153] and either strongly connected structurally balanced coupling [5, 157], quasi-strongly connected coupling with an in-isolated structurally balanced subgraph [165], or the existence of a spanning tree on the coupling graph [136] for disagreement.

As we prove in Section 9.4 for our model, agreement is always possible for generic strongly connected (balanced or unbalanced) graphs, and disagreement only requires a weak and provable

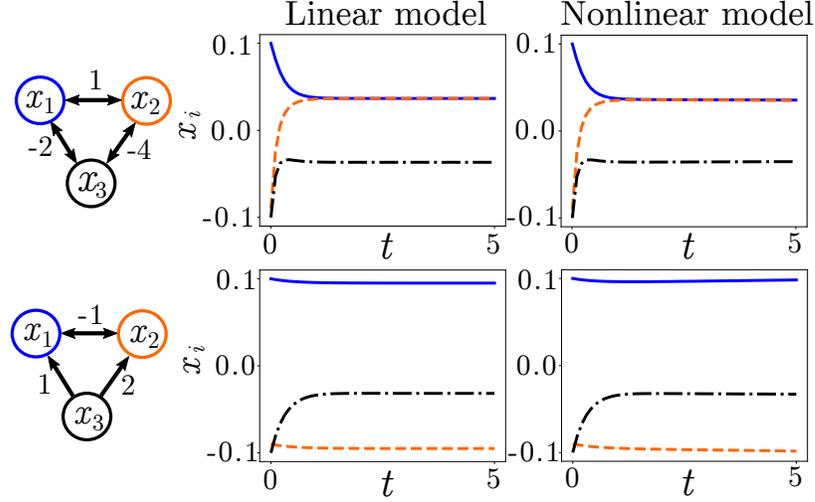


Figure 9.2: For small initial conditions $\mathbf{x}(0) = (0.1, -0.09, -0.1)$, trajectories of the linear model from [5] (left) approximate trajectories of the nonlinear model of Example 9.3.4.B (right) with $N_a = 3$, $u = 1.01$, $b_i = 0$, and $\hat{S}_1 = \tanh$. Top: bipartite consensus on a strongly connected and structurally balanced graph, as in [5, Example 1]. Bottom: polarized opinions on a quasi-strongly connected graph containing an in-isolated structurally balanced subgraph, as in [165, Example 1]. Arrows point to neighbors per the convention of Fig. 9.1.

condition on the spectral properties of the adjacency matrix (satisfied even for networks with homogeneous weights). It follows that our model recovers the behavior of linear models when one of the above conditions is satisfied (Figure 9.2) but avoids the conservativeness of linear model predictions under more general coupling topologies (Figure 9.3). In Figure 9.3, for a network with all negative edges weights, the linear model predicts neutrality whereas our model predicts disagreement (see also Figure 9.4). In the linear model, the repulsion felt among agents balances the resistance, whereas in the nonlinear model, the repulsion dominates the resistance, destabilizing the neutral solution and driving different agents to form a strong opinion for different available options.

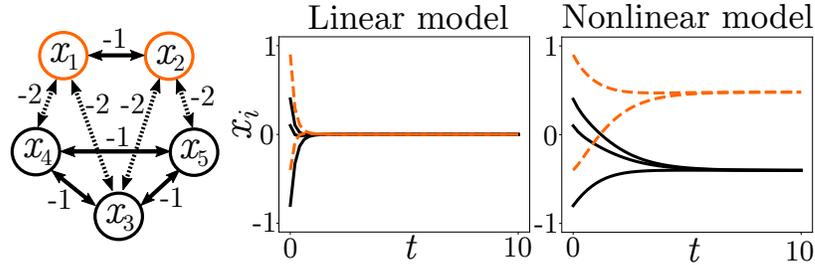


Figure 9.3: The model from [5] (left) and nonlinear model (9.6) (right) with $N_a = 5$, $u_i = 0.5$, $d_i = 1$, $b_i = 0$ for all $i = 1, \dots, N_a$, $\hat{S}_1 = \tanh$, initial conditions $\mathbf{x}(0) = (0.9, -0.4, 0.4, 0.1, -0.8)$, and same adjacency matrix given by $\gamma_{ik} = -1$ for $i, k \in \mathcal{I}_p$, $i \neq k$, and $\gamma_{ik} = -2$ for $i \in \mathcal{I}_p$, $k \in \mathcal{I}_s$, $p \neq s$ for clusters with indices $\mathcal{I}_1 = \{1, 2\}$ and $\mathcal{I}_2 = \{3, 4, 5\}$ – see network diagram for illustration of the interconnection topology. The linear model converges to the neutral solution. The nonlinear model, however, converges to a stable clustered dissensus state, as follows from Remark 9.4.5.

9.3.5 Clustering and Model Reduction

The opinion states \mathbf{Z}_i of the model (9.3) can either represent individual agents or alternatively the average opinion of a subgroup. The latter perspective can be advantageous, for example, in designing methodology for robotic swarm activities where subgroups of robots needs to make consensus decisions, in studying cognitive control where the behavior of competing subpopulations of neurons determines task switching [143], and in modeling and investigating mechanisms that explain sociopolitical processes such as political polarization [120]. In this section we prove a sufficient condition for *cluster synchronization* of the opinions on the network with the opinion dynamics (9.3), in which the network trajectories converge to a lower-dimensional manifold on which agents within each cluster have identical opinions, whereas agents in different clusters can have different opinions.

The cluster synchronization problem has been extensively studied in dynamical systems with diffusive coupling, as in [163, 217]. More broadly, cluster synchronization has been linked to graph symmetries and graph structure called *external equitable partitions* [75, 154, 177, 188, 191]. In the following theorem we show that such a network structure constitutes a sufficient condition for a network of agents to form opinion clusters – see Appendix 9.7.3 for the proof.

Theorem 9.3.5 (Model Reduction with Opinion Clusters). *Consider N_c clusters with N_p agents in cluster p such that $\sum_{p=1}^{N_c} N_p = N_a$. Let \mathcal{I}_p be the set of indices for agents in cluster p . Assume for every $p = 1, \dots, N_c$: 1) $u_i = \hat{u}_p$, $d_i = d_p$, $b_{ij} = b_{pj}$ for $i \in \mathcal{I}_p$; 2) within a cluster $\alpha_i = \bar{\alpha}_p$, $\gamma_{ik} = \tilde{\alpha}_p$, $\beta_i = \bar{\beta}_p$, $\delta_{ik} = \tilde{\beta}_p$ for $i, k \in \mathcal{I}_p$, and $i \neq k$; 3) between clusters $\gamma_{ik} = \tilde{\gamma}_{ps}$, $\delta_{ik} = \tilde{\delta}_{ps}$ for $i \in \mathcal{I}_p$, $k \in \mathcal{I}_s$ $s = 1, \dots, N_c$ and $s \neq p$. Define bounded set $K_q \subset \mathbb{R}^{>0}$, $q = 1, 2$, as the image of the derivative of the saturating function S'_q of (9.3). If the following condition holds:*

$$\sup_{\kappa_1 \in K_1, \kappa_2 \in K_2} \left\{ -d_p + u_p \kappa_1 (\bar{\alpha}_p - \tilde{\alpha}_p) + u_p \kappa_2 (\bar{\beta}_p - \tilde{\beta}_p) \right\} < 0, \quad (9.7)$$

for all $p = 1, \dots, N_c$, then every trajectory of (9.3) converges exponentially to the $N_c(N_c - 1)$ -dimensional manifold

$$\mathcal{E} = \{ \mathbf{Z} \in V \mid z_{ij} = z_{kj} \quad \forall i, k \in \mathcal{I}_p, p = 1, \dots, N_c \}. \quad (9.8)$$

The dynamics on \mathcal{E} reduce to (9.3) with N_c agents with opinion states \hat{z}_{pj} , $p = 1, \dots, N_c$, and with

coupling weights

$$\hat{\alpha}_p = \bar{\alpha}_p + (N_p - 1)\tilde{\alpha}_p, \quad \hat{\gamma}_{ps} = N_s \tilde{\gamma}_{ps}, \quad (9.9a)$$

$$\hat{\beta}_p = \bar{\beta}_p + (N_p - 1)\tilde{\beta}_p, \quad \hat{\delta}_{ps} = N_s \tilde{\delta}_{ps}. \quad (9.9b)$$

Whenever conditions of Theorem 9.3.5 are met, the group of N_a agents will converge to a clustered group opinion state. This can happen for a broad class of interaction networks including an all-to-all network with interaction weights that all have the same sign. The sufficient condition can, for example, inform network design for technological systems where each of several groups must make a different collaborative decision. See Figure 9.3 for an illustration of opinion trajectories with two clusters, membership in which is defined by the network.

9.3.6 A Minimal Opinion Network Model

Several of the results characterizing opinion formation in (9.3) will be proved in the *homogeneous regime* defined by

$$\begin{aligned} b_{ij} &= 0, \quad d_i = d > 0, \quad u_i = u \geq 0, \quad \alpha_i = \alpha \geq 0 \\ \beta_i &= \beta \in \mathbb{R}, \quad \gamma_{ik} = \gamma a_{ik}, \quad \delta_{ik} = \delta a_{ik}, \quad A = [a_{ik}] \end{aligned} \quad (9.10)$$

where $\alpha, \beta, \gamma, \delta \in \mathbb{R}$ and $a_{ik} \in \{0, 1\}$, $a_{ii} = 0$ for all $i, k = 1, \dots, N_a$, $k \neq i$, so that A is an unweighted adjacency matrix without self-loops.

With this choice of parameters, the nonlinear model is minimal in the following sense. The matrix A with elements a_{ik} defines the influence network topology. The set of four interactions gains $\alpha, \beta, \gamma, \delta$ is minimal because in general there are four distinct types of arrows in a multi-option opinion network. The (global) attention parameter u and resistance parameter d tune an agent's attention to other agents' opinions and they jointly determine the occurrence of opinion-formation bifurcations, as we prove in Section 9.4.

We show that our model, even in the fully homogeneous regime, exhibits extremely rich and analytically provable opinion-formation behaviors. We further build upon the results proved for the homogeneous model to study, either analytically or numerically, the effects of heterogeneity and perturbations.

9.4 Agreement and disagreement opinion formation

We show key results on opinion formation for dynamics (9.3):

1. Opinion formation can be modeled as a bifurcation, an intrinsically nonlinear dynamical phenomenon. Opinions form rapidly through bifurcation-induced instabilities rather than slow linear integration of evidence. Opinions can form even in the absence of input, as long as attention (urgency or susceptibility, etc.) is sufficiently high.
2. The way opinions form at a bifurcation depends on the eigenstructure of the matrix $\Gamma - \Delta$ defined by (9.4).
3. In the homogenous regime defined by (9.10), cooperative agents ($\gamma > \delta$) always form agreement opinions, whereas under suitable assumptions on the eigenstructure of the adjacency matrix A , competitive agents ($\gamma < \delta$) always form disagreement opinions.
4. At the bifurcation, there are multiple stable solutions, and opinion formation breaks deadlock, that is, the situation in which every agent remains neutral, and therefore undecided, about all the options.
5. Near the bifurcation, opinion formation is ultra-sensitive to input.
6. Away from the bifurcation, opinion formation is robust to small heterogeneity in parameter values and small inputs.
7. In the absence of inputs, multistable agreement solutions *and* multistable disagreement solutions emerge generically at opinion-forming bifurcations.
8. In the presence of inputs, the opinion-forming bifurcation unfolds (i.e., multistability is partially or completely broken) in a such a way that the opinion states favored by inputs attract most of the initial conditions close to the bifurcation. The network structure governs the relative influence of inputs, which leads to a formal notion of centrality indices for agreement and disagreement.
9. Agreement and disagreement can *co-exist*, revealing the possibility of easy transition between them.
10. With sufficient symmetry, agreement specializes to consensus and disagreement to dissensus.

9.4.1 Agreement and Disagreement States

We say the agents *agree*, i.e., are in an *agreement state*, when $\text{sign}(z_{ij}) = \text{sign}(z_{kj})$ for all $i, k = 1, \dots, N_a$, $j = 1, \dots, N_o$. This means that all agents unanimously favor or disfavor each of the options, although they may differ on the magnitude of their opinions. Agreement specializes to *consensus* when $\mathbf{Z}_i = \mathbf{Z}_k$ for all $i, k = 1, \dots, N_a$. We say the agents *disagree*, i.e., are in a *disagreement state*, when $\text{sign}(z_{ij}) \neq \text{sign}(z_{kj})$ for at least one pair of agents $i, k = 1, \dots, N_a$, $i \neq k$, and at least one option j . Disagreement specializes to *dissensus* when the average opinion of the group is neutral, i.e., $\sum_{i=1}^{N_a} \mathbf{Z}_i = 0$. For a network with clustered coupling (Theorem 9.3.5), *clustered consensus*, *dissensus*, *agreement*, and *disagreement* are defined by applying the definitions to the average states of each cluster \hat{z}_{pj} in the reduced model.

Remark 9.4.1. *In the presence of nonzero inputs b_{ij} , agents will generically have nonzero opinions about options as follows from Lemma 9.3.3. For realistic applications, small opinions formed in a linear response to inputs should be distinguished from large opinions which arise from a nonlinear response. To make this distinction we say agents are **opinionated** when their opinions are large, and **unopinionated** when their opinions are close to zero. In this paper we keep this distinction qualitative. A precise bound between opinionated and unopinionated magnitudes depends on the application and can be defined when necessary.*

9.4.2 Opinions Form through a Bifurcation

Steady-state bifurcations of the opinion dynamics (9.3) result in large opinions even for zero input. Theorem 9.4.2, proved in Appendix 9.7.4, provides sufficient conditions under which opinions form through a bifurcation from the neutral equilibrium $\mathbf{Z} = 0$ and formulas to compute the kernel along which the bifurcation appears. Let interaction matrices Γ, Δ be as in (9.4).

Theorem 9.4.2 (Opinion Formation as a Bifurcation). *Consider model (9.3) with $b_{ij} = 0$, $d_i = d$, $u_i = u$, $\alpha_i = \alpha$, and $\beta_i = \beta$, for all $i = 1, \dots, N_a$. Let J be the Jacobian of the system evaluated at neutral equilibrium $\mathbf{Z} = 0$. Define λ to be the eigenvalue of $\Gamma - \Delta$ with largest real part. Assume that λ is real, $\alpha - \beta + \lambda > 0$, and that $\text{Re}[\mu] \neq \lambda$ for any eigenvalue $\mu \neq \lambda$ of $\Gamma - \Delta$. Then $\mathbf{Z} = 0$ is locally exponentially stable for $0 < u < u^*$, and unstable for $u > u^*$, with*

$$u^* = \frac{d}{\alpha - \beta + \lambda}. \quad (9.11)$$

If λ is simple², at $u = u^*$ an opinion-forming steady-state bifurcation happens along $\ker J = \mathbb{R}\{\mathbf{v}^*\} \otimes \mathbf{1}_{N_o}^\perp$, where \mathbf{v}^* is the right unit eigenvector associated to λ . More precisely, generically for each bifurcation branch, there exists $\mathbf{v}_{ax} \in \mathbf{1}_{N_o}^\perp$ such that the branch is tangent at $\mathbf{Z} = 0$ to the one-dimensional subspace $\mathbb{R}\{\mathbf{v}^* \otimes \mathbf{v}_{ax}\}$.

Remark 9.4.3. The vector \mathbf{v}_{ax} can be computed as the generator of the fixed-point subspace of an axial subgroup [83, Section 1.4] of the (irreducible) action of \mathfrak{S}_{N_o} on $\ker J$.

Theorem 9.4.2 reveals how agents can become opinionated even without input: opinions form when attention u is greater than threshold u^* . This means that deadlock can be avoided even when there is little or no evidence to distinguish among options. The value of the threshold is determined from the structure of the communication network. Additionally, from this result we can deduce how agreement and disagreement solutions are informed by the network structure. In particular, the equilibrium opinions of each agent near the bifurcation are directly proportional to the vector \mathbf{v}_{ax} , scaled by the entries of \mathbf{v}^* . When all of the entries of \mathbf{v}^* have the same sign, the agents will be in an agreement state. If \mathbf{v}^* contains mixed-sign entries, the agents will necessarily be in a disagreement state. This provides a straightforward connection between the spectral properties of the effective inter-agent communication graph $\Gamma - \Delta$ and the opinion configurations that arise from the opinion dynamics (9.3). The entries of the vector \mathbf{v}_{ax} determine the relative preference associated to the various options. In the following corollary we show how in the homogeneous regime (9.10) Theorem 9.4.2 specializes to simple conditions for agreement and disagreement.

Corollary 9.4.4 (Agreement and Disagreement). *Consider model (9.3) with homogeneous parameters as in (9.10) on a strongly connected graph. Let $\lambda_{max} > 0$ be the largest real-part eigenvalue of A . Let $\lambda_{min} < 0$ be the smallest real-part eigenvalue of A . Assume λ_{min} is real, simple, and for all eigenvalues $\xi \neq \lambda_{min}$ of A , $\text{Re}[\xi] \neq \lambda_{min}$.*

*A. **Cooperative agents.** Suppose that $\gamma - \delta > 0$ and $\alpha - \beta + \lambda_{max}(\gamma - \delta) > 0$. Then the steady-state bifurcation predicted by Theorem 9.4.2 happens when attention $u = u^*$, where*

$$u^* := u_a = \frac{d}{\alpha - \beta + \lambda_{max}(\gamma - \delta)} \quad (9.12)$$

and close to bifurcation all the bifurcation branches are made of agreement solutions.

*B. **Competitive agents.** Suppose $\gamma - \delta < 0$ and that $\alpha - \beta + \lambda_{min}(\gamma - \delta) > 0$. Then the steady-state*

²This result can be generalized to networks for which λ is not simple.

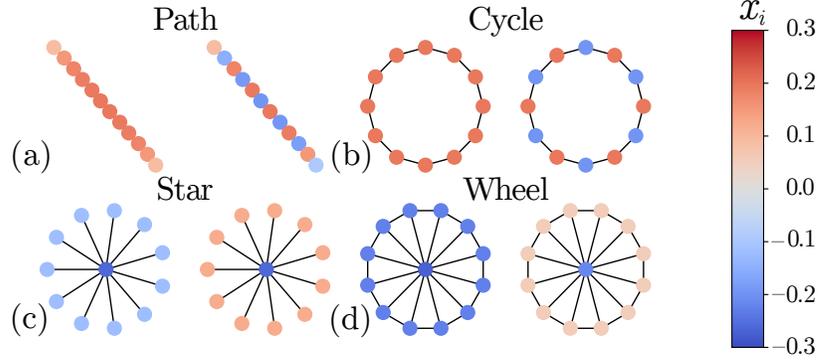


Figure 9.4: Adapted from [20]. Agreement (left) and disagreement (right) opinion configurations at steady state from simulation of two-option opinion dynamics (9.14) and four different undirected graph types, with attention u slightly above the critical value u^* in each case. Color of each node i corresponds to opinion x_i at $t = 500$. For all graphs $\gamma = 1.3$ (left) and $\gamma = -1.3$ (right), $d = 1$, $\alpha = 1.2$, and $b_i = 0$ for all $i = 1, \dots, N_a$. For the path and cycle graphs, $u = 0.31$, and for the star and wheel graphs, $u = 0.26$. Randomized initial opinions are drawn from the distribution $U(-1, 1)$.

bifurcation predicted by Theorem 9.4.2 happens when attention $u = u^*$, where

$$u^* := u_d = \frac{d}{\alpha - \beta + \lambda_{\min}(\gamma - \delta)}. \quad (9.13)$$

Moreover, whenever \mathbf{v}_{\min} , the right unitary eigenvector associated to λ_{\min} , has mixed-sign entries, close to bifurcation all the bifurcation branches are made of disagreement solutions.

We emphasize that the assumption about eigenvalues λ of Theorem 9.4.2 and λ_{\min} of Corollary 9.4.4 being simple often holds, and can be easily verified numerically for various graph structures. Furthermore, the eigenvector \mathbf{v}_{\min} of Corollary 9.4.4 typically has mixed-sign entries, and competition between agents therefore tends to result in network disagreement. For example, on undirected networks \mathbf{v}_{\min} always has mixed-sign entries since \mathbf{v}_{\max} , the unitary right eigenvector associated to λ_{\max} , i.e., the Perron-Frobenius eigenvector, is positive and $\langle \mathbf{v}_{\max}, \mathbf{v}_{\min} \rangle = 0$. For example, see Figure 9.4 for patterns of agreement and disagreement solutions for $N_o = 2$ and several representative undirected graphs.

An important feature of the opinion dynamics (9.3) is the multistability of opinion configurations at the bifurcations described by Theorem 9.4.2 and Corollary 9.4.4. When agents cooperate and $\ker J$ is made of agreement vectors, if agreement in favor of one option is stable then agreement in favor of each other option is stable, and likewise for disagreement solutions. There is a deadlock when $u < u_a$ ($u < u_d$) and breaking of deadlock when $u > u_a$ ($u > u_d$).

At the bifurcation the linearization is singular, and the model is ultra-sensitive at transition from neutral to opinionated. Even infinitesimal perturbations (e.g., tiny difference in option values) are

sufficient to destroy multistability at bifurcation by selecting a subset of stable equilibria (e.g., those corresponding to higher-valued options), a phenomenon known as forced-symmetry breaking and widely exploited in nonlinear decision-making models [87, 156, 168].

Generically, stable equilibria that appear at the bifurcation are hyperbolic, and thus they and their basin of attraction are robust to perturbations, a key property that ensures stability of opinion formation despite (sufficiently small) changes in inputs, heterogeneity in parameters, and perturbations in the communication network. Robustness bounds can be derived using methods like those used for Hopfield networks in [106]. Robust multistability of equilibria gives the opinion-forming process hysteresis, and thus memory, between different opinion states: once an opinion is formed in favor of an option, a large change in the inputs is necessary for a switch.

Remark 9.4.5. *Under the clustering conditions of Theorem 9.3.5, we can apply Theorem 9.4.2 and Corollary 9.4.4 with N_c agents and coupling parameters defined by (9.9).*

Remark 9.4.6 (Mode Interaction and Coexistence of Agreement and Disagreement). *When $\gamma = \delta$, there is **mode interaction** [82], and agreement and disagreement bifurcations appear at the same critical value of u . This regime is especially interesting because it allows for co-existence of stable agreement and disagreement solutions, which can result in agents easily transitioning between the two in response to changing conditions. However, additional primary solution branches not captured by the analysis presented here can appear in this regime; we leave exploring this regime to future work.*

9.4.3 Patterns of Opinion Formation for Two Options

We examine the ultra-sensitivity of the network opinion dynamics to inputs or biases of individual agents when operating near its bifurcation point. We consider the two-option opinion dynamics (9.6) with homogeneous parameters (9.10), relaxing the assumption of zero inputs:

$$\dot{x}_i = -dx_i + u\hat{S}_1 \left(\alpha x_i + \gamma \sum_{k=1, k \neq i}^{N_a} a_{ik} x_k \right) + b_i. \quad (9.14)$$

The next corollary follows from Corollary 9.4.4 and [21, Theorems IV.1 and IV.2]. It recognizes the opinion-forming bifurcations of (9.14) as agreement and disagreement pitchfork bifurcations and predicts their unfolding in response to distributed inputs as a function of network structure. In other words, it predicts the location of the two symmetric agreement (or disagreement) solutions and how the input-driven unfolding selects one of the two solutions (see Figure 9.5).

Corollary 9.4.7. Consider (9.14) and suppose that adjacency matrix A is irreducible, i.e., the associated graph is strongly connected. Let $\lambda_{max} > 0$ be the largest real-part eigenvalue of A , i.e. the Perron-Frobenius eigenvalue, with associated unitary positive right eigenvector \mathbf{v}_{max} and unitary positive left eigenvector \mathbf{w}_{max} . Let $\lambda_{min} < 0$ be the smallest real-part eigenvalue of A . Assume λ_{min} is real, simple, and for all eigenvalues $\xi \neq \lambda_{min}$ of A , $\text{Re}[\xi] \neq \lambda_{min}$. Let \mathbf{v}_{min} and \mathbf{w}_{min} be the right and left unitary eigenvectors associated to λ_{min} with $\langle \mathbf{v}_{min}, \mathbf{w}_{min} \rangle > 0$.

A. Cooperative agents. If $\gamma > 0$, inputs satisfy $\langle \mathbf{b}, \mathbf{w}_{max} \rangle = 0$, and $\alpha + \lambda_{max}\gamma > 0$, model (9.14) undergoes a supercritical pitchfork bifurcation for $u = u^* = \frac{d}{\alpha + \lambda_{max}\gamma}$ at which opinion-forming bifurcation branches emerge from $\mathbf{x} = 0$. The associated bifurcation branches are tangent at $\mathbf{x} = 0$ to $\mathbb{R}\{\mathbf{v}_{max}\}$. The pitchfork unfolds in the direction given by $\langle \mathbf{b}, \mathbf{w}_{max} \rangle$, i.e., if $\langle \mathbf{b}, \mathbf{w}_{max} \rangle > 0 (< 0)$, then the only stable equilibrium \mathbf{x}^* for u close to u^* satisfies $\langle \mathbf{x}^*, \mathbf{v}_{max} \rangle > 0 (< 0)$.

B. Competitive agents. If $\gamma < 0$, inputs satisfy $\langle \mathbf{b}, \mathbf{w}_{min} \rangle = 0$, and $\alpha + \lambda_{min}\gamma > 0$, model (9.14) undergoes a supercritical pitchfork bifurcation for $u = u^* = \frac{d}{\alpha + \lambda_{min}\gamma}$ at which opinion-forming bifurcation branches emerge from $\mathbf{x} = 0$. The associated bifurcation branches are tangent at $\mathbf{x} = 0$ to $\mathbb{R}\{\mathbf{v}_{min}\}$. The pitchfork unfolds in the direction given by $\langle \mathbf{b}, \mathbf{w}_{min} \rangle$, i.e., if $\langle \mathbf{b}, \mathbf{w}_{min} \rangle > 0 (< 0)$, then the only stable equilibrium \mathbf{x}^* for u close to u^* satisfies $\langle \mathbf{x}^*, \mathbf{v}_{min} \rangle > 0 (< 0)$.

Remark 9.4.8. For (9.5) with homogeneous parameters (9.10) an analogous result to Corollary (9.4.7) holds, except with $u^* = \frac{d}{\alpha - \beta + \lambda_{max/min}(\gamma - \delta)}$.

The symmetric opinion-forming pitchfork bifurcation predicted by Corollary 9.4.7 in the case of trivial or balanced inputs $\langle \mathbf{b}, \mathbf{w}_{max/min} \rangle = 0$ constitutes the simplest instance of multi-stability (bistability in this case) between different possible equilibrium opinion states (see Figure 9.5 left for the disagreement case and [21, Figure 1] for the identical figure in the agreement case). For u greater than critical value u^* (the bifurcation point), the group of agents can converge to either of the two stable opinion states depending on initial conditions as well as unmodelled uncertainties and disturbances.

In the agreement regime, solutions on the upper branch correspond to agents agreeing on option 1 and on the lower branch to agents agreeing on option 2. In the disagreement regime, solutions on the upper branch correspond to one subgroup favoring option 1 and the second subgroup favoring option 2 and the lower branch to the first subgroup favoring option 2 and the second subgroup favoring option 1. Both the sign and relative magnitudes of the agent opinions are predicted by \mathbf{v}_{max} in the agreement regime and \mathbf{v}_{min} in the disagreement regime – see Figure 9.4 for an illustration for four types of graphs. Observe that for the highly symmetric cycle graph, the group splits evenly in the

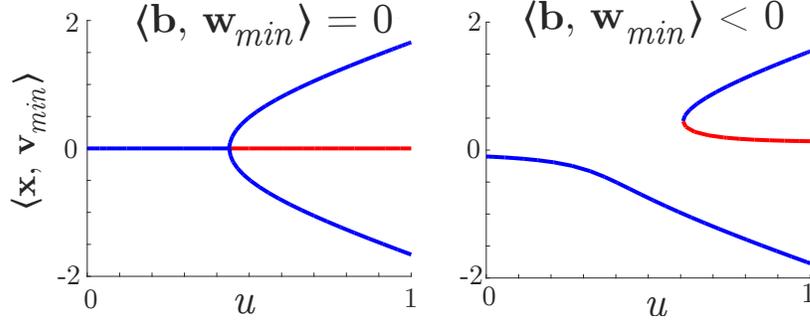


Figure 9.5: Bifurcation diagrams showing the symmetric pitchfork bifurcation (left) and its unfolding (right) for two-option opinion dynamics (9.14) with $d = \alpha = 1$ in the disagreement regime ($\gamma = -1$) for three agents communicating over an undirected line graph. Blue (red) curves represent stable (unstable) equilibria. The vertical axis is the projection of the system equilibria \mathbf{x} onto \mathbf{w}_{min} ($\mathbf{w}_{min} = \mathbf{v}_{min}$ since the graph is undirected). Left: $\mathbf{b} = (0.2, 0, -0.2)$; right: $\mathbf{b} = -0.1\mathbf{w}_{min} + (0.2, 0, -0.2)$. Bifurcation diagrams generated with help of MatCont [47]. In the agreement regime, the diagrams look qualitatively the same with $\mathbf{w}_{min}/\mathbf{v}_{min}$ replaced with $\mathbf{w}_{max}/\mathbf{v}_{max}$ and \mathbf{b} modified appropriately (see Figure 1 in [21]).

disagreement case, whereas in the star and wheel graphs, the center node disagrees with all of the peripheral nodes. These results are easily predicted using well-known results on the eigenstructure of the adjacency matrix for these graphs. See [20] for details.

The symmetric pitchfork unfolds (Figure 9.5 right) such that only one solution (predicted by the sign $\langle \mathbf{b}, \mathbf{w}_{max/min} \rangle$) is stable close to the symmetric bifurcation point; this follows from unfolding theory for a pitchfork bifurcation ([82, Chapter I]). For larger values of the attention parameter, the other solution also regains stability in a saddle-node bifurcation but the input-driven asymmetry is still reflected in the relative sizes of the basin of attraction of the two solutions. The left eigenvectors of the adjacency matrix $\mathbf{w}_{max/min}$ define agreement/disagreement *centrality indices* because the unfolding formula $\langle \mathbf{b}, \mathbf{w}_{max/min} \rangle \leq 0$ implies that the larger $[\mathbf{w}_{max/min}]_i$ the larger the effect of a nonzero input b_i on the agreement/disagreement pitchfork unfolding. Agreement and disagreement centrality indices can thus naturally be used to control opinion forming behavior via distributed inputs. By augmenting our opinion dynamics with an attention feedback mechanism, these centrality indices determine distributed thresholds for the triggering of opinion cascade, as illustrated in the next section (see also [67] for numerical illustrations on large random graphs with $N_o > 2$ and application to task allocation in robot swarms). All the results in this section generalize to the case $N_o > 2$. This generalization requires the computation of the vector v_{ax} appearing in Theorem 9.4.2 using equivariant bifurcation theory methods (see Remark 9.4.3), a direction that we leave for future extensions of this work.

9.4.4 Consensus and Dissensus Generic for Transitive Symmetry

In Section 9.4.2 we have shown how graph structure can inform what types of opinion configurations arise in the group. Here we consider, for the homogeneous regime (9.10), how the presence of symmetry in the communication graph can further constrain opinion configurations. We show how consensus and dissensus emerge for dynamics (9.3) with two examples of transitive symmetry. We first introduce a few technical definitions from group theory and equivariant bifurcation theory.

Let \mathfrak{G} be a compact Lie group acting on \mathbb{R}^n . Consider a dynamical system $\dot{\mathbf{x}} = \mathbf{h}(\mathbf{x})$ where $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{h} : \mathbb{R}^n \rightarrow \mathbb{R}^n$. Then $\rho \in \mathfrak{G}$ is a *symmetry* of the system, equivalently \mathbf{h} is ρ -*equivariant*, if $\rho\mathbf{h}(\mathbf{x}) = \mathbf{h}(\rho\mathbf{x})$. If \mathbf{h} is ρ -equivariant for all $\rho \in \mathfrak{G}$, then \mathbf{h} is \mathfrak{G} -*equivariant* [83]. \mathfrak{G} -equivariance means elements of symmetry group \mathfrak{G} send solutions to solutions.

The compact Lie group associated with permutation symmetries of n objects is the *symmetric group on n symbols* \mathfrak{S}_n , which is the set of all bijections of $\Omega_n := \{1, \dots, n\}$ to itself (i.e., all permutations of ordered sets of n elements). The opinion dynamics (9.3) with homogeneous parameters (9.10) and all-to-all coupling are *maximally symmetric*, i.e. $(\mathfrak{S}_{N_o} \times \mathfrak{S}_{N_a})$ -equivariant, where elements of \mathfrak{S}_{N_a} permute the N_a -element set of agents and elements of \mathfrak{S}_{N_o} permute the N_o -element set of options [68]. Maximally symmetric opinion dynamics are unchanged under any permutation of agents or options.

A subgroup $\mathfrak{G}_n \subset \mathfrak{S}_n$ is *transitive* if the orbit $\mathfrak{G}_n(i) = \{\rho(i), \rho \in \mathfrak{G}_n\} = \Omega$, for some (and thus all) $i \in \Omega$. $(\mathfrak{S}_{N_o} \times \mathfrak{S}_{N_a})$ -equivariant opinion dynamics, with transitive \mathfrak{G}_{N_a} , are still *highly symmetric* since any pair of agents, while not necessarily interchangeable by arbitrary permutations, can be mapped into each other by the symmetry group action. The following are examples of transitive subgroups of \mathfrak{S}_{N_a} :

- \mathfrak{D}_{N_a} , dihedral group of order N_a ; symmetries correspond to N_a rotations and N_a reflections. \mathfrak{D}_{N_a} -equivariant opinion dynamics are unchanged if agents are permuted by a rotation or a reflection, e.g., if agents communicate over a network defined by an undirected cycle.
- \mathfrak{Z}_{N_a} , cyclic group of order N_a ; symmetries correspond to N_a rotations (and no reflections). \mathfrak{Z}_{N_a} -equivariant opinion dynamics are unchanged if agents are permuted by a rotation, e.g., if agents communicate over a network defined by a directed cycle.

The system opinion state space decomposes as $V = W_c \oplus W_d$, where W_c is the *multi-option consensus space* defined as

$$W_c = \{(\mathbf{Z}_1, \dots, \mathbf{Z}_{N_a}) \mid \mathbf{Z}_i = \tilde{\mathbf{Z}} \in \mathbf{1}_{N_o}^\perp, \forall i\}, \quad (9.15)$$

and W_d is the *multi-option dissensus space* defined as

$$W_d = \{(\mathbf{Z}_1, \dots, \mathbf{Z}_{N_a}) \mid \mathbf{Z}_1 + \dots + \mathbf{Z}_{N_a} = \mathbf{0}\}. \quad (9.16)$$

On the consensus space W_c , agents have identical opinions. On the dissensus space W_d , agent opinions are balanced over the options such that the average opinion of the group is neutral.

Model-independent results [68, Theorem 4.6 and Remark 4.7] ensure that, in the presence of transitive symmetry, $\ker J = W_c$ or $\ker J = W_d$. I.e., if (9.3) is symmetric with respect to a group Γ_a that acts by swapping the agent indices transitively, then generically $\ker J = W_c$ or $\ker J = W_d$. In the homogeneous regime (9.10), agent symmetry of (9.3) is fully determined by A as proved in the following proposition for the maximally symmetric case $\mathfrak{G}_a = \mathfrak{S}_{N_a}$ and the highly symmetric case $\mathfrak{G}_a = \mathfrak{D}_{N_a}$ (see Appendix 9.7.5 for proof). The same result holds, with similar proof, for other transitive agent symmetries, e.g., $\mathfrak{G}_a = \mathfrak{Z}_{N_a}$.

Proposition 9.4.9. *Consider model (9.3) in the homogeneous regime defined by (9.10). Then the following hold true:*

1. *Model (9.3) is $(\mathfrak{S}_{N_o} \times \mathfrak{S}_{N_a})$ -equivariant if and only if A is the adjacency matrix of an all-to-all graph;*
2. *If A is the adjacency matrix of an undirected cycle graph, then model (9.3) is $(\mathfrak{S}_{N_o} \times \mathfrak{D}_{N_a})$ -equivariant.*

Remark 9.4.10. *More generally, the symmetry group of the opinion dynamics is determined by the automorphism group of the graph associated to A . The proof follows as for Proposition 9.4.9.*

The next corollary follows from Theorem 9.4.2 and [68, Theorem 4.6 and Remark 4.7]. The two types of opinion-formation behaviors proved in this corollary, i.e., consensus for cooperative agents and dissensus for competitive agents, respectively, constitute an opinion-formation analogue of consensus and balanced (split) states in coupled phase oscillators (see, e.g., [153, 183, 196]).

Corollary 9.4.11 (Consensus from Cooperation and Dissensus from Competition). *Consider model (9.3) in the homogeneous regime (9.10). Suppose that the graph associated to adjacency matrix A is either all-to-all or an undirected cycle. Let u_a and u_d be defined by (9.12) and (9.13).*

A. Cooperative agents and consensus. *If agents are cooperative ($\gamma - \delta > 0$), then opinion formation appears as a bifurcation along the consensus space at $u = u_a$ with $\lambda_{max} = N_a - 1$ for the all-to-all case and $\lambda_{max} = 2$ for the cycle case.*

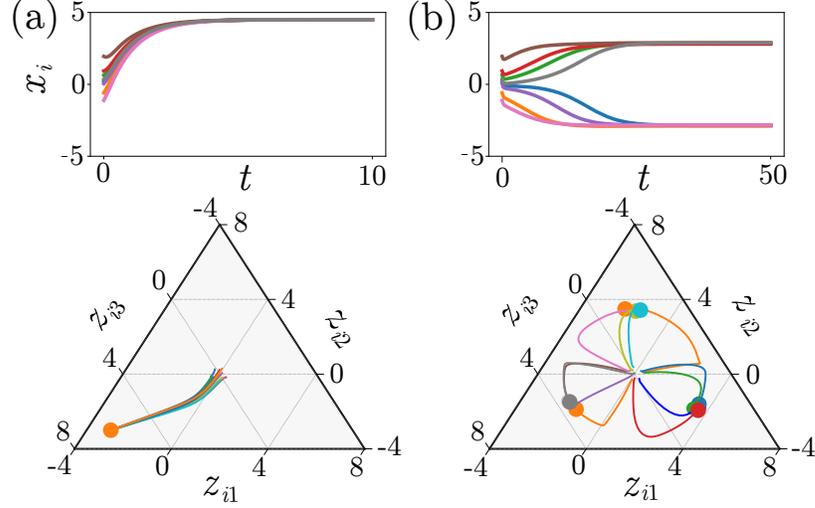


Figure 9.6: Simulations for $N_o = 2$ options and $N_a = 8$ agents (top) and $N_o = 3$ options and $N_a = 12$ agents (bottom) with $A = \mathbf{11}^T - \mathcal{I}$. Opinions form (a) consensus when agents are cooperative: $\gamma = 0.2$, $\delta = -0.1$; (b) dissensus when agents are competitive: $\gamma = -0.1$, $\delta = 0.2$. In each plot, $\alpha = 0.2$, $\beta = 0.1$, $d = 1$, $u = 3$, $\hat{b} = 0$, and random initial conditions are the same. Communication weights $\alpha, \beta, \gamma, \delta$ were perturbed with small random additive perturbations drawn from a normal distribution with variance (a) 0.01, (b) 0.001. Ternary plots for three options generated with the help of [54].

B. Competitive agents and dissensus. If agents are competitive ($\gamma - \delta < 0$), then opinion formation appears as a bifurcation along the dissensus space at $u = u_d$ with $\lambda_{min} = -1$ for the all-to-all case, $\lambda_{min} = -2$ for the cycle case, when N_a is even, and $\lambda_{min} = 2 \cos(\pi(N_a - 1)/N_a)$, when N_a is odd.

As an illustration of Corollary 9.4.11, representative consensus and dissensus trajectories of the opinion dynamics for two and three options on all-to-all graphs are shown in Fig. 9.6.

Remark 9.4.12 (Stability of Consensus and Dissensus). *Consensus and dissensus solution branches predicted for the symmetric networks in Corollary 9.4.11 are a consequence of the Equivariant Branching Lemma [83, Section 1.4], and are made of hyperbolic equilibria. Their stability can be proved using the tools in [84, Section XIII.4] and [83, Section 2.3].*

9.5 Attention dynamics and tunable sensitivity

We have established that existence of agreement and disagreement equilibria and multistability of opinion formation outcomes arise from bifurcations of the general opinion dynamic model (9.3). In this section we explore how *ultra-sensitivity to inputs* b_{ij} , *robustness to changes in inputs*, and *opinion cascade dynamics* also arise from bifurcations. With the addition of dynamic state feedback for model parameters in (9.3), the opinion formation process can reliably amplify arbitrarily small

inputs b_{ij} , reject small changes in input as unwanted disturbance, facilitate an opinion cascade even if only one agent gets an input, and enable groups to move easily between agreement and disagreement. The choice of feedback design parameters determine implicit thresholds that make all of these behaviors tunable.

The addition of dynamic state feedback for parameters in our model is similar in spirit to the extension of the linear weighted-average model with nonlinear state-feedback update rules for the coupling gains, as in bounded confidence models [44, 93, 94] and biased assimilation models [43, 218]. However, our motivation, rather than to capture a specific sociological phenomenon, is to make our model adaptable to inputs and flexibly responsive to changing environments. This is achieved by ensuring tunable sensitivity of opinion formation to inputs. We illustrate our ideas and prove our results for the case $N_o = 2$. The multi-option extension is left for future work.

9.5.1 Dynamic State Feedback Law for Attention

It is natural to consider each agent's attention u_i in (9.3) as a quantity that evolves in time in response to signals from others or from the environment ([87, 226]). This might happen when the agents delay making a collective decision until some task-relevant signal is detected on the network. To study this, we augment the opinion dynamics with an attention update

$$\tau_u \dot{u}_i = -u_i + f_{u,i}(\mathbf{Z}), \quad (9.17)$$

where $\tau_u > 0$ is a time scale, which can be freely chosen, and $f_{u,i} : \mathbb{R}^{N_a N_o} \rightarrow \mathbb{R}$ is a state feedback law, which can take different functional forms depending on the application. To study how attention feedback dynamics can enable a small local signal to excite a cascade of opinions across the entire network, we define $f_{u,i}$ to drive agent i to increase its attention when its neighbors form a strong opinion about any option, i.e., agent i engages when it observes its neighbors engaging:

$$\tau_u \dot{u}_i = -u_i + S_u \left(\frac{1}{N_o} \sum_{k=1}^{N_a} \sum_{l=1}^{N_o} (\bar{a}_{ik} z_{kl})^2 \right). \quad (9.18)$$

S_u is a smooth saturating function, satisfying $S_u(0) = 0$, $S_u(y) \rightarrow \bar{u} > 0$ as $y \rightarrow \infty$, $S'_u(y) > 0$ for all $y \in \mathbb{R}$, and $S''_u(y) > 0$ for all $y > 0$. We define S_u as a Hill function

$$S_u(y) = \underline{u} + (\bar{u} - \underline{u}) \frac{y^n}{(y_{th})^n + y^n}, \quad (9.19)$$

where threshold $y_{th} > 0$ and $n > 0$. In (9.19) we constrain \bar{u} and \underline{u} such that $\bar{u} > u_c \geq \underline{u} > 0$, with $u_c = u_a$ (u_d) when $\gamma > 0$ (< 0) and u_a, u_d are defined by (9.12),(9.13). For the remainder of this section we consider the homogeneous regime (9.10), except for the u_i , which are heterogeneous, and for nonzero inputs. The attention coupling matrix \bar{A} with elements \bar{a}_{ik} can be distinct from the opinion coupling matrix A , but here we let $\bar{A} = A + \mathcal{I}_{N_a}$. For $N_o = 2$ the attention feedback dynamics (9.18) simplify to

$$\tau_u \dot{u}_i = -u_i + S_u \left(\sum_{k=1}^{N_a} (\bar{a}_{ik} x_k)^2 \right). \quad (9.20)$$

9.5.2 Tunable Sensitivity and Robustness for a Single Agent

In this section we first consider a single agent with dynamic opinions (9.5) and dynamic attention (9.20) with no neighbors, i.e., $a_{ik} = 0$ for all $k = 1, \dots, N_a$. As shown in Figures 9.7 and 9.8, the equilibria of the coupled opinion and attention dynamics can be graphically represented as the intersection of the x_i -nullcline $\{\dot{x}_i = 0\}$ (black solid) and u_i -nullcline $\{\dot{u}_i = 0\}$ (red dashed) in the (u_i, x_i) plane. Corollary 9.4.7 defines the shape of the x_i -nullcline as a pitchfork bifurcation which unfolds with nonzero input b_i , analogous to Figure 9.5.

For model (9.5),(9.20), define agent i to be *strongly opinionated* when its attention is close to its upper saturation value, i.e., $u_i \simeq \bar{u}$, and *weakly opinionated* when its attention is close to its lower saturation value, i.e., $u_i \simeq \underline{u}$. What we refer to as *tunable sensitivity of opinion formation to input b_i* can then be understood by comparing the plots of Figure 9.7, where the opinion trajectory for agent i is plotted on the left for $b_i = 0.5$ and on the right for $b_i = 1$. For the given parameters and $b_i = 0.5$, the nullclines intersect at three points in the positive half-plane. For unopinionated initial conditions, the opinion state is attracted to the point corresponding to a weakly opinionated equilibrium: agent i *rejects the input* $b_i = 0.5$ and does not form a strong opinion. For the same parameters and $b_i = 1$, the nullclines intersect at only one point, corresponding to a strongly opinionated equilibrium. Thus, for the same initial conditions, agent i *accepts the input* $b_i = 1$ and forms a strong opinion. The implicit *sensitivity threshold*³ that distinguishes rejected from accepted inputs can be tuned by using parameters n, y_{th} in (9.19). Changing their value changes the shape of the u_i -nullcline and thereby varies how strong of an input b_i is required to reduce the number of nullcline intersections from three to one, as in Figure 9.7.

Tunable robustness of opinion formation to changes in input b_i can be understood by comparing

³Quantifying the exact relationship between the design parameters in the saturation function (9.19) and the implicit thresholds described in this section is a straightforward but lengthy calculation, which involves taking implicit derivatives of the equilibria of the coupled system with respect to the design parameters. Due to space constraints we leave out this analysis here.

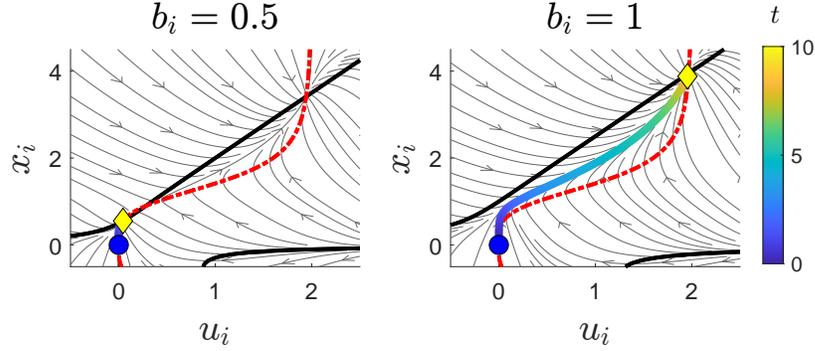


Figure 9.7: Sensitivity of opinion formation to input magnitude. (u_i, x_i) -phase plane and trajectories of (9.5),(9.20); $n = 2, y_{th} = 0.4, \alpha_i = 2, \beta_i = -1, \gamma_{ik} = \delta_{ik} = 0, d_i = 1, \tau_u = 1, \underline{u} = 0, \bar{u} = 2$ for $b_i = 0.5$ (left) and $b_i = 1$ (right). Initial state $(u_i(0), x_i(0)) = (0, 0)$ is a blue circle, and final state a yellow diamond. Nullclines of (9.5) are black solid and (9.20) are red dashed. Gray arrows show flow streamlines. Color scale is time.

the sequence of plots in the top and bottom halves of Figure 9.8. The plots on the left show agent i forming a strong opinion in the direction of the input $b_i = 1$. The plots on the right show what happens to agent i 's opinion when the input switches to $b_i = -1$, i.e., an input that is in opposition to the original input. In the top sequence, when $\bar{u} = 1$, agent i *accepts the change of input* and forms a strong opinion in the direction of the new input. In the bottom sequence, when $\bar{u} = 2.5$, agent i *rejects the change of input* and retains a strong opinion in the direction of the original input. The implicit *robustness threshold* that distinguishes rejected from accepted changes in input can be tuned by design parameter \bar{u} .

9.5.3 Opinion Cascades with Tunable Distributed Sensitivity

The following corollary shows that our feedback attention dynamics create a distributed threshold for the opinion dynamics below which the agents remain weakly opinionated and above which agents converge to a strongly opinionated equilibrium. The transition from a weakly opinionated to a strongly opinionated equilibrium in response to inputs is called an *opinion cascade*. The threshold is defined in terms of the inner product of the vector of inputs \mathbf{b} and suitable eigenvectors of the opinion network adjacency matrix. In other words, the threshold is distributed across the agents and the spectral properties of the adjacency matrix determine highly sensitive and weakly sensitive directions in the input vector space. As in Section 9.5.2 for single agents, the threshold can be tuned with parameters of the attention dynamics.

In the following theorem, we let $\lambda_{max}, \mathbf{w}_{max}$ and $\lambda_{min}, \mathbf{w}_{min}$ satisfy the assumptions of Corollary 9.4.7.⁴

⁴The proof of Theorem 9.5.1 follows from [21, Theorem V.3] and from geometric arguments based on implicit

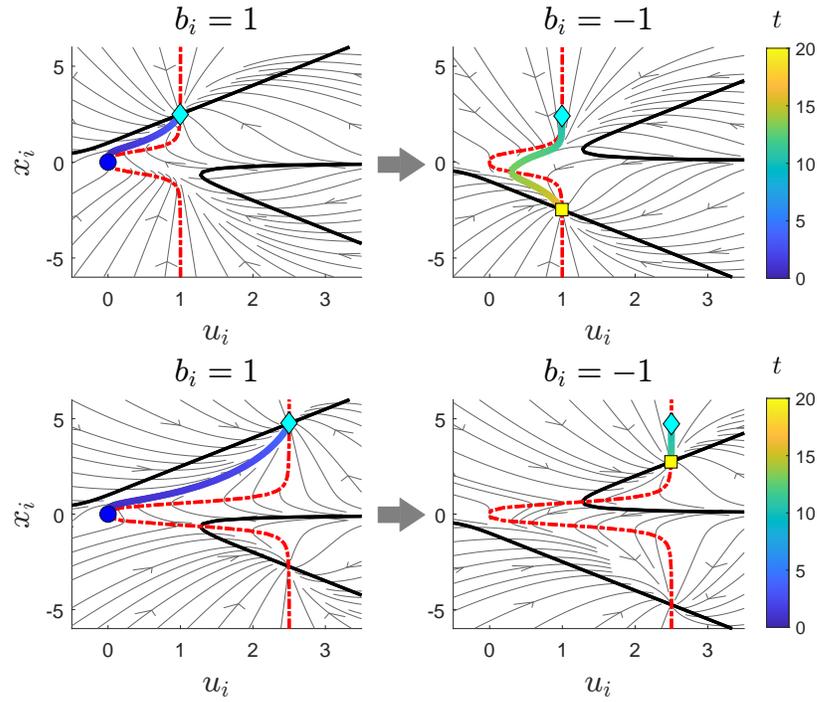


Figure 9.8: Robustness of opinion formation to changes in input. (u_i, x_i) -phase plane and trajectories of (9.5),(9.20); $n = 2, y_{th} = 0.4, \alpha_i = 2, \beta_i = -1, \gamma_{ik} = \delta_{ik} = 0, d_i = 1, \tau_u = 1, \underline{u} = 0$. (Left) Input is $b_i = 1$, initial state $(u_i(0), x_i(0)) = (0, 0)$ is a blue circle, and final state is a cyan diamond. (Right) Input changes to $b_i = -1$, initial state is final state on left and final state is yellow square. Top: $\bar{u} = 1$, and agent changes opinion in direction of new input. Bottom: $\bar{u} = 2.5$, and agent retains opinion in original direction. Nullclines, streamlines, and time are drawn as in Figure 9.7.

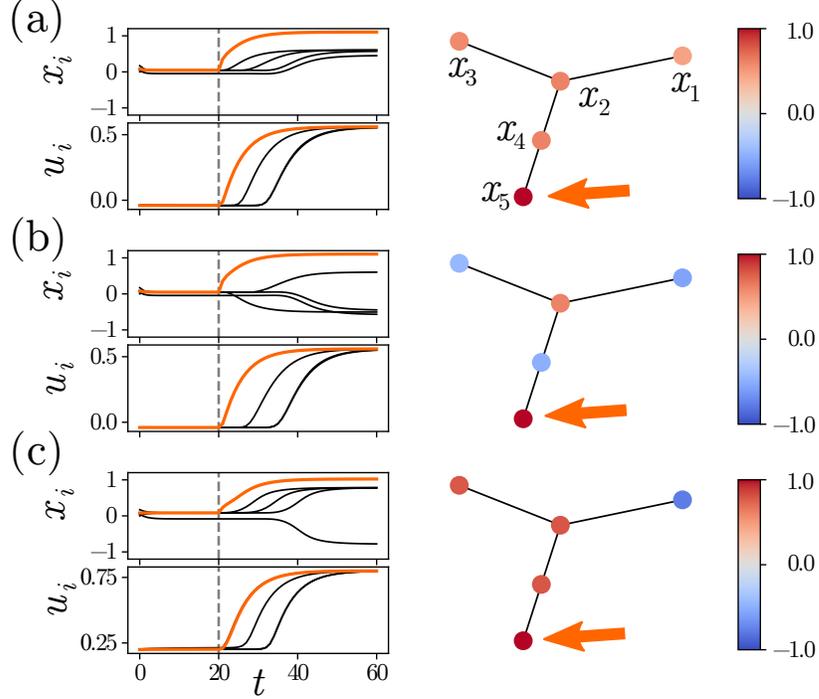


Figure 9.9: Opinion cascades with opinion and attention dynamics defined in Theorem 9.5.1. For $t < 20$, $\mathbf{b} = (-0.05, 0.05, 0.05, 0.05, 0.05)$ for all three simulations. At $t = 20$ the input to agent 5 (indicated by the arrow) increases to $b_5 = 0.25$, which triggers an opinion cascade on the network. Plots show opinion and attention trajectories of the agents with agent 5 in orange. Network diagrams on the right show the opinion strength of each agent at $t = 60$ of the simulation. (a) Agreement cascade with $\gamma = 1$; the network chooses the positive option following the informed agent. (b) Disagreement cascade with $\gamma = -1$; agents' opinions on the network disperse following the sign structure of \mathbf{v}_{min} . (c) Agents are coupled through the attention dynamics only (i.e. $\gamma = 0$); opinion cascade causes each agent to amplify its small input and commit to a strong opinion. Other parameters: $\alpha = 2$, $n = 3$, $y_{th} = 0.1$, $\tau_u = 5$, $d = 1$, $\bar{u} = u^* + 0.3$, $\underline{u} = u^* - 0.3$, $u_i(0) = \underline{u}$ for all $i = 1, \dots, N_a$. $\mathbf{x}(0)$ generated randomly from a uniform distribution between -0.2 and 0.2 ; the same initial condition was used for all three simulations.

Theorem 9.5.1. Consider the coupled system (9.6),(9.18) with $d_i = d$, $\alpha_i = \alpha$, and $\gamma_{ik} = \gamma a_{ik}$, where $A = [a_{ik}]$ is a symmetric and irreducible adjacency matrix. Let $u_c = \frac{d}{\alpha + \lambda_{max} \gamma}$, $\mathbf{w}_c = \mathbf{w}_{max}$ if $\gamma > 0$ and $u_c = \frac{d}{\alpha + \lambda_{min} \gamma}$, $\mathbf{w}_c = \mathbf{w}_{min}$ if $\gamma < 0$. There exists $\varepsilon > 0$ such that for $u_c > \underline{u}$, $y_{th} < \varepsilon$, and n sufficiently large, the following generically hold. There exists $p = p(y_{th}) > 0$ satisfying $\frac{\partial p}{\partial y_{th}} > 0$ such that, for $|\langle \mathbf{w}_c, \mathbf{b} \rangle| < p$, model (9.6),(9.18) possesses a weakly opinionated locally exponentially stable equilibrium. This equilibrium loses stability in a saddle-node bifurcation for $|\langle \mathbf{w}_c, \mathbf{b} \rangle| = p$. No weakly opinionated equilibria exist for $|\langle \mathbf{w}_c, \mathbf{b} \rangle| > p$ and all trajectories converge to a strongly opinionated agreement (disagreement) equilibrium for $\gamma > 0$ ($\gamma < 0$). For $\gamma = 0$, with $\alpha > 0$, the strongly opinionated equilibrium $(\mathbf{x}^*, \mathbf{u}^*)$ satisfies $\text{sign}(x_i^*) = \text{sign}(b_i)$.

Figure 9.9 illustrates the predictions of Theorem 9.5.1. It shows that the arrival of a suprathreshold input at $t = 20$ triggers an opinion cascade. Independently of the entries of the input vector \mathbf{b} , differentiation, similarly to the single-agent case of Section 9.5.2. It is omitted for space constraints.

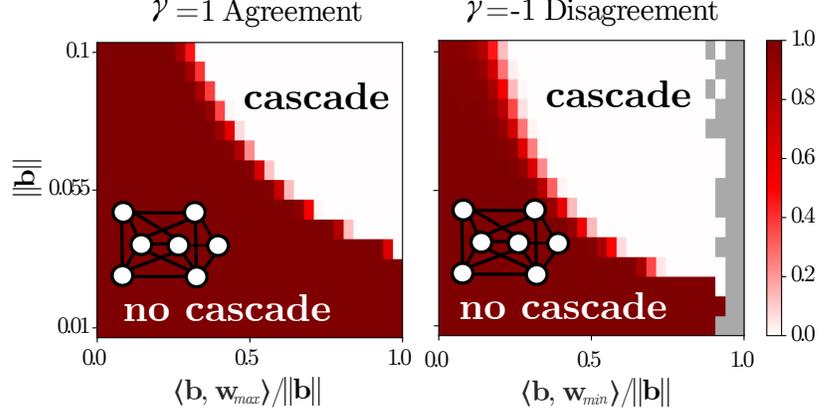


Figure 9.10: Adapted from [21]. Frequency of agreement (left) and disagreement (right) cascades for opinion and attention dynamics defined in Theorem 9.5.1. Color represents proportion of simulations in the given parameter range that did not result in a network cascade by $t = 500$. Dark red corresponds to no cascades, white to always a cascade, and grey to bins with no datapoints. Each plot shows the results of 1.5×10^5 distinct simulations with $\tau_u = 10$, $y_{th} = 0.2$, $\underline{u} = u_a - 0.01$ for $\gamma = 1$ (left) and $\underline{u} = u_d - 0.01$ for $\gamma = -1$ (right). For every simulation, initial conditions were $x_i = 0$, $u_i = 0$ for all $i = 1, \dots, N_a$ and b_i were drawn from $\mathcal{N}(0, 1)$ with \mathbf{b} normalized to a desired constant magnitude. 10000 simulations were performed for each constant input magnitude, with 15 magnitudes sampled uniformly spaced between 0 and 0.1.

the cascade goes to a strongly opinionated agreement equilibrium for $\gamma > 0$ (Figure 9.9a) and to a strongly opinionated disagreement equilibrium for $\gamma < 0$ (Figure 9.9b). Conversely, for $\gamma = 0$, the pattern of opinions at the strongly opinionated equilibrium is determined by the sign of the entries of the input vector. Figure 9.10 makes these observations more quantitative by showing the cascade threshold predicted by Theorem 9.5.1 as a joint function of the norm of the input vector and of the cosine of the angle between the input vector and the relevant eigenvector of the adjacency matrix. As predicted by the theorem, when the input vector is misaligned with respect to the adjacency matrix eigenvector, large-magnitude inputs are necessary to robustly trigger an opinion cascade. Conversely, as the two vectors align, an opinion cascade can be triggered with much smaller inputs.

9.5.4 Dynamics on weights: agreement-disagreement transitions

We illustrate how feedback dynamics of social influence weights in the two-option opinion dynamics (9.5) can be used to facilitate transitions between agreement and disagreement on the network. Suppose agents comprise two clusters of size N_1 and N_2 with index sets \mathcal{I}_1 and \mathcal{I}_2 . Let $b_i = b_p$ for $i \in \mathcal{I}_p$ and $\hat{x}_p = \frac{1}{N_p} \sum_{i \in \mathcal{I}_p} x_i$, where $p \in \{1, 2\}$. We define intra-cluster coupling as $\alpha_i = \gamma_{ik} = \alpha/N_p > 0$ and $\beta_i = \delta_{ik} = \beta/N_p < 0$, $l \neq j$, $p = 1, 2$, $d_i = d$ for all $i, k \in \mathcal{I}_p$, and agent attention dynamics by (9.20) with $\bar{a}_{ik} = 1$ for all i, k .

The influence network between the clusters is dynamic. Let $\gamma_{ik}(t) = \gamma_i(t)/N_s$, $\delta_{ik}(t) = \delta_i(t)/N_s$

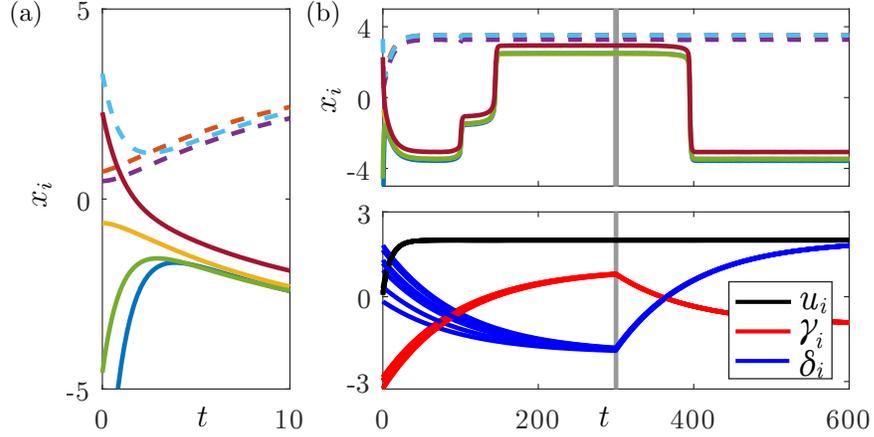


Figure 9.11: (a) Transient opinion trajectories settling to the clustered attractive manifold from random initial conditions in a simulation; (b) Full simulation. Top: opinion trajectories; Bottom: parameter trajectories. Seven agents form two clusters of sizes $N_1 = 3$ (dashed-line opinion trajectories), $N_2 = 4$ (solid-line opinion trajectories). $d = 1$, $\alpha = 1$, $\beta = -1$, $b_1 = 0.5$, $b_2 = -0.5$, $\tau_u = 10$, $\tau_\gamma = \tau_\delta = 100$, $\gamma_f = 2$, $\delta_f = 1$, $\underline{u} = 2$, $g_\gamma = g_\delta = 10$, $y_m = 1.5$. $x_i(0)$ are drawn from $\mathcal{N}(0, 2)$, $u_i(0)$ from $\mathcal{N}(0, 0.3)$, $\gamma_i(0)$ from $\mathcal{N}(-3, 0.3)$, and $\delta_i(0)$ from $\mathcal{N}(1, 0.3)$. d_i , α_i , β_i , b_i have additive perturbations drawn from $\mathcal{N}(0, 0.1)$ independently for each agent i . For $t < 300$, $\sigma = 1$ and for $t \geq 300$, $\sigma = -1$.

if $i \in \mathcal{I}_p, k \in \mathcal{I}_s, s \neq p$. Define feedback dynamics for inter-cluster coupling as

$$\tau_\gamma \dot{\gamma}_i = -\gamma_i + \sigma S_\gamma(\hat{x}_1 \hat{x}_2) \quad (9.21a)$$

$$\tau_\delta \dot{\delta}_i = -\delta_i - \sigma S_\delta(\hat{x}_1 \hat{x}_2) \quad (9.21b)$$

where $\sigma \in \{1, -1\}$, $\tau_\gamma, \tau_\delta > 0$ are time scales, $S_\gamma(y) = \gamma_f \tanh(g_\gamma y)$, $S_\delta(y) = \delta_f \tanh(g_\delta y)$, and $\gamma_f, \delta_f, g_\gamma, g_\delta > 0$.

The sign of design parameter σ in (9.21) determines whether the system tends towards agreement or disagreement, and switching the sign can reliably trigger a transition between agreement and disagreement. Figure 9.11 illustrates the opinion formation of 7 agents that form two clusters, one with 3 agents and the other with 4 agents. One cluster has input favoring option 1 and the second favoring option 2. Initially, $\gamma - \delta < 0$ on average and the clusters evolve to a dissensus state⁵ which is informed by the agents' inputs. However, because $\sigma = 1$, the two clusters eventually evolve towards a consensus state once $\gamma - \delta > 0$ despite the inputs favoring disagreement. At time $t = 300$, σ switches sign to $\sigma = -1$ and the two clusters evolve back towards a clustered dissensus state once $\gamma - \delta < 0$.

⁵Equilibria differ slightly from clustered consensus/dissensus due to parameter perturbations, simulated to illustrate robustness to uncertainty.

9.6 Final Remarks

Our opinion dynamics provide a new modeling framework for studying a variety of phenomena in which opinion formation is the governing behavior. In contrast to previous models, our approach focuses on the intrinsic nonlinear nature of opinion exchanges and thus on bifurcations as the key mechanism for analyzing and controlling opinion formation. Our model exhibits the flexibility, adaptability and robustness of natural opinion-forming systems, including deadlock-breaking and tunable sensitivity to changing inputs. A special instance of our model was motivated by modeling decision making in honeybee communities [87]. The analytical tractability of our model makes it possible to tackle its rich dynamical behavior constructively. This has allowed us to make novel predictions about the role of the opinion network structure in determining the emerging patterns of opinion formations and the sensitivity of the network to exogenous inputs, as well as to design adaptive feedback control laws for the model parameters.

The applicability of our model to real-world problems has recently been confirmed by our recent contributions in sociopolitical problems [120], the design of task-allocation algorithms in robot swarms [67], cognitive control [143], and game theory [158]. Other possible applications include decision making in biological and artificial neural networks, epidemiology and disease spread, and decision making in groups, from humans and robots to bacteria and animals on the move.

Acknowledgements

We acknowledge the contributions of Ayanna Matthews and Timothy Sorochkin to Figures 9.4 and 9.10, respectively.

9.7 Appendix

9.7.1 Extension to Heterogeneous Inter-option Coupling

In future applications of the opinion dynamics model (9.3) it may be useful to consider scenarios in which there is a heterogeneous level of influence between different options, i.e., in addition to the inter-agent interaction network there is an *inter-option interaction network*. Thus, we introduce the *adjacency tensor* with entries A_{ik}^{jl} that capture the weight of influence agent k 's opinion on option l

has on agent i 's opinion on option j . This leads to the generalized opinion dynamics:

$$\begin{aligned}\dot{\mathbf{Z}}_i &= P_0 \mathbf{F}_i(\mathbf{Z}) \\ F_{ij}(\mathbf{Z}) &= -d_i z_{ij} + u_i \sum_{l=1}^{N_o} S_l \left(\sum_{k=1}^{N_a} A_{ik}^{jl} z_{kl} \right) + b_{ij}.\end{aligned}$$

The model studied in this paper is recovered when S_l is S_1 for same-option interactions and S_2 for inter-option interactions, and $A_{ii}^{jj} = \alpha_i$, $A_{ik}^{jj} = \gamma_{ik}$, $A_{ii}^{jl} = \beta_i$, and $A_{ik}^{jl} = \delta_{ik}$ for all $i, k = 1, \dots, N_a$, $j, l = 1, \dots, N_o$, $i \neq k, j \neq l$.

9.7.2 Well-definedness of Model

The general model (9.3) is well defined since V is forward invariant for (9.3) (Lemma 9.7.1) and solutions are bounded (Theorem 9.7.2). Let $D = \text{diag}\{d_1, \dots, d_{N_a}\} \otimes \mathcal{I}_{N_o}$.

Lemma 9.7.1. *V is forward invariant for (9.3).*

Proof. For all $i = 1, \dots, N_a$, $\sum_{j=1}^{N_o} \dot{z}_{ij} = 0$, so if $z_{i1}(0) + \dots + z_{iN_o}(0) = 0$, $z_{i1}(t) + \dots + z_{iN_o}(t) = 0$ for all $t > 0$. \square

Theorem 9.7.2 (Boundedness). *Let \bar{U} be a compact subset of \mathbb{R} . There exists $R > 0$ such that, $\forall u_i, d_i, \alpha_i, \beta_i, \gamma_{ik}, \delta_{ik}, b_{ij} \in \bar{U}$, $i, k = 1, \dots, N_a$, $j, l = 1, \dots, N_o$, the set $V \cap \{|z_{ij}| \leq R, i = 1, \dots, N_a, j = 1, \dots, N_o\}$ is forward invariant for (9.3). This implies that the solutions $\mathbf{Z}(t)$ of the dynamics (9.3) are bounded for all time $t \geq 0$.*

Proof. By boundedness of $S_p(\cdot)$, there exists $\tilde{R} > 0$ such that, for all $u_i, d_i, \alpha_i, \beta_i, \gamma_{ik}, \delta_{ik}, b_{ij} \in \bar{U}$, $F_{ij}(\mathbf{Z}) = -d_i z_{ij} + C_{ij}(\mathbf{Z})$, with $|C_{ij}(\mathbf{Z})| \leq \tilde{R}$. For all $\mathbf{Z} \in V$, it holds that

$$\begin{aligned}\frac{d}{dt} \frac{1}{2} \|\mathbf{Z}\|^2 &= \sum_{i=1}^{N_a} \sum_{j=1}^{N_o} z_{ij} \dot{z}_{ij} = \sum_{i=1}^{N_a} \sum_{j=1}^{N_o} z_{ij} \left(-d_i z_{ij} + C_{ij}(\mathbf{Z}) + \frac{1}{N_o} \sum_{l=1}^{N_o} (d_{il} z_{il} - C_{il}(\mathbf{Z})) \right) \\ &= \mathbf{Z}^T D \mathbf{Z} + \sum_{i=1}^{N_a} \sum_{j=1}^{N_o} z_{ij} \left(C_{ij}(\mathbf{Z}) - \frac{1}{N_o} \sum_{l=1}^{N_o} C_{il}(\mathbf{Z}) \right) \leq \mathbf{Z}^T D \mathbf{Z} + N_a N_o \tilde{R} \|\mathbf{Z}\|\end{aligned}$$

where we have used $\sum_{j=1}^{N_o} z_{ij} = 0$ for all i . We compute

$$\mathbf{Z}^T D \mathbf{Z} = \sum_{i=1}^{N_a} \sum_{j=1}^{N_o} (-d_i z_{ij}^2) + \frac{1}{N_o} \sum_{i=1}^{N_a} \sum_{l=1}^{N_o} d_{il} z_{il} \left(\sum_{j=1}^{N_o} z_{ij} \right) = \sum_{i=1}^{N_a} \sum_{j=1}^{N_o} -d_i z_{ij}^2 \leq -\min_i \{d_i\} \|\mathbf{Z}\|^2.$$

Then, for all $\|\mathbf{Z}\| \geq \frac{N_a N_o \tilde{R}}{\min_i \{d_i\}}$, it follows that $\frac{d}{dt} \frac{1}{2} \|\mathbf{Z}\|^2 \leq -\|\mathbf{Z}\| \left(\min_i \{d_i\} \|\mathbf{Z}\| - N_a N_o \tilde{R} \right) \leq 0$. The result follows by [109, Theorem 4.18]. \square

These results connect the opinion vector $\mathbf{Z}_i \in \mathbf{1}_{N_o}^\perp$ to a simplex vector $\mathbf{y}_i = (y_{i1}, \dots, y_{iN_o})$, where $y_{ij} \geq 0$ for all i, j and $y_{i1} + \dots + y_{iN_o} = r$, $r > 0$, i.e. $\mathbf{y}_i \in \Delta$ where Δ is a $(N_o - 1)$ -dimensional simplex. Let $\mathcal{V} = \Delta \times \dots \times \Delta$.

Corollary 9.7.3. Mapping to the Simplex Product \mathcal{V} . *Given a bounded set $\bar{U} \subset \mathbb{R}$, assume $u_i, d_i, \alpha_i, \gamma_{ik}, \beta_i, \delta_{ik}, b_{ij} \in \bar{U}$, $i, k = 1, \dots, N_a$, $j, l = 1, \dots, N_o$. Then, the vector field of (9.3) can be mapped from the forward invariant region $V \cap \{|z_{ij}| \leq R, i = 1, \dots, N_a, j = 1, \dots, N_o\}$ to the product of simplex \mathcal{V} by the affine change of coordinates $L : V \cap \{|z_{ij}| \leq R, i = 1, \dots, N_a, j = 1, \dots, N_o\} \rightarrow \mathcal{V}$, $\mathbf{Z} \mapsto \frac{r}{N_o R} \mathbf{Z} + \frac{r}{N_o}$, $r > 0$.*

The simplex product space \mathcal{V} is often associated with models of opinion dynamics, e.g., in [105, 128, 186]. Under the mapping proposed in Corollary 9.7.3 or any other bijective mapping to the simplex product space (e.g. using the standard softmax function), the system state $\mathbf{y} = (\mathbf{y}_1, \dots, \mathbf{y}_{N_a}) \in \mathcal{V}$ can be interpreted as the *absolute opinions* of agents that have equal voting capacity in the collective decision [68], or as probabilities of choosing a particular option [158].

9.7.3 Proof of Theorem 9.3.5

Opinion dynamics (9.3) of agent $i \in \mathcal{I}_p$ are defined by

$$\begin{aligned} F_{ij}(\mathbf{Z}) = & -d_p z_{ij} + b_{pj} + \\ & u_p (S_1(\bar{\alpha}_p z_{ij} + \tilde{\alpha}_p \sum_{k \in \mathcal{I}_p \setminus \{i\}} z_{kj} + \sum_{s=1}^{N_c} \sum_{k \in \mathcal{I}_s} \tilde{\gamma}_{ps} z_{kj}) + \\ & \sum_{l=1}^{N_o} S_2(\bar{\beta}_p z_{il} + \tilde{\beta}_p \sum_{k \in \mathcal{I}_p \setminus \{i\}} z_{kl} + \sum_{s=1}^{N_c} \sum_{k \in \mathcal{I}_s} \tilde{\delta}_{ps} z_{kl})). \end{aligned} \quad (9.23)$$

Let $V_T(\mathbf{Z}) = \sum_{p=1}^{N_c} V_p(\mathbf{Z})$, $V_p(\mathbf{Z}) = \frac{1}{2} \sum_{i, k \in \mathcal{I}_p} \sum_{j=1}^{N_o} (z_{ij} - z_{kj})^2$. Let $F_{ij}(\mathbf{Z}) = -d_i z_{ij} + C_{ij}(\mathbf{Z})$. Then $\dot{V}_p(\mathbf{Z}) = -\sum_{i \in \mathcal{I}_p} \sum_{k \in \mathcal{I}_p} d_p (\mathbf{Z}_i - \mathbf{Z}_k)^T (\mathbf{Z}_i - \mathbf{Z}_k) + \sum_{i \in \mathcal{I}_p} \sum_{k \in \mathcal{I}_p} \sum_{j=1}^{N_o} (z_{ij} - z_{kj}) (C_{ij}(\mathbf{Z}) - C_{kj}(\mathbf{Z})) - \frac{1}{N_o} \sum_{i \in \mathcal{I}_p} \sum_{k \in \mathcal{I}_p} \sum_{j=1}^{N_o} \sum_{l=1}^{N_o} (z_{ij} - z_{kj}) (C_{il}(\mathbf{Z}) - C_{kl}(\mathbf{Z}))$. The last term is zero because $\sum_{j=1}^{N_o} z_{ij} = 0$ on V . By the Mean Value Theorem, we can write $C_{ij}(\mathbf{Z}) - C_{kj}(\mathbf{Z})$ in the second term as $u_p \left(\kappa_1 (\bar{\alpha}_p - \tilde{\alpha}_p) - \kappa_2 (\bar{\beta}_p - \tilde{\beta}_p) \right) (z_{ij} - z_{kj})^2$, where $\kappa_1 \in K_1$ and $\kappa_2 \in K_2$. Then we find that $\dot{V}_p(\mathbf{Z}) \leq \sup_{\kappa_1 \in K_1, \kappa_2 \in K_2} \left\{ -d_p + u_p \kappa_1 (\bar{\alpha}_p - \tilde{\alpha}_p) + u_p \kappa_2 (\bar{\beta}_p - \tilde{\beta}_p) \right\} 2V_p(\mathbf{Z})$. When (9.7) is satisfied, by LaSalle's invariance principle [109, Theorem 4.4] every trajectory of (9.3) converges asymptotically in time to \mathcal{E} , the largest invariant set of $V_T(\mathbf{Z}) = 0$. Let $\hat{z}_{pj} = z_{ij}$ for any $i \in \mathcal{I}_p$. The dynamics

(9.23) on \mathcal{E} reduce to (9.3) with $N_a = N_c$ and weights (9.9).

Remark 9.7.4. *This proof could be carried out using a group-theoretic approach outlined in [188], which would guarantee local stability of the clustered manifold. The Lyapunov function approach presented here provides a global stability guarantee.*

9.7.4 Proof of Theorem 9.4.2

$J = ((-d + u(\alpha - \beta))\mathcal{I}_{N_a} + u(\Gamma - \Delta)) \otimes P_0$, with eigenvalues $\xi_i \lambda_o$, for ξ_i an eigenvalue of $(-d + u(\alpha - \beta))\mathcal{I}_{N_a} + u(\Gamma - \Delta)$ and λ_o an eigenvalue of P_0 restricted to V . So $\lambda_o = 1$ and $\xi_i = -d + u(\alpha - \beta) + u\lambda_i$, λ_i , $i = 1, \dots, N_a$, an eigenvalue of $\Gamma - \Delta$. Thus, whenever $\alpha - \beta + \lambda > 0$, all eigenvalues of J are negative for $u < u^*$, zero is an eigenvalue of J for $u = u^*$ (with multiplicity $(N_o - 1)N_\lambda$, where N_λ is the multiplicity of λ), and there exist positive eigenvalues for $u > u^*$. The form of the eigenvectors of J corresponding to its zero eigenvalue for $u = u^*$ follows since the eigenvectors of the Kronecker product of matrices is the Kronecker product of the eigenvectors. For simple λ , the statement follows from the Equivariant Branching Lemma [83, Section 1.4].

9.7.5 Proof of Proposition 9.4.9

The proof of (1) follows analogously to that of [68, Theorem 2.5] with the additional coefficient d_i on the linear terms.

To prove (2), it is sufficient to show equivariance of the dynamics under the action of generators of $S_{N_o} \times D_{N_a}$. Element $\sigma \in S_{N_o}$ acts on V by permuting the elements of each agent's opinion \mathbf{Z}_i . Generators of S_{N_o} are N_o transpositions σ_j where each σ_j swaps adjacent elements j and $j + 1$ (or N_o and 1 when $j = N_o$). Let $\mathbf{F}_i(\mathbf{Z}) = (F_{i1}(\mathbf{Z}), \dots, F_{iN_o}(\mathbf{Z}))$ and observe that $\sigma_j \mathbf{F}_i(\mathbf{Z}) = (F_{i1}(\mathbf{Z}), \dots, F_{i(j+1)}(\mathbf{Z}), F_{ij}(\mathbf{Z}), \dots, F_{iN_o}(\mathbf{Z}))$. Computing $\mathbf{F}_i(\sigma_j \mathbf{Z})$, only F_{ij} and $F_{i(j+1)}$ are changed, with $F_{ij}(\sigma_j \mathbf{Z}) = -dz_{i(j+1)} + u \left(S_1(\alpha z_{i(j+1)} + \gamma z_{(i-1)(j+1)} + \gamma z_{(i+1)(j+1)}) + \sum_{l=1}^{N_o} S_2(\beta z_{il} + \delta z_{(i-1)(j+1)} + \delta z_{(i+1)(j+1)}) \right) + \hat{b}$. Thus, $\sigma_j \mathbf{F}_i(\mathbf{Z}) = \mathbf{F}_i(\sigma_j \mathbf{Z}) \forall j = 1, \dots, N_o$, $i = 1, \dots, N_a$, and the dynamics are equivariant under the action of S_{N_o} . $\rho \in D_{N_a}$ acts on V by permuting the order of the agent vectors \mathbf{Z}_i in the total system vector $\mathbf{Z} = (\mathbf{Z}_1, \dots, \mathbf{Z}_{N_a})$. The generators of D_{N_a} are the reflection element ρ_1 which reverses the order of elements in \mathbf{Z} , and a rotation ρ_2 which cycles forward the vector by one element, mapping each element i to $i+1$ (and N_a to 1). Let $\mathbf{F}(\mathbf{Z}) = (\mathbf{F}_1(\mathbf{Z}), \dots, \mathbf{F}_{N_a}(\mathbf{Z}))$ and observe that $\rho_1 \mathbf{F}(\mathbf{Z}) = (\mathbf{F}_{N_a}(\mathbf{Z}), \mathbf{F}_{N_a-1}(\mathbf{Z}), \dots, \mathbf{F}_2(\mathbf{Z}), \mathbf{F}_1(\mathbf{Z}))$ and $\rho_2 \mathbf{F}(\mathbf{Z}) = (\mathbf{F}_{N_a}(\mathbf{Z}), \mathbf{F}_1(\mathbf{Z}), \mathbf{F}_2(\mathbf{Z}), \dots, \mathbf{F}_{N_a-1}(\mathbf{Z}))$. For compactness we leave out the full expression for $F_{ij}(\rho_p \mathbf{Z})$.

Chapter 10

Patterns of Nonlinear Opinion Formation on Networks

ANASTASIA BIZYAEVA, AYANNA MATTHEWS, ALESSIO FRANCI,
AND NAOMI EHRICH LEONARD

When communicating agents form opinions about a set of possible options, agreement and disagreement are both possible outcomes. Depending on the context, either can be desirable or undesirable. We show that for nonlinear opinion dynamics on networks, and a variety of network structures, the spectral properties of the underlying adjacency matrix fully characterize the occurrence of either agreement or disagreement. We further show how the corresponding eigenvector centrality, as well as any symmetry in the network, informs the resulting patterns of opinion formation and agent sensitivity to input that triggers opinion cascades.

10.1 Introduction

Multi-agent systems that perform distributed control tasks in uncertain or dynamic contexts benefit when agents use network communications to form and change opinions about context-dependent options. For example, network communications can help autonomous multi-robot teams make better navigational choices among alternative routes or better allocation choices among alternative tasks. Mathematical models of opinion dynamics over networks are often used to bring a group to a desired opinion configuration. In a task-allocation context, agreement is not necessarily the only desirable opinion configuration, as sometimes agents are better off exploring different routes or doing different

tasks.

A general model of opinion dynamics for distributed agents on a network was recently introduced in [19, 68]. In this multi-agent multi-option model, real-valued opinions evolve in continuous time according to a nonlinear update rule that saturates network exchanges. A key feature of the model is the emergence of consensus and dissensus as equilibrium opinion configurations, even when agents are homogeneous, receive no input, and communicate over an all-to-all network. The emergence of consensus and dissensus depends on a small number of parameters that distinguish the interactions between agents as cooperative or competitive.

In [19] the behavior of the general model is examined with particular attention to all-to-all and vertex-transitive cycle network topologies. The work identifies parameter regimes that correspond to consensus and dissensus (a heterogeneous opinion configuration with neutral average opinion). In the present paper we take first steps to analyze opinion formation on other classes of networks. We consider opinion formation for two options and homogeneous agents that communicate over networks that include K -regular, bipartite, and strongly connected directed graphs. We examine how network structure influences the group outcome of the opinion formation process, and we prove that, generically, agreement and disagreement arise on these networks. We show that the parameter regimes associated with consensus and dissensus for complete graphs in [19] correspond precisely to agreement and disagreement regimes for more general networks.

The engineering literature on distributed opinion dynamics typically associates formation of opinions in continuous time to the spectral properties of the Laplacian matrix of the network graph. In linear consensus protocols, the governing equations take on the structure of the graph Laplacian, and consensus is achieved as opinions converge to its kernel [153]. Analogous distributed Laplacian schemes are considered with antagonistic (signed) interconnections in, e.g., [5, 127, 184]. When graphs are structurally balanced, i.e., when the signed graph Laplacian has a zero eigenvalue [5], such schemes give rise to clustered disagreement on the network. A nonlinear model of distributed consensus formation with saturated network interactions studied in [1, 61, 62, 70, 87] also relies on Laplacian-like structure of the governing equations, with each agent weighting its opinion state based on its in-degree on the network graph. Linearization of this model about the unopinionated state yields linear Laplacian consensus dynamics.

We study opinion dynamics of homogeneous agents with governing equations that do not necessarily have a Laplacian structure. We show how spectral properties of the *adjacency matrix* of the underlying graph, rather than those of its Laplacian, play a key role in characterizing the opinion formation process. In [61] the spectral properties of the adjacency matrix are tied to a necessary and

sufficient condition for existence of multi-stable consensus equilibria. However, unlike our model, the model in [61] relies on a Laplacian structure, and so the dynamics do not inherit the full symmetry properties of the network graph. For our model, we show that the largest and smallest eigenvalues of the adjacency matrix determine if nonzero opinions form, and the associated eigenspaces select the sign and relative magnitude of resultant opinions on the network. With the addition of dynamics on an attention parameter, the eigenvectors also determine which agents in the group are maximally sensitive to inputs. We illustrate how this can be used to trigger opinion cascades.

In Section 10.2 we present the opinion dynamics model of [19] on a network in the case of homogeneous agents and two options. In Section 10.3 we prove that agreement and disagreement equilibria, of which consensus and dissensus are special cases, arise on networks as bifurcations from an unopinionated equilibrium. We prove relationships between the magnitude of an agent’s equilibrium opinion and its centrality on the network in Sections 10.4, and to the symmetry properties of the network graph in Section 10.5. In Section 10.6 we introduce attention feedback dynamics and demonstrate how the group’s sensitivity to input relates to the adjacency eigenvectors. We conclude in Section 10.7 .

10.2 Opinion Dynamics Model

We study a nonlinear model of N_a homogeneous agents forming opinions about two options, a specialization of the general heterogeneous multi-agent, multi-option, model introduced in [19, 68]. The *opinion of each agent i* is captured by a real-valued variable $x_i \in \mathbb{R}$. When $x_i = 0$ agent i has a *neutral opinion*, and when $x_i > 0$ (< 0) agent i *favours* option A (option B). A greater magnitude $|x_i|$ corresponds to a stronger commitment of agent i to one of the options. We define a threshold $\vartheta > 0$ to formalize what is meant by “close”; the value of ϑ can be chosen for purposes of interpretation. We call an agent *opinionated* when $|x_i| > \vartheta > 0$, and *unopinionated* otherwise.

The *directed* graph $\tilde{G} = (V, \tilde{E})$ encodes which agents can communicate with which other agents. The set of vertices $V = \{1, \dots, N_a\}$ represents the set of N_a agents, and edges $\tilde{E} \subseteq V \times V$ represent interactions between agents. If edge $\tilde{e}_{ik} \in \tilde{E}$, then agent k is a *neighbor* of agent i . When inter-agent communication is bidirectional, G is *undirected*: if $\tilde{e}_{ik} \in \tilde{E}$, then $\tilde{e}_{ki} \in \tilde{E}$. The graph adjacency matrix $\tilde{A} \in \mathbb{R}^{N_a \times N_a}$ encodes interaction weights with element $\tilde{a}_{ik} \neq 0$ if and only if $\tilde{e}_{ik} \in \tilde{E}$. The sign of inter-agent weights determines whether agents are *cooperative* ($\tilde{a}_{ik} > 0$) or *competitive* ($\tilde{a}_{ik} < 0$). In this paper we specialize $\tilde{a}_{ik} \in \{0, 1\}$.

Let $\mathbf{x} = (x_1, \dots, x_{N_a}) \in \mathbb{R}^{N_a}$ be the *opinion state of the group*. When $\mathbf{x} = \mathbf{0}$, the group is in the

neutral state. When all agents are unopinionated, the group is in an *unopinionated state*. A pair of opinionated agents i, k *agree* if both share the same qualitative opinion state, i.e., $\text{sign}(x_i) = \text{sign}(x_k)$. When all agents agree, the group is in an *agreement state*. The *consensus state* is a special type of agreement state in which opinions are close in value, i.e. $|x_i - x_k| < \vartheta$ for all $i, k \in V$. A pair of opinionated agents *disagrees* if each has a different qualitative opinion state. If at least one pair of agents disagrees, then the group is in a *disagreement state*. The *dissensus* state is a special type of disagreement state in which individual agents may be opinionated but the group is unopinionated on average, i.e., $\left| \sum_{i=1}^{N_a} x_i \right| < \vartheta N_a$.

Each agent's opinion is updated in continuous time as a function of three key terms: a linear damping term, a nonlinear network interaction term that includes self-reinforcement, and an additive input term:

$$\dot{x}_i = -dx_i + u_i S \left(\alpha x_i + \gamma \sum_{\substack{k=1 \\ k \neq i}}^{N_a} \bar{a}_{ik} x_k \right) + b_i := h_i(\mathbf{x}). \quad (10.1)$$

$S : \mathbb{R} \rightarrow \mathbb{R}$ is an odd saturating function satisfying $S(0) = 0$, $S'(0) = 1$, $\text{sign}(S''(z)) = -\text{sign}(z)$. The saturation function is applied to the sum of two terms. The first is a self-reinforcement term with weight $\alpha \geq 0$. The second is the sum of opinions of agent i 's neighbors with weight $\gamma \in \mathbb{R}$. That the network interactions are saturated using S in (10.1) means the opinion dynamics of agent i are proportionally sensitive to changes in these interactions when the corresponding opinion magnitudes are small, but that the influence of these interactions levels off when the corresponding opinion magnitudes become large. In all simulations we let $S = \tanh$; however the results hold qualitatively for more general sigmoidal functions. For the general model, motivation, and relation to the broader literature see [19, 68].

The parameter $d > 0$ is the *damping coefficient*. In the absence of network interactions, agent i 's opinion x_i will converge exponentially to b_i/d at a rate determined by d . The control parameter $u_i > 0$ is the *attention* of agent i to network interactions. Attention u_i governs the influence of the saturated network interactions, relative to damping, on the opinion dynamics of agent i . When u_i is above a critical threshold that grows with d , the magnitude of x_i grows nonlinearly to a value much larger than b_i/d . Attention u_i can be fixed, time-varying, or defined to evolve according to state-dependent (closed-loop) dynamics. The *input* $b_i \in \mathbb{R}$ is an external signal or internal bias that stimulates an agent in favor of option A (option B) when $b_i > 0$ (< 0). For most of this paper we focus on the zero-input case: $b_i = 0$ for all i . We consider nonzero input in Section 10.6.

Let $W(\lambda_i)$ be the generalized eigenspace of \bar{A} relative to its eigenvalue λ_i . Let λ_{max} and λ_{min}

be the eigenvalues of \bar{A} with largest and smallest real parts. In the following lemma we state several useful properties of λ_{min} , λ_{max} , and their associated eigenspaces.

Lemma 10.2.1. A. *When \bar{G} is strongly connected (directed), $\lambda_{max} > 0$ is real with multiplicity 1, and for any nonzero vector $\mathbf{v} = (v_1, \dots, v_{N_a}) \in W(\lambda_{max})$, $v_i \neq 0$ and $\text{sign}(v_i) = \text{sign}(v_k)$ for all $i, k \in V$;*

B. *When \bar{G} is connected (undirected), $\lambda_{min} < 0$ is real and for any nonzero vector $\mathbf{w} = (w_1, \dots, w_{N_a}) \in W(\lambda_{min})$, $\text{sign}(v_i) = -\text{sign}(v_k)$ for at least one pair of $i, k \in V$.*

Proof. Observe that $\text{tr } \bar{A} = 0$ so $\text{Re}(\lambda_{min}) < 0$. When \bar{G} is undirected, \bar{A} is symmetric and λ_{min} is real. Further, \bar{A} is a nonnegative matrix. Parts A and B follow from the Perron-Frobenius Theorem [57, Theorem 11]. \square

10.3 Agreement and Disagreement

In this section we study opinion dynamics (10.1) with static $u_i := u \geq 0$ for all $i \in V$ and show how cooperative agents ($\gamma > 0$) give rise to agreement, whereas competitive agents ($\gamma < 0$) give rise to disagreement. In the following theorem, we expand upon the result in [19, Theorem IV.1] for two-option networks, and describe the steady-state solutions that arise from (10.1) in different parameter regimes.

Theorem 10.3.1. *The following hold true for (10.1) with $u_i := u \geq 0$ and $b_i = 0$ for all $i = 1, \dots, N_a$:*

A. Cooperation leads to agreement: *Let \bar{G} be a strongly connected directed graph. If $\gamma > 0$, the neutral state $\mathbf{x} = \mathbf{0}$ is a locally exponentially stable equilibrium for $0 < u < u_a$ and unstable for $u > u_a$, with*

$$u_a = \frac{d}{\alpha + \gamma \lambda_{max}}. \quad (10.2)$$

At $u = u_a$, branches of agreement equilibria, $x_i \neq 0$, $\text{sign}(x_i) = \text{sign}(x_k)$ for all $i, k \in V$, emerge in a steady-state bifurcation off of $\mathbf{x} = \mathbf{0}$ along $W(\lambda_{max})$;

B. Competition leads to disagreement: *Let \bar{G} be a connected undirected graph. If $\gamma < 0$ the neutral state $\mathbf{x} = \mathbf{0}$ is a locally exponentially stable equilibrium for $0 < u < u_d$ and unstable for $u > u_d$, with*

$$u_d = \frac{d}{\alpha + \gamma \lambda_{min}}. \quad (10.3)$$

At $u = u_d$, branches of disagreement equilibria, $\text{sign}(x_i) = -\text{sign}(x_k)$ for at least one pair $i, k \in V$,

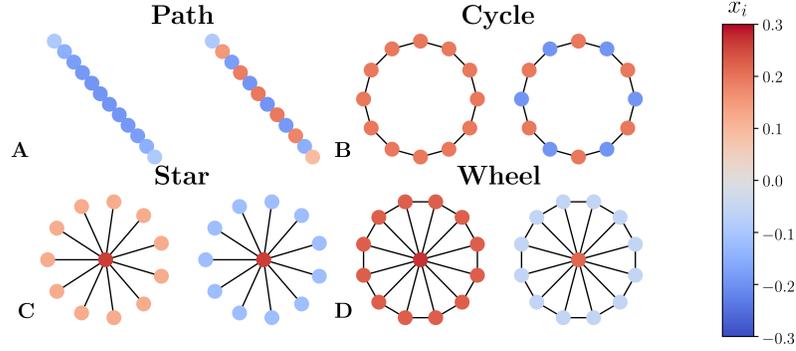


Figure 10.1: Steady-state patterns of agreement with $\gamma = 1.3$ (left) and disagreement with $\gamma = -1.3$ (right) from simulation of opinion dynamics (10.1) for four undirected graph types. Node color represents opinion x_i at $t = 500$. All nodes have $b_i = 0$ and randomized initial opinions drawn from $U(-1, 1)$. Parameters: $d = 1$, $\alpha = 1.2$, $u = 0.31$ for path and cycle, $u = 0.26$ for star and wheel.

$i \neq k$, emerge in a steady-state bifurcation off of $\mathbf{x} = \mathbf{0}$ along $W(\lambda_{min})$.

Proof. The Jacobian of (10.1) evaluated at $\mathbf{x} = \mathbf{0}$ is $J(\mathbf{0}) = (u\alpha - d)\mathcal{I} + u\gamma\bar{A}$, where \mathcal{I} is the identity matrix. The eigenvalues of $J(\mathbf{0})$ are $\mu_i = u(\alpha + \gamma\lambda_i) - d$ where λ_i is an eigenvalue of \bar{A} . When $0 \leq u < \min_i \frac{d}{\alpha + \gamma \operatorname{Re}(\lambda_i)}$, $\operatorname{Re}(\mu_i)$ is negative for all $i \in V$ and $\mathbf{x} = \mathbf{0}$ is locally exponentially stable. For values of u above this bound the origin is unstable. When $\gamma > 0$ this bound u_a corresponds to $\lambda_i = \lambda_{max}$ and is given by (10.2). When $\gamma < 0$ the bound u_d corresponds to $\lambda_i = \lambda_{min}$ and is given by (10.3). Thus, when $\gamma > 0$ ($\gamma < 0$) a steady-state bifurcation happens at $u = u_a$ ($u = u_d$) along $W(\lambda_{max})$ ($W(\lambda_{min})$). The rest follows from Lemma 10.2.1. \square

Remark 10.3.2. *Due to space constraints, we defer to a future publication detailed analysis of stability of the agreement and disagreement equilibria emerging at agreement and disagreement opinion-forming bifurcations. For one-dimensional kernels, we expect these equilibria to be generically stable, as easily verifiable using center-manifold reduction arguments.*

A main takeaway of Theorem 10.3.1 is that the spectral properties of \bar{A} inform the opinion formation outcomes on the network. Characterizing the eigenvalues λ_{min} , λ_{max} along with their associated eigenspaces $W(\lambda_{min})$, $W(\lambda_{max})$ is equivalent to characterizing the primary branches of opinionated steady-state solutions of (10.1) emerging at bifurcations from the neutral state. We use this approach to systematically classify the opinion patterns that arise for various networks with spectral properties that are well known or easily established. For larger and less structured networks these quantities can be easily computed numerically. Figures 10.1 and 10.2 illustrate agreement and disagreement equilibria of (10.1) on a variety of graphs, with u slightly above the bifurcation point.

Network consensus and dissensus are special cases of agreement and disagreement described in Theorem 10.3.1. Let the *consensus space* be $W_c := \operatorname{span}\{\mathbf{1}\}$ and the *dissensus space* be $W_d := \{\mathbf{x} \in$

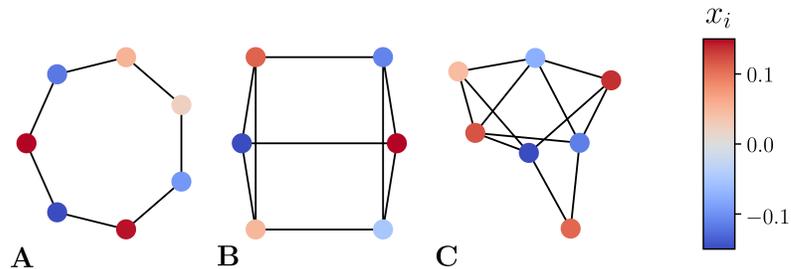


Figure 10.2: Disagreement patterns on odd cycle (A), 3-regular (B) and randomly generated (C) graphs. Parameters are $d = 1$, $\alpha = 0.5$, $\gamma = -0.5$, $u = u_d + 0.01$. All else is as in Fig. 10.1.

$\mathbb{R}^{N_a} : \sum_{i=1}^{N_a} x_i = 0$. Observe that W_c and W_d are orthogonal complements in \mathbb{R}^{N_a} . Outside a ϑ -neighborhood of the origin, the ϑ -neighborhood of W_c is made of consensus solutions, whereas the ϑ -neighborhood of W_d is made of dissensus solutions. We show next that for graphs in which all agents have the same number of neighbors, the agreement and disagreement equilibria of Theorem 10.3.1 correspond to consensus and dissensus.

Definition 10.3.3. $\hat{G} = (\hat{V}, \hat{E})$ is a K -regular graph if every vertex $i \in \hat{V}$ has exactly K neighbors.

Lemma 10.3.4. If \bar{G} is undirected, connected, and K -regular, all vectors $\mathbf{x} = (x_1, \dots, x_{N_a}) \in W(\lambda_{min})$ satisfy $\sum_{i=1}^{N_a} x_i = 0$.

Proof. Observe that for a connected, K -regular graph, $\lambda_{max} = K$ and $W(\lambda_{max}) = W_c$. Because \bar{A} is symmetric, all its generalized eigenspaces are orthogonal and, thus, $W(\lambda_{min}) \subseteq W_d$. The statement of the lemma follows. \square

Theorem 10.3.5 (Consensus and Dissensus). If \bar{G} is undirected, connected, and K -regular, the agreement bifurcations at $u = u_a$ with $\gamma > 0$ give rise to consensus solutions $|x_i - x_k| < \vartheta$ for all $i, k \in V$, and the disagreement bifurcations at $u = u_d$ with $\gamma < 0$ give rise to dissensus solutions $\left| \sum_{i=1}^{N_a} x_i \right| < \vartheta N_a$.

Proof. From Theorem 10.3.1 and Lemma 10.3.4, an opinion-forming bifurcation emerges along the consensus space for $\gamma > 0$ and dissensus space for $\gamma < 0$. This means that the resulting equilibrium solutions are arbitrarily close to the consensus (dissensus) space as $u \searrow u_a$ ($u \searrow u_d$), where the notation \searrow signifies approaching in value from above. \square

Figure 10.1B and Figure 10.2A,B are examples of consensus and dissensus on 2-regular graphs (a cycle with an even and odd number of nodes) and a 3-regular graph.

10.4 Agent Centrality

In this section we examine how equilibrium opinions of agents depend on their location in the graph. We show that at an opinion-forming bifurcation, an agent's opinion strength is often determined by its relative location in the network as quantified by a suitable centrality measure.

A *centrality measure* ranks how central each node is in a network, i.e., measures its influence over some emergent network property. We recall the definition of a well known network centrality measure originally proposed in [27]:

Definition 10.4.1 (Eigenvector Centrality). *The entries of the normalized positive left eigenvector corresponding to the eigenvalue λ_{max} of \hat{A} for a (directed or undirected) graph \hat{G} provide a centrality measure for the nodes of the graph.*

It is shown in [28] that eigenvector centrality, deriving from the adjacency matrix and not the Laplacian, is particularly useful for graphs on which agents with high in-degree have many neighbors of low in-degree, and is related to several other common graph centrality measures. Let $\mathbf{v}^c = (v_1^c, \dots, v_{N_a}^c)$ be the centrality eigenvector for \bar{A} from Definition 10.4.1. For undirected graphs, this centrality vector determines the opinion strength of each agent at an agreement equilibrium predicted by Theorem 10.3.1A, i.e., the larger an agent's centrality, the stronger its opinion at agreement.

Theorem 10.4.2. *Consider opinion dynamics (10.1) with undirected, connected \bar{G} , $u_i := u \geq 0$, $b_i = 0$, and $\alpha > 0$. Agreement equilibria $\mathbf{x} = (x_1, \dots, x_{N_a})$ described in Theorem 10.3.1A satisfy $|x_i| < |x_k|$ if $v_i^c < v_k^c$ and $|x_i| = |x_k|$ if $v_i^c = v_k^c$ for all $i, k = 1, \dots, N_a$.*

Proof. Let $\mathbf{v} = (v_1, \dots, v_{N_a})$ be the normalized right λ_{max} -eigenvector of \bar{A} . By symmetry of \bar{A} , $v_i = v_i^c$ for all $i = 1, \dots, N_a$ and the theorem follows from Theorem 10.3.1A. \square

We next state an analogous result for disagreement equilibria on a common class of graphs called *bipartite graphs*.

Definition 10.4.3. *Undirected $\hat{G} = (\hat{V}, \hat{E})$ is a bipartite graph if \hat{V} can be subdivided into disjoint subsets \hat{V}_1, \hat{V}_2 such that every edge $\hat{e}_{ik} \in \hat{E}$ connects a vertex in \hat{V}_1 to one in \hat{V}_2 .*

In the following, we show that for disagreement on bipartite graphs, an agent's partition membership determines the sign of its equilibrium opinion and an agent's centrality in the network determines the strength of its equilibrium opinion.

Lemma 10.4.4 (λ_{min} -Eigenvector of Bipartite Graph). *Suppose \bar{G} is a bipartite graph and let \hat{V}_1, \hat{V}_2 be the two vertex subsets of V from Definition 10.4.3. Let $\mathbf{w} = (w_1, \dots, w_{N_a})$ be the eigenvector*

corresponding to λ_{min} of \bar{A} . Then $w_i, w_k \neq 0$ and $\text{sign}(w_i) = -\text{sign}(w_k)$ for all $i \in \hat{V}_1, k \in \hat{V}_2$. Moreover, $|w_i| = v_i^c$ for all $i = 1, \dots, N_a$.

Proof. For all bipartite graphs, $\dim W(\lambda_{min}) = 1$. By symmetry of \bar{A} , the λ_{max} -eigenvector is equal to \mathbf{v}^c . The lemma then follows from [190, Theorem 1.2], which states that the terms of the λ_{min} -eigenvector are equal in magnitude to the terms of the λ_{max} -eigenvector, with the sign structure reflecting the bipartition. \square

Theorem 10.4.5 (Disagreement Opinion Strength Reflects Agent Centrality on Bipartite Graphs).

Consider (10.1) with undirected bipartite \bar{G} , $u_i := u \geq 0$, $b_i = 0$, $\alpha > 0$, and $\gamma < 0$. Let \hat{V}_1, \hat{V}_2 be the two subsets of V from Definition 10.4.3. Disagreement equilibria $\mathbf{x} = (x_1, \dots, x_{N_a})$ described in Theorem 10.3.1.B satisfy $|x_i| < |x_k|$ if $v_i^c < v_k^c$ and $|x_i| = |x_k|$ if $v_i^c = v_k^c$ for all $i, k = 1, \dots, N_a$. Moreover, $\text{sign}(x_i) = -\text{sign}(x_k)$ for all $i \in \hat{V}_1, k \in \hat{V}_2$.

Proof. This follows by Theorem 10.3.1B and Lemma 10.4.4. \square

All graphs shown in Figure 10.1 have a simple λ_{min} . Further, the cycle, path, and star are bipartite graphs and the sign distribution of nodes across options reflects their bipartition. Observe that the magnitude of opinions reflects relative centrality on all the shown graphs in both the agreement and disagreement parameter regimes, including the wheel which is not bipartite. This suggests that the λ_{min} -eigenvector can sometimes be related to a notion of agent centrality even when it does not precisely equal \mathbf{v}^c . In contrast, graphs in Figure 10.2 are not bipartite, and graphs A, B have a two-dimensional $W(\lambda_{min})$. The distribution of disagreement opinions on these graphs is more heterogeneous in magnitude, despite the first two graphs being regular and all agents being equally central. Additional equilibrium distributions of opinions are possible in the disagreement regime.

10.5 Graph Symmetry

In this section we relate the patterns of emergent opinions on solution branches described in Theorem 10.3.1 to the symmetry of the underlying graph. We first state several key definitions.

Recall that a *graph automorphism* of \bar{G} is a permutation of vertices that preserves adjacency (i.e., maps edges to edges and non-edges to non-edges). The *automorphism group* $\text{Aut}(\bar{G}) := \Gamma$ is the set of all graph automorphisms of \bar{G} . Define the group of permutations of a set of n symbols $\{1, \dots, n\}$ as S_n . Then for graph \bar{G} with N_a agent vertices, $\Gamma \subseteq S_{N_a}$. The graph automorphism

group of \bar{G} can also be interpreted as the group of permutation matrices which commute with its adjacency matrix \bar{A} . Each element $\rho \in \Gamma$ can be represented by a permutation matrix P_ρ which satisfies $P_\rho \bar{A} = \bar{A} P_\rho$ [41, Proposition 3.8.1]. We commonly refer to ρ as a *symmetry* of \bar{G} and Γ as its *symmetry group*.

A connected notion is the *equivariance* of a dynamical system with respect to a symmetry group. Consider the opinion dynamics (10.1) as a dynamical system $\dot{\mathbf{x}} = \mathbf{h}(\mathbf{x})$, where the map $\mathbf{h} : \mathbb{R}^{N_a} \rightarrow \mathbb{R}^{N_a}$ is $\mathbf{h}(\mathbf{x}) = (h_1(\mathbf{x}), \dots, h_{N_a}(\mathbf{x}))$. Let Σ be a compact Lie group with elements σ that act on \mathbb{R}^{N_a} . Then \mathbf{h} is σ -equivariant for some $\sigma \in \Sigma$ if $\sigma \mathbf{h}(\mathbf{x}) = \mathbf{h}(\sigma \mathbf{x})$, and \mathbf{h} is Σ -equivariant if this holds true for all $\sigma \in \Sigma$ [83, Definition 1.7]. In the following theorem we show that the symmetry group of the graph \bar{G} is also a symmetry group of the opinion dynamics (10.1) with zero input.

Theorem 10.5.1 (Γ -equivariance). *Consider opinion dynamics (10.1) with $b_i = 0$ for $i = 1, \dots, N_a$. Let Γ be the automorphism group of \bar{G} . The opinion dynamics are Γ -equivariant.*

Proof. Let ρ be an element of Γ , and define the function $\hat{\mathbf{S}} : \mathbb{R}^{N_a} \rightarrow \mathbb{R}^{N_a}$ as $(S(y_1), \dots, S(y_{N_a}))$. Every permutation ρ can be decomposed into a product of transpositions ρ_{ik} that interchange elements in positions i and k of a set. For an arbitrary transposition, it is easy to see that $P_{\rho_{ik}} \hat{\mathbf{S}}(\mathbf{y}) = \hat{\mathbf{S}}(P_{\rho_{ik}} \mathbf{y})$, and by iteratively applying transpositions we see that for all permutations ρ , $P_\rho \hat{\mathbf{S}}(\mathbf{y}) = \hat{\mathbf{S}}(P_\rho \mathbf{y})$. Using this and the definition of graph automorphism we get

$$P_\rho \mathbf{h}(\mathbf{x}) = -d P_\rho \mathbf{x} + u \hat{\mathbf{S}}(P_\rho(\alpha \mathcal{I} + \gamma \bar{A}) \mathbf{x}) = -d P_\rho \mathbf{x} + u \hat{\mathbf{S}}((\alpha \mathcal{I} + \gamma \bar{A}) P_\rho \mathbf{x}) = \mathbf{h}(P_\rho \mathbf{x}).$$

This holds for all $\rho \in \Gamma$, and the theorem follows. \square

In the remainder of this section we illustrate how graph symmetry dictates patterns of opinions on the network.

Definition 10.5.2. *Let Γ be the automorphism group of \bar{G} , and $i \in V$ a vertex. An orbit of i is $O_i = \{k \in V \mid k = \rho i \text{ for some } \rho \in \Gamma\}$. The orbits are equivalence classes that partition V through the equivalence relation*

$$i \sim k \text{ if } k = \rho i \text{ for some } \rho \in \Gamma. \quad (10.4)$$

Theorem 10.5.3. *Consider opinion dynamics (10.1) with $u_i := u \geq 0$ and $b_i = 0$ for all $i \in V$. Let Γ be the automorphism group of the undirected graph \bar{G} , and for any two vertices $i, k \in V$ define the equivalence relation $i \sim k$ as in (10.4).*

A. Suppose $\gamma > 0$. For the agreement equilibria $\mathbf{x} = (x_1, \dots, x_{N_a})$ from Theorem 10.3.1.A, if $i \sim k$, then $x_i = x_k$.

B. Suppose $\gamma < 0$ and λ_{min} has multiplicity 1. For the disagreement equilibria $\mathbf{x} = (x_1, \dots, x_{N_a})$ from Theorem 10.3.1.B, if $i \sim k$, then $|x_i| = |x_k|$.

Proof. In A and B, the solutions \mathbf{x} appear along the 1-dimensional subspace $\ker J(\mathbf{0})$. For any $\rho \in \Gamma$ and $\mathbf{v} \in \ker J(\mathbf{0})$, $P_\rho \mathbf{v} \in \ker J(\mathbf{0})$ [83, Remark 1.25], and $P_\rho \mathbf{v} = \pm \mathbf{v}$ [41, Lemma 3.8.2]. The only way for this to be true is if for any $i, k \in V$ for which $i \sim k$, $|v_i| = |v_k|$. \square

All graphs shown in Figure 10.1 have a nontrivial symmetry and a simple λ_{min} . The vertices contained in the same orbit under the action of its symmetry have equal magnitude opinion in both agreement and disagreement parameter regimes (e.g. all the vertices of a cycle, the outer vertices of the star and wheel). The randomly generated graph in Figure 10.2C also has a simple λ_{min} but only a trivial symmetry. Each node is its own orbit and the equilibrium opinion magnitude is different at each node. In contrast, both graphs in Figure 10.2A,B have a nontrivial symmetry but the dimension of $W(\lambda_{min})$ is 2 and the conditions of Theorem 10.5.1 are not met. The disagreement opinion magnitudes on these graphs do not reflect the orbit structure of the graph's symmetry group.

10.6 Agent Sensitivity and Opinion Cascades

In [19, Section VI] a state feedback law is introduced for the attention parameter u_i in (10.1), which enables the group to become opinionated in response to input b_i . Each agent's attention parameter u_i tracks a saturated norm of the system state observed by agent i , with dynamics defined by

$$\tau_s \dot{u}_i = -u_i + S_u \left(x_i^2 + \sum_{k=1}^{N_a} (\bar{a}_{ik} x_k)^2 \right). \quad (10.5)$$

Here, $\tau_s > 0$ is the time scale of the integration and $S_u(y) = u_f(F(g(y - y_m)) - F(-gy_m))$, with $F(x) = \frac{1}{1+e^{-x}}$. Design parameters $g, u_f, y_m > 0$ tune the system response. The positive feedback between opinion strength and attention defined by dynamics (10.5) provides the resulting opinion-attention dynamics with a threshold for the input strength needed to let an originally weakly opinionated ($|x_i| \ll \vartheta$, for all $i = 1, \dots, N_a$), weakly attentive ($|u_i| \ll \vartheta$, for all $i = 1, \dots, N_a$) network transition to a strongly opinionated ($|x_i| \gg \vartheta$, for all $i = 1, \dots, N_a$), strongly attentive ($|u_i| \gg \vartheta$, for all $i = 1, \dots, N_a$) state [67].

Such an *opinion cascade* is illustrated in Figures 10.3 and 10.4 top left. An opinion cascade is

of agreement (disagreement) type if the final opinion configuration is of agreement (disagreement) type.

A rigorous analysis of opinion cascade over networks is the subject of ongoing work based on unfolding theory [82] of agreement and disagreement bifurcations. Here, we illustrate with numerical examples. We show in particular how the spectral properties of the adjacency matrix predict which agents are more likely to trigger a cascade when excited by an input.

We conjecture that the norm of the projection of the vector of inputs \mathbf{b} onto the eigenspaces $W(\lambda_{max})$ or $W(\lambda_{min})$, depending on the cooperative or competitive regime, is directly related to whether or not the opinion cascade gets triggered. To illustrate this, in Figures 10.3 and 10.4 we show two examples of a disagreement cascade for graphs with a simple λ_{min} . In Figure 10.3 the graph is a balanced tree, which is bipartite. Thus, the magnitude of each entry of its λ_{min} -eigenvector corresponds to the corresponding agent's relative centrality, as proved in Lemma 10.4.4. When an input is given to the most central agent an opinion cascade gets triggered, whereas when the same input is given to an agent located on an outer leaf, the group remains neutral.

Analogously, in Figure 10.4 a cascade is triggered on a randomly generated network when an input is given to the agent with the largest magnitude entry in the λ_{min} -eigenvector, and is not triggered when the same input is given to the agent with the smallest entry. The post-cascade pattern of opinions on this network is very close to the pattern for the solution of the zero-input system shown in Figure 10.2C. In this way, the spectral properties of \bar{A} can be used not only to predict the patterns of opinion formation for opinion dynamics (10.1) both with a static $u_i = u$ and coupled with (10.5), but also to inform which agent should receive a control signal or carry a sensor in order to have maximal likelihood of an informed opinion cascade spreading on the network.

10.7 Discussion

We have shown how patterns of nonlinear opinion formation on two-option networks can be studied systematically using spectral properties of the graph adjacency matrix. We proved that agent centrality and symmetry of the underlying graph select the pattern of opinions in agreement and disagreement for a large set of networks. We illustrated how the eigenvectors of the adjacency matrix inform sensitivity of individual nodes to input that triggers opinion cascades; proving the influence of distributed input on opinion cascades is the subject of ongoing work. Other future work includes the natural generalization to multi-option networks with opinion dynamics defined in [19].

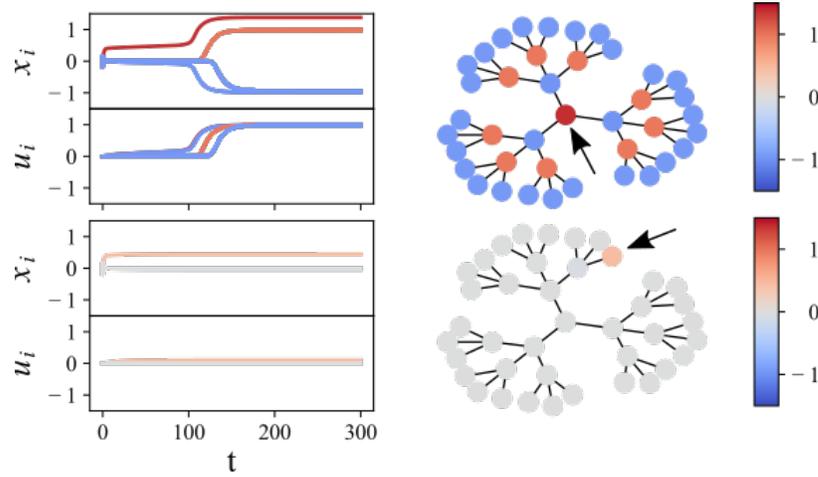


Figure 10.3: Triggering a cascade on the balanced tree. Input $b_i = 0.4$ for agent marked with arrow; $b_i = 0$ for all other agents. Simulations of (10.1) start from small random initial opinions drawn from $\mathcal{N}(0, 0.1)$. Left: time trajectories of the dynamics. Right: final opinion, indicated by color, of simulation at $t = 300$. Parameters: $d = 1$, $\alpha = 1$, $\gamma = -1$, $u_i(0) = 0$ for all $i \in V$, $g = 10$, $y_m = 0.4$, $u_f = 1$, $\tau_s = 10$. Only in the top plots does the model undergo a disagreement opinion cascade, with $\vartheta = 0.5$.

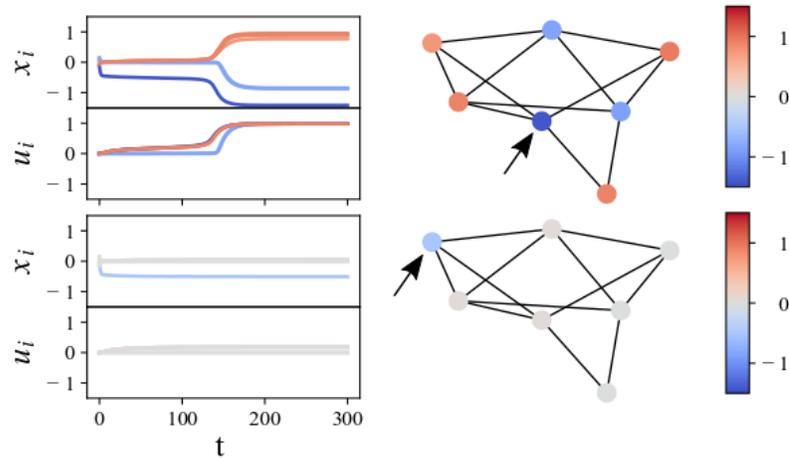


Figure 10.4: Triggering a cascade on a randomly generated graph. Input $b_i = -0.45$ for agent marked with arrow; $b_i = 0$ for all other agents. Simulations of (10.1) start from small random initial opinions drawn from $\mathcal{N}(0, 0.1)$. Left: time trajectories of the dynamics. Right: final opinion, indicated by color, of simulation at $t = 300$. Parameters: $d = 1$, $\alpha = 0.5$, $\gamma = -0.5$, $u_i(0) = 0$ for all $i \in V$, $g = 10$, $y_m = 0.4$, $u_f = 1$, $\tau_s = 10$. Only in the top plots does the model undergo a disagreement opinion cascade, with $\vartheta = 0.5$.

Chapter 11

Control of Agreement and Disagreement Cascades with Distributed Inputs

ANASTASIA BIZYAEVA, TIMOTHY SOROCHKIN, ALESSIO FRANCI,
AND NAOMI EHRICH LEONARD

For a group of autonomous communicating agents to carry out coordinated objectives, it is paramount that they can distinguish meaningful input from disturbance, and come rapidly and reliably to agreement or disagreement in response to that input. We study how opinion formation cascades through a group of networked decision makers in response to a distributed input signal. Using a nonlinear opinion dynamics model with dynamic feedback modulation of an attention parameter, we prove how the triggering of an opinion cascade and the collective decision itself depend on both the distributed input and node agreement and disagreement centrality indices, determined by the spectral properties of the network graph. Moreover, we show how the attention dynamics introduce an implicit threshold that distinguishes between distributed inputs that trigger cascades and ones that are rejected as disturbance.

11.1 Introduction

Emerging technologies rely on network communications and sensor input to make coherent collective decisions. For example, autonomous multi-robot teams must cooperate to move as a group, avoid

collisions, and perform collective tasks in potentially dynamic and uncertain environments. These objectives necessarily involve on-the-fly collective decision making about context-dependent options, such as which of multiple available paths to take, in which direction to turn, or how to distribute available tasks among team members. There is urgent need for a unified design framework that enables autonomous teams to rapidly and reliably coordinate decisions across different contexts in a distributed manner.

Mathematical models of networked opinion dynamics, e.g. [62, 105, 160, 161, 220], can be useful for this purpose, in part due to their analytical tractability. However, most existing models rely on a linear weighted-average opinion updating process, which imposes limits on the range of behaviors exhibited. Notably, special network structure or asymmetry is needed to produce solutions other than consensus, whereas applications that require groups to split among locations or tasks warrant more generically enabled disagreement solutions.

We present new results for the nonlinear opinion dynamics model [19, 20, 67, 68], which provides an analytically tractable generalization of models that rely on linear weighted-averaging by applying a saturation function to opinion exchanges. The saturation makes the opinion update process fundamentally nonlinear, which has a number of important consequences. First, the model yields multi-stability of disagreement solutions as well as agreement solutions, each in easily identifiable parameter regimes, even for homogeneous agents. Second, opinions form through a bifurcation in which the neutral solution becomes unstable and agreement or disagreement solutions become stable, independent of the number of agents or options [19] and across network topologies [20]. This means solutions are reached rapidly and reliably – strength of opinions grow nonlinearly even with little or no input.

Our contributions yield a rigorous and systematic method for designing distributed inputs to control opinion formation and opinion cascades. We specialize the model to opinions on two options here, but results extend naturally to an arbitrary number of options. First, using Lyapunov-Schmidt reduction methods [82, Chapter VII], we prove that opinions generically form through a supercritical pitchfork bifurcation where the two stable branches are either agreement solutions or disagreement solutions, which we can fully characterize. Second, we prove that the agreement (disagreement) centrality of a node, which depends only on the spectral properties of the network adjacency matrix, determines the influence an input to the node has on the agreement (disagreement) bifurcation behavior. Third, when the opinion dynamics are coupled with the feedback attention dynamics introduced in [19], sufficiently large inputs can trigger an opinion cascade, depending on where in the network they are introduced. We show how agreement and disagreement centrality indices

predict the sensitivity of opinion cascades to distributed inputs: The more aligned the input vector is with the centrality vector, the smaller the inputs need to be to trigger a cascade.

We present the model in Section 11.2 and review Lyapunov-Schmidt reduction in Section 11.3. We prove the pitchfork bifurcations and the role of distributed input on opinion formation behavior for constant attention in Section 11.4 and for dynamic feedback controlled attention in Section 11.5. We conclude in Section 11.6.

11.2 Opinion Dynamics Model

We study a model of N_a agents communicating over a network and forming opinions on two options through a nonlinear process specialized from the multi-option general model in [19],[68]. As in [20], we specialize to agents that are homogeneous with respect to three fixed parameters in the dynamics: the rate of forgetting (damping coefficient $d > 0$), the edge weight in the communication network ($\gamma \in \mathbb{R}$), and the strength of self-reinforcement of opinion ($\alpha \geq 0$). In [20], we focused on the zero-input setting, i.e., the case in which there is no stimulus, evidence or bias that informs the agents about the relative merits of the options. Instead, here, we consider an input $b_i \in \mathbb{R}$, for each agent $i = 1, \dots, N_a$, and allow the inputs to be heterogeneously distributed over the network of agents. We further model heterogeneity over the agents in their attention to network exchange.

Agent interactions are encoded in graph $G = (V, E)$ where $V = \{1, \dots, N_a\}$ is the index set of vertices. Vertex $i \in V$ represents agent i , and edge set $E \subseteq V \times V$ represents agent interactions. $A = (a_{ik})$, $i, k \in V$, is the unweighted graph adjacency matrix with elements satisfying $a_{ik} = 1$ if and only if $e_{ik} \in E$, and $a_{ik} = 0$ otherwise. We let $a_{ii} = 0$ for all $i \in V$. G is an *undirected* graph if $a_{ik} = a_{ki}$ for all $i, k \in V$. Let λ_q , $q = 1, \dots, N_a$, be the eigenvalues of A and $W(\lambda_q)$ the generalized eigenspace associated to λ_q . We define λ_{max} and λ_{min} to be the λ_q with largest and smallest real parts, respectively, and \mathbf{v}_{max} (\mathbf{w}_{max}) and \mathbf{v}_{min} (\mathbf{w}_{min}) to be the corresponding unit right (left) eigenvectors.

With two options, the opinion of each agent i is a real-valued scalar $x_i \in \mathbb{R}$. The sign of x_i corresponds to agent i favoring option 1 ($x_i > 0$) or favoring option 2 ($x_i < 0$). The magnitude of the opinion variable x_i describes the strength of agent i 's commitment. The vector of agents' opinions $\mathbf{x} = (x_1, \dots, x_{N_a}) \in \mathbb{R}^{N_a}$ is the *network opinion state*.

Agent i has a *neutral* opinion when $x_i = 0$, and we say it is *opinionated* otherwise. Furthermore we say that any pair of agents $i, k \in V$ *agree* (*disagree*) when they are opinionated and favor the same option (different options), i.e. $\text{sign}(x_i) = \text{sign}(x_k)$ ($\text{sign}(x_i) \neq \text{sign}(x_k)$). The group is in an

agreement state when all agents agree, and in a *disagreement state* when at least one pair of agents disagree.

Each agent i updates its own opinion state x_i in continuous time according to the nonlinear update rule:

$$\dot{x}_i = -dx_i + u_i S \left(\alpha x_i + \gamma \sum_{k=1}^{N_a} a_{ik} x_k \right) + b_i. \quad (11.1)$$

The rule has three parts: a damping term with coefficient $d > 0$, a nonlinear interaction term that includes inter-agent exchanges with weight $\gamma \in \mathbb{R}$ and a self-reinforcement term with weight $\alpha \geq 0$, and an additive input $b_i \in \mathbb{R}$.

The nonlinearity applied to the inter-agent exchanges and self-reinforcement is defined by an odd sigmoidal saturating function S which satisfies $S(0) = 0$, $S'(0) = 1$, and $\text{sign}(S''(z)) = -\text{sign}(z)$. This is motivated from the literature and means that agent i is more influenced by opinion fluctuations in its neighbors when their average opinion is close to neutral, and as neighbors' opinions grow large on average their influence levels off. In simulations and analysis throughout this paper we use $S = \tanh$. We purposely leave the sigmoid more general in (11.1) because the results in this paper generalize to arbitrary odd sigmoidal functions with minor modifications in the algebraic details of the proofs.

We begin by specializing a result from [20].

Proposition 11.2.1 ([20], Theorem 1). *The following hold true for (11.1) with $u_i := u \geq 0$ and $b_i = 0$ for all $i = 1, \dots, N_a$:*

A. Cooperation leads to agreement: *Let G be a connected undirected graph. If $\gamma > 0$, the neutral state $\mathbf{x} = \mathbf{0}$ is a locally exponentially stable equilibrium for $0 < u < u_a$ and unstable for $u > u_a$, with $u_a = \frac{d}{\alpha + \gamma \lambda_{max}}$. At $u = u_a$, branches of agreement equilibria, $x_i \neq 0$, $\text{sign}(x_i) = \text{sign}(x_k)$ for all $i, k \in V$, emerge in a steady-state bifurcation off of $\mathbf{x} = \mathbf{0}$ along $W(\lambda_{max})$;*

B. Competition leads to disagreement: *Let G be a connected undirected graph. If $\gamma < 0$ the neutral state $\mathbf{x} = \mathbf{0}$ is a locally exponentially stable equilibrium for $0 < u < u_d$ and unstable for $u > u_d$, with $u_d = \frac{d}{\alpha + \gamma \lambda_{min}}$. At $u = u_d$, branches of disagreement equilibria, $\text{sign}(x_i) = -\text{sign}(x_k)$ for at least one pair $i, k \in V$, $i \neq k$, emerge in a steady-state bifurcation off of $\mathbf{x} = \mathbf{0}$ along $W(\lambda_{min})$.*

11.3 Lyapunov-Schmidt Reduction

To systematically characterize the equilibria of the opinion dynamics model as a function of parameters, we leverage Lyapunov-Schmidt reduction and its use in computing bifurcation diagrams. Consider the n -dimensional dynamical system $\dot{\mathbf{y}} = \Phi(\mathbf{y}, \mathbf{p})$, where $\Phi : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^n$ is a smooth parameterized vector field, $\mathbf{y} \in \mathbb{R}^n$ is a state vector, and $\mathbf{p} \in \mathbb{R}^m$ is a vector of parameters. Let $\mathbf{r}_s \in \mathbb{R}^n$, $s = 1 \dots N$. The N^{th} order derivative of Φ at $(\mathbf{y}^*, \mathbf{p}^*)$ is

$$(d^N \Phi)_{\mathbf{y}^*, \mathbf{p}^*}(\mathbf{r}_1, \dots, \mathbf{r}_N) = \left. \frac{\partial}{\partial t_1} \dots \frac{\partial}{\partial t_N} \Phi \left(\mathbf{y}^* + \sum_{s=1}^N t_s \mathbf{r}_s, \mathbf{p}^* \right) \right|_{t_1 = \dots = t_N = 0}. \quad (11.2)$$

The equilibria of $\dot{\mathbf{y}} = \Phi(\mathbf{y}, \mathbf{p})$ are the level sets $\Phi(\mathbf{y}, \mathbf{p}) = 0$, which defines the *bifurcation diagram* of the system.

The Jacobian of the system is the matrix J with elements $J_{ij} = \frac{\partial \Phi(\mathbf{y}, \mathbf{p})}{\partial y_{ij}}$. When J evaluated at an equilibrium point $(\mathbf{y}^*, \mathbf{p}^*)$ is degenerate (i.e. has rank $n - m$ where $0 < m < n$), the local bifurcation diagram can be described using m variables and the point is a *singular point*. The *Lyapunov-Schmidt reduction* of $\Phi(\mathbf{y}, \mathbf{p})$ is an m -dimensional system of equations that captures the structure of the local bifurcation diagram of the system near $(\mathbf{y}^*, \mathbf{p}^*)$. The procedure for deriving the Lyapunov-Schmidt reduction [82, Chapter VII] involves projecting the Taylor expansion of $\Phi(\mathbf{y}, \mathbf{p})$ onto the kernel of its Jacobian at the singularity. The Implicit Function Theorem is used to solve for $n - m$ variables as function of the remaining m , thus approximating the local vector field in the directions orthogonal to the kernel.

The *normal form* for a bifurcation is the simplest equation that captures all qualitative features of the bifurcation diagram. Systems with an odd state symmetry $\Phi(-\mathbf{y}, \mathbf{p}) = -\Phi(\mathbf{y}, \mathbf{p})$ often exhibit a *pitchfork bifurcation*. A normal form for the pitchfork bifurcation universal unfolding is

$$\dot{y} = p_1 y \pm y^3 + p_2 + p_3 y^2 \quad (11.3)$$

where $y \in \mathbb{R}$ is the reduced state, p_1 is a *bifurcation parameter* and p_2, p_3 are *unfolding parameters*. When $p_2 = p_3 = 0$, the symmetric pitchfork normal form is recovered in (11.3). When one of the unfolding parameters is nonzero, it follows from unfolding theory [82, Chapter III] that the bifurcation diagram changes locally to one of four possible topologically distinct configurations (see Fig. 11.1).

11.4 Constant Attention: Sensitivity to Input

In this section, we investigate how a vector of constant inputs \mathbf{b} informs the outcome of the opinion formation process (11.1) when attention is constant and $u_i := u \in \mathbb{R}$ for all $i = 1, \dots, N_a$. The Jacobian of (11.1) evaluated at $\mathbf{x} = 0$ is

$$J_x = (u\alpha - d)\mathcal{I} + u\gamma A \quad (11.4)$$

with identity matrix \mathcal{I} . The dynamics (11.1) in vector form are

$$\dot{\mathbf{x}} = -d\mathbf{x} + u\mathbf{S}((\alpha\mathcal{I} + \gamma A)\mathbf{x}) + \mathbf{b} := F(\mathbf{x}, u, \mathbf{b}) \quad (11.5)$$

where $\mathbf{S}(\mathbf{y}) = (S(y_1), \dots, S(y_n))$, $\mathbf{y} \in \mathbb{R}^n$, and $\mathbf{b} = (b_1, \dots, b_{N_a})$. The following theorem generalizes [87, Theorem 1] to describe bifurcations of the opinion dynamics of homogeneous agents. The theorem shows that any bifurcation of $\mathbf{x} = 0$ of (11.1) that is generated by a simple eigenvalue of the adjacency matrix A must be a pitchfork bifurcation.

Theorem 11.4.1 (Pitchfork Bifurcation). *Consider (11.1) and define $u^* = \frac{d}{\alpha + \lambda\gamma}$, where λ is a simple real eigenvalue of adjacency matrix A for a strongly connected graph G . Let $\mathbf{v} = (v_1, \dots, v_{N_a})$ and $\mathbf{w} = (w_1, \dots, w_{N_a})$ be right and left unit eigenvectors, respectively, corresponding to λ . Assume that (i) for all eigenvalues $\xi \neq \lambda$ of A , $\text{Re}[\xi] \neq \lambda$; (ii) $\alpha + \lambda\gamma \neq 0$; (iii) $\langle \mathbf{w}, \mathbf{v}^3 \rangle \neq 0$. Let $f(z, u, \mathbf{b})$ be the Lyapunov-Schmidt reduction of $F(\mathbf{x}, u, \mathbf{b})$ at $(0, u^*, 0)$.*

A. *Bifurcation problem $f(z, u, 0) = 0$ has a symmetric pitchfork singularity at $(z, u, \mathbf{b}) = (0, u^*, 0)$. For values of $u > u^*$ and sufficiently small $|u - u^*|$, two branches of equilibria branch off from $\mathbf{x} = 0$ in a pitchfork bifurcation along a manifold tangent at $\mathbf{x} = 0$ to $\text{span}\{\mathbf{v}\}$.*

When $\text{sign}\{\langle \mathbf{w}, \mathbf{v}^3 \rangle / \langle \mathbf{w}, \mathbf{v} \rangle\}(\alpha + \lambda\gamma) > 0$ (< 0) the bifurcation happens supercritically (subcritically) with respect to u .

B. *Bifurcation problem $f(z, u, \mathbf{b}) = 0$ is an N_a -parameter unfolding of the symmetric pitchfork, and $\frac{\partial f}{\partial b_i}(z, u, \mathbf{b}) = w_i$.*

Proof. The eigenvalues of J_x (11.4) are $\mu = u\alpha - d + u\gamma\lambda$, and so, at $u = u^*$, J_x has a single zero eigenvalue. Observe that the left and right null eigenvectors of J_x are precisely \mathbf{w} and \mathbf{v} . Following the procedure outlined in [82, Chapter I, 3.(e)] we derive $f(z, u, \mathbf{b})$. We derive the coefficients of the polynomial expansion of $f(z, u, \mathbf{b})$ [82, Chapter I, Equations 3.23(a)-(e)] through third order in the state variable. Note that $(d^2F)_{0, u^*, 0}(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3) = 0$ for any \mathbf{v}_i because $S''(0) = 0$, which implies

that $f_{zz} = 0$ by [82, Chapter I, Equation 3.23(b)]. Also, $f_z(0, u^*, 0) = 0$ by [82, Chapter I, Equation 3.23(a)]. The nonzero coefficients in the expansion are

$$\begin{aligned} f_{xxx} &= \langle \mathbf{w}, (d^3 F)_{0, u^*, 0}(\mathbf{v}, \mathbf{v}, \mathbf{v}) \rangle = -2d(\alpha + \lambda\gamma)^2 \langle \mathbf{w}, \mathbf{v}^3 \rangle \\ f_{b_i} &= \left\langle \mathbf{w}, \frac{\partial F}{\partial b_i}(0, u^*, 0) \right\rangle = w_i \\ f_{\hat{u}x} &= \left\langle \mathbf{w}, \left(d \frac{\partial F}{\partial \hat{u}} \right)_{0, u^*, 0}(\mathbf{v}, \mathbf{v}) \right\rangle = (\alpha + \lambda\gamma) \langle \mathbf{w}, \mathbf{v} \rangle \end{aligned}$$

where $\hat{u} = u - u^*$ and $\langle \cdot, \cdot \rangle$ denotes the standard vector inner product. Also, observe that we can align the left and right eigenvectors to satisfy $\langle \mathbf{w}, \mathbf{v} \rangle = k_1 > 0$ (the inner product is nonzero by duality). Then $\langle \mathbf{w}, \mathbf{v}^3 \rangle := k_2 = \sum_{i=1}^{N_a} w_i v_i^3$. The Lyapunov-Schmidt reduction of (11.1) about $(0, u^*, 0)$ is thus

$$\dot{z} = k_1(\alpha + \lambda\gamma)\hat{u}z - 2k_2d(\alpha + \lambda\gamma)^2z^3 + \langle \mathbf{w}, \mathbf{b} \rangle + h.o.t. \quad (11.6)$$

Part A of the lemma follows by (11.6), by the recognition problem for the pitchfork bifurcation [82, Chapter II, Proposition 9.2], as well as by the definition of a center manifold. Part B follows by the definition of an unfolding and by (11.6). \square

From Theorem 11.4.1 we can describe many of the bifurcations of $\mathbf{x} = 0$ of (11.5) from the spectrum of A . In particular, if A has $N \leq N_a$ simple eigenvalues λ_q , we expect $\mathbf{x} = 0$ to exhibit N distinct pitchfork bifurcations at critical values of the parameter $u_q^* = d/(\alpha + \lambda_q\gamma)$. Locally near the bifurcation point the left eigenvector \mathbf{w}_q corresponding to λ_q informs the sign structure of the emergent equilibria, as explored in [20]. For undirected graphs we can deduce the direction in which the bifurcation branches appear.

Corollary 11.4.2. *Suppose G is an undirected graph. When $u_q^* = d/(\alpha + \lambda_q\gamma) > 0 (< 0)$ the pitchfork bifurcation at u_q^* happens supercritically (subcritically).*

Proof. Let \mathbf{v}_q and \mathbf{w}_q be the right and left eigenvectors of A corresponding to λ_q . For an undirected graph, $\mathbf{w}_q = \mathbf{v}_q$. Then $\langle \mathbf{w}_q, \mathbf{v}_q \rangle = \langle \mathbf{v}_q, \mathbf{v}_q \rangle > 0$ and $\langle \mathbf{w}_q, \mathbf{v}_q^3 \rangle = \langle \mathbf{v}_q, \mathbf{v}_q^3 \rangle = \sum_{k=1}^{N_a} (\mathbf{v}_q)_k^4 > 0$. The criticality condition from Theorem 11.4.1 becomes $(\alpha + \lambda_i\gamma) > 0 (< 0)$ for supercritical (subcritical) pitchfork bifurcation. Since $d > 0$ the result follows. \square

Using these general results on the bifurcation behavior of the opinion dynamics, the next theorem establishes that the agreement and disagreement bifurcations in Proposition 11.2.1 are supercritical pitchfork bifurcations in which $\mathbf{x} = 0$ loses stability and new branches of locally stable solutions

appear.

Theorem 11.4.3 (Agreement and Disagreement Pitchforks). *Consider (11.1) and let $u_i := u \geq 0$. The agreement and disagreement bifurcations in Proposition 11.2.1 are supercritical pitchfork bifurcations. Additionally, the two steady-state solutions, which appear for $u > u_a(u_d)$, are locally exponentially stable for $|u - u_a|(|u - u_d|)$ sufficiently small.*

Proof. Supercriticality of the bifurcating branches of equilibria follows for the undirected case from Corollary 11.4.2. For a directed graph and $\gamma > 0$ it follows from the Perron-Frobenius theorem that \mathbf{v}_{max} and \mathbf{w}_{max} have strictly positive components, i.e., $\langle \mathbf{w}_{max}, \mathbf{v}_{max} \rangle > 0$ and $\langle \mathbf{w}_{max}, \mathbf{v}_{max}^3 \rangle > 0$. Supercriticality then follows from Theorem 11.4.1. The two nontrivial fixed points are locally exponentially stable by analytic continuity of eigenvalues: $N_a - 1$ negative eigenvalues are shared with $\mathbf{x} = 0$ and the bifurcating eigenvalue is negative by [82, Chapter I, Theorem 4.1] because $\partial f / \partial z > 0$ for the Lyapunov-Schmidt reduction (11.6). \square

The results presented in this section provide rigorous predictions of the influence of inputs on the opinion formation bifurcation behavior. We define the *node agreement (disagreement) centrality index for node i* to be w_i , the i th component of \mathbf{w}_{max} (\mathbf{w}_{min}) (see also [67]). It follows by Theorem 11.4.1B and Theorem 11.4.3 that the influence of an input b_i to node i on the network opinion formation behavior is exactly node i 's agreement or disagreement centrality w_i . This allows us to predict in which direction the agreement or disagreement pitchfork unfolds as a function of the locations and strengths of distributed inputs. If $\langle \mathbf{b}, \mathbf{w}_{max} \rangle = 0$ the pitchfork does not unfold. If $\langle \mathbf{b}, \mathbf{w}_{max} \rangle < 0$ ($\langle \mathbf{b}, \mathbf{w}_{max} \rangle > 0$) the pitchfork unfolds in a such a way that it exhibits a lower (upper) smooth branch of equilibria. For example, in Fig. 11.1 the diagram on the left receives a nonzero input which is orthogonal to \mathbf{w}_{max} , and the symmetry of the pitchfork bifurcation is unbroken. On the right, $\langle \mathbf{b}, \mathbf{w}_{max} \rangle = 0.1$ and near the singular point of the symmetric diagram, the unfolded diagram favors the positive solution branch which corresponds to agents agreeing on the positive option.

11.5 Dynamic Attention: Cascades and Tunable Sensitivity to Input

In this section we show how distributed state feedback dynamics in the attention parameters of the opinion dynamics (11.1) give rise to agreement and disagreement cascades with tunable sensitivity to

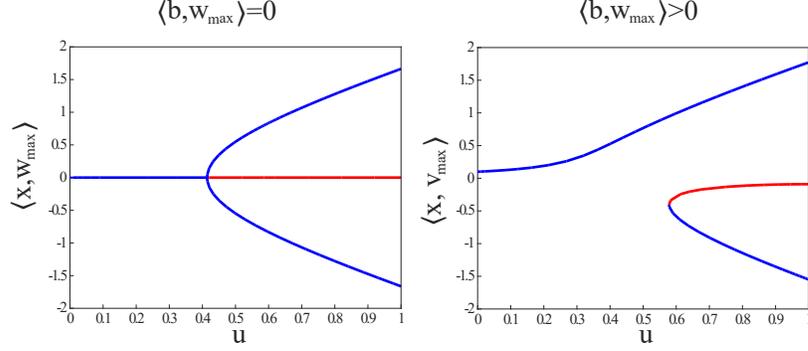


Figure 11.1: Symmetric pitchfork bifurcation and its unfolding for opinion dynamics (11.1) in the agreement regime ($\gamma > 0$) with three agents communicating over an undirected line graph. Blue (red) curves correspond to stable (unstable) equilibria. Vertical axis is the projection of equilibria onto $W(\lambda_{max})$. $d = \alpha = \gamma = 1$. Left: $\mathbf{b} = (0.05, 0, -0.05)$; right: $\mathbf{b} = 0.1\mathbf{v}_{max} + (0.05, 0, -0.05)$. Bifurcation diagrams generated with help of MatCont [47].

distributed input. We show that the magnitude of the distributed input vector, and its orientation relative to the centrality eigenvector \mathbf{w}_{max} (\mathbf{w}_{min}) when $\gamma > 0$ (< 0) provide control parameters for triggering cascades over the network. A single design parameter in the attention feedback dynamics can be used to tune the threshold above which inputs trigger a cascade.

As in [19] we define state feedback dynamics for the attention parameter u_i of each agent i to track the saturated norm of the opinions observed by agent i :

$$\tau_u \dot{u}_i = -u_i + S_u \left(x_i^2 + \sum_{k=1}^{N_a} (a_{ik} x_k)^2 \right). \quad (11.7)$$

S_u takes the form of the Hill activation function:

$$S_u(y) = \underline{u} + (\bar{u} - \underline{u}) \frac{y^n}{(y_{th})^n + y^n}, \quad (11.8)$$

where threshold $y_{th} > 0$. We constrain \bar{u} and \underline{u} such that $\bar{u} > u_c \geq \underline{u} > 0$, with $u_c = u_a$ (u_d) when $\gamma > 0$ (< 0). As in [67], we define an *opinion cascade* as a network transition from a weakly to a strongly opinionated state, where in a *weakly (strongly) opinionated* state, the agents' attention is close to its lower (upper) saturation bound, i.e. $u \cong \underline{u}$ ($u \cong \bar{u}$). See Fig. 11.2 for an example of an opinion cascade in an agreement ($\gamma > 0$) and disagreement ($\gamma < 0$) regime.

Assumption 11.5.1. G is undirected.

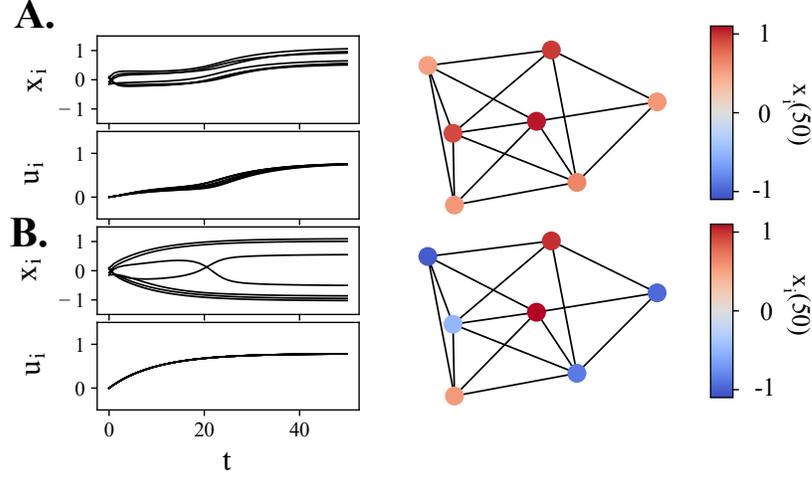


Figure 11.2: A. Agreement cascade, $\gamma = 1$, $\underline{u} = u_a - 0.01$, $\bar{u} = u_a + 0.6$; B. Disagreement cascade, $\gamma = -1$, $\underline{u} = u_d - 0.01$, $\bar{u} = u_d + 0.6$. Left) Opinion and attention trajectories. Right) Graph with node i color equal to $x_i(50)$. $d = 1$, $n = 3$, $y_{th} = 0.4$, $\tau_u = 10$, $\alpha = 1$, $d = 1$. For each i , $x_i(0) \in \mathcal{N}(0, 0.1)$, $u_i(0) = 0$, $b_i \in \mathcal{N}(0, 0.2)$.

In vector form, coupled dynamics (11.1),(11.7) become

$$\begin{pmatrix} \dot{\mathbf{x}} \\ \dot{\mathbf{u}} \end{pmatrix} = - \begin{pmatrix} d\mathbf{x} \\ \mathbf{u} \end{pmatrix} + \begin{pmatrix} \mathbf{u} \odot \mathbf{S}((\alpha\mathcal{I} + \gamma A)\mathbf{x}) + \mathbf{b} \\ \mathbf{S}_u((\mathcal{I} + A)\mathbf{x}^2) \end{pmatrix} \quad (11.9)$$

where $\mathbf{S}_u(\mathbf{y}) = (S_u(y_1), \dots, S_u(y_n))$, $\mathbf{y} \in \mathbb{R}^n$, $\mathbf{x}^2 = (x_1^2, \dots, x_{N_a}^2)$, and \odot is the element-wise product of vectors. The Jacobian of (11.9) at equilibrium point $(\mathbf{x}_s, \mathbf{u}_s)$ is

$$J_{(\mathbf{x}, \mathbf{u})} = \begin{pmatrix} -d\mathcal{I} + (\text{diag}\{\mathbf{u}_s\}(\alpha\mathcal{I} + \gamma A)) \odot K_1 & K_2 \\ (\mathcal{I} + A) \text{diag}\{\mathbf{x}_s\} \odot K_3 & -\mathcal{I} \end{pmatrix}, \quad (11.10)$$

$K_1 = \mathbf{S}'((\alpha\mathcal{I} + \gamma A)\mathbf{x}_s)\mathbf{1}^T$, $K_2 = \text{diag}\{\mathbf{S}((\alpha\mathcal{I} + \gamma A)\mathbf{x}_s)\}$, $K_3 = 2\mathbf{S}'_u((\mathcal{I} + A)\mathbf{x}_s^2)\mathbf{1}^T$, and $\mathbf{1} = (1, \dots, 1) \in \mathbb{R}^{N_a}$. Let $G(\mathbf{y}, \mathbf{b})$ be the right hand side of (11.9) with $\mathbf{y} = (\mathbf{x}, \mathbf{u})$.

Lemma 11.5.2 (Stability of $\mathbf{x} = 0$). *Consider (11.9) with $\mathbf{b} = 0$. The point $(\mathbf{x}_s, \mathbf{u}_s) = (\mathbf{0}, \underline{u}\mathbf{1})$ is an equilibrium point of the coupled dynamics. When either $\gamma > 0$ and $\underline{u} < u_a$ or $\gamma < 0$ and $\underline{u} < u_d$, it is locally exponentially stable.*

Proof. Plugging the state values into the coupled dynamics (11.9) easily verifies that $(\dot{\mathbf{x}}, \dot{\mathbf{u}}) = 0$ at $(\mathbf{x}_s, \mathbf{u}_s)$. Evaluated at this point, (11.10) simplifies to the block diagonal matrix

$$J_{(0, \underline{u})} = \begin{pmatrix} -d\mathcal{I} + \underline{u}(\alpha\mathcal{I} + \gamma A) & 0 \\ 0 & -\mathcal{I} \end{pmatrix}. \quad (11.11)$$

When $0 < \underline{u} < u_a$ (u_d), (11.11) has $2N_a$ eigenvalues with negative real part, and the stability conclusion follows. \square

Lemma 11.5.3 (Small Input Approximates Equilibrium Opinion). *Let $(\mathbf{x}_s, \mathbf{u}_s)$ be an equilibrium of (11.9) with inputs \mathbf{b} . Let $\underline{u} < u_c$ where $u_c = u_a$ if $\gamma > 0$ or $u_c = u_d$ if $\gamma < 0$. Define $\mathbf{w} = \mathbf{w}_{max}$ if $\gamma > 0$ or $\mathbf{w} = \mathbf{w}_{min}$ if $\gamma < 0$. Then*

$$\left. \frac{\partial \|\mathbf{x}_s\|}{\partial |\langle \mathbf{w}, \mathbf{b} \rangle|} \right|_{\mathbf{b}=\mathbf{x}_s=0} > 0. \quad (11.12)$$

Proof. Since $\mathbf{x} = 0$ is an equilibrium of the system with $\mathbf{b} = 0$, $(\mathbf{x}_s, \mathbf{u}_s)$ can be approximated by the linearization

$$J_{(0,\underline{u})} \begin{pmatrix} \mathbf{x}_s \\ \mathbf{u}_s \end{pmatrix} + \begin{pmatrix} \mathbf{b} \\ \mathbf{0} \end{pmatrix} = \mathbf{0}. \quad (11.13)$$

$J_{(0,\underline{u})}$ is symmetric and invertible so its inverse has the same eigenvectors. Thus, to the first order, it holds that

$$\mathbf{x}_s = -J_x^{-1} \mathbf{b} = -\sum_{q=1}^{N_a} \frac{1}{\mu_q} \langle \mathbf{w}_q, \mathbf{b} \rangle \mathbf{w}_q \quad (11.14)$$

where J_x is (11.4) with $u = \underline{u}$, \mathbf{w}_q is an eigenvector of A corresponding to λ_q , and $\mu_q = d + \underline{u}(\alpha + \lambda_q \gamma)$. Eigenvectors are orthogonal, so $\|\mathbf{x}_s\| = \sum_{i=1}^{N_a} \frac{1}{\mu_i^2} \langle \mathbf{w}_i, \mathbf{b} \rangle^2$. \square

Theorem 11.5.4 (Saddle-Node Bifurcation). *Consider (11.9) with a nonzero input vector \mathbf{b} and define $u_c = u_a$ if $\gamma > 0$ and $u_c = u_d$ if $\gamma < 0$. Let $\mathbf{w}_c = \mathbf{w}_{max}$ or \mathbf{w}_{min} respectively. Suppose $u_{th} \ll 1$ and $\underline{u} < u_c$ with $|\underline{u} - u_c|$ sufficiently small. There exists $p > 0$ such that when $|\langle \mathbf{w}_c, \mathbf{b} \rangle| = p$ there exists an equilibrium $(\mathbf{x}_p, \mathbf{u}_p)$ of (11.9) such that, if*

$$\langle \mathbf{w}_c, \mathbf{u}_p \odot \mathbf{v}_c^2 \odot \mathbf{S}''((\alpha \mathcal{I} + \gamma A) \mathbf{x}_p) \rangle > 0 \quad (11.15)$$

$$k_q \langle \mathbf{w}_c, \mathbf{u}_p \odot \mathbf{v}_c \odot \mathbf{v}_q \odot \mathbf{S}''((\alpha \mathcal{I} + \gamma A) \mathbf{x}_p) \rangle < 0 \quad (11.16)$$

is verified for all q at $(\mathbf{x}_p, \mathbf{u}_p)$ with $\lambda_q \neq \lambda_c$ an eigenvalue of A with corresponding left (right) eigenvector \mathbf{w}_q (\mathbf{v}_q), where $k_q = (\alpha + \gamma \lambda_q) / (d + \underline{u}(\alpha + \gamma \lambda_q))$: (i) There exists a smooth curve of equilibria in $\mathbb{R}^{2N_a} \times \mathbb{R}$ passing through $(\mathbf{x}_p, \mathbf{u}_p, p)$, tangent to the hyperplane $\mathbb{R}^{2N_a} \times \{p\}$; (ii) There are no equilibria near $(\mathbf{x}_p, \mathbf{u}_p, p)$ when $|\langle \mathbf{w}_c, \mathbf{b} \rangle| > p$ and two equilibria near $(\mathbf{x}_p, \mathbf{u}_p, p)$ for each $|\langle \mathbf{w}_c, \mathbf{b} \rangle| < p$; (iii) The two equilibria near $(\mathbf{x}_p, \mathbf{u}_p, p)$ are hyperbolic and have stable manifolds of dimensions N_a and $N_a - 1$ respectively.

Proof. (11.10) depends continuously on the model parameters and on the state. So, by [108, Chapter II, Theorem 5.1] the eigenvalues and eigenvectors of (11.10) change continuously for $\|\mathbf{x}_s\|$ sufficiently small. Leaving the full development of the matrix perturbation theory for future work, we conjecture that if $\|\mathbf{x}_s\|$ is sufficiently small then the eigenvectors of (11.11) are a good approximation of the eigenvectors of (11.10). Since the origin of (11.9) with $\mathbf{b} = \mathbf{0}$ is stable by Lemma 11.5.2 and because λ_{min} and λ_{max} are simple eigenvalues, if an eigenvalue of $J_{(\mathbf{x}_s, \mathbf{u}_s)}$ crosses zero for some $\|\mathbf{b}\|$ it must also be simple. This eigenvalue corresponds to a perturbation of $-d + \underline{u}(\alpha + \gamma\lambda_c)$ where $\lambda_c = \lambda_{max}$ or λ_{min} respectively.

By Lemma 11.5.3, if $\mathbf{b} \neq \mathbf{0}$ then at equilibrium $\|\mathbf{x}\| \neq 0$. Define $\tilde{\mathbf{v}}_c = (\mathbf{v}_c, \mathbf{0})$ and $\tilde{\mathbf{w}}_c = (\mathbf{w}_c, \mathbf{0})$. Let $g(z, \mathbf{b})$ be the Lyapunov-Schmidt reduction of (11.9) at an equilibrium (\mathbf{x}, \mathbf{u}) for sufficiently small inputs. We have

$$d^2G_{\mathbf{y}_p, \mathbf{b}_p}(\tilde{\mathbf{v}}_c, \tilde{\mathbf{v}}_c) = \sum_{j=1}^{N_a} \sum_{k=1}^{N_a} \frac{\partial^2(G)_i}{\partial x_j \partial x_k} (\mathbf{v}_c)_j (\mathbf{v}_c)_k \Big|_{(\mathbf{y}_p, \mathbf{b}_p)} = (\alpha + \lambda_c \gamma)^2 \begin{pmatrix} \mathbf{u}_p \\ \mathbf{0} \end{pmatrix} \odot \begin{pmatrix} \mathbf{v}_c^2 \\ \mathbf{0} \end{pmatrix} \odot \begin{pmatrix} \mathbf{S}''((\alpha \mathcal{I} + \gamma A)\mathbf{x}_p) \\ \mathbf{0} \end{pmatrix}. \quad (11.17)$$

The second derivative in the Lyapunov-Schmidt reduction is

$$g_{zz} = \langle \tilde{\mathbf{w}}_c, d^2G_{\mathbf{y}_p, \mathbf{b}_p}(\tilde{\mathbf{v}}_c, \tilde{\mathbf{v}}_c) \rangle = (\alpha + \lambda_c \gamma)^2 \times \sum_{i=1}^{N_a} (\mathbf{u}_p)_i (\mathbf{w}_c)_i^3 S'' \left(\alpha (\mathbf{x}_p)_i + \gamma \sum_{\substack{k=1 \\ k \neq i}}^{N_a} a_{ik} (\mathbf{x}_p)_k \right) > 0$$

by assumption (11.15). Additionally, the term $\langle \mathbf{w}_c, \mathbf{b} \rangle$ appears in $g(z, p)$ since $g_{b_i} = \left\langle \tilde{\mathbf{w}}_c, \frac{\partial G}{\partial b_i} \right\rangle = (\tilde{\mathbf{w}}_c)_i$.

Finally, we compute the coefficient of the cross-term $g_{z\tilde{b}}$ in the Lyapunov-Schmidt reduction. For convenience, we express $\mathbf{b} = \sum_{q=1}^{N_a} \beta_q \mathbf{w}_q$ where each $\beta_q := \langle \mathbf{w}_q, \mathbf{b} \rangle$. Coefficients of the cross-terms $z\beta_q$ in $g(z, \mathbf{b})$ simplify to

$$g_{z\beta_q} = \left\langle \tilde{\mathbf{w}}_c, -d^2G_{\mathbf{y}_p, \mathbf{b}_p} \left(\tilde{\mathbf{v}}_c, J_{(0, \mathbf{u})}^{-1} E \left(\frac{\partial G}{\partial \beta_q} \right) \right) \right\rangle. \quad (11.18)$$

E is a projection onto the range of $J_{(0, \mathbf{u})}$ and $J_{(0, \mathbf{u})}^{-1} : \mathbf{v}_c^\perp \mapsto \mathbb{R}^{N_a}$ is the inverse of to restriction of $J_{(0, \mathbf{u})}$ to the orthogonal complement to \mathbf{v}_c . We find that $J_{(0, \mathbf{u})}^{-1} E \left(\frac{\partial G}{\partial \beta_q} \right) = \frac{1}{\mu_q} (\mathbf{v}_q, \mathbf{0})$ and $g_{z\beta_q} = -(\alpha + \lambda_c \gamma) K_q$ where each K_q is the quantity in (11.16). Since $g_{z\beta_q} > 0$ for all q , we conclude that

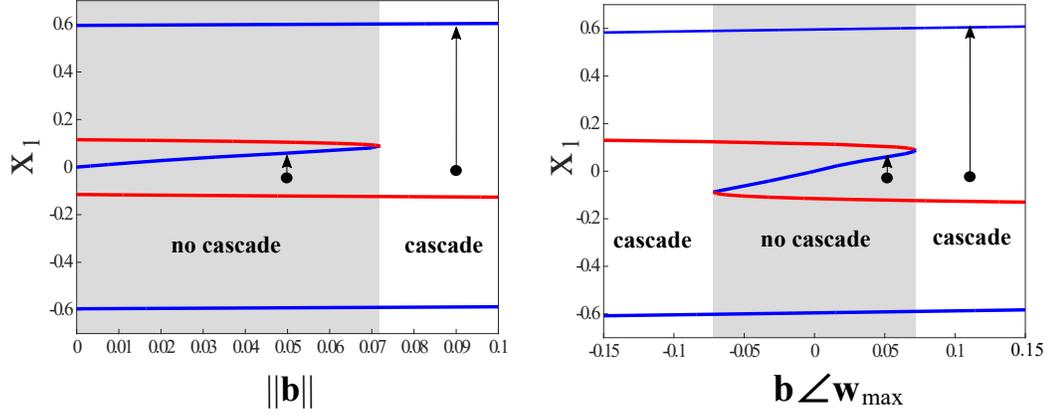


Figure 11.3: Blue (red) lines track the first coordinate of the stable (unstable) equilibrium solutions of the coupled dynamics (11.9) on a 3-agent undirected line graph. Parameters: $u_{th} = 0.1$, $\gamma = 1$, $n = 3$, $\mathbf{b} = \|\mathbf{b}\| \cdot |\mathbf{b} \angle \mathbf{w}_{max}| \cdot \mathbf{w}_{max}$; left: $|\mathbf{b} \angle \mathbf{w}_{max}| = 0.1$, right: $\|\mathbf{b}\| = 0.1$ Bifurcation diagrams generated using MatCont [47].

the eigenvalue of the equilibrium is monotonically increasing with $|\langle \mathbf{w}_c, \mathbf{b} \rangle|$. By continuity of eigenvalues of the perturbed Jacobian, it follows that the leading eigenvalue necessarily crosses zero as input is increased. By [90, Theorem 3.4.1] this singularity must be a saddle-node bifurcation point, with bifurcation parameter $\hat{b} = \langle \mathbf{w}_c, \mathbf{b} \rangle$ and properties outlined in the statement of the theorem. \square

Corollary 11.5.5. *The input magnitude $\|\mathbf{b}\|$ and its relative orientation $\mathbf{b} \angle \mathbf{w}_c := \langle \mathbf{w}_c, \mathbf{b} \rangle / \|\mathbf{b}\|$ can be used as controls to trigger a network opinion cascade.*

Proof. This follows by factoring out the magnitude of the input vector from the bifurcation parameter $\langle \mathbf{w}_c, \mathbf{b} \rangle$. \square

Figure 11.3 illustrates the prediction of Corollary 11.5.5, showing bifurcation diagrams with stable and unstable equilibria of the opinion dynamics in the agreement regime on a small network. The two diagrams illustrate the saddle-node bifurcation predicted by Theorem 11.5.4 with $\|\mathbf{b}\|$ and $\mathbf{b} \angle \mathbf{w}$ as bifurcation parameters. Opinion cascades are activated when the bifurcation parameter passes the critical value. Although the predictions of the results in this section assume inputs are small, in simulation and through numerical continuation of the dynamics on different networks we observe that this result is quite robust. The existence of a saddle-node bifurcation, and therefore a threshold which differentiates between inputs which trigger a cascade and ones which do not, persists across network structures and for large inputs.

A consequence of Theorem 11.5.4 is that also for opinion cascades the node centrality indices are the key determinant of the effect of inputs on the coupled attention-opinion dynamics (11.9): The smaller the angle between the input vector and the agreement or disagreement centrality vector,

the smaller the needed input strength to trigger an agreement or a disagreement opinion cascade. Figure 11.4 numerically illustrates our theoretical prediction. The transition line from the red (no cascade) to the white (cascade) regions correspond to the threshold, i.e., the saddle node bifurcation predicted by Theorem 11.5.4, at which the opinion cascade is ignited. It shows, for different network topologies and agreement and disagreement opinion cascades, that as the angle between the input vector and the centrality vector decreases, the norm of the input needed to trigger a cascade gets smaller. In the cascade region of the simulations in Figure 11.4, the centrality eigenvector accurately predicts the sign distribution among the nodes. Rigorously proving this is subject of future work.

The cascade threshold is implicitly defined by the design parameter y_{th} in the attention saturation function (11.8). In future work we will explore how the sensitivity of the group to distributed input can be tuned with this parameter.

11.6 Final Remarks

We have derived and proved a systematic method for designing distributed inputs to control opinion formation and opinion cascades for both agreement and disagreement among distributed agents that communicate over a network. Future directions include expanding the analysis presented here to multi-option cascades using the general formulation of nonlinear multi-option opinion dynamics of [19].

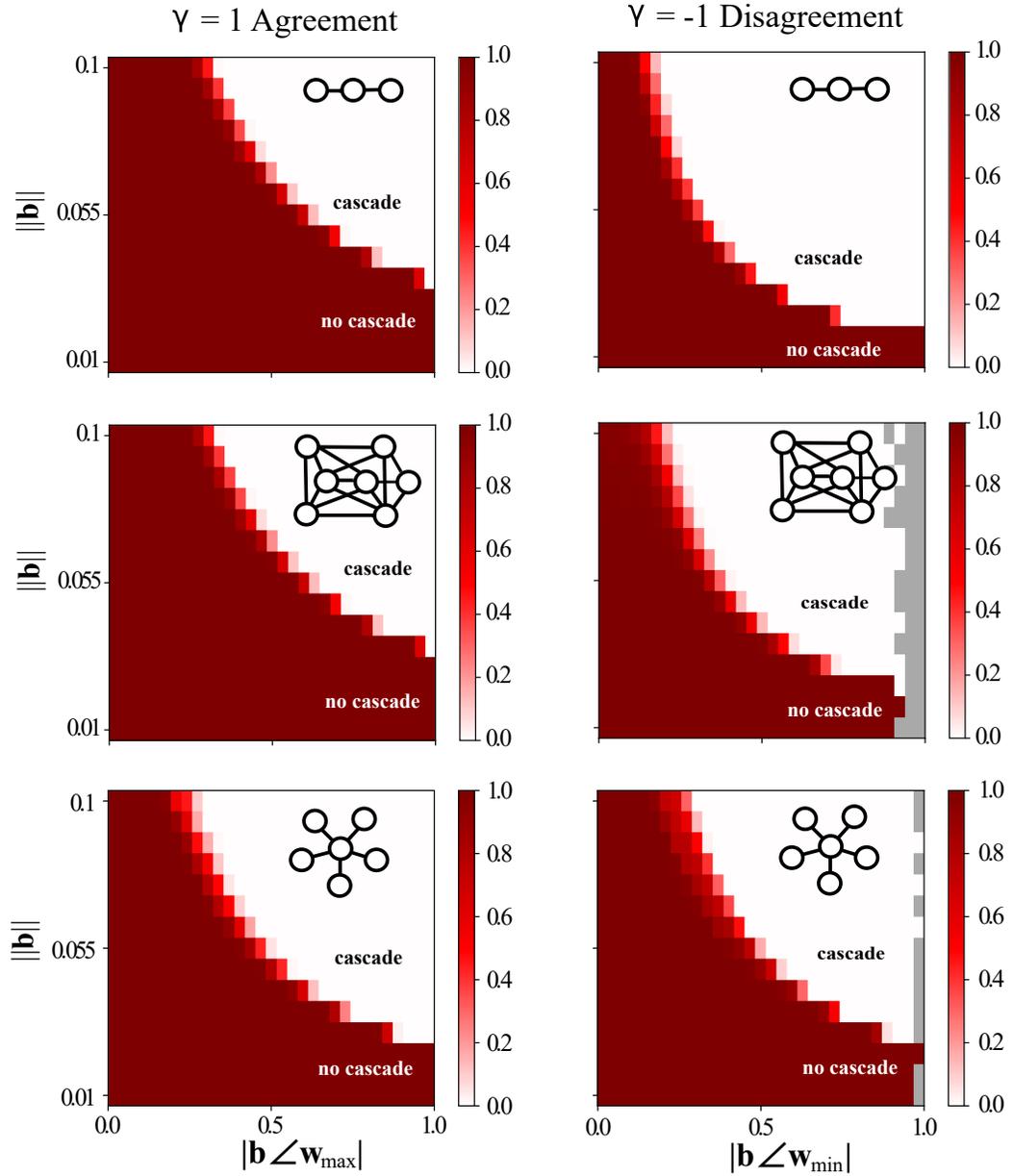


Figure 11.4: Heatmaps with color corresponding to proportion of simulations in the given parameter range that did not result in a network cascade by $t = 500$. Dark red corresponds to no cascades, white to there always being a cascade. Grey squares are bins with no datapoints. Each plot corresponds to 1.5×10^5 distinct simulations on an undirected graph shown in the diagram. Simulation parameters: $\tau_u = 10$, $u_{th} = 0.2$, $\underline{u} = u_a - 0.01$ for $\gamma = 1$ (left plots) and $\underline{u} = u_d - 0.01$ for $\gamma = -1$ (right plots). For each simulation, inputs b_i were drawn from $\mathcal{N}(0, 1)$ and the input vector \mathbf{b} was normalized to a desired magnitude. There were 10000 simulations performed at each constant input magnitude, with 15 magnitudes sampled uniformly spaced between 0 and 0.1. The initial conditions for each simulation were $x_i = 0$, $u_i = 0$ for all $i = 1, \dots, N_a$.

Chapter 12

Switching transformations for decentralized control of opinion patterns in signed networks: application to dynamic task allocation

ANASTASIA BIZYAEVA, GIOVANNA AMORIM, MARÍA SANTOS, ALESSIO FRANCI,
AND NAOMI EHRICH LEONARD

We propose a new decentralized design to control opinion patterns on signed networks of agents making decisions about two options and to switch the network from any opinion pattern to a new desired one. Our method relies on *switching transformations*, which switch the sign of an agent's opinion at a stable equilibrium by flipping the sign of the interactions with its neighbors. The global dynamical behavior of the switched network can be predicted rigorously when the original, and thus the switched, networks are structurally balanced. Structural balance ensures that the network dynamics are monotone, which makes the study of the basin of attraction of the various opinion patterns amenable to monotone systems theory. We illustrate the utility of the approach through scenarios motivated by multi-robot coordination and dynamic task allocation.

12.1 Introduction

Modern networked technologies require decentralized mechanisms for decision-making and allocation of tasks. For example, systems such as smart power grids, cloud computing services, or multi-robot teams, call for strategies that dynamically distribute tasks among individual units to optimize system performance even as task requirements change or units experience failure.

We use the model of networked nonlinear opinion dynamics of [19, 87] to illustrate how network interconnection topology can be designed so a group of decision makers converges to a desired opinion pattern and how the network can be transformed so the group switches to a desired alternative opinion pattern. When all agents commit to the same option the network is in ‘agreement’, while for any other opinion configuration it is in ‘disagreement’. The emergence of agreement and disagreement in nonlinear opinion networks has been studied in [19–21]. However, the analysis in those works has assumed that all network interactions are either positive or negative. In this paper we add to this body of analysis by allowing mixed-sign interactions.

Decision-making with signed interactions has been studied on linear networks with averaging dynamics [5, 127], as well as with nonlinear consensus models [61, 63] and biased assimilation models [209]. *The novelty of our approach is to use signed interactions on a network as a design tool.* Our design methodology drives a distributed system to a desired network state and allows any individual agent to respond to local contextual changes and adjust its allocation by dynamically adjusting the sign of interaction with its neighbors. Since the strategy relies only on pairwise interactions between neighboring agents, it is decentralized and agnostic to the global topology of the network communication graph.

Our contributions are as follows. First, we prove that a network system can be easily and intuitively controlled to any agreement or disagreement opinion pattern using standard tools from signed graph theory grounded in switching transformations of graphs. Second, we prove a sufficient condition for the networked state to converge to one of two available equilibrium configurations. Third, we show how a pattern of equilibrium opinions can be changed dynamically through local updates of the network weights that follow the structure of a switching transformation. Fourth, we validate the theory with simulation examples.

In Section II we introduce notation. Section III describes the opinion dynamics model and summarizes some of its properties. In Section IV we present new analysis of the model on signed graphs and propose a systematic design approach for agent allocation across two tasks. In Section V we describe the asymptotic dynamics of trajectories on structurally balanced graphs. Section VI

relates the features of the approach in the context of multi-robot task allocation. Final remarks are included in Section VII.

12.2 Notation and Mathematical Preliminaries

For any vectors $\mathbf{x} = (x_1, \dots, x_N) \in \mathbb{R}^N$, $\mathbf{y} = (y_1, \dots, y_N) \in \mathbb{R}^N$, the standard Euclidean inner product is $\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^N x_i y_i$. Let $\mathbf{x} \succeq \mathbf{y}$ if $x_i \geq y_i$ for all $i = 1, \dots, N$, and $\mathbf{x} \succ \mathbf{y}$ if $x_i > y_i$ for all $i = 1, \dots, N$. Define the operation \odot as the element-wise product of vectors, $\mathbf{x} \odot \mathbf{y} = (x_1 y_1, \dots, x_N y_N)$. $\mathbf{0}_N$ denotes the vector with all zero entries in \mathbb{R}^N , and \mathcal{I}_N the identity matrix in $\mathbb{R}^{N \times N}$.

We study networks of N agents with a signed communication graph $\mathcal{G} := (\mathcal{V}, \mathcal{E}, \sigma)$ where $\mathcal{V} = \{1, \dots, N\}$ is the *vertex set*, \mathcal{E} is the *edge set*, and $\sigma : \mathcal{E} \rightarrow \{1, -1\}$ is a *sign function* or *signature* of the graph \mathcal{G} . We use the sensing convention such that $e_{ik} \in \mathcal{E}$ denotes a directed edge in \mathcal{G} that points from vertex i to vertex k , indicating that k is a neighbor of i . We assume that the unsigned directed graph $\Gamma = (\mathcal{V}, \mathcal{E})$ underlying \mathcal{G} is *simple*, i.e. contains no self-loops $e_{ii} \notin \mathcal{E}$ for all $i \in \mathcal{V}$, and there is at most one edge e_{ik} in \mathcal{E} that begins at vertex i and ends at vertex k for all $i, k \in \mathcal{V}$. We say that the graph \mathcal{G} is *strongly connected* if the edges contained in \mathcal{E} form a path between any two nodes.

Define $A = (a_{ik})$ to be the $N \times N$ *signed adjacency matrix* of \mathcal{G} whose entries $a_{ik} \in \{0, 1, -1\}$ satisfy $a_{ik} = 0$ if $e_{ik} \notin \mathcal{E}$ and $a_{ik} = \sigma(e_{ik})$ if $e_{ik} \in \mathcal{E}$. We use the symbol λ^* to distinguish, when it exists, the real and unique eigenvalue of A that satisfies $\text{Re}(\lambda^*) > \text{Re}(\lambda_i)$ for all eigenvalues $\lambda_i \neq \lambda^*$ of A . We denote the right and left eigenvectors of A corresponding to λ^* as \mathbf{v}^* and \mathbf{w}^* , respectively. We always assume $\mathbf{v}^*, \mathbf{w}^*$ are normalized to satisfy $\langle \mathbf{w}^*, \mathbf{v}^* \rangle = 1$. We adapt the statement of the standard Perron-Frobenius theorem, e.g. as presented in [195], to specialize to adjacency matrices of graphs with an all-positive signature.

Proposition 12.2.1 (Perron-Frobenius). *Suppose $\sigma(e_{ik}) = 1$ for all $e_{ik} \in \mathcal{E}$ for some strongly connected graph \mathcal{G} . Then the following hold: 1) λ^* exists; 2) $\lambda^* > 0$; 3) we can choose $\mathbf{v}^*, \mathbf{w}^*$ to satisfy $\mathbf{v}^* \succ \mathbf{0}_N$ and $\mathbf{w}^* \succ \mathbf{0}_N$.*

12.3 Nonlinear Opinion Dynamics Model

The evolution of the opinion of N agents on a signed network choosing between two options is modeled in this paper according to the continuous-time multi-agent, multi-option nonlinear opinion

dynamics model in [19].

Let $x_i \in \mathbb{R}$ denote the opinion of agent i , where the magnitude of x_i determines the agent's commitment to one of the two options such that a stronger (weaker) commitment to an option corresponds to a larger (smaller) $|x_i|$. If $x_i = 0$, the agent is said to be unopinionated, and, if $x_i > 0$ (< 0), agent i prefers option 1 (option 2). We define the opinion state of the network as $\mathbf{x} = (x_1, \dots, x_N) \in \mathbb{R}^N$, with $\mathbf{x} = \mathbf{0}_N$ being the neutral state of the group. The network is in an agreement state when $\text{sign}(x_i) = \text{sign}(x_k)$ for all $i, k \in \{1, \dots, N\}$ (i.e., if all the agents commit to the same option), and in a disagreement state otherwise.

The evolution of agent i 's opinion is determined by a linear damping term and a saturated network interaction term

$$\dot{x}_i = -d x_i + u_i S \left(\alpha x_i + \gamma \sum_{\substack{k=1 \\ k \neq i}}^N a_{ik} x_k \right), \quad (12.1)$$

where $d > 0$ is the damping coefficient, $u_i > 0$ regulates the relative strength of the two terms, and the odd saturating function $S : \mathbb{R} \rightarrow \mathbb{R}$ acts on network interactions. Furthermore, S satisfies $S(0) = 0$, $S'(0) = 1$, $\text{sign}(S''(x)) = -\text{sign}(x)$.¹ Network interactions comprise self-reinforcement interactions, weighted by $\alpha \geq 0$, and neighbor interactions, weighted by $\gamma > 0$. The sign of network interactions is determined by the signed adjacency weight $a_{ik} \in \{0, 1, -1\}$. Agent i cooperates (competes) with agent k when $a_{ik} = 1$ ($= -1$) and is indifferent to agent k 's opinion when $a_{ik} = 0$. In vector form, dynamics (12.1) are

$$\dot{\mathbf{x}} = -d \mathbf{x} + U S(\alpha \mathbf{x} + \gamma \mathbf{A} \mathbf{x}), \quad (12.2)$$

where $S(\mathbf{y}) := (S(y_1), \dots, S(y_N))$ for any $\mathbf{y} \in \mathbb{R}^N$ and $U = \text{diag}(u_1, \dots, u_N)$.

12.3.1 Network opinion formation through bifurcation

The following proposition, adapted from [19, Theorem IV.1] and [21, Theorem IV.1] and stated without proof, summarizes several key features of the opinion dynamics (12.1).

Proposition 12.3.1 (Opinion formation as a pitchfork bifurcation). *Consider (12.1) on a graph \mathcal{G} with $u_i = u \geq 0$, $\alpha \geq 0$, $\gamma > 0$, $d > 0$ for all $i = 1, \dots, N$, and assume a simple, real largest eigenvalue λ^* exists. Suppose $\alpha + \gamma \lambda^* > 0$ and $\langle \mathbf{w}^*, (\mathbf{v}^*)^3 \rangle > 0$, where $(\mathbf{v}^*)^3 = \mathbf{v}^* \odot \mathbf{v}^* \odot \mathbf{v}^*$. Then*

¹The presence of non-smooth (piece-wise linear) saturation functions can be tackled using methods from non-smooth analysis [119] and recent bifurcation-theoretical tools for linear complementarity systems [138].

1) for $0 \leq u < u^* := d/(\alpha + \gamma\lambda^*)$, the neutral equilibrium $\mathbf{x} = \mathbf{0}_N$ is locally exponentially stable; 2) for $u > u^*$, $\mathbf{x} = \mathbf{0}_N$ is unstable, and two branches of locally exponentially stable equilibria $\mathbf{x} = \mathbf{x}_1^*$, \mathbf{x}_2^* branch off from $(\mathbf{x}, u) = (\mathbf{0}_N, u^*)$ along a manifold tangent at $\mathbf{x} = \mathbf{0}_N$ to $\text{span}(\mathbf{v}^*)$. The two nonzero equilibria differ by a sign, i.e. $\mathbf{x}_2^* = -\mathbf{x}_1^*$.

Proposition 12.3.1 shows that a group of decision-makers with opinion dynamics (12.1) can break deadlock and commit to an opinionated configuration when their level of attention u_i is sufficiently large. Fig. 12.1 provides a graphical illustration of the equilibria and their stability as a function of u in the form of the pitchfork bifurcation described by the proposition. Next we state other useful properties of (12.1).

Corollary 12.3.2 (Sufficient condition for agreement). *When \mathcal{G} is strongly connected with an all-positive signature, conditions of Proposition 12.3.1 are always satisfied. For $u > u^*$, one of the two new stable equilibria satisfies $\mathbf{x}^* \succ \mathbf{0}_N$.*

Proof. The corollary follows from Proposition 12.3.1 and Proposition 12.2.1, since λ^* is the Perron-Frobenius eigenvalue, and eigenvectors $\mathbf{w}^*, \mathbf{v}^*$ have all-positive entries. \square

We next show that the equilibria predicted by Proposition 12.3.1 are the only equilibria admitted by dynamics (12.1) for a range of values of u , following similar arguments as those used for Laplacian-weighted nonlinear consensus networks in [61]. We first state a necessary lemma.

Lemma 12.3.3 (Boundedness). *Any compact set $\Omega_r \subset \mathbb{R}^N$ of the form $\Omega_r = \{\mathbf{x} \in \mathbb{R}^N \text{ s.t. } |x_i| < r \max_j \{u_j\}/d, \forall i, j \in \mathcal{V}\}$ with $r > 1$ is forward-invariant for (12.1).*

Proof. The lemma follows directly from the more general result [19, Theorem A.2]. \square

Corollary 12.3.4 (Uniqueness of Equilibria). *Suppose conditions of Proposition 12.3.1 are satisfied, and let λ_2 be an eigenvalue of A satisfying $\text{Re}(\lambda_2) \geq \text{Re}(\lambda_i)$ for all eigenvalues $\lambda_i \neq \lambda^*$ of A . 1) $\mathbf{x} = \mathbf{0}_N$ is globally asymptotically stable on a forward-invariant compact set $\Omega \subset \mathbb{R}^n$ containing the origin $\mathbf{x} = \mathbf{0}_N$, for all $u \in [0, u^*)$; 2) when $\text{Re}(\lambda_2) \geq -\alpha/\gamma$, $u \in (u^*, u_2)$, the only equilibria the system admits are $\mathbf{0}_N$, \mathbf{x}_1^* , and \mathbf{x}_2^* , where $u_2 = d/(\alpha + \gamma \text{Re}(\lambda_2))$; 3) when $\text{Re}(\lambda_2) < -\alpha/\gamma$, the only equilibria the system admits in Ω for all $u > u^*$ are $\mathbf{0}_N$, \mathbf{x}_1^* , and \mathbf{x}_2^* .*

Proof. 1) Existence of Ω , and thereby boundedness of solutions of (12.1), is established in Lemma 12.3.3. Define $\tilde{A} = \alpha\mathcal{I}_N + \gamma A$ with components \tilde{a}_{ij} , and let $f_i(\mathbf{x}) = \sum_{j=1}^N \tilde{a}_{ij}x_j$. Consider the continuously differentiable function $V(\mathbf{x}) = \sum_{i=1}^N \int_0^{f_i(\mathbf{x})} S(\eta)d\eta$. Along trajectories of (12.1), $\dot{V}(\mathbf{x}) = S(\tilde{A}\mathbf{x})^T \tilde{A}\dot{\mathbf{x}} = S(\tilde{A}\mathbf{x})^T \tilde{A}(-d\mathbf{x} + uS(\tilde{A}\mathbf{x})) = -dS(\tilde{A}\mathbf{x})^T \tilde{A}\mathbf{x} + uS(\tilde{A}\mathbf{x})^T \tilde{A}S(\tilde{A}\mathbf{x}) \leq -S(\tilde{A}\mathbf{x})^T (d\mathcal{I}_N -$

$u\tilde{A})S(\tilde{A}\mathbf{x})$ (using $|S(y)| \leq |y|$ and $\text{sign}(S(y)) = \text{sign}(y)$). Since $d\mathcal{L}_N - u\tilde{A}$ is positive definite for $u \in [0, u^*)$,

$$\dot{V}(\mathbf{x}) \leq -(d - u(\alpha + \gamma\lambda^*))S(\tilde{A}\mathbf{x})^T S(\tilde{A}\mathbf{x}) \leq 0. \quad (12.3)$$

The set on which (12.3) is exactly zero is $\mathcal{N}(\tilde{A}) = \{\mathbf{x} \in \mathbb{R}^N \text{ s.t. } \tilde{A}\mathbf{x} = \mathbf{0}_N\}$. By LaSalle's invariance principle [109, Theorem 4.4] we conclude that the trajectories $\mathbf{x}(t)$ approach the largest invariant set in $\mathcal{N}(\tilde{A})$ as $t \rightarrow \infty$. If $\mathcal{N}(\tilde{A}) = \{\mathbf{0}_N\}$, the corollary follows trivially. Let $\mathbf{x} \in \mathcal{N}(\tilde{A})$ and suppose $\mathbf{x} \neq \mathbf{0}_N$. Then $\dot{\mathbf{x}} = -d\mathbf{x}$, i.e. all trajectories that start in $\mathcal{N}(\tilde{A})$ decay to the origin exponentially in time, and the corollary follows. Under the assumptions on u stated in 2) and 3), the Jacobian matrix $J(\mathbf{x}) = -d\mathcal{L}_N + u \text{diag}(S'(\tilde{A}\mathbf{x}))\tilde{A}$ is Hurwitz for all $\mathbf{x} \in \mathbb{R}^N \setminus \{\mathbf{0}_N\}$, the proof of which follows closely the argument presented in [63, Lemma 6] and we omit its details. By Proposition 12.3.1, for values of u in a small neighborhood above u^* exactly three equilibria exist. Since the Jacobian is nonsingular for all $\mathbf{x} \in \Omega$ and all $u \in (u^*, u_2)$, by the implicit function theorem, the number of equilibria remains unchanged. \square

12.4 Switching Transformation as a Design Tool for Synthesis of Opinion Patterns

When the communication graph \mathcal{G} contains edges with a negative signature and its adjacency matrix A has a simple leading eigenvalue, the opinion-forming bifurcation of Proposition 12.3.1 results in disagreement network equilibria. We describe a simple synthesis procedure for generating a signed adjacency matrix that results in a desired pattern of opinions among the decision-makers following opinion dynamics (12.1). We first introduce a few important concepts from the theory of signed graphs; for more details on the theory we refer the reader to [223] and [224].

12.4.1 Signed graphs and switching

Let $\mathcal{W} \subset \mathcal{V}$ be a subset of nodes on a signed graph \mathcal{G} . *Switching* a set \mathcal{W} on the graph \mathcal{G} refers to a mapping of the graph \mathcal{G} to $\mathcal{G}^{\mathcal{W}} = (\mathcal{V}, \mathcal{E}, \sigma_{\mathcal{W}})$ where the signature of all the edges in \mathcal{E} between nodes in \mathcal{W} and nodes in its complement $\mathcal{V} \setminus \mathcal{W}$ reverses sign. We introduce the *switching function* $\theta : \mathcal{V} \rightarrow \{1, -1\}$, where for any $i \in \mathcal{V}$, $\theta(i) = -1$ if $i \in \mathcal{W}$ and $\theta(i) = 1$ otherwise. Then the signature of the switched graph $\mathcal{G}^{\mathcal{W}}$ is generated as

$$\sigma_{\mathcal{W}}(e_{ik}) = \theta(i)\sigma(e_{ik})\theta(k) \quad (12.4)$$

for all $e_{ik} \in \mathcal{E}$. From (12.4) we see that the signature update for an edge between agents i and k depends only on their membership in the switching set \mathcal{W} . Thus, the edges between i and k flip sign if and only if exactly one of i, k is in the switching set \mathcal{W} , and does not change sign if i, k are both in \mathcal{W} or in $\mathcal{V} \setminus \mathcal{W}$. Importantly, switching a set \mathcal{W} all at once generates the same graph $\mathcal{G}^{\mathcal{W}}$ as sequentially switching individual vertices in \mathcal{W} . If \mathcal{G} can be transformed into $\mathcal{G}^{\mathcal{W}}$ by switching, \mathcal{G} and $\mathcal{G}^{\mathcal{W}}$ are *switching equivalent graphs*.

Let θ be the function for switching from graph \mathcal{G} to $\mathcal{G}^{\mathcal{W}}$, with adjacency matrices A and $A^{\mathcal{W}}$, respectively. Define the *switching matrix* $\Theta = \text{diag}(\theta(1), \theta(2), \dots, \theta(N))$. The adjacency matrices of \mathcal{G} and its switching $\mathcal{G}^{\mathcal{W}}$ are related as

$$A^{\mathcal{W}} = \Theta^{-1} A \Theta. \quad (12.5)$$

Since Θ is diagonal and $\theta(i) = \pm 1$, $\Theta^{-1} = \Theta$. We refer to (12.5) as a *switching transformation* of the adjacency matrix A , and A and $A^{\mathcal{W}}$ as *switching equivalent adjacency matrices*.

Proposition 12.4.1. *Suppose \mathcal{G} , $\mathcal{G}^{\mathcal{W}}$ are switching equivalent with adjacency matrices A and $A^{\mathcal{W}}$ and associated switching matrix Θ . Then 1) A and $A^{\mathcal{W}}$ are isospectral, i.e. have the same set of eigenvalues; 2) \mathbf{v} (\mathbf{w}) is a right (left) eigenvector of A corresponding to eigenvalue λ if and only if $\Theta \mathbf{v}$ ($\Theta \mathbf{w}$) is a right (left) eigenvector of $A^{\mathcal{W}}$ with the same eigenvalue.*

Proof. The proposition follows from the standard properties of a matrix similarity transformation, since A and $A^{\mathcal{W}}$ are related through a similarity transformation (12.5). \square

Proposition 12.4.1 implies that the eigenvectors of the switched adjacency matrix $A^{\mathcal{W}}$ are obtained from the eigenvectors of the original adjacency matrix A by flipping the sign of each entry that corresponds to a node which is being switched. We will take advantage of this observation in our design of nonlinear opinion patterns on a network.

12.4.2 Nonlinear opinion patterns on switch equivalent graphs

In this section we show that a switching transformation of the nonlinear opinion dynamics (12.1) is effectively a coordinate change, and two switching equivalent networks generate topologically equivalent flow and bifurcation diagrams.

Theorem 12.4.2 (Diffeomorphism between trajectories of switching equivalent systems). *Consider switching equivalent graphs \mathcal{G} , $\mathcal{G}^{\mathcal{W}}$ with adjacency matrices A and $A^{\mathcal{W}}$ and with switching matrix Θ . The trajectory $\mathbf{x}(t)$ is a solution to (12.1) on \mathcal{G} if and only if $\Theta \mathbf{x}(t)$ is a solution of (12.1) on $\mathcal{G}^{\mathcal{W}}$.*

Proof. Suppose $\mathbf{x}(t)$ is a solution of (12.1) on \mathcal{G} . Multiplying both sides of (12.2) by the switching matrix Θ yields

$$\begin{aligned} \frac{d}{dt}(\Theta\mathbf{x}(t)) &= \Theta(-d\mathbf{x}(t) + US(\alpha\mathbf{x}(t) + \gamma A\mathbf{x}(t))) \\ &= -d\Theta\mathbf{x}(t) + \Theta US(\alpha\mathbf{x}(t) + \gamma A\mathbf{x}(t)) \\ &= -d\Theta\mathbf{x}(t) + US(\alpha\Theta\mathbf{x}(t) + \gamma A^{\mathcal{W}}\Theta\mathbf{x}(t)), \end{aligned}$$

where the last step follows since $\Theta U = U\Theta$ and $-S(y) = S(-y)$. This shows that $\Theta\mathbf{x}(t)$ is a solution of (12.1) on $\mathcal{G}^{\mathcal{W}}$. The other direction follows by an identical proof. \square

Corollary 12.4.3 (Switching a graph “rotates” a pitchfork bifurcation). *Consider (12.1) with $u_i = u \geq 0$ for all $i = 1, \dots, N$ on the graphs described in Theorem 12.4.2. Suppose \mathcal{G} satisfies the conditions of Proposition 12.3.1. Then $\mathcal{G}^{\mathcal{W}}$ also satisfies the conditions of Proposition 12.3.1. Furthermore, \mathbf{x}^* is an equilibrium on the bifurcation diagram on \mathcal{G} at some u if and only if $\Theta\mathbf{x}^*$ is an equilibrium on the bifurcation diagram of $\mathcal{G}^{\mathcal{W}}$ at the same u .*

Proof. For (12.1) on $\mathcal{G}^{\mathcal{W}}$, λ^* is simple and $\alpha + \gamma\lambda^* > 0$ because $\mathcal{G}, \mathcal{G}^{\mathcal{W}}$ are isospectral. Additionally, $\langle \Theta\mathbf{w}^*, (\Theta\mathbf{v}^*)^3 \rangle = \sum_{i=1}^N \theta(i)^4 w_i^* (v_i^*)^3 = \sum_{i=1}^N w_i^* (v_i^*)^3 = \langle \mathbf{w}^*, \mathbf{v}^* \rangle > 0$ since $\theta(i) = \pm 1$ for all $i \in \mathcal{V}$ and therefore the condition of Proposition 12.3.1 are satisfied. The rest of the corollary statement follows as a direct consequence of Theorem 12.4.2. \square

We illustrate the intuition of Corollary 12.4.3 in Fig. 12.1.

Theorem 12.4.4 (Switching complementary vertex sets generates the same flow). *Consider two switching equivalent graphs $\mathcal{G}^{\mathcal{W}}, \mathcal{G}^{\mathcal{V}\setminus\mathcal{W}}$, generated by switching a set of vertices \mathcal{W} or its complement $\mathcal{V}\setminus\mathcal{W}$ on graph \mathcal{G} . Let the switching matrices in relation to \mathcal{G} of these two graphs be $\Theta^{\mathcal{W}}$ and $\Theta^{\mathcal{V}\setminus\mathcal{W}}$ respectively. The trajectory $\mathbf{x}(t)$ is a solution of (12.1) on $\mathcal{G}^{\mathcal{W}}$ if and only if it is also a solution of (12.1) on $\mathcal{G}^{\mathcal{V}\setminus\mathcal{W}}$.*

Proof. Suppose $\mathbf{x}(t)$ is a solution of (12.1) on $\mathcal{G}^{\mathcal{W}}$. Then by Theorem 12.4.2, $\Theta^{\mathcal{W}}\mathbf{x}(t)$ is an equilibrium of (12.1) on \mathcal{G} . Applying the complementary switching transformation, and observing that $\Theta^{\mathcal{V}\setminus\mathcal{W}}\Theta^{\mathcal{W}} = -\mathcal{I}_N$, we see that $\Theta^{\mathcal{V}\setminus\mathcal{W}}\Theta^{\mathcal{W}}\mathbf{x}(t) = -\mathbf{x}(t)$ is a solution of (12.1) on $\mathcal{G}^{\mathcal{V}\setminus\mathcal{W}}$. By odd symmetry of the dynamic equations (12.1), $\mathbf{x}(t)$ is also a solution of (12.1) on $\mathcal{G}^{\mathcal{V}\setminus\mathcal{W}}$. The proof of the converse follows the same steps in opposite order. \square

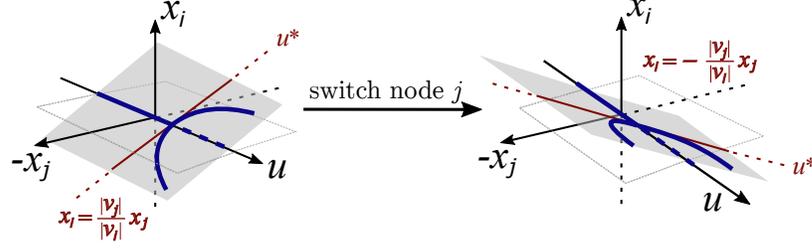


Figure 12.1: Illustration of Corollary 12.4.3. The bifurcation diagram of the switched system is a “rotated” version of the original diagram because the sign of v_j flips.

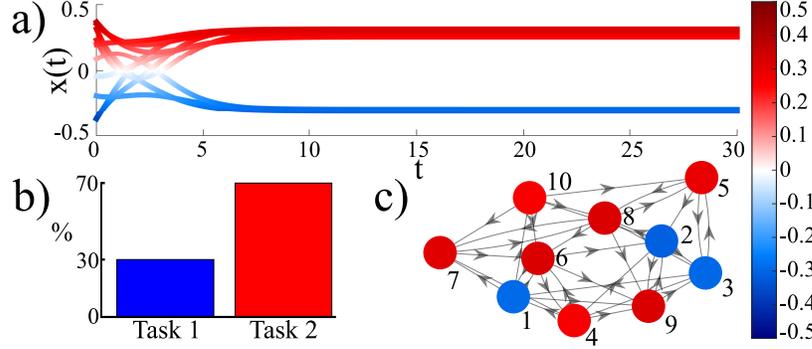


Figure 12.2: Assigning 10 agents to a 30-70% distribution by switching agents 1, 2 and 3. (a) Time trajectory of the opinion dynamics. (b) Final agent distribution. (c) Network diagram with the opinion of each agent at $t = 30$. Parameters: $S = \tanh$, $d = 1$, $\alpha = 1.2$, $\gamma = 1.3$, $u = 0.324$. Arrows on the graph defined by the sensing convention.

12.4.3 Synthesis of nonlinear opinion patterns

The theoretical results of Section 12.4.2, lead to a design procedure to build a signed adjacency matrix that ensures a desired allocation of agents across the two options. **Step 1.** Start with a strongly connected \mathcal{G} with an all-positive signature, i.e. $a_{ik} \in \{0, 1\}$ for all $i, k \in \mathcal{V}$. By Corollary 12.3.2, (12.1) on \mathcal{G} has an all-positive stable equilibrium \mathbf{x}_1^* and all-negative stable equilibrium \mathbf{x}_2^* . **Step 2.** Define the switching set \mathcal{W} . In this step, the designer chooses which nodes are grouped together. The two partitions correspond to the two tasks. **Step 3.** Update edge signatures of \mathcal{G} locally as $a_{ik}^{\mathcal{W}} = \theta(i)a_{ik}\theta(k)$. This edge signature update generates the switch-equivalent graph $G^{\mathcal{W}}$ and groups all nodes in \mathcal{W} and all nodes in $\mathcal{V} \setminus \mathcal{W}$ together by sign, i.e. the dynamics (12.1) on $G^{\mathcal{W}}$ is bistable with stable equilibria $\Theta\mathbf{x}_1^*$, $\Theta\mathbf{x}_2^*$. If $|\mathcal{W}| = M$, the equilibrium $\Theta\mathbf{x}_1^*$ corresponds to M negative nodes, and $\Theta\mathbf{x}_2^*$ to $N - M$ positive nodes. We illustrate this in Fig. 12.2. Step 3 can also be implemented in a decentralized manner since it only relies on the pairwise switching states of neighboring agents.

12.5 Dynamic Switching

We next investigate the asymptotic opinion dynamics of (12.1) when the underlying communication graph \mathcal{G} instantaneously changes to a switching equivalent graph $\mathcal{G}^{\mathcal{W}}$.

12.5.1 Monotonicity and structural balance

First, we introduce some relevant definitions from the study of monotone systems. Let \mathcal{K} be an orthant of \mathbb{R}^N , $\mathcal{K} = \{\mathbf{x} \in \mathbb{R}^N \text{ s.t. } (-1)^{m_i} x_i \geq 0, i = 1, \dots, N\}$ with each $m_i \in \{0, 1\}$. The orthant \mathcal{K} generates a partial ordering “ $\leq_{\mathcal{K}}$ ” on \mathbb{R}^N where if $\mathbf{x}, \mathbf{y} \in \mathbb{R}^N$, $\mathbf{y} \leq_{\mathcal{K}} \mathbf{x}$ if and only if $\mathbf{x} - \mathbf{y} \in \mathcal{K}$. We say a system $\dot{\mathbf{x}} = f(\mathbf{x})$ on $\mathcal{U} \subseteq \mathbb{R}^N$ is *type \mathcal{K} monotone* if its flow preserves the partial ordering $\leq_{\mathcal{K}}$, i.e. if $\mathbf{x}_1(0) \leq_{\mathcal{K}} \mathbf{x}_2(0)$ implies $\mathbf{x}_1(t) \leq_{\mathcal{K}} \mathbf{x}_2(t)$ for all $t > 0$.

Lemma 12.5.1. *Consider (12.1) on a signed graph \mathcal{G} . It is a type \mathcal{K} monotone system if and only if \mathcal{G} is switching equivalent to \mathcal{G}^+ , for which $\sigma(e_{ik}) = 1$ for all $e_{ik} \in \mathcal{E}$, i.e. \mathcal{G} is **structurally balanced**.*

Proof. The off-diagonal terms of the Jacobian matrix $J(\mathbf{x})$ are $u\gamma \text{diag}(S'((\alpha\mathcal{I}_N + \gamma A)\mathbf{x}))A$. Let Θ be the switching matrix between \mathcal{G} and \mathcal{G}^+ . Since $S'(y) > 0$ for all $y \in \mathbb{R}$, the matrix $u\gamma\Theta \text{diag}(S'((\alpha\mathcal{I}_N + \gamma A)\mathbf{x}))A\Theta$ has nonnegative components, and the lemma follows by [187, Lemma 2.1]. \square

12.5.2 Instantaneous switching

Suppose \mathbf{x}^* is a hyperbolic equilibrium of (12.1), i.e. the linearization of the system at \mathbf{x}^* has m unstable eigenvalues and $N - m$ stable eigenvalues. Then by [90, Theorem 1.3.2], there exist smooth local unstable and stable manifolds $W_{loc}^u(\mathbf{x}^*)$, $W_{loc}^s(\mathbf{x}^*)$ of dimensions m , $N - m$ that are tangent to the unstable and stable eigenspaces of the linearized systems at \mathbf{x}^* and invariant under the dynamics. Global stable and unstable manifolds $W^s(\mathbf{x}^*)$, $W^u(\mathbf{x}^*)$ invariant under the dynamics can be obtained by continuing the trajectories in their local counterparts forwards or backwards in time.

Assumption 12.5.2 (Stable manifold of origin is bounded; Fig. 12.3). *Consider (12.1) on some structurally balanced graph \mathcal{G} with $u_i = u > u^*$ and $u < u_2$ when appropriate, as defined in Corollary 12.3.4. Let $\mathcal{U}' \subset \mathbb{R}^N$ be an open neighborhood containing the origin, and let $\mathbf{x} \in W^s(\mathbf{0}) \cap \mathcal{U}'$. 1) $|\langle \mathbf{w}^*, \mathbf{x} \rangle| < \varepsilon \|\mathbf{x}\|^2$ for some $0 < \varepsilon < 1$; 2) for equilibria $\mathbf{x}_k^* \neq \mathbf{0}$ of Proposition 12.3.1 with $k \in \{1, 2\}$, $|\langle \mathbf{w}^*, \mathbf{x}_k^* \rangle| > \varepsilon \|\mathbf{x}_k^*\|^2$.*

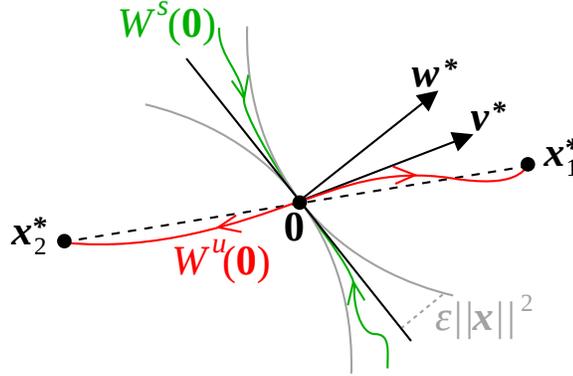


Figure 12.3: Geometric intuition behind Assumption 12.5.2. The one-dimensional unstable manifold $W^u(\mathbf{0})$ of the origin (shown in red) forms heteroclinic orbits with the stable equilibria $\mathbf{x}_1^*, \mathbf{x}_2^*$, as is generically the case for monotone systems - see [187, Theorem 2.8].

Estimating the ε bound described above requires a lengthy computation of the stable manifold approximation (see [90, p.132] for an example of an invariant manifold approximation) which we do not carry out for space considerations. However, this assumption should hold at least locally as a consequence of the (Un)Stable Manifold Theorem [90, Theorem 1.3.2] and monotonicity of the flow. We verified the assumption numerically for several graphs. For example, numerically we find $\varepsilon = 0.05$ to be a valid bound for the graph and parameter values in Fig. 12.5 with \mathbf{w}^* normalized to unit norm; in general ε will vary with u, d, α, γ .

Lemma 12.5.3 (Regions of attraction). *Consider (12.1) on some structurally balanced graph \mathcal{G} with $u_i = u > u^*$ for all $i = 1, \dots, N$, on an open and bounded neighborhood Ω_r as defined in Lemma 12.3.3. Let $\mathbf{x}_1^*, \mathbf{x}_2^*$ be the nonzero equilibria described in Proposition 12.3.1 with $\langle \mathbf{w}^*, \mathbf{x}_1^* \rangle > 0$. Consider an initial condition $\mathbf{x}(0)$ at $t = 0$. If $\langle \mathbf{w}^*, \mathbf{x}(0) \rangle > \varepsilon \|\mathbf{x}(0)\|^2 (< -\varepsilon \|\mathbf{x}(0)\|^2)$ then as $t \rightarrow \infty$, $\mathbf{x}(t) \rightarrow \mathbf{x}_1^* (\mathbf{x}_2^*)$.*

Proof. We established in Corollary 12.3.4 that the only equilibria the system admits are $\mathbf{0}, \mathbf{x}_1^*, \mathbf{x}_2^*$, and Ω_r is positively invariant by Lemma 12.3.3. Let $B(\mathbf{x}_i^*)$ be the basin of attraction of equilibrium \mathbf{x}_i in Ω_r . By monotonicity (Lemma 12.5.1) and [187, Theorem 2.6], the set $\text{Int}(B(\mathbf{x}_1^*)) \cup \text{Int}(B(\mathbf{x}_2^*))$ is open and dense in Ω_r , where Int signifies the interior points. Then following Assumption 12.5.2, the stable manifold partitions Ω_r into the basins of attraction of the two locally asymptotically stable equilibria. The sets $U_+ = \{\mathbf{x} \in \Omega_r \text{ s.t. } \langle \mathbf{w}^*, \mathbf{x} \rangle > \varepsilon \|\mathbf{x}\|^2\}$, $U_- = \{\mathbf{x} \in \Omega_r \text{ s.t. } \langle \mathbf{w}^*, \mathbf{x} \rangle < -\varepsilon \|\mathbf{x}\|^2\}$ do not intersect the center manifold and are therefore positively invariant. Then since $\mathbf{x}_1 \in U_+$ and $\mathbf{x}_2 \in U_-$, we get that $U_+ \subset B(\mathbf{x}_1^*)$ and $U_- \subset B(\mathbf{x}_2^*)$. \square

Remark 12.5.4. *In practice, without a precise value for the bound ε from Assumption 12.5.2, for*

most points $\mathbf{x}(0) \in \mathbb{R}^N$ it is sufficient to check whether the projection of $\mathbf{x}(0)$ onto \mathbf{w}^* is positive or negative to determine which region of attraction the points belongs to, i.e. $\langle \mathbf{w}^*, \mathbf{x}(0) \rangle > 0 (< 0)$ where $>$ implies convergence to \mathbf{x}_1^* and $<$ to \mathbf{x}_2^* . This is because the stable manifold that partitions the space of possible opinion configurations occurs near the plane of points normal to \mathbf{w}^* at the origin; see Fig. 12.3 for illustration. As long as $\mathbf{x}(0)$ is not too close to this plane, the projection is a reliable heuristic for the asymptotic dynamics of the network opinions.

Theorem 12.5.5. Consider (12.1) on some \mathcal{G} and let $\mathbf{x}_1^*, \mathbf{x}_2^*$ be the nonzero equilibria described in Proposition 12.3.1, with $\langle \mathbf{w}^*, \mathbf{x}_1^* \rangle > 0$. Let $\mathcal{G}^{\mathcal{W}}$ be switch equivalent to \mathcal{G} with the associated switching matrix Θ . Suppose at $t = 0$, $\mathbf{x}(0)$ is close to \mathbf{x}_i^* where $i = 1$ or 2 . If $|\langle \Theta \mathbf{w}^*, \mathbf{x}_i^* \rangle| > \varepsilon \|\mathbf{x}_i^*\|^2$ and $\langle \Theta \mathbf{w}^*, \mathbf{x}_i \rangle > 0 (< 0)$ then for (12.1) on $\mathcal{G}^{\mathcal{W}}$ as $t \rightarrow \infty$, $\mathbf{x}(t) \rightarrow \Theta \mathbf{x}_i (\rightarrow -\Theta \mathbf{x}_i)$.

Proof. Without loss of generality, let $\|\mathbf{x}(0) - \mathbf{x}_1^*\| < \mu$ so that $\langle \Theta \mathbf{w}^*, \mathbf{x}(0) \rangle > \varepsilon \|\mathbf{x}(0)\|^2$ and $w_i^* x_i(0) > 0$ for all $i \in \mathcal{V}$ (these are true at \mathbf{x}_1 by assumption; sufficiently close nearby points will satisfy the conditions by continuity). By Theorem 12.4.2 we know that that for (12.1) on $\mathcal{G}^{\mathcal{W}}$, $\Theta \mathbf{x}_1$ is an equilibrium, and the vector $\Theta \mathbf{w}^*$ is normal to the stable eigenspace at the origin. The theorem follows by Lemma 12.5.3. \square

Theorem 12.5.5 shows that instantaneously changing a structurally balanced graph \mathcal{G} to its switching equivalent $\mathcal{G}^{\mathcal{W}}$ results in a predictable transition of the system state. Namely, if the number of nodes in \mathcal{W} is small in comparison with the cardinality of \mathcal{V} , we expect that all nodes in \mathcal{W} will change sign, and all of the nodes in $\mathcal{V} \setminus \mathcal{W}$ will not. A simulation example of this behavior is shown in Fig. 12.4. The precise number of nodes that can be switched simultaneously to generate this behavior depends on the eigenvector \mathbf{w}^* of the graph adjacency matrix, the value of the equilibrium \mathbf{x}_1^* , and the bound ε . In practice, it is often sufficient that $|\mathcal{W}| < \frac{1}{2}|\mathcal{V}|$. For the graph and parameter values in Fig. 12.5, and the numerical estimate $\varepsilon = 0.05$, any combination of 4 or fewer nodes can indeed be switched simultaneously.

The analysis in this section reveals that dynamics (12.1) should be well-behaved if the transition between \mathcal{G} and $\mathcal{G}^{\mathcal{W}}$ is driven by smooth dynamics, e.g., a suitably designed feedback law. We consider an example of such a smoothly-driven transition in the following section.

12.6 Applications to Multi-Robot Task Allocation

We illustrate how our method can be used to change the proportion of robots dedicated to a task, how it can be decentralized, how it is robust to individual robot failures or additions, and how we

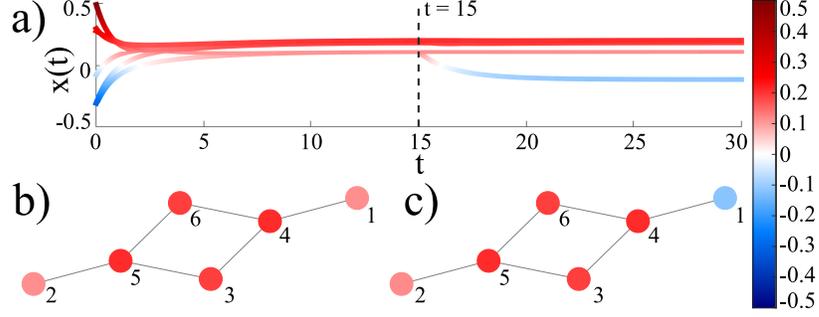


Figure 12.4: Applying a switching transformation to agent 1 at $t = 15$. (a) Time trajectory of the opinion dynamics (b)-(c) Network diagram with the opinion of each agent at $t = 10$, $t = 30$. Parameters: $S = \tanh$, $d = 1$, $\alpha = 1.2$, $\gamma = 1.3$, $u = 0.294$.

can ensure that robots switch when triggered locally.

Task distributions Many multi-robot applications need subteams to be assigned to different tasks in a certain proportion (see e.g. [110] for a review on multi-robot task allocation). As illustrated in Fig. 12.2, switching transformations can be used to distribute a team of agents among two tasks in predetermined proportion.

In scenarios where robots can be divided in subteams, and each subteam can indiscriminately contribute to one of the tasks, our method guarantees a predetermined proportion of agents among tasks, but does not control which subteam is assigned to which task. For example, in Fig. 12.2, the team composed of agents 1, 2, and 3 could execute either task 1 or 2. This affords flexibility from the application point of view, where the initial condition on the opinion (e.g., how close a robot is to an area where the demand for a task is abundant) can determine the final distribution of agents.

Local flexibility Our approach provides flexibility by letting single agents/robots make individual decisions in a decentralized way without disconnecting from the network. This is important for multi-robot teams that should change their allocation across tasks in response to changing environmental conditions that are observed only by some agents; see [159] for the case of globally available environmental cues in a multi-robot trash pick-up problem. This method is also useful for long-duration autonomy applications where team performance should not be affected by failures or individuals that stop contributing to tasks, e.g. to charge batteries [148]. See Fig. 12.4, where one agent switches between tasks without affecting the task preference of its neighbors in the graph.

To illustrate suppose that agent i wants to switch options. Agent i alerts its neighbors that $\theta_i = -1$. We define

$$\tau_a \dot{a}_{ik} = -a_{ik} + a_{ik}(0)\theta_i\theta_k, \quad (12.6)$$

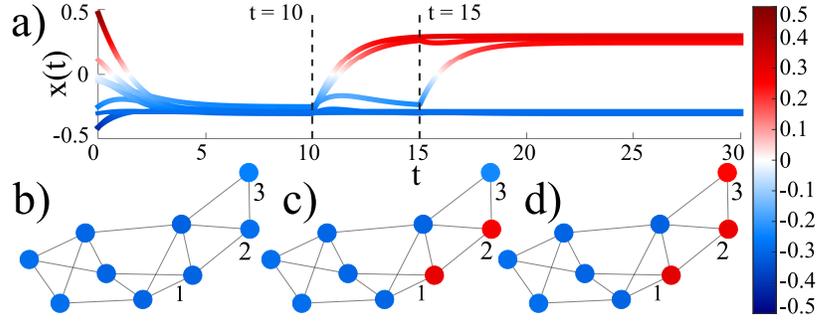


Figure 12.5: Locally switching agents 1, 2 and 3 using (12.6). Agents 1, 2 switch at $t = 10$ and agent 3 at $t = 15$. (b)-(d) Network diagram with the opinion of each agent at $t = 10$, $t = 15$, $t = 30$. Parameters: $S = \tanh$, $d = 1$, $\alpha = 1.2$, $\gamma = 1.3$, $u = 0.315$, $\tau_a = 0.01$

where τ_a is a time-scale parameter and $a_{ik}(0) \in \{0, 1, -1\}$ is the initial signature of the edge between agents i and k . As seen in Fig. 12.5, the dynamics (12.6) allow locally originated switches to take place simultaneously between neighbors. This property can be useful in applications involving switching cascades, where a robot switching to a new task can trigger its neighbors to switch. This feature is relevant in dynamic task allocation for multi-robot systems where robots can assign themselves to new tasks as a result of interacting with the environment or with their neighboring robots [122].

12.7 Final Remarks

We analyzed the nonlinear networked opinion dynamics (12.1) on signed graphs and proposed a novel approach for dynamic and decentralized allocation of a group of agents across two tasks. In future work, we aim to generalize the results in Section 12.5 to graphs that are not structurally balanced, to derive an estimate for the ε bound from Assumption 12.5.2, and to extend this analysis to the more general multi-option opinion dynamics of [19].

Appendix A

Supporting material for Chapter 4

A.1 Proof of Proposition 4.4.1

Recall that we are examining the dynamical system

$$\dot{x}_i = -dx_i + u\left(\hat{S}_1(\alpha x_i + cx_j) - \hat{S}_2(\beta x_i + cx_j)\right) =: f_i(x_1, x_2) \quad i, j \in \{1, 2\}, \quad i \neq j.$$

with $d > 0$, $\alpha, \beta \geq 0$, and $c \in \mathbb{R}$. Furthermore, $\hat{S}_i(x) = \frac{1}{a_i} \tanh(x)$. Local exponential stability of $\mathbf{x} = \mathbf{0}$ follows from linearization. The Jacobian of (4.9) is diagonal with a repeated eigenvalue equal to $-d + u(\alpha - \beta)$. This quantity is negative for $u < ud/(\alpha - \beta)$ and positive for $u > d/(\alpha - \beta)$.

To prove the rest of the statement of the Proposition we rely on results from singularity theory outlined in [82, Chapter X]. In particular, we verify that the governing equations for the opinion dynamics (4.9) satisfy the conditions for local topological equivalence to a $(Z_2 \times Z_2)$ -equivariant nondegenerate normal form, the dynamical behavior of which is well understood.

First, we check that the equations commute with the symmetry group $S_2 \times S_2 \cong Z_2 \oplus Z_2$. Define the action $(\tau, \sigma) \in S_2 \times S_2$ where σ is a permutation of the agents and τ is a permutation of options for both agents ($x_i \rightarrow -x_i \forall i$). Also define $e \in S_2$ as the identity element of the permutation group on two symbols. The group elements of $S_2 \times S_2$ are $\{(\tau, \sigma), (\tau, e), (e, \sigma), (e, e)\}$. If the dynamical system is $(S_2 \times S_2)$ -equivariant, it should commute with the action of each nontrivial group element, which means we must check the following:

- The dynamics are equivariant under permutation of agents:

$$(e, \sigma) \cdot (f_1(x_1, x_2), f_2(x_1, x_2)) = (f_1((e, \sigma) \cdot (x_1, x_2)), f_2((e, \sigma) \cdot (x_1, x_2))) \quad (\text{A.1})$$

- The dynamics are equivariant under permutation of options:

$$(\tau, e) \cdot (f_1(\theta_1, \theta_2), f_2(\theta_1, \theta_2)) = (f_1((\tau, e) \cdot (\theta_1, x_2)), f_2((\tau, e) \cdot (x_1, x_2))) \quad (\text{A.2})$$

- The dynamics are equivariant under permutation of both agents and options:

$$(\tau, \sigma) \cdot (f_1(x_1, x_2), f_2(x_1, x_2)) = (f_1((\tau, \sigma) \cdot (x_1, x_2)), f_2((\tau, \sigma) \cdot (x_1, x_2))) \quad (\text{A.3})$$

As a recap, $(e, \sigma) \cdot (x, y) = (y, x)$, $(\tau, e) \cdot (x, y) = (-x, -y)$, and $(\tau, \sigma) \cdot (x, y) = (-y, -x)$. First we consider the action (e, σ) :

$$(e, \sigma) \cdot \begin{pmatrix} f_1(x_1, x_2) \\ f_2(x_1, x_2) \end{pmatrix} = \begin{pmatrix} f_2(x_1, x_2) \\ f_1(x_1, x_2) \end{pmatrix} = \begin{pmatrix} -dx_2 + u\hat{S}_1(\alpha x_2 + cx_1) - u\hat{S}_2(\beta x_2 + cx_1) \\ -dx_1 + \hat{S}_1(\alpha x_2 + cx_2) - u\hat{S}_2(\beta x_1 + cx_2) \end{pmatrix} \quad (\text{A.4})$$

$$\begin{pmatrix} f_1((e, \sigma) \cdot (x_1, x_2)) \\ f_2((e, \sigma) \cdot (x_1, x_2)) \end{pmatrix} = \begin{pmatrix} f_1(x_2, x_1) \\ f_2(x_2, x_1) \end{pmatrix} = \begin{pmatrix} -dx_2 + u\hat{S}_1(\alpha x_2 + cx_1) - u\hat{S}_2(\beta x_2 + cx_1) \\ -dx_1 + u\hat{S}_1(\alpha x_1 + cx_2) - u\hat{S}_2(\beta x_1 + cx_2) \end{pmatrix} \quad (\text{A.5})$$

The two statements above are equivalent and the dynamics are equivariant under the permutation of agents. Then we consider the action (τ, e) :

$$(\tau, e) \cdot \begin{pmatrix} f_1(x_1, x_2) \\ f_2(x_1, x_2) \end{pmatrix} = \begin{pmatrix} -f_1(x_1, x_2) \\ -f_2(x_1, x_2) \end{pmatrix} = \begin{pmatrix} dx_1 - u\hat{S}_1(\alpha x_1 + cx_2) + u\hat{S}_2(\beta x_1 + cx_2) \\ dx_2 - u\hat{S}_1(\alpha x_2 + cx_1) + u\hat{S}_2(\beta x_2 + cx_1) \end{pmatrix} \quad (\text{A.6})$$

$$\begin{pmatrix} f_1((\tau, e) \cdot (x_1, x_2)) \\ f_2((\tau, e) \cdot (x_1, x_2)) \end{pmatrix} = \begin{pmatrix} f_1(-x_1, -x_2) \\ f_2(-x_1, -x_2) \end{pmatrix} = \begin{pmatrix} dx_1 + u\hat{S}_1(-\alpha x_1 - cx_2) - u\hat{S}_2(-\beta x_1 - cx_2) \\ dx_2 + u\hat{S}_1(-\alpha x_2 - cx_1) - u\hat{S}_2(-\beta x_2 - cx_1) \end{pmatrix} \quad (\text{A.7})$$

Since the sigmoid functions \hat{S}_i are odd, we can conclude that the two statements above are equivalent and the dynamics are equivariant under permutation of options. Finally, we consider the action

(τ, σ) :

$$(\tau, \sigma) \cdot \begin{pmatrix} f_1(x_1, x_2) \\ f_2(x_1, x_2) \end{pmatrix} = \begin{pmatrix} -f_2(x_1, x_2) \\ -f_1(x_1, x_2) \end{pmatrix} = \begin{pmatrix} dx_2 - u\hat{S}_1(\alpha x_2 + cx_1) + u\hat{S}_2(\beta x_2 + cx_1) \\ dx_1 - u\hat{S}_1(\alpha x_1 + cx_2) + u\hat{S}_2(\beta x_1 + cx_2) \end{pmatrix} \quad (\text{A.8})$$

$$\begin{pmatrix} f_1((\tau, \sigma) \cdot (x_1, x_2)) \\ f_2((\tau, \sigma) \cdot (x_1, x_2)) \end{pmatrix} = \begin{pmatrix} f_1(-x_2, -x_1) \\ f_2(-x_2, -x_1) \end{pmatrix} = \begin{pmatrix} dx_2 + u\hat{S}_1(-\alpha x_2 - cx_1) - u\hat{S}_2(-\beta x_2 - cx_1) \\ dx_1 + u\hat{S}_1(-\alpha x_1 - cx_2) - u\hat{S}_2(-\beta x_1 - cx_2) \end{pmatrix} \quad (\text{A.9})$$

Once again, the two statements above are equivalent and so the dynamics are equivariant under simultaneous permutation of agents and options.

We have established that the dynamical system we are working with has the same symmetry as the normal form. Next, for convenience of calculations, we transform the coordinate system to align the two axes with the primary modes of the bifurcation. The coordinates on the consensus and dissensus subspaces are $x_c = \frac{1}{2}(x_1 + x_2)$ (average opinion of the agents) and $x_d = \frac{1}{2}(x_1 - x_2)$ (average difference in the opinions of the agents). Transformed into these consensus and dissensus coordinates, the model equations become

$$\begin{aligned} \dot{x}_c = -dx_a + \frac{1}{2}u\hat{S}_1((\alpha+c)x_c + (\alpha-c)x_d) + \frac{1}{2}u\hat{S}_1((\alpha+c)x_c + (c-\alpha)x_d) - \frac{1}{2}u\hat{S}_2((\beta+c)x_c + (\beta-c)x_d) \\ - \frac{1}{2}u\hat{S}_2((\beta+c)x_c + (\beta-c)x_d), \quad (\text{A.10}) \end{aligned}$$

$$\begin{aligned} \dot{x}_d = -dx_d + \frac{1}{2}u\hat{S}_1((\alpha+c)x_c + (\alpha-c)x_d) - \frac{1}{2}u\hat{S}_1((\alpha+c)x_c + (c-\alpha)x_d) - \frac{1}{2}u\hat{S}_2((\beta+c)x_c + (c-\beta)x_d) \\ + \frac{1}{2}u\hat{S}_2((\beta+c)x_c + (c-\beta)x_d). \quad (\text{A.11}) \end{aligned}$$

Next, we compute a series expansion of (A.10),(A.11)

$$\begin{aligned} \dot{x}_c &= A_1 r x_c + A_2 x_c^3 + A_3 x_c x_d^2 + h.o.t. \\ \dot{x}_d &= B_1 r x_d + B_2 x_d^3 + B_3 x_d x_c^2 + h.o.t. \end{aligned} \quad (\text{A.12})$$

where the higher order terms are of the form $x_c^u x_d^s u^t$ and satisfy one of (a) $u + s \geq 5$, (b) $t =$

1, $u + s \geq 3$, or (c) $t \geq 2$. The coefficients in the expansion are

$$\begin{aligned}
 A_1 &= \alpha - \beta, & B_1 &= \alpha - \beta, \\
 A_2 &= \frac{d}{3(\alpha - \beta)} (-a_1^2(c + \alpha)^3 + a_2^2(c + \beta)^3) \\
 A_3 &= \frac{d}{\alpha - \beta} (-a_1^2(c - \alpha)^2(c + \alpha) + a_2^2(c - \beta)^2(c + \beta)), \\
 B_2 &= \frac{d}{3(\alpha - \beta)} (a_1^2(c - \alpha)^3 - a_2^2(c - \beta)^3) \\
 B_3 &= \frac{d}{\alpha - \beta} (a_1^2(c - \alpha)(c + \alpha)^2 - a_2^2(c - \beta)(c + \beta)^2).
 \end{aligned}$$

From these coefficients the nondegeneracy conditions stated in the proposition are computed using [82, Chapter X, Definition 2.2]. Equivalence to the normal form then follows by [82, Chapter X, Proposition 2.3], assuming the nondegeneracy conditions are satisfied.

Appendix B

Supporting material for Chapter 5

For proofs of Theorems 5.2.2 and 5.3.1 we rely on Lyapunov-Schmidt reduction and singularity theory techniques, for a thorough development of which we refer the interested reader to [82].

B.1 Directional derivatives

In this section we compute the directional derivatives of (5.1) at $(\mathbf{Z}, u) = (\mathbf{0}, u^*)$, for use in the Lyapunov-Schmidt reduction calculations in the following two sections. We compute the directional derivatives which appear in the Lyapunov-Schmidt reduction coefficients (2.15). Recall the directional derivative formula (2.13). Let $\hat{\mathbf{v}} = \mathbf{v} \otimes \tilde{\mathbf{v}}$, with $\mathbf{v} \in \mathbb{R}^{N_a}$ and $\tilde{\mathbf{v}} \in \mathbb{R}^{N_o}$. For

First, we compute the first order directional derivative for $\frac{\partial \mathbf{F}^h}{\partial u}(\mathbf{Z})$:

$$d\left(\frac{\partial \mathbf{F}^h}{\partial u}\right)(\hat{\mathbf{v}}) = S'_1(0)(\alpha \mathcal{I}_{N_a} + \gamma A_a)\mathbf{v} \otimes \tilde{\mathbf{v}} + S'_2(0)(\beta \mathcal{I}_{N_a} + \delta A_a)\mathbf{v} \otimes A_o \tilde{\mathbf{v}}.$$

When \mathbf{v} is an eigenvector of A_a with eigenvalue λ , and $\tilde{\mathbf{v}}$ is an eigenvector of A_o with eigenvalue μ , this derivative further simplifies to

$$d\left(\frac{\partial \mathbf{F}^h}{\partial u}\right)(\hat{\mathbf{v}}) = (S'_1(0)(\alpha + \gamma\lambda) + S'_2(0)\mu(\beta + \delta\lambda))\hat{\mathbf{v}}. \quad (\text{B.1})$$

Let $\mathbf{x}^2 = \mathbf{x} \odot \mathbf{x}$. With a similar calculation at second order, we find

$$d^2 \mathbf{F}^h(\hat{\mathbf{v}}, \hat{\mathbf{v}}) = u^* S''_1(0)((\alpha \mathcal{I}_{N_a} + \gamma A_a)\mathbf{v})^2 \otimes \tilde{\mathbf{v}}^2 + u^* S''_2(0)((\beta \mathcal{I}_{N_a} + \delta A_a)\mathbf{v})^2 \otimes (A_o \tilde{\mathbf{v}})^2$$

which, under the eigenvector assumption simplifies to

$$d^2\mathbf{F}^h(\hat{\mathbf{v}}, \hat{\mathbf{v}}) = u^* (S_1''(0)(\alpha + \gamma\lambda)^2 + S_2''(0)\mu^2(\beta + \delta\lambda)^2) \hat{\mathbf{v}}^2. \quad (\text{B.2})$$

Next, defining $\mathbf{x}^3 = \mathbf{x} \odot \mathbf{x} \odot \mathbf{x}$ at third order we find

$$d^3\mathbf{F}^h(\hat{\mathbf{v}}, \hat{\mathbf{v}}, \hat{\mathbf{v}}) = u^* S_1'''(0)((\alpha\mathcal{I}_{N_a} + \gamma A_a)\mathbf{v})^3 \otimes \hat{\mathbf{v}}^3 + u^* S_2'''(0)((\beta\mathcal{I}_{N_a} + \delta A_a)\mathbf{v})^3 \otimes (A_o\hat{\mathbf{v}})^3$$

which simplifies to

$$d^3\mathbf{F}^h(\hat{\mathbf{v}}, \hat{\mathbf{v}}) = u^* (S_1'''(0)(\alpha + \gamma\lambda)^3 + S_2'''(0)\mu^3(\beta + \delta\lambda)^3) \hat{\mathbf{v}}^3. \quad (\text{B.3})$$

Define $P_N = \mathcal{I}_N - \hat{\mathbf{v}}\hat{\mathbf{v}}^T$ to be the projection from \mathbb{R}^N onto the range of J at $u = u^*$. Suppose $S_m(y) = \hat{S}(y + g_m(y))$ with $\hat{S}(-y) = -\hat{S}(y)$, $\hat{S}'_m(0) > 0$, and bounded small perturbation g satisfying $g(0) = 0$, $g'(0) = 0$, $g(-y) \neq -g(y)$. Then $S''_m(0) = g''_m(0)$ and

$$J^{-1}P_N d^2\mathbf{F}^h(\hat{\mathbf{v}}, \hat{\mathbf{v}}) = u^* (g_1''(0)(\alpha + \gamma\lambda)^2 + g_2''(0)\mu^2(\beta + \delta\lambda)^2) J^{-1}P_N \hat{\mathbf{v}}^2$$

where J^{-1} is the inverse of the restriction of (11.4) to its range. Finally we write down the expression for the following directional derivative

$$\begin{aligned} d^2\mathbf{F}^h(\hat{\mathbf{v}}, J^{-1}P_N d^2\mathbf{F}^h(\hat{\mathbf{v}}, \hat{\mathbf{v}})) &= (u^*)^2 (g_1''(0)(\alpha + \gamma\lambda)^2 + g_2''(0)\mu^2(\beta + \delta\lambda)^2) \\ &\quad (g_1''(0)(\alpha + \gamma\lambda)\hat{\mathbf{v}} \odot ((\alpha\mathcal{I}_{N_a} + \gamma A_a) \otimes \mathcal{I}_{N_o})(J^{-1}P\hat{\mathbf{v}}^2) \\ &\quad + g_2''(0)\mu(\beta + \delta\lambda)\hat{\mathbf{v}} \odot ((\beta\mathcal{I}_{N_a} + \delta A_a) \otimes A_o)(J^{-1}P\hat{\mathbf{v}}^2)). \quad (\text{B.4}) \end{aligned}$$

B.2 Proof of Theorem 5.2.2

1) The loss of stability of $\mathbf{Z} = \mathbf{0}$ at $u = u^*$ is established in Theorem 5.1.2. Since η_{max} is simple, at $u = u^*$ $\eta_{max} = 0$ equilibria along $\text{span}\{\mathbf{v}_a \otimes \mathbf{v}_o\}$. Following the Lyapunov-Schmidt reduction procedure outlined in [82, Chapter I] we derive the coefficients of the polynomial expansion of a Lyapunov-Schmidt reduction of (2.14) $f(y, u, \mathbf{b})$ through third order. With the odd symmetry

assumption on S_m , it follows that $S_m''(0) = 0$ and

$$f(y, u, \mathbf{0}) = (\alpha + \gamma\lambda_a + \beta\mu_o + \delta(\lambda_a\mu_o - \lambda_c\mu_c))(u - u^*)y + \frac{d(S_1'''(0)(\alpha + \gamma\lambda_a)^3 + S_2'''(0)(\beta\mu_o + \delta(\lambda_a\mu_o - \lambda_c\mu_c))^3)}{\alpha + \gamma\lambda_a + \beta\mu_o + \delta(\lambda_a\mu_o - \lambda_c\mu_c)} \langle \mathbf{w}_a, \mathbf{v}_a^3 \rangle \langle \mathbf{w}_o, \mathbf{v}_o^3 \rangle y^3. \quad (\text{B.5})$$

The conclusion follows by the recognition problem for a pitchfork bifurcation [82, Chapter II, Proposition 9.2]. Linearization of the system about each of the two bifurcating fixed points shares the $N - 1$ stable eigenvalues with the origin, and the bifurcating eigenvalue is negative (or positive) under the conditions in the theorem statement by [82, Chapter I, Theorem 4.1].

2) The first statement follows by definition of an unfolding. With nonzero \mathbf{b}, g_1, g_2 , the coefficients of the polynomial expansion of $f(y, u, \mathbf{b})$ through third order take the form

$$f(y, u, \mathbf{b}) = f(y, u, \mathbf{0}) + \langle \mathbf{w}_a \otimes \mathbf{w}_o, \mathbf{b} \rangle + C_1 y^3 + \frac{d(g_1''(0)(\alpha + \gamma\lambda_a)^2 + g_2''(0)(\beta\mu_o + \delta(\lambda_a\mu_o - \lambda_c\mu_c))^2)}{\alpha + \gamma\lambda_a + \beta\mu_o + \delta(\lambda_a\mu_o - \lambda_c\mu_c)} \langle \mathbf{w}_a, \mathbf{v}_a^2 \rangle \langle \mathbf{w}_o, \mathbf{v}_o^2 \rangle y^2 \quad (\text{B.6})$$

where C_1 is the quantity defined in (B.4). Let C_0 denote the coefficient of the cubic term in (B.5). When $g_1''(0) = 0$ and $g_2''(0) = 0$, the total coefficient of the cubic term in (B.6) is C_0 . Therefore by continuity there exist small positive $\varepsilon_1, \varepsilon_2$ such that $\text{sign}(C_0 + C_1) = \text{sign}(C_0)$ as long as $|g_1''(0)| < \varepsilon_1$ and $|g_2''(0)| < \varepsilon_2$. We assume $g_m''(0)$ for $m = 1, 2$ is sufficiently small so that C_0 determines the sign of the cubic term, and do not compute (B.4).

For compactness let $f(y, u, \mathbf{b}) = a_1(u - u^*)x + a_2x^2 + a_3x^3 + a_4 + h.o.t.$ where a_1, a_2, a_3, a_4 are the coefficients in (B.5),(B.6). Following [82, Chapter III, Proposition 4.4]. We confirm that the following determinant is nonzero:

$$\det \begin{pmatrix} 0 & 0 & a_1 & a_3 \\ 0 & a_1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = -a_1^2 = -(\alpha + \gamma\lambda_a + \beta\mu_o + \delta(\lambda_a\mu_o - \lambda_c\mu_c))^2 \neq 0.$$

Then if we treat the quantities a_2, a_4 as unfolding parameters, (B.6) realizes a universal unfolding of the pitchfork bifurcation by [82, Chapter III, Proposition 4.4] and the theorem statement describes the expected persistent bifurcation diagrams, as in [82, Chapter 1].

B.3 Proof of Theorem 5.3.1

1) To establish existence of periodic orbits we check that the system (5.1) under the stated assumptions satisfies the conditions of the Hopf bifurcation theorem [90, Theorem 3.4.2]. When $u = u^*$, the leading eigenvalues of (11.4) are a simple purely imaginary pair

$$\eta_{\pm}(u^*) = \pm i \frac{d(\gamma\lambda_c - \beta\mu_c + \delta(-\lambda_a\mu_c + \lambda_c\mu_o))}{\alpha + \gamma\lambda_a + \beta\mu_o + \delta(\lambda_a\mu_o + \lambda_c\mu_c)}$$

which satisfies the eigenvalue condition (H1) of the Hopf theorem. Next, we check that the leading eigenvalues cross the imaginary axis with nonzero speed as u is varied, i.e.

$$\frac{d}{du} \operatorname{Re}(\eta_{\pm}(u)) = \alpha + \gamma\lambda_a + \beta\mu_o + \delta(\lambda_a\mu_o + \lambda_c\mu_c) \neq 0$$

which satisfies the nonzero crossing speed condition (H2) of the Hopf theorem. Existence of periodic orbits directly follows by the Hopf theorem. By this theorem and by the definition of a center manifold [90, Theorem 3.2.1], the solutions appear along a unique W^s which is tangent at $u = u^*$ to $\mathcal{N}(J(\mathbf{0}, u^*)) = \operatorname{span}\{\operatorname{Re}(\mathbf{v}_a \otimes \mathbf{v}_o), \operatorname{Im}(\mathbf{v}_a \otimes \mathbf{v}_o)\}$.

To show 2) and 3) we first compute the coefficients of the Lyapunov-Schmidt reduction of (5.1), following the Lyapunov-Schmidt reduction for a Hopf bifurcation [82, Chapter VIII, Proposition 3.3], also outlined in Chapter 2.3 of this dissertation.

$$f(y, u) = \left(\alpha + \gamma\lambda_a + \beta\mu_o + \delta(\lambda_a\mu_o + \lambda_c\mu_c) \right) y(u - u^*) + by^3 + h.o.t. \quad (\text{B.7})$$

where

$$b = \frac{1}{16} u^* \operatorname{Re} \left(\left(S_1'''(0) ((\alpha + \gamma\lambda_a)^2 + (\gamma\lambda_c)^2) (\alpha + \gamma(\lambda_a + i\lambda_c)) \right. \right. \\ \left. \left. + S_2'''(0) (\mu_o^2 + \mu_c^2) ((\beta + \delta\lambda_a)^2 + (\delta\lambda_c)^2) (\mu_o(\beta + \delta\lambda_a) - \delta\mu_c\lambda_c \right. \right. \\ \left. \left. + i(\mu_c(\beta + \delta\lambda_a) + \delta\lambda_c\mu_o) \right) \langle \mathbf{w}_a, |\mathbf{v}_a|^2 \odot \mathbf{v}_a \rangle \langle \mathbf{w}_o, |\mathbf{v}_o|^2 \odot \mathbf{v}_o \rangle \right) + C_2. \quad (\text{B.8})$$

where all of the terms grouped in C_2 have a dependence on $g_1''(0)$ and/or $g_2''(0)$, with $C_2 = 0$ whenever $g_1''(0) = 0$ and $g_2''(0) = 0$. By an analogous continuity argument to that presented in the pitchfork bifurcation proof in Appendix B.2 we assume $g_m''(0)$ is sufficiently small so the first term in the sum in (B.8) dominates and determines the sign of the cubic coefficient b . Then as long as $b \neq 0$, by

[82, Chapter VIII, Theorems 2.1 and 3.2] the reduced bifurcation equation $f(y, u)$ is Z_2 -equivalent to the pitchfork $\text{sign}(b)x^3 + x(u - u^*)$, i.e. when $|u - u^*|$ is small, solutions to $f(y, u) = 0$ are in one-to-one correspondence with orbits of small amplitude periodic solutions to the system (5.1) with period near $2\pi/|\gamma\lambda_c + \beta\mu_c + \delta(\lambda_a\mu_c + \lambda_c\mu_o)| =: 1/\omega$. For u near u^* , the small amplitude oscillations can be approximated to first order as scalar multiples of $e^{i\omega t}\mathbf{v}_a \otimes \mathbf{v}_o$ from which the conclusions on phase and amplitude difference between agents follow. When $b < 0$, the bifurcating periodic solutions are stable by [82, Chapter VIII, Theorem 4.1].

Bibliography

- [1] P. U. Abara, F. Ticozzi, and C. Altafini. Spectral conditions for stability and stabilization of positive equilibria for a class of nonlinear cooperative systems. *IEEE Transactions on Automatic Control*, 63(2):402–417, 2018.
- [2] M. A. Aguiar and A. P. S. Dias. Synchronization and equitable partitions in weighted networks. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 28(7):073105, 2018.
- [3] S. Alava, D. Frau-Meigs, and G. Hassan. *Youth and violent extremism on social media: mapping the research*. UNESCO Publishing, 2017.
- [4] L. Albantakis and G. Deco. The encoding of alternatives in multiple-choice decision making. *Proceedings of the National Academy of Sciences*, 106(25):10308–10313, 2009.
- [5] C. Altafini. Consensus problems on networks with antagonistic interactions. *IEEE Trans. Aut. Ctrl.*, 58(4):935–946, 2013.
- [6] S. Amador, S. Okamoto, and R. Zivan. Dynamic multi-agent task allocation with spatial and temporal constraints. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 28, 2014.
- [7] P. W. Anderson. *The economy as an evolving complex system*. CRC Press, 2018.
- [8] F. Antoneli and I. Stewart. Symmetry and synchrony in coupled cell networks 2: Group networks. *International Journal of Bifurcation and Chaos*, 17(03):935–951, 2007.
- [9] V. I. Arnold. *Singularity theory*, volume 53. Cambridge University Press, 1981.
- [10] R. Axelrod. *The Evolution of Cooperation*. Basic Books, 1984.
- [11] R. Axelrod, J. J. Daymude, and S. Forrest. Preventing extreme polarization of political attitudes. *Proceedings of the National Academy of Sciences*, 118(50):e2102139118, 2021.

- [12] A. Baronchelli. The emergence of consensus: a primer. *Royal Society open science*, 5(2):172189, 2018.
- [13] T. Başar and G. J. Olsder. *Dynamic noncooperative game theory*. SIAM, 1998.
- [14] F. Baumann, P. Lorenz-Spreen, I. M. Sokolov, and M. Starnini. Modeling echo chambers and polarization dynamics in social networks. *Physical Review Letters*, 124(4):048301, 2020.
- [15] F. Baumann, P. Lorenz-Spreen, I. M. Sokolov, and M. Starnini. Emergence of polarized ideological opinions in multidimensional topic spaces. *Physical Review X*, 11(1):011012, 2021.
- [16] L. Bayındır. A review of swarm robotics tasks. *Neurocomputing*, 172:292–321, 2016.
- [17] W. Bialek, A. Cavagna, I. Giardina, T. Mora, E. Silvestri, M. Viale, and A. M. Walczak. Statistical mechanics for natural flocks of birds. *Proceedings of the National Academy of Sciences*, 109(13):4786–4791, 2012.
- [18] A. Bizyaeva, G. Amorim, M. Santos, A. Franci, and N. E. Leonard. Switching transformations for decentralized control of opinion patterns in signed networks: application to dynamic task allocation. *IEEE Control Systems Letters*, 2022.
- [19] A. Bizyaeva, A. Franci, and N. E. Leonard. Nonlinear opinion dynamics with tunable sensitivity. *IEEE Transactions on Automatic Control*, 2022.
- [20] A. Bizyaeva, A. Matthews, A. Franci, and N. E. Leonard. Patterns of nonlinear opinion formation on networks. In *American Control Conference*, pages 2739–2744, 2021.
- [21] A. Bizyaeva, T. Sorochnik, A. Franci, and N. E. Leonard. Control of agreement and disagreement cascades with distributed inputs. In *IEEE Conference on Decision and Control*, pages 4994–4999, 2021.
- [22] V. Blondel, J. M. Hendrickx, and J. N. Tsitsiklis. On Krause’s multi-agent consensus model with state-dependent connectivity. *IEEE Transactions on Automatic Control*, 54(11):2586–2597, 2009.
- [23] R. Bogacz and K. Gurney. The basal ganglia and cortex implement optimal decision making between alternative actions. *Neural Computation*, 19(2):442–477, 2007.
- [24] R. Bogacz, M. Usher, J. Zhang, and J. L. McClelland. Extending a biologically inspired model of choice: multi-alternatives, nonlinearity and value-based multidimensional choice.

- Philosophical Transactions of the Royal Society B: Biological Sciences*, 362(1485):1655–1670, 2007.
- [25] E. Bonabeau. Swarm intelligence. *A Primer on Multiple Intelligences*, page 211, 2003.
- [26] E. Bonabeau, G. Theraulaz, M. Dorigo, G. Theraulaz, D. d. R. D. F. Marco, et al. *Swarm intelligence: from natural to artificial systems*. Number 1. Oxford university press, 1999.
- [27] P. Bonacich. Factoring and weighting approaches to status scores and clique identification. *Journal of Mathematical Sociology*, 2(1):113–120, 1972.
- [28] P. Bonacich. Some unique properties of eigenvector centrality. *Social networks*, 29(4):555–564, 2007.
- [29] M. Brambilla, E. Ferrante, M. Birattari, and M. Dorigo. Swarm robotics: a review from the swarm engineering perspective. *Swarm Intelligence*, 7(1):1–41, 2013.
- [30] M. Caponigro, A. C. Lai, and B. Piccoli. A nonlinear model of opinion formation on the sphere. *Discrete & Continuous Dynamical Systems - A*, 35(9):4241–4268, Sept. 2015.
- [31] S. B. S. D. d. Castro. Mode interactions with symmetry. *Dynamics and Stability of Systems*, 10(1):13–31, 1995.
- [32] C. Cathcart, S. Park, A. Bizyaeva, and N. E. Leonard. Robot navigation around oncoming movers using bio-inspired opinion dynamics. In *International Conference on Intelligent Robots and Systems, workshop on robotics-inspired biology.*, 2020.
- [33] C. Cathcart, M. Santos, S. Park, and N. E. Leonard. Nonlinear opinion dynamics for human-robot navigation. *In preparation*.
- [34] D. Centola, J. Becker, D. Brackbill, and A. Baronchelli. Experimental evidence for tipping points in social convention. *Science*, 360(6393):1116–1119, 2018.
- [35] G. Chen, J. L. Moiola, and H. O. Wang. Bifurcation control: theories, methods, and applications. *International Journal of Bifurcation and Chaos*, 10(03):511–548, 2000.
- [36] G. Cicogna and G. Gaeta. Symmetry invariance and centre manifolds for dynamical systems. *Il Nuovo Cimento B (1971-1996)*, 109(1):59–76, 1994.
- [37] M. Cinelli, G. De Francisci Morales, A. Galeazzi, W. Quattrociocchi, and M. Starnini. The echo chamber effect on social media. *Proceedings of the National Academy of Sciences*, 118(9):e2023301118, 2021.

- [38] P. Cisneros-Velarde, K. S. Chan, and F. Bullo. Polarization and fluctuations in signed social networks. *IEEE Transactions on Automatic Control*, 66:3789–3793, 2021.
- [39] I. D. Couzin, C. C. Ioannou, G. Demirel, T. Gross, C. J. Torney, A. Hartnett, L. Conradt, S. A. Levin, and N. E. Leonard. Uninformed individuals promote democratic consensus in animal groups. *Science*, 334(6062):1578–1580, 2011.
- [40] I. D. Couzin, J. Krause, N. R. Franks, and S. A. Levin. Effective leadership and decision-making in animal groups on the move. *Nature*, 433(7025):513–516, 2005.
- [41] D. Cvetkovic, S. Simic, and P. Rowlinson. *An Introduction to the Theory of Graph Spectra*. Cambridge University Press, 2009.
- [42] R. Dale, R. Fusaroli, N. D. Duran, and D. C. Richardson. The self-organization of human interaction. In *Psychology of learning and motivation*, volume 59, pages 43–95. Elsevier, 2013.
- [43] P. Dandekar, A. Goel, and D. T. Lee. Biased assimilation, homophily, and the dynamics of polarization. *Proceedings of the National Academy of Sciences*, 110(15):5791–5796, 2013.
- [44] G. Deffuant, D. Neau, F. Amblard, and G. Weisbuch. Mixing beliefs among interacting agents. *Advances in Complex Systems*, 03(01n04):87–98, 2000.
- [45] M. H. DeGroot. Reaching a consensus. *Journal of the American Statistical Association*, 69(345):118–121, 1974.
- [46] J.-L. Deneubourg, S. Aron, S. Goss, and J. M. Pasteels. The self-organizing exploratory pattern of the argentine ant. *Journal of insect behavior*, 3(2):159–168, 1990.
- [47] A. Dhooge, W. Govaerts, Y. Kuznetsov, H. Meijer, and B. Sautois. New features of the software MatCont for bifurcation analysis of dynamical systems. *Math Comp Mod Dyn Syst*, 14(2):147–175, 2008.
- [48] M. B. Dias, R. Zlot, N. Kalra, and A. Stentz. Market-based multirobot coordination: A survey and analysis. *Proceedings of the IEEE*, 94(7):1257–1270, 2006.
- [49] L. Ding, W. X. Zheng, and G. Guo. Network-based practical set consensus of multi-agent systems subject to input saturation. *Automatica*, 89:316–324, 2018.
- [50] M. Dorigo, G. Theraulaz, and V. Trianni. Reflections on the future of swarm robotics. *Science Robotics*, 5(49):eabe4385, 2020.

- [51] M. Dorigo, G. Theraulaz, and V. Trianni. Swarm robotics: past, present, and future [point of view]. *Proceedings of the IEEE*, 109(7):1152–1165, 2021.
- [52] F. Ducatelle, G. A. Di Caro, A. Förster, M. Bonani, M. Dorigo, S. Magnenat, F. Mondada, R. O’Grady, C. Pinciroli, P. Rétornaz, et al. Cooperative navigation in robotic swarms. *Swarm Intelligence*, 8(1):1–33, 2014.
- [53] R. Elhabyan, W. Shi, and M. St-Hilaire. Coverage protocols for wireless sensor networks: Review and future directions. *Journal of Communications and Networks*, 21(1):45–60, 2019.
- [54] M. H. et al. python-ternary: Ternary plots in python. *Zenodo 10.5281/zenodo.594435*.
- [55] X. Fan, W. Sayers, S. Zhang, Z. Han, L. Ren, and H. Chizari. Review and classification of bio-inspired algorithms and their applications. *Journal of Bionic Engineering*, 17(3):611–631, 2020.
- [56] J. J. Faria, S. Krause, and J. Krause. Collective behavior in road crossing pedestrians: the role of social information. *Behavioral ecology*, 21(6):1236–1242, 2010.
- [57] L. Farina and S. Rinaldi. *Positive linear systems: theory and applications*, volume 50. John Wiley & Sons, 2011.
- [58] E. Fehr and S. Gächter. Reciprocity and economics: The economic implications of homo reciprocans. *European Economic Review*, 42(3-5):845–859, 1998.
- [59] J. Fernández-Gracia, K. Suchecki, J. J. Ramasco, M. San Miguel, and V. M. Eguíluz. Is the voter model a model for voters? *Physical review letters*, 112(15):158701, 2014.
- [60] E. L. Fink, S. A. Kaplowitz, and S. M. Hubbard. Oscillation in beliefs and decisions. *The Persuasion Handbook: Developments in Theory and Practice. Thousand Oaks, CA: Sage Publications*, pages 17–38, 2002.
- [61] A. Fontan and C. Altafini. Multiequilibria analysis for a class of collective decision-making networked systems. *IEEE Transactions on Control of Network Systems*, 5(4):1931–1940, 2017.
- [62] A. Fontan and C. Altafini. Achieving a decision in antagonistic multi agent networks: frustration determines commitment strength. In *IEEE Conference on Decision and Control*, pages 109–114, 2018.
- [63] A. Fontan and C. Altafini. The role of frustration in collective decision-making dynamical processes on multiagent signed networks. *IEEE Transactions on Automatic Control*, 2021.

- [64] A. Fontan and C. Altafini. A signed network perspective on the government formation process in parliamentary democracies. *Scientific reports*, 11(1):1–17, 2021.
- [65] S. Fortunato, V. Latora, A. Pluchino, and A. Rapisarda. Vector opinion dynamics in a bounded confidence consensus model. *International Journal of Modern Physics C*, 16:1535–1551, 2005.
- [66] J. H. Fowler. Connecting the congress: A study of cosponsorship networks. *Political Analysis*, 14(4):456–487, 2006.
- [67] A. Franci, A. Bizyaeva, S. Park, and N. E. Leonard. Analysis and control of agreement and disagreement opinion cascades. *Swarm Intelligence*, 15:47–82, 2021.
- [68] A. Franci, M. Golubitsky, A. Bizyaeva, and N. E. Leonard. A model-independent theory of consensus and dissensus decision making. *arXiv:1909.05765 [math.OC]*, Sept. 2020.
- [69] A. Franci, M. Golubitsky, I. Stewart, A. Bizyaeva, and N. E. Leonard. Breaking indecision in multi-agent, multi-option dynamics, 2022.
- [70] A. Franci, V. Srivastava, and N. E. Leonard. A realization theory for bio-inspired collective decision making. *arXiv:1503.08526*, Mar. 2015.
- [71] N. E. Friedkin and E. C. Johnsen. Social influence networks and opinion change. In *Advances in Group Processes*, volume 16, page 1–29. Emerald Group Publishing Limited, 1999.
- [72] N. E. Friedkin, A. V. Proskurnikov, R. Tempo, and S. E. Parsegov. Network science on belief system dynamics under logic constraints. *Science*, 354(6310):321–326, 2016.
- [73] D. Fudenberg and J. Tirole. *Game theory*. MIT press, 1991.
- [74] L. G. Gajewski, J. Sienkiewicz, and J. A. Hołyst. Transitions between polarization and radicalization in a temporal bilayer echo-chamber model. *Physical Review E*, 105(2):024125, 2022.
- [75] L. V. Gambuzza and M. Frasca. A criterion for stability of cluster synchronization in networks with external equitable partitions. *Automatica*, 100:212–218, 2019.
- [76] S. Garnier, J. Gautrais, and G. Theraulaz. The biological principles of swarm intelligence. *Swarm intelligence*, 1(1):3–31, 2007.
- [77] B. P. Gerkey and M. J. Matarić. A formal analysis and taxonomy of task allocation in multi-robot systems. *The International journal of robotics research*, 23(9):939–954, 2004.

- [78] J. Ghaderi and R. Srikant. Opinion dynamics in social networks with stubborn agents: Equilibrium and convergence rate. *Automatica*, 50:3209–3215, 2014.
- [79] P. W. Glimcher and E. Fehr. *Neuroeconomics: Decision making and the brain*. Academic Press, 2013.
- [80] M. Golubitsky and R. Lauterbach. Bifurcations from synchrony in homogeneous networks: linear theory. *SIAM Journal on Applied Dynamical Systems*, 8(1):40–75, 2009.
- [81] M. Golubitsky, M. Pivato, and I. Stewart. Interior symmetry and local bifurcation in coupled cell networks. *Dynamical Systems*, 19(4):389–407, 2004.
- [82] M. Golubitsky and D. G. Schaeffer. *Singularities and Groups in Bifurcation Theory*, volume 51 of *Applied Mathematical Sciences*. Springer-Verlag, New York, NY, 1985.
- [83] M. Golubitsky and I. Stewart. *The Symmetry Perspective*, volume 200 of *Progress in Mathematics*. Birkhäuser Basel, New-York, 1 edition, 2002.
- [84] M. Golubitsky, I. Stewart, and D. Shaeffer. *Singularities and Groups in Bifurcation Theory, Vol. II*, volume 69 of *Applied Mathematical Sciences*. Springer-Verlag, New York, NY, 1988.
- [85] A. E. Goodenough, N. Little, W. S. Carpenter, and A. G. Hart. Birds of a feather flock together: Insights into starling murmuration behaviour revealed using citizen science. *PloS one*, 12(6):e0179277, 2017.
- [86] A. W. Gouldner. The norm of reciprocity: A preliminary statement. *American Sociological Review*, 25(2):161–178, 1960.
- [87] R. Gray, A. Franci, V. Srivastava, and N. E. Leonard. Multiagent decision-making dynamics inspired by honeybees. *IEEE Transactions on Control of Network Systems*, 5(2):793–806, 2018.
- [88] R. A. Gray. *Designing collective decision-making dynamics for multi-agent systems with inspiration from honeybees*. PhD thesis, Princeton University, 2019.
- [89] S. Grey. Numbers and beyond: The relevance of critical mass in gender research. *Politics & Gender*, 2(4):492–502, 2006.
- [90] J. Guckenheimer and P. Holmes. *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*, volume 42. Springer Science & Business Media, 2013.

- [91] R. Guimera, L. Danon, A. Diaz-Guilera, F. Giralt, and A. Arenas. Self-similar community structure in a network of human interactions. *Physical review E*, 68(6):065103, 2003.
- [92] M. Hechter and K.-D. Opp. *Social Norms*. Russell Sage Foundation, 2001.
- [93] R. Hegselmann and U. Krause. Opinion dynamics and bounded confidence models, analysis, and simulations. *Journal of Artificial Societies and Social Simulation*, 5(3):121–132, 2002.
- [94] R. Hegselmann and U. Krause. Opinion dynamics driven by various ways of averaging. *Computational Economics*, 25(4):381–405, 2005.
- [95] J. P. Hespanha. *Linear systems theory*. Princeton university press, 2018.
- [96] L. Heuer and A. Orland. Cooperation in the prisoner’s dilemma: An experimental comparison between pure and mixed strategies. *Royal Society Open Science*, 6(7):182142, 2019.
- [97] M. W. Hirsch and H. Smith. Monotone dynamical systems. In *Handbook of differential equations: ordinary differential equations*, volume 2, pages 239–357. Elsevier, 2006.
- [98] J. J. Hopfield. Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences*, 79(8):2554–2558, 1982.
- [99] J. J. Hopfield. Neurons with graded response have collective computational properties like those of two-state neurons. *Proceedings of the National Academy of Sciences*, 81(10):3088–3092, 1984.
- [100] R. A. Horn and C. R. Johnson. Topics in matrix analysis, 1991. *Cambridge University Press, Cambridge*, 37:39, 1991.
- [101] R. A. Horn and C. R. Johnson. *Matrix analysis*. Cambridge university press, 2012.
- [102] A. Hu, J. Cao, M. Hu, and L. Guo. Event-triggered group consensus for multi-agent systems subject to input saturation. *Journal of the Franklin Institute*, 355(15):7384–7400, 2018.
- [103] T. Hu and Z. Lin. *Control Systems with Actuator Saturation: Analysis and Design*. Springer Science & Business Media, 2001.
- [104] A. Jevtic, A. Gutiérrez, D. Andina, and M. Jamshidi. Distributed bees algorithm for task allocation in swarm of robots. *IEEE Systems Journal*, 6(2):296–304, 2011.
- [105] P. Jia, A. MirTabatabaei, N. E. Friedkin, and F. Bullo. Opinion dynamics and the evolution of social power in influence networks. *SIAM Review*, 57(3):367–397, 2015.

- [106] K. Wang and N. Michel. Robustness and perturbation analysis of a class of nonlinear systems with applications to neural networks. *IEEE Trans. Circuits Syst. I*, 41(1):24–32, Jan 1994.
- [107] J. W. Kable and P. W. Glimcher. The neurobiology of decision: consensus and controversy. *Neuron*, 63(6):733–745, 2009.
- [108] T. Kato. *Perturbation theory for linear operators*, volume 132. Springer Science & Business Media, 2013.
- [109] H. Khalil. *Nonlinear systems*. Pearson Education Int. Inc., third edition, 2000.
- [110] A. Khamis, A. Hussein, and A. Elmogy. Multi-robot task allocation: A review of the state-of-the-art. In A. Koubâa and J. Martínez-de Dios, editors, *Cooperative Robots and Sensor Networks*, pages 31–51. Springer, 2015.
- [111] C. A. Kitts and M. Egerstedt. Design, control, and applications of real-world multirobot systems [from the guest editors]. *IEEE Robotics & Automation Magazine*, 15(1):8–8, 2008.
- [112] A. Kolling, P. Walker, N. Chakraborty, K. Sycara, and M. Lewis. Human interaction with robot swarms: A survey. *IEEE Transactions on Human-Machine Systems*, 46(1):9–26, 2016.
- [113] P. Kollock. Social dilemmas: The anatomy of cooperation. *Annual Review of Sociology*, 24(1):183–214, 1998.
- [114] G. A. Korsah, A. Stentz, and M. B. Dias. A comprehensive taxonomy for multi-robot task allocation. *The International Journal of Robotics Research*, 32(12):1495–1512, 2013.
- [115] M. J. Krieger, J.-B. Billeter, and L. Keller. Ant-like task allocation and recruitment in cooperative robots. *Nature*, 406(6799):992–995, 2000.
- [116] R. Kühn. Spectra of sparse random matrices. *Journal of Physics A: Mathematical and Theoretical*, 41(29):295002, 2008.
- [117] Y. A. Kuznetsov, I. A. Kuznetsov, and Y. Kuznetsov. *Elements of applied bifurcation theory*, volume 112. Springer, 1998.
- [118] A. Landherr, B. Friedl, and J. Heidemann. A critical review of centrality measures in social networks. *Business & Information Systems Engineering*, 2(6):371–385, 2010.
- [119] D. Leenaerts and V. W. Bokhoven. *Piecewise Linear Modeling and Analysis*. Springer Science & Business Media, 1998.

- [120] N. E. Leonard, K. Lipsitz, A. Bizyaeva, A. Franci, and Y. Lelkes. The nonlinear feedback dynamics of asymmetric political polarization. *Proceedings of the National Academy of Sciences*, 118(50), 2021.
- [121] N. E. Leonard, T. Shen, B. Nabet, L. Scardovi, I. D. Couzin, and S. A. Levin. Decision versus compromise for animal groups in motion. *Proc. National Academy of Sciences*, 109(1):227–232, 2012.
- [122] K. Lerman, C. Jones, A. Galstyan, and M. J. Matarić. Analysis of dynamic task allocation in multi-robot systems. *The International Journal of Robotics Research*, 25(3):225–241, 2006.
- [123] A. Li, S. P. Cornelius, Y.-Y. Liu, L. Wang, and A.-L. Barabási. The fundamental advantages of temporal networks. *Science*, 358(6366):1042–1046, 2017.
- [124] C.-K. Li and F. Zhang. Eigenvalue continuity and ger\{s\} gorin’s theorem. *Electronic Journal of Linear Algebra*, 35:619–625, 2019.
- [125] L. Li, A. Scaglione, A. Swami, and Q. Zhao. Consensus, polarization and clustering of opinions in social networks. *IEEE Journal on Selected Areas in Communications*, 31(5):1072–1083, 2013.
- [126] D. Liu and A. N. Michel. *Dynamical Systems with Saturation Nonlinearities: Analysis and Design*. Springer-Verlag, 1994.
- [127] J. Liu, X. Chen, T. Başar, and M. A. Belabbas. Exponential convergence of the discrete- and continuous-time altafini models. *IEEE Transactions on Automatic Control*, 62(12):6168–6182, 2017.
- [128] J. Lorenz. Continuous opinion dynamics of multidimensional allocation problems under bounded confidence: More dimensions lead to better chances for consensus. In *Eur. J. Econ. Soc. Syst.*, volume 19, pages 213–227, 2006.
- [129] R. Ma, H.-H. Chen, Y.-R. Huang, and W. Meng. Smart grid communication: Its challenges and opportunities. *IEEE transactions on Smart Grid*, 4(1):36–46, 2013.
- [130] K. Macarthur, R. Stranders, S. Ramchurn, and N. Jennings. A distributed anytime algorithm for dynamic task allocation in multi-agent systems. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 25, pages 701–706, 2011.

- [131] M. W. Macy, M. Ma, D. R. Tabin, J. Gao, and B. K. Szymanski. Polarization and tipping points. *Proceedings of the National Academy of Sciences*, 118(50):e2102144118, 2021.
- [132] A. Mao, L. Dworkin, S. Suri, and D. J. Watts. Resilient cooperators stabilize long-run cooperation in the finitely repeated prisoner’s dilemma. *Nature Communications*, 8(1):13800, Jan 2017.
- [133] A. Matthews. The role of network structure and heterogeneity in nonlinear multi-agent opinion formation. Undergraduate thesis, Princeton University, Princeton NJ, 2020.
- [134] W. Mei, F. Bullo, G. Chen, J. M. Hendrickx, and F. Dörfler. Rethinking the micro-foundation of opinion dynamics: Rich consequences of an inconspicuous change. *IFAC-PapersOnLine*, 53(5):307–310, 2020.
- [135] W. Mei, F. Bullo, G. Chen, J. M. Hendrickx, and F. Dörfler. Micro-foundation of opinion dynamics: Rich consequences of the weighted-median mechanism. *Physical Review Research*, 4(2):023213, 2022.
- [136] D. Meng, M. Du, and Y. Jia. Interval bipartite consensus of networked agents associated with signed digraphs. *IEEE Transactions on Automatic Control*, 61(12):3755–3770, 2016.
- [137] M. Mesbahi and M. Egerstedt. *Graph theoretic methods in multiagent networks*. Princeton University Press, 2010.
- [138] F. Miranda-Villatoro, F. Castañón, and A. Franci. Equivalence of linear complementarity problems: Theory and application to nonsmooth bifurcations. *arXiv preprint arXiv:2108.06917*, 2021.
- [139] R. E. Mirollo and S. H. Strogatz. Jump bifurcation and hysteresis in an infinite-dimensional dynamical system of coupled spins. *SIAM J. Appl. Math.*, 50(1):108–124, 1990.
- [140] R. E. Mirollo and S. H. Strogatz. Synchronization of pulse-coupled biological oscillators. *SIAM Journal on Applied Mathematics*, 50(6):1645–1662, 1990.
- [141] A. J. Morales, J. Borondo, J. C. Losada, and R. M. Benito. Measuring political polarization: Twitter shows the two sides of venezuela. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 25(3):033114, 2015.
- [142] M. Morrison, J. N. Kutz, and M. Gabbay. Transitions between peace and systemic war as bifurcations in a signed network dynamical system. *arXiv preprint arXiv:2203.04451*, 2022.

- [143] S. Musslick, A. Bizyaeva, S. Agaron, N. E. Leonard, and J. D. Cohen. Stability-flexibility dilemma in cognitive control: a dynamical system perspective. In *Proc. Annual Meeting Cog. Sci. Soc.*, Montreal, 2019.
- [144] B. Nabet, N. E. Leonard, I. D. Couzin, and S. A. Levin. Dynamics of decision making in animal group motion. *Journal of Nonlinear Science*, 19:399–435, 2009.
- [145] Y. Nakamura, K. Torii, and T. Munakata. Neural-network model composed of multidimensional spin neurons. *Physical Review E*, 51(2):1538–1546, 1995.
- [146] I. Navarro and F. Matía. A survey of collective movement of mobile robots. *International Journal of Advanced Robotic Systems*, 10(1):73, 2013.
- [147] A. Nedić and B. Touri. Multi-dimensional Hegselmann-Krause dynamics. In *IEEE Conference on Decision and Control*, pages 68–73, 2012.
- [148] G. Notomista and M. Egerstedt. Persistification of robotic tasks. *IEEE Trans. Ctrl. Sys. Tech.*, 29(2):756–767, 2021.
- [149] D. Noutsos. On perron–frobenius property of matrices having some negative entries. *Linear Algebra and its Applications*, 412(2-3):132–153, 2006.
- [150] H. Oh, A. R. Shirazi, C. Sun, and Y. Jin. Bio-inspired self-organising multi-robot pattern formation: A review. *Robotics and Autonomous Systems*, 91:83–100, 2017.
- [151] K.-K. Oh, M.-C. Park, and H.-S. Ahn. A survey of multi-agent formation control. *Automatica*, 53:424–440, 2015.
- [152] R. Olfati-Saber. Flocking for multi-agent dynamic systems: Algorithms and theory. *IEEE Transactions on automatic control*, 51(3):401–420, 2006.
- [153] R. Olfati-Saber and R. M. Murray. Consensus problems in networks of agents with switching topology and time-delays. *IEEE Transactions on Automatic Control*, 49(9):1520–1533, 2004.
- [154] N. O’Clery, Y. Yuan, G.-B. Stan, and M. Barahona. Observability and coarse graining of consensus dynamics through the external equitable partition. *Physical Review E*, 88(4):042805, 2013.
- [155] R. Pagliara, D. M. Gordon, and N. E. Leonard. Regulation of harvester ant foraging as a closed-loop excitable system. *PLoS computational biology*, 14(12):e1006200, 2018.

- [156] D. Pais, P. M. Hogan, T. Schlegel, N. R. Franks, and N. E. Leonard. A mechanism for value-sensitive decision-making. *PLoS One*, 8:e73216, 2013.
- [157] L. Pan, H. Shao, M. Mesbahi, Y. Xi, and D. Li. Bipartite consensus on matrix-valued weighted networks. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 66(8):1441–1445, 2019.
- [158] S. Park, A. Bizyaeva, M. Kawakatsu, A. Franci, and N. Leonard. Tuning cooperative behavior in games with nonlinear opinion dynamics. *IEEE Control Systems Letters*, 6:2030–2035, 2022.
- [159] S. Park, Y. D. Zhong, and N. E. Leonard. Multi-robot task allocation games in dynamically changing environments. In *IEEE International Conference on Robotics and Automation (ICRA)*, pages 8678–8684, Xi’an, China, 2021.
- [160] S. E. Parsegov, A. V. Proskurnikov, R. Tempo, and N. E. Friedkin. Novel multidimensional models of opinion dynamics in social networks. *IEEE Trans. Aut. Ctrl.*, 62(5):2270–2285, 2017.
- [161] G. D. Pasquale and M. E. Valcher. Consensus for clusters of agents with cooperative and antagonistic relationships. *arXiv:2008.12398 [eess.SY]*, Mar. 2020.
- [162] I. Pinkoviezky, I. D. Couzin, and N. Gov. Collective conflict resolution in groups on the move. *Phys. Rev. E*, 97:032304, 2018.
- [163] A. Pogromsky, G. Santoboni, and H. Nijmeijer. Partial synchronization: from symmetry towards stability. *Physica D: Nonlinear Phenomena*, 172(1-4):65–87, 2002.
- [164] M. Prokopenko. *Guided self-organization: Inception*, volume 9. Springer Science & Business Media, 2013.
- [165] A. V. Proskurnikov, A. S. Matveev, and M. Cao. Opinion dynamics in social networks with hostile camps: Consensus vs. polarization. *IEEE Transactions on Automatic Control*, 61(6):1524–1536, 2015.
- [166] R. Ramprasad. Designing and implementing strategies for human-robot interaction using a nonlinear model of opinion dynamics. Undergraduate thesis, Princeton University, Princeton NJ, 2022.
- [167] A. Reina, J. A. R. Marshall, V. Trianni, and T. Bose. Model of the best-of- N nest-site selection process in honeybees. *Physical Review E*, 95(5):052411, 2017.

- [168] A. Reina, J. A. R. Marshall, V. Trianni, and T. Bose. Model of the best-of- N nest-site selection process in honeybees. *Phys. Rev. E*, 95:052411, 2017.
- [169] P. Reverdy and D. E. Koditschek. A dynamical system for prioritizing and coordinating motivations. *SIAM Journal on Applied Dynamical Systems*, 17(2):1683–1715, 2018.
- [170] P. B. Reverdy, V. Vasilopoulos, and D. E. Koditschek. Motivation dynamics for autonomous composition of navigation tasks. *IEEE Transactions on Robotics*, 37(4):1239–1251, 2021.
- [171] Y. Rizk, M. Awad, and E. W. Tunstel. Cooperative heterogeneous multi-robot systems: A survey. *ACM Computing Surveys (CSUR)*, 52(2):1–31, 2019.
- [172] F. Roces. Individual complexity and self-organization in foraging by leaf-cutting ants. *The Biological Bulletin*, 202(3):306–313, 2002.
- [173] F. Rossi, S. Bandyopadhyay, M. Wolf, and M. Pavone. Review of multi-agent algorithms for collective behavior: a structural taxonomy. *IFAC-PapersOnLine*, 51(12):112–117, 2018.
- [174] A. Roxin and A. Ledberg. Neurobiological models of two-choice decision making can be reduced to a one-dimensional nonlinear diffusion equation. *PLoS Computational Biology*, 4(3):e1000046, 2008.
- [175] M. Rubenstein, C. Ahler, and R. Nagpal. Kilobot: A low cost scalable robot system for collective behaviors. In *2012 IEEE international conference on robotics and automation*, pages 3293–3298. IEEE, 2012.
- [176] N. Sabor, S. Sasaki, M. Abo-Zahhad, and S. M. Ahmed. A comprehensive survey on hierarchical-based routing protocols for mobile wireless sensor networks: Review, taxonomy, and future directions. *Wireless Communications and Mobile Computing*, 2017, 2017.
- [177] M. T. Schaub, N. O’Clery, Y. N. Billeh, J.-C. Delvenne, R. Lambiotte, and M. Barahona. Graph partitions and cluster synchronization in networks of oscillators. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 26(9):094821, 2016.
- [178] M. Schranz, M. Umlauft, M. Sende, and W. Elmenreich. Swarm robotic behaviors and current applications. *Frontiers in Robotics and AI*, 7:36, 2020.
- [179] T. D. Seeley and S. C. Buhrman. Nest-site selection in honey bees: how well do swarms implement the” best-of-n” decision rule? *Behavioral Ecology and Sociobiology*, 49(5):416–427, 2001.

- [180] T. D. Seeley and P. K. Visscher. Quorum sensing during nest-site selection by honeybee swarms. *Behavioral Ecology and Sociobiology*, 56(6):594–601, 2004.
- [181] T. D. Seeley, P. K. Visscher, T. Schlegel, P. M. Hogan, N. R. Franks, and J. A. R. Marshall. Stop signals provide cross inhibition in collective decision-making by honeybee swarms. *Science*, 335(6064):108–111, 2012.
- [182] R. Sepulchre. Consensus on nonlinear spaces. *IFAC Proceedings Volumes*, 43(14):1029 – 1039, 2010.
- [183] R. Sepulchre, D. A. Paley, and N. E. Leonard. Stabilization of planar collective motion with limited communication. *IEEE Transactions on Automatic Control*, 53(3):706–719, 2008.
- [184] G. Shi, C. Altafini, and J. S. Baras. Dynamics over signed networks. *SIAM Review*, 61(2):229–257, 2019.
- [185] A. F. Siegenfeld and Y. Bar-Yam. Negative representation and instability in democratic elections. *Nature Phys.*, 16:186–190, 2020.
- [186] A. Sirbu, V. Loreto, V. D. P. Servedio, and F. Tria. Opinion dynamics with disagreement and modulated information. *Journal of Statistical Physics*, 151(1):218–237, 2013.
- [187] H. L. Smith. Systems of ordinary differential equations which generate an order preserving flow. a survey of results. *SIAM Review*, 30(1):87–113, 1988.
- [188] F. Sorrentino, L. M. Pecora, A. M. Hagerstrom, T. E. Murphy, and R. Roy. Complete characterization of the stability of cluster synchronization in complex dynamical networks. *Science Advances*, 2(4):e1501737, 2016.
- [189] V. H. Sridhar, L. Li, D. Gorbonos, M. Nagy, B. R. Schell, T. Sorochnik, N. S. Gov, and I. D. Couzin. The geometry of decision-making in individuals and collectives. *Proceedings of the National Academy of Sciences*, 118(50):e2102157118, 2021.
- [190] D. Stevanovic. *Spectral radius of graphs*. Academic Press, 2014.
- [191] I. Stewart, M. Golubitsky, and M. Pivato. Symmetry groupoids and patterns of synchrony in coupled cell networks. *SIAM Journal on Applied Dynamical Systems*, 2(4):609–646, 2003.
- [192] J. Stimson. *Public Opinion in America: Moods, Cycles, and Swings*. Routledge, 2018.
- [193] J. Stimson. Data from “Public Policy Mood”, 2021.

- [194] A. Strandburg-Peshkin, D. R. Farine, I. D. Couzin, and M. C. Crofoot. Shared decision-making drives collective movement in wild baboons. *Science*, 348(6241):1358–1361, 2015.
- [195] G. Strang. *Introduction to Linear Algebra*. Cambridge, fifth edition, 2016.
- [196] S. H. Strogatz. From Kuramoto to Crawford: Exploring the onset of synchronization in populations of coupled oscillators. *Physica D: Nonlinear Phenomena*, 143(1-4):1–20, 2000.
- [197] Taixiong Zheng and Jiongqiu Li. Multi-robot task allocation and scheduling based on fish swarm algorithm. In *2010 8th World Congress on Intelligent Control and Automation*, pages 6681–6685, 2010.
- [198] T. Tao. *Topics in random matrix theory*, volume 132. American Mathematical Soc., 2012.
- [199] C. Thompson and P. Reverdy. Multi-agent recurrent rendezvous using drive-based motivation. *arXiv preprint arXiv:2109.13408*, 2021.
- [200] W. Tian and J. Liu. An improved artificial fish swarm algorithm for multi robot task scheduling. In *2009 Fifth International Conference on Natural Computation*, volume 4, pages 127–130, 2009.
- [201] P. Törnberg, C. Andersson, K. Lindgren, and S. Banisch. Modeling the emergence of affective polarization in the social media society. *Plos one*, 16(10):e0258259, 2021.
- [202] P. Trautman and A. Krause. Unfreezing the robot: Navigation in dense, interacting crowds. In *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 797–803. IEEE, 2010.
- [203] M. L. Tuballa and M. L. Abundo. A review of the development of smart grid technologies. *Renewable and Sustainable Energy Reviews*, 59:710–725, 2016.
- [204] K. Tunström, Y. Katz, C. C. Ioannou, C. Huepe, M. J. Lutz, and I. D. Couzin. Collective states, multistability and transitional behavior in schooling fish. *PLoS Comput Biol*, 9(2):e1002915, 2013.
- [205] M. Usher and J. L. McClelland. The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review*, 108(3):550–592, 2001.
- [206] G. Valentini, E. Ferrante, and M. Dorigo. The best-of-n problem in robot swarms: Formalization, state of the art, and novel perspectives. *Frontiers in Robotics and AI*, 4:9, 2017.

- [207] P. Van Mieghem. *Graph spectra for complex networks*. Cambridge University Press, 2010.
- [208] A. Vanderbauwhede. Lyapunov–schmidt method for dynamical systems. In *Mathematics of Complexity and Dynamical Systems*. 2011.
- [209] L. Wang, Y. Hong, G. Shi, and C. Altafini. A biased assimilation model on signed graphs. In *IEEE Conference on Decision and Control*, pages 494–499, 2020.
- [210] L. Wen, X. Liang, J. Zhong, and B. Long. Oscillation of the nco model. *Physica A: statistical mechanics and its applications*, 391(11):3300–3307, 2012.
- [211] H. R. Wilson and J. D. Cowan. Excitatory and inhibitory interactions in localized populations of model neurons. *Biophysical Journal*, 12(1):1–24, 1972.
- [212] H. R. Wilson and J. D. Cowan. A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. *Kybernetik*, 13(2):55–80, 1973.
- [213] A. F. Winfield, C. J. Harper, and J. Nembrini. Towards the application of swarm intelligence in safety critical systems. In *2006 1st IET International Conference on System Safety*, pages 89–95. IET, 2006.
- [214] A. Witmer. Development of robot navigation systems. Undergraduate thesis, Princeton University, Princeton NJ, 2020.
- [215] A. Woolcock, C. Connaughton, Y. Merali, and F. Vazquez. Fitness voter model: Damped oscillations and anomalous consensus. *Physical Review E*, 96(3):032313, 2017.
- [216] M. Wooldridge. *An introduction to multiagent systems*. John wiley & sons, 2009.
- [217] W. Xia and M. Cao. Clustering in diffusively coupled networks. *Automatica*, 47(11):2395–2405, 2011.
- [218] W. Xia, M. Ye, J. Liu, M. Cao, and X.-M. Sun. Analysis of a nonlinear opinion dynamics model with biased assimilation. *Automatica*, 120:109113, 2020.
- [219] X.-S. Yang, Z. Cui, R. Xiao, A. H. Gandomi, and M. Karamanoglu. *Swarm intelligence and bio-inspired computation: theory and applications*. Newnes, 2013.
- [220] M. Ye, J. Liu, B. D. O. Anderson, C. Yu, and T. Başar. Evolution of social power in social networks with dynamic topology. *IEEE Trans. Aut. Ctrl.*, 63(11):3793–3808, 2018.

- [221] M. Ye, J. Liu, L. Wang, B. D. Anderson, and M. Cao. Consensus and disagreement of heterogeneous belief systems in influence networks. *IEEE Transactions on Automatic Control*, 65(11):4679–4694, 2019.
- [222] G. F. Young, L. Scardovi, A. Cavagna, I. Giardina, and N. E. Leonard. Starling flock networks manage uncertainty in consensus at low cost. *PLoS computational biology*, 9(1):e1002894, 2013.
- [223] T. Zaslavsky. Signed graphs. *Discrete Applied Mathematics*, 4(1):47–74, 1982.
- [224] T. Zaslavsky. Matrices in the theory of signed simple graphs. Number 13, pages 207–229, 2010.
- [225] D. Zhang, P. Shi, Q.-G. Wang, and L. Yu. Analysis and synthesis of networked control systems: A survey of recent advances and challenges. *ISA transactions*, 66:376–392, 2017.
- [226] Y. D. Zhong and N. E. Leonard. A continuous threshold model of cascade dynamics. In *2019 IEEE 58th Conference on Decision and Control (CDC)*, pages 1704–1709. IEEE, 2019.