

**Behaviour and Stability of Interconnected Systems
From Biological Applications to Opinion Dynamics**

Devia Pinzon, C.A.

DOI

[10.4233/uuid:eb5c1ae4-f580-4fa0-9faa-3cbb94f04ee8](https://doi.org/10.4233/uuid:eb5c1ae4-f580-4fa0-9faa-3cbb94f04ee8)

Publication date

2023

Document Version

Final published version

Citation (APA)

Devia Pinzon, C. A. (2023). *Behaviour and Stability of Interconnected Systems: From Biological Applications to Opinion Dynamics*. [Dissertation (TU Delft), Delft University of Technology]. <https://doi.org/10.4233/uuid:eb5c1ae4-f580-4fa0-9faa-3cbb94f04ee8>

Important note

To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

BEHAVIOUR AND STABILITY OF INTERCONNECTED SYSTEMS

FROM BIOLOGICAL APPLICATIONS TO OPINION DYNAMICS

BEHAVIOUR AND STABILITY OF INTERCONNECTED SYSTEMS

FROM BIOLOGICAL APPLICATIONS TO OPINION DYNAMICS

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus prof. dr. ir. T. H. J. J. van der Hagen,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen op
maandag 12 juni 2023 om 12:30 uur

door

Carlos Andrés DEVIA PINZON

Master of Electronic Engineering,
Pontificia Universidad Javeriana, Bogotá, Colombia,
geboren te Bogotá, Colombia.

Dit proefschrift is goedgekeurd door de promotoren.

Samenstelling promotiecommissie bestaat uit:

Rector Magnificus,	voorzitter
Prof. dr. T. Keviczky,	Technische Universiteit Delft, promotor
Prof. dr. ir. G. Giordano,	Università degli Studi di Trento, Italië, promotor

Onafhankelijke leden:

Prof. dr. ir. M. Cao,	Rijksuniversiteit Groningen
Prof. dr. ir. B. De Schutter,	Technische Universiteit Delft
Prof. dr. ir. P.F.A. Van Mieghem,	Technische Universiteit Delft
Dr. W.S. Rossi,	Randstad Sourceright



Keywords: Interconnected systems, robust stability, agent based opinion formation models, classification-based opinion formation, linear systems, nonlinear systems, network dynamics, dynamical networks, opinion dynamics, social systems.

Printed by: Proefschriftspecialist

Front & Back: Carlos Andrés Devia Pinzon

Copyright © 2023 by C.A. Devia Pinzon

ISBN 978-94-6384-440-6

An electronic version of this dissertation is available at
<http://repository.tudelft.nl/>.

*I knew exactly what to do.
But in a much more real sense, I had no idea what to do.*

— Michael Scott, The Office, Season 5: Stress Relief

CONTENTS

Summary	xi
Samenvatting	xiii
1 Introduction	1
1.1 On Interconnected Systems	2
1.2 Structure of the Dissertation	3
1.3 Main Contributions	4
1.3.1 Part 1: Analysis of Interconnected MIMO Systems	4
1.3.2 Part 2: Modelling and Analysis of Opinion Formation Dynamics	5
1.4 Data and code availability.	6
1.5 Terminology clarification	6
I Part 1: Analysis of Interconnected MIMO Systems	9
2 Background on Interconnected MIMO Systems	11
2.1 Introduction	13
2.1.1 Families of Networked Systems and Uncertainties	13
2.1.2 Stability Conditions	15
2.2 Notation and Preliminaries	17
3 Frequency Domain Analysis	21
3.1 Main Results	22
3.1.1 Topology-independent nominal stability.	23
3.1.2 Topology-independent robust stability with homogeneous uncertainties.	24
3.1.3 Topology-independent robust stability with heterogeneous uncertainties.	25
3.2 Examples	27
3.2.1 α -Convergence of a Suspension Bridge	29
3.3 Proofs	33
3.3.1 Lemma 3.	33
3.3.2 Proof of Theorem 3.	34
3.3.3 Proof of Theorem 4.	34
3.3.4 Proof of Lemma 2	35
3.3.5 Proof of Theorem 5.	35
3.3.6 Proof of Theorem 6.	36
3.3.7 Proof of Corollary 1	37
3.3.8 Proof of Corollary 2	37
3.3.9 Proof of Theorem 7.	37

3.3.10	Proof of Theorem 8.	39
3.3.11	Proof of Theorem 9.	39
3.4	Conclusions.	39
4	Time Domain Analysis	41
4.1	Main Results	42
4.2	Example: An Application to Cancer Biology.	45
4.3	Proofs.	47
4.3.1	Lemma 4.	47
4.3.2	Proof of Theorem 10.	48
4.3.3	Proof of Theorem 11.	49
4.3.4	Proof of Theorem 12.	50
4.3.5	Proof of Proposition 4.	51
4.3.6	Proof of Proposition 3.	51
4.4	Conclusions.	52
II	Part 2: Modelling and Analysis of Opinion Formation Dynamics	53
5	Background on Opinion Formation Models	55
5.1	Introduction, Motivation, and Contribution	56
5.2	Graphs and Opinions	62
5.3	Opinion Formation Models	62
5.3.1	French-DeGroot Model	62
5.3.2	Weighted-Median Model.	62
5.3.3	Bounded Confidence Model	63
5.3.4	Quantum Game Model.	63
5.3.5	Friedkin-Johnsen Model	63
5.3.6	Backfire Effect and Biased Assimilation Model	64
5.4	Digraph Topology	64
5.5	Additional Notes on the Simulation Results	67
5.5.1	Model Names	67
5.5.2	Histogram-Based Sorting Algorithm Difference	67
5.5.3	Use of World Values Survey Data	67
5.5.4	Collections of Underlying Digraphs	68
5.5.5	Agent Opinion and Parameter Allocation	69
6	Analysis of Agent-Based Opinion Formation Models	71
6.1	Opinion Distribution Classes and Transition Tables.	73
6.1.1	Opinion Distribution Sorting	73
6.1.2	Model Predictions vs. Real Opinions: a Framework for Systematic Comparison	75
6.1.3	Real Transition Tables Highlight Features of Opinion Evolution	77
6.1.4	Predicted Transition Tables and Model Comparison	77

6.2	Transition Table Analysis	80
6.3	Agreement Plot Analysis	81
6.3.1	Basic Idea	81
6.3.2	Agreement Plot of Time Evolution	84
6.3.3	Agreement Plot of Steady State	91
6.4	Probabilistic Analysis	97
6.4.1	Basic Idea	97
6.4.2	Implementation	99
6.5	Agreement Plot and Probabilistic Analysis Applied to the BEBA and Bounded Confidence Models	108
6.5.1	Backfire Effect and Biased Assimilation Model	108
6.5.2	Bounded Confidence Model	122
6.6	Conclusions.	130
7	Classification-Based Opinion Formation Model Embedding Agents' Psycho- logical Traits	133
7.1	The Classification-Based Model.	134
7.1.1	Model Parameters	136
7.2	Simulation Results	139
7.2.1	Simulations in Simple Cases	140
7.2.2	Parameter Sensitivity Analysis	141
7.2.3	Model Validation with Real Data	146
7.2.4	Comparison with the Friedkin-Johnsen Model.	154
7.2.5	Agreement Plot and Probabilistic Analysis	157
7.3	Summary and Conclusions	170
7.4	Appendix A: Network Metrics	170
7.5	Appendix B: Simulation Process.	172
7.6	Appendix D: Countries and Questions	173
7.7	Appendix E: Tables	174
8	Conclusion	181
	Acknowledgements	185
	Bibliography	202
	Curriculum Vitae	203
	List of Publications	205

SUMMARY

An interconnected system is composed of multiple well-defined self-contained subsystems that interact among them and that together create collective behaviours. We can find many examples of interconnected systems in real life. Ranging from biological systems, such as the growth and interaction of populations in diverse and spatially distributed environments, to electric grids connecting power-generating sources, buildings and infrastructures in a country. When studying interconnected systems, a fundamental and natural question is how the properties and characteristics of the individual subsystems and the way they are connected relate to the collective behaviour of the complete system. That is the driving question of the present dissertation.

Given that interconnected systems can be found in a wide variety of contexts, their representation and specific research interests can be equally varied. Because of this, it is impossible to answer the aforementioned question uniquely for all interconnected systems, and specific cases must be considered. In this dissertation, we consider two types of interconnected systems: a general class of uncertain multiple-input-multiple-output (MIMO) systems, and agent-based opinion formation models.

The investigation of uncertain MIMO interconnected systems is focused on providing topology-independent conditions for robust stability. The primary motivation for this approach is that, in real systems, it is costly or even impossible to have complete and accurate information on the network topology and subsystem parameters and dynamics. However, it is of critical interest to guarantee the system's stability. Therefore we need stability conditions that require only partial information about the network and the subsystems to ensure the system's stability. By studying these systems both in the time and frequency domain, we are able to provide conditions that meet these requirements.

As for agent-based opinion formation models, we assume that each individual (or agent) in a population has an opinion about a statement. By exchanging opinions among themselves, the agents update their own internal opinion, resulting in a collective dynamic of opinion evolution. When studying these systems, the interests shift from stability conditions, to a characterisation of the relation between the agents' individual traits and qualitative properties of the opinion distribution in the population. Several techniques and approaches to analyse opinion formation models are proposed and applied to multiple models, one of which is new to this dissertation.

The collective study of the previously mentioned interconnected systems requires the use of multiple and diverse analysis techniques and approaches, from analytical methods based on the Nyquist criterion, Bauer-Fike theorem, and Lyapunov functions to qualitative and numerical analysis techniques like histograms and binomial proportion confidence intervals. It is our hope that some of the presented results, methods, or ideas may advance the knowledge frontier in this scientific field, spark new research directions, and either directly or indirectly prove some value to society.

SAMENVATTING

Een onderling verbonden systeem is gemaakt uit meerdere goed gedefinieerde afzonderlijke subsystemen die op elkaar inwerken en samen collectief gedrag vertonen. We kunnen veel voorbeelden vinden van zulke systemen in het echte leven. Variërend van biologische systemen, zoals de groei en interactie van populaties in diverse en ruimtelijk verspreide omgevingen, tot elektriciteitsnetten geschakeld aan energie producerende bronnen, gebouwen en infrastructuur in een land. Wanneer je deze systemen onderzoekt, is een fundamentele en natuurlijke vraag hoe de eigenschappen en karakteristieken van de individuele subsystemen en de manier waarop ze verbonden zijn verband houden met het collectieve gedrag van het complete systeem. Dat is de drijvende vraag van dit proefschrift.

Omdat deze systemen gevonden kunnen worden in verschillende soorten contexten, kunnen de representatie en het specifieke onderzoeksdoel evenveel verschillen. Hierdoor is het onmogelijk om de hiervoor genoemde vraag voor alle systemen gelijk te beantwoorden en zullen specifieke gevallen beschouwd moeten worden. In dit proefschrift beschouwen we twee types van systemen: een generieke klasse van onzekere multiple-input-multiple-output (MIMO) systemen en agent-based meningsvormingmodellen.

Het onderzoek van onzekere MIMO-systemen richt zich op het verstrekken van topologisch onafhankelijke condities voor robuuste stabiliteit. De primaire motivatie voor deze benadering is dat, in echte systemen, het kostbaar en zelfs onmogelijk is om volledige en accurate informatie van de netwerktopologie en parameters van de subsystemen en de dynamica te hebben. Echter, het is cruciaal om de stabiliteit van het systeem te garanderen. Daarom hebben we stabiliteitsvoorwaarden nodig die alleen gedeeltelijke informatie van het netwerk en de subsystemen vereisen om de stabiliteit van het systeem te verzekeren. Bij het bestuderen van deze systemen zowel in het tijd- als frequentiedomein, kunnen wij de voorwaarden verstrekken die aan deze eisen voldoen.

Wat betreft agent-based meningsvormingmodellen, nemen we aan dat elke individu (of agent) in een populatie een mening heeft over een statement. Bij het uitwisselen van meningen binnen de groep updaten de agents hun eigen interne mening, resulterend in een collectieve dynamica van meningsontwikkeling. Wanneer je deze systemen onderzoekt, verschuift de belangstelling van de stabiliteit van de voorwaarden naar een karakterisering van de relaties tussen de individuele kenmerken van de agents en de kwalitatieve eigenschappen van de meningsverdeling in de populatie. Verscheidene technieken en benaderingen om meningsvormingmodellen te analyseren zijn voorgesteld en toegepast in meerdere modellen, waarvan één nieuw is voor dit proefschrift.

Het gezamenlijke onderzoek van de hiervoor genoemde onderling verbonden systemen vereist het gebruik van meerdere en diverse analysetechnieken en benaderingen, van analytische methodes gebaseerd op de Nyquist criterion, Bauer-Fike theorem en Lyapunov functions tot kwalitatieve en numerieke analysetechnieken zoals histogrammen en binomial proportion confidence intervals. Het is onze hoop dat sommige ge-

presenteerde resultaten, methodes, of ideeën de grens van onze kennis doet opschuiven in dit onderzoeksveld, nieuwe onderzoeksrichtingen inspireren en direct of indirect waarde zullen opleveren voor de maatschappij.

1

INTRODUCTION

What is thy bidding, my master ?

Darth Vader, Star Wars Episode V: The Empire Strikes Back

This initial chapter explains the broad motivation, main contributions, and structure of the dissertation. The dissertation comprises two self-contained parts that study different classes of interconnected systems from various perspectives. This chapter aims to elaborate on the connection of these classes and present them in the broader field of interconnected systems. This chapter also comments on the data sets and code availability and contains a dedicated section to clarify the used terminology.

1.1. ON INTERCONNECTED SYSTEMS

The overall theme of this dissertation is the study of interconnected systems and their properties. An interconnected system is a collection of individual systems or agents that communicate between them by exchanging information about their state [85, 74, 27]. This definition can be applied to an incredibly wide range of systems in many disciplines. Some examples include supply networks, where the subsystems are storage facilities and their interaction is the movement of goods [190, 119, 131]; epidemic models, where the subsystems are disease hosts and their interaction leads to disease transmission [210, 229, 43]; sensor networks, where the subsystems are individual sensors that exchange estimates of some variables [193, 21, 149, 219, 194, 121]; and collective group behaviour, where the subsystems are biological or artificial individuals (animals or robots, for instance) and they exchange information about their state [72, 142, 84].

Given the wide range of contexts and applications, interconnected systems can be represented differently and studied with various research goals in mind. For instance, supply networks can be modelled by discrete-event dynamics and simulations [225, 88, 64], and their study may focus on control of material flows [138] measuring real-time demand [221], or developing stabilisation strategies for a demand-driven supply network [220]; epidemic models can use ordinary differential equations [29, 182, 50, 169, 228, 12], Markov chains [4, 38, 39], or discrete-time systems [11, 206, 76, 227], and their study may aim to find regions of stability for the disease-free equilibrium [211, 145, 230, 146, 224, 44]; differential equations can represent individual robot dynamics [181, 113, 13, 67], and their collective study may seek to find synchronisation algorithms [33, 130, 141, 63].

This dissertation approaches the study of interconnected systems from two different perspectives, reflected in the two parts of the manuscript. The first part studies the robust asymptotic stability of uncertain systems formed by the interconnection of linear time-invariant (LTI), multiple-input-multiple-output (MIMO) subsystems [61, 60, 62]. This first body of work is at the core of classical control theory, both in terms of system representation (LTI systems in the time and the frequency domain) and research interest (robustness and stability conditions). Reflecting the fact that in real systems, knowledge of the exact network topology and subsystem dynamics is rarely available, the provided stability conditions do not require complete information about the system in terms of both interconnection topology and subsystem parameters. Consequently, the main results of this first part are topology-independent robust stability conditions for uncertain systems. Thanks to the generality of the considered class of models and to the topology-independent and robust characteristics of the results, the conditions are very general and can be applied in a wide range of fields.

The second part of the dissertation studies a very different class of interconnected systems: agent-based opinion formation models [159, 57, 58, 59]. Here the subsystems are individuals or agents that hold an opinion about a statement and communicate among themselves. Because of this communication, each agent changes its opinion according to some opinion update law. These systems are often highly nonlinear and can be represented in continuous or discrete time by differential or difference equations [179, 180, 159]. Given the nature of these types of models, their behaviour and our ensuing research interest in them are very different from those related to models studied in the first part of the manuscript. Social systems are not physical systems; no “univer-

sal laws” accurately model people’s behaviour; individuals can be complicated, chaotic, unpredictable, and unreasonable. There is no equivalent to Newton’s law to how people behave. Accurately measuring people’s opinions, attitudes, values, influence, and perception is nearly impossible and will never be standardised due to the subjective nature of definitions. None of these concepts has units or unanimous ways of being measured.

Hence, the second part of the dissertation does not aim to find stability conditions or mathematically prove properties of opinion formation models. Instead, it focuses on characterising the models in terms of opinion transitions they can predict, possible model outcomes, and their relation with real opinion evolution [57, 58, 59]. Despite the lack of a unified framework to study and model opinions and the high number of subjective assumptions we make when proposing models, making formal and insightful statements about these models is possible. Which conditions are needed for a population to achieve polarisation? How probable is it that opinions will remain constant in a given time interval? How do opinions in real societies evolve? All these are essential questions that can be approximately answered by studying agent-based opinion formation models and their qualitative behaviours, which are the subject of the second part of this dissertation.

A common theme in Parts 1 and 2 is the motivation of “saying as much as possible, with as little quantitative information as possible”. In Part 1, this is translated into providing stability conditions that do not depend on knowledge of the complete network topology and subsystem parameters. In Part 2, this becomes giving probabilities that the opinion of a population will evolve in a certain qualitative way given incomplete information about the agent opinions, parameters, and interconnections. In both cases, this objective (and to some degree necessity, due to the reality of actual applications) of providing answers in spite of incomplete information is both challenging and rewarding: the amount of information or guarantees the results can provide is limited, while at the same time being applicable to a broader range of realistic cases.

Because the interconnected systems considered in Parts 1, and 2 are so different, each part is self-contained and can be read independently. The structure of the dissertation is detailed in Section 1.2. Section 1.3 provides a summary of the main contributions of each part, while Section 1.4 provides additional information on the data and code availability to reproduce the results presented in this dissertation. Finally, Section 1.5 clarifies some possible misunderstandings caused by different terminology.

1.2. STRUCTURE OF THE DISSERTATION

Part 1 of the dissertation studies the interconnection of LTI MIMO systems. It is composed of three chapters. Chapter 2 provides the preliminary information used in the next two chapters, this includes a specific introduction to the subject and highlights the main contributions. Chapters 3 and 4 study the interconnection of LTI MIMO systems in the frequency and in the time domains, respectively. Because of the difference in representation and analysis tools available for each domain, the precise model structure of the studied systems is not exactly the same: in Chapter 3, the subsystems and their interconnections have associated transfer function matrices [62, 60], whereas in Chapter 4, each subsystem has a different state-space representation and their interconnections are matrices of the appropriate size [61]. Also, due to this different representation, the

resulting conditions require different information on the subsystems and their interconnection. Chapters 3 and 4 are independent and self-contained. To improve clarity, both chapters have the same structure: main results are presented first, followed by examples, proofs, and then conclusions.

Part 2 of the dissertation analyses agent-based opinion formation models and is also composed of three chapters. Chapter 5 introduces the field of opinion formation modelling, emphasises the viewpoint and focus of this dissertation, and presents some existing opinion formation models and notation. It is followed by Chapters 6 and 7. Unlike Part 1, these chapters are sequential, as Chapter 6 introduces a novel methodology for the analysis of agent-based opinion formation models [57, 59], exemplified in existing models [45, 79, 110, 54, 55, 163], and Chapter 7 presents a new model [58], which is then analysed with the methodology proposed in Chapter 6. This part also includes an appendix with details on the simulation process, which are not necessary to understand the main results. The methodology presented in Chapter 6 is composed of four techniques: *Histogram Sorting Algorithm*, *Transition Tables*, *Agreement Plot*, and *Probabilistic Analysis*. The chapter explains each of these techniques, their relation with one another and their application to existing models in the literature. The model presented in Chapter 7 can be seen as building upon a classical opinion-based model: the Friedkin-Johnsen model [78, 79]. It combines three well-known psychological traits, namely, conformism, radicalism, and stubbornness. Simulation results presented in this chapter show that the proposed model has the potential to mimic opinion evolutions seen in real societies.

Finally, Chapter 8 presents the conclusion of the dissertation.

The overall structure of the dissertation can be seen in Figure 1.1.

1.3. MAIN CONTRIBUTIONS

This section summarises the main contributions of the dissertation for both parts, each preceded by the associated goal.

1.3.1. PART 1: ANALYSIS OF INTERCONNECTED MIMO SYSTEMS

Goal: provide stability conditions, for networks of interconnected LTI MIMO subsystems, that do not depend on complete knowledge of the network topology or subsystem dynamics.

Main results:

- Topology-independent sufficient condition for nominal stability in the frequency domain (Theorem 5) [62].
- Topology-independent sufficient condition for robust stability in the presence of homogeneous uncertainties in the frequency domain (Theorem 6) [62].
- Topology-independent sufficient condition for robust stability in the presence of heterogeneous uncertainties in the frequency domain (Theorem 8) [60].
- Topology-independent necessary conditions for robust stability in the presence of heterogeneous uncertainties in the time domain (Theorem 12 and Proposition 2) [61].

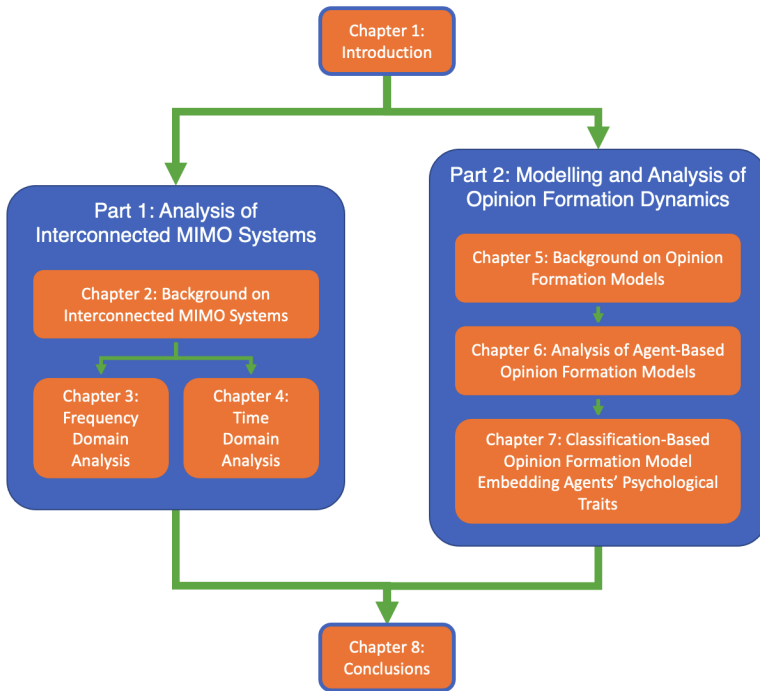


Figure 1.1: Structure of the dissertation: it is divided into two self-contained parts. Part 1 contains three chapters which are also self-contained and partially sequential. Part 2 contains three sequential chapters.

- Topology-independent sufficient conditions for robust stability in the presence of heterogeneous uncertainties in the time domain (Theorem 11 and Proposition 1) [61].

1.3.2. PART 2: MODELLING AND ANALYSIS OF OPINION FORMATION DYNAMICS

Goal: explore the capability of agent-based opinion formation models to (i) produce a certain opinion evolution, and in particular characterise which qualitative opinion evolutions can occur and under which circumstances; (ii) look for global patterns of opinion behaviours predicted by the models; and (iii) gain as much insight as possible into opinion evolution when only limited information is available.

Main results:

- A *Histogram-based Sorting Algorithm* [57], which sorts opinion sets into five different qualitative classes: perfect consensus, consensus, polarisation, clustering, and dissensus, based on formal definitions of each of these classes.
- Introduction of *Transition Tables* [57], which enable the high-level study of opinion evolutions for different models, agent parameters, and initial opinions. Transition Tables allow the study and characterisation of opinion evolution in a real so-

ciety and comparison with model predictions, pointing to possible model changes to make them more accurate.

- The representation of opinion sets by the *Agreement Plot* [59], which provides a visual and intuitive representation of a set of opinions as a single point in the Cartesian plane. This, in turn, enables the characterisation of real opinions, model predictions and model outcomes.
- A *Probabilistic Analysis* [59] that provides information on opinion prediction and evolution even when only incomplete information is available. This is especially important in the context of opinion dynamics, where measurement and estimation of model parameters are impossible or extremely challenging.
- A novel *Classification-based opinion formation model* [58], an agent-based model that combines the psychological mechanisms of conformism, radicalism, and stubbornness and accounts for the imperfect assessment of the opinions of others. The resulting opinion formation model has the potential of reproducing opinion evolution measured in real societies and producing a wide range of opinion distributions.

1.4. DATA AND CODE AVAILABILITY

All data, scripts, and additional information to reproduce the results presented in this dissertation can be found here:

<https://giuliagiordano.dii.unitn.it/docs/DissertationCarlosAndres.zip>

Or, in the 4TU Repository, with DOI

<https://doi.org/10.4121/1fbd5ecd-d0af-4b77-a28a-0aca7b38f45f>

The *.zip* file contains a file describing the code and how to obtain the presented results. The scripts and datasets are separated in different folders.

The code used for the ternary plots is based on the scripts in [189] by Carl Sandrock. Some scripts by David Holdaway [114] were used for the Probabilistic Analysis. All the code is written in Matlab 2021.

1.5. TERMINOLOGY CLARIFICATION

In different sections of the dissertation, a variety of terminology is used to refer to similar concepts. Here possible misunderstandings are clarified:

- The term ‘interconnected system’ refers to a system formed by subsystems that communicate among themselves. It can refer to an abstract system, belonging to a general class as in Part 1, or a more specific system related to a precise application, as in Part 2.
- The terms ‘digraph’ and ‘network’ are mainly used interchangeably and are assumed to have the same meaning, which we formally provide in Chapters 2, and 5. This equivalence also applies to ‘underlying digraph’ and ‘underlying network’.
- All graphs are assumed to be digraphs, so the terms ‘graph’ and ‘digraph’ are also used interchangeably. The terms ‘interaction graph’ and ‘interaction digraph’ also have the same meaning as ‘digraph’.
- Formally, the *topology* of a network or digraph refers to the way vertices are connected by edges. A digraph can be signed, weighted and time-varying, whereas a topology is unsigned

and unweighted and can be time-varying. The topology of a digraph is simply the digraph obtained by setting all the weights to 1.

- When a network/graph is considered, the terms ‘node’ and ‘arc’ have the same meaning as ‘vertices’ and ‘edges’, respectively. The first ones are primarily used in Part 1, whereas the second ones are mainly used in Part 2.
- In Part 2, the vertices are associated with agents that have an opinion. Because of this association, the phrasing ‘the opinion of vertex i ’ should be read as ‘the opinion of the agent associated with vertex i ’.
- In Part 2, the set of opinions of a population is called the *opinion distribution*. When it is clear from context, this term is omitted and instead, this set of opinions is called the ‘agents’ opinions’ or the ‘opinion of the population’.
- In Part 2, to be consistent with the papers where the Backfire Effect and Biased Assimilation (BEBA) model [45] and Classification-based (CB) model [58] were first proposed, the two agent parameters entrenchment (for the BEBA model) and radical trait weight (for the CB model) use the same letter (β). The two models are analysed separately (the BEBA model in Chapter 6 and the CB model in Chapter 7), and from the context it is clear to which model the statements refer.
- In some tables and numerical results the base 10 exponents are preceded by an **e** (this is the notation used by Matlab, the program used throughout this dissertation). An example of the equivalence between that notation and the scientific notation is $4.7\text{e} - 14$ instead of $4.7 \cdot 10^{-14}$.

I

PART 1: ANALYSIS OF INTERCONNECTED MIMO SYSTEMS

2

BACKGROUND ON INTERCONNECTED MIMO SYSTEMS

So it begins.

King Théoden, Lord of the Rings: The Two Towers

This chapter introduces the state of the art, main contributions, and preliminary concepts related to Part 1 of the dissertation, focused on the search for topology-independent robust stability conditions for networked systems formed by the interconnection of uncertain LTI MIMO subsystems. Because LTI systems can be represented both in the frequency and in the time domains, different conditions can be stated for the robust stability of these networked systems, resulting from tools available in each of the two domains. The frequency and time domain analyses are presented in Chapter 4 and Chapter 5, respectively.

The resulting conditions can be applied to two classes of interconnected systems. The first class considers systems, represented in the frequency domain, where both the nodes and arcs have internal dynamics, which are respectively identical for all the nodes and all arcs. For the second class, corresponding to systems represented in the time domain, the dynamics are present only in the nodes (this is not restrictive, since arc dynamics can be absorbed by node dynamics); however, they are not constrained to be identical and not even to have the same number of inputs and outputs.

Regarding topology information, some of the conditions depend only on the maximum connectivity degree of the network, resulting in stability conditions that can be evaluated locally. Other conditions depend on bounds of the eigenvalues of the Laplacian matrix, which for many types of topolo-

Parts of this chapter have been published in the papers “Topology-independent robust stability for networks of homogeneous MIMO systems” (2020) [62], “Topology-independent robust stability conditions for uncertain MIMO networks” (2021) [61], and “MIMO networks with heterogeneous uncertainties: topology-independent robust stability and α -convergence” (2021) [60] by Carlos Andrés Devia and Giulia Giordano.

gies are known. As a result, the provided conditions are topology-independent and also independent of the number of nodes or arcs.

The fact that the provided stability conditions account for subsystem uncertainties and depend only on partial network information makes them very versatile and amenable to be used in a wide variety of fields. As an example of this versatility, we present a suspension bridge application in Chapter 4 and a cancer therapy application in Chapter 5.

2.1. INTRODUCTION

Large-scale networks formed by the interconnection of interacting dynamical subsystems can be found in a wide variety of fields, including smart-grids [223], opinion dynamics [10], biological systems, pharmacokinetics, epidemiology [118, 34, 103, 115], and multi-agent systems where we wish to enforce control [158, 32, 150], estimation [91, 207], consensus [172], or synchronisation [204].

As with any dynamic system, stability is a fundamental property of interest. Assessing it by exploiting the network structure of the overall system gives rise to the question: *under which conditions does the connection of the individual subsystems guarantee the stability of the whole network?* Furthermore, it is common that the exact parameter values of the subsystems are unknown, as is the complete network topology. Therefore, it is crucial that the answer to this question does not depend on knowledge of the complete topology and that it also considers parameter uncertainty. This can be done by providing **topology-independent conditions for the robust stability of interconnected systems**, which is the problem addressed in Part 1 of the dissertation.

For interconnections of SISO (single-input and single-output) LTI (linear time invariant) systems, robust stability conditions are provided in [127] and [123] for particular topologies; in [102, 104] and [115] in the generalised frequency variable framework; and in [143] using the multivariable Nyquist criterion proposed in [56] and the concept of S-hull. This type of conditions is further refined in [176], including a partial extension to interconnections of multiple-input-multiple-output (MIMO) systems.

The MIMO case is fully addressed in [129, 8] using Integral Quadratic Constraints, also used in [175] for control design: these conditions are scalable because, to ensure stability of the whole network, it is enough to satisfy a local condition at each node. In [87] the MIMO case is also addressed in the frequency-domain.

Also an approach based on the generalised frequency variable, proposed in [107], was used to obtain necessary and sufficient stability conditions for networks of homogeneous MIMO LTI systems in [103], where the robustness analysis encompasses heterogeneous systems of almost equal agents for specific topologies.

The above approaches either require the knowledge of the network topology, which is typically not available for complex and large-scale networks, or don't consider uncertainty. Topology-independent sufficient conditions and necessary conditions for robust stability of dynamic networks with both dynamic nodes and dynamic interconnections [170] are considered in [31, 30] for the SISO case: topology-independent robust stability conditions in the frequency domain are obtained for nominally homogeneous node and arc dynamics, with homogeneous or heterogeneous uncertainties. These conditions are based on Nyquist-type arguments.

In this first part of the dissertation, the robust stability of families of interconnected uncertain MIMO LTI systems is studied both in the frequency and in the time domain (Chapters 3 and 4 respectively). We provide topology-independent necessary conditions and sufficient conditions that guarantee robust stability for this type of interconnected systems.

In this first background chapter we comment on the context and importance of the presented results, then preliminary results and notation are presented. Chapters 3 and 4 have identical structure, namely, first constructing the main results, then providing examples, followed by proofs, and finally conclusions.

2.1.1. FAMILIES OF NETWORKED SYSTEMS AND UNCERTAINTIES

All the results of this first part are stated in terms of families of networked systems, which are defined as follows:

Definition 1 (Family of networked systems). *A family $\mathcal{N}$ of networked systems is a collection of*

LTI systems formed by the interconnection of LTI subsystems according to a network structure. The generic system in the family has an underlying graph structure $\mathcal{G} = \{\mathcal{N}, \mathcal{A}\}$, where each node in $\mathcal{N}$, labelled with an integer number in the set $\{1, \dots, N\}$ is associated with a LTI MIMO system and each arc in $\mathcal{A}$, labelled with an integer number in the set $\{1, \dots, M\}$, is associated with a LTI MIMO system or an interconnection matrix.

Definition 1 encompasses a wide variety of systems, mainly because it does not make any assumptions about the individual subsystems or the network topology. Thus, it may refer to any number of possibly different and unstable LTI subsystems connected in any possible configuration. The results presented in this dissertation focus on two specific classes of networked systems:

1. **Networks with identical dynamic nodes and identical dynamic arcs:** For this class of networked systems, all nodes and all arcs have the same dynamics. This characterisation could apply, for instance, to a network of identical sensors that communicate among themselves to estimate spatially-distributed variables. In that example, the nodes would be the sensors, which are assumed to be identical, and the arcs, the communication channels, which also are assumed to be identical. Another example is the coordinated movement of identical agents, like robots. In this case, all the nodes are identical robots, and the arcs are the communication channels. A final example is the modelling of spatially distributed systems via discretisation; for instance, a bridge can be modelled as the interconnection of many identical 'bridge sections' represented by simpler lumped dynamics.
2. **Networks with different dynamic nodes and static arcs:** For this class of networked systems, all the nodes can represent possibly different dynamics and all arcs different static interconnections between nodes. This characterisation could apply, for instance, to connected population subsystems, in which each node has dynamics corresponding to populations in different environments, or to flow networks in which the buffers in the nodes have different geometry and, therefore, different dynamics.

The second class of networked systems is more general than the first, as every system in the first class also belongs to the second class: one simply needs to represent the arc dynamics as another type of node dynamics, which is allowed since, for systems in the second class, the node dynamics are allowed to be different. However, the distinction is relevant since systems belonging to the first class have a representation that allows for more analysis tools to be used.

Even though Definition 1 does not refer to subsystem properties, an assumption that is made for all results presented in Part 1 is the stability of the individual subsystems. This is because, although the interconnection of unstable subsystems may produce a stable system, the intention of providing topology-independent conditions implicitly implies that the empty graph (that is, isolated nodes) is a possibility, and stability for this type of systems requires stability of the individual subsystems.

The first class of networked systems (networks where all the nodes have identical dynamics and all the arcs have identical dynamics) is analysed in the frequency domain, and the second (networks with heterogeneous dynamic nodes and static arcs) in the time domain. Each class of networked systems is associated with a different form of subsystem uncertainties: **additive dynamic uncertainties** in the frequency domain and **additive parametric uncertainties** in the time domain.

Additive dynamic uncertainties occur when unmodelled dynamics are present in the subsystems; a simple example is the parasitic capacitance in electronic circuits. Additive parametric uncertainties are caused by the tolerance in the value of real-life elements that form parts of a system; a simple example is the tolerance of resistance values in electronic circuits. These two forms of uncertainty are present in every real-life system. The degree to which they may be relevant depends on the nature of the system itself.

For dynamic uncertainties, we consider two cases: homogeneous uncertainties (the same for all nodes and for all arcs) and heterogeneous uncertainties (potentially different for each node and for each arc). Although the case of heterogeneous uncertainties is more general and includes as a particular case homogeneous uncertainties, the stability condition for systems with homogeneous uncertainties is easier to check, making it a valuable option when the subsystems are such that all the node uncertainties and all the arc uncertainties can be assumed to be identical. For parametric uncertainties, only the heterogeneous case is considered, and the ensuing results can also be applied when the uncertainties are homogeneous.

The existence of dynamic and parametric uncertainties is the main reason why families of systems are considered. Imagine a connection of electronic circuits. By itself, it is a single interconnected system. However, suppose the value of a single component is not known precisely but lies in a given interval. In that case, each of the values in that interval corresponds to a slightly different system. The collection of all these possible systems, caused by all the values a single component can have, generates in itself a family of systems. The same happens with dynamic uncertainties: all the possible unmodelled dynamics that each subsystem has creates a family of systems. When stability is proven to hold for every member of the family created by uncertainties (whether dynamical or parametric), it is said that the family is *robustly stable*. If at least one family member is unstable, then we cannot say that the family is stable.

2.1.2. STABILITY CONDITIONS

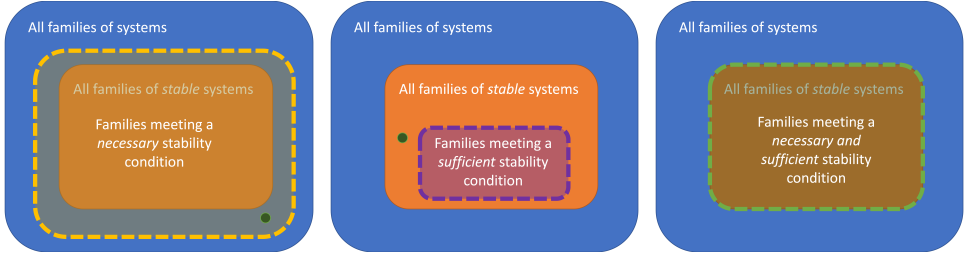
There are three types of stability conditions for families of systems: **necessary conditions**, **sufficient conditions**, and **necessary and sufficient conditions**. When a family of systems meets necessary stability conditions, it does not mean that every element of the family is stable, but when it does not meet the necessary stability conditions, then at least one member of the family is guaranteed to be unstable (See Figure 2.1a). When a family of systems meets sufficient stability conditions, it means that every member of the family is guaranteed to be stable, while if it does not meet the sufficient conditions, it does not mean that there are unstable elements in the family (See Figure 2.1b). Finally, if a family of systems meets necessary and sufficient stability conditions, it means that every member is stable; if it does not meet the condition, it means that the family contains at least one unstable element (See Figure 2.1c).

Both necessary conditions and sufficient conditions are said to be conservative when there are families of systems that have the corresponding property but that cannot be guaranteed by the condition. Necessary conditions are conservative if there exist families of systems that meet the condition but contain unstable members (therefore, the condition is not precise enough to ‘detect’ the unstable members). Analogously, sufficient conditions are said to be conservative if there exist families of systems that do not meet the condition but still all the elements are stable (therefore, the condition is not precise enough to ‘detect’ that indeed all members are stable). Conservativeness can be thought of as ‘asking more than strictly required to guarantee an outcome’.

The less conservative a condition is, the more information it provides on the family of systems. In the limit of zero conservativeness, necessary conditions and sufficient conditions become necessary *and* sufficient conditions. If a necessary condition is not conservative at all, it means that there are no family of systems that meet the condition and contain unstable members, which means that it is also a sufficient condition. On the other hand, a sufficient condition that is not conservative at all means that there are no families of systems that do not meet the condition and still contain only stable members; therefore, it is also a necessary condition.

Based on frequency and time domain analysis the main contributions of Part 1 in the form of necessary conditions and sufficient conditions for robust stability are:

- Topology-independent sufficient condition for nominal stability in the frequency domain (Theorem 5).



(a) Representation of a *necessary* stability condition. All families that do not meet the condition contain at least one unstable member, but there may be families that meet the condition and include unstable members (like the green dot). The level of conservativeness is proportional to the blue area under the yellow area.

(b) Representation of a *sufficient* stability condition. All families that meet the condition are stable, but there may be families that do not meet the condition and are also stable (like the green dot). The level of conservativeness is proportional to the orange area outside the purple area.

(c) Representation of a necessary and sufficient condition. All families that meet the condition are stable. All families that do not meet the condition include unstable members. A necessary and sufficient condition is non conservative.

Figure 2.1: Visual representation of necessary conditions, sufficient conditions and necessary and sufficient conditions. In the diagrams, the blue area contains all families of systems, the orange area the families of stable systems (all their members are stable), the yellow area contains families that meet a *necessary stability condition*, the purple area contains families meeting a *sufficient stability condition*, and the green area contains families that meet a *necessary and sufficient stability condition*.

- Topology-independent sufficient condition for robust stability in the presence of homogeneous uncertainties in the frequency domain (Theorem 6).
- Topology-independent sufficient condition for robust stability in the presence of heterogeneous uncertainties in the frequency domain (Theorem 8).
- Topology-independent necessary conditions for robust stability in the presence of heterogeneous uncertainties in the time domain (Theorem 12 and Proposition 2).
- Topology-independent sufficient conditions for robust stability in the presence of heterogeneous uncertainties in the time domain (Theorem 11 and Proposition 1).

All of the conditions provided in the frequency domain are sufficient with varying degrees of conservativeness. In the time domain, we present two necessary and two sufficient conditions. The degree of the conservativeness of these necessary conditions and sufficient conditions is measured by a single parameter $\chi \geq 1$. We show that when $\chi = 1$, these four conditions become two necessary and sufficient conditions. This happens for the interconnection of diagonal systems. For systems where $\chi > 1$, we provide a result on the minimal χ to minimise the level of conservativeness.

The provided conditions are said to be topology-independent in the sense that they do not require complete information on the network topology. In many cases, the only information required about the network is the maximum connectivity degree, i.e., the maximum number of nodes connected to a single node in the network. This is of particular importance because the degree of a node can be computed locally, therefore if a bound on the maximum connectivity degree is set (to meet a specific stability condition), nodes can be added or removed from the network while ensuring that the bound is met by simply using local information. The fact that the required topology information is the maximum connectivity degree not only allows for conditions that can be checked locally but also for complete scalability: the conditions and the way to check them do not depend on the number of nodes in the network.

All the provided conditions are in the form of inequalities. These inequalities depend on subsystem dynamics and the network topology. For all the main results, except for Theorem 8, the maximum connectivity degree is the only information required for the interconnection network. This means that except for Theorem 8, all the conditions can be checked locally without additional constraints. For Theorem 8, an upper bound on the maximum eigenvalue of the Laplacian matrix is needed. For an arbitrary topology, this may be challenging. However, these bounds can be found in the literature for a variety of well-known and well-studied topologies. In a way, this additional constraint makes Theorem 8 less topology-independent. Nevertheless, the condition does not explicitly require knowledge of the complete topology, and if a bound can be found (even if it is conservative), the condition is still topology-independent.

From the main results, we derive additional conditions. These conditions are generally more conservative but easier to check, which may be helpful in specific applications. Another additional result obtained in the frequency domain is that of topology-independent α -convergence. A system is α -convergent if it is stable and its spectral abscissa (i.e., the maximum real part of the eigenvalues) is smaller than $-\alpha$, with $\alpha > 0$ real. In practice, this means that the system converges with a minimum guaranteed speed. This result can be valuable in some applications where not only system stability is required, but also fast enough convergence is essential (for instance, in a car suspension).

2.2. NOTATION AND PRELIMINARIES

A **directed graph** (digraph) with N nodes and M arcs is represented by the pair $\mathcal{G} = \{\mathcal{N}, \mathcal{A}\}$, where the set of nodes $\mathcal{N}$ is indexed by the set $\{1, \dots, N\}$ and $\mathcal{A} \subset \mathcal{N} \times \mathcal{N}$ is the arc set, with $|\mathcal{A}| = M$ and $\mathcal{A}$ indexed by the set $\{1, \dots, M\}$. The element $(i, j) \in \mathcal{A}$ denotes an arc that leaves node i and enters node j . Two nodes are assumed to be connected by at most one directed arc (meaning, that the set $\mathcal{A}$ does not contain repeated elements). Each node $i \in \{1, \dots, N\}$ has an **outward (respectively, inward) connectivity degree** δ_i^{out} (resp. δ_i^{in}), defined as the number of arcs that leave (resp. enter) the node.

The **maximum outward (respectively, inward) connectivity degree** of the digraph is

$$\mathcal{D}^{\text{out}} = \max_{i \in \mathcal{N}} (\delta_i^{\text{out}}) \quad \mathcal{D}^{\text{in}} = \max_{i \in \mathcal{N}} (\delta_i^{\text{in}})$$

A **bidirectional graph** is a directed graph with the property that if $(i, j) \in \mathcal{A}$, then $(j, i) \in \mathcal{A}$. For this type of graphs the outward and inward connectivity degrees of every node are the same, therefore simply called the **connectivity degree** of the node δ_i , and the **maximum connectivity degree** is defined as

$$\mathcal{D} = \max_i (\delta_i)$$

A **path** from node i to node j is a sequence of nodes $i = i_0, i_1, \dots, i_\ell = j$ such that $(i_{h-1}, i_h) \in \mathcal{A}$ for $h = 1, \dots, \ell$. The graph is **strongly connected** if each pair of nodes is connected by a path. It is **weakly connected** if the associated *bidirectional graph*, where the presence of arc (i, j) implies the presence of arc (j, i) (namely, each existing arc can be crossed in both directions), is strongly connected.

Let $X \in \{-1, 0, 1\}^{N \times M}$ denote the **generalised node-arc incidence matrix** defined as

$$[X]_{ih} = \begin{cases} 1, & \text{if arc } h \text{ enters node } i, \\ -1, & \text{if arc } h \text{ leaves node } i, \\ 0, & \text{otherwise,} \end{cases}$$

For bidirectional graphs, the generalised **Laplacian matrix** of the graph is¹ $L = XX^\top$ [91]. Its diagonal entries are $L_{ii} = \mathfrak{d}_i$, for $i = 1, \dots, N$, while its off-diagonal entries L_{ij} , $i \neq j$, take the value -1 if there is an arc either from node i to node j or from node j to node i , and 0 otherwise. Matrix L is symmetric, hence its eigenvalues $\{\gamma_k\}_{k=1}^N$ are real, and it can be diagonalised by a unitary matrix W :

$$W^{-1}LW = \text{diag}(\gamma_k)_{k=1}^N.$$

By the Gershgorin circle theorem [89], the real eigenvalues of L are located inside a circle of radius $\mathfrak{D}$ with centre in $(\mathfrak{D}, 0)$, since $\mathfrak{D}$ is the maximum element along the diagonal of L . Hence,

$$\sigma(L) = \{\gamma_k\}_{k=1}^N \subset \{z \in \mathbb{R} : 0 \leq z \leq 2\mathfrak{D}\}. \quad (2.1)$$

The p -th norm of matrix A is

$$\|A\|_p = \sup_{v \neq 0} \frac{\|Av\|_p}{\|v\|_p}$$

and $\otimes$ is the Kronecker product of matrices. If $p = 2$, the subscript is omitted in the matrix norm for simplicity.

For the 2-norm it holds that

$$\|A\| = \sup_{v \neq 0} \frac{\|Av\|}{\|v\|} = \sqrt{\lambda_{\max}(A^\top A)},$$

where $\lambda_{\max}(S)$ denotes the largest eigenvalue of a symmetric matrix S , and

$$\|A \otimes Y\| = \|A\| \|Y\| \quad \text{and} \quad \|AY\| \leq \|A\| \|Y\|, \quad (2.2)$$

as discussed in [136]. Also, $\|A\| = \|A^\top\|$.

This implies for the Laplacian matrix L that

$$\|L\| \leq \|X\| \|X^\top\| \leq 2\mathfrak{D}, \quad (2.3)$$

since $\|X\| = \|X^\top\| = \sqrt{\lambda_{\max}(L)} \leq \sqrt{2\mathfrak{D}}$.

The **spectrum** of a square matrix Y is denoted by $\sigma(Y)$ and its *condition number* [25] is defined as

$$\mathcal{K}_p(Y) = \|Y\|_p \|Y^{-1}\|_p.$$

Also for the condition number, we omit the subscript when $p = 2$. For unitary matrices, the condition number is one; therefore if W is a matrix that diagonalises the Laplacian matrix of a bidirectional graph, then $\mathcal{K}(W) = 1$.

The $n \times n$ identity matrix is denoted I_n .

We now consider the 1-norm and the ∞ -norm, defined as

$$\|X\|_1 = \max_{1 \leq j \leq m} \sum_{i=1}^n |X_{ij}| \quad \|X\|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^m |X_{ij}|,$$

for a matrix $X \in \mathbb{C}^{n \times m}$.

¹This is a slightly different meaning from the one found in other works in the literature, such as [20], where generalised Laplacian matrix (sometimes also called the grounded Laplacian matrix) refers to a Laplacian matrix where one row and corresponding column have been removed, to represent a reference (or ground) node.

Theorem 1 (Properties of 1-norm and ∞ -norm [136]). *Given complex matrices A and B of compatible dimensions, $\|AB\|_* \leq \|A\|_* \|B\|_*$ and $\|A \otimes B\|_* = \|A\|_* \|B\|_*$, where the subscript $*$ denotes either always 1 or always ∞ .*

Lemma 1 (Norm of block-diagonal matrices). *The complex block-diagonal matrix $X = \text{diag}(X_k)_{k=1}^K$ has norm $\|X\|_* = \max_{k=1, \dots, K} \{\|X_k\|_*\}$, where the subscript $*$ denotes either always 1 or always ∞ .*

We denote by $\mathcal{H}$ the space of stable, linear, time invariant and continuous-time transfer functions and by $\mathcal{H}^{q \times m}$ the space of $q \times m$ matrices with entries in $\mathcal{H}$.

Given a set $\mathcal{S}$, the set $\zeta(\mathcal{S})$ is defined as

$$\zeta(\mathcal{S}) = \left\{ s : -\frac{1}{s} \in \mathcal{S} \right\}.$$

Finally, we need a bound for the eigenvalues of uncertain matrices, which is provided by the Bauer-Fike theorem.

Theorem 2 (Bauer-Fike theorem[22]). *Consider the two matrices $A, M \in \mathbb{R}^{n \times n}$, with A diagonalisable, that is, $V^{-1}AV = \text{diag}(\lambda_1, \dots, \lambda_n)$ for some $V \in \mathbb{C}^{n \times n}$ and $\lambda_1, \dots, \lambda_n \in \mathbb{C}$. For every (complex) eigenvalue β of $A + M$, there exists an index $i \in \{1, \dots, n\}$ such that $|\beta - \lambda_i| \leq \mathcal{K}_p(V) \|M\|_p$.*

With these preliminaries, we are now ready to present the main results of Part 1.

3

FREQUENCY DOMAIN ANALYSIS

Sometimes, life is like this dark tunnel. You can't always see the light at the end of the tunnel, but if you just keep moving, you will come to a better place.

Uncle Iroh, Avatar: The Last Airbender, Season 2, Episode 20: The Crossroads of Destiny

This chapter considers network systems where the node dynamics are described by identical LTI MIMO subsystems with transfer-function matrix $F(s)$, while the dynamic interactions associated with the bidirectional arcs are described by identical LTI MIMO subsystems with transfer-function matrix $G(s)$; the dynamics of the individual nodes and arcs are affected by either homogeneous or heterogeneous, norm-bounded uncertainties. We provide a topology-independent condition for the robust stability of all possible network systems with a maximum connectivity degree, regardless of their size and interconnection structure. We also give a topology-independent condition that robustly guarantees not only stability, but also α -convergence (i.e. all poles having real part less than a negative $-\alpha$). The proposed frequency-domain conditions are scalable and can be evaluated locally, also for large-scale networks where nodes and arcs can be added or removed in real time. Multiple examples showcase the proposed conditions, including an application of robust α -convergence to a suspension bridge system of arbitrary size.

This chapter is based on the papers “Topology-independent robust stability for networks of homogeneous MIMO systems” (2020) [62], and “MIMO networks with heterogeneous uncertainties: topology-independent robust stability and α -convergence” (2021) [60] by Carlos Andrés Devia and Giulia Giordano.

All the results in this chapter (and the next one) are related to the concept of **Family of networked systems**, whose definition is recalled below. All the networks in this chapter are assumed to have identical node dynamics $F(s)$, identical arc dynamics $G(s)$, and be affected by norm bounded uncertainties in the frequency domain. Given that in this chapter all the systems are represented as transfer function matrices, the main tools used to prove the robust stability conditions will be the Nyquist stability criterion, algebraic properties of the Laplacian matrix, and the Bauer-Fike theorem [22]. To improve the readability of the chapter, first the main results are stated, followed by application examples, proofs, and the conclusions.

Definition 1 (Family of networked systems). *A family $\mathcal{N}$ of networked systems is a collection of LTI systems formed by the interconnection of LTI subsystems according to a network structure. The generic system in the family has an underlying graph structure $\mathcal{G} = \{\mathcal{N}, \mathcal{A}\}$, where each node in $\mathcal{N}$, labelled with an integer number in the set $\{1, \dots, N\}$ is associated with a LTI MIMO system and each arc in $\mathcal{A}$, labelled with an integer number in the set $\{1, \dots, M\}$, is associated with a LTI MIMO system or an interconnection matrix.*

3.1. MAIN RESULTS

We start with some basic assumptions to formally define the type of families considered in this chapter.

Assumption 1. *For each system in the family of networked systems $\mathcal{N}$, the underlying graph is bidirectional with Laplacian L and has maximum connectivity degree $\mathfrak{D}$.*

Assumption 2. *For all elements in the family $\mathcal{N}$, each node in $\mathcal{N}$ (resp. arc in $\mathcal{A}$) is associated with a stable MIMO linear subsystem represented by the transfer-function matrix $F(s) \in \mathcal{H}^{r \times n}$ (resp. $G(s) \in \mathcal{H}^{n \times r}$), which describes its nominal dynamics (n and r are the number of inputs of the nodes and arcs respectively). Furthermore, the generic i -th node (resp. h -th arc) dynamics is affected by the uncertainty $\Delta F_i(s) \in \mathcal{H}^{r \times n}$, $i \in \{1, \dots, N\}$ (resp. $\Delta G_h(s) \in \mathcal{H}^{n \times r}$, $h \in \{1, \dots, M\}$).*

Theorem 3. *Under Assumption 2, a generic element in the family of networked systems $\mathcal{N}$ has the following characteristic polynomial:*

$$p(s) = \det(I_{Nr} + L \otimes H(s) + \mathcal{D}(s)), \quad (3.1)$$

where $L = XX^\top$ is the Laplacian matrix, $H(s) = F(s)G(s)$ and

$$\mathcal{D} = (X \otimes F)\mathcal{D}_G(X^\top \otimes I_r) + \mathcal{D}_F(X \otimes I_n)\mathcal{D}_G(X^\top \otimes I_r) + \mathcal{D}_F(XX^\top \otimes G). \quad (3.2)$$

If all the node and arc uncertainties are homogeneous, i.e. $\Delta F_i(s) = \Delta F(s)$ for all $i \in \mathcal{N}$, and $\Delta G_h(s) = \Delta G(s)$ for all $h \in \mathcal{A}$, then the characteristic polynomial becomes:

$$p(s) = \det(I_{Nr} + L \otimes (H(s) + \Delta_H(s))), \quad (3.3)$$

where

$$\Delta_H(s) = F(s)\Delta_G(s) + \Delta_F(s)G(s) + \Delta_F(s)\Delta_G(s). \quad (3.4)$$

Finally, if there is no uncertainty the characteristic polynomial becomes

$$p(s) = \det(I_{Nr} + L \otimes H(s)). \quad (3.5)$$

Assumption 3. *The transfer-function matrix $H(s) = F(s)G(s)$ does not have poles in the closed right half plane.*

The eigenvalues of $H(s)$, $\sigma(H(s)) = \{\lambda_i(s)\}_{i=1}^r$, are not rational transfer functions: they are not a quotient of polynomials in s . In general, they are complex functions of the variable s . Therefore, the poles of $\lambda_i(s)$ are not the roots of a polynomial but the set of complex numbers $\tilde{p} \in \mathbb{C}$ such that $\lambda_i^{-1}(\tilde{p}) = 0$. We have the following result.

Theorem 4. Consider the transfer-function matrix $H(s) \in \mathcal{H}^{r \times r}$ and its eigenvalues $\{\lambda_i(s)\}$. Let $\tilde{p} \in \mathbb{C}$ be a pole of the complex function $\lambda_i(s)$, for some $i \in \{1, \dots, r\}$. Then, $\tilde{p}$ is a pole of the transfer-function matrix $H(s)$.

Remark 1. In view of Theorem 4, if Assumption 3 is satisfied, then the complex functions $\lambda_i(s)$ are stable: there exists some $\epsilon < 0$ such that, for every pole $\tilde{p}$ of $\lambda_i(s)$, $\text{Re}(\tilde{p}) \leq \epsilon < 0$. In fact, we can pick ϵ as the largest real part of all the poles of $H(s)$, which must be strictly negative in view of Assumption 3. This is the condition we will exploit in the following results.

Computing the poles of the transfer-function matrix $H(s)$, which are the roots of the denominator polynomial, is much easier than computing the poles of the generic complex functions $\lambda_i(s)$. The results presented in this chapter do not require to analytically compute the functions $\lambda_i(s)$, nor their poles.

Lemma 2. Consider the scalar complex function $h(s) = \bar{h}(s) + \delta_h(s)$, where the nominal function $\bar{h}(s)$ is stable and $\delta_h(s)$ is an uncertainty bounded as $|\delta_h(j\omega)| \leq \delta_h^{\max}(\omega)$, where $\delta_h^{\max}(\omega)$ is an assigned real weighting function. Consider also the scalar coefficient $\mu \in \sigma(L_{\mathcal{D}})$, where $L_{\mathcal{D}}$ is a Laplacian matrix with maximum connectivity degree $\mathcal{D}$. Then, the feedback complex function $h^{\text{feed}}(s) = \mu h(s)(1 + \mu h(s))^{-1}$ is robustly stable for all $\mu \in \sigma(L_{\mathcal{D}})$ and for all possible realisations of the uncertainty if, for all $\omega \in \mathbb{R}^+$,

$$\min_{\rho \leq -(2\mathcal{D})^{-1}} |\bar{h}(j\omega) - \rho| > \delta_h^{\max}(\omega). \quad (3.6)$$

Lemma 2 will be the foundation upon which the first results will be derived. The basic idea will be to transform the MIMO systems so that the subsystems are decoupled and we can then apply Lemma 2.

3.1.1. TOPOLOGY-INDEPENDENT NOMINAL STABILITY

In this section we provide a topology-independent condition for the stability of the nominal members of a family (namely, in the absence of uncertainty on the dynamics). First it is necessary to formally define what a nominal member of a family is.

Definition 2 (Nominal networked system). An element in the family of networked systems $\mathcal{N}$, under Assumption 2, is a nominal networked system if $\Delta_{F_i}(s) = 0$ for all $i \in \{1, \dots, N\}$ and $\Delta_{G_h}(s) = 0$ for all $h \in \{1, \dots, M\}$.

By Definition 2 a nominal system is a member of a family of networked systems whose dynamics are not affected by uncertainty.

Theorem 5. Consider the family of networked systems $\mathcal{N}$, under Assumptions 1, 2, and 3. Then, stability is ensured for all the nominal network systems in $\mathcal{N}$ if, for all $i \in \{1, \dots, r\}$ and $\omega \in \mathbb{R}^+$,

$$\min_{\rho \leq -(2\mathcal{D})^{-1}} |\lambda_i(j\omega) - \rho| > 0 \quad \text{where} \quad \sigma(H(s)) = \{\lambda_i(s)\}_{i=1}^r. \quad (3.7)$$

Condition (3.7) can be checked locally, regardless of the network size and topology, and guarantees the stability of the whole dynamic network also if new nodes and arcs are added or removed, as long as the maximum connectivity degree is $\mathcal{D}$. This ensures robustness to online modifications, and scalability.

Definition 3 (Family of nominally stable networked systems). *A family of nominally stable networked systems is a family of networked systems $\mathcal{N}$ where all the nominal networked systems as per Definition 2 are stable.*

Note that Definition 3 does not require that every element in the family is stable. Theorem 5 provides a way to guarantee that a give family $\mathcal{N}$ is a family of nominally stable networked systems.

The results to guarantee stability of every member in a family of networked systems, including those affected by uncertainties, are built upon Definition 3. This means that stability of all the nominal members of a family will be a requisite for stability of all the other members of the family. First, in Section 3.1.2 the case of homogeneous uncertainties will be considered, followed by the case of heterogeneous uncertainties in Section 3.1.3.

3.1.2. TOPOLOGY-INDEPENDENT ROBUST STABILITY WITH HOMOGENEOUS UNCERTAINTIES

We start by assuming that the transfer function matrix $H(s)$ can be diagonalised (which is true for important classes of systems, as pointed out in Remark 2 below) and that the uncertainties are homogeneous.

Assumption 4. *The transfer-function matrix $H(s) = F(s)G(s)$, with eigenvalues $\sigma(H(s)) = \{\lambda_i(s)\}_{i=1}^r$, can be diagonalised by the change-of-basis matrix $V(s)$, so that*

$$V(s)^{-1}H(s)V(s) = \text{diag}(\lambda_1(s), \dots, \lambda_r(s)).$$

Remark 2. *For the important classes of MISO and SIMO systems, Assumption 4 is automatically satisfied. In fact, if $F(s)$ is a row vector and $G(s)$ is a column vector, then $H(s)$ is a scalar function; if $F(s)$ is a column vector and $G(s)$ is a row vector, then $H(s)$ is a rank-one matrix, hence it is diagonalisable.*

Assumption 5. *Both node and arc uncertainties are homogeneous, i.e. $\Delta_{F_i}(s) = \Delta_F(s)$ for all $i \in \{1, \dots, N\}$, and $\Delta_{G_h}(s) = \Delta_G(s)$ for all $h \in \{1, \dots, r\}$, and they satisfy the bound $\|\Delta_H(j\omega)\| \leq \Delta_H^{\max}(\omega)$, where*

$$\Delta_H(s) = F(s)\Delta_G(s) + \Delta_F(s)G(s) + \Delta_F(s)\Delta_G(s). \quad (3.8)$$

We can now state the main result of this section.

Theorem 6. *Consider the family of nominally stable networked systems $\mathcal{N}$, under Assumptions 1, 2, 3, 4, and 5. Then, stability is ensured for all networks in $\mathcal{N}$ if, for all $i \in \{1, \dots, r\}$ and $\omega \in \mathbb{R}^+$,*

$$\min_{\rho \leq -(2\mathfrak{D})^{-1}} |\lambda_i(j\omega) - \rho| > \mathcal{K}(V(j\omega))\Delta_H^{\max}(\omega). \quad (3.9)$$

where $\mathcal{K}(X) = \|X\| \|X^{-1}\|$ is the condition number of matrix X .

Condition (3.9) for robust stability is scalable, because it can be checked locally and it is size- and topology-independent. It allows stability-preserving plug-and-play modifications [26, 184] to the network as long as the maximum connectivity degree remains $\mathfrak{D}$, which can be checked only by the newly added nodes or arcs.

Some conservativeness is introduced by the Bauer-Fike theorem [22], on which the proof of the result relies (see Section 3.3.6 for the proof).

Let us now consider a suitable upper bound $\phi(\omega)$ for the spectral radius of $H(j\omega)$,

$$|\lambda_i(j\omega)| \leq \phi(\omega) \quad \text{for all } i, \quad \text{for } \omega \in \mathbb{R}_+, \quad (3.10)$$

and a suitable upper bound $\xi(\omega)$ for the condition number $\mathcal{K}(V(j\omega))$,

$$\mathcal{K}(V(j\omega)) \leq \xi(\omega) \quad \text{for } \omega \in \mathbb{R}_+. \quad (3.11)$$

Bounds of these types have been well studied in the literature; for instance, [46] presents a comprehensive survey on bounds for condition numbers.

Then, a more conservative sufficient condition, which however allows to assess robust topology-independent stability without the need of computing the eigenvalues and eigenvectors of $H(s)$, is the following.

Corollary 1. *Consider the family of nominally stable networked systems $\mathcal{N}$, under Assumptions 1, 2, 3, 4, and 5. Then, stability is ensured for all networks in $\mathcal{N}$ if*

$$^c\mathcal{C} + \mathcal{M} < (2\mathfrak{D})^{-1}, \quad (3.12)$$

with

$$^c\mathcal{C} = \sup_{\omega \in \mathbb{R}^+} \{\phi(\omega)\},$$

where $\phi(\omega)$ is the bound in (3.10), and

$$\mathcal{M} = \sup_{\omega \in \mathbb{R}^+} \{\xi(\omega) \Delta_H^{max}(\omega)\},$$

where $\xi(\omega)$ is the bound in (3.11) and $\|\Delta_H(j\omega)\| \leq \Delta_H^{max}(\omega)$.

Corollary 2. *Consider the family of nominally stable networked systems $\mathcal{N}$, under Assumptions 1, 2, 3, 4, and 5. Then, stability is ensured for all networks in $\mathcal{N}$ if each node i satisfies $\mathfrak{d}_i < \mathcal{T}$, where*

$$\mathcal{T} = \frac{1}{2(^c\mathcal{C} + \mathcal{M})}, \quad (3.13)$$

and $^c\mathcal{C}$ and $\mathcal{M}$ are defined as in Corollary 1.

Corollary 2 gives *fully local* sufficient conditions for robust stability, which are *independent* of the network size and topology and do not even rely on the shared knowledge of the maximum connectivity degree. As long as each node satisfies the local condition, new arcs and nodes can be added, or removed, and the overall networked system remains stable. Furthermore two separate stable dynamic networks can be connected and, as long as all the connecting nodes satisfy the local condition, the resulting dynamic network is stable.

Remark 3. *The conditions in Theorems 5 and 6 allow us to verify if, given a maximum connectivity degree $\mathfrak{D}$, we have topology-independent stability for all networks with node dynamics $F(s)$ and arc dynamics $G(s)$, possibly in the presence of homogeneous bounded uncertainties. The conditions in Corollaries 1 and 2 can be alternatively interpreted as providing the largest $\mathfrak{D}$ such that all networks with maximum connectivity degree $\mathfrak{D}$, node dynamics $F(s)$ and arc dynamics $G(s)$ are guaranteed to be (robustly) stable.*

3.1.3. TOPOLOGY-INDEPENDENT ROBUST STABILITY WITH HETEROGENEOUS UNCERTAINTIES

When considering the case of heterogeneous uncertainties, we start by assuming that the uncertainties are bounded.

Assumption 6. *The nominal node and arc transfer function matrices never vanish, i.e. $\|F(j\omega)\| \neq 0$ and $\|G(j\omega)\| \neq 0 \forall \omega \in \mathbb{R}^+$, and node and arc uncertainties are bounded as*

$$\frac{\|\mathcal{D}_F(j\omega)\|}{\|F(j\omega)\|} \leq K_F(j\omega) \quad \text{and} \quad \frac{\|\mathcal{D}_G(j\omega)\|}{\|G(j\omega)\|} \leq K_G(j\omega),$$

where $\mathcal{D}_F = \text{diag}(\Delta_{F_i})_{i=1}^N$ and $\mathcal{D}_G = \text{diag}(\Delta_{G_h})_{h=1}^M$. Equivalently, since $\|\text{diag}(X_i)\| = \max_i \|X_i\|$, we can assume the local uncertainty bounds

$$\begin{aligned} \frac{\|\Delta_{F_i}(j\omega)\|}{\|F(j\omega)\|} &\leq K_F(j\omega) \quad \forall i \in \{1, \dots, N\}, \\ \frac{\|\Delta_{G_h}(j\omega)\|}{\|G(j\omega)\|} &\leq K_G(j\omega) \quad \forall h \in \{1, \dots, M\}. \end{aligned}$$

With the uncertainties bounded by Assumption 6 we can state the first result of this section.

Theorem 7. *Consider the family of nominally stable networked systems $\mathcal{N}$, under Assumptions 1, 2, 3, 4, and 6. Then, all systems in $\mathcal{N}$ are robustly stable if the inequality*

$$|\gamma_k \lambda_i(j\omega) + 1| > 2\mathfrak{D}\zeta(F, G)K(j\omega) \quad (3.14)$$

holds for all $k \in \{1, \dots, N\}$, for all $i \in \{1, \dots, r\}$, and for all $\omega \in \mathbb{R}^+$, where $\{\lambda_i(s)\}_{i=1}^r$ are the eigenvalues of $H(s)$, γ_k are the eigenvalues of the corresponding Laplacian matrix, and

$$\zeta(F, G) = \mathcal{K}(V(j\omega))\|F(j\omega)\|\|G(j\omega)\|, \quad (3.15)$$

$$K(j\omega) = K_F(j\omega) + K_G(j\omega) + K_F(j\omega)K_G(j\omega). \quad (3.16)$$

The condition (3.14) in Theorem 7 can be easily checked numerically. However, it is topology-dependent: complete knowledge of the network is required to compute the Laplacian eigenvalues of H_k . The result can be refined, yielding a *topology-independent* condition.

Theorem 8. *Consider the family of nominally stable networked systems $\mathcal{N}$, under Assumptions 1, 2, 3, 4, and 6. Then, topology-independent stability is robustly guaranteed for every element in $\mathcal{N}$ if the inequality*

$$\min_{i \in \{1, \dots, r\}} \{\phi_i(j\omega)\} > 2\mathfrak{D}\zeta(F, G)K(j\omega) \quad (3.17)$$

holds for all $\omega \in \mathbb{R}^+$, where

$$\phi_i(j\omega) = \begin{cases} 1 & \text{if } 0 \leq \text{Re}(\lambda_i(j\omega)), \\ \frac{|\text{Im}(\lambda_i(j\omega))|}{|\lambda_i(j\omega)|} & \text{if } -\rho|\lambda_i(j\omega)|^2 \leq \text{Re}(\lambda_i(j\omega)) < 0, \\ |\rho\lambda_i(j\omega) + 1| & \text{if } \text{Re}(\lambda_i(j\omega)) < -\rho|\lambda_i(j\omega)|^2, \end{cases} \quad (3.18)$$

ρ is an upper bound for the maximum eigenvalue of the Laplacian matrix, $\zeta(F, G)$ is defined as in (3.15) and $K(j\omega)$ is defined as in (3.16).

Theorem 8 avoids the need of computing the eigenvalues of the Laplacian matrix and provides a *topology-independent* condition: all the required information about the topology is the maximum connectivity degree $\mathfrak{D}$ and an upper bound ρ for the maximum Laplacian eigenvalue. The general upper bound provided in (2.1) is $\rho = 2\mathfrak{D}$, but for many common topologies tighter bounds exist.

So far, the main objective has been to guarantee stability of the overall network. However, in many instances it is desirable to have fast enough convergence, or equivalently an upper bound on the settling time, as per the next definition.

Definition 4. A LTI system with pole set $\mathcal{P}$ is α -convergent if $\text{Re}(p) < -\alpha < 0$ for all $p \in \mathcal{P}$.

Based on Theorem 8, we can obtain *topology-independent* sufficient conditions to robustly certify not only stability, but also α -convergence.

Theorem 9. Consider the family of nominally stable networked systems $\mathcal{N}$, under Assumptions 1, 2, 3, 4, and 6. Then, topology-independent α -convergence is robustly guaranteed for every element in $\mathcal{N}$ if the inequalities

$$\min_{\xi < -(2\mathfrak{D})^{-1}} \{|\lambda_i(j\omega - \alpha) - \xi|\} > 0, \quad i \in \{1, \dots, r\}, \quad (3.19)$$

$$\min_{i \in \{1, \dots, r\}} \{\phi_i(j\omega - \alpha)\} > 2\mathfrak{D}\hat{\zeta}(F, G)K(j\omega - \alpha), \quad (3.20)$$

hold for all $\omega \in \mathbb{R}^+$, with ϕ_i as in (3.18),

$$\hat{\zeta}(F, G) = \mathcal{H}(V(j\omega - \alpha))\|F(j\omega - \alpha)\|\|G(j\omega - \alpha)\|$$

and K as in (3.16).

The suitable value of α is problem-dependent and can be selected based on the desired settling time, which can be approximated as $4/\alpha$.

3.2. EXAMPLES

First we provide some numerical examples where we check the stability of nominal interconnected systems according to Theorem 5.

Example 1. Consider an arbitrary network where the node and arc nominal transfer-function matrices are

$$F(s) = \frac{1}{s^2 + 3.412s + 2.871} \begin{bmatrix} 0.1307s - 0.08404 \\ -0.1105s + 0.06774 \end{bmatrix},$$

$$G(s) = \frac{1}{s^2 + 1.805s + 0.4837} \begin{bmatrix} 0.1485s - 0.753 \\ 0.4924s + 0.329 \end{bmatrix}^T,$$

while the maximum connectivity degree is $\mathfrak{D} = 5$. Numerically evaluating the eigenvalues $\lambda_1(s)$ and $\lambda_2(s)$ of $H(s)$ for $s = j\omega$ is enough to check that the stability condition (3.7) in Theorem 5 is satisfied. The blue lines in Figure 3.1 represent the values $\min_{\rho \leq -(2\mathfrak{D})^{-1}} |\lambda_i(j\omega) - \rho|$ for $i = 1, 2$. Since the blue lines are always greater than zero the condition in Equation (3.7) is satisfied.

Example 2. Consider the node and arc nominal transfer-function matrices

$$F(s) = \frac{1}{s^2 + 3.338s + 2.613} \begin{bmatrix} 0.08071s + 2.308 \\ 0.6187s + 1.78 \end{bmatrix},$$

$$G(s) = \frac{1}{s^2 + 0.9124s + 0.2505} \begin{bmatrix} 0.8022s + 0.2204 \\ 0.1696s + 0.2161 \end{bmatrix}^T.$$

We wish to check whether stability is guaranteed for all networks with maximum connectivity degree $\mathfrak{D} = 5$. The stability condition (3.7) in Theorem 5 is violated: in Figure 3.2, the blue line representing $\min_{\rho \leq -(2\mathfrak{D})^{-1}} |\lambda_1(j\omega) - \rho|$ is zero for $\omega \approx 2.5$. There **could exist** at least one dynamic network with the given node and arc nominal transfer-function matrices and with maximum connectivity

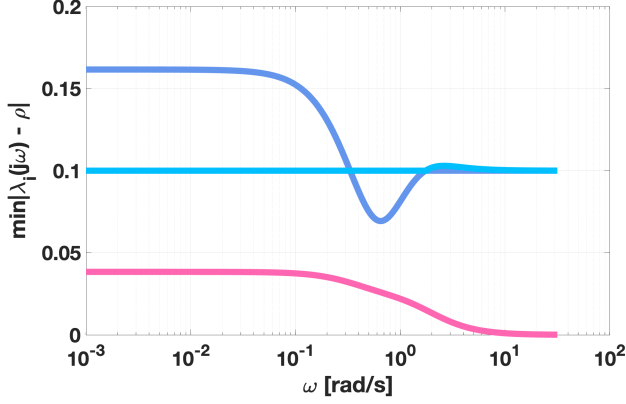


Figure 3.1: Visualisations of the stability conditions (3.7) in Theorem 5 and (3.9) in Theorem 6, for the systems in Example 1. The two blue lines are $\min_{p \leq -(2\mathfrak{D})-1} |\lambda_i(j\omega) - \rho|$ for $i = 1, 2$, with $\mathfrak{D} = 5$. Nominal topology-independent stability requires that they are strictly above zero, while robust topology-independent stability requires that they are strictly above the bound $\mathcal{K}(V(j\omega))\Delta_H^{max}(j\omega)$ shown in magenta.

degree 5 that is not stable: take, for instance, the network with 5 nodes and 9 arcs whose topology is described by the generalised Laplacian matrix

$$L = \begin{bmatrix} 3 & -1 & 0 & -1 & -1 \\ -1 & 3 & -1 & 0 & -1 \\ 0 & -1 & 3 & -1 & -1 \\ -1 & 0 & -1 & 3 & -1 \\ -1 & -1 & -1 & -1 & 5 \end{bmatrix}.$$

The overall interconnected system is unstable because it has positive-real-part poles.

We now propose examples to showcase the stability of interconnected systems with homogeneous uncertainty (Corollary 2)

Example 3. In an arbitrary network, assume that all node and arc dynamics are nominally as in Example 1 and are affected by suitably bounded, but unknown, homogeneous uncertainties. The bound $\mathcal{K}(V(j\omega))\Delta_H^{max}(\omega)$ is reported in magenta in Figure 3.1, which shows that the robust stability condition (3.9) in Theorem 6 is satisfied for all networks with maximum connectivity degree $\mathfrak{D} = 5$. The condition is satisfied up to $\mathfrak{D} = 8$, but violated for $\mathfrak{D} \geq 9$.

For the given transfer functions, we have the bounds

$$\mathcal{C} = 0.0616 \text{ and } \mathcal{M} = 0.0384.$$

Therefore, by Corollary 2, topology-independent stability is guaranteed for all networks where the connectivity degree is $\mathfrak{d}_i < 5$ for each $i \in \{1, \dots, N\}$. This shows that the condition in Corollary 2 is more conservative.

In the absence of uncertainties, $\mathcal{M} = 0$, Corollary 2 allows for a maximum connectivity degree of 8: each node could be connected to 8 other nodes and the network would remain stable. This is conservative, because the condition in Theorem 5 is satisfied for values of $\mathfrak{D}$ up to 17.

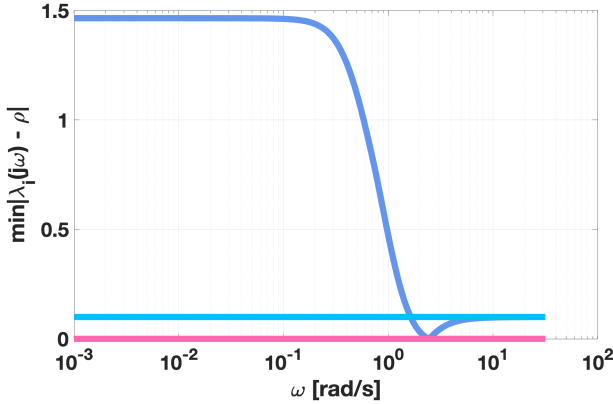


Figure 3.2: Visualisation of the stability condition (3.7) in Theorem 5, for the system in Example 2. The two blue lines are $\min_{\rho \leq -(2\mathfrak{D})-1} |\lambda_i(j\omega) - \rho|$ for $i = 1, 2$, with $\mathfrak{D} = 5$. The condition for nominal topology-independent stability is that they are strictly above zero, which is violated.

3.2.1. α -CONVERGENCE OF A SUSPENSION BRIDGE

For a suspension bridge, it is important not only to ensure stability within suitable uncertainty bounds, but also to guarantee a short enough settling time, so as to prevent long-lasting oscillations.

A suspension bridge can be modelled as a network of interconnected systems (see Figure 3.3): the nodes correspond to the cables that hold the bridge road and the arcs correspond to the discretisation of the bridge road that connects the cables. The resulting graph is known as a *ladder graph*, for which the maximum connectivity degree is always $\mathfrak{D} = 3$.

It is worth stressing that the results in this section hold for *any* type of graph with $\mathfrak{D} = 3$, regardless of its topology (information about the topology is not needed).

Figure 3.4 shows an example of a ladder graph with 8 nodes and 10 arcs, having incidence

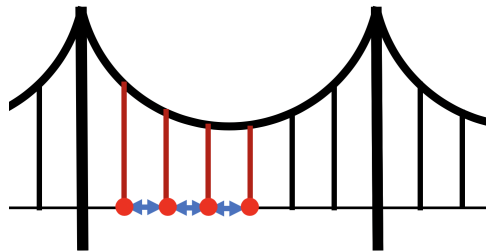


Figure 3.3: Suspension bridge (side view). The vertical cables are the nodes (red) and the road discretisation segments are the arcs (blue). The side view shows only one side of the graph. The upper view is similar to the complete graph in Figure 3.4.

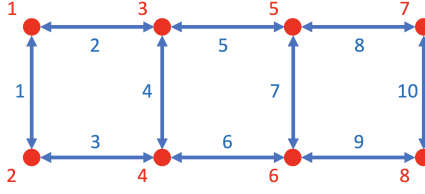


Figure 3.4: Example of ladder graph, corresponding to the incidence matrix X in (3.21). The red dots are nodes and the blue arrows are arcs. The arrows are double headed to denote bidirectional interactions.

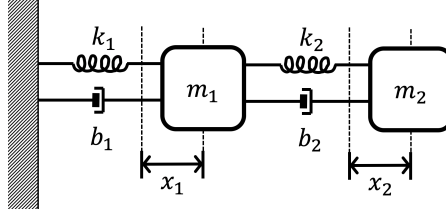


Figure 3.5: Simplified dynamic system at the nodes. The cable mass is m_1 and the road mass is m_2 ; k_i denote spring coefficients and b_i denote damping coefficients. The system is modelled as a double oscillator.

matrix

$$X = \begin{pmatrix} -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & -1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix}. \quad (3.21)$$

Node and arc dynamics. Assume that an affine transformation was used to remove gravitational effects. Then each node can be modelled as a two mass-spring-damper system (see Figure 3.5). The state vector is $x = [x_1, x_2, x_3, x_4]^\top$, where x_1 (resp. x_2) corresponds to the displacement of mass m_1 (resp. m_2) from its equilibrium location, while x_3 (resp. x_4) represents the velocity of mass m_1 (resp. m_2). The nodes have a single input u_1 , which is a force acting on m_2 , and the outputs are the state variables x_2 and x_4 . The system matrices are therefore:

$$A_F = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -\frac{(k_1+k_2)}{m_1} & \frac{k_2}{m_1} & -\frac{(b_1+b_2)}{m_1} & \frac{b_2}{m_1} \\ \frac{k_2}{m_2} & -\frac{k_2}{m_2} & \frac{b_2}{m_2} & -\frac{b_2}{m_2} \end{pmatrix} \quad B_F = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \frac{1}{m_2} \end{pmatrix} \quad (3.22)$$

$$C_F = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad D_F = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

The arcs can be represented by a mass-spring-damper, with only one mass (M) and two identical dampers (B) and springs (K) at each side. The state variables are the position and velocity of

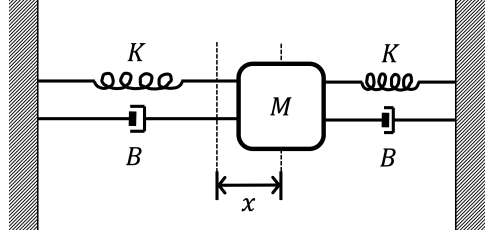


Figure 3.6: Simplified dynamic system at the arcs, corresponding to a discretisation of the road. The spring coefficient K and the damping coefficient B account for a rotational spring and damper.

3

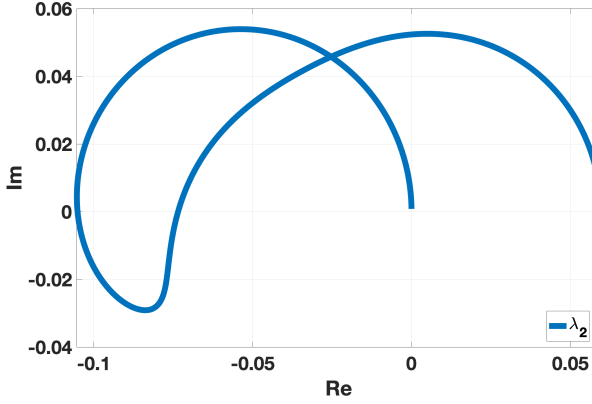


Figure 3.7: Nyquist plot of the non-zero eigenvalue $\lambda_2(j\omega - \alpha)$ of $H(j\omega - \alpha)$. Since it is far from the point $-1/(2\mathfrak{D}) = -1/6$, condition (3.19) is satisfied.

the mass M , while the inputs are the difference between the position and velocity at each side (see Figure 3.6). The resulting system matrices are

$$\begin{aligned} A_G &= \begin{pmatrix} 0 & 1 \\ -\frac{2K}{M} & -\frac{2B}{M} \end{pmatrix} & B_G &= \begin{pmatrix} 0 & 0 \\ -\frac{K}{M} & -\frac{B}{M} \end{pmatrix} \\ C_G &= (-2K \quad -2B) & D_G &= (-K \quad -B) \end{aligned} \quad (3.23)$$

Topology-independent robust α -convergence. We wish to determine the maximum uncertainty magnitude for which the overall bridge system is robustly α -convergent, with $\alpha = 0.4$, regardless of the network size. Using a second order approximation, this means that the settling time is at most $T \approx 4/\alpha = 10$ seconds. To assess topology-independent robust α -convergence, we can apply Theorem 9 with $\rho = 2$ and $\mathfrak{D} = 6$.

For the simulation results, we use the parameter values: $k_1 = 200$, $m_1 = 800$, $b_1 = 400$, $k_2 = 200$, $m_2 = 1000$, $b_2 = 800$, $K = 800$, $M = 200$, $B = 800$.

First, we show that condition (3.19) is satisfied. In this case there are only two eigenvalues λ_i , $i = 1, 2$: λ_1 is zero, because H is a rank-one matrix, thus it satisfies the inequality; Figure 3.7 shows the Nyquist plot of λ_2 , which is far from the point $-1/(2\mathfrak{D}) = -1/6$. Hence, condition (3.19) is satisfied.

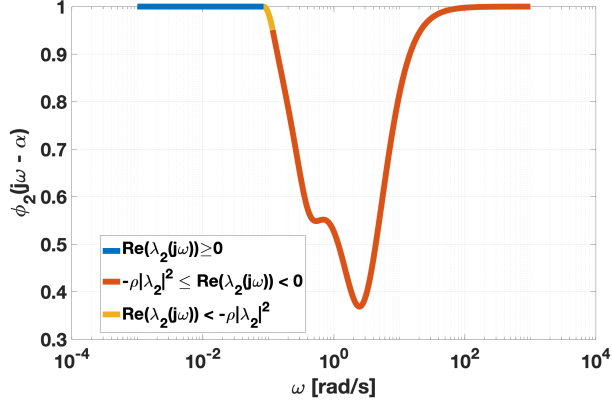


Figure 3.8: Plot of $\phi_2(j\omega - \alpha)$; colour denotes the different cases in (3.18).

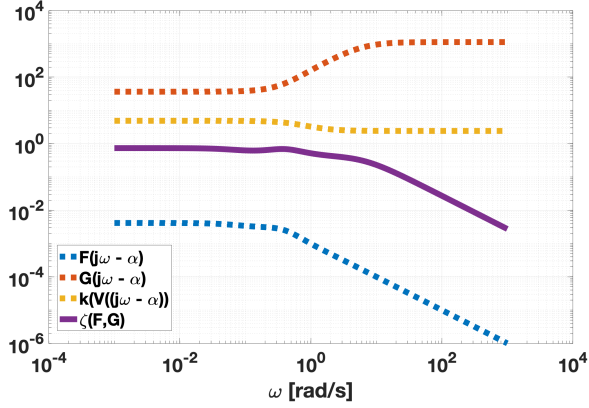


Figure 3.9: Frequency response of $\mathcal{K}(V(j\omega - \alpha))$, $\|F(j\omega - \alpha)\|$, $\|G(j\omega - \alpha)\|$, and $\hat{\zeta}(F, G)$.

Then, we can determine the upper bound for the uncertainty K that guarantees robust α -convergence for all topologies with $\mathfrak{D} = 3$. Rearranging inequality (3.20) yields

$$K(j\omega - \alpha) < \theta(j\omega - \alpha) \doteq \frac{\min_{i \in \{1,2\}} \{\phi_i(j\omega - \alpha)\}}{2\mathfrak{D}\hat{\zeta}(F, G)}. \quad (3.24)$$

Let us analyse function $\theta(j\omega - \alpha)$ in (3.24).

On the numerator is $\min\{\phi_1(j\omega - \alpha), \phi_2(j\omega - \alpha)\}$. Since $\lambda_1 = 0$, $\phi_1 = 1$. Regarding λ_2 , for different values of ω it satisfies all the cases in (3.18): Figure 3.8 shows the value of $\phi_2(j\omega - \alpha)$, indicating the frequency intervals for each of the cases. From Figure 3.8, it can be seen that in this case we have $\min\{\phi_1(j\omega - \alpha), \phi_2(j\omega - \alpha)\} = \phi_2(j\omega - \alpha)$.

The denominator consists of the product $2\mathfrak{D}\mathcal{K}(V(j\omega - \alpha))\|F(j\omega - \alpha)\|\|G(j\omega - \alpha)\|$, where $\mathfrak{D} = 3$ for a ladder graph. Figure 3.9 shows $\mathcal{K}(V(j\omega - \alpha))$, $\|F(j\omega - \alpha)\|$, $\|G(j\omega - \alpha)\|$, and the product $\hat{\zeta}(F, G)$ as a function of ω .

As expected, since the input-output matrix $D_G \neq 0$, $\|G\|$ is non-zero at high frequencies. This

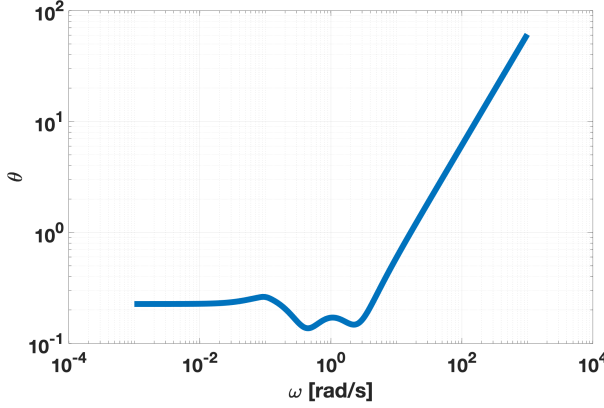


Figure 3.10: Largest admissible value for the uncertainty $K(j\omega - \alpha)$ that guarantees robust topology-independent stability, according to equation (3.24).

is reflected also in the frequency response of $\mathcal{K}(V)$. On the other hand, $F(s)$ has a low-pass-filter behaviour, hence $\|F\|$ and $\hat{\zeta}(F, G)$ tend to zero for high frequencies.

Finally, taking the ratio between ϕ_2 and $2\mathfrak{D}\hat{\zeta}(F, G)$ gives the upper bound in (3.24), whose plot is shown in Figure 3.10.

According to Figure 3.10, the minimum value of the upper bound for $K(j\omega - \alpha)$ is around 0.137, when $\omega = 0.426$ rad/s. This means that, even if the uncertainties were about 6%, i.e. $K_F(j\omega - \alpha) \leq 0.06$ and $K_G(j\omega - \alpha) \leq 0.06$, it would be $K(j\omega - \alpha) \leq 0.1236$; then, $K(j\omega - \alpha)$ would still satisfy the inequality (3.24) at all frequencies. Hence, the system would be stable and α -convergent with $\alpha = 0.4$. Note that, since $\phi_2 \leq 1$ and $\hat{\zeta}(F, G)$ has a low-pass-filter response, the uncertainties can be very large at high frequencies without compromising stability and α -convergence.

Given that the expression of K couples the uncertainties affecting the node transfer function matrix, F , and the arc transfer function matrix, G , it is not possible to immediately deduce separate upper bounds for the individual uncertainties. However, if for instance $K_G \approx 0$, then $K \approx K_F$, which yields an upper bound for the node uncertainties. An analogous result is obtained if $K_F \approx 0$.

3.3. PROOFS

3.3.1. LEMMA 3

The following lemma will be used in the proof of Theorem 7.

Lemma 3. *Let a, b and u be three complex numbers. If the inequalities $|a - b| \leq \varphi$ and $|b - (-u)| > \varphi$ are satisfied for some real $\varphi > 0$, then $a \neq -u$.*

Proof. The inequality $|a - b| \leq \varphi$ can be rewritten as $|(a+1) - (b+1)| \leq \varphi$. Hence, the reverse triangle inequality $||x| - |y|| \leq |x - y|$ ensures

$$||a+1| - |b+1|| \leq \varphi. \quad (3.25)$$

Then, we can distinguish two cases. If $|a+1| \geq |b+1|$, the result immediately follows because $|a+1| \geq |b+1| > \varphi > 0$. If $|a+1| < |b+1|$, then (3.25) becomes $|b+1| - |a+1| \leq \varphi$, and $|a+1| \geq |b+1| - \varphi > 0$ since $|b+1| > \varphi$. $\square$

3.3.2. PROOF OF THEOREM 3

Proof. Consider a generic element in the family of networked systems $\mathcal{N}$, let vectors $Y_i(s)$ and $U_h(s)$ represent the output of the i th node and h th arc respectively. The dynamics of node i is

$$Y_i(s) = [F(s) + \Delta_{F_i}(s)] \sum_{h=1}^M [B]_{ih} U_h(s), \quad i \in \mathcal{N},$$

while the dynamics of arc $h = (i, j) \in \mathcal{A}$ is

$$U_h(s) = [G(s) + \Delta_{G_h}(s)][Y_i(s) - Y_j(s)], \quad h \in \mathcal{A}.$$

We can stack the output and input vectors as $Y(s) = [Y_1(s)^\top, \dots, Y_N(s)^\top]^\top$ and $U(s) = [U_1(s)^\top, \dots, U_M(s)^\top]^\top$ and write the complete system dynamics as

$$Y(s) = [(I_N \otimes F(s)) + \mathcal{D}_F(s)](X \otimes I_n)U(s), \quad (3.26)$$

$$U(s) = -[(I_M \otimes G(s)) + \mathcal{D}_G(s)](X^\top \otimes I_r)Y(s), \quad (3.27)$$

where I_k denotes the identity matrix of size k , while $\mathcal{D}_F = \text{diag}(\Delta_{F_i})_{i=1}^N$ and $\mathcal{D}_G = \text{diag}(\Delta_{G_h})_{h=1}^M$. It follows that, the characteristic equation of the complete network is

$$\begin{aligned} p(s) &= \det \left(I_{Nr} + [(I_N \otimes F(s)) + \mathcal{D}_F(s)](X \otimes I_n) \right. \\ &\quad \left. [(I_M \otimes G(s)) + \mathcal{D}_G(s)](X^\top \otimes I_r) \right) \\ &= \det \left(I_{Nr} + L \otimes H(s) + \mathcal{D}(s) \right), \end{aligned} \quad (3.28)$$

where $L = XX^\top$, $H(s) = F(s)G(s)$ and $\mathcal{D} = (X \otimes F)\mathcal{D}_G(X^\top \otimes I_r) + \mathcal{D}_F(X \otimes I_n)\mathcal{D}_G(X^\top \otimes I_r) + \mathcal{D}_F(XX^\top \otimes G)$. Equations (3.3) and (3.5) follow from simplifying Equations (3.1) and (3.2) $\square$

3.3.3. PROOF OF THEOREM 4

Proof. Since the complex function $\lambda_i(s)$ is an eigenvalue of $H(s)$, it must satisfy the characteristic equation

$$\det(\lambda_i(s)I - H(s)) = 0. \quad (3.29)$$

The transfer-function matrix $H(s)$ can be written as

$$H(s) = \frac{1}{d(s)} R(s), \quad (3.30)$$

where $d(s)$ is the pole polynomial and $R(s)$ is a matrix with polynomial entries. For any $s \in \mathbb{C}$ such that $d(s) \neq 0$, replacing (3.30) into (3.29) gives

$$\det(\lambda_i(s)d(s)I - R(s)) = 0. \quad (3.31)$$

By contradiction, assume that $\bar{p}$ is a pole of the complex function $\lambda_i(s)$ but not of the transfer-function matrix $H(s)$. Then, by continuity, $\lim_{s \rightarrow \bar{p}} \{d(s)\} \rightarrow d(\bar{p}) \neq 0$ and $\lim_{s \rightarrow \bar{p}} \{R(s)\} \rightarrow R(\bar{p})$, which is a matrix with finite entries. At the same time, $\lim_{s \rightarrow \bar{p}} \{\lambda_i(s)\} \rightarrow \infty$. This in turn implies that $\lim_{s \rightarrow \bar{p}} \{\det(\lambda_i(s)d(s)I - R(s))\} \rightarrow \infty$, which contradicts equation (3.31). Hence, it must be $\lim_{s \rightarrow \bar{p}} \{d(s)\} \rightarrow d(\bar{p}) = 0$, namely, $\bar{p}$ must be a root of $d(s)$, hence a pole of $H(s)$. $\square$

3.3.4. PROOF OF LEMMA 2

Proof. First, we prove stability in the nominal case. The Nyquist stability criterion requires $Z = N + P$, where P is the number of unstable poles of the complex function $\bar{h}(s)$, Z is the number of unstable zeros of the complex function $1 + \mu \bar{h}(s)$ and N is the number of times the Nyquist diagram encircles the point $-1/\mu$ clockwise. Because the nominal open loop complex function is assumed to be stable ($P = 0$), the closed loop system is stable ($Z = 0$) if and only if $N = 0$, that is, if and only if the Nyquist diagram does not encircle the point $-1/\mu$.

Recall that $\sigma(L_{\mathfrak{D}}) \subset \mathcal{S} = \{z \in \mathbb{R} : 0 \leq z \leq 2\mathfrak{D}\}$. Then, the Nyquist diagram $h(j\omega)$ cannot encircle the point $-1/\mu$ if, for all $\omega \in \mathbb{R}^+$, it has an empty intersection with the set

$$\zeta(\mathcal{S}) = \left\{ z \in \mathbb{R} : z \leq \frac{-1}{2\mathfrak{D}} \right\}. \quad (3.32)$$

To enforce stability robustly in the presence of a bounded uncertainty, we resort to the Zero Exclusion principle, which requires $|\mu \bar{h}(j\omega) + 1| > \mu \delta_h^{max}(\omega)$, or equivalently $|\bar{h}(j\omega) + \frac{1}{\mu}| > \delta_h^{max}(\omega)$, for all possible choices of μ . Namely, it must be $|\bar{h}(j\omega) - \rho| > \delta_h^{max}(\omega)$ for any ρ in the set $\zeta(\mathcal{S})$. This condition can be interpreted geometrically with the help of Figure 3.11: the distance of $\bar{h}(j\omega)$ from the set $\zeta(\mathcal{S})$ must be larger than the uncertainty, $\min_{\rho \leq -(2\mathfrak{D})^{-1}} |\bar{h}(j\omega) - \rho| > |\delta_h(j\omega)|$, for any possible uncertainty $|\delta_h(j\omega)| \leq \delta_h^{max}(\omega)$, thus leading to (3.6):

$$\min_{\rho \leq -(2\mathfrak{D})^{-1}} |\bar{h}(j\omega) - \rho| > \delta_h^{max}(j\omega). \quad (3.33)$$

□

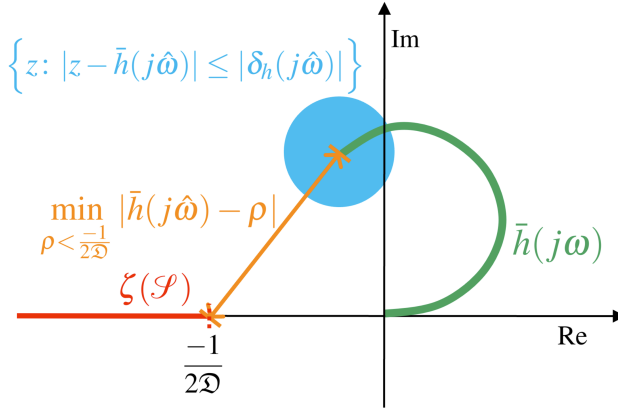


Figure 3.11: **Visualisation of the robust stability condition in Lemma 2.** Nyquist diagram of the nominal transfer function $\bar{h}(j\omega)$ (green), uncertainty disk centred at $\bar{h}(j\hat{\omega})$ representing $\delta_h(j\hat{\omega})$ at the frequency $\hat{\omega}$ (cyan), set $\zeta(\mathcal{S})$ (red) and distance of $\bar{h}(j\hat{\omega})$ from the set $\zeta(\mathcal{S})$ (orange). For all $\omega \in \mathbb{R}^+$, the distance of $\bar{h}(j\omega)$ from the set $\zeta(\mathcal{S})$ needs to be larger than the uncertainty radius $\delta_h^{max}(\omega) \geq |\delta_h(j\omega)|$.

3.3.5. PROOF OF THEOREM 5

Proof. Consider a generic nominal networked system from the family of networked systems $\mathcal{N}$. By Theorem 3 the characteristic polynomial is given by Equation (3.5).

$$p(s) = \det(I_{Nr} + L \otimes H(s)).$$

The system will be stable if and only if all the roots of the characteristic polynomial have negative real part.

The generalised Laplacian can be diagonalised as $W^{-1}LW = \text{diag}\{\gamma_k\}_{k=1}^N$ and the transfer function matrix can be triangularised as $V(s)^{-1}H(s)V(s) = \Lambda(s)$, where $\Lambda(s)$ is a triangular matrix carrying on the diagonal the eigenvalues $\{\lambda_k(s)\}_{k=1}^r$ of $H(s)$. Then, we can pre-multiply the characteristic polynomial $p(s) = \det(I_{Nr} + L \otimes H(s))$ by $\det((W \otimes V(s))^{-1})$ and post-multiply it by $\det(W \otimes V(s))$ to get

$$p(s) = \prod_{i=1}^r \prod_{k=1}^N (1 + \gamma_k \lambda_i(s)). \quad (3.34)$$

The characteristic polynomial (3.34) is stable (all its roots have negative real part) if and only if each polynomial $1 + \gamma_k \lambda_i(s)$ is stable (all their roots have negative real part).

Assumption 3 and Theorem 4 guarantee that the complex functions $\lambda_i(s)$ are stable (all the poles have negative real part); see Remark 1. Therefore, we can apply Lemma 2 where $\tilde{h}(s) = \lambda_i(s)$ and $\delta_h^{max} \equiv 0$ to conclude that the polynomials $1 + \gamma_k \lambda_i(s)$ are stable (all their roots have negative real part) if, for all i and all $\omega \in \mathbb{R}^+$, (3.7) holds. $\square$

3.3.6. PROOF OF THEOREM 6

Proof. Consider a generic networked system from the family of nominally stable networked systems $\mathcal{N}$. By Theorem 3 the characteristic polynomial when all the uncertainties are homogeneous is given by Equation (3.3),

$$p(s) = \det(I_{Nr} + L \otimes (H(s) + \Delta_H(s))),$$

where

$$\Delta_H(s) = F(s)\Delta_G(s) + \Delta_F(s)G(s) + \Delta_F(s)\Delta_G(s).$$

The system will be stable if and only if all the roots of the characteristic polynomial have negative real part.

Recalling that $W^{-1}LW$ is a diagonal matrix carrying the eigenvalues $\{\gamma_k\}_{k=1}^N$ in the diagonal, pre-multiplying the characteristic polynomial $p(s) = \det(I_{Nr} + L \otimes (H(s) + \Delta_H(s)))$ by $\det((W \otimes I_r)^{-1})$ and post-multiplying it by $\det(W \otimes I_r)$ yields

$$p(s) = \prod_{k=1}^N \det(I_r + \gamma_k (H(s) + \Delta_H(s))).$$

Let $\{\beta_q(s)\}_{q=1}^r$ denote the eigenvalues of $H(s) + \Delta_H(s)$. Then the characteristic polynomial can be written as

$$p(s) = \prod_{q=1}^r \prod_{k=1}^N (1 + \gamma_k \beta_q(s)), \quad (3.35)$$

which is stable (all its roots have negative real part) if and only if each polynomial $1 + \gamma_k \beta_q(s)$ is stable (all their roots have negative real part).

So now we need to show that under the theorem assumptions and hypothesis all the roots of the polynomials $1 + \gamma_k \beta_q(s)$ have negative real part.

In view of Assumption 4, we can apply the Bauer-Fike theorem; hence, there exists an index $i \in \{1, \dots, r\}$ such that $\beta_q(s) = \lambda_i(s) + \delta_{\lambda_i}(s)$, where

$$\begin{aligned} |\beta_q(j\omega) - \lambda_i(j\omega)| &= |\delta_{\lambda_i}(j\omega)| \leq \mathcal{K}(V(j\omega)) \|\Delta_H(j\omega)\| \\ &\leq \mathcal{K}(V(j\omega)) \Delta_H^{max}(\omega). \end{aligned}$$

We now can apply Lemma 2, with $h(s) = \beta_q(s)$, $\bar{h}(s) = \lambda_i(s)$, $\delta_h(s) = \delta_{\lambda_i}(s)$ and $\delta_h^{max}(\omega) = \mathcal{K}(V(j\omega)) \Delta_H^{max}(\omega)$, for all $i \in \{1, \dots, r\}$. Lemma 2 can be applied because Assumption 3 and Theorem 4 guarantee that the nominal complex functions $\lambda_i(s)$ are stable (Remark 1), and γ_k are the eigenvalues of the generalised Laplacian matrix of any network with maximum connectivity degree $\mathfrak{D}$.

As a result of applying Lemma 2 to all $i \in \{1, \dots, r\}$, we can say that all the roots of the polynomials $1 + \gamma_k \beta_q(s)$ have negative real part if inequality (3.9) holds. $\square$

3.3.7. PROOF OF COROLLARY 1

Proof. Using the same decomposition as in the proof of Theorem 6, the eigenvalues of matrix $H(s) + \Delta_H(s)$ can be written as $\beta_q(j\omega) = \lambda_i(j\omega) + \delta_{\lambda_i}(j\omega)$ for some $i \in \{1, \dots, r\}$, where

$$|\delta_{\lambda_i}(j\omega)| < \mathcal{K}(V(j\omega)) \|\Delta_H(j\omega)\| \leq \xi(\omega) \Delta_H^{max}(\omega).$$

Hence

$$|\beta_q(j\omega)| \leq |\lambda_i(j\omega)| + |\delta_{\lambda_i}(j\omega)| \leq \mathcal{C} + \mathcal{M}, \quad \forall \omega \in \mathbb{R}^+.$$

If

$$\mathcal{C} + \mathcal{M} < \min_{z \in \zeta(\mathcal{S})} |z| = \frac{1}{2\mathfrak{D}}, \quad (3.36)$$

where $\mathcal{S} = \{z \in \mathbb{R} : 0 \leq z \leq 2\mathfrak{D}\}$, then the distance between $\lambda_i(j\omega) + \delta_{\lambda_i}(j\omega)$ and the set $\zeta(\mathcal{S})$ is larger than zero for all i and for all possible bounded realisations of the uncertainty, which ensures topology-independent robust stability of all networks with maximum connectivity degree $\mathfrak{D}$. $\square$

3.3.8. PROOF OF COROLLARY 2

Proof. With $\mathcal{T}$ as defined in (3.13), the inequality (3.12) becomes $\mathfrak{D} < \mathcal{T}$ and is of course satisfied if and only if $\mathfrak{d}_i < \mathcal{T}$ for all $i = 1, \dots, N$, since $\mathfrak{D} = \max_{i=1, \dots, N} \{\mathfrak{d}_i\}$. $\square$

3.3.9. PROOF OF THEOREM 7

Proof. Consider a generic networked system from the family of nominally stable networked systems $\mathcal{N}$. By Theorem 3 the characteristic polynomial when all the uncertainties are homogeneous is given by Equation (3.1),

$$p(s) = \det(I_{Nr} + L \otimes H(s) + \mathcal{D}(s)), \quad (3.37)$$

where

$$\mathcal{D} = (X \otimes F) \mathcal{D}_G(X^\top \otimes I_r) + \mathcal{D}_F(X \otimes I_n) \mathcal{D}_G(X^\top \otimes I_r) + \mathcal{D}_F(X X^\top \otimes G). \quad (3.38)$$

The system will be stable if and only if all the roots of the characteristic polynomial have negative real part.

Denoting by $\beta_q(s)$, $q \in \{1, \dots, rN\}$, the eigenvalues of the matrix $[(L \otimes H(s)) + \mathcal{D}(s)]$, the characteristic polynomial (3.1) can be rewritten as

$$p(s) = \prod_{q=1}^{rN} (1 + \beta_q(s)).$$

By the zero-exclusion theorem [19], since the nominal interconnected system is stable by assumption, robust stability of the uncertain system is guaranteed if, for all possible $\mathcal{D}_F(s)$ and $\mathcal{D}_G(s)$ within the bounds, $p(j\omega) \neq 0$ for all $\omega \in \mathbb{R}^+$, which is equivalent to

$$\beta_q(j\omega) \neq -1, \quad \forall q \in \{1, \dots, rN\} \quad \text{and} \quad \forall \omega \in \mathbb{R}^+. \quad (3.39)$$

The eigenvalues of $L \otimes H(s)$ are the products of the eigenvalues of L and of $H(s)$, $\{\gamma_k \lambda_i(s)\}$, and its diagonalization matrix is $W \otimes V(s)$, where W and $V(s)$ are the diagonalization matrices for L and $H(s)$ respectively. Hence, in view of the *Bauer-Fike Theorem*, for every $q = 1, \dots, rN$ there is a pair of indices $(k, i) \in \{1, \dots, N\} \times \{1, \dots, r\}$ such that

$$|\beta_q(j\omega) - \gamma_k \lambda_i(j\omega)| \leq \mathcal{K}(W \otimes V(j\omega)) \|\mathcal{D}(j\omega)\|. \quad (3.40)$$

Using the properties of the 2-norm and of the Kronecker product we have that

$$\begin{aligned} \mathcal{K}(W \otimes V(j\omega)) &= \|(W \otimes V(j\omega))^{-1}\| \|W \otimes V(j\omega)\| \\ &= \|W^{-1} \otimes V^{-1}(j\omega)\| \|W \otimes V(j\omega)\| \\ &= \|W^{-1}\| \|V^{-1}(j\omega)\| \|W\| \|V(j\omega)\| \\ &= \mathcal{K}(W) \mathcal{K}(V(j\omega)) = \mathcal{K}(V(j\omega)). \end{aligned}$$

Also,

$$\begin{aligned} \|\mathcal{D}(j\omega)\| &\leq \|X\| \|X^\top\| \left(\|F(j\omega)\| \|\mathcal{D}_G(j\omega)\| + \|G(j\omega)\| \|\mathcal{D}_F(j\omega)\| + \|\mathcal{D}_F(j\omega)\| \|\mathcal{D}_G(j\omega)\| \right) \\ &\leq \|X\| \|X^\top\| \|F(j\omega)\| \|G(j\omega)\| K(j\omega), \end{aligned}$$

with $K(j\omega)$ defined in (3.16):

$$K(j\omega) = K_F(j\omega) + K_G(j\omega) + K_F(j\omega) K_G(j\omega).$$

Recall from equation (2.3) that $\|X\| \|X^\top\| \leq 2\mathfrak{D}$ to get

$$\mathcal{K}(W \otimes V(j\omega)) \|\mathcal{D}(j\omega)\| \leq 2\mathfrak{D} \zeta(F, G) K(j\omega), \quad (3.41)$$

with $\zeta(F, G)$ defined in (3.15)

$$\zeta(F, G) = \mathcal{K}(V(j\omega)) \|F(j\omega)\| \|G(j\omega)\|. \quad (3.42)$$

Inequalities (3.40) and (3.41) yield

$$|\beta_q(j\omega) - \gamma_k \lambda_i(j\omega)| \leq 2\mathfrak{D} \zeta(F, G) K(j\omega). \quad (3.43)$$

Now we can apply Lemma 3 to (3.43) and (3.14), setting $\varphi = 2\mathfrak{D} \zeta(F, G) K(j\omega)$, $a = \beta_q(j\omega)$, $b = \gamma_k \lambda_i(j\omega)$ and $u = 1$. We get that $\beta_q(j\omega) \neq -1$, hence condition (3.39) is satisfied and the network system is robustly stable. $\square$

3.3.10. PROOF OF THEOREM 8

Proof. To show that inequality (3.17) implies inequality (3.14) for all $k \in \{1, \dots, N\}$ and $i \in \{1, \dots, r\}$, write the complex number $\lambda_i(j\omega)$ as $\lambda_i = \alpha_i + j\beta_i$, where the argument is omitted for clarity. Now define the convex function

$$D(\gamma_k) = |\gamma_k \lambda_i + 1| = \sqrt{(\gamma_k \alpha_i + 1)^2 + \gamma_k^2 \beta_i^2}.$$

Taking into account that $\gamma_k \in [0, \rho]$, the minimum of $D(\gamma_k)$ is obtained for $\gamma_k = \gamma_k^*$ given by

$$\gamma_k^* = \begin{cases} 0 & \text{if } 0 \leq \alpha_i \\ \frac{-\alpha_i}{|\lambda_i|^2} & \text{if } -\rho|\lambda_i|^2 \leq \alpha_i < 0 \\ \rho & \text{if } \alpha_i < -\rho|\lambda_i|^2 \end{cases}$$

In other words, $\gamma_k^* = \operatorname{argmin}\{D(\gamma_k) \text{ s.t. } \gamma_k \in [0, \rho]\}$. Hence, the minimum value of $D(\gamma_k)$ depends on $\lambda_i(j\omega)$ as follows

1. if $0 \leq \alpha_i$, then $D(\gamma_k^*) = 1$,
2. if $-\rho|\lambda_i|^2 \leq \alpha_i < 0$, then $D(\gamma_k^*) = \frac{|\operatorname{Im}(\lambda_i(j\omega))|}{|\lambda_i(j\omega)|}$,
3. if $\alpha_i < -\rho|\lambda_i|^2$, then $D(\gamma_k^*) = |\rho \lambda_i(j\omega) + 1|$.

Since $|\gamma_k \lambda_i(j\omega) + 1| \geq D(\gamma_k^*)$, each case gives a different lower bound for $|\gamma_k \lambda_i(j\omega) + 1|$. Thus, by construction, $\phi_i(j\omega)$ satisfies $|\gamma_k \lambda_i(j\omega) + 1| \geq \phi_i(j\omega)$.

Taking the minimum over all $i \in \{1, \dots, r\}$ makes inequality (3.17) imply inequality (3.14) for all $k \in \{1, \dots, N\}$ and $i \in \{1, \dots, r\}$. $\square$

3.3.11. PROOF OF THEOREM 9

Proof. Take the characteristic polynomial $p(s)$ of the complete network, given in (3.1), and define $\hat{p}(s)$ as the shifted polynomial $\hat{p}(s) = p(s - \alpha)$.

If $\hat{p}(s)$ is stable, then $p(s)$ is α -convergent.

By Theorem 5, inequality (3.19) guarantees that the nominal networked system associated with $\hat{p}(s)$ is stable, thus Theorem 8 can be applied to check that, if inequality (3.20) is satisfied, then $\hat{p}(s)$ is stable, hence $p(s)$ is α -convergent. $\square$

3.4. CONCLUSIONS

We have investigated the stability of homogeneous and heterogeneous dynamic networks where both the nodes and the arcs have MIMO dynamics, described by the uncertain transfer-function matrices $F(s) + \Delta_{F_i}(s)$ and $G(s) + \Delta_{G_h}(s)$ respectively. The transfer-function matrices $F(s)$ and $G(s)$ are assumed to be stable and the uncertainties bounded. In the nominal case, we have provided a stability condition that is topology-independent and exclusively relies on the knowledge of the maximum connectivity degree of the network. The condition constrains the Nyquist diagram of the poles of the transfer-function matrix $H(s) = F(s)G(s)$. In the presence of homogeneous and heterogeneous uncertainties, the topology-independent condition for robust stability depends on the magnitude of the uncertainty and relies on the bound for the eigenvalues of uncertain matrices given by the Bauer-Fike theorem. An advantage of the obtained conditions, which guarantee robust stability regardless of the network size and topology, is that they can be checked locally to ensure stability of the network also when nodes and arcs are added or removed online. In addition to stability conditions, a topology-independent condition to robustly guarantee fast enough convergence was provided and applied to a suspension bridge example.

4

TIME DOMAIN ANALYSIS

The last enemy that shall be destroyed is death.

Harry Potter and the Deathly Hallows

We give a sufficient and a necessary condition for the topology-independent robust stability of networked systems formed by uncertain MIMO systems. Both conditions involve constants associated with the nominal node dynamics and arc interconnection matrices, the uncertainty bounds, and the maximum connectivity degree of the network; they are scalable (they can be checked locally), independent of the network topology and even of the number of nodes and arcs, and hold for networks of heterogeneous MIMO subsystems and interconnection matrices, with heterogeneous uncertainties. The dual cases of 1-norm and ∞ -norm bounds are considered. In both cases, if the systems at the nodes are diagonal, we get a necessary and sufficient condition. We apply our results to the topology-independent robust stability analysis of a case-study from cancer biology.

This chapter is based on the paper “Topology-independent robust stability conditions for uncertain MIMO networks” (2021) [61] by Carlos Andrés Devia and Giulia Giordano.

All the results in this chapter are related to the concept of **Family of networked systems**, whose definition is recalled below. All the networks in this chapter are assumed to have dynamic (possibly different) nodes and static (possibly different) arcs of the appropriate size. Given that in this chapter all the systems are represented in a state-space form, the main tools used to prove the robust stability conditions will be algebraic properties of norms and eigenvalues. To improve the readability of the chapter first the main results are stated, followed by application examples, proofs, and the conclusions.

Definition 1 (Family of networked systems). *A family $\mathcal{N}$ of networked systems is a collection of LTI systems formed by the interconnection of LTI subsystems according to a network structure. The generic system in the family has an underlying graph structure $\mathcal{G} = \{\mathcal{N}, \mathcal{A}\}$, where each node in $\mathcal{N}$, labelled with an integer number in the set $\{1, \dots, N\}$ is associated with a LTI MIMO system and each arc in $\mathcal{A}$, labelled with an integer number in the set $\{1, \dots, M\}$, is associated with a LTI MIMO system or an interconnection matrix.*

4

4.1. MAIN RESULTS

We start with some basic assumptions to formally define the type of families considered in this chapter.

Assumption 7. *For all elements in the family $\mathcal{N}$, every node in the graph $\mathcal{G}$ is associated with a LTI MIMO subsystem of the form*

$$\dot{x}^{(i)} = A_i x^{(i)} + B_i u^{(i)}, \quad y^{(i)} = C_i x^{(i)}, \quad i \in \{1, \dots, N\},$$

where the system matrices are the sum of a nominal and an uncertain part: $A_i = \bar{A}_i + \hat{A}_i$, $B_i = \bar{B}_i + \hat{B}_i$, $C_i = \bar{C}_i + \hat{C}_i$. Let p_i and q_i be the number of inputs and outputs of subsystem i , respectively. Every arc in the graph $\mathcal{G}$ is associated with an interconnection matrix G_h , of the proper size which is the sum of a nominal and an uncertain part: $G_h = \bar{G}_h + \hat{G}_h$, $h \in \{1, \dots, M\}$.

Figure 4.1 shows an example of a member of a family with the structure described in Assumption 7.

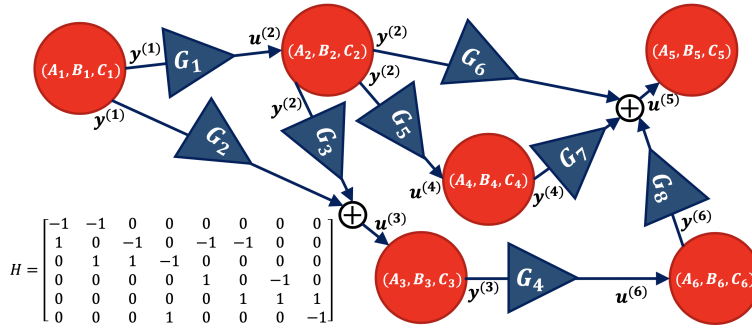


Figure 4.1: Example of an uncertain networked system in the family $\mathcal{N}$, with $N = 6$ node systems, $M = 8$ arcs and incidence matrix (in this figure denoted by H).

Theorem 10. *Under Assumption 7, a generic element in the family of networked systems $\mathcal{N}$ is modelled by the following matrix differential equation*

$$\dot{x} = [A + B\mathcal{P}G\mathcal{R}C]x \doteq \tilde{A}x, \quad (4.1)$$

where

$$\begin{aligned} x &= [x^{(1)\top} \dots x^{(N)\top}]^\top & A &= \text{diag}(A_i)_{i \in \mathcal{N}} & B &= \text{diag}(B_i)_{i \in \mathcal{N}} \\ C &= \text{diag}(C_i)_{i \in \mathcal{N}} & G &= \text{diag}(G_h)_{h \in \mathcal{A}}, \end{aligned}$$

and the matrices $\mathcal{P}$ and $\mathcal{R}$ are built as follows: first, define the matrices P and R as $P = \max\{X, 0\}$ and $R = -\min\{X, 0\}$ element-wise, where X is the incidence matrix associated with the networked system; then the block matrix in the position (i, j) of $\mathcal{P}$ is the square scaled identity matrix $P_{ij}I_{p_i}$ if $P_{ij} = 1$, while if $P_{ij} = 0$ it is a rectangular matrix of zeros of the appropriate size. For $\mathcal{R}$, the block matrix in position (i, j) is $R_{ji}I_{q_j}$ if $R_{ji} = 1$, while it is a rectangular zero matrix if $R_{ji} = 0$.

If all the subsystems have the same number of inputs and outputs, p and q , respectively, then Equation (4.1) becomes

$$\dot{x} = [A + B(P \otimes I_p)G(R^\top \otimes I_q)C]x \doteq \tilde{A}x. \quad (4.2)$$

Example 4. Consider a network composed of 3 nodes, with $p_1 = 2$, $p_2 = 3$, $p_3 = 2$ inputs and $q_1 = 1$, $q_2 = 2$, $q_3 = 1$ outputs. Let the incidence matrix be

$$X = \begin{bmatrix} -1 & 0 & 1 & 1 \\ 1 & -1 & 0 & -1 \\ 0 & 1 & -1 & 0 \end{bmatrix},$$

that is, the network has 4 arcs with matrices $G_1 \in \mathbb{R}^{3 \times 1}$, $G_2 \in \mathbb{R}^{2 \times 2}$, $G_3 \in \mathbb{R}^{2 \times 1}$, and $G_4 \in \mathbb{R}^{2 \times 2}$. Then

$$P = \begin{bmatrix} 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \quad \text{and} \quad \mathcal{P} = \left[\begin{array}{ccc|ccc|ccc} 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ \hline 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{array} \right].$$

Assumption 8. For each element in the family $\mathcal{N}$, the maximum outward (resp. inward) connectivity degree of the underlying graph is at most $\mathcal{D}^{out}$ (resp. $\mathcal{D}^{in}$).

Assumption 9. For each element in the family $\mathcal{N}$, Assumption 7 holds, and all the node systems and interconnection matrices have the following properties:

All node systems (A_i, B_i, C_i) , for $i \in \{1, \dots, N\}$, are such that $A_i = \bar{A}_i + \hat{A}_i$, $B_i = \bar{B}_i + \hat{B}_i$, $C_i = \bar{C}_i + \hat{C}_i$, where, denoting with $*$ either always 1 or always ∞ ,

- $\max_{\lambda \in \sigma(\bar{A}_i)} \{\text{Re}(\lambda)\} \leq \alpha$, for a given $\alpha < 0$;
- $\|W_i\|_* \leq \chi$ and $\|W_i^{-1}\|_* \leq \chi$, for a given $\chi \geq 1$, where W_i is some eigenmatrix that diagonalises $\bar{A}_i$;
- $\|\bar{B}_i\|_* \leq \mu_B$ and $\|\bar{C}_i\|_* \leq \mu_C$, for given $\mu_B, \mu_C > 0$;
- $\|\hat{A}_i\|_* \leq \xi_A$, $\|\hat{B}_i\|_* \leq \xi_B$ and $\|\hat{C}_i\|_* \leq \xi_C$, for given $\xi_A, \xi_B, \xi_C \geq 0$.

All interconnection matrices G_h , for $h \in \mathcal{A}$, are such that $G_h = \bar{G}_h + \hat{G}_h$, with $\|\bar{G}_h\|_* \leq \mu_G$ and $\|\hat{G}_h\|_* \leq \xi_G$, for given $\mu_G, \xi_G \geq 0$, where the subscript $*$ denotes either always 1 or always ∞ .

Assumption 9 implies $\mathcal{K}_*(W_i) = \|W_i\|_* \|W_i^{-1}\|_* \leq \chi^2$.

We are then ready to state our main results for the case of uncertain networked systems with 1-norm bounds; the proofs are given in Section 4.3.

Theorem 11 (Sufficient condition for topology-independent robust stability). *Consider the family of networked systems $\mathcal{N}$, under Assumptions 8, and 9 with 1-norm bounds. Then, all systems in $\mathcal{N}$ are stable if*

$$\alpha + \chi^2 [\xi_A + \mathcal{D}^{out}(\mu_B + \xi_B)(\mu_G + \xi_G)(\mu_C + \xi_C)] < 0. \quad (4.3)$$

Theorem 12 (Necessary condition for topology-independent robust stability). *Consider the family of networked systems $\mathcal{N}$, under Assumptions 8, and 9 with 1-norm bounds. A necessary condition for all systems in $\mathcal{N}$ to be stable is*

$$\alpha + [\xi_A + \mathcal{D}^{out}(\mu_B + \xi_B)(\mu_G + \xi_G)(\mu_C + \xi_C)] < 0. \quad (4.4)$$

By duality, our main results still hold if the 1-norm is replaced by the ∞ -norm, and $\mathcal{D}^{out}$ is replaced by $\mathcal{D}^{in}$. Since the proofs are essentially unchanged, we just report the results.

Proposition 1 (Dual of Theorem 11). *Consider the family of networked systems $\mathcal{N}$, under Assumptions 8, and 9 with ∞ -norm bounds. Then, all systems in $\mathcal{N}$ are stable if*

$$\alpha + \chi^2 [\xi_A + \mathcal{D}^{in}(\mu_B + \xi_B)(\mu_G + \xi_G)(\mu_C + \xi_C)] < 0. \quad (4.5)$$

Proposition 2 (Dual of Theorem 12). *Consider the family of networked systems $\mathcal{N}$, under Assumptions 8, and 9 with ∞ -norm bounds. Then, a necessary condition for all systems in $\mathcal{N}$ to be stable is that*

$$\alpha + [\xi_A + \mathcal{D}^{in}(\mu_B + \xi_B)(\mu_G + \xi_G)(\mu_C + \xi_C)] < 0. \quad (4.6)$$

For diagonal systems, the topology-independent robust stability condition becomes necessary and sufficient.

Proposition 3 (Diagonal systems). *Consider the family of networked systems $\mathcal{N}$, under Assumptions 8, and 9 with 1-norm (resp. ∞ -norm) bounds. Assume that, for each element of the family, all the systems at the nodes have a diagonal state matrix A_i . Then, all systems in $\mathcal{N}$ are stable if and only if inequality (4.3) (resp. (4.5)) holds.*

Our results highlight the crucial role of the condition number χ^2 for topology-independent stability: for non-diagonal systems, it leads to a gap between the sufficient and the necessary condition, thus introducing conservativeness. To have the tightest gap, we wish to compute the *minimum* value of χ^2 . Consider the nominal matrix $\bar{A}$ corresponding to the system associated with a *single node*; being diagonalisable, it has distinct eigenvectors. Then, the columns of its eigenmatrix W can be scaled independently with the positive diagonal matrix $D = \text{diag}(D_i)$ and we can find

$$(\chi^2)^{\text{opt}} = \min_{D \in \text{diag}(D_i), D_i > 0} \|WD\|_* \|D^{-1}W^{-1}\|_* \quad (4.7)$$

where the subscript $*$ denotes either always 1 or always ∞ .

This optimisation problem has a neat solution.

Proposition 4 (Minimum χ^2). *The optimal $(\chi^2)^{\text{opt}}$ in (4.7) is obtained when D is such that: all columns of $\tilde{W} = DW$ have unitary 1-norm, with 1-norm bounds; all rows of $\tilde{W}^{-1} = D^{-1}W^{-1}$ have unitary 1-norm, with ∞ -norm bounds.*

After stating the main results we can show how they can be applied in the next section.

4.2. EXAMPLE: AN APPLICATION TO CANCER BIOLOGY

Consider a multi-compartment evolutionary model describing growth, mutation and metastasis of a heterogeneous tumour cell population [92], where a set of mutant cell lines $\mathcal{M}$ can spread in a set of body compartments $\mathcal{J}$ and the d available drugs are differently effective against different mutants in different compartments. The mutants can settle just in some compartments: $\mathcal{M}_k$ denotes the set of mutants in compartment k . Denoting by r_i^k the growth rate of mutant i in compartment k , q_{ji}^k the mutation rate from mutant j to i in compartment k , μ_i^{ck} the migration rate from compartment c to k of mutant i ($\mu_i^{ck} = 0$ if there is no migration path), $\phi_{s,i}^k$ the effect of drug s on mutant i in compartment k , and ℓ_s the constant amount of drug s , the concentration x_i^k of mutant $i \in \mathcal{M}_k$ in compartment $k \in \mathcal{J}$ evolves as

$$\dot{x}_i^k = \sum_{j \in \mathcal{M}_k} r_i^k q_{ji}^k x_j^k + \sum_{c \in \mathcal{J}} r_i^k \mu_i^{ck} x_i^c - \sum_{j \in \mathcal{M}_k} q_{ij}^k x_i^k - \sum_{c \in \mathcal{J}} \mu_i^{kc} x_i^k - \sum_{s=1}^d \phi_{s,i}^k \ell_s x_i^k.$$

We can see this model as a networked system with compartments (nodes), including a set of mutants, connected by possible migration routes (arcs). Compartment k is associated with the linear system $\dot{x}^k = A_k x^k + B_k u^k$, $y^k = x^k$, where $x^k = (x_i^k)_{i \in \mathcal{M}_k}$ includes all mutant lines in compartment k and $u^k = (u_i^k)_{i \in \mathcal{M}_k}$, where u_i^k is the sum of all cells of mutant i migrating to compartment k . For the state matrix,

$$[A_k]_{ii} = r_i^k q_{gg}^k - \sum_{\substack{j \in \mathcal{M}_k \\ g \neq j}} q_{gj}^k - \sum_{\substack{c \in \mathcal{J} \\ c \neq k}} \mu_g^{kc} - \sum_{s=1}^d \phi_{s,g}^k \ell_s,$$

with $g = \mathcal{M}_k(i)$, while

$$[A_k]_{ij} = r_g^k q_{fg}^k,$$

with $g = \mathcal{M}_k(i)$, $f = \mathcal{M}_k(j)$. The nonzero entries of B_k are $[B_k]_{ii} = r_g^k$, with $g = \mathcal{M}_k(i)$. The nonzero entries of the interconnection matrix G_h , associated with the arc from compartment k to compartment c , are $[G_h]_{ij} = \mu_i^{kc}$ if $g = \mathcal{M}_k(i) = \mathcal{M}_c(j)$.

As in all biological systems, the parameter values are subject to huge uncertainties. The network topology, and even the number of affected compartments, are not known exactly. However, if we assume that the overall networked system belongs to the family $\mathcal{N}$ satisfying Assumptions 8, and 9 with 1-norm bounds, with $\mathcal{D}^{\text{out}} = 3$ (mutants in a compartment can migrate to 3 other compartments at most), and values

$$\begin{aligned} \alpha &= -25.0227 & \chi &= 1.2236 & \mu_B &= 6.2 & \mu_C &= 1 & \mu_G &= 0.3 \\ \xi_A &= 8.5268 & \xi_B &= 0.93 & \xi_C &= 0 & \xi_G &= 0.045, \end{aligned}$$

then condition (4.3) is satisfied:

$$\alpha + \chi^2 [\xi_A + \mathcal{D}^{\text{out}}(\mu_B + \xi_B)(\mu_G + \xi_G)(\mu_C + \xi_C)] = -1.2 < 0.$$

As long as the networked system belongs to this class, *stability is robustly guaranteed* (namely, the adopted cancer therapy successfully reduces the tumour size) *for all topologies* with maximum degree 3, *regardless of the number of nodes* (affected body compartments) *and arcs* (possible migration paths), *and even of the actual number of inputs, outputs and states for each node* (number of mutants in each compartment). For comparative simulations, we consider

$$\mathcal{J} = \{1, 2, 3, 4\} \quad \text{and} \quad \mathcal{M} = \{1, 2, 3\}$$

with

$$\mathcal{M}_1 = \{1, 2\} \quad \mathcal{M}_2 = \{2, 3\} \quad \mathcal{M}_3 = \{1, 2, 3\} \quad \mathcal{M}_4 = \{1, 3\}.$$

We take the uncertain parameters in the same intervals for all compartments:

$$\begin{aligned} r_1 q_{11} &\in [2.4, 3.3], & r_2 q_{12} &\in [0.61, 0.82], & r_3 q_{13} &\in [0.76, 1], & r_1 q_{21} &\in [0.24, 0.33], \\ r_2 q_{22} &\in [3.1, 4.1], & r_3 q_{23} &\in [1.1, 1.5], & r_1 q_{31} &\in [0.73, 0.98], & r_2 q_{32} &\in [0.31, 0.41], \\ r_3 q_{33} &\in [3.8, 5.1], & q_{11}, q_{22}, q_{33} &\in [0.65, 0.78], & q_{12}, q_{13} &\in [0.13, 0.16], \\ q_{21}, q_{32} &\in [0.065, 0.078], & q_{23}, q_{31} &\in [0.2, 0.23], \\ r_1 &\in [2.8, 5.2], & r_2 &\in [3.5, 6.5], & r_3 &\in [4.3, 8.1], \\ \mu_1 &\in [0.19, 0.21], & \mu_2 &\in [0.29, 0.31], & \mu_3 &\in [0.099, 0.1], \\ \phi_{1,1} &\in [0.3981, 0.4019], & \phi_{2,1} &\in [0.497, 0.503], & \phi_{1,2} &\in [0.1592, 0.1608], \\ \phi_{2,2} &\in [0.1988, 0.2012], & \phi_{1,3} &\in [0.199, 0.201], & \phi_{2,3} &\in [0.2485, 0.2515]. \end{aligned}$$

Figure 4.2 shows the graph representation of the system with all the possible 12 mutation paths: each mutation path can be active or inactive (hence $\mu_i = 0$), leading to 4096 different graph topologies. With 2 available drugs, we compare four different therapies:

$$T_1 = \{\ell_1 = 1.957, \ell_2 = 21.137\} \quad T_2 = \{\ell_1 = 2.571, \ell_2 = 26.453\}$$

$$T_3 = \{\ell_1 = 3.531, \ell_2 = 29.302\} \quad T_4 = \{\ell_1 = 11.76, \ell_2 = 133.229\}.$$

Only with therapy T_4 the uncertain networked system satisfies the sufficient condition (4.3).

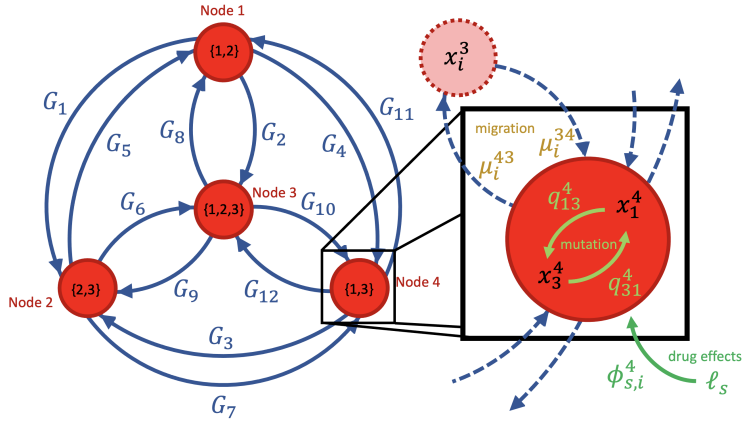


Figure 4.2: Example of possible graph topology for multi-compartment cancer evolution, with 4 nodes (body compartments) and 12 directed arcs (migration paths); illustration of the mutation, migration and drug selection dynamics.

As shown in Table 4.1, T_1 stabilises the nominal disconnected systems, but can fail in the presence of uncertainties and/or interconnections; T_2 guarantees robust stability of the disconnected systems, but can fail when the systems are interconnected; T_3 guarantees stability also of all the interconnected systems, but not robustly; finally, T_4 guarantees topology-independent robust stability, as expected.

Figure 4.3, showing the time evolution of the total number of cancer cells, and Figure 4.4, showing the eigenvalue distribution, confirm that only therapy T_4 ensures stability for all network topologies and all uncertainty realisations; for all other therapies, at least one system realisation is unstable, meaning that the chosen treatment fails and the tumour grows unboundedly.

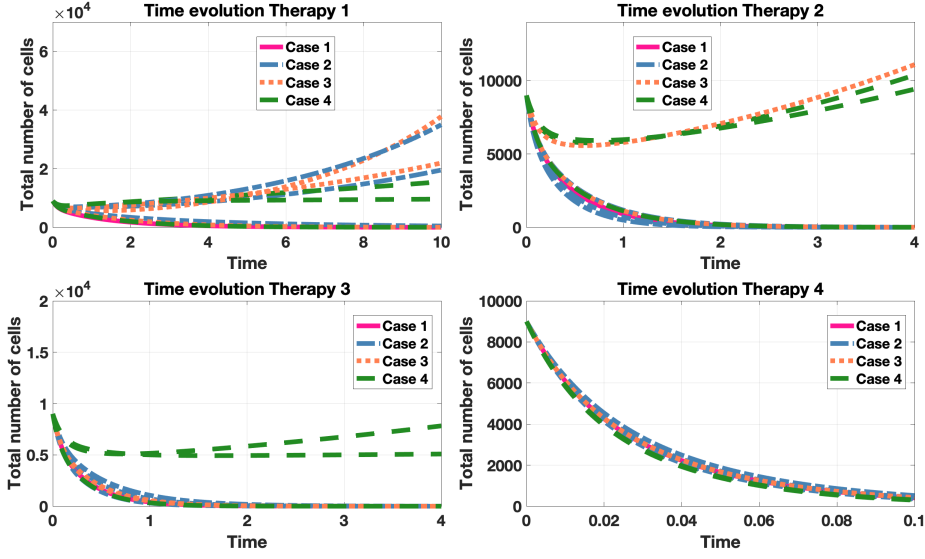


Figure 4.3: Time evolution for the system cases and therapies in Table 4.1.

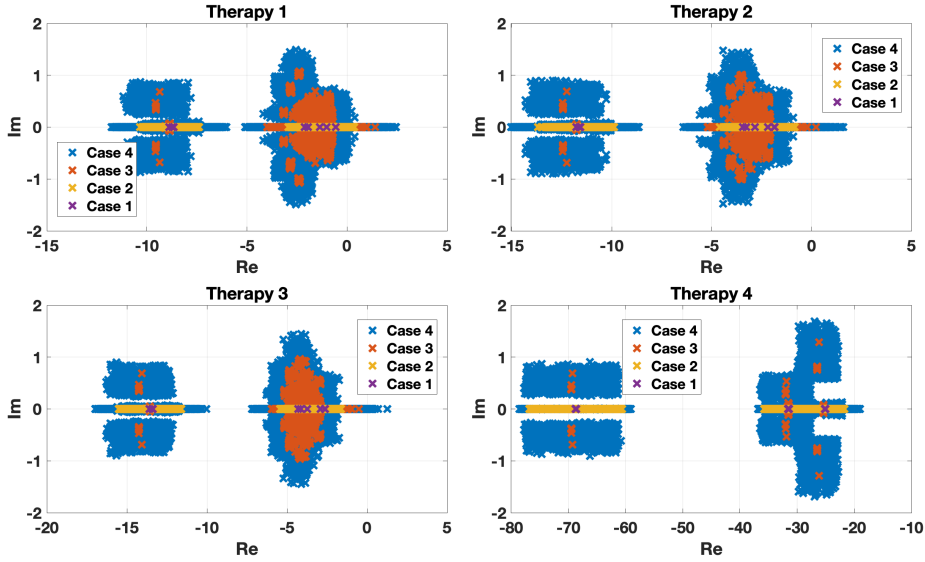


Figure 4.4: Eigenvalue distribution for the system cases and therapies in Table 4.1.

4.3. PROOFS

4.3.1. LEMMA 4

Lemma 4. Consider three matrices $\tilde{X}$, $\tilde{X}$, $\delta X \in \mathbb{R}^{n \times n}$ such that $\tilde{X} = \tilde{X} + \delta X$ and let $Z \in \mathbb{C}^{n \times n}$ be an eigenmatrix that diagonalises $\tilde{X}$. Given the scalars $\rho = \max_{\lambda \in \sigma(\tilde{X})} \{\text{Re}(\lambda)\}$ and $\kappa \geq \mathcal{K}_p(Z) \|\delta X\|_p$, matrix $\tilde{X}$ is Hurwitz stable if

	T_1	T_2	T_3	T_4
Case 1: disconnected, nominal	S	S	S	S
Case 2: disconnected, uncertain	U	S	S	S
Case 3: connected, nominal	U	U	S	S
Case 4: connected, uncertain	U	U	U	S

Table 4.1: Effect of the therapies in different cases. **S**: stability is guaranteed for all systems in the case. **U**: at least one system in the case is unstable. Simulations with 150 random parameter variations for uncertain disconnected systems and with all the 4096 possible interconnection topologies for connected systems (each with 2 parameter variations in the uncertain case).

$$\rho + \kappa < 0. \quad (4.8)$$

Proof. Let $D(x, r)$ denote the closed disk with center $x \in \mathbb{C}$ and radius r . By the Bauer-Fike theorem (Theorem 2), all the eigenvalues of $\tilde{X}$ are located in the set $Y = \bigcup_{\lambda \in \sigma(\tilde{X})} D(\lambda, \kappa)$. Since $\max_{\varphi \in Y} \{\operatorname{Re}(\varphi)\} = \max_{\lambda \in \sigma(\tilde{X})} \{\operatorname{Re}(\lambda)\} + \kappa = \rho + \kappa$, all the eigenvalues of $\tilde{X}$ have negative real part if $\rho + \kappa < 0$. $\square$

Since $\kappa \geq 0$, condition (4.8) requires $\rho < 0$, i.e. Hurwitz stability of the nominal $\tilde{X}$.

4.3.2. PROOF OF THEOREM 10

Proof. Under Assumption 7, the overall dynamics for the disconnected node systems is

$$\dot{x} = Ax + Bu, \quad y = Cx, \quad (4.9)$$

where $x = [x^{(1)\top} \dots x^{(N)\top}]^\top$, $u = [u^{(1)\top} \dots u^{(N)\top}]^\top$, $y = [y^{(1)\top} \dots y^{(N)\top}]^\top$, $A = \operatorname{diag}(A_i)_{i \in \mathcal{N}}$, $B = \operatorname{diag}(B_i)_{i \in \mathcal{N}}$, $C = \operatorname{diag}(C_i)_{i \in \mathcal{N}}$. Splitting nominal and uncertain parts, $A = \bar{A} + \hat{A}$, where $\bar{A} = \operatorname{diag}(\bar{A}_i)_{i \in \mathcal{N}}$, $\hat{A} = \operatorname{diag}(\hat{A}_i)_{i \in \mathcal{N}}$; $B = \bar{B} + \hat{B}$, where $\bar{B} = \operatorname{diag}(\bar{B}_i)_{i \in \mathcal{N}}$, $\hat{B} = \operatorname{diag}(\hat{B}_i)_{i \in \mathcal{N}}$; $C = \bar{C} + \hat{C}$, where $\bar{C} = \operatorname{diag}(\bar{C}_i)_{i \in \mathcal{N}}$, $\hat{C} = \operatorname{diag}(\hat{C}_i)_{i \in \mathcal{N}}$.

The node systems are connected through the incidence matrix $X \in \{-1, 0, 1\}^{N \times M}$ of $\mathcal{G}$, defined as $X_{ih} = 1$ if the arc $h \in \mathcal{A}$ enters node $i \in \mathcal{N}$; $X_{ih} = -1$ if the arc h leaves node i ; and $X_{ih} = 0$ otherwise. In particular, we define matrices $P = \max\{X, 0\}$ and $R = -\min\{X, 0\}$ element-wise, so that $P_{ih} = 1$ if arc h enters node i and $R_{ih} = 1$ if arc h leaves node i . These scalar matrix entries match nodes and arcs according to the interconnection topology.

Then, the input to node $i \in \mathcal{N}$ is

$$u^{(i)} = \sum_{h=1}^M P_{ih} G_h \left(\sum_{j=1}^N R_{jh} y^{(j)} \right),$$

where just one of the scalars R_{jh} , with $j = 1, \dots, N$, is nonzero and selects the node output “feeding” arc h .

Let $G = \operatorname{diag}(G_h)_{h \in \mathcal{A}}$, then the vector of inputs can be written as

$$u = \mathcal{P} G \mathcal{R} y, \quad (4.10)$$

where the matrices $\mathcal{P}$ and $\mathcal{R}$ are as described in the theorem (Theorem 10). Merging (4.9) and (4.10) produces the networked system

$$\dot{x} = [A + B \mathcal{P} G \mathcal{R} C] x \doteq \tilde{A} x, \quad (4.11)$$

If all the subsystems have the same number of inputs and outputs ($p_i = p$ and $q_i = q$ for all $i \in \mathcal{N}$), then the matrices $\mathcal{P}$ and $\mathcal{R}$ simplify to $(P \otimes I_p)$ and $(R^\top \otimes I_q)$ respectively, and Equation (4.11) becomes

$$\dot{x} = [A + B(P \otimes I_p)G(R^\top \otimes I_q)C]x \doteq \tilde{A}x. \quad (4.12)$$

□

4.3.3. PROOF OF THEOREM 11

Proof. To assess stability of the networked system (4.1), we rewrite matrix $\tilde{A}$ as the sum of three matrices

$$\tilde{A} = \bar{A} + \delta A_1 + \delta A_2,$$

where

$$\bar{A} = \text{diag}(\bar{A}_i)_{i \in \mathcal{N}}$$

is the block diagonal matrix of state matrices,

$$\delta A_1 = \hat{A} = \text{diag}(\hat{A}_i)_{i \in \mathcal{N}}$$

represents the uncertainty in the state dynamics, and

$$\delta A_2 = B\mathcal{P}G\mathcal{R}C$$

includes the uncertainty due to the input and output matrices and to the interconnection.

Thanks to its particular block-diagonal structure, the nominal matrix $\bar{A}$ can be diagonalised as $\bar{A} = W^{-1}\Lambda W$, where $\Lambda = \text{diag}(\Lambda_i)_{i \in \mathcal{N}}$ has on the diagonal the blocks $\Lambda_i = \text{diag}(\lambda)_{\lambda \in \sigma(\bar{A}_i)}$ including the eigenvalues of the individual systems at the nodes, while $W = \text{diag}(W_i)_{i \in \mathcal{N}}$ has on the diagonal the eigenmatrices W_i of $\bar{A}_i$ that satisfy $\|W_i\|_1 \leq \chi$ and $\|W_i^{-1}\|_1 \leq \chi$ as per Assumption 9 with 1-norm bounds.

The stability of $\tilde{A}$ can be checked by applying Lemma 4 with

$$\tilde{X} = \tilde{A} \quad \bar{X} = \bar{A} \quad \delta X = \delta A_1 + \delta A_2 \quad Z = W \quad \varrho = \alpha,$$

and

$$\kappa = \chi^2 [\xi_A + \mathcal{D}^{\text{out}}(\mu_B + \xi_B)(\mu_G + \xi_G)(\mu_C + \xi_C)]. \quad (4.13)$$

In fact, since $\max_{\lambda \in \sigma(\bar{A})} \{\text{Re}(\lambda)\} \leq \alpha$, which is negative in view of Assumption 9, the nominal state matrices are Hurwitz stable.

To make sure that the assumptions of Lemma 4 are all satisfied, we must show that

$$\mathcal{K}_1(W) \|\delta A_1 + \delta A_2\|_1 \leq \kappa, \quad (4.14)$$

with κ as in Equation (4.13). We have $\mathcal{K}_1(W) \leq \chi^2$ in view of Lemma 1, while

$$\|\delta A_1 + \delta A_2\|_1 \leq \|\delta A_1\|_1 + \|\delta A_2\|_1$$

can be upper bounded by exploiting Theorem 1 and, in view of the block structure of matrix δA_1 , Lemma 1:

$$\|\delta A_1\|_1 = \|\hat{A}\|_1 = \max_{i \in \mathcal{N}} \{\|\hat{A}_i\|_1\} \leq \xi_A$$

and

$$\begin{aligned}\|\delta A_2\|_1 &\leq \|\mathcal{P}\|_1 \|\mathcal{R}\|_1 \|B\|_1 \|G\|_1 \|C\|_1 \\ &\leq \mathcal{D}^{\text{out}}(\mu_B + \xi_B)(\mu_G + \xi_G)(\mu_C + \xi_C),\end{aligned}$$

where the last inequality holds because

$$\|\mathcal{P}\|_1 \|\mathcal{R}\|_1 = \|P\|_1 \|R^\top\|_1 \leq \mathcal{D}^{\text{out}} \quad (4.15)$$

by construction, and

$$\|B\|_1 = \|\bar{B} + \hat{B}\|_1 \leq \|\bar{B}\|_1 + \|\hat{B}\|_1 = \max_{i \in \mathcal{N}} \{\|\bar{B}_i\|_1\} + \max_{i \in \mathcal{N}} \{\|\hat{B}_i\|_1\} \leq \mu_B + \xi_B,$$

$$\|G\|_1 = \|\bar{G} + \hat{G}\|_1 \leq \|\bar{G}\|_1 + \|\hat{G}\|_1 = \max_{h \in \mathcal{A}} \{\|\bar{G}_h\|_1\} + \max_{h \in \mathcal{A}} \{\|\hat{G}_h\|_1\} \leq \mu_G + \xi_G,$$

$$\|C\|_1 = \|\bar{C} + \hat{C}\|_1 \leq \|\bar{C}\|_1 + \|\hat{C}\|_1 = \max_{i \in \mathcal{N}} \{\|\bar{C}_i\|_1\} + \max_{i \in \mathcal{N}} \{\|\hat{C}_i\|_1\} \leq \mu_C + \xi_C.$$

Since inequality (4.14) is proven, Lemma 4 can be applied and guarantees that matrix $\tilde{A}$ in (4.1) is Hurwitz stable if the sufficient condition (4.3) is satisfied. $\square$

4.3.4. PROOF OF THEOREM 12

Proof. If condition (4.4) is violated, hence

$$\alpha + (\xi_A + \mathcal{D}^{\text{out}}(\mu_B + \xi_B)(\mu_G + \xi_G)(\mu_C + \xi_C)) \geq 0, \quad (4.16)$$

then there exists an unstable system structure in the family $\mathcal{N}$. We show that this structure is associated with a circulant matrix, for which the following result [98, Sec. 3.1] holds.

Theorem 13 (Spectrum of a circulant matrix [98]). *The eigenvalues of a circulant matrix $C \in \mathbb{R}^{n \times n}$ with coefficients $\{c_0, c_1, \dots, c_{n-1}\}$,*

$$C = \begin{bmatrix} c_0 & c_1 & \dots & c_{n-1} \\ c_{n-1} & c_0 & \dots & c_{n-2} \\ \vdots & \vdots & \ddots & \vdots \\ c_1 & c_2 & \dots & c_0 \end{bmatrix},$$

are

$$\psi_m = \sum_{k=0}^{n-1} c_k \rho_m^k, \quad (4.17)$$

where $\rho_m = \exp(\frac{-2\pi i m}{n})$, $m \in \{0, \dots, n-1\}$.

Without loss of generality, since the networked systems in the family $\mathcal{N}$ can have node systems of *any size*, we consider a networked system where each node is the same scalar system $\dot{x}^{(i)} = ax^{(i)} + bu^{(i)}$, $y^{(i)} = cx^{(i)}$, $i \in \mathcal{N}$, and all the arcs are associated with the same interconnection scalar g . Assume that $\mathcal{D}^{\text{out}}$ arcs leave each node to reach the previous $\mathcal{D}^{\text{out}}$ nodes: there is an arc leaving node i to node $(i-k) \bmod M$ for $i \in \mathcal{N}$ and $k = 1, \dots, \mathcal{D}^{\text{out}}$.

The networked system has the following state matrix:

$$\tilde{A} = aI_N + (bI_N)(P)(gI_M)(R^\top)(cI_N) = aI_N + bgcPR^\top,$$

where $[PR^\top]_{ij} = 1$ if there is an arc going from node j to i , $[PR^\top]_{ij} = 0$ otherwise. For this graph structure, $\tilde{A}$ is a circulant matrix satisfying Theorem 13, where $c_0 = a$, $c_k = bgc$ for $k = 1, \dots, \mathcal{D}^{\text{out}}$ and the other coefficients are zero. Then, equation (4.17) with $m = 0$ gives $\psi_0 = \sum_{k=0}^{n-1} c_k = a + \mathcal{D}^{\text{out}} bgc = \alpha + \xi_A + \mathcal{D}^{\text{out}}(\mu_B + \xi_B)(\mu_G + \xi_G)(\mu_C + \xi_C)$, where the last equality is obtained by splitting the *nominal* and the *uncertain* part in $a = \alpha + \xi_A$, $b = \mu_B + \xi_B$, $c = \mu_C + \xi_C$, and $g = \mu_G + \xi_G$.

By inequality (4.16) the eigenvalue ψ_0 of matrix $\tilde{A}$ is nonnegative, therefore the networked system is unstable. Since this system belongs to the family $\mathcal{N}$, this proves the necessity of condition (4.4). $\square$

4.3.5. PROOF OF PROPOSITION 4

Proof. Set $U = W^{-1}$ and denote by W_j the j -th column of W , by U_i the i -th row of U . Then, $\|WD\|_1 \|D^{-1}U\|_1 = \max_j D_j \|W_j\|_1 \max_h \sum_i \frac{|U_{ih}|}{D_i} = \max_j \frac{1}{z_j} \max_h v_h^\top z$, where the last equality follows by assuming without restriction that $\|W_j\|_1 = 1$ (which can be obtained via pre-scaling) and denoting by z the vector with i -th component $z_i = 1/D_i$ and by v_h the vector with i -th component $|U_{ih}|$. In the dual case, assuming $\|U_h\|_1 = 1$ without restriction, $\|WD\|_\infty \|D^{-1}U\|_\infty = \max_j \sum_i |W_{ij}| D_j \max_h \frac{\|U_h\|_1}{D_h} = \max_i v_i^\top z \max_h \frac{1}{z_h}$, where vector z has i th component $z_i = D_i$ and vector v_i has j th component $|W_{ij}|$.

Then the function to be minimised can be written in the form

$$\phi(z) = \max_j \left\{ \frac{1}{z_j} \right\} \max_h \{v_h^\top z\},$$

where v_h are non-negative vectors and $z > 0$ component-wise. Since ϕ is positively homogeneous of order 0 (i.e., $\phi(\lambda z) = \phi(z)$ for any $\lambda > 0$), we can find its minimum assuming the additional constraint

$$\max_j \left\{ \frac{1}{z_j} \right\} = 1. \quad (4.18)$$

Indeed, if $z^{\text{opt}} > 0$ is a minimum, then we can take the maximum $1/z_{i^*}^{\text{opt}} = \max_j 1/z_j^{\text{opt}}$ and set $\lambda \doteq 1/z_{i^*}^{\text{opt}} \geq 1/z_j^{\text{opt}}$ for all j . Now, λz^{opt} produces the same minimum value, since $\phi(\lambda z^{\text{opt}}) = \phi(z^{\text{opt}})$, and satisfies (4.18). Therefore, the additional constraint does not change the result. The surface in (4.18) can be split into n faces: $\mathcal{F}_i = \{z: z_i = 1, z_j \geq 1, j \neq i\}$, for $i = 1, \dots, n$. So we need to consider n problems of the form $\min_z \max_h \{v_h^\top z\}$ with constraints $z_i = 1$ for $i = 1, \dots, n$ and $z_j \geq 1$ for $j = 1, \dots, n, j \neq i$, which can be converted into linear programs. Since all the components of v_h are non-negative, the minimum of the i th problem with $z_i = 1$ is immediately achieved by choosing the smallest possible value for all other components: $z_j = 1$ for all $j \neq i$. Hence the initial pre-scaling, with $\|W_j\|_1 = 1$ in the 1-norm case and $\|U_h\|_1 = 1$ in the ∞ -norm case, was already optimal. $\square$

4.3.6. PROOF OF PROPOSITION 3

Proof. For diagonal systems, $\chi^2 = 1$. Then, the result directly follows from Theorems 11 and 12 in the 1-norm case, and from Propositions 1 and 2 in the ∞ -norm case. $\square$

4.4. CONCLUSIONS

We deal with the topology-independent robust stability analysis of uncertain networked systems with completely unknown topology (but known maximum connectivity degree). Both the necessary and the sufficient condition provided here are easy to verify in a state-space framework and are fully scalable, since they can be checked locally and do not depend on the number of nodes and arcs. Both the systems at the nodes and their uncertainties, as well as the uncertain interconnection matrices at the arcs, can be heterogeneous, thus making these conditions applicable to a general class of systems.

However, our results are conservative, because they cannot exploit the knowledge of the physical structure of the uncertainties in the system parameters and interconnections, and tight norm bounds on the system matrices are hard to obtain.

Since we seek *topology-independent* results, another unavoidable source of conservativeness is that we cannot exploit the knowledge of the interconnection and its possibly stabilising effects. Hence, requiring stability of the individual subsystems is necessary for topology-independent stability: the system with disconnected nodes is a possible topology. The interconnection may compromise the stability of the node systems (as shown in the case-study in Section 4.2). The maximum connectivity degree naturally appears in our conditions; intuitively, a smaller degree facilitates topology-independent stability, because it limits the number of possible topologies among which the worst case must be considered.

For given nominal systems and uncertainty bounds, conditions (4.3) or (4.5) allow to find the maximum connectivity degree ensuring topology-independent robust stability; the stability of the networked system is robust to online modifications of the network, in a plug-and-play framework [26, 184], provided that the maximum connectivity degree is not exceeded.

Inequalities (4.3) and (4.5) can then be seen as a balance between the stable systems at the nodes, on the one hand, and the uncertainties and interconnections that can potentially destabilise the overall system, on the other hand. The sufficient condition may not be satisfied because the spectral abscissa α of the nominal systems at the nodes is not negative enough to counteract the possibly destabilising effect of interconnections and uncertainties. Then, local controllers can be added to move the eigenvalues further to the left of the complex plane, until the sufficient condition is met.

II

PART 2: MODELLING AND ANALYSIS OF OPINION FORMATION DYNAMICS

5

BACKGROUND ON OPINION FORMATION MODELS

What happened, happened,
and couldn't have happened any other way.

Morpheus, The Matrix Reloaded

This chapter introduces fundamental concepts and sets the context for Part 2 of the dissertation, which focuses on the analysis and development of agent-based opinion formation models. This chapter also elaborates on the connection between Chapters 6 and 7 and provides details related to the simulation results.

Parts of this chapter are based on the manuscripts "A framework to analyze opinion formation models" (2022) [57], and "Classification-Based Opinion Formation Model Embedding Agents' Psychological Traits" (2023) [58] by Carlos Andrés Devia and Giulia Giordano.

5.1. INTRODUCTION, MOTIVATION, AND CONTRIBUTION

The development and analysis of opinion formation models has been an active field of research since the introduction of the first opinion formation models by French, DeGroot, and Harary [77, 105, 106, 54]. In recent years the study of opinion formation has attracted growing attention [195, 179, 180, 159, 75, 9, 183]. An *opinion formation model* is a mathematical model aimed at reproducing the evolution of opinions within a population in a given time interval. Increasingly more sophisticated models have been developed by embedding different concepts such as *susceptibility* [78, 79], *stubbornness* [111, 160], *leaders* [124, 125], *emotions* [199, 47], *trust* [222, 133], *bounded confidence* [110], *coevolving networks* [201, 197], *biases* [196], *polarity* [154], *assimilation* [49, 83, 154, 18], *tolerance* [65], *mass media* [42], *controversy* [23], *weighted balance theory* [191], *curation algorithms* (or *recommender systems*) [185], among others.

In the resulting models, opinions can be represented by continuous [52, 110] or discrete [132, 202] variables, and can evolve in discrete [77, 54] or in continuous [2, 1] time in a deterministic [54] or stochastic [152] way, over an underlying interaction graph that can be time-varying [117, 151, 177], directed, weighted, or signed [6]. Although there may be different reasons to construct mathematical models of opinion formation [69], the ultimate goal is typically to capture the mechanisms behind opinion change in society and accurately predict the evolution of real-life opinions [203, 205]. The literature on opinion formation models is very varied and rich. In this dissertation we focus on a type of model called agent-based models (ABMs). These models offer a high degree of versatility and flexibility, by allowing individual agents to update their opinions following a (possibly unique) opinion update law. At the same time, the model evolves over a graph which can also be analysed and showcase interesting properties.

Agent-based models (ABMs), such as the French-DeGroot model [54], are very common in the opinion formation literature. In an ABM, every individual holds a different opinion (or vector of opinions) and interacts with the other agents according to a given function over a network that can be directed, weighted, or signed. Some examples of agent-based models are those by Hegselmann and Krause [112], Salzarulo [188], and Deffuant [52], among many others [209, 3, 51, 162, 208]. An extensive literature [159] proposes and analyses opinion formation models for different types of agent interactions and network characteristics.

In this dissertation, each agent has a single time-varying absolute opinion, denoted $x_i[k]$ for agent i at time $k \in \mathbb{N}$. It is assumed that each opinion belongs to the interval $[-1, 1]$ and that its value represents the level of agreement or disagreement the agent has with a particular statement (an alternative, equivalent, interpretation is that of *polar opinions*, where both opinions represent two competing alternative [7]). Opinion values of $+1$, -1 , and 0 represent complete agreement, complete disagreement, and indifference, respectively; and intermediate values correspond to less extreme opinions, i.e. the opinions are absolute, in contrast with *relative* opinions [41].

When studying the behaviours emerging from these models, the focus is not on individual opinions but on the overall evolution of opinions in the entire population. Denoting as *opinion distribution* the collection of all the opinions within a population at a given time instant, the study of opinion formation models addresses four main questions:

1. **How does the model evolve opinion distributions?** (informally, what does the model do to the opinions?)
2. **What opinion distributions can the model produce, and under which conditions?** (informally, what can the model do?)
3. **What can be said about the model outcomes when only incomplete information is known?**
4. **How realistic is the model?**

Let us examine each one of these questions more closely and answer them for two of the simplest opinion formation models: the Null model, and the French-DeGroot (FG) model.

1. *How does the model evolve opinion distributions?:* consider the set of all possible opinion distributions. Any opinion formation model is a map from this set to itself. This question basically asks, *how does this map look like?* For the Null model, the answer is simple: the opinion distributions remain the same, and there is no transformation. Therefore, the map is the identity. For the FG model, the opinions move closer together and eventually (asymptotically) become identical.
2. *Which opinion distributions can the model produce and under which conditions?:* regarding an opinion formation model as a map, this question asks about the range of the map and how it changes depending on the model parameters. Related relevant questions include: *for which model parameters is the range invariant?* and *can the model produce any possible opinion distribution (or opinion distribution type) for suitable model parameters?* For the Null model, which has no parameters, the range is the same as the domain, and every opinion distribution belongs to it. For the FG model, if the underlying digraph is strongly connected, then the range is contained in the domain. This is because this model cannot produce for instance, opinion distributions in which half of the opinions are -1 and half are 1 .
3. *What can be said about the model outcomes when only incomplete information is known?:* one characteristic of social systems is that measuring or estimating their parameters is remarkably challenging. In a realistic scenario, only partial information about the system is known, for instance, the set of initial opinions (but not which opinion is held by which agent, i.e., the allocation of opinions, see Section 5.5.5), the set of agent parameters (but not their allocation), properties of the initial opinion set, agent parameter set, or underlying digraph. When only incomplete information is available, what can be said about the model outcomes? For the Null model, all the properties of the initial opinions also hold for the final opinions. For the FG model, it is known that the final opinions will always belong to the convex hull of the initial opinions. So, for instance, if the only known information is that the initial opinions are positive for both models, we can say that the final opinions will also be positive, independent of any other model parameter.
4. *How realistic is the model?:* this question can be rephrased as *to what degree can the model reproduce opinion transitions observed in real life?* And to answer it, real-life opinion information is required. The question also depends on the type of opinion formation phenomenon under consideration. For instance, if the model aims to reproduce the opinion evolution in a team or group, the FG model outcomes are accurate in that teams often reach a consensus. In contrast, if it aims to reproduce the opinion evolution of a population, it is often inaccurate, as consensus is not often seen at the broad societal level.

There are two main challenges associated with answering these questions. The first one is about *representation*, and the second one is about *data analysis*:

- **Representation problem:** For simplicity, assume that the set of opinions an agent can have is $O = \{-1, -1 + \delta, \dots, 1 - \delta, 1\}$ where $\delta = 2/K$ for some integer K and that we are analysing models with N agents. Therefore, the set of all possible opinion distributions is O^N , a finite set. In this case, an opinion distribution is a set of N numbers from O . This complete representation of an opinion distribution is not practical to use because it is not intuitive and is difficult to manage and interpret. With this representation, even comparing two opinion distributions becomes difficult. To formulate questions and statements about opinion distributions, a more intuitive and practical way to represent them is necessary: we need a representation that is formal and retains the most relevant information.

- **Data analysis:** Assume that the model we analyse guarantees that the predicted opinion distribution belongs to O^N for every possible prediction horizon and model parameters. Suppose the elements in O^N are given an order. Then, for any prediction horizon and model parameters, the model can be represented by a look-up table that relates the initial opinions with the predicted opinions. Moreover, suppose the collection of every possible prediction horizon and model parameters is also ordered. In that case, the opinions predicted by the model can be arranged in a multidimensional look-up table containing all the information about the model. In principle, this multidimensional look-up table contains the answers to every question we might have about the model. However, even if it could be calculated (which cannot be because it would require too many computations), the answers are hidden in a tremendous amount of data. So the challenge is to select which simulations to run and how to arrange and interpret the results.

In this section we propose a methodology to solve these problems and answer the four aforementioned questions for agent-based opinion formation models. The methodology relies on four techniques: *Histogram-based Sorting Algorithm*, *Transition Tables*, *Agreement Plot*, and *Probabilistic Analysis*.

Histogram-based Sorting Algorithm and Transition Tables: We identified *perfect consensus* (all individuals share the very same opinion), *consensus* (all individuals have almost the same opinion), *polarisation* (presence of two opposed opinion groups), *clustering* (presence of several distinct opinion groups) and *dissensus* (a practically uniform distribution of the opinions) as qualitative categories of opinion distributions that emerge in real life. They recurrently appear in the results of the World Values Survey [99, 101, 100] (WVS), a research project that explores people's values and beliefs by conducting global surveys approximately every 5 years; in particular, we monitored the answers to 30 questions (regarding values, behaviour, and ethics) asked to participants in 25 countries in three occasions separated by roughly 5 years, these three sets of survey answers are labeled as *wave 5*, *wave 6*, and *wave 7* (the years in which each survey wave was conducted varies with country, however, there are approximately 5 years between two consecutive waves in the same country: wave 5 spans the years 2005-2009, wave 6 the years 2010-2014, and wave 6 the years 2017-2022).

These categories are also present in the outcomes produced by different opinion formation models. For instance, the French-DeGroot model is guaranteed to asymptotically achieve *perfect consensus* if the graph is strongly connected [179]. For a structurally balanced digraph, the Altafini model predicts *polarisation* if the digraph is strongly connected [6, 166] and *consensus* near the origin, if it has a spanning tree [165, 164]. When bounded confidence is added to the model, then *clustering* is a likely outcome [109], and if the confidence radius is small enough opinions can remain in *dissensus*.

The proposed *Histogram-based Sorting Algorithm* associates an opinion distribution with a qualitative category. Histogram-based sorting has been used in many fields, especially related to image processing [187, 28]; yet, to the best of our knowledge, this is the first time it is adopted in an opinion-dynamics setting. Second, we construct a *Transition Table* to visualise how opinion distributions evolve over time from an *initial* to a possibly different *final* qualitative category.

Transition Tables are used to assess and analyse known opinion formation models and compare their prediction with real data. To do this, we compare real-world data, representing opinion distributions at different sampling times (waves) taken from the World Values Survey [99, 101, 100], with the predictions provided by various well-known opinion formation models: French-DeGroot [77, 54], Weighted Median [163], Bounded Confidence [53, 215, 216], and Quantum Game [55] models.

An extensive literature deals with model comparison based on Bayesian analysis: Bayes factor [37, 139, 128], Bayesian evidence [122], and Bayesian methods such as the L_p measure [148].

Model comparison statistics [147] can be based on the information criterion [212] or the deviance information criterion [200]. These statistical model comparisons aim at selecting the model that fits at best a given data set. Conversely, simulation-based model comparison directly compares the outcomes of two or more models without fitting real data [48, 186].

As a main novelty, we propose a peculiar simulation-based framework that compares opinion formation models not in their ability to reproduce a given data set, but in their capability to generate *a spectrum of qualitative behaviours that is as broad as the one observed in real life*.

Our results provide insight into real-life opinion evolution and comparatively assess different opinion formation models. They reveal that, while *all* transitions between qualitative categories occur in reality, existing models can only yield *some* peculiar transitions and are characterised by a *bias towards consensus* that cannot be found in real opinion data. This observation, and other qualitative comparisons, can also be identified by the tools presented in the next chapters.

A specific application for the Histogram-based Sorting Algorithm and Transition Tables is to characterise opinion evolutions seen in real life for populations with different backgrounds and try to relate these evolutions with societal context. So, for instance, in a society that benefits from significant economic growth, the opinion regarding government economic policies may shift from polarisation to consensus, reflecting the fact that most individuals approve these policies.

Agreement Plot: besides labelling an opinion distribution using the Histogram-based Sorting Algorithm, another intuitive and helpful representation of an opinion distribution is by its mean and the mean of the opinion absolute values. Plotting this point in the Cartesian plane results in what we call the Agreement Plot. The mean of the opinion absolute values measures the population's interest in a statement. If it is near 1, then the agent's opinions are near -1 or 1 , meaning strong agreement or disagreement, i.e. the agents care a lot about the statement. On the contrary, if the mean of the opinion absolute values is close to 0, then the agents are mostly indifferent. On the other hand, the mean of the opinions gives information on what the general average opinion is.

A crucial advantage of the Agreement Plot is that, by reducing a complete opinion distribution to a single point in the Cartesian plane, it is possible to plot multiple opinions at once. This allows for plotting and comparing multiple opinion evolutions for different initial opinions, agent parameters, and underlying networks. This flexibility enables a high-level analysis of the model behaviour for various parameters and, consequently, the identification of patterns and model properties.

A specific application for the Agreement Plot is to give a measure of the capacity that a given opinion formation model has to produce different opinions. This can be done thanks to the fact that all opinion distributions have a representation in the triangle with vertices $(0, 0)$, $(1, 1)$, and $(1, -1)$. Therefore, 'measuring' how much 'area' an opinion formation model can cover enables us to assess the model versatility and capacity to produce rich results.

Probabilistic Analysis: in the context of opinion dynamics, we often have access to only limited information. For instance, only the mean of the initial opinions is known, or only the variance of the agent parameters. In these cases, making statements on the model outcomes is still possible. For instance, for the French-DeGroot model, the variance of the predicted opinions will be lower than that of the initial opinions, independent of the initial opinions' exact value.

When limited information is available, no precise statements can be made about the final opinion distribution. Instead, the statements become probabilistic; starting with incomplete information, there is a certain probability that the predicted opinions will have a given property or qualitative characteristic. This is the approach taken in the Probabilistic Analysis technique.

A specific application for the Probabilistic Analysis is to identify combinations of agent parameters and initial opinions that produce results unexpected from certain opinion formation models. For instance, the Friedkin-Johnsen model is known for its tendency to produce consensus, so it is

interesting to know whether, for suitable parameters, the model can produce polarisation or clustering.

The Agreement Plot and Probabilistic Analysis techniques are used to study and analyse three agent-based opinion formation models: the Friedkin-Johnsen, Bounded Confidence, and Backfire Effect and Biased Assimilation models [79, 110, 45]. The results provide insight into the intrinsic properties of the models and exemplify the application of these techniques to agent-based models.

The application of the proposed methodology and the obtained results are detailed and explained in Chapter 6. After introducing this methodology and applying it to known opinion formation models, we present our own agent-based opinion formation model, the **Classification-based model**. Two novel features characterise this agent-based model.

1. Even if the agents communicate, and openly express their real opinion, it is impossible for an agent to exactly measure and quantify the opinion of another. Therefore, the model introduces a classification-based approach, supported by the empirical finding that the assessment of the opinion of others depends on the perceived distance to those others [191]: each agent classifies its neighbours in different groups according to their *perceived* opinion, distinguishing between those that agree *much more*, or *more*, or *comparably*, or *less*, or *much less* (than itself) with a given statement.

The fact that agents don't have access to the exact opinion of their neighbours with infinite resolution and accuracy has been taken into account by models with quantised opinions [167, 40], while threshold models [95, 96] could be seen as adopting a classification approach because the opinion update law depends on the number of neighbours expressing a particular opinion or action. Our classification-based approach is based not on the opinion of an agent's neighbours, but on the weighted difference between the opinion of the agent and of its neighbours, accounting for the finite resolution with which agents perceive the opinions of their neighbours. Also in opinion formation models with private and public opinions [168, 10, 192, 65, 17] the agents cannot have perfect access to the real opinion of their neighbours. However, there is a critical difference. In these models, the agents can choose which public opinion they show, with certainty that it will be the opinion perceived by others, and hide their true private opinion: the misperception is intentional. Conversely, in our model, the misperception is unintentional and unavoidably caused by the inability to communicate with infinite accuracy: the agents wish to show as openly as possible their opinion, which still cannot be perceived with infinite resolution, and the other agents can only perceive the range in which the opinion falls, which depends on both the agent that expresses the opinion and the one that assesses it. Therefore an agent cannot know with certainty how its opinion is perceived by others. Also in the Continuous Opinions and Discrete Actions model [156], the mismatch between real and perceived opinions is intentional and due to the agents purposefully hiding their actual opinion to others (each agent controls the action it takes and consequently how its opinion is perceived by its neighbours), while in the classification-based model the mismatch is due to the intrinsically imperfect perception mechanism.

To reflect imperfect communication in the model, our proposed solution of classifying the opinion of others in one of five categories is inspired by the field of psychometrics: in questionnaires, the responses quantify opinions according to discrete scales. Likert scales are a standard psychometric scale used to analyse surveys, which in turn are the typical approach to measure the opinions of individuals in a population. In our model, the process of agent i assessing the opinion of agent j yields, at each time step, the answer to a five-point Likert question, which asks how much agent j agrees with a statement, compared with agent i , where the possible answers are: *Agrees much more*, *Agrees more*, *Agrees the same*, *Agrees less*, and *Agrees much less*. For certain specific questions and specific social groups and connec-

tions, the perception may be sharper, while in other cases it may be less sharp; also, some agents may have a sharper perception than others. Five levels are chosen as a compromise resolution to account for the perception skills of the *average agent* interacting with an *average neighbour*. Still, the model could be modified to consider more than five levels, thus accounting for agents with a sharper average perception, and differentiating the sharpness of perception for different agents could also be interesting future work.

2. Each agent behaves according to a combination of three *internal traits* based on well studied sociological and psychological concepts: conformisms, radicalism, and stubbornness.
 - Conformism: agents tend to agree with their neighbours. This behaviour was first shown in the conformity experiment by Asch [14, 15, 16] and evolved into social conformity theory [137]. A similar behaviour is supported by the cognitive dissonance theory [73, 161].
 - Radicalism: agents do not care if their opinion is different from their neighbours'. On the contrary, their opinion is strengthened by the presence of agents with a similar opinion, which reinforce their beliefs; this is known as the persuasive argument theory, which supports several polarisation models [157, 135, 153, 83, 178].
 - Stubbornness: agents refuse to change their opinion; this type of behaviour has been often present in opinion formation models starting from those by Friedkin and Johnsen [81, 79].

In our proposed model, the behaviour of each agent is determined by a *combination* of these three traits: in fact, actual people are not completely conformist, radical, or stubborn, but everyone is characterised by a peculiar blend of these three traits. The inclusion of the radical trait can be seen as an extension of the model by Friedkin and Johnsen [81, 79], which includes both conformist and stubborn traits.

The proposed model evolves over an invariant, directed, signed and unweighted network. Signed edges are interpreted as in structural balance theory [5, 218, 36]: an edge from agent j to agent i is positive if agent i approves, trusts, or follows agent j , whereas it is negative if agent i disapproves, distrusts, or antagonises agent j .

Despite significant research efforts in the development and analysis of opinion formation models, empirical validation is often lacking, and has been identified as one of the frontiers of opinion modelling [75]. In most cases, just an analytical or numerical characterisation of possible opinion evolutions is provided and, with some exceptions (most notably the Friedkin-Johnsen model [79, 81]), there are no systematic comparisons with real world behaviours.

The problem of identifying individual-level parameters (in the model we propose, the agent inner traits, see Section 7.2.3) from population-level data (survey results in the problem considered in Section 7.2.3) is known as the *inverse problem* [126] and arises, in the context of opinion dynamics, for any agent-based model, also when estimating agent interactions [155] and underlying networks [108] from data. An approach to solve the inverse problem using survey results relies on evolutionary algorithms [65]; other papers taking into account survey results or empirical data in the study of opinion formation models include those by Banisch and Shamon [18], Chattoe-Brown [42], Baumann, Lorenz-Spreen, Sokolov, and Starnini [23], and Martins [156].

In Chapter 7, we describe our Classification-based model and we assess its potential to predict opinion evolution in real-life settings using data from the World Values Survey [99, 101]. Our main purpose is to present a new opinion formation model; through the comparison with real data, we identify parameter choices showing that the model has the *potential* to accurately predict real opinions starting from a variety of different initial opinions, but this does not fully or univocally solve the inverse problem [126].

5.2. GRAPHS AND OPINIONS

A directed graph (*digraph*) $\mathcal{G}$ with N vertices can be represented by a set of vertices $\mathcal{V} = \{1, \dots, N\}$ and a set of edges $\mathcal{E} \subseteq \{(i, j) : i, j \in \mathcal{V}\}$, where the pair (i, j) represents an edge from vertex j to vertex i .

In the context of opinion dynamics, the vertices correspond to individuals, each associated with a time-varying opinion (individual i at time k has opinion $x_i[k]$) and the edges correspond to interactions, each with a possibly time-varying sign and weight ($w_{ij}[k]$ represents the influence that individual j exerts over individual i at time k).

For all the considered models, the opinions will belong to the interval $[-1, 1]$, namely, $x_i \in [-1, 1]$ for all $i \in \mathcal{V}$. The edge properties (sign, weight, time-dependence) depend on the model, for details see the information on each model provided in Section 5.3.

For every model, each individual is assumed to have self-confidence $w_{ii} > 0$ (for all $i \in \mathcal{V}$), which represents the persistence of belief in its own opinion. In the remainder of the Part 2, the terms agent and vertex will be used interchangeably. All the simulations evolve over strongly connected digraphs with different topologies, including: (Generalised) Ring, Complete, Small-World, Scale-Free, Lattice, and Random graphs. More on these graph topologies will be explained in Section 5.4.

The edge weights can be arranged in a $\mathcal{V} \times \mathcal{V}$ matrix W , where the coefficient in row i , column j is $W_{i,j} = w_{i,j}$. For edges not present in $\mathcal{E}$, the weight is assumed to be zero. This representation is particularly useful when expressing opinion formation laws in matrix form, or when computing network metrics. In the remainder of the dissertation, this matrix representation of the digraph will be used. Matrix W is the weighted adjacency matrix. Every digraph has a topology, associated with the adjacency matrix with all non-zero weights set to one (in other words, an unsigned and unweighted adjacency matrix). The digraph topology provides information on how the agents are connected. For more information about the considered topologies see, Section 5.4.

5.3. OPINION FORMATION MODELS

All the opinion formation models we consider evolve in discrete time. Here we provide an overview of the opinion update rules for the considered models: French-DeGroot [179, 77, 54], Weighted-Median [163], Bounded Confidence [110], Quantum Game [55], Friedkin-Johnsen [79], and Back-fire Effect and Biased Assimilation [45]. This is not a comprehensive explanation or description of the models: for more details, see the cited papers.

5.3.1. FRENCH-DEGROOT MODEL

At every time step, the opinions of each individual are updated according to the rule:

$$x_i[t+1] = \sum_{j \in \mathcal{V}} w_{ij} x_j[t], \quad \forall i \in \mathcal{V}.$$

The associated weight adjacency matrix W is time-invariant and row stochastic, meaning that all the coefficients belong to the interval $[0, 1]$ and the row sum is 1 for every row. This has the interpretation of ‘partitioning’ the total influence or attention a given agent can have. For more details see [179, 77, 54].

5.3.2. WEIGHTED-MEDIAN MODEL

At every time step, only the opinion of a single individual is updated. This individual is chosen randomly and moves to the opinion of another individual selected as follows:

$$x_i[k+1] = x^*,$$

where $x^* \in \{x_1[k], \dots, x_N[k]\}$ is the opinion satisfying

$$\sum_{j: x_j[k] < x^*} w_{ij} \leq \frac{1}{2}, \quad \text{and} \quad \sum_{j: x_j[t] > x^*} w_{ij} \leq \frac{1}{2}.$$

If more than one opinion satisfies these inequalities, then x^* is taken as the opinion closest to $x_i[t]$.

For this model, matrix W has the same properties as the one for the French-DeGroot model. For more details see [163].

5.3.3. BOUNDED CONFIDENCE MODEL

This model is similar to the French-DeGroot model; however, at every time step, agent i is influenced by agent j if and only if $|x_i - x_j| \leq r_i$, where r_i is the confidence radius of agent i . Mathematically the model evolves according to the following equation:

$$x_i[k+1] = \sum_{j \in \mathcal{V}} w_{ij} x_j[k], \quad w_{ij}[k] = \begin{cases} \frac{1}{|N_i|} & \text{if } |x_i[k] - x_j[k]| \leq r_i \\ 0 & \text{otherwise} \end{cases} \quad \forall i \in \mathcal{V},$$

where $N_i = \{j \in \mathcal{V} : |x_i[k] - x_j[k]| \leq r_i\}$, and $|N_i|$ is its cardinality. The Bounded Confidence model can be interpreted as the French-DeGroot model where the digraph is time-varying. At each time step, matrix W is still row-stochastic. Edges can be created and removed, therefore the whole network topology changes with time, not only the edge weights. For more details see [110].

5.3.4. QUANTUM GAME MODEL

In this model, at each time step two randomly chosen agents interact pairwise. In each interaction the agents have three options: *Keep* (keep their opinion), *Change* (take the other agent opinion), and *Agree* (take an intermediate opinion). The action each agent decides to take depends on the payoff. In this quantum model the payoff matrices depend on parameters a , b , c and the initial entangled state $|\psi_{\text{in}}\rangle$. For the simulations shown in the dissertation, the values of the parameters are $a = 1$, $b = 3$, $c = 1$, and $|\psi_{\text{in}}\rangle = \sqrt{1/2}|00\rangle + (1/2)|11\rangle + (1/2)|22\rangle$. Denoting by d the opinion distance between agents, the opinion formation law reduces to the following:

- If $d < 0.25$, then the unique Nash Equilibrium is to *Agree*, hence the new opinion of both agents is the mean of their previous opinion.
- If $d \geq 0.25$, then the Nash Equilibrium with greatest payoff is to *Keep*, hence the new opinion of both agents is the same as their previous opinion.

For more details, see the referenced paper.

This type of model does not have an underlying digraph: two agents are chosen at random to interact, without regard of a possible underlying network structure.

5.3.5. FRIEDKIN-JOHNSEN MODEL

This model generalises the French-DeGroot model, by including for each agent a parameter called *susceptibility*. The susceptibility of agent $i \in \mathcal{V}$ is denoted by $\lambda_i \in [0, 1]$. According to the Friedkin-Johnsen model, the opinion of agent i evolves as

$$x_i[k+1] = \lambda_i \left(\sum_{j \in \mathcal{V}} w_{ij} x_j[k] \right) + (1 - \lambda_i) u_i, \quad \forall i \in \mathcal{V},$$

where u_i is called the *prejudice* of agent i . Throughout this dissertation, $u_i = x_i[0]$. If $\lambda_i = 1$ for all agents, then we get the French-DeGroot model. Correspondingly, the digraphs used in the Friedkin-Johnsen model are time-invariant and row-stochastic. For more details see [79].

5.3.6. BACKFIRE EFFECT AND BIASED ASSIMILATION MODEL

This model is also a generalisation of the French-DeGroot model, modified to represent backfire effect and biased assimilation. Backfire effect occurs when, after agents i and j communicate, their resulting opinions are more distant than before their interaction. Biased assimilation is a form of homophily, in that it represents the tendency to be more influenced by neighbours that think like us.

Both these mechanisms are represented by the *entrenchment* parameter, a positive number associated with each agent: $\beta_i > 0$. If β_i is small then it means that the agent's backfire and biased assimilation tendencies are low, and as β_i increases so do the strength of these tendencies.

The opinion of agent i evolves as follows:

$$x_i[k+1] = \begin{cases} \text{sgn}(x_i[k]) & \text{if } \sum_{j \in \mathcal{V}} w_{ij}[k] \leq 0 \\ \frac{\sum_{j \in \mathcal{V}} w_{ij}[k] x_j[k]}{\sum_{j \in \mathcal{V}} w_{ij}[k]} & \text{otherwise} \end{cases}$$

and the edge weights evolve as follows:

$$w_{ij}[k] = \begin{cases} \beta_i x_i[k] x_j[k] + 1 & \text{if } w_{ij}[0] \neq 0, \\ 0 & \text{if } w_{ij}[0] = 0. \end{cases}$$

Hence, edges that were not present in the initial digraph cannot be added, and the edges that were present may change sign and weight (for some specific opinions, the edge could become zero, effectively removing the edge for that time step), meaning that the *topology* remains approximately the same. For more details see [45].

5.4. DIGRAPH TOPOLOGY

As previously mentioned, the topology of a digraph determines how the agents are connected among themselves. In Part 2 of the dissertation, only digraphs with one of the following topologies are considered:

1. **(Generalised) Ring:** a Ring digraph is the same as a Directed Cycle graph. In these digraphs, all the agents can be arranged in a circle and connected to the agent in front (moving clockwise). We can generalise this concept in several ways, for instance connecting to the N agents in front or connecting to *some* of the agents in front. The connections can also be made bidirectional. If all the connections are bidirectional, we say that the digraph is undirected. In this dissertation, all Ring topologies are generalised: a Ring topology can always be a Generalised Ring topology. Examples of these digraphs can be seen in Figures 5.1b, 5.1c, and 5.1d.
2. **Complete:** every agent is connected to all the others. Equivalently, matrix W has no zero coefficients. See Figure 5.1a.
3. **Small-World:** this type of topology is built starting with a Generalised Ring and then going through all the edges and changing their end-point with a given probability. This is an extension of the Watts-Strogatz mechanism to construct directed graphs. A high clustering coefficient and low average path length characterise these topologies. Some examples can be seen in Figures 5.1e and 5.1i.
4. **Scale-Free:** this type of topology is built starting with an initially connected digraph. As new vertices are added, these are connected to the existing vertices with a probability proportional to the in-degree of the existing nodes (this is the Barabási-Albert model). In these topologies, the clustering coefficient follows a power-law of the number of vertices. For an example, see Figure 5.1h.

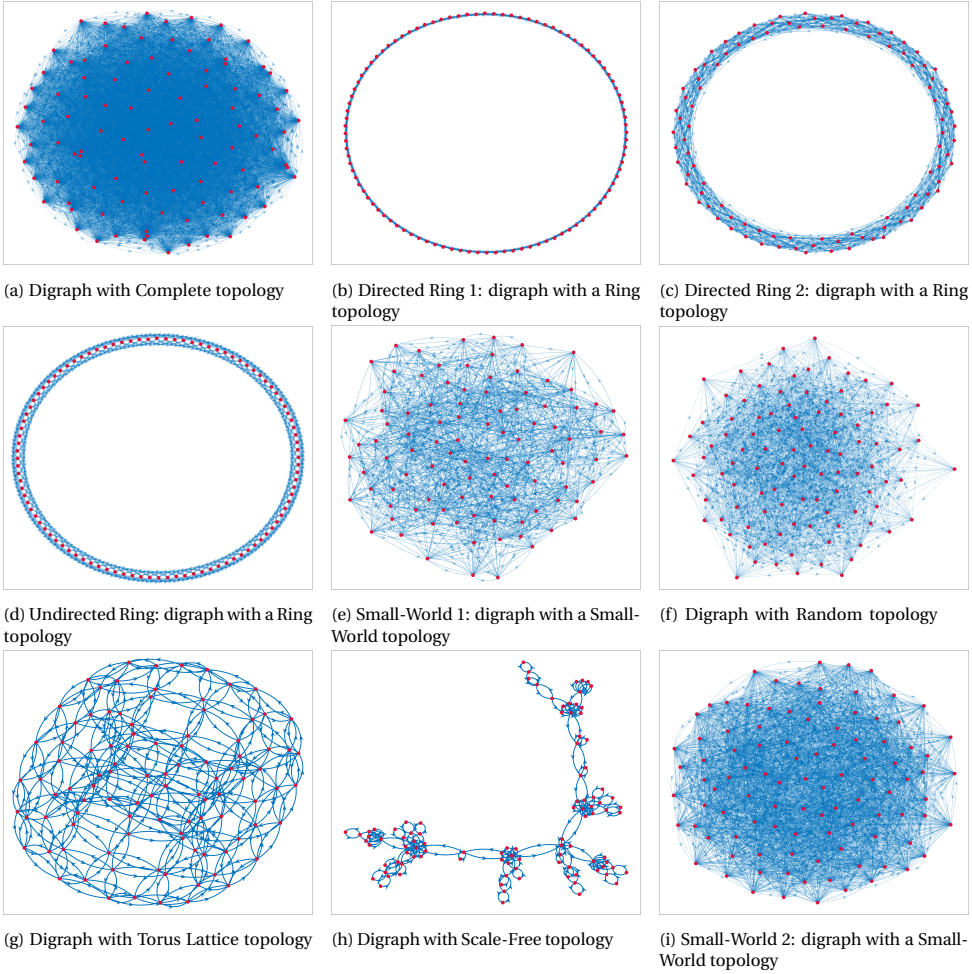


Figure 5.1: Examples of digraphs with the topologies considered in Part 2. Some topology metrics of these digraphs are presented in Table 5.1. An explanation about how these metrics are computed can be found in Appendix A (Section 7.4). All the presented digraphs have 100 vertices. All the simulation results presented in Part 2 of this dissertation use 100 agents, except for Sections 6.1 and 6.2, for more information see Section 5.5.4.

5. **(Torus) Lattice:** Lattice topologies follow a regular pattern. For every Lattice digraph in this dissertation, vertices are arranged in a square grid and connected to the nearest four neighbours. A standard regular lattice has a boundary, and the agents at the border have fewer neighbours. An option to solve this is to connect the boundaries so that the system is ‘closed’ and every agent has the same number of connections. We call the result a torus lattice. For an example, see Figure 5.1g.
6. **Random:** as the name implies, in these topologies, the edges are created randomly. Every possible edge exists with a given probability. See Figure 5.1f.

Figure 5.1 shows examples of digraphs with the topologies mentioned earlier.

Digraph	APL	D	$\overline{CC}$	$\sigma(CC)$	#E	$\overline{\delta^{out}}$	$\sigma(\delta^{out})$	$\overline{\delta^{in}}$	$\sigma(\delta^{in})$	BC
Complete	1	1	1	0	10000	100	0	100	0	1
Directed Ring 1	10.404	20	0.5	0	600	6	0	6	0	0
Directed Ring 2	5.4545	10	0.5	0	1100	11	0	11	0	0
Undirected Ring	5.4545	10	0.66667	0	1100	11	0	11	0	1
Small-World 1	2.0365	3	0.17538	0.0020628	1435	14.35	6.553	14.35	6.553	0.69513
Random Network	1.8129	3	0.20122	0.00057076	2083	20.83	17.4961	20.83	15.7385	0.21785
Torus Lattice	5.0505	10	0	0	500	5	0	5	0	1
Scale-Free	5.6893	14	0.67	0.22333	298	2.98	5.0905	2.98	5.0905	1
Small-World 2	1.6778	2	0.32651	0.00036845	3290	32.9	15.2828	32.9	14.6162	0.65078

Table 5.1: Metrics for the different digraph topologies considered in Sections 5.4 and 6.3, and shown in Figure 5.1. The reported metrics are Average Path Length (APL), Diameter (D), Mean clustering ($\overline{CC}$), Clustering variance ($\sigma(CC)$), Number of edges (#E), Mean out-degree ($\overline{\delta^{out}}$), Out-degree variance ($\sigma(\delta^{out})$), Mean in-degree ($\overline{\delta^{in}}$), In-degree variance ($\sigma(\delta^{in})$), and Bidirectional Coefficient (BC). All the metrics correspond to digraphs with 100 vertices. All the simulation results presented in Part 2 of this dissertation use 100 agents, except for Sections 6.1 and 6.2, for more information see Section 5.5.4.

5

Although we consider digraphs with all these topologies in Part 2 of the dissertation, we use only Small-World topologies when investigating the model capabilities to predict opinion evolutions seen in real life (Section 7.2.3). A significant body of literature backs up the decision to evolve the models over Small-World networks [134, 120, 86], because the two main features of Small-World networks (high clustering coefficient and low average path distance) are typically observed in real-life interactions.

As mentioned, the various topologies have specific features that can be highlighted by computing some metrics, such as the clustering coefficient, average path length, and diameter. Appendix A (Section 7.4) explains how we calculated these metrics. Table 5.1 reports the values of these metrics for examples of networks with the previously described topologies.

A brief explanation of the meaning of these metrics is as follows:

- Average path length (APL): the average number of edges connecting any two vertices. It is a positive number greater than or equal to one. The only topology for which it is one is the Complete topology, where every agent is connected to all the others. In general, the greater this number, the more dispersed the agents are.
- Diameter (D): the maximum length of a path connecting any two agents. It is a positive integer. As with the APL, the greater this number, the more dispersed the agents are.
- Clustering mean and variance ($\overline{CC}$ and $\sigma(CC)$ respectively): each agent has an individual clustering coefficient, which measures the proportion of neighbours that are neighbours among them. Since each agent has a single coefficient, it is possible to compute the mean and variance of these numbers to obtain the clustering mean and variance. The mean of the agent's clustering coefficients is also called the *clustering mean* and measures how interconnected the agents are. It is a number between 0 and 1, where 0 means no two connected agents have a common neighbour and 1 means complete connectivity (this can only happen in the Complete topology). The clustering variance measures how diverse are the individual clustering coefficients in the network. A low variance indicates a largely homogeneous network, whereas a high variance means a diverse network.
- Number of edges (#E): it is the number of edges of the digraph. It is an indirect measure of how connected the digraph is.
- Mean degree ($\overline{\delta^{out}}$ and $\overline{\delta^{in}}$): each agent has an individual in-degree and out-degree, the number of edges that point to the agent or that exit the agent, respectively. The two mean

degrees give information on how interconnected the agents are. If edges are allowed to enter and exit vertices ‘outside the system’, these two numbers can be different. If every edge is constrained to enter or leave one of the nodes in the system, both numbers are equal.

- Degree variance ($\sigma(\delta^{out})$ and $\sigma(\delta^{in})$): the variance of in- and out-degrees provides an idea of how distributed the influences are in the network. If the variance is low, every agent communicates with approximately the same number of agents; thus, they all have comparable influence. On the contrary, if a subgroup of agents has a much higher degree, they can be perceived as highly influential over the network.
- Bidirectional coefficient (BC): the proportion of edges for which there exists an edge connecting the same vertices in the opposite direction.

It is essential to distinguish between *topology metrics* and *digraph metrics*. Every digraph has a topology; therefore, it is possible to compute topology metrics for every digraph. However, because the edges in a digraph can have signs and weights (unlike in a topology), not every digraph metric is topological. For instance, if the edges are signed, it is possible to compute the balance index. Because different signed digraphs with the same topology can have distinct balance indices, the balance index is a digraph metric but not a topological metric. Every metric shown in Table 5.1 is topological.

5

5.5. ADDITIONAL NOTES ON THE SIMULATION RESULTS

There are five possible sources of confusion that we take the opportunity to clarify now:

5.5.1. MODEL NAMES

In Part 2 of the dissertation 7 opinion formation modes are analysed. Besides their complete name, all of them also have a shorthand name that is used mainly in tables, figures, or to make the text more legible. The names of the models and their corresponding shorthand notation are, in no particular order, French-DeGroot (FG), Friedkin-Johnsen (FJ), Quantum Game (QG), Bounded Confidence (BC), Backfire Effect and Biased Assimilation (BEBA), Weighted-Median (WM), and Classification-based (CB). The first 6 models are presented in Section 5.3, whereas the last model is introduced and explained in Chapter 7.

5.5.2. HISTOGRAM-BASED SORTING ALGORITHM DIFFERENCE

In Chapter 6 the Histogram-based Sorting Algorithm is introduced. When comparing this algorithm with the one published in [57], in the latter the last step is missing. This is because in the time between the publication of that paper and this dissertation that last step was added. This can be confusing, as the results presented both in the paper and the dissertation are the same. This is because that additional step introduces a change that is only relevant for a particular class of histograms, which is not present in the histograms analysed in [57]. Therefore the results remain the same.

5.5.3. USE OF WORLD VALUES SURVEY DATA

We used the results of the WVS data on two occasions: first, in the computation of real Transition Tables (Figure 6.3) and then in the *Model Validation with Real Data* (Section 7.2.3). In both cases, we considered the same 30 questions of the questionnaires (the questions reported in Table 7.19); however, in the two cases, we considered different countries (See Tables 7.17 and 7.18)

For the Transition Tables, we considered the answers from 25 countries. However, not all $30 \times 25 = 750$ possible combinations of question-country pairs had answers for the three waves (Waves

5, 6, and 7), which is why Figure 6.3 considers only 675 question-country pairs (only question-country pairs with answers in the three waves were used). On the other hand, Section 7.2.3 uses the answers from waves 5 and 6. In this case, we selected 26 countries such that every question-country pair had answers in these two waves. This is reflected by the fact that Table 7.24 does not have empty cells. The list of 25 (respectively, 26) countries considered in the first (resp. second) occasion is reported in Table 7.17 (resp. Table 7.18).

It is important to note that for every question-country pair the number of survey answers was not the same. To address this, and make meaningful comparisons, the opinion distributions were rescaled so that the overall distribution remained approximately the same (the form of the histogram did not change significantly) for a different number of agents. This rescaling is valid if the original histograms are assumed to be accurate representations of the society opinion, since in that case, sampling a different number of people should not significantly change the resulting opinion distributions.

5.5.4. COLLECTIONS OF UNDERLYING DIGRAPHS

Different reported simulation results use separate collections of underlying digraphs. Here we distinguish them:

- To construct the Transition Tables reported in Figure 6.4, 15 initial digraphs were used. Of these 15 digraphs, 5 have 100 vertices, 5 have 500 vertices, and 5 have 1000 vertices. All digraphs are strongly connected and were constructed based on the Watts-Strogatz Small-World graph model. All the associated matrices are row-stochastic.
- For all the results presented in Section 6.3, a collection of 45 different digraphs were used (each with 100 vertices). This collection can be separated into 9 groups of 5 digraphs. Every digraph in each group has the same topology and differs only in the edge weights (all the associated matrices are row-stochastic, since the model is Friedkin-Johnsen). The considered topologies are Complete, Directed Ring (two different ones), Undirected Ring, Small-World (two different ones), Random, Torus Lattice, and Scale-Free. We present some metrics for these topologies in Table 5.1.
- The results presented in Section 6.4 use a collection of 1000 different strongly connected digraphs with Small-World network topology (each with 100 vertices). Figure 6.14 shows some of the characteristics of these digraphs.
- In Section 6.5.1, the Agreement Plot analysis uses 9 digraphs (each with 100 vertices), which have the same topology as the digraphs in Section 6.3. This is because the digraph topology for the BEBA model is constant, while the weights and signs can change; therefore, for each topology, only one digraph is necessary. The Probability Analysis in Section 6.5.1 uses a collection of 1000 digraphs with the same topology as the collection used in Section 6.4.
- In Section 7.2.3, we use a collection of 35 different digraphs (each with 100 vertices). All the digraphs are strongly connected and have Small-World topology. This collection can be divided into 7 groups, each containing digraphs with the same topology and different edge signs. We report the metrics for this collection of 35 digraphs in Table 7.1.
- In Section 7.2.5, the Agreement Plot analysis uses 45 digraphs (each with 100 vertices) which can be divided into 9 groups, each with the same topology as the digraphs in Section 6.3. The Probability Analysis in Section 7.2.5 uses a collection of 1000 digraphs with the same topology as the collection used in Section 6.4, and the considered digraphs are signed and unweighted.

5.5.5. AGENT OPINION AND PARAMETER ALLOCATION

When simulating an opinion formation model with N agents, five elements are required:

1. *Interaction network*: it encodes how the N agents communicate. The interaction network is represented by a graph with N vertices, where each vertex corresponds to an agent. In general, we denote this graph by $\mathcal{G}$. In the Friedkin-Johnsen case, it is a directed, unsigned, weighted graph associated with a $N \times N$ row-stochastic adjacency matrix.
2. *Set of initial opinions*: a set of N real numbers representing the initial opinions of the population.
3. *Set of agent parameters*: a set of N tuples, each corresponding to the parameters describing the behaviour of agent. In the Friedkin-Johnsen model, it is a set of N numbers in the interval $[0, 1]$ representing the agent's susceptibility.
4. *Initial opinion allocation*: it dictates which initial opinion corresponds to each agent, by means of a one-to-one correspondence between elements in the initial opinion set and vertices in $\mathcal{V}$.
5. *Agent parameter allocation*: it designates which parameters each agent has, by means of a one-to-one correspondence between elements in the agent parameters set and vertices in $\mathcal{V}$.

Usually, the initial opinion and agent parameter allocations are implicit: in code implementations, the vertices, by default, have an ordering and managing the sets of initial opinions and agent parameters using vectors also gives them an order. So the allocation is simply the identity. Also, when referring to initial opinions and agent parameters, we do not mention allocation since it is assumed that the opinions and parameters are assigned to the network vertices in some irrelevant way. Opinion allocation is not limited to the initial opinions but has the same interpretation throughout the complete opinion evolution process: opinion allocation is the association between opinion values and vertices in the network structure.

The allocation, however, is an essential piece of information, since it can radically change the outcome of an opinion formation model. Besides being critically important, measuring or estimating it in real societies is exceptionally challenging. This is why we mention it in Section 6.4 in the Probability analysis, as it typically is part of the information not available when making predictions on the qualitative outcomes of opinion formation models.

6

ANALYSIS OF AGENT-BASED OPINION FORMATION MODELS

Choose not to be harmed, and you won't feel harmed.
Don't feel harmed and you haven't been.

Marcus Aurelius

This chapter introduces a methodology to analyse the evolution of opinion formation models. The proposed methodology is based on four techniques: Histogram-based Sorting Algorithm, Transition Tables, Agreement Plot, and Probabilistic Analysis. Collectively, these techniques allow for a comprehensive qualitative characterisation of agent-based opinion formation models. This characterisation includes the opinion transformations the model predicts, the capability of the model to produce different qualitative types of opinion distributions, and their relation with real data. Furthermore, the proposed methodology can be applied also to real opinion data, to gain insight into how opinions evolve in real societies.

The chapter starts by introducing the Histogram-based Sorting Algorithm, which is a way to classify sets of opinions into one of five qualitative opinion categories that were identified from the World Values Survey data: perfect consensus, consensus, polarisation, clustering, or dissensus. This sorting allows for an intuitive interpretation and simplified handling of opinion sets made up of possibly hundreds of numbers. Next, the sorting algorithm is applied in the Transition Tables, which capture the qualitative evolution of the opinion distribution between an initial and a final time. The use of Transition Tables for survey data reveals patterns in the evolution of real opinions and their comparison with model predictions allows for improvement and validation of the model.

Following the Transition Tables, the concept of Agreement Plot is introduced. This third technique is an intuitive visual representation of a collection of opinion sets, which can be used in a variety of

This chapter is based on the paper “A framework to analyze opinion formation models” (2022) [57], by Carlos Andrés Devia and Giulia Giordano, and on the work-in-progress paper “Graphical and Probabilistic Analysis of Agent-Based Opinion Formation Models” [59], by Carlos Andrés Devia and Giulia Giordano

ways. It can provide further insight into the sorting algorithm by showing that the different qualitative categories can be seen too in the Agreement Plot. It visualises 'where' real opinions can be found, and shows which opinions a model can produce. The fourth technique, Probabilistic Analysis, aims to describe the possible transitions an opinion distribution may undertake when only incomplete information on the opinions, agent parameters, and network is known.

All the techniques are applied to a variety of opinion formation models, from classical models such as the French-DeGroot, Friedkin-Johnsen, Bounded Confidence model, to more recent models such as the Weighted-Median, Quantum Game, and Backfire Effect and Biased Assimilation model. The results presented in this chapter not only illustrate the use of the presented techniques but also provide insight into the evolution of real-life opinions and highlight key directions to improve opinion formation models.

6.1. OPINION DISTRIBUTION CLASSES AND TRANSITION TABLES

We denote as *opinion* the level of agreement with a statement. The opinions that an individual can have belong to the interval $[-1, 1]$, where the values -1 , 0 , and 1 respectively denote complete disagreement, indifference, and complete agreement with the statement. Given a population of N individuals, each having an opinion about a statement, the set containing the opinions of all the individuals in the population is called an *opinion distribution*, which belongs to one of our identified qualitative categories, exemplified in Figure 6.1: perfect consensus, consensus, polarisation, clustering, and dissensus. Their mathematical definitions, provided in the following, are inspired by these informal definitions (where by majority we denote more than 50% of the population):

- *perfect consensus*: the majority chooses the very same opinion.
- *consensus*: the majority chooses approximately the same opinion.
- *polarisation*: the majority is split between two 'distant' opinions.
- *clustering*: the majority is split between two or more groups.
- *dissensus*: the majority of the opinions are uniformly distributed.

These categories of opinion distributions capture an increasing level of inhomogeneity. When all the individuals have the exact same opinion (perfect consensus), there is perfect homogeneity. Starting from perfect consensus, progressively increasing inhomogeneity leads to consensus, polarisation, clustering, and lastly dissensus. When every opinion is held by the same number of people (perfect dissensus), inhomogeneity is maximal and no preference whatsoever can be identified.

6

6.1.1. OPINION DISTRIBUTION SORTING

An opinion distribution can be visualised as a histogram with M bins, which can then be sorted (namely, "classified") so as to determine to which qualitative category the opinion distribution belongs. This process is performed by our proposed *Histogram-based Sorting Algorithm*:

1. Input the positive integers M , $B < M$ and $K \leq M - 2$, and the thresholds T_1 , T_2 with $0 < T_2 < 50 \leq T_1 < 100$.
2. Partition the $[-1, 1]$ interval in M bins of equal width.
3. Count how many opinions fall in each bin. Denote by $H(k)$ the number of opinions in bin k ($1 \leq k \leq M$).
4. Normalise the bin counts so they add up to 100. Denote the normalised bin counts by $\tilde{H}(k)$.
5. Classify each bin as *green*, *blue*, or *red*: bin k is green if $\tilde{H}(k) > T_1$; blue if $\tilde{H}(k) < T_2$; red otherwise.
6. Compute the *number of groups*; a group is formed by consecutive green or red bins. For each group, compute the *number of bins*, and the *normalised group count*, which is the sum of all the normalised bin counts of the bins belonging to the group.
7. Sort the histogram according to the following criteria:
 - *perfect consensus* if there is a green bin;
 - *consensus* if there is one group, with at most B bins, and with normalised group count larger than 50;
 - *polarisation* if there are two groups, each with at most B bins, with at least K bins in between, whose normalised group counts add up to more than 50;
 - *clustering* if there are two or more groups, each with at most B bins, whose normalised group counts add up to more than 50;
 - *dissensus* otherwise.

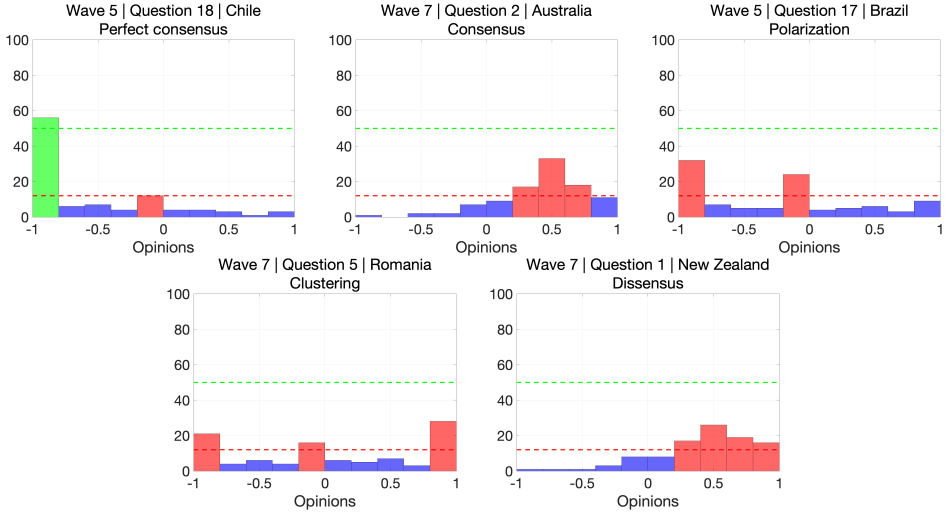


Figure 6.1: Examples of real-life opinion distribution histograms, taken from the World Values Survey data, sorted as perfect consensus, consensus, polarisation, clustering, and dissensus by the proposed *Histogram-based Sorting Algorithm* (the title shows the corresponding wave, question number, country, and type). The vertical axis represents the normalised bin counts $\bar{H}$. The dotted lines mark thresholds T_1 (green) and T_2 (red). The bins are coloured according to the Histogram Sorting Algorithm: green if the normalised bin count is larger than T_1 ; blue if it is smaller than T_2 ; red if it is between T_1 and T_2 . The parameter values we adopted to perform the sorting are: $M = 10$, $B = 3$, $K = 4$, $Y = 6$, $T_1 = 50$, $T_2 = 12$, and $T_3 = 40$.

8. The only exception to the previous criteria is if the histogram has only two non-empty bins with at least Y empty bins in between and each of these two non-empty bins has a normalised group count larger than T_3 . Then the opinion distribution is also classified as polarisation.

The last exception has been added because, when the histogram has only two non-empty bins, it will always be classified as perfect consensus, except in the rare occasion in which both bins have the exact same height (i.e. 50). However, intuitively, if there are only two non-empty bins at a significant distance, of comparable height, the histogram should be classified as polarisation.

The parameters M , B , K , Y , T_1 , T_2 , and T_3 allow the proposed sorting to be tuned according to the problem at hand, thus taking into account possible differences in the interpretation of our proposed qualitative categories.

In our case study, $M = 10$ is a natural choice, since the data from the World Values Survey [99, 101, 100] comes from Likert 10-scale questions. Parameter B represents the ‘level of closeness’ required to state that a group of people share a ‘similar’ opinion: we set $B = 3$. Polarisation is defined as the presence of two groups with significantly opposing views, the ‘level of opposition’ is encoded by the parameter K . Two groups at a distance less than K would represent clustering, since the opinions are not very different, while two groups at a distance K or more represent two opposing views and hence polarisation. The value $K = M - 2$ would mean that extreme opposition is needed to define polarisation; in this dissertation, we choose $K = 4$. The threshold T_1 defines perfect consensus: we choose $T_1 = 50$ to capture all instances where the majority (more than 50%) shares a single opinion. The threshold T_2 discriminates between significantly numerous opinion groups and ‘white noise’ residual opinions. A low T_2 leads to the appearance of multiple groups with more than B bins, while a high T_2 leads to interpreting significant opinion groups as white

noise: in both cases, the sorting is biased towards dissensus. After repeated numerical experiments, the intermediate value of $T_2 = 12$ was selected and can be seen as a robust choice, because varying T_2 between 10 and 14 gave comparable classification results. For the polarisation exception, the values $Y = 6$, and $T_3 = 40$ were chosen, because with these values we interpret the corresponding histograms as polarisation. The parameter T_2 is the only one for which a sensitivity analysis was done, given that the other parameters have a more straightforward interpretation and choice of value.

Figure 6.1 shows examples of real-life opinion distribution histograms, taken from the World Values Survey data, representative of our proposed qualitative categories.

From Figure 6.1 it can also be seen that histograms with limited difference may be categorised in substantially different categories (for instance, Consensus and Dissensus). This is an inevitable consequence of the sorting approach. Since the possible categories are much less than the number of possible histograms, one could think of the histograms as being ‘continuous’ and the sorting generates a partition of this ‘continuous’ set. As a consequence, two elements that are arbitrarily close can belong to two different subsets, because the algorithm is very sensitive to small variations regarding some features. Although this is not a desirable feature, it is unavoidable when associating a unique label with each histogram, which is the only way to implement the Transition Tables (as will be explained in the next section). A different approach would be to compute a ‘grade’ or coefficient for each category and for each histogram, expressing how much the histogram fits into the category; in this scenario, a single histogram has five measures of how well it represents each category. This approach would be less sensitive, but since it does not categorise the histograms, it cannot be used in order to build Transition Tables.

6.1.2. MODEL PREDICTIONS VS. REAL OPINIONS: A FRAMEWORK FOR SYSTEMATIC COMPARISON

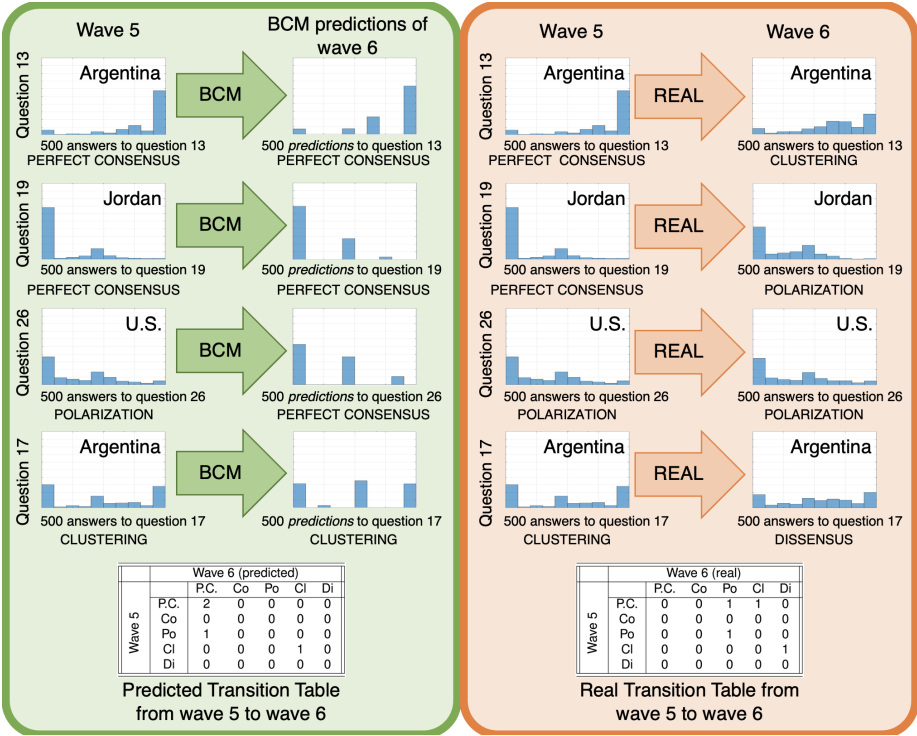
The proposed *histogram-based sorting* approach allows us to systematically associate a given opinion distribution, which can be either real (e.g. survey data) or predicted by an opinion formation model, with a qualitative category. An opinion distribution is a static snapshot; to study opinion formation, we need to understand how opinion distributions can evolve over time. We introduce *Transition Tables* to capture the possible qualitative categories of final opinion distribution that can be obtained, after some time, starting from various categories of initial opinion distribution. A transition table is a matrix whose rows (respectively, columns) are associated with the qualitative category of initial (respectively, final) opinion distribution: entry (i, j) represents the number of opinion distributions that evolve from an initial configuration belonging to category i to a final configuration belonging to category j , where i and j can be either perfect consensus, consensus, polarisation, clustering, or dissensus. To systematically compare the outcome of a given opinion formation model with real opinion data collected at two different time instants, we proceed as follows:

1. classify the *real* initial opinions,
2. let them evolve according to the opinion formation model, and produce the *predicted* final opinions,
3. sort the *predicted* final opinions,
4. using the sorting of *real* initial opinions and *predicted* final opinions, construct the *predicted transition table*,
5. sort the *real* final opinions,
6. using the sorting of *real* initial opinions and *real* final opinions, construct the *real transition table*, and
7. compare the two Transition Tables.

As an example, we assess the Bounded Confidence model (BC) [110, 53, 52, 215, 216], with confidence radius 0.3 (all agents have the same confidence radius), along with the answers, pro-

vided by 500 people, to four questions of the World Values Survey both in wave 5 and wave 6. The four initial opinion distributions (wave 5) are sorted by our algorithm as perfect consensus, perfect consensus, polarisation, and clustering, respectively. Taking these opinion distributions as initial conditions, the Bounded Confidence model yields *predicted* opinion distributions that our algorithm respectively sorts as perfect consensus, perfect consensus, perfect consensus, and clustering. Hence, two opinion distributions are predicted to remain perfect consensus, one to change from polarisation to perfect consensus, and one to remain clustering, as summarised in the *predicted* transition table in Figure 6.2 (left). The *real* transition table in Figure 6.2 (right) can be constructed by considering the *real* final opinion distributions (wave 6) for the same four questions, which are sorted as clustering, polarisation, polarisation, and dissensus, respectively.

Comparing real and predicted Transition Tables allows us to evaluate the model, identify its shortcomings and suggest ways to improve its realism. Furthermore, real Transition Tables provide qualitative understanding of how the actual opinion distributions can evolve within the population in the considered time interval.



Question 17

Argentina

500 answers to question 17

CLUSTERING

BCM

Question 17

Argentina

500 predictions to question 17

CLUSTERING

		Wave 6 (predicted)				
		P.C.	Co	Po	Cl	Di
Wave 5	P.C.	2	0	0	0	0
	Co	0	0	0	0	0
	Po	1	0	0	0	0
	Cl	0	0	0	1	0
Di	0	0	0	0	0	

Predicted Transition Table
from wave 5 to wave 6

Wave 5

Question 13

Argentina

500 answers to question 13

PERFECT CONSENSUS

REAL

Wave 6

Question 13

Argentina

500 answers to question 13

CLUSTERING

Question 19

Jordan

500 answers to question 19

PERFECT CONSENSUS

REAL

Question 19

Jordan

500 answers to question 19

POLARIZATION

Question 26

U.S.

500 answers to question 26

POLARIZATION

REAL

Question 26

U.S.

500 answers to question 26

POLARIZATION

Question 17

Argentina

500 answers to question 17

CLUSTERING

REAL

Question 17

Argentina

500 answers to question 17

DISSENSUS

		Wave 6 (real)				
		P.C.	Co	Po	Cl	Di
Wave 5	P.C.	0	0	1	1	0
	Co	0	0	0	0	0
	Po	0	0	1	0	0
	Cl	0	0	0	0	1
Di	0	0	0	0	0	

Real Transition Table from
wave 5 to wave 6

Figure 6.2: Simple example to illustrate the proposed approach: based on the answers to 4 questions administered to 500 people in two consecutive survey waves, the accuracy of the Bounded Confidence Model (BC) can be assessed by comparing the *predicted* transition table (left) with the *real* transition table (right). In the tables P.C. is perfect consensus, Co is consensus, Po is polarisation, Cl is clustering, and Di is dissensus. The BC model was selected for this example because it is a widely used and well-studied model; any other model could have been chosen.

6.1.3. REAL TRANSITION TABLES HIGHLIGHT FEATURES OF OPINION EVOLUTION

We analysed in total 2025 real opinion distributions, corresponding to World Values Survey answers to 30 questions in 25 countries in waves 5, 6, and 7, approximately separated by 5 years [99, 101, 100]; not all questions were asked in all countries, hence there are 675 opinion distributions for each wave. The orange panels in Figure 6.3 show the qualitative sorting of all the opinion distributions in each wave. The number of opinion distributions belonging to each qualitative category does not change significantly in different waves and a recurrent pattern can be observed: dissensus is consistently the most common outcome, followed by perfect consensus and then by clustering, by consensus and finally by polarisation, which is invariably the least common outcome. Figure 6.3 also reports the *real* Transition Tables from wave 5 to 6, and 6 to 7, which evidence that, in spite of the observed recurring pattern, opinion distributions themselves do not tend to remain in the same category. On the contrary, there are several examples of opinion distributions that move from a category to almost any of the others: the real Transition Tables indicate that, in real life, opinion distributions can evolve from any category to any other. The likelihood of evolving towards a different qualitative category can be assessed by comparing the sum of diagonal and off-diagonal entries in the Transition Tables: from wave 5 to wave 6, these numbers are 368 and 307 respectively, indicating that around 45% of the opinion distributions move to a different qualitative category; from wave 6 to wave 7, these numbers are 381 and 294 respectively, hence the probability of change has decreased to roughly 44%. Interesting similarities emerge in the patterns of the two Transition Tables: corresponding entries often have close values, or at least the same order of magnitude, which seems to suggest that the likelihood of evolving from a qualitative category to another changes slowly over time.

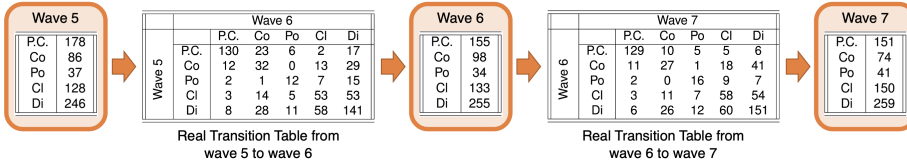


Figure 6.3: Real Transition Tables: The 675 real opinion distributions emerging from the World Values Survey [99, 101, 100] waves 5, 6, and 7 are qualitatively classified as perfect consensus (P.C.), consensus (Co), polarisation (Po), clustering (Cl), and dissensus (Di) in the orange panels. The Transition Tables show the qualitative evolution of opinion distributions between these waves, highlighting how each qualitative category of opinion distributions could evolve into the various other categories.

6.1.4. PREDICTED TRANSITION TABLES AND MODEL COMPARISON

Starting from the opinion distributions in wave 5 and wave 6, we generated the next wave results predicted by six different opinion formation models: French-DeGroot (FG), Weighted-Median (WM), Bounded Confidence with confidence radius 0.1 (BC1), 0.3 (BC2), and 0.7 (BC3) (all the agents have the same confidence radius), and Quantum Game (QG). Figure 6.4 shows the obtained *average* predicted Transition Tables, from wave 5 to 6 (left) and from wave 6 to 7 (right): each reported transition table is the average over 75 tables, computed for different population size, directed graph topology, and initial opinion assignment. The shade of blue quantifies the variability across the 75 tables represented by the difference between the maximum and minimum value for each cell across all 75 Transition Tables.

All simulations of the French-DeGroot and Bounded Confidence models evolved for 50 time steps, since the time difference between waves is about 5 years, and it is assumed that the agents

can update their opinions regarding values about 10 times per year. For the Weighted-Median (respectively, Quantum Game), since at each time step only one agent (resp. two agents) update the opinion, the number of time steps is $50N$ (resp. $\frac{1}{2}50N$), where N is the number of agents. This is so that *on average*, each agent updates its opinion 50 times during the complete simulation. This limitation on time steps was implemented so that the opinion transitions are comparable.

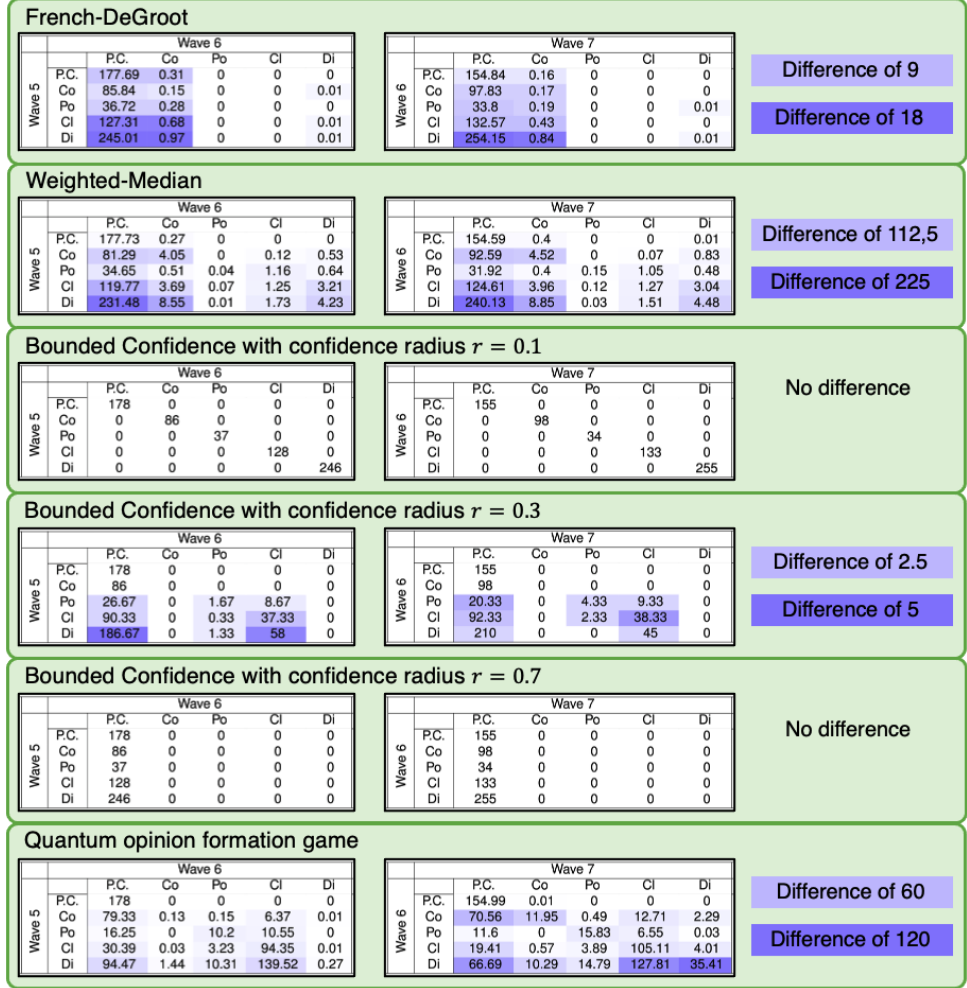


Figure 6.4: Average predicted Transition Tables from wave 5 to 6 (left), and 6 to 7 (right) for the six considered models. Each table entry contains the average of the corresponding values in the 75 computed Transition Tables, obtained for different population sizes, graph topologies, and initial opinion assignments. The variability of these values, in terms of the difference between the maximum and minimum value across all 75 tables, is denoted by the shade of blue. A cell with light blue colour represents half the maximum difference for that model, and a cell with dark blue colour represents the maximum difference for the model. The value of the difference represented by these shades of blue is reported to the right. A white cell means that all values are identical for all 75 tables. Summing the rows (respectively, columns) results in the *average* starting (resp. ending) number of histograms in each category.

The French-DeGroot (FG) model behaves as expected, with a clear trend towards perfect consensus, evidenced not only by the average transitions but also by the low difference. However, it is interesting to note that not all final opinions result in perfect consensus: some produce consensus, and in exceptional circumstances dissensus. There are two explanations for these cases. First, if the opinion towards which the agents converge is in the middle of two histogram bins, it may happen that the two adjacent bins to that converging opinion are equally populated, thus resulting in consensus and not perfect consensus. The second reason is that, in large interaction graphs it takes more time to achieve perfect consensus because, even if the graph is strongly connected, only few edges may be responsible for that strong connectivity, thus the graph could have two or more 'pseudo clusters'.

The Bounded Confidence models with confidence radius $r = 0.1$ (BC1) and $r = 0.7$ (BC3) also behave as expected: for the BC1 model, the confidence radius is so small that most of the edges vanish in the first step and then the opinions remain the same, hence the transition table is a diagonal matrix. On the other hand, the BC3 model produces exclusively perfect consensus. The reason why BC3 produces perfect consensus always, while the same does not happen with FG, is that BC3 creates more edges in the interaction graph. Hence it is possible to assume that after a few time steps the interaction graph is Complete, and then the convergence to a single opinion is much faster, resulting in perfect consensus. It is also interesting to note that BC1 and BC3 are the models showing the smallest variability (difference of zero, see Figure 6.4).

The Bounded Confidence model with an intermediate value of confidence radius $r = 0.3$ (BC2) is biased towards perfect consensus, but allows some instances of polarisation, clustering and dissensus to evolve into polarisation or clustering. A larger confidence radius (with respect to BC1) yields strongly connected subgroups of individuals that achieve perfect consensus among them: if there are only two subgroups with sufficiently distant opinions, polarisation occurs, otherwise the model produces clustering, which is the most likely outcome of the two. No consensus outcomes are generated, because, once the opinions are sufficiently close, they evolve into perfect consensus. For this model, varying population size, graph topology, and initial opinion assignment appears to have little impact, as seen in a maximum difference of 5.

The Weighted Median (WM) model exhibits a very rich behaviour. Although biased towards perfect consensus, it can produce every transition except the ones from perfect consensus and consensus to polarisation, and from perfect consensus to clustering. This wide range of outcomes is accompanied by a high sensitivity with respect to varying population size, graph topology, and initial opinion assignment, which is the highest across all considered models. The bias towards perfect consensus is expected, given the fact that the WM model is based on the cognitive dissonance theory and conformist tendencies.

Finally, the Quantum Game (QG) model presents a very interesting transition table. There is a tendency to consensus, which is coherent with the fact that agents can only *Change*, *Keep*, or *Agree*. Hence there is no disagreement mechanism. However, the randomness with which agents are chosen to interact, and the dependence of the payoff matrices on the opinion distance, produces also a clustering behaviour (a bounded confidence effect). The result is that, when the initial distribution is perfect consensus, most agents will interact with each other but the final opinion will be almost the same, resulting in perfect consensus; when the initial distribution is of consensus type, then there is a tendency to stay perfect consensus, but the agents that are not in the consensus bins can move other agents away resulting in some clustering; starting from polarisation there is a greater chance of producing polarisation or clustering due to the bounded confidence effect; this pattern is also present when the initial conditions correspond to clustering and dissensus. Another interesting observation is that this model can evolve with a lower 'speed of change'. The fact that at each time step only two agents are chosen to interact and that they may not change opinion creates the possibility that in a considerable percentage of time steps the opinions don't change,

in contrast with the other models where the opinions are constantly changing. The lower ‘speed of change’ already takes into account that only two agents interact each time step, thus it is a model property.

6.2. TRANSITION TABLE ANALYSIS

Several opinion formation models have been proposed in the literature, often based on well-studied sociological and psychological principles, such as social conformity theory [66, 14], credibility [226], biases [196, 90, 49], trust [133, 174], strong and weak ties [94, 97], moral foundations [93], expertise [222], stimulus-response theory [213], stimulus-object-response theory [35], ‘back-fire’ effect [171, 144] and ‘boomerang’ effect [116], among many others. Significant effort has been devoted to analysing opinion formation models, but their predictions are rarely compared with real data. A notable exception is the Friedkin-Johnsen model, which has been validated on numerous experiments with *small* and *medium*-size groups [81, 80, 82, 9, 183]. However, comparing the model results with *large scale data* is a challenging task for several reasons, including the difficulty in collecting large amounts of reliable opinion distributions at subsequent time instants and in gathering information about the topology of the corresponding interaction network [183], as well as the lack of systematic approaches for a high-level comparison between qualitative model outcomes and real data. Despite the difficulties, comparison with real data is crucial to assess the usefulness of a model and to identify possible directions for improving it.

The Histogram Sorting Algorithm and Transition Tables are a first step towards systematically comparing model predictions with large-scale data sets taken from real-life surveys.

Qualitative sorting of opinion distributions into five categories that account for increasing heterogeneity (*perfect consensus*, *consensus*, *polarisation*, *clustering* and *dissensus*) and its use in the computation of *Transition Tables* allows us to capture how an initial opinion distribution, belonging to a given category, can evolve over time into a final opinion distribution belonging to a possibly different category. The accuracy of an opinion formation model can be evaluated by comparing the *real transition table*, which displays the evolution between survey data taken in two subsequent occasions, with the *predicted transition table*, which displays the prediction generated by the model starting from initial survey data.

The analysis of real opinion data from the World Values Survey [99, 101, 100], shown in Figure 6.3, provides insight into the evolution of real-life opinions, and in particular reveals that:

1. In real life there are examples of all possible transitions.
2. The number of opinion distributions belonging to each type appears to remain almost constant in each wave.
3. About half of the opinion distributions remain of the same type.

Therefore, a fully realistic opinion formation model should be able to produce, with suitably chosen parameters, opinion distributions that recreate these three features. However, while the real Transition Tables are almost full matrices (Figure 6.3), the predicted Transition Tables are sparse (Figure 6.4): the models are inherently unable to yield some of the transitions observed in real data. Among the considered models, the Bounded Confidence model with intermediate confidence radius and the recently proposed Weighted-Median and Quantum Game models appear to be the most flexible, able to generate the richest variety of transitions and behaviours. However, there is still room for improvement.

The comparison between real and predicted Transition Tables highlights that improved opinion formation models should include flexible mechanisms able to both leave the opinion distribution category unchanged and produce any of the other distribution categories, under appropriate circumstances. We summarise some key observations, pointing at directions to improve existing models so as to match opinion transitions observed in real life.

1. Most models exhibit a strong agreement bias, resulting in an unrealistic tendency towards (perfect) consensus. This tendency could be mitigated by considering, e.g., the Friedkin-Johnsen model [79], which takes into account not only individual self-confidences, but also individual susceptibilities to social influence; these additional parameters are however extremely challenging to estimate, especially in large-scale interaction networks.
2. In the studied models, there is no direct mechanism to produce dissensus, clustering or polarisation starting from (perfect) consensus; however, these transitions do happen in real life. At the expense of the simplicity of the model, stochastic and random effects could be introduced through a noise component, representing the individuals' free will and the unpredictability of their decisions [183]. Other mechanism that can increase the heterogeneity of the opinion distribution are signed edges. Clustering and polarisation can also be produced by 'Kinetic Theory for Active Particles' models [24], this is thanks to the competition mechanism that characterises these models.
3. Most often, in the models the opinions change fast and significantly, typically converging towards an equilibrium state, whereas in real data there are plenty of examples of opinion distributions that remain *almost* constant and continue to change very slowly. This suggests that, as recently observed [198], most of the actual social dynamics lead to transient, non-equilibrium phenomena: *ad-hoc* models should be developed to capture this effect. In this context, understanding the timescale of phenomena influencing opinion formation is crucial to map the time of model simulation to the time of real-world opinion evolution, a still unresolved challenge [198].
4. Random initial opinion distributions are typically used when analysing, or numerically simulating, opinion formation models. However, as shown in survey results [99, 101, 100], this is not realistic: opinions tend to have characteristic qualitative distributions, which should be taken into account when evaluating opinion formation models. The critical role of initial conditions in opinion formation models has been pointed out as a long-overlooked problem [198]: different initial conditions can lead to grossly different final states and it is then fundamental to assign the initial opinion distribution appropriately, which remains an open challenge, in particular when large-scale interaction networks are considered.

6.3. AGREEMENT PLOT ANALYSIS

This section introduces the **Agreement Plot** concept. Based on this concept six different plots are defined, the Agent Parameter Time Evolution (**APTE**), Underlying Digraph Time Evolution (**UDTE**), Initial Opinion Time Evolution (**IOTE**), Agent Parameter Steady State (**APSS**), Underlying Digraph Steady State (**UDSS**), and Initial Opinion Steady State (**IOSS**) plots. Collectively, these six plots can be used to analyse models and identify behaviour patterns and intrinsic characteristics. The technique is exemplified with the Friedkin-Johnsen model.

6.3.1. BASIC IDEA

Given an opinion distribution $x = (x_i)_{i=1}^N$, the *general agreement* of x , denoted by $\pi(x)$ and defined as

$$\pi(x) = (\overline{|x|}, \bar{x}), \quad \text{where} \quad \overline{|x|} = \frac{1}{N} \sum_{i=1}^N |x_i| \quad \text{and} \quad \bar{x} = \frac{1}{N} \sum_{i=1}^N x_i, \quad (6.1)$$

represents how much the population cares about and agrees with a statement. Since $x_i \in [-1, 1]$ for all i , $\pi(x)$ is contained in the triangle with vertices in the points $(0, 0)$, $(1, -1)$, and $(1, 1)$. Therefore, the general agreement of every possible opinion distribution can be plotted in this triangle, independent on the number of agents. The plot of the general agreement of one or

more opinion distributions in the Cartesian plane creates what we call the **Agreement Plot**. Figure 6.5 shows the Agreement Plot of five different opinion distributions and their corresponding histograms.

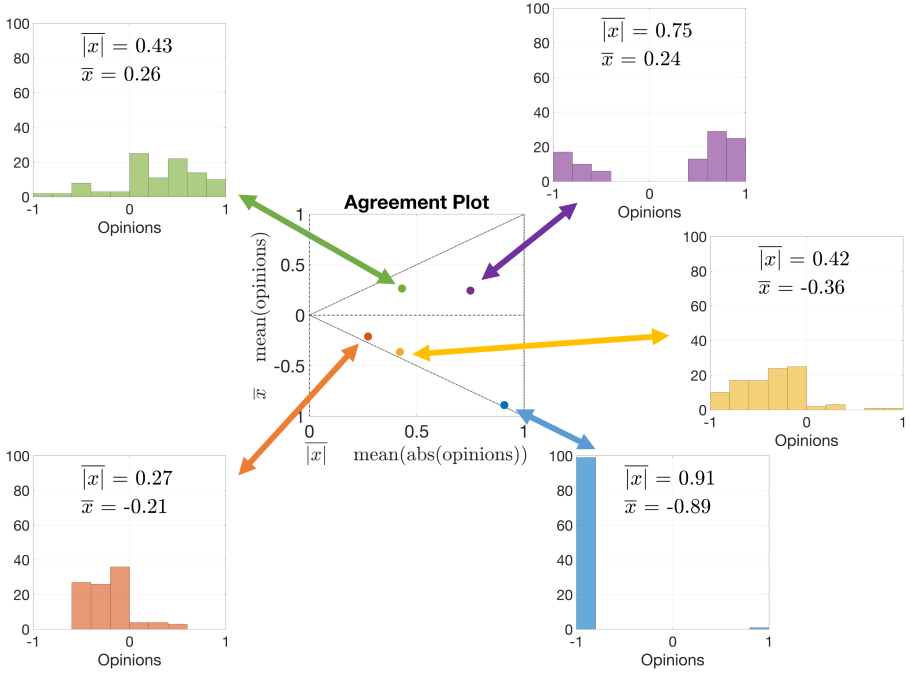


Figure 6.5: Example of five opinion distributions and their general agreement plotted in the Cartesian plane, creating what we call the *Agreement Plot*. The location of a point in the Agreement Plot provides important information on the opinion distribution, making it a visual and intuitive way to represent one or more opinion distributions.

The use of $|x|$ and $\bar{x}$ to represent complete opinion distributions consisting of possibly hundreds of agent opinions is motivated by two main reasons:

- The value of each quantity has a simple and direct interpretation, and therefore, also its representation in the Cartesian Plane. In addition to this, combined, they can provide information on the opinion distribution (as explained in the following paragraphs),
- the points $\pi(x)$ are located in a subset of the Cartesian Plane that is easy to characterise: the triangle with vertices $(0,0)$, $(1,1)$, $(1,-1)$. This is important because when plotting multiple resulting opinion distributions it is intuitively clear how much of this plane they cover.

As previously mentioned, the location of the point $\pi(x)$ provides useful information about the opinion distribution. For instance, if the point is located near one of the corners $(1,-1)$ or $(1,1)$, the population cares significantly about that statement, as $|x| \approx 1$ means most of the opinions are either -1 or 1 and therefore extremely strong opinions are common. Furthermore, near these corners $\bar{x} \approx 1$ or $\bar{x} \approx -1$, so almost every agent must have the same opinion, either ‘strongly agree’ or ‘strongly disagree’.

If the point $\pi(x)$ is in the neighbourhood of the corner $(0,0)$, then $|x| \approx 0$, so most agents are indifferent about that particular statement, and the average opinion $\bar{x} \approx 0$ is neutral. If $\pi(x) \approx (1,0)$,

the agents care about that statement but the mean is low, therefore two opposite groups must exist, with almost equal number of agents, i.e. polarisation. When $\pi(x)$ is located near the lines $y = \pm x$, almost all the population either agrees or disagrees to some degree. On the other hand, if $\pi(x)$ is located near the line $y = 0$, there is almost equal amount of agreement and disagreement.

Figure 6.6 shows the location in the Agreement Plot of different opinion distributions, and their corresponding category according to the sorting algorithm. Although it is not possible to partition the triangle with vertices $(0, 0)$, $(1, -1)$, and $(1, 1)$ into regions containing exclusively opinion distributions belonging to one category, there is a strong correlation between the location of the points in the Agreement Plot and their corresponding category, especially for perfect consensus, consensus, and polarisation. This plot further shows how arguably similar opinion distributions may correspond to different categories (two close points of different colours), which, as explained before, is a consequence of the sorting algorithm.

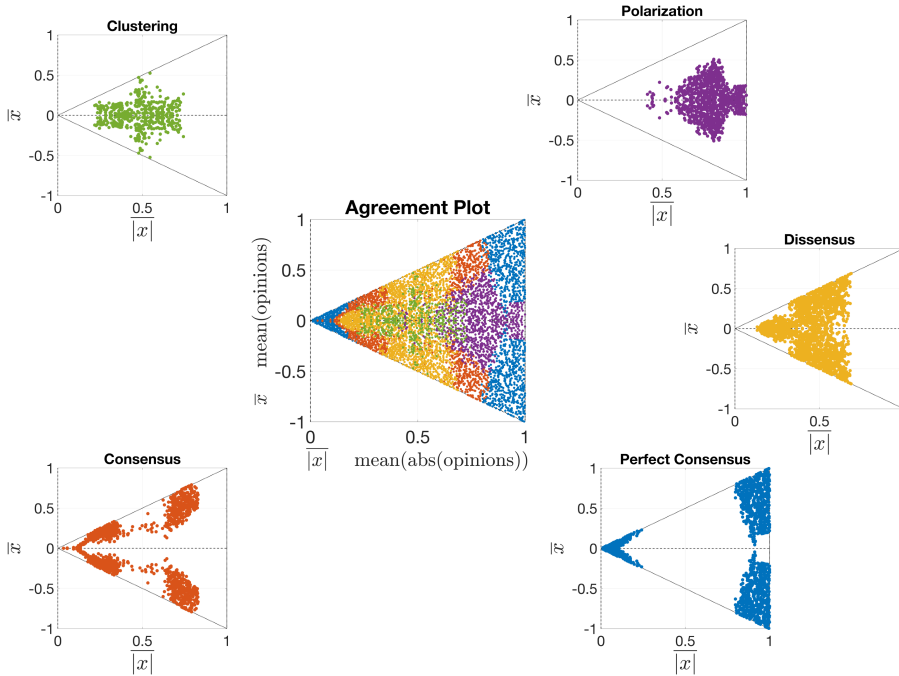


Figure 6.6: Agreement Plot of multiple opinion distributions and their corresponding category. Each point corresponds to an opinion distribution that belongs to a certain category according to the sorting algorithm. The point colour represents to which category the opinions belong: blue for perfect consensus, red for consensus, purple for polarisation, green for clustering, and yellow for dissensus. The five surrounding plots show only points belonging to one of the five categories. Although the triangle with vertices $(0, 0)$, $(1, -1)$, and $(1, 1)$ cannot be partitioned in regions containing only one single category, there is a strong correlation between the location of the opinion distribution in the Agreement Plot and its category.

The possibility to represent any opinion distribution by a single point with simple and intuitive interpretation allows for new types of model analysis. Specifically, the Agreement Plot can be used in two distinct ways: *time evolution* (see Section 6.3.2), and *steady state* (see Section 6.3.3) analysis.

Colour coding: In order to maximise the information obtained by the different plots, the plotted lines and points will be colour coded. Unless otherwise specified, the colour coding will rep-

represent the *mean agent parameters* of a population. For the Friedkin-Johnsen (FJ), Backfire Effect and Biased Assimilation (BEBA), and Bounded Confidence (BC) model each agent has a single parameter, represented by a number in a given interval ($[0, 1]$ for the FJ model, $(0, 7)$ for the BEBA model, and $[0, 2]$ for the BC model). Thus, the agent parameters of a population of N individuals are simply a set of N numbers belonging to the corresponding interval. The mean of these values is the parameter an average agent of the population would have. Each set of agent parameters has a single mean that can be colour coded in lines or points representing that population.

For different models, different colours are used based on the following convention:

- Friedkin-Johnsen model: the agent parameter is *susceptibility* and is a number in the interval $[0, 1]$. Less susceptible populations have a dark magenta colour, and more susceptible populations have a teal colour.
- Backfire Effect and Biased Assimilation model: the agent parameter is *entrenchment* and is a number in the interval $(0, 7)$. Less entrenched populations have a green colour, and more entrenched populations have an orange colour.
- Bounded Confidence model: the agent parameter is *confidence radius* and is a number in the interval $[0, 2]$. Less open populations have a lavender colour, and more open populations have a pink colour.

When the model is clear from context, the term ‘agent parameter’ is interchangeable with the agent parameter name. For instance, the statements ‘higher parameters’ mean’ is the same as ‘higher susceptibility mean’, if from the context it is clear that the statements refer to the Friedkin-Johnsen model. The colour coding extends also to histograms of the agent parameters.

6

6.3.2. AGREEMENT PLOT OF TIME EVOLUTION

The first use of the Agreement Plot is to plot one or more parametric curves (with the parameter being time) corresponding to the temporal evolution of one or more different populations.

Figure 6.7 shows an example of the complete evolution of a population's opinion represented by a single parametric curve in the Agreement Plot. On the upper left figure we can see the opinion evolution of every single agent over time. The bottom left figure shows the time-dependence of the mean of the opinion absolute values and of the opinion mean for the complete population. We can represent these two lines as a parametric curve in the Agreement Plot, as shown in the right plot. Visualising the time evolution of the individual opinions provides some information, however collective behaviours can be hidden in the complex evolution of all individual agents. The Agreement Plot allows us to better visualise collective behaviours.

If the populations differ in only one aspect, for instance, agent parameters, underlying digraph, or initial opinions, then the time evolution Agreement Plots show the overall effect these changes have on the opinion evolution. Figure 6.8 shows the three approaches applied to the Friedkin-Johnsen model.

Figure 6.8a shows 15 parametric curves. In these plots the initial opinion and underlying digraph are constant and only the agent parameters change (this is represented by the change in the line colour). It can be seen that lines with more dark magenta colour tend to be shorter than lines with more teal colour. This is to be expected, as for the first lines the agents are less susceptible and therefore the opinions change less. On the other hand, teal lines correspond to populations with high average susceptibility and therefore opinions change significantly more.

Figure 6.8b shows 9 curves corresponding to models with the same initial opinions and agent parameters evolving over different underlying digraphs. Although all the lines follow different paths, they tend to have the same overall behaviour of moving to the left, i.e. decreasing the mean of the opinion absolute values.

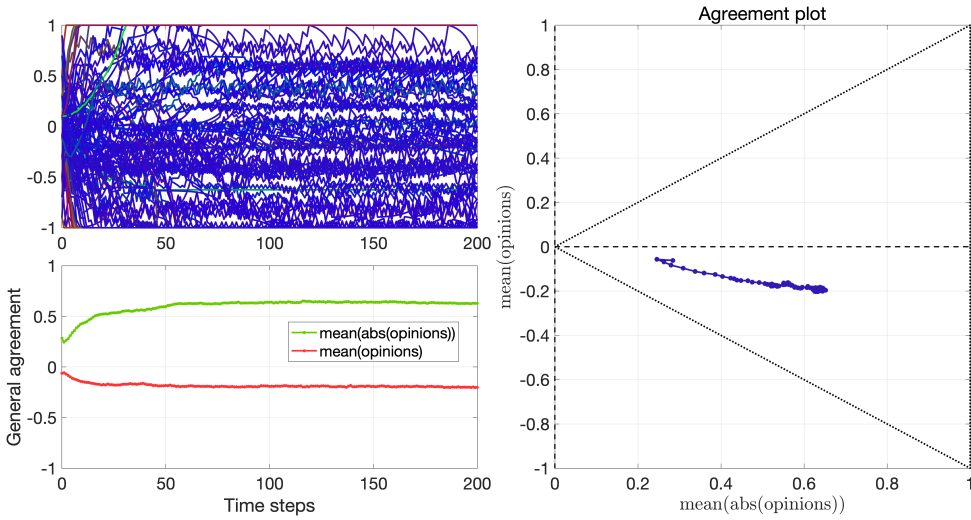


Figure 6.7: Example of how a parametric curve in the Agreement Plot can represent the opinion evolution of a complete population. The top left plot shows the time evolution of all the agents in a population (evolving with the Classification-based model, see Chapter 7). The bottom left plot shows the time-dependence of the mean of the opinion absolute values and of the opinion mean for the complete population. The right plot shows the corresponding parametric curve in the Agreement Plot.

6

Figure 6.8c shows 15 parametric curves corresponding to opinions evolving with the same agent parameters and underlying digraph, and different initial opinions (represented by the orange dots). The agent parameter number can be seen in the upper left corner. Figure 6.9 shows a histogram of the corresponding agent parameters. The digraph name can be seen in the bottom left corner, the corresponding topology is shown in Figure 5.1, and some metrics are presented in Table 6.1. Independent of the initial opinions, there is a tendency to move towards the left, and opinions close to the lines $y = \pm x$ tend to remain in the same place.

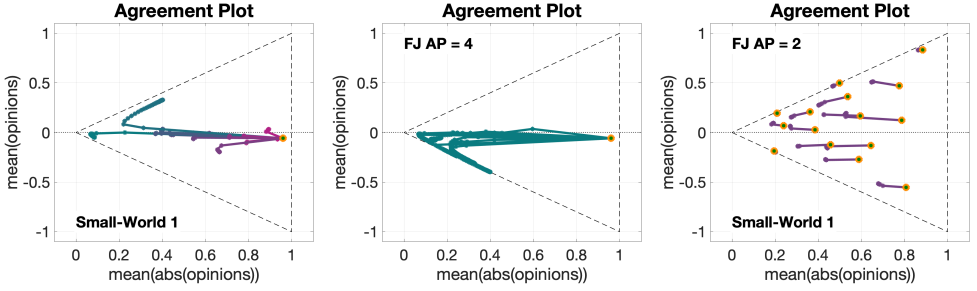
Although plots in Figure 6.8 evidence certain behaviours, a valid question to ask is if these observations depend on the chosen initial opinions, agent parameters, and underlying digraphs. For instance, how would Figure 6.8b change if the agent parameters were changed to the ones in Figure 6.8c? Or if the digraph topology changed from Small-World to Complete? These questions can be answered by repeating similar plots for a variety of different initial opinions, agent parameters, and underlying digraphs.

For simplicity, plots of parametric curves with different agent parameters and constant initial opinions and underlying digraph (like the one shown in Figure 6.8a) will be referred to as Agent Parameter Time Evolution (**APTE**) plots. Plots of parametric curves with different underlying digraph, and constant initial opinions and agent parameters (like the one shown in Figure 6.8b), will be referred to as Underlying Digraph Time Evolution (**UDTE**) plots. Plots of parametric curves with different initial opinions, and constant agent parameters and underlying digraph (like the one shown in Figure 6.8c), will be referred to as Initial Opinion Time Evolution (**IOTE**) plots.

Figure 6.9 shows the histograms corresponding to the agent parameters used for the **UDTE**, **IOTE**, **UDSS**, and **IOSS** plots for the Friedkin-Johnsen model.

Table 6.2 shows 12 different **APTE** plots for 3 choices of the initial opinions and 4 different underlying digraphs. The metrics for these four digraphs are summarised in Table 6.1.

Comparing the plots in Table 6.2, it is possible to conclude that, although the overall qualitative



(a) Parametric curves for 15 different opinion evolutions with the same initial opinions (orange dot) and underlying digraph, and different agent parameters. The digraph name can be seen in the bottom left corner, the corresponding topology is shown in Figure 5.1, and some metrics are presented in Table 6.1. This type of plot is called Agent Parameter Time Evolution (APTE).

(b) Parametric curves for 9 different opinion evolutions with the same initial opinions (orange dot) and agent parameters, and different underlying digraph. The agent parameter number can be seen in the upper left corner. Figure 6.9 shows a histogram of the corresponding agent parameters. This type of plot is called Underlying Digraph Time Evolution (UDTE).

(c) Parametric curves for 15 different opinion evolutions with the same agent parameters and underlying digraph, and different initial opinions. The agent parameter number can be seen in the upper left corner. Figure 6.9 shows a histogram of the corresponding agent parameters. The digraph name can be seen in the bottom left corner, the corresponding topology is shown in Figure 5.1, and some metrics are presented in Table 6.1. This type of plot is called Initial Opinion Time Evolution (IOTE).

Figure 6.8: Use of the Agreement Plot to identify opinion evolution behaviours produced by the Friedkin-Johnsen model. In each of the panels, multiple parametric curves (where the parameter is time) are plotted that correspond to populations that differ only in the agent parameters (Figure 6.8a), underlying digraph (Figure 6.8b), or initial opinions (Figure 6.8c). For all the plots the line colour represents the average susceptibility of the agents (magenta is less susceptible, teal is more susceptible). The initial opinions are the orange dots, and when either the agent parameters or underlying digraph are constant, their name or number is also shown in the plot. All simulations were for 50 time steps and 100 agents.

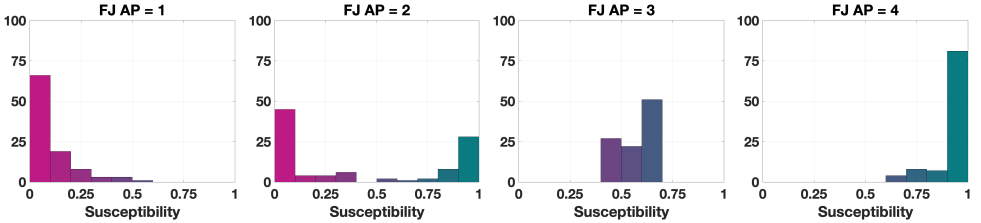


Figure 6.9: Histograms of the agent parameters used in the UDTE, IOTE, UDSS, and IOSS plots for the Friedkin-Johnsen model (Tables 6.3, 6.4, 6.6, and 6.7 respectively).

behaviour is the same as seen in Figure 6.8a, the network topology appears to have an important effect on the opinion evolution. Specifically, for all the initial opinions it can be noted that the opinions change less for the Scale-Free digraph, whereas for the other three digraphs there appears to be no significant difference (this is particularly true for row 2). A possible explanation for this phenomenon is that the high network diameter and clustering variance, combined, lead to a network that has a few central, highly influential vertices connecting otherwise disconnected sub-networks. Because of this, opinions take more time to spread, possibly explaining why opinions change less than with the other digraphs for all initial opinions.

It is also interesting to remark that the other three topologies are quite different, and still, sim-

Digraph	APL	D	$\overline{CC}$	$\sigma(CC)$	#E	$\overline{\delta^{out}}$	$\sigma(\delta^{out})$	$\overline{\delta^{in}}$	$\sigma(\delta^{in})$	BC
Directed Ring 2	5.4545	10	0.5	0	1100	11	0	11	0	0
Scale-Free	5.6893	14	0.67	0.22333	298	2.98	5.0905	2.98	5.0905	1
Small-World 1	2.0365	3	0.17538	0.0020628	1435	14.35	6.553	14.35	6.553	0.69513
Complete	1	1	1	0	10000	100	0	100	0	1

Table 6.1: Digraph information for the digraph topologies considered in the agreement plot analysis. This information can also be found in Table 5.1. The metrics are Average Path Length (APL), Diameter (D), Mean clustering ($\overline{CC}$), Clustering variance ($\sigma(CC)$), Number of edges (#E), Mean out-degree ($\overline{\delta^{out}}$), Out-degree variance ($\sigma(\delta^{out})$), Mean in-degree ($\overline{\delta^{in}}$), In-degree variance ($\sigma(\delta^{in})$), and Bidirectional Coefficient (BC).

ulation results are very similar. This is specially true for the Ring and Complete digraphs (columns 1 and 4), even though, as seen in Table 6.1, for some metrics one has the highest value, while the other has the lowest value (for instance average path length, or clustering coefficient). This indicates that, for the Friedkin-Johnsen model, the opinion evolution is independent of these network metrics.

Table 6.3 presents 12 **UDTE** plots for 4 choices of the agent parameters and 3 different initial opinions. The fact that multiple lines are close together indicates that all the evolutions have more or less the same behaviour for all different digraphs we considered (shown in Figure 5.1). This could be surprising, taking into account that Table 6.2 showed examples where the digraph has a significant effect on the opinion evolution. This apparent contradiction (of the digraphs not having a significant effect on the opinion evolution in one table, and then having a significant effect on another table) can be explained by noting that the digraph effect observed in Table 6.2 was on the speed of opinion change, but not the direction of the curves, and that the plots in Table 6.3 are showing exactly the same directions for parametric curves of different length.

Regarding the agent parameters, the behaviour observed in columns 1 and 4 is consistent with what is known about the model properties, i.e. if most agents have a low susceptibility (column 1), then opinions will change slightly, and on the contrary, if most agents have a high susceptibility (column 4), then opinions will change significantly. The most interesting plots in this regard are the ones in columns 2 and 3. They correspond to agent parameters with similar mean susceptibility value, but different variance. Column 2 describes to a society where approximately half the population is very susceptible and half is not, while column 3 represents a society where all agents have approximately the same medium susceptibility. The plots in these two columns show that then all the agents have approximately the same susceptibility, their opinions change more, indicated by the length of the curves. This behaviour is independent of the location of the initial opinion distribution in the Agreement Plot.

Table 6.4 shows 12 **IOTE** plots for 3 choices of the agent parameters and 4 different underlying digraphs. These plots provide further evidence that the observations made from Tables 6.2 and 6.3 are true for opinions starting from any initial opinion distribution. This is important to check, as it may not always be true. It is also interesting to note that the parametric curves for the plot in row 3, column 2 (Scale-Free digraph with highly susceptible agents) show a subtle peculiar behaviour in that they don't immediately move to the lines $y = \pm x$, like the other plots with these same agent parameters do. This can possibly be explained by the structure of the Scale-Free network. In contrast to the other digraphs which are made of a single 'entity', the 'different compartments' of the Scale-Free network disperse and slow down the tendency of the Friedkin-Johnsen model (with high agent susceptibility) to create perfect consensus.

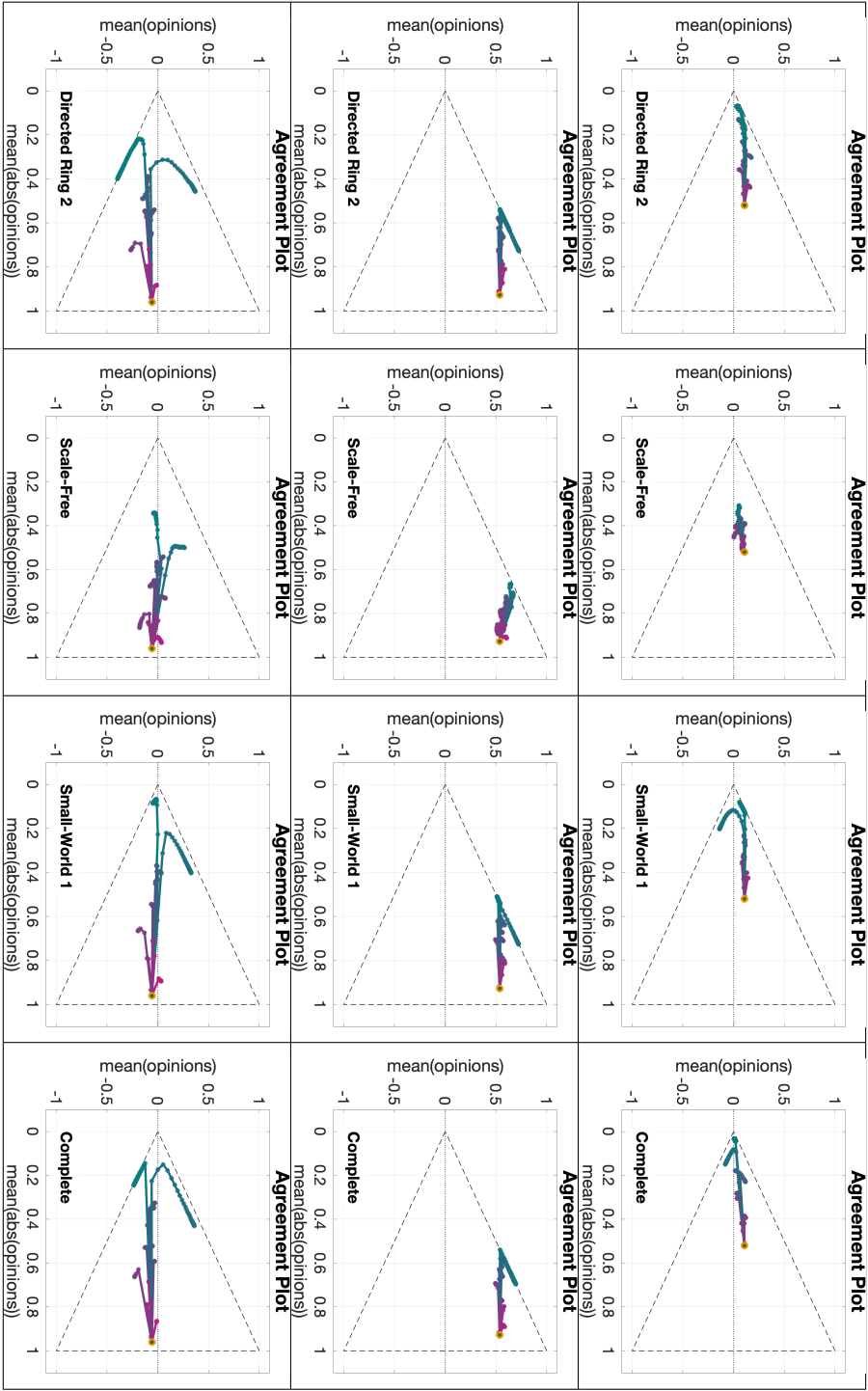


Table 6.2: **APTE** plots for the Friedkin-Johnsen model. Each of the 12 **APTE** plots includes 15 curves with different choices of agent parameters and constant initial opinion distributions and underlying digraphs. Plots in the same row start from the same initial opinion distribution (represented by the orange dot). Plots along the same column have the same underlying digraph, (from left to right, Directed Ring 2, Scale-Free, Small-World 1, and Complete, for metrics see Table 6.1). The digraph name can be seen in the bottom left corner, the corresponding topology is shown in Figure 5.1, and some metrics are presented in Table 6.1. All simulations are with 100 agents for 50 time steps.

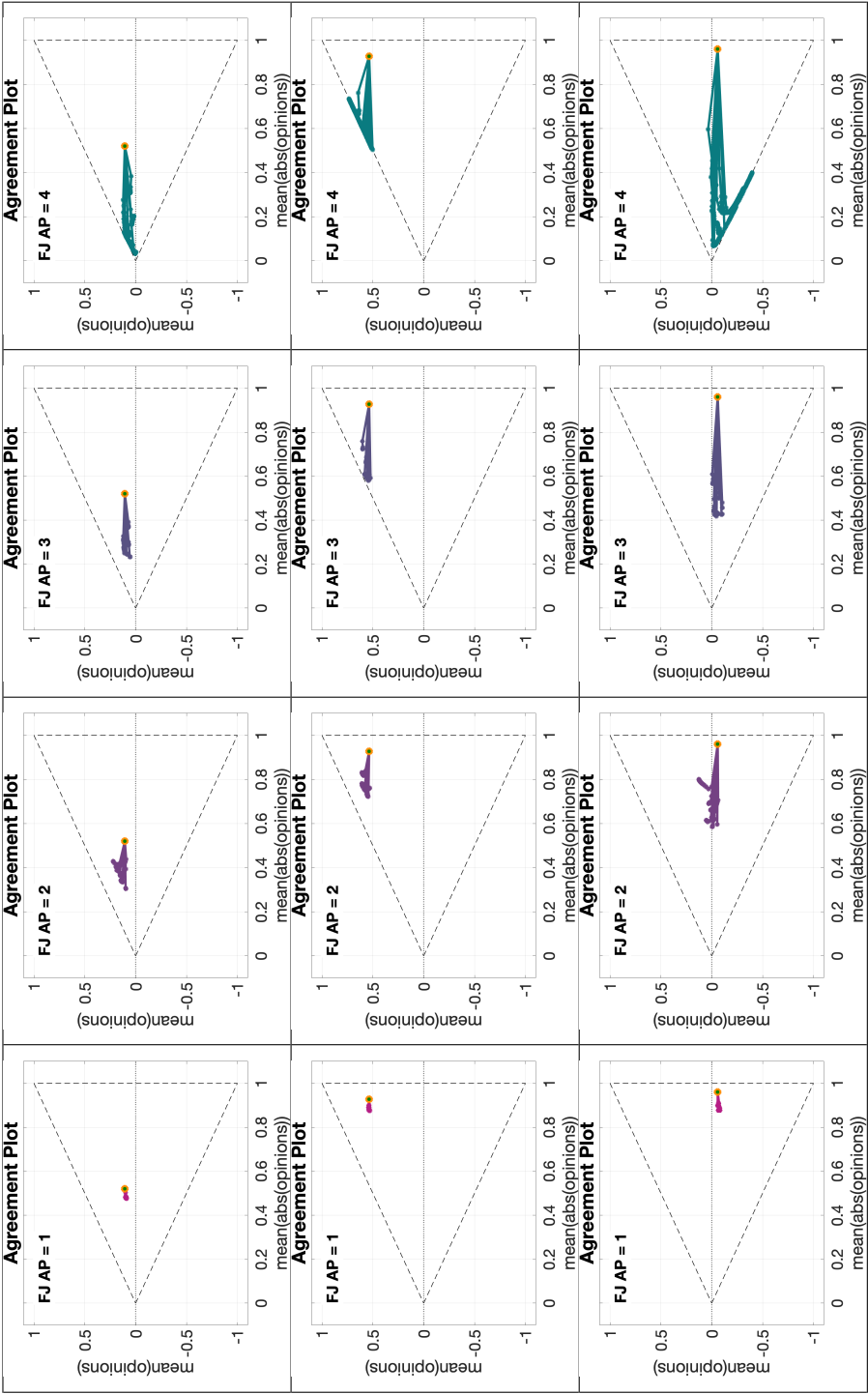


Table 6.3: **UDTE** plots for the Friedkin-Johnsen model. Each of the 12 **UDTE** plots includes 45 curves with different choices of underlying digraphs and constant initial opinion distributions and agent parameters. Plots in the same row start from the same initial opinion distribution (represented by the orange dot). Plots along the same column have the same agent parameters. The agent parameter number can be seen in the upper left corner. Figure 6.9 shows a histogram of the corresponding agent parameters. All simulations are with 100 agents for 50 time steps.

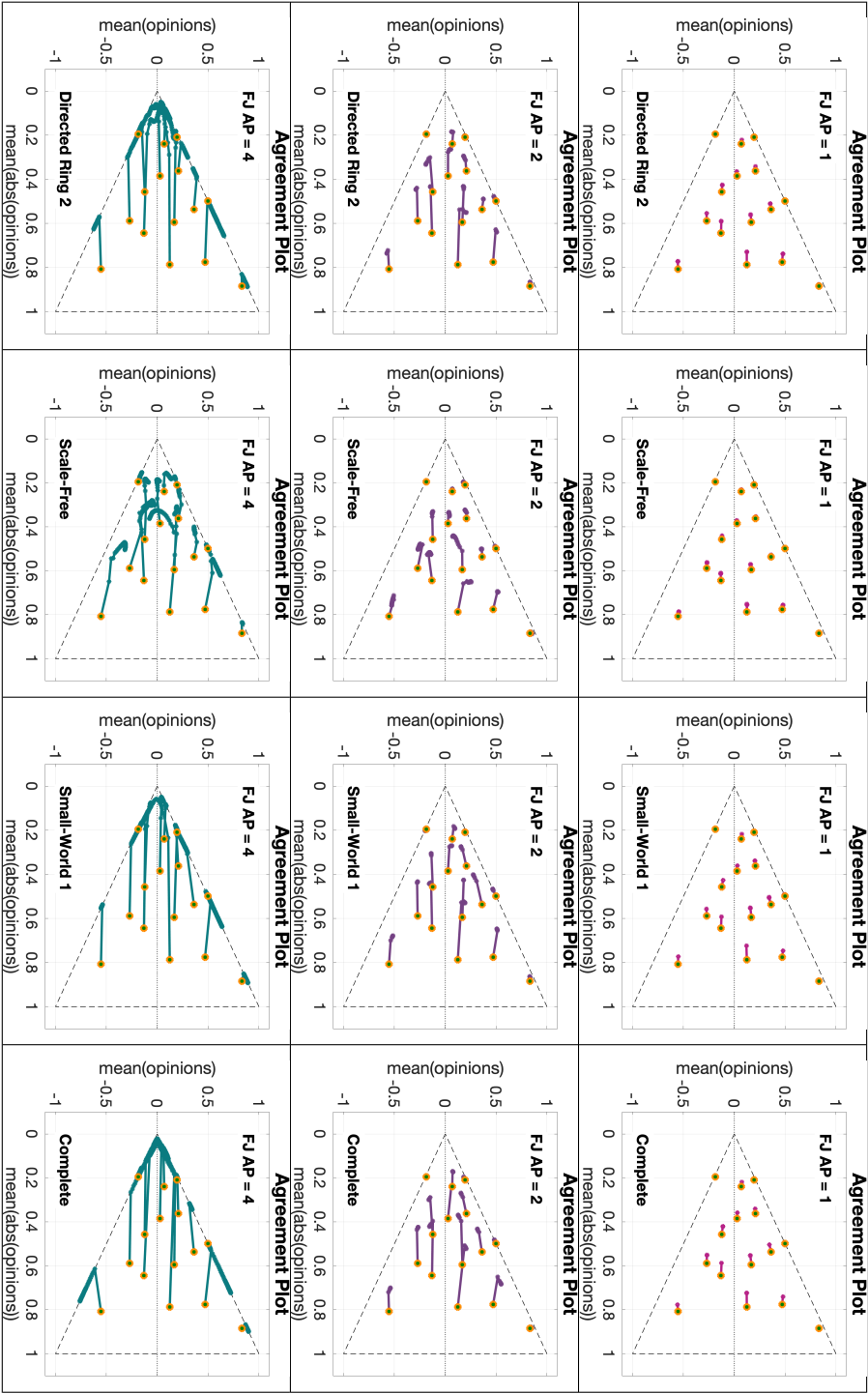
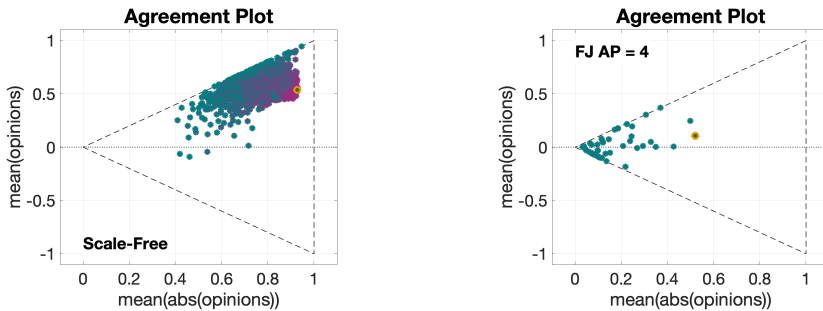


Table 6.4: *IOTE* plots for the Friedkin-Johnsen model. Each of the 12 *IOTE* plots includes 15 curves with different choices of initial opinion distributions and constant agent parameters and underlying digraphs. Plots in the same row have the same agent parameters. Plots along the same column have the same underlying digraph, (from left to right, Directed Ring 2, Scale-Free, Small-World 1, and Complete, for metrics see Table 6.1.) The agent parameter number can be seen in the upper left corner. Figure 6.9 shows a histogram of the corresponding agent parameters. The digraph name can be seen in the bottom left corner, the corresponding topology is shown in Figure 5.1, and some metrics are presented in Table 6.1. All simulations are with 100 agents for 50 time steps.

6.3.3. AGREEMENT PLOT OF STEADY STATE

If we only care about the steady state, it is possible to skip the parametric curve and simply plot the initial and final opinions. Doing this for a wide collection of agent parameters, underlying digraphs, and initial opinions (while keeping the other model parameters constant) can highlight what kind of opinion distributions the model can produce starting from a given point and also how the model transforms the opinions. For these plots, the number of time steps can also be increased, so that the models are given enough time to show the opinion changes caused by slow dynamics. For these plots, the number of time steps is increased from 50 to 1000.

We first consider the case of multiple agent parameters and underlying digraphs. These plots are quite straightforward: we simply repeat the **APTE** and **UDTE** plots but without the parametric curve, and with (possibly) a higher number of agent parameters or underlying networks. Figure 6.10 shows an example for the Friedkin-Johnsen model.



(a) Agreement Plot of the predicted opinions for 3528 different opinion evolutions with the same initial opinions (orange dot) and underlying digraph, and different agent parameters. The digraph name can be seen in the bottom left corner, the corresponding topology is shown in Figure 5.1, and some metrics are presented in Table 6.1. This type of plot is called Agent Parameter Steady State (APSS).

(b) Agreement Plot of the predicted opinions for 45 different opinion evolutions with the same initial opinions (orange dot) and agent parameters, and different underlying digraph. The agent parameter number can be seen in the upper left corner. Figure 6.9 shows a histogram of the corresponding agent parameters. This type of plot is called Underlying Digraph Steady State (UDSS).

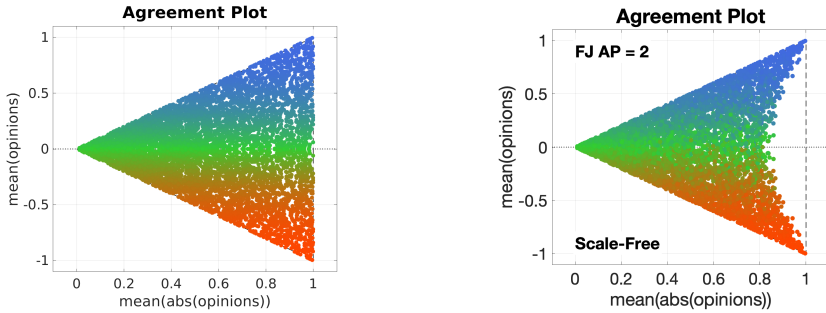
Figure 6.10: Use of the Agreement Plot to display the capacity of the Friedkin-Johnsen model to create a wide range of opinion distributions starting with the same initial opinions (orange dot), for different agent parameters (Figure 6.10a) and underlying digraphs (Figure 6.10b). For all the plots the point colour represents the average susceptibility of the agents (magenta is less susceptible, teal is more susceptible). The initial opinions are the orange dots, and when either the agent parameters or underlying digraph are constant, their name or number is also shown in the plot. All simulations were for 1000 time steps and 100 agents.

Figure 6.10a shows the Agreement Plot of 3528 predicted opinions starting with the same initial opinion distribution (orange dot), underlying digraph, and different agent parameters. The digraph name can be seen in the bottom left corner, the corresponding topology is shown in Figure 5.1, and some metrics are presented in Table 6.1. Besides observing behaviours consistent with previous simulations (such that the dark magenta points are closer to the orange dot, in contrast to the teal points), it is interesting to note that very few opinions can be found below the x-axis or to the right of the initial opinions. This seems to suggest that, independent of the agent parameters, the sign of the mean will most likely remain constant, and also that the mean of the opinion absolute values is unlikely to increase for this model.

Figure 6.10b shows the Agreement Plot of 45 predicted opinions starting with the same initial opinion distribution (orange dot), agent parameters, and different underlying digraphs. The agent parameter number can be seen in the upper left corner. Figure 6.9 shows a histogram of the corresponding agent parameters. This plot is specially interesting: given the fact that all agents have

a very high susceptibility, and that the prediction horizon is 1000 time steps, one would expect all the agents to have the same opinion, which would correspond to the prediction points being along the $y = \pm x$ lines. However that is not the case, at least with some of the points. This further signals the importance of the digraph structure and how the model can produce different and non-obvious predictions depending on the underlying digraph.

Adapting **IOTE** plots to represent the opinion evolution of a wide variety of initial opinions for a fixed agent parameter set and underlying digraph cannot be done as with **APSS** and **UDSS** plots, simply because there is not enough space. Instead what can be done is to start with a reference plot, like the one in Figure 6.11a, where each dot corresponds to a different initial opinion. Then, evolve the model, with the same agent parameters and underlying digraph, for each of the initial opinions, and present the resulting opinions in a separate plot, as shown in Figure 6.11b.



(a) Reference image for the location in the Agreement Plot of the 5314 initial opinions used in the Initial Opinion Steady State (IOSS) plot.

(b) Agreement Plot of the predicted opinions for 5314 different opinion evolutions with the same agent parameters, underlying digraph, and different initial opinions, shown in Figure 6.11a. The agent parameter number can be seen in the upper left corner. Figure 6.9 shows a histogram of the corresponding agent parameters. The digraph name can be seen in the bottom left corner, the corresponding topology is shown in Figure 5.1, and some metrics are presented in Table 6.1. This type of plot is called Initial Opinion Steady State (IOSS).

Figure 6.11: Evolution of the Friedkin-Johnsen model for different initial opinions. Dots with the same colour in Figures 6.11a and 6.11b correspond to initial and final opinion distributions for the same simulation (in these plots the colours don't have a meaning on themselves, they are only used to visually connect one dot in Figure 6.11a with a dot in Figure 6.11b). Figure 6.11a shows the initial opinion, whereas Figure 6.11b shows the predicted opinions. All simulations were for 1000 time steps and 100 agents.

Figure 6.11b shows how the 5314 initial opinions from Figure 6.11a transform according to the Friedkin-Johnsen model when the agent parameters and underlying digraph are the ones in the top and bottom left corners of Figure 6.11b, respectively. Several observations can be made about Figure 6.11b. The contraction of the initial opinions towards the y -axis is evident, and so is the fact that the contraction is more pronounced near the x -axis. This indicates that very polarised opinions are transformed into less polarised opinions, which is consistent with the averaging tendency caused by the susceptibility trait in the Friedkin-Johnsen model. The colours also allow to see that the mean opinion remains almost constant, this is hinted by the colour gradient seen in Figure 6.11a, which is still distinguishable in Figure 6.11b.

As with **APTE**, **UDTE**, and **IOTE** plots, the analysis possibilities are increased when multiple plots of the same type are arranged together, as shown in Tables 6.5, 6.6, and 6.7. This allows to determine if the observed behaviour is intrinsic for the model, or if (and how) it depends on the model parameters. It is worth clarifying that, for all **IOSS** plots, Figure 6.11a will always be the reference for the starting initial opinions.

The **APSS** 12 plots, for 3 different initial opinions, and 4 different underlying digraphs in Table 6.5 indicate that, for all initial opinions and digraphs, the qualitative effect of increasing the mean susceptibility is the same: as the mean susceptibility increases, the dark magenta points form a 'cone' to the left of the initial point. The higher the mean susceptibility, the more the predicted opinions move to the left. This trend continues until they arrive to the $y = \pm x$ lines, at which point they move along these lines, which is where most of the teal points can be found.

This also implies that the higher the mean susceptibility, the more likely the final opinions are to form perfect consensus, which is not surprising. What is surprising, is the noticeable difference between the second column and the others. As in Table 6.2 we see the Scale-Free digraph effect in that opinions don't change as much as with the other topologies. This is better evidenced in the plots of row 3.

The 12 **UDSS** plots, for 3 different initial opinions, and 4 different agent parameters in Table 6.6 show that the observation made in Table 6.3, that societies where most agents have approximately the same susceptibility change more than societies with a similar mean susceptibility but with higher susceptibility variance. This can be seen by comparing the plots in columns 2 and 3 of Table 6.6. It is also important to recall that the simulations seen in Table 6.6 are for 1000 time steps, indicating that points closer to the initial opinion not only change slower, but also change less in the asymptotic behaviour. Besides this important observation, the behaviour seen in these plots is coherent with the one presented in the previous tables.

Finally, Table 6.7 presents 12 **IOSS** plots which show how the reference initial opinions seen in Figure 6.11a evolve for 3 different choices of the agent parameters and 4 different underlying digraphs. These plots showcase familiar behaviours, while at the same time providing additional information on the model. It is not surprising that every plot contracts, as seen initially in Figure 6.11b: every plot so far shows this pattern. Nevertheless, comparing the plots evidences a hidden nonlinearity in the 'contraction factor'. Indeed, the agent parameters in row 2 are in a way the middle point between the agent parameters in row 1 and 3, however, the difference between plots in rows 1 and 2 is less significant than the difference between plots in rows 2 and 3. Regarding the underlying digraph, row 1 shows that for models with very low mean susceptibility the result is almost independent of the digraph topology. The opposite can be seen in rows 2 and 3, where the Scale-Free digraph again causes a significantly different behaviour.

After analysing Tables 6.2 to 6.7, we can draw the following conclusions regarding the opinion transformation and predictive capabilities of the Friedkin-Johnsen model:

- independent on the initial opinions, agent parameters, and underlying digraph the opinions have the tendency to move towards the lines $y = \pm x$, meaning that all the agents have (almost) the same opinion, they either agree or disagree,
- when the mean susceptibility is below about 0.5, or when the underlying digraph has a Scale-Free topology, the opinions change more slowly and the final asymptotic value is closer to the initial value,
- as the mean susceptibility increases, or the average path length and clustering variance decreases, the opinions change faster and their final asymptotic value is closer to the line $y = \pm x$,
- when the mean susceptibility is high, and the average path length and clustering variance are low, the opinions reach the $y = \pm x$ line and move parallel to it.
- the final predicted opinions are very unlikely to have a higher mean of the opinion absolute values than the initial opinions, and, when that happens, the opinions are along the $y = \pm x$ line,
- in a society with low susceptibility variance opinions will change more than another society with similar susceptibility mean and higher susceptibility variance.

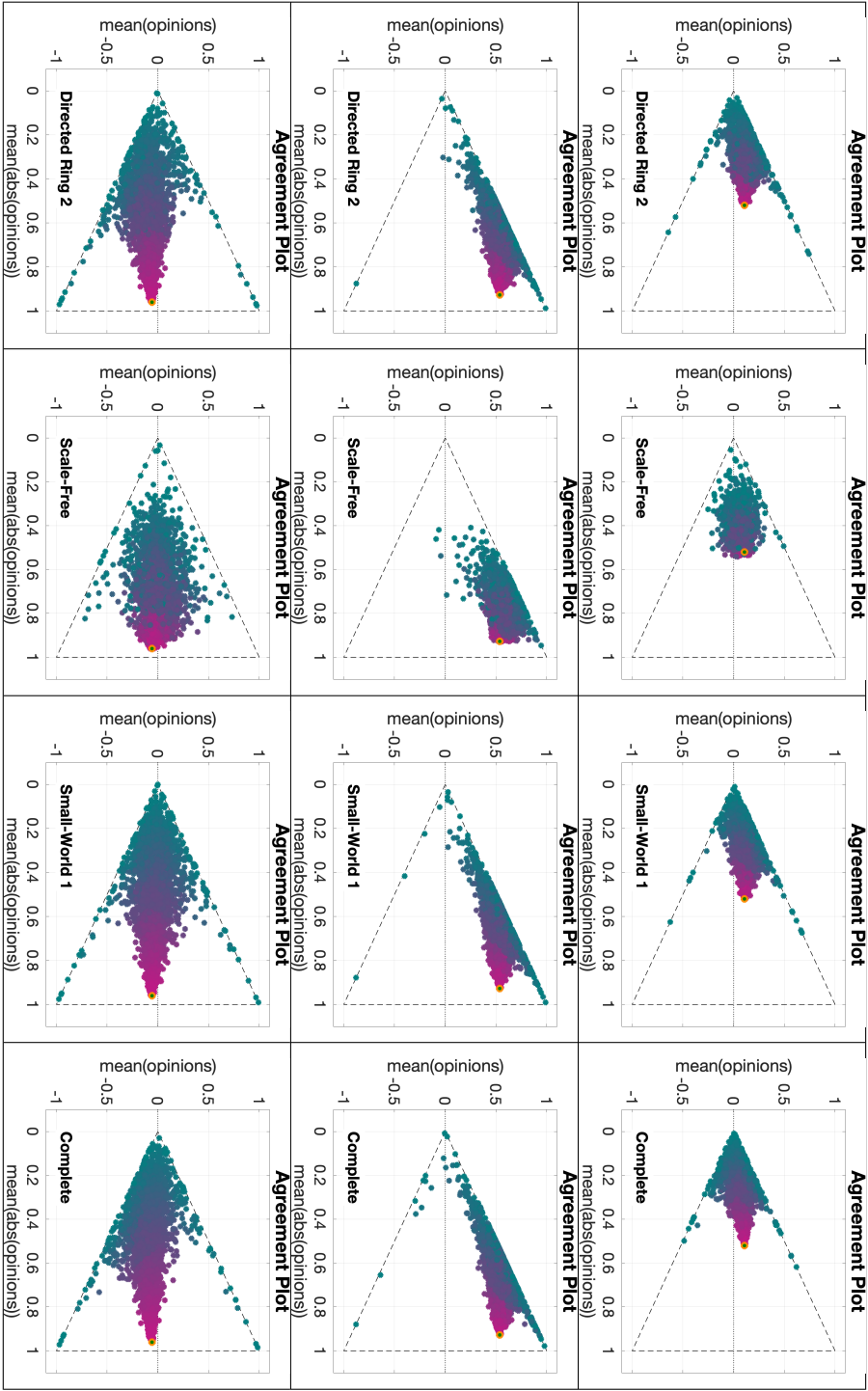


Table 6.5: **APSS** plots for the Friedkin-Johnsen model. Each of the 12 **APSS** plots includes 3528 points with different choices of agent parameters and constant initial opinion distributions and underlying digraphs. Plots in the same row start from the same initial opinion distribution (represented by the orange dot). Plots along the same column have the same underlying digraph, (from left to right, Directed Ring 2, Scale-Free, Small-World 1, and Complete, for metrics see Table 6.1). The digraph name can be seen in the bottom left corner, the corresponding topology is shown in Figure 5.1, and some metrics are presented in Table 6.1. All simulations are with 100 agents for 1000 time steps.

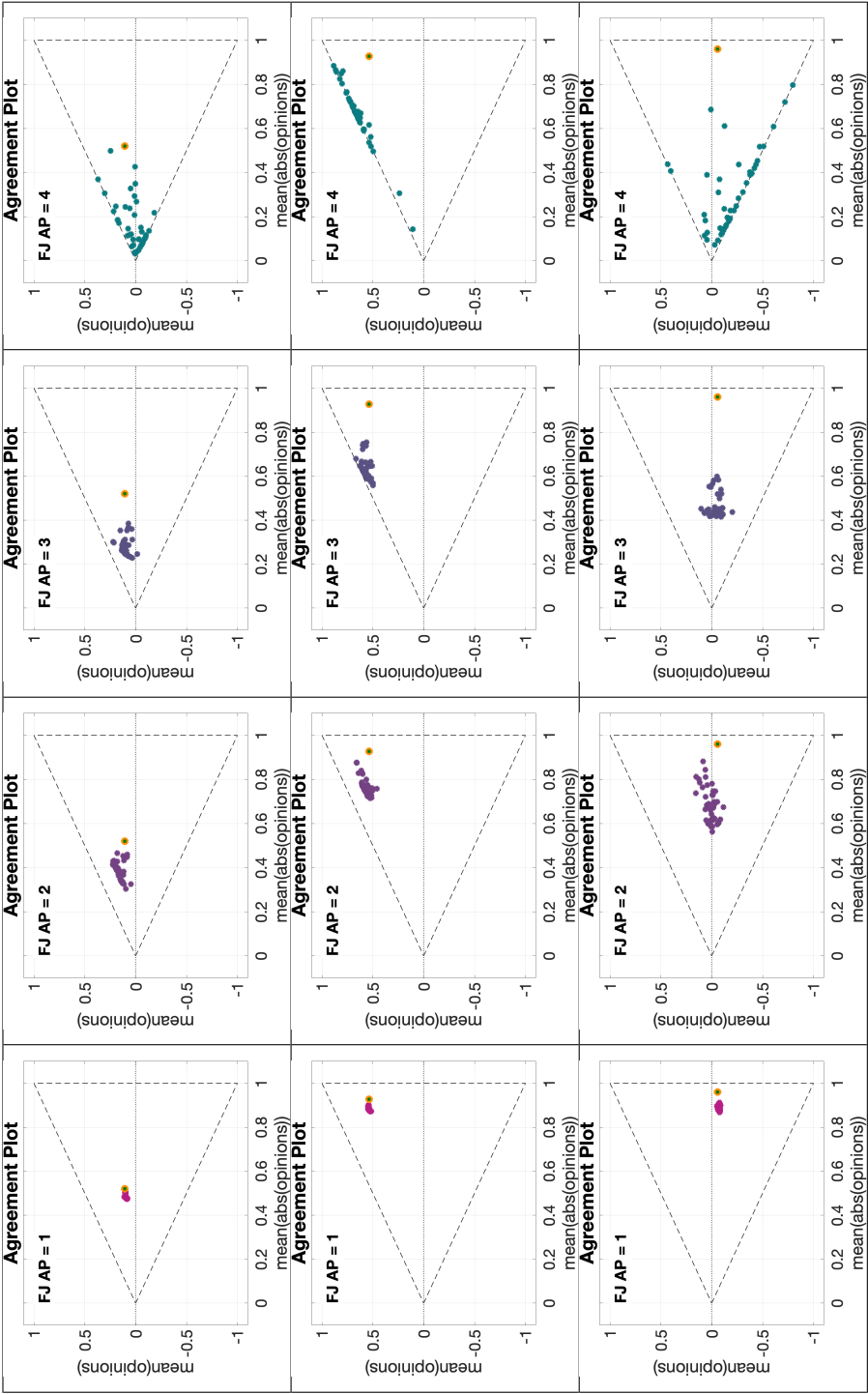


Table 6.6: **UDSS** plots for the Friedkin-Johnsen model. Each of the 12 **UDSS** plots includes 45 points with different choices of underlying digraphs and constant initial opinion distributions and agent parameters. Plots in the same row start from the same initial opinion distribution (represented by the orange dot). Plots along the same column have the same agent parameters. The agent parameter number can be seen in the upper left corner. Figure 6.9 shows a histogram of the corresponding agent parameters. All simulations are with 100 agents for 1000 time steps.

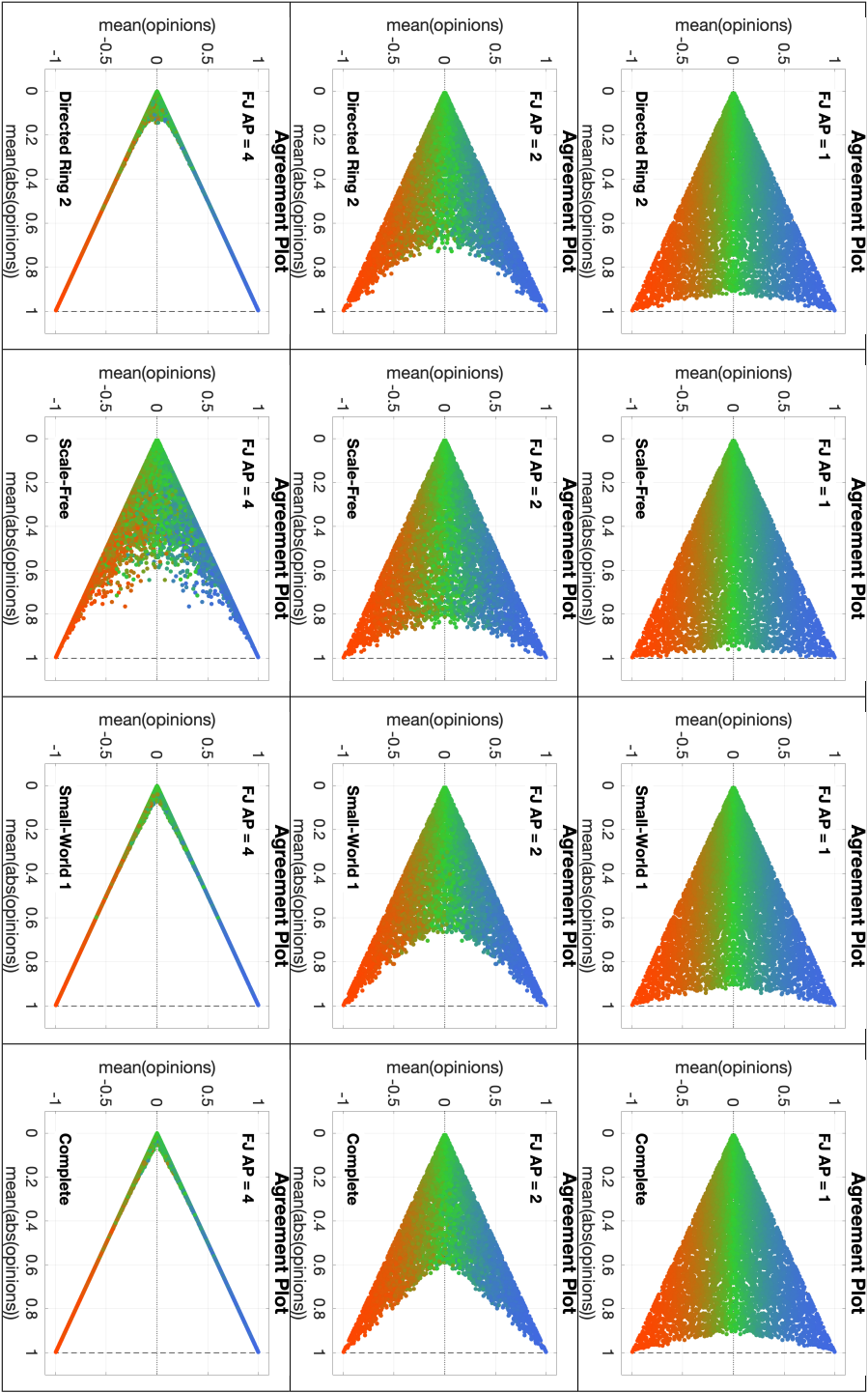


Table 6.7: IOSS plots for the Friedkin-Johnsen model. Each of the 12 IOSS plots includes 5314 points with different choices of initial opinion distributions and constant agent parameters and underlying digraphs. Plots in the same row have the same agent parameters. Plots along the same column have the same underlying digraph, (from left to right, Directed Ring 2, Scale-Free, Small-World 1, and Complete, for metrics see Table 6.1.) The agent parameter number can be seen in the upper left corner. Figure 6.9 shows a histogram of the corresponding agent parameters. The digraph name can be seen in the bottom left corner, the corresponding topology is shown in Figure 5.1, and some metrics are presented in Table 6.1. All simulations are with 100 agents for 1000 time steps.

The analysis of the Friedkin-Johnsen model using the Agreement Plot by means of Tables 6.2 to 6.7 may seem unnecessary and repetitive: after all, the same patterns were seen in almost every plot. However, it is important to keep in mind that the Friedkin-Johnsen model is relatively simple and intuitive, and is also well known, and hence the expected behaviour is showcased in the Agreement Plot analysis. Still, even in this case, the Agreement Plot analysis provided insight on the intrinsic properties of the model that can be missed otherwise, for example the critical effect of the average path length and clustering variance in the model predictions, or the fact that the opinions in populations with similar susceptibility mean change more or less, depending on the susceptibility variance.

We will see that, when the Agreement Plot is used to analyse more intricate agent-based opinion formation models, with less intuitive dynamics, each of the plot types can provide unique valuable insight on the model intrinsic properties. The information provided by, and possible interpretations of, the six previously introduced plots is summarised in Table 6.8.

Plot name	What is plotted	Constant model parameters	Variable model parameters	Line or point colour meaning	Example figure
APTE	Parametric curves (the parameter is time) corresponding to the Agreement Plot of the system evolution (50 time steps).	Initial opinions and underlying digraph	Agent parameters	The average agent parameter: susceptibility for the FJ model ($\bar{\lambda} = 0$ is dark magenta, and $\bar{\lambda} = 1$ is teal); entrenchment for the BEBA model ($\bar{\beta} = 0$ is green, and $\bar{\beta} = 7$ is orange); confidence radius for the BC model ($\bar{r} = 0$ is lavender, and $\bar{r} = 2$ is pink); and trait weights for the CB model (this will be explained in Section 7.2).	Figure 6.8a
UDTE		Initial opinions and agent parameters	Underlying digraph		Figure 6.8b
IOTE		Agent parameters and underlying digraph	Initial opinions		Figure 6.8c
APSS	Agreement Plot of the steady state opinions resulting from the system evolution (1000 time steps).	Initial opinions and underlying digraph	Agent parameters	The colour itself has no meaning, but it is used to relate the initial and final points corresponding to the same system evolution. The reference image is shown in Figure 6.11a	Figure 6.10a
UDSS		Initial opinions and agent parameters	Underlying digraph		Figure 6.10b
IOSS		Agent parameters and underlying digraph	Initial opinions		Figure 6.11b

Table 6.8: Summary of the six plot types that make use of the Agreement Plot to analyse agent-based opinion formation models.

6.4. PROBABILISTIC ANALYSIS

This section introduces and explains the Probabilistic Analysis technique. This technique is based on what we call the Qualitative Outcome Likelihood Tables (**QOL Tables**), a collection of five tables containing the probabilities that a certain qualitative outcome will be obtained. The **QOL Tables** can be used to make statements about the model outcomes when limited information is available. The technique is exemplified with the Friedkin-Johnsen model.

6.4.1. BASIC IDEA

Consider the following question: **Q1** *Given a population of 100 agents where opinions evolve according to the Friedkin-Johnsen model, if the initial opinion distribution has its mean in the interval $[-0.1, 0.4]$, the mean agent susceptibility is in the interval $[0, 0.3]$, and the underlying digraph is strongly connected, what is the probability that after 50 time steps the final opinion distribution forms perfect consensus?*

This is a natural question to ask when only limited or incomplete information on the social system is available, which is the case in reality. When the information is not complete, it is not possible to precisely compute the final opinions and their allocation, so evaluating the probability

that they belong to a certain category, is the most that can be done. If, for instance, the answer to **Q1** is 99%, then knowing the exact initial opinions, and agent parameters, along with their allocation, and the exact underlying digraph is not very relevant, as the qualitative outcome of the model is almost surely perfect consensus.

In principle, providing a quantitative answer to questions like **Q1** can be done with a four step process:

1. Translate the known information about the initial opinions, agent parameters, and underlying digraphs into well defined mathematical objects. This is done by building the set of initial opinion distributions $\mathcal{O}$, the set of agent parameter sets $\mathcal{P}$, and the set of underlying digraphs $\mathcal{N}$. In the example of **Q1** these sets are:

- The set of initial opinion distributions $\mathcal{O}$ (this is the same as a collection of initial opinion sets, as by definition, an opinion distribution is a set of opinions) contains all the initial opinion sets that meet the known information:

$$\mathcal{O} = \left\{ x = \{x_i\}_{i=1}^{100} \mid x_i \in [-1, 1] \quad \forall i, \quad \text{and} \quad \bar{x} \in [-0.1, 0.4] \right\}$$

- The set of agent parameter sets $\mathcal{P}$ contains all the agent parameter sets that meet the known information:

$$\mathcal{P} = \left\{ \lambda = \{\lambda_i\}_{i=1}^{100} \mid \lambda_i \in [0, 1] \quad \forall i, \quad \text{and} \quad \bar{\lambda} \in [-0.1, 0.4] \right\}$$

- The set of underlying digraphs $\mathcal{N}$ contains all the digraphs (represented by the corresponding adjacency matrices N) that meet the known information:

$$\mathcal{N} = \left\{ N \in [0, 1]^{100 \times 100} \mid \sum_{j=1}^{100} N_{ij} = 1 \quad \forall i, \quad \text{and} \quad N \text{ is strongly connected} \right\}$$

2. Create three algorithms that randomly sample from the sets $\mathcal{O}$, $\mathcal{P}$, and $\mathcal{N}$ uniformly.
3. For a number of events N_e do the following:

- (a) Randomly select uniformly an element from $\mathcal{O}$, $\mathcal{P}$, and $\mathcal{N}$,
- (b) evolve the system with the selected initial opinion distribution (random allocation), agent parameter set (random allocation), and underlying digraph according to the opinion model for the designated time steps,
- (c) sort the final opinions, if the sorting is perfect consensus, then label this event as success.

4. The previous sequence of N_e events has the structure of a Bernoulli process, where the random variable is a success if the final opinions belong to the category on the question. Because of this, the number of events N_e and successes N_s can be used to approximate the probability that the final opinions form perfect consensus, i.e. the answer to question **Q1**.

Denote by $\mathcal{P}_B(\mathcal{O}, \mathcal{P}, \mathcal{N}, K)$ the probability that, if the initial opinion distribution belongs to the set $\mathcal{O}$, the agent parameter set belongs to the set $\mathcal{P}$, and the underlying digraph belongs to the set $\mathcal{N}$, the final predicted opinions after K time steps is sorted into the category $B \in \{P.C., Co, Po, Cl, Di\}$. This probability can be approximated by the Wilson score interval [217]

$$\mathcal{P}_B(\mathcal{O}, \mathcal{P}, \mathcal{N}, K) \in [p - \delta, p + \delta] \quad \text{with probability 0.95} \quad (6.2)$$

$$\text{with } p = \frac{N_s + \frac{1}{2}z^2}{N_e + z^2} \quad \delta = \frac{z}{N_e + z^2} \sqrt{\frac{(N_e - N_s)N_s}{N_e} + \frac{z^2}{4}}$$

where $z = 1.96$ is the z-value for 95% confidence level, and N_e and N_s are respectively the number of events and of successes of a Bernoulli process where the Bernoulli trial is the answer to the question: *Does the final opinion belong to category B?* It can be seen that the higher the number of events, the better the approximation (since, as $N_e \rightarrow \infty$, $\delta \rightarrow 0$). This procedure can be adjusted to account for different forms of available information. For instance, if some correlation between initial opinion assignation and agent parameters is known, it can be added as a constraint to the process. The only requirement is that the random variables are independent and identically distributed, which means that every possible initial configuration should be equally likely to be chosen so as to be evolved and produce final opinions that are then sorted.

The meaning of the 95% confidence level is that the real probability of the Bernoulli process $\mathcal{P}_B(\mathcal{O}, \mathcal{P}, \mathcal{N}, K)$ is in the interval $[p - \delta, p + \delta]$ with a probability of 0.95 (meaning that Equation (6.2) is true with probability 95%).

6.4.2. IMPLEMENTATION

Although conceptually simple, there is one step in the previously explained process that is challenging, and that is step 2. Creating an algorithm that is guaranteed to uniformly sample elements from the sets $\mathcal{O}$, $\mathcal{P}$, and $\mathcal{N}$ is a challenge. That is why in the implementation the sets $\mathcal{O}$, $\mathcal{P}$, and $\mathcal{N}$ are slightly modified as explained in this section. For the creation of the sets $\mathcal{O}$ and $\mathcal{P}$ the scripts by David Holdaway [114] were used.

As with the rest of simulation results in Part 2 of the dissertation, all the simulations related to the Probabilistic Analysis evolve 100 agents.

SET OF INITIAL OPINION DISTRIBUTIONS

So far, every agent could have an opinion in the interval $[-1, 1]$. For the Probabilistic Analysis (Sections 6.4 and 7.2.5) this assumption will change. The set of opinions an agent can have will be $O = \{-1, -1 + \Delta, -1 + 2\Delta, \dots, 1 - 2\Delta, 1 - \Delta, 1\}$ where $\Delta = 2/14$. That is, the set O contains only 15 possible opinions an agent can have. We label these opinions as $O = \{X_k\}_{k=1}^{15}$. Next, the set of possible initial opinion distributions will be constrained to the following:

$$\mathcal{X} = \left\{ x = \{x_i\}_{i=1}^{100} \mid x_i \in O \text{ for } i = 1, 2, \dots, 100 \text{ and} \right. \quad (6.3)$$

$$\left. \left| \{ \tilde{x} \in x \mid \tilde{x} = X_k \} \right| \in \{0, 10, 20, \dots, 100\} \text{ for } k = 1, \dots, 15 \right\}$$

In other words, besides the possible opinions an agent can have being discrete (the 15 options in the set O), the number of agents in the opinion distribution x that have each of the possible opinions in O needs to be a multiple of 10. Thanks to these constraints, the set of possible opinion distributions $\mathcal{X}$ becomes much more manageable. If on one hand the original set of possible opinion distributions has infinite elements, the set $\mathcal{X}$ as defined in Equation 6.3 contains “only” 1961256 elements.

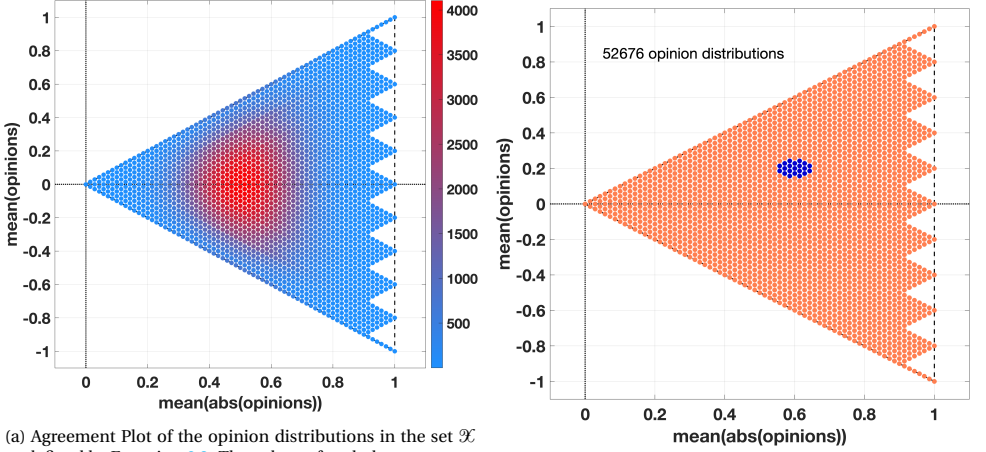
Following the intuitive interpretation of the points in the Agreement Plot, the limited information on the initial opinion distributions will be assumed to be on the opinion mean and on the

mean of the opinion absolute values. Representing every element in $\mathcal{X}$ by its mean and mean of its absolute values produces the Agreement Plot of the set $\mathcal{X}$, shown in Figure 6.12a.

If the mean $\bar{x}$ and the mean of the absolute values $\overline{|x|}$ of the initial opinion distribution x are only approximately known with uncertainty radius ϵ_o , then the collection of initial opinion distributions that takes into account this uncertainty can be defined as

$$\mathcal{O}(\bar{x}, \overline{|x|}, \epsilon_o) = \left\{ z \in \mathcal{X} \mid \sqrt{(\bar{x} - \bar{z})^2 + (\overline{|x|} - \overline{|z|})^2} \leq \epsilon_o \right\} \quad (6.4)$$

The collection defined in Equation (6.4) will be the one used for the Probabilistic Analysis. Figure 6.12b shows an example of how a set of this looks like in the Agreement Plot.



(a) Agreement Plot of the opinion distributions in the set $\mathcal{X}$ as defined by Equation 6.3. The colour of each dot represents the number of elements in $\mathcal{X}$ that have the same mean and absolute value mean.

(b) The highlighted points correspond to the set $\mathcal{O}(0.2, 0.6, 0.05)$ according to Equation (6.4).

Figure 6.12: Agreement Plot of the opinion distributions in the collection $\mathcal{X}$ according to Equation (6.3). The vertical saw-tooth pattern is due to the way the sets in $\mathcal{X}$ are constructed, in particular because of the restriction in the values the opinions can take, and also the requirement that the number of agents in the opinion distributions that have the same opinion must be a multiple of 10.

SET OF AGENT PARAMETERS SETS

The definition of this set depends on the model and the type of agent parameters it uses. For now we consider the Friedkin-Johnsen model. In this case, since the agent parameter is a single number (susceptibility, λ_i) in the interval $[0, 1]$, then the approach is basically the same as before. Instead of considering all the possible sets of agent parameters, we consider the constrained collection

$$\mathcal{Y} = \left\{ \lambda = \{\lambda_i\}_{i=1}^{100} \mid \lambda_i \in P \text{ for } i = 1, 2, \dots, 100 \text{ and} \right. \quad (6.5)$$

$$\left. \left\{ \tilde{\lambda} \in \lambda \mid \tilde{\lambda} = P_k \right\} \in \{0, 10, 20, \dots, 100\} \text{ for } k = 1, \dots, 15 \right\}$$

where $P = \{P_k\}_{k=1}^{15}$ and $P_k = (k-1)/14$. As before, the set $\mathcal{Y}$ contains only 1961256 elements. This time, the relevant properties of the agent parameter set will be the mean and variance (as

the mean of the absolute value is the same as the mean since the susceptibility is always non-negative). Representing every element in $\mathcal{Y}$ by its mean and variance produces the plot shown in Figure 6.13a.

If the mean $\bar{\lambda}$ and variance $\sigma(\lambda)$ of the agent parameters (susceptibilities, λ) are only approximately known, with uncertainty radius ϵ_p , then the set of agent parameter sets that represents this uncertainty can be defined as

$$\mathcal{P}(\bar{\lambda}, \sigma(\lambda), \epsilon_p) = \left\{ q \in \mathcal{Y} \mid \sqrt{(\bar{\lambda} - \bar{q})^2 + (\sigma(\lambda) - \sigma(q))^2} \leq \epsilon_p \right\} \quad (6.6)$$

The collection defined in Equation (6.6) will be the one used for the Probabilistic Analysis. Figure 6.13b shows an example of how this collection looks like.

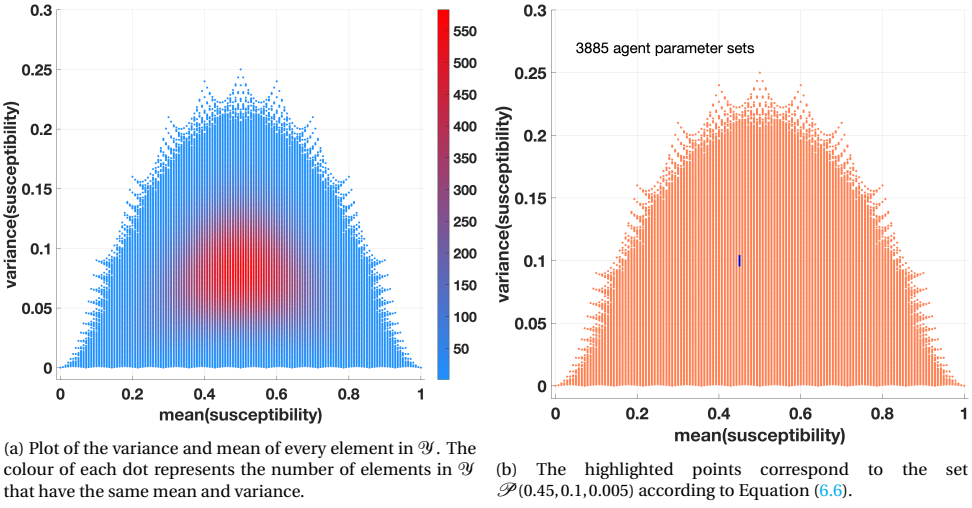


Figure 6.13: Plot of mean and variance of the susceptibility (agent parameter) sets in the collection $\mathcal{Y}$ according to Equation (6.5).

SET OF UNDERLYING DIGRAPHS

The set of all possible underlying digraphs also depends on the model, for instance the Friedkin-Johnsen model is associated with the set of all $N \times N$ row-stochastic weighted adjacency matrices, whereas the digraphs in the Altafini model belong to the set $\{-1, 0, 1\}^{N \times N}$ since in that model the digraphs are signed but unweighted. Call $\mathcal{L}$ the set of adjacency matrices for all possible underlying digraphs admitted by the model, with no other constraints. So, for the Friedkin-Johnsen model $\mathcal{L}$ is the set of all $N \times N$ row-stochastic matrices:

$$\mathcal{L} = \left\{ N \in [0, 1]^{100 \times 100} \mid \sum_{j=1}^{100} N_{ij} = 1 \text{ for } i = 1, 2, \dots, 100 \right\}. \quad (6.7)$$

It is possible to add additional constraints to the set $\mathcal{L}$, for instance topological constraints, by requiring all nodes to have a self-loop, the digraph to be strongly connected, or having certain average path length, clustering coefficient, or diameter. For the Probabilistic Analysis simulations, two constraints were taken into account: the digraphs need to be strongly connected and have a Small-World network topology. Hence, the set of considered underlying digraphs (represented by

adjacency matrices) $\mathcal{N} \subset \mathcal{L}$ is composed of 1000 different strongly connected Small-World network digraphs. Plots showing metrics for these digraphs are presented in Figure 6.14. The different digraphs were made by varying the parameters used in the creation of directed Small-World networks, such as the number of edges from the starting Ring digraph, the rewiring probability, and the probability of having bidirectional edges. From this variation, we can obtain digraphs with the same type of topology but significantly different metrics, shown in Figure 6.14.

For all the Probabilistic Analysis simulations presented in this dissertation, the set $\mathcal{N}$ will contain 1000 digraphs. For each model the digraphs will be slightly different, as some models use weighted, or signed digraphs. However, for all models the topology of the digraphs will remain the same. That is, if $\mathcal{N}_{BC}$, $\mathcal{N}_{BEB}$, $\mathcal{N}_{FJ}$, and $\mathcal{N}_{CB}$ are the sets of digraphs used in the Probabilistic Analysis for the Bounded Confidence, Backfire Effect and Biased Assimilation, Friedkin-Johnsen, and Classification-based (explained in Chapter 7) models, then $|\mathcal{N}_{BC}| = |\mathcal{N}_{BEB}| = |\mathcal{N}_{FJ}| = |\mathcal{N}_{CB}| = 1000$ and for each digraph in any of these sets, there is one digraph in each of the other sets with the exact same *topology*.

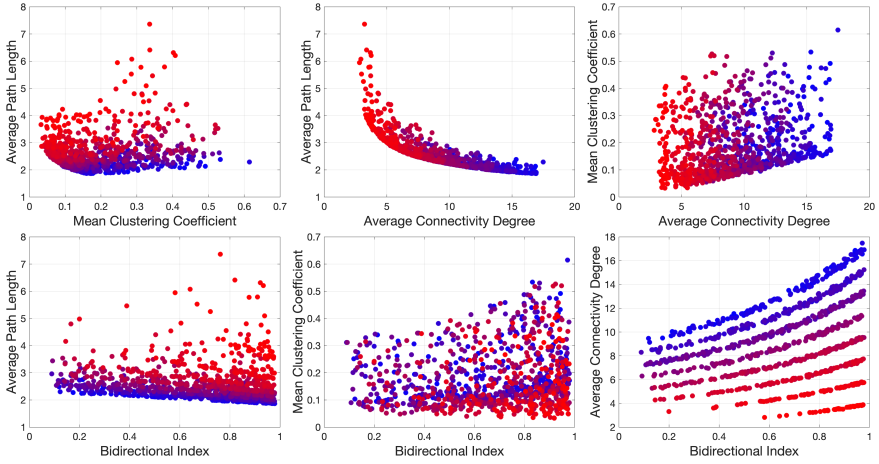


Figure 6.14: Plots showing the relations between four metrics computed for all the digraphs in the set $\mathcal{N}$. The metrics are: *average path length*, *clustering coefficient*, *average connectivity degree*, and *bidirectional coefficient*. The set $\mathcal{N}$ contains 1000 strongly connected digraphs with Small-World topology, therefore each plot has 1000 points. Points with the same colour correspond to the same digraph. Details on how these topological metrics are computed can be found in Section 7.4.

The great advantage of approximating the sets $\mathcal{O}$, $\mathcal{P}$, and $\mathcal{N}$ as in Equations (6.4), (6.6), and as discussed above is that their cardinality is finite and relatively easy to handle. For instance the collection $\mathcal{O}(0.2, 0.6, 0.05)$ has 52676 elements, as shown in blue in Figure 6.12b; the collection $\mathcal{P}(0.45, 0.1, 0.005)$ has 3885 elements, as shown in blue in Figure 6.13b; and the set $\mathcal{N}$ has 1000 elements, with metrics shown in Figure 6.14. Therefore, choosing one element of these sets with uniform probability can be done simply by ordering the elements and uniformly choosing a random integer in the appropriate interval.

QUALITATIVE OUTCOME LIKELIHOOD TABLES

With this construction in place, it is straightforward to compute the interval $[p - \delta, p + \delta]$ that (with a 95% probability) contains the probability $\mathcal{P}_B = \mathcal{P}_B(\mathcal{O}, \mathcal{P}, \mathcal{N}, K)$ according to Equation (6.2). At this point, it is important to clearly explain the exact relation between p and $\mathcal{P}_B$. $\mathcal{P}_B$ is the **real** probability, what we want to know; while, p is a value we can compute that **approximates** the

real probability. For all the simulations done in the Probabilistic Analysis section, the number of events N_e is 10000. As shown in Figure 6.15, this means that, for any possible number of successes $N_s \in \{0, 1, \dots, 9999, 10000\}$, the uncertainty δ is less than 0.01, hence **for all the results in the Probabilistic Analysis** $|\mathcal{P}_B - p| < 0.01$ **with 95% probability, where p is as defined in Equation (6.2)**. For clarity and simplicity, in the remainder of the chapter we will identify p with $\mathcal{P}_B$, meaning that if we say ‘we can compute the probability $\mathcal{P}_B$ ’ the precise statement would be ‘we can compute the probability p for which $|\mathcal{P}_B - p| < 0.01$ with 95% probability’.

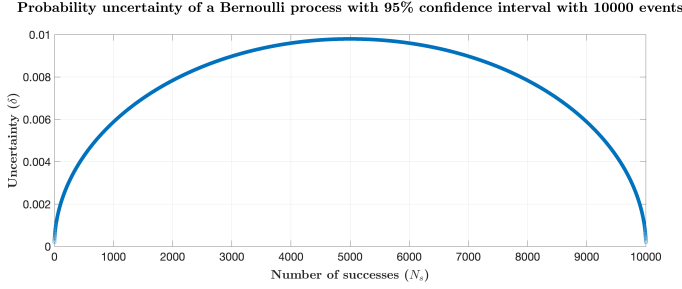


Figure 6.15: Probability uncertainty δ of the Bernoulli process when the number of events N_e is 10000 with a confidence level of 95%, given by Equation (6.2) (with $z = 1.96$). Regardless of the number of successes the uncertainty is always lower than 0.01.

We have thus established the groundwork to compute and interpret the probability $\mathcal{P}_B(\mathcal{O}, \mathcal{P}, \mathcal{N}, K)$, with $\mathcal{O} = \mathcal{O}(\bar{x}, |\bar{x}|, \epsilon_o)$ and $\mathcal{P} = \mathcal{P}(\bar{\lambda}, \sigma(\lambda), \epsilon_p)$, that approximates the probability that the predicted opinions after K time steps starting from an initial distribution with mean $\bar{x}$ and mean absolute value $|\bar{x}|$, with agent parameter (susceptibility) mean $\bar{\lambda}$ and variance $\sigma(\lambda)$, evolving over a strongly connected, Small-World network digraph in $\mathcal{N}$, belongs to the category B , with $B \in \{P.C., Co, Po, Cl, Di\}$. This probability can be computed with the single equation

$$\mathcal{P}_B(\bar{x}, |\bar{x}|, \bar{\lambda}, \sigma(\lambda), \epsilon_o, \epsilon_p, \mathcal{N}, K) = \mathcal{P}_B\left(\mathcal{O}(\bar{x}, |\bar{x}|, \epsilon_o), \mathcal{P}(\bar{\lambda}, \sigma(\lambda), \epsilon_p), \mathcal{N}, K\right). \quad (6.8)$$

The value ϵ_p and the set $\mathcal{N}$ depend on the model (for all models, all the sets $\mathcal{N}$ have the same number of digraphs with the same topology, but depending on the model the edges may be weighted or signed). Therefore, for a given model ϵ_p and $\mathcal{N}$ are fixed. For all Probabilistic Analysis results shown in the dissertation, $\epsilon_o = 0.05$ and $K = 50$. Consequently, since the only changing parameters are $\bar{x}$, $|\bar{x}|$, $\bar{\lambda}$, and $\sigma(\lambda)$, it is more practical to redefine

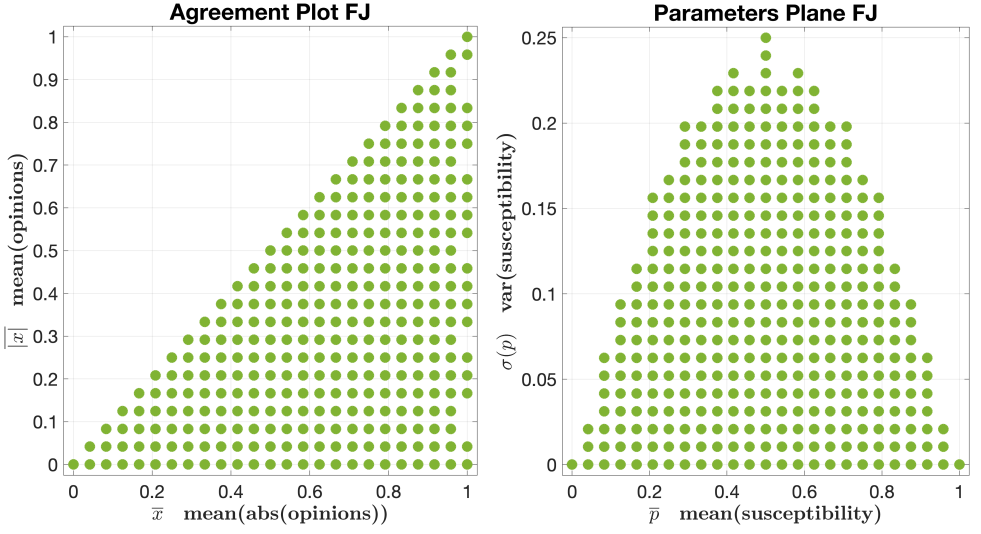
$$\mathcal{P}_B(\bar{x}, |\bar{x}|, \bar{\lambda}, \sigma(\lambda)) = \mathcal{P}_B\left(\mathcal{O}(\bar{x}, |\bar{x}|, \epsilon_o), \mathcal{P}(\bar{\lambda}, \sigma(\lambda), \epsilon_p), \mathcal{N}, K\right), \quad (6.9)$$

which is valid only when computing probabilities for the same model (since ϵ_p and $\mathcal{N}$ are omitted).

Thanks to Equation (6.9) if values for $\bar{x}$, $|\bar{x}|$, $\bar{\lambda}$, and $\sigma(\lambda)$ are given such that the sets $\mathcal{O}$, $\mathcal{P}$ are not empty, it is possible to compute the probability $\mathcal{P}_B$. Although a single application of Equation (6.9) may be useful in particular cases, in order to uncover behaviour patterns of the model it is necessary to take a broader and more systematic approach.

Figure 6.16 shows values of $\bar{x}$, $|\bar{x}|$, $\bar{\lambda}$, and $\sigma(\lambda)$ for which the sets $\mathcal{O}$ and $\mathcal{P}$ are not empty for the Friedkin-Johnsen model.

Any pair of points in Figures 6.16a and 6.16b can be associated with a number in the interval $[0, 1]$, representing the probability $\mathcal{P}_B$ (computed with Equation (6.9)) for the corresponding values of $\bar{x}$, $|\bar{x}|$, $\bar{\lambda}$, $\sigma(\lambda)$. By comparing Figure 6.16a with Figure 6.12a, it can be seen that Figure



(a) Pairs of opinion mean $\bar{x}$ and of mean of opinion absolute values $|\bar{x}|$ that produce non empty sets $\mathcal{O}$.

(b) Pairs of susceptibility mean $\bar{\lambda}$ and variance $\sigma(\lambda)$ that produce non empty sets $\mathcal{P}$ for the Friedkin-Johnsen model.

Figure 6.16: Non empty sets $\mathcal{O}$ and $\mathcal{P}$ for the Friedkin-Johnsen model. Taking any point to the left (corresponding to collection $\hat{\mathcal{O}}$) and any point to the right (corresponding to collection $\hat{\mathcal{P}}$) produces a probability $\mathcal{P}_B$ that initial opinion distributions from the collection $\hat{\mathcal{O}}$ evolving with agent parameters in $\hat{\mathcal{P}}$ produce predicted opinions sorted in category B . The ‘missing’ dots at the right of Figure 6.16a are a consequence of the saw-tooth pattern seen in Figure 6.12a. Because of this saw-tooth pattern, the pairs $(\bar{x}, |\bar{x}|)$ corresponding to the missing dots create empty $\mathcal{O}$ sets and that is why they are omitted.

6.16a only takes into account opinions with a positive mean. This is because, due to symmetry, the Probabilistic Analysis results obtained when the opinions’ mean is negative is a reflection of the results when the opinions’ mean is positive. Therefore, we only need to analyse one case.

Unfortunately, since $\mathcal{P}_B$ depends on four parameters, it is not possible to plot it directly. However, if the points in Figures 6.16a and 6.16b are given an order, it is possible to represent the values $\mathcal{P}_B$ resulting from all combinations of points in Figures 6.16a and 6.16b in a meaningful way. Figure 6.17 shows a schematic example of how this would work.

As shown in Figure 6.17, by ordering the points in Figures 6.16a and 6.16b it is possible to arrange all the probabilities $\mathcal{P}_B$ in a Table. Since $B \in \{P.C., Co, Po, Cl, Di\}$, this procedure actually creates five different tables, one per qualitative type, which are collectively called the Qualitative Outcome Likelihood Tables (**QOL Tables**). In this case, since there are 318 points in Figure 6.16a and 371 in Figure 6.16b, the **QOL Tables** have 318 rows and 371 columns. For each table, every one of the 117978 cells contains a number between 0 and 1. Therefore, instead of showing each number, it is more practical and insightful to represent visually the tables by an image. The colour of each cell depends on its value, probabilities near 0 (respectively, 0.5, and 1) have a blue (resp. green, and red) colour. This colour convention is used for all the Probabilistic Analysis and can be seen in the colourbars of the figures.

The ordering of the points in Figures 6.16a and 6.16b is directly related to the order of rows and columns in the **QOL Tables**, and as such, a suitable ordering may reveal patterns in the form of clusters or regions in the table with high or low probability.

It is also possible to plot the histogram of the cell values in the **QOL Tables**, this histogram is

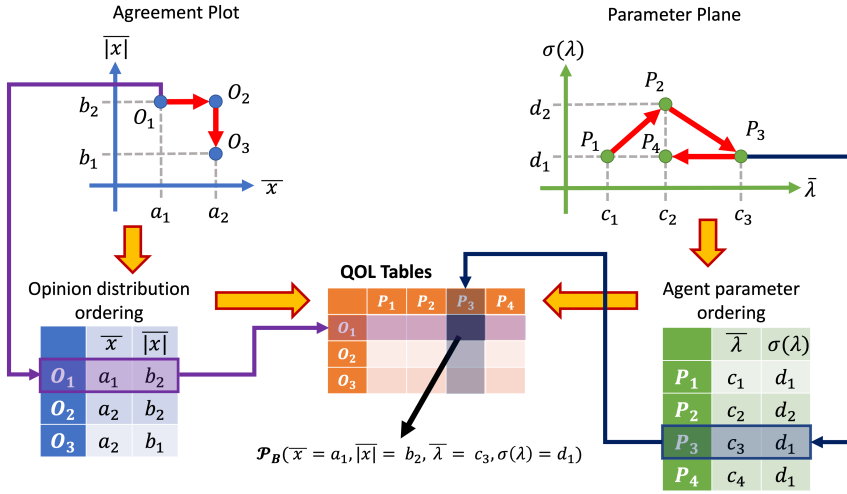


Figure 6.17: Construction of the **QOL Tables**: start with points in the Agreement Plot $\{O_1, O_2, O_3\}$, each of these points corresponds to a pair $(\bar{x}, |\bar{x}|)$ for which the set $\mathcal{O}$, given by Equation (6.4), is non-empty. These points can be ordered (see the red arrows) to produce an opinion distribution ordering. An analogous process can be done for points in the Parameter Plane $\{P_1, P_2, P_3, P_4\}$, each of these points corresponds to a pair $(\bar{\lambda}, \sigma(\lambda))$ for which the set $\mathcal{P}$, given by Equation (6.6), is non-empty. These points can also be ordered (see the red arrows) to produce an agent parameter ordering. Rows (respectively, columns) of the **QOL Tables** are associated to points in the Agreement Plot (reps. Parameter Plane) in the corresponding order. Therefore, a cell in the **QOL Tables** is related to a point in the Agreement Plot $O = (\bar{x}, |\bar{x}|)$ and a point in the Parameter Plane $P = (\bar{\lambda}, \sigma(\lambda))$, and its value between 0 and 1 is the probability $\mathcal{P}_B$ that opinions starting from the set $\mathcal{O}(\bar{x}, |\bar{x}|)$ with agent parameters in the set $\mathcal{P}(\bar{\lambda}, \sigma(\lambda))$ evolving over an underlying digraph from the set $\mathcal{N}$ produce opinions that are of qualitative type B . For analysis purposes the **QOL Tables** can be plotted. For the Friedkin-Johnsen model, the opinion distribution and agent parameter ordering are shown in Figure 6.18 (instead of arrows the order is indicated by point colour), and the **QOL Tables** plots in Figure 6.19.

invariant with respect to row and column permutations and represents the overall probability that a certain predicted category is obtained.

Figure 6.18 shows the ordering of points, and Figure 6.19 shows the corresponding images of the **QOL Tables** and their histograms, for the Friedkin-Johnsen model.

Before starting to analyse Figure 6.19 it is worth to first look at Figure 6.18. Looking at the Agreement Plot ordering (order of rows in the **QOL Tables**), it can be seen that the point order goes from the corners $(0,0)$ and $(1,1)$ to the corner $(1,0)$. Based on the interpretation of the Agreement Plot, the intuition built so far and Figure 6.6 (the figure where different opinion distributions are plotted in the Agreement Plot and colour-coded by their qualitative type) this can be interpreted as: the first rows of the **QOL Tables** correspond to opinion distributions with perfect consensus, which in the following rows becomes consensus, then clustering, and then dissensus, and eventually polarisation in the last rows.

Regarding the agent parameters, the first columns correspond to societies with very high susceptibility values that in the following columns become societies with a lower susceptibility, and the last columns correspond to societies where every agent has the same value of susceptibility which is around 0.3. Notice that the order appears to be somewhat independent of the variance, which was also observed during the Agreement Plot analysis.

This interpretation of the meaning of rows and columns can be combined to further provide meaning to different parts of the **QOL Tables**. In Figure 6.19 it can be seen for instance that perfect

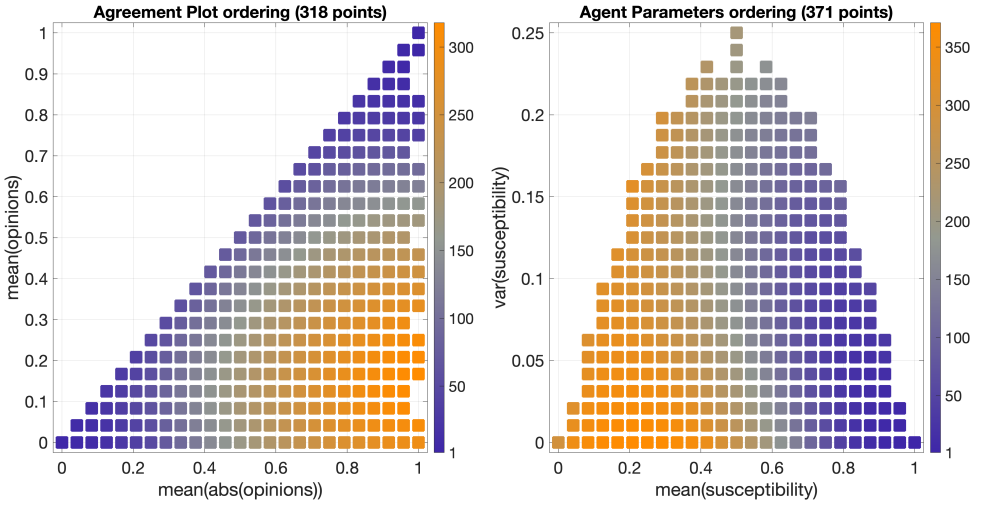


Figure 6.18: Ordering of the points in the Agreement Plot and Parameter Plane seen in Figure 6.16, for the Friedkin-Johnson model. Each point in the Agreement Plot corresponds to a pair $(\bar{x}, |\bar{x}|)$ that creates a non-empty set of opinion distributions $\mathcal{O}$. Each point in the Parameter Plane corresponds to a pair $(\bar{\lambda}, \sigma(\lambda))$ that creates a non-empty set of agent parameter sets $\mathcal{P}$. The ordering is indicated by the colorbars and will be used in the plotting of the corresponding **QOL Tables** in Figure 6.19 as explained in Figure 6.17.

6

consensus is an almost guaranteed outcome for the first rows. This makes sense: if the societies in these rows are located around the corners $(0, 0)$ and $(1, 1)$, then they already have perfect consensus, and because the Friedkin-Johnsen model has no mechanism to diversify these opinions, they remain as perfect consensus. Perfect consensus is also almost sure for cells in the left columns, corresponding to societies with very high susceptibility. This also is consistent: if all agents are highly susceptible, they will reach perfect consensus, because in these cases the model is basically the French-DeGroot model.

Figure 6.19 also shows that polarisation can be expected only when agents start from an already polarised opinion distribution and the susceptibility is very low and uniform (seen in the red bottom right corner in the third table). This basically means that the agents remain with their initial opinion, which was polarised from the start. Regarding other opinion distribution categories, they also have regions in which they are more likely. However, these regions have a less intuitive interpretation.

The histograms below the **QOL Tables** provide additional information independent on the row and column ordering. For instance, the only category which has almost guaranteed outcome is perfect consensus, as only for this table the right bin (in the neighbourhood of 1) has a significant height. It can also be seen that the least likely outcome is polarisation, as the corresponding histogram has the largest left bin (in the neighbourhood of 0). It is also interesting to note that other histograms, especially consensus, have a more uniform distribution. This means that the model is very susceptible to missing information.

Imagine now that agent parameter (respectively, initial opinion) information is not available. In this case the same analysis can be done, setting $\mathcal{O} = \mathcal{X}$ (resp. $\mathcal{P} = \mathcal{Y}$). An advantage of this approach is that in this case $\mathcal{P}_B$ depends only on $\bar{\lambda}$ and $\sigma(\lambda)$ (resp. $\bar{x}$ and $|\bar{x}|$), and therefore the resulting probabilities can be visualised in the Agreement Plot, Figure 6.16a (resp. Parameter Plane, Figure 6.16b) without having to resort to the **QOL Tables**.

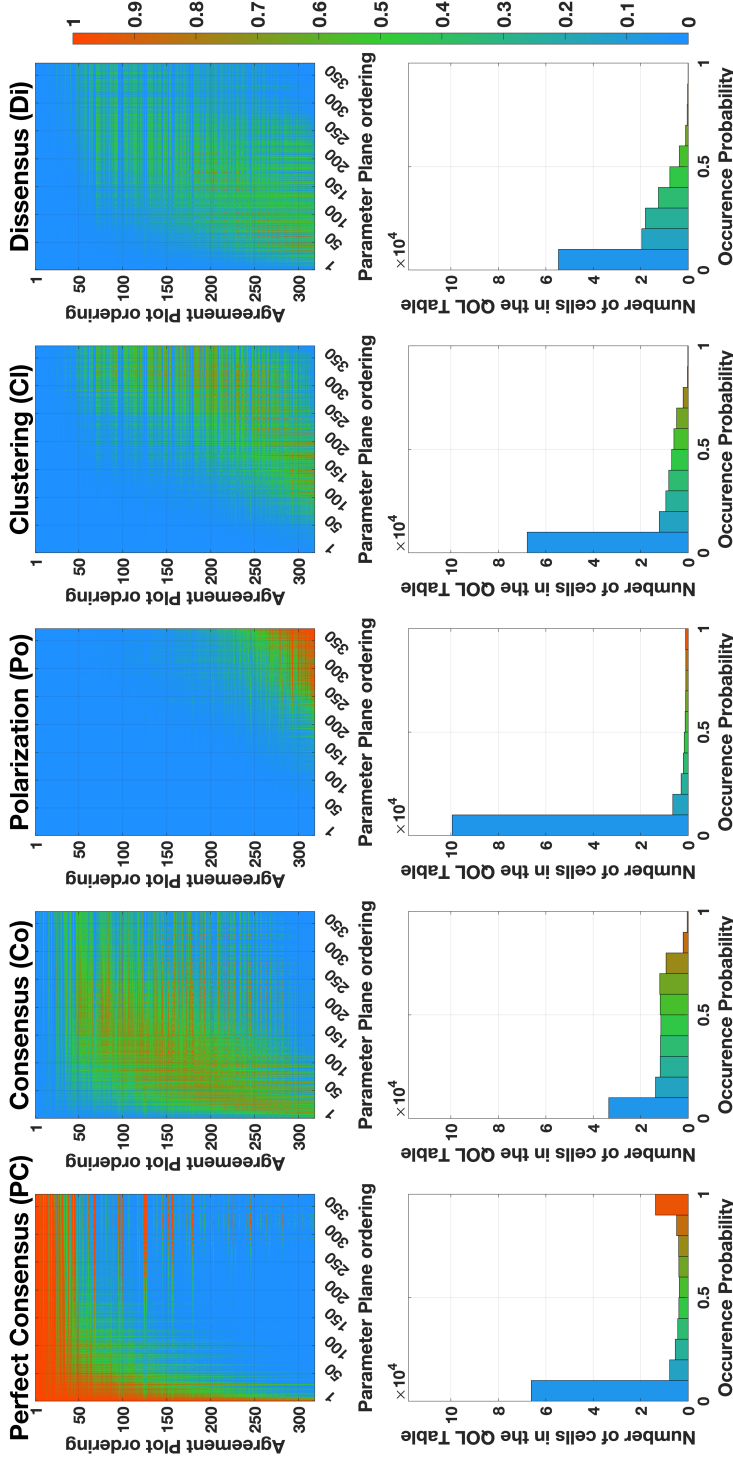


Figure 6.19: **QOL Tables** for the Friedkin-Johnson model for $B \in \{PC, Co, Po, Cl, Di\}$. The rows corresponds to points $(\bar{x}, \bar{x})$ in the Agreement Plot ordered according to Figure 6.18. The columns corresponds to points $(\bar{\lambda}, \sigma(\bar{\lambda}))$ in the Parameter Plane ordered according to Figure 6.18. Each cell is related to an initial opinion distribution and agent parameters. The cell colour represents the probability $\mathcal{P}_B$ that the resulting opinion distribution will be of qualitative type $B \in \{PC, Co, Po, Cl, Di\}$ as per Equation 6.2 (for details see Figure 6.17). Second row: the histograms corresponding to the **QOL Tables** presenting the number of cells in the table that have probability $\mathcal{P}_B$ in a given interval.

Figure 6.20 shows the results for $B \in \{P.C, Co, Po, Cl, Di\}$. This different approach shows information that supports the conclusions drawn from Figures 6.18 and 6.19

According to Figure 6.20, perfect consensus is the only outcome that is almost guaranteed (very high probability) for some regions of the Agreement Plot and the Parameter Plane. This happens when the opinions' general agreement is the neighbourhood of corners (0,0), or (1,1), which makes sense because opinions starting from perfect consensus will most likely remain in that category, as the Friedkin-Johnsen model has no mechanism to move opinions away from each other. Looking at the Parameter Plane, perfect consensus is also almost guaranteed when the mean susceptibility is very high. This also makes sense, since in these cases the model behaves almost like the French-DeGroot model, which is known to lead to perfect consensus.

In addition to these clear trends, analysing the other plots of Figure 6.20 reveals interesting observations. For instance, it is notable that for both the Agreement Plot and Parameter Plane, the probability of obtaining consensus is significant (around 50%) for a wide variety of initial opinions and agent parameters. This is evidenced by the high amount of green present in the consensus plots. The fact that most opinions form either perfect consensus or consensus means that the probabilities of obtaining other opinion categories are low. The polarisation plots are to be expected, as polarisation can only occur in the Friedkin-Johnsen model when the opinions start polarised and the susceptibility is low. Similarly, clustering can only happen when the opinions start either in clusters or dissensus and the agent parameters are allocated in an 'uniform' way, such that in the proximity of different opinions some agents have a low susceptibility (these agents will remain with that opinion) and others have high susceptibility (these agents will move towards the low susceptibility agents). This combination of circumstances is unlikely, but still noticeable in the cluster plots, especially in the green colour of agent parameters with high variance and low mean susceptibility.

Some of the observations drawn from Figures 6.18, 6.19, and 6.20 may appear repetitive, in part because of the simplicity of the Friedkin-Johnsen model. However, we will see that, for less intuitive models, our proposed techniques can provide valuable insight into the behaviour of opinion formation models that cannot be obtained via other means.

6.5. AGREEMENT PLOT AND PROBABILISTIC ANALYSIS APPLIED TO THE BEBA AND BOUNDED CONFIDENCE MODELS

The previous two sections have introduced the Agreement Plot and Probabilistic Analysis as techniques to study and discover behaviours and patterns of agent-based opinion formation models. So far, these techniques have been explained using the Friedkin-Johnsen model, resulting in a comprehensive understanding of the model behaviour and characteristics for a wide range of initial opinions, agent parameters, and underlying digraphs, also in the presence of uncertainty. In this section, the properties of two other opinion formation models, the Backfire Effect and Biased Assimilation (BEBA) model and the Bounded Confidence model, are studied using the aforementioned techniques.

6.5.1. BACKFIRE EFFECT AND BIASED ASSIMILATION MODEL

Agreement Plot analysis. Starting with the Agreement Plot technique, Tables 6.9 to 6.14 present multiple plots in a way analogous to Tables 6.2 to 6.7 in Section 6.3. For Tables 6.9, 6.11, 6.12, and 6.14, the underlying digraphs have the same topology as the digraphs used in the Agreement Plot analysis of the Friedkin-Johnsen model. Therefore, the topology metrics are shown in Table 6.1.

Figure 6.21 shows the histograms corresponding to the agent parameters used for the UDTE,

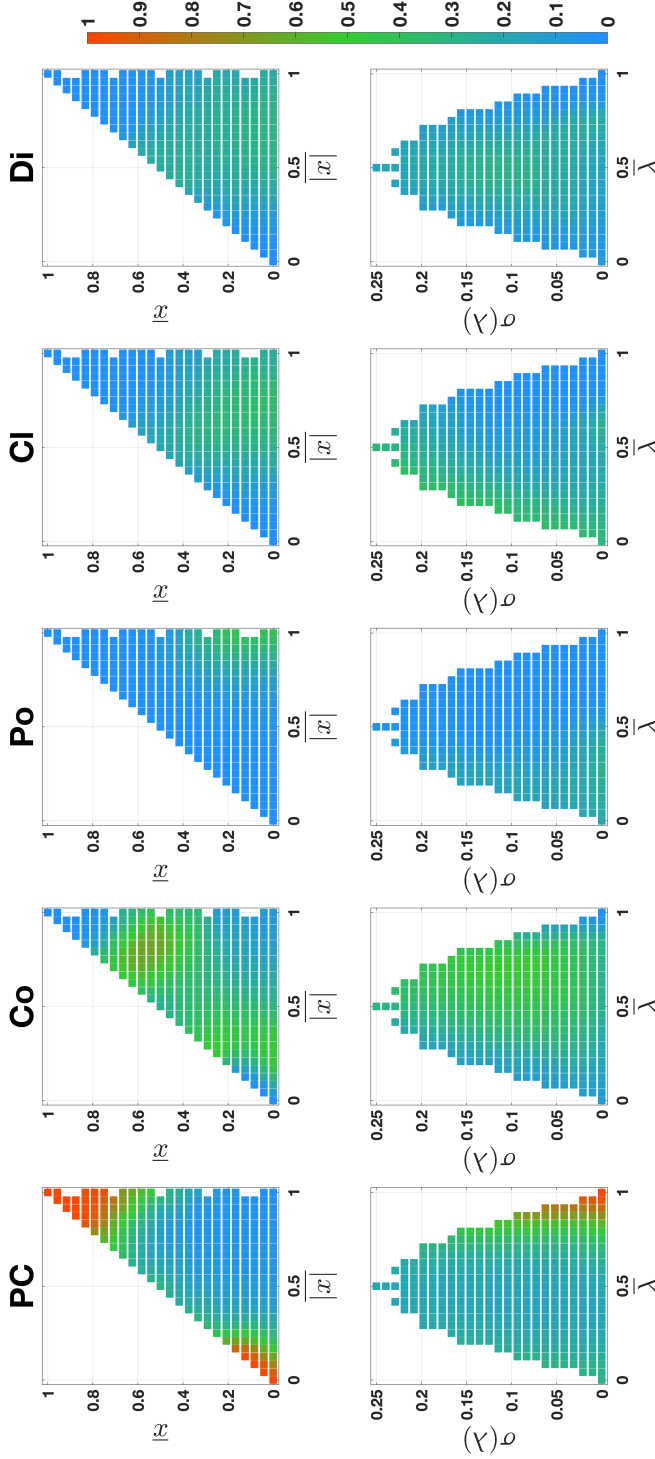


Figure 6.20: For the Friedkin-Johnson model, when agent parameter information (respectively, initial opinion) information is not available the probabilities $\mathcal{P}_B$ of the final opinions being of a qualitative type $B \in \{PC, Co, Po, Ci, Di\}$ can be plotted in the Agreement Plot (resp. Parameter Plane) shown in Figure 6.18. These probability results are shown in the first (resp. second) row.

IOTE, **UDSS**, and **IOSS** plots for the BEBA model.

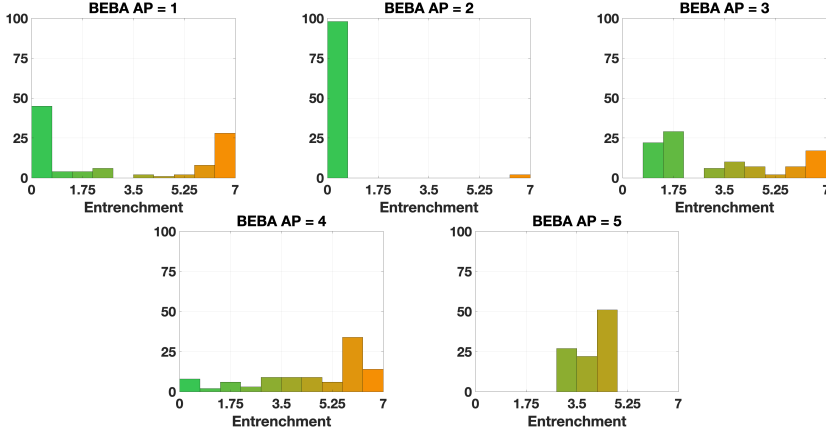


Figure 6.21: Histograms of the agent parameters used in the **UDTE**, **IOTE**, **UDSS**, and **IOSS** plots for the BEBA model (Tables 6.10, 6.11, 6.13, and 6.14 respectively).

6

As in the previous section, the line and point colours in the plots on Tables 6.9 to 6.14 indicate the mean agent parameter, which for the BEBA model is the *mean entrenchment* $\bar{\beta}$. Although originally $\beta_i > 0$, for the purpose of this analysis (and the Probabilistic Analysis) the individual entrenchment values β_i will be restricted to $\beta_i \in (0, 7)$ for all agents. Low values of the mean entrenchment are represented by green and generally translate to more open-minded agents, more willing to accept diverse opinions. Higher values of mean entrenchment are represented by orange and correspond to agents that will react negatively to other agents with a slightly different opinion, and that are positively influenced only by agents with very similar views.

Table 6.9 shows 12 **APTE** plots for 3 different initial opinions and 4 different underlying digraphs. Immediately we can notice a new behaviour, previously unseen in the Friedkin-Johnsen model. The initial opinions in rows 2 and 3 have a mean of the opinion absolute values of approximately 0.5, that is, they are more or less at the centre of the x -axis, and still several parametric curves move to the right of the axis. The explanation for this phenomenon can be found in the lines' colour. The curves that move to the right have orange shade or are orange, and thus correspond to a highly entrenched society where biased assimilation and backfire effect are predominant. It makes sense, then, that these societies will move towards regions in the Agreement Plot where agents are more interested, whether to create polarisation or perfect consensus.

On the other end of the spectrum, the green lines tend to move towards the y -axis. This behaviour is similar to the French-DeGroot model, and it also makes sense. It is known that in the limiting case where $\beta \rightarrow 0$ the BEBA model behaves like the French-DeGroot model. The digraph topology effect is similar to what was seen previously with the Friedkin-Johnsen model, i.e., opinions change more slowly in the Scale-Free digraph and faster on the other digraphs. It remains to be seen if the digraph topology has more effects on the model evolution.

Table 6.10 shows 12 **UDTE** plots for 3 different initial opinions and 4 different choices of the agent parameters. From this table it is possible to make two observations. First, besides the opinion change speed, different digraph topologies do not dramatically affect the model evolution. Some sporadic examples where not all lines follow the same direction can be seen in row 2 and 3, and column 2. Nevertheless, most curves appear to follow the same trajectory. Second, the mean entrenchment appears to be a good predictor of the system behaviour, like with the Friedkin-

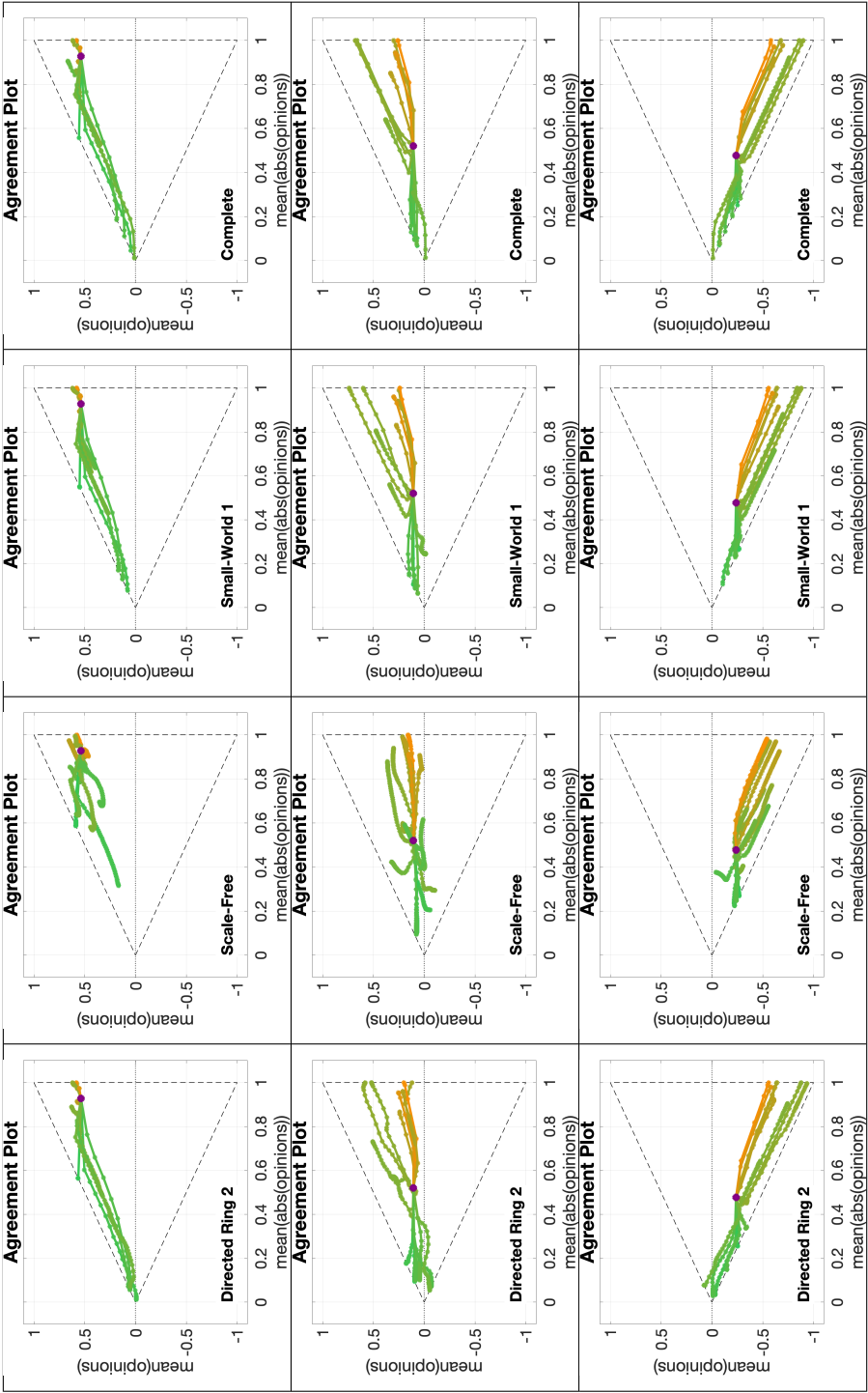


Table 6.9: APTE plots for the Backfire Effect and Biased Assimilation model. Each of the 12 APTE plots includes 15 curves with different choices of agent parameters and constant initial opinion distributions and underlying digraphs. Plots in the same row start from the same initial opinion distribution (represented by the purple dot). Plots along the same column have the same underlying digraph, (from left to right, Directed Ring 2, Scale-Free, Small-World 1, and Complete, for metrics see Table 6.1). The digraph name can be seen in the bottom left corner, the corresponding topology is shown in Figure 5.1, and some metrics are presented in Table 6.1. All simulations are with 100 agents for 50 time steps.

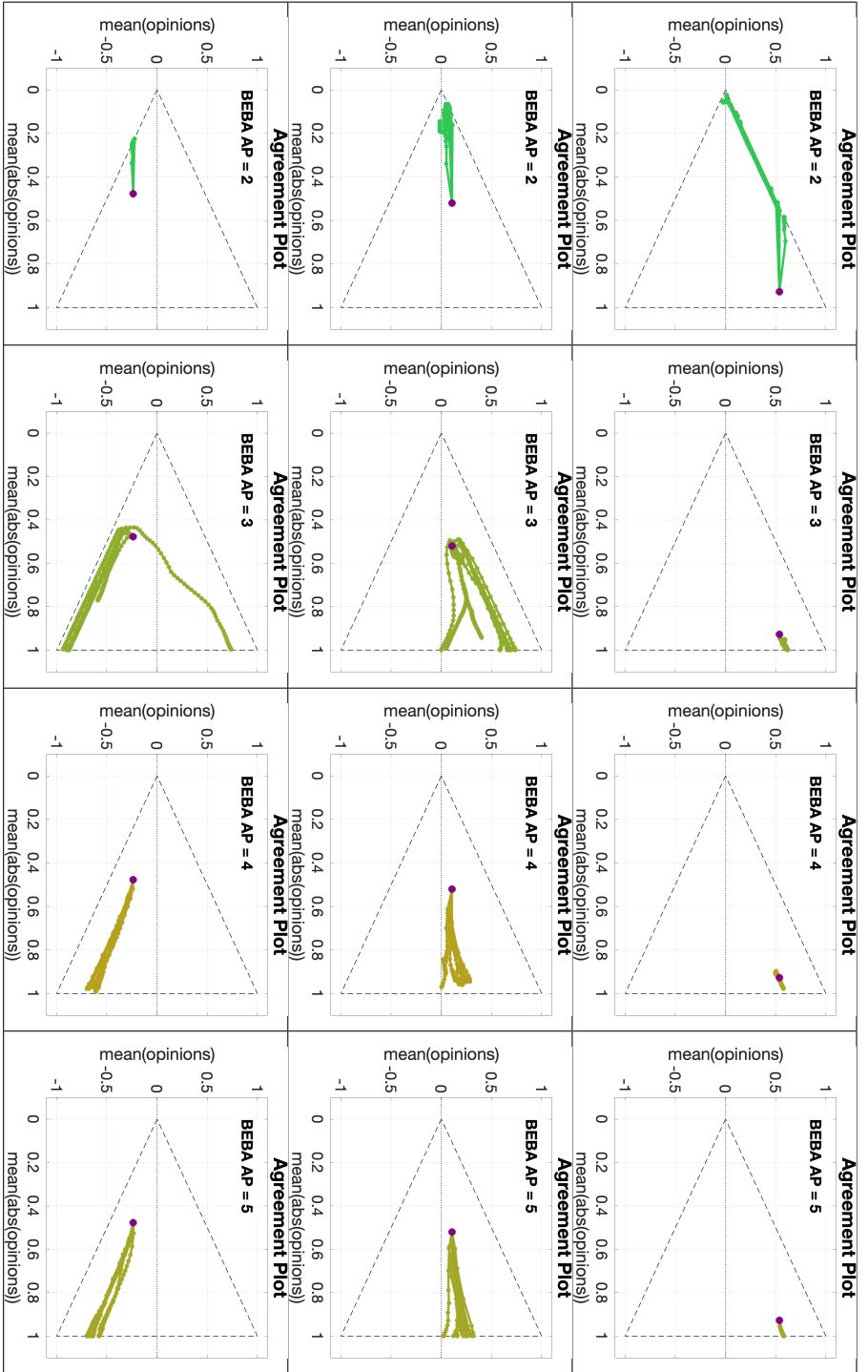


Table 6.10: UDTE plots for the Backfire Effect and Biased Assimilation model. Each of the 12 UDTE plots includes 9 curves with different choices of underlying digraphs and constant initial opinion distributions and agent parameters. Plots in the same row start from the same initial opinion distribution (represented by the purple dot). Plots along the same column have the same agent parameters. The agent parameter number can be seen in the upper left corner. Figure 6.21 shows a histogram of the corresponding agent parameters. All simulations are with 100 agents for 50 time steps.

Johnsen model. This can be seen in columns 2, 3, and 4, where the agent parameters were different but had comparable mean (seen in the similarity of the line colours) and the behaviour is approximately qualitatively comparable.

Table 6.11 shows 12 **IOTE** plots for 3 different choices of the agent parameters and 4 different underlying digraphs. We can immediately see a behaviour present in all the plots: some curves go to the left and others go to the right. Since in these plots the agent parameters and underlying network are constant, this means that the tendency of the opinions to go to extremes or to consensus depends on the location of the initial opinions in the Agreement Plot. Closer inspection reveals that some plots have a ‘threshold’ in the x -axis (the mean of the opinion absolute values) that separates the curves that move to the left or to the right. One such plot is shown in row 3, column 4. For other plots this ‘threshold’ is not so clear. See for instance the plot in row 1, column 3.

Here it is important to take into account that there are two types of curves that move to the right. One type starts moving to the right immediately, while the other moves initially to the left, and then turns to the right. One possible explanation for this is that for the first type the backfire effect is predominant and causes polarisation and seeking of extreme opinions from the start; while for the second type opinions tend to form consensus and are then reinforced, thus moving towards consensus or perfect consensus.

Table 6.12 shows 12 **APSS** plots for 3 different initial opinions and 4 different underlying digraphs. This table is very interesting, as it clearly shows the capacity of the BEBA model to produce a wide variety of predicted opinions. Several observations are relevant: (i) almost every point along the line $x = 1$ is reached, meaning that it can achieve predicted opinions where all the opinions are extreme (-1 or 1) with any proportion (the mean ranges from 1 to -1); (ii) not every orange dot is located near the point $(1, 0)$, meaning that complete entrenchment does not automatically mean polarisation; (iii) looking at the rows, it is clear that the digraph topology has an effect on the possible model outcomes. However, unlike with the Friedkin-Johnsen model, it is not only the Scale-Free digraph that behaves differently. For the BEBA model every digraph produces a significantly different plot; and (iv) although predicted opinions can be found in almost every point inside the triangle, the area near the point $(0.1, 0)$ is more empty. This may indicate that when opinions are located in that area they will not remain there, but most likely move to the left until they reach the triangle boundary.

Table 6.13 shows 12 **UDSS** plots for 3 different initial opinions and 4 different choices of the agent parameters. Plots in Table 6.13 show that, in most cases, the initial opinion and the agent parameters have a greater effect than the underlying digraph over the location of the predicted opinions: for most plots, the predicted points are closely located. Combining this observations, with the observations from Table 6.12 we can say that different underlying digraphs enable different model predictions, but that these different model predictions are ultimately achieved by the agent parameters themselves.

Table 6.14 shows 12 **IOSS** plots for 3 different choices of the agent parameters and 4 different underlying digraphs. This is one of the most interesting tables, as it highlights the rich and diverse behaviour of the BEBA model. The plot in row 1, column 1 evidences that, for fixed agent parameters and underlying digraph, the model can produce extreme opinions (recall that some opinions along the $y = \pm x$ lines belong to the perfect consensus category, while opinions located along the line $x = 1$ belong to perfect consensus or polarisation). Table 6.14 once again evidences that the Scale-Free topology leads to a peculiar model behaviour, as can be seen in every row, in particular rows 1 and 3. Row 2 is very intriguing, in particular column 4. The dots in the plots of row 2 can be separated in two groups: the first group is composed of points located along the $y = \pm x$ lines, for which either all agents agree or disagree. The second group corresponds to the ‘arc’ which is more or less defined depending on the digraph. A similar arc was seen in the Friedkin-Johnsen model

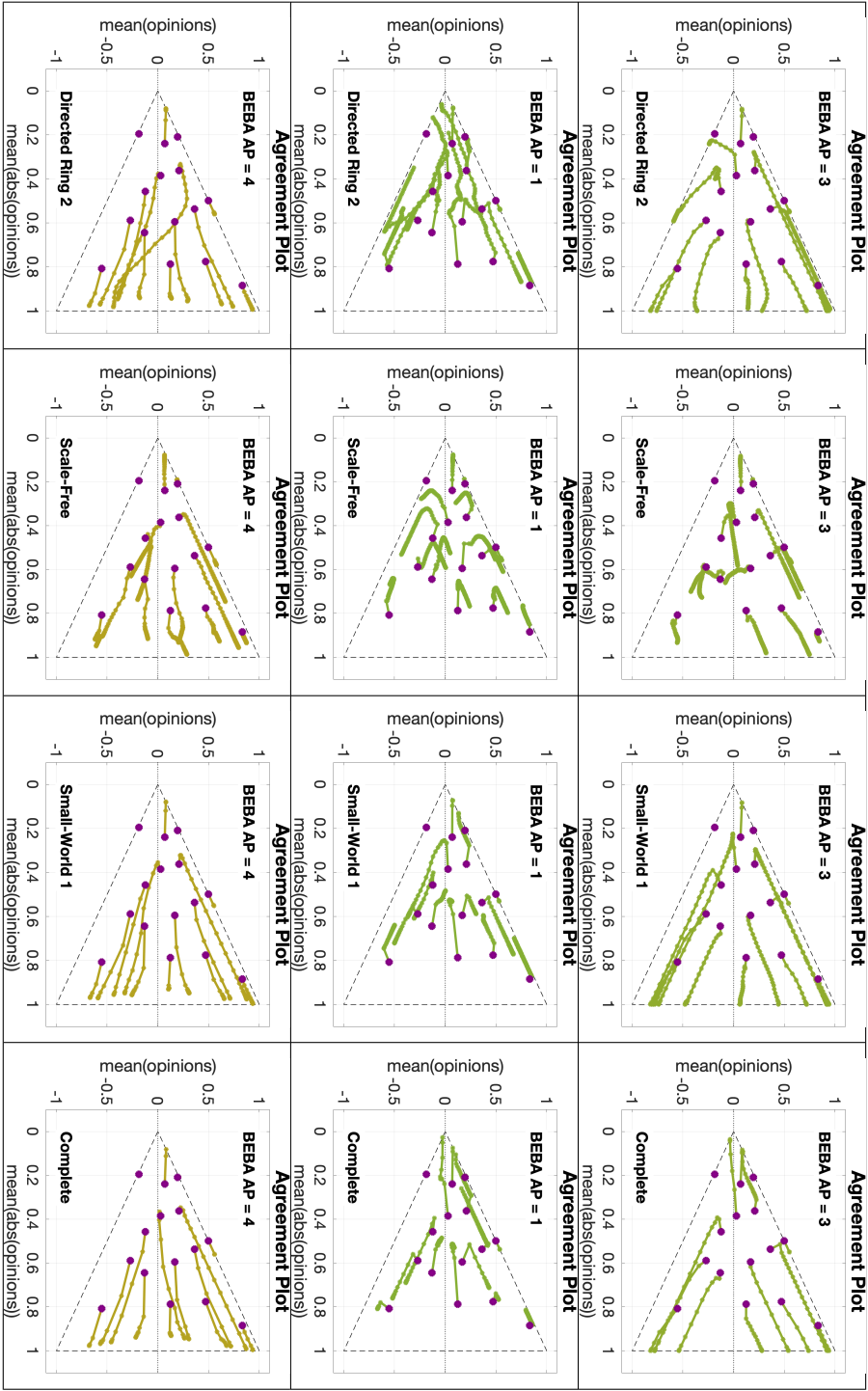


Table 6.1: *IOTE* plots for the Backfire Effect and Biased Assimilation model. Each of the 12 *IOTE* plots includes 15 curves with different choices of initial opinion distributions and constant agent parameters and underlying digraphs. Plots in the same row have the same agent parameters. Plots along the same column have the same underlying digraph, (from left to right, Directed Ring 2, Scale-Free, Small-World 1, and Complete, for metrics see Table 6.1). The agent parameter number can be seen in the upper left corner. Figure 6.21 shows a histogram of the corresponding agent parameters. The digraph name can be seen in the bottom left corner, the corresponding topology is shown in Figure 5.1, and some metrics are presented in Table 6.1. All simulations are with 100 agents for 50 time steps.

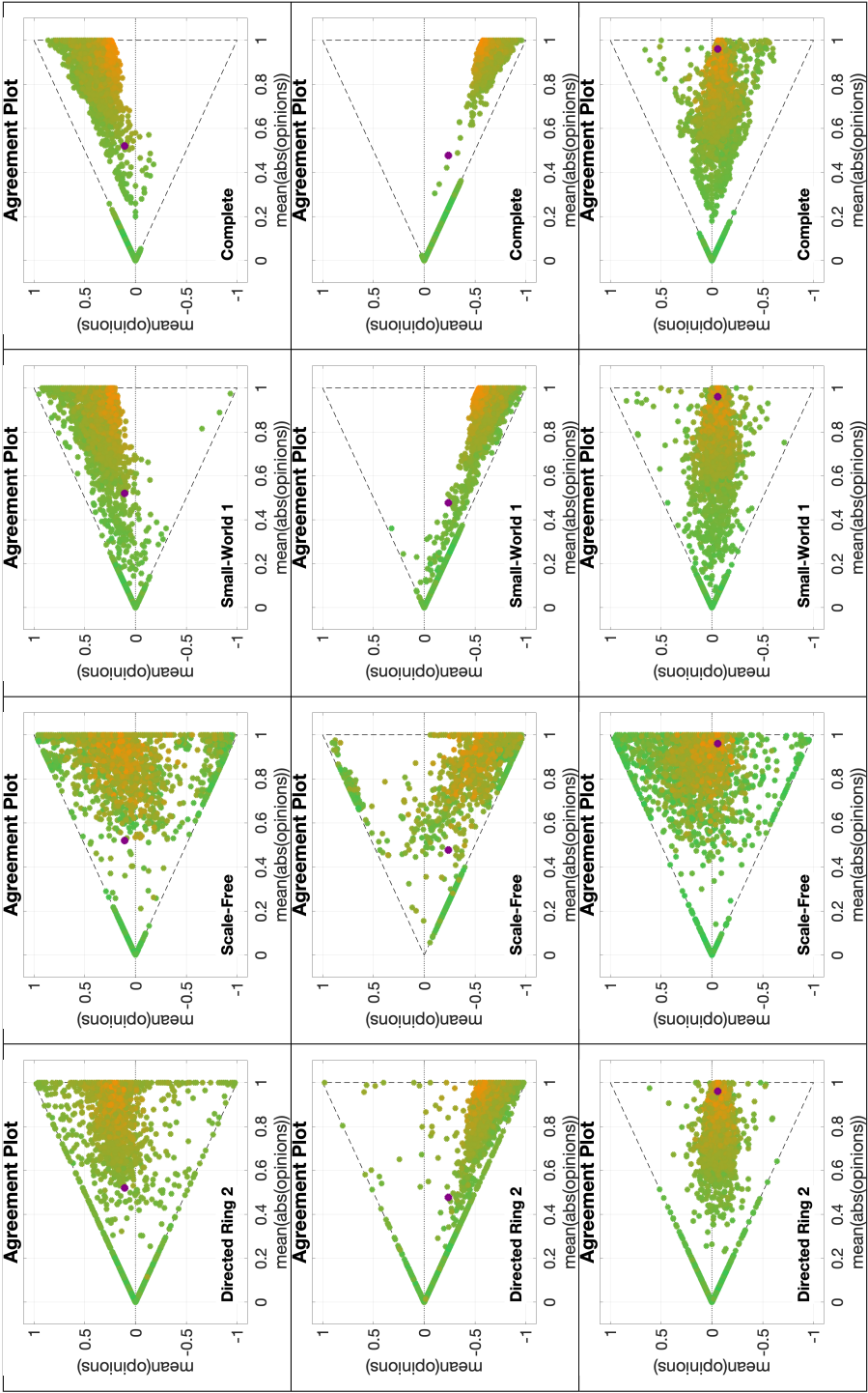


Table 6.12: APSS plots for the Backfire Effect and Biased Assimilation model. Each of the 12 APSS plots includes 3528 points with different choices of agent parameters and constant initial opinion distributions and underlying digraphs. Plots in the same row start from the same initial opinion distribution (represented by the purple dot). Plots along the same column have the same underlying digraph, (from left to right, Directed Ring 2, Scale-Free, Small-World 1, and Complete, for metrics see Table 6.1). The digraph name can be seen in the bottom left corner, the corresponding topology is shown in Figure 5.1, and some metrics are presented in Table 6.1. All simulations are with 100 agents for 1000 time steps.

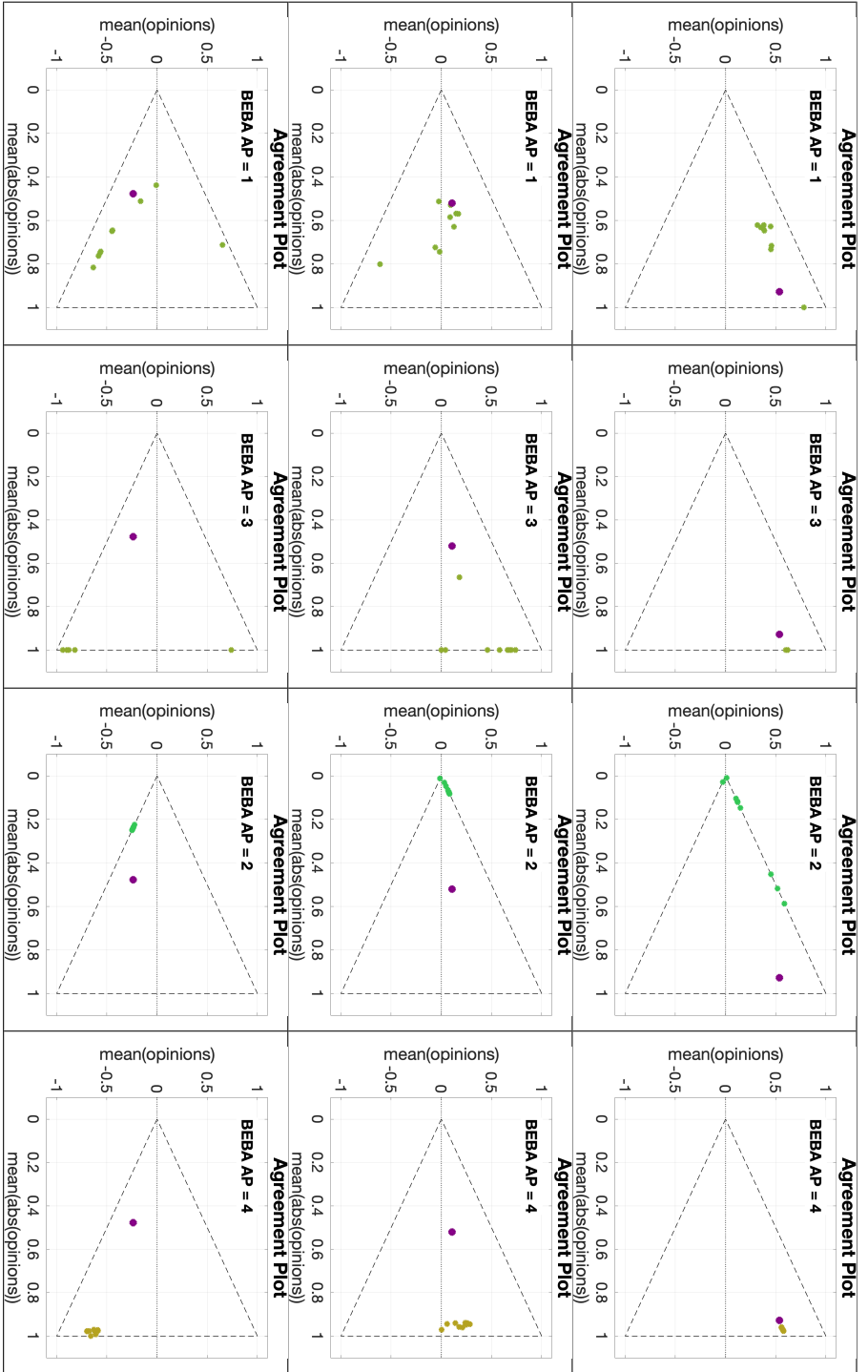


Table 6.13: UDSS plots for the Backfire Effect and Biased Assimilation model. Each of the 12 UDSS plots includes 9 points with different choices of underlying digraphs and constant initial opinion distributions and agent parameters. Plots in the same row start from the same initial opinion distribution (represented by the purple dot). Plots along the same column have the same agent parameters. The agent parameter number can be seen in the upper left corner. Figure 6.21 shows a histogram of the corresponding agent parameters. All simulations are with 100 agents for 1000 time steps.

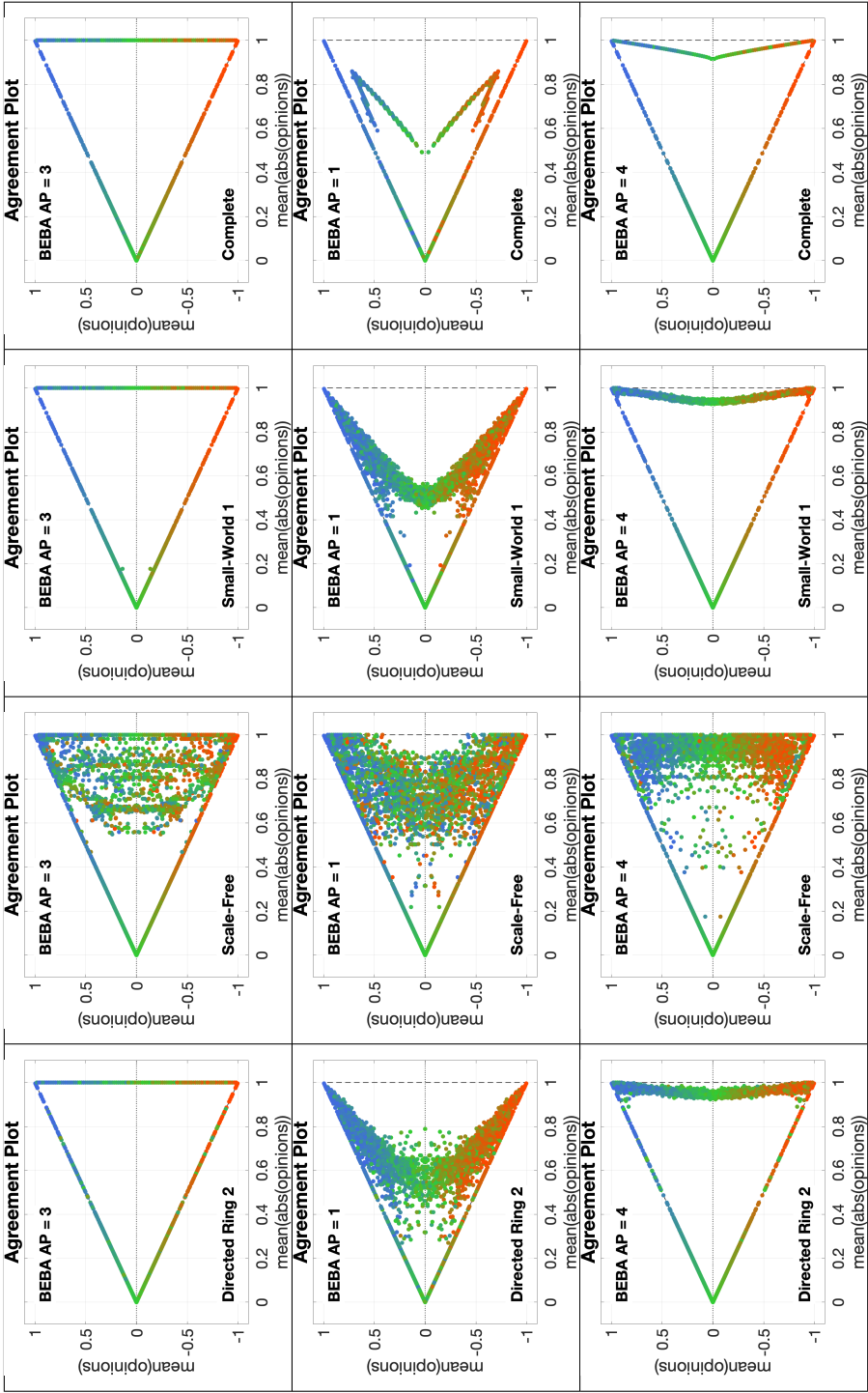


Table 6.14: IOSS plots for the Backfire Effect and Biased Assimilation model. Each of the 12 IOSS plots includes 5314 points with different choices of initial opinion distributions and constant agent parameters and underlying digraphs. Plots in the same row have the same agent parameters. Plots along the same column have the same underlying digraph, (from left to right, Directed Ring 2, Scale-Free, Small-World 1, and Complete, for metrics see Table 6.1). The agent parameter number can be seen in the upper left corner. Figure 6.21 shows a histogram of the corresponding agent parameters. The digraph name can be seen in the bottom left corner, the corresponding topology is shown in Figure 5.1, and some metrics are presented in Table 6.1. All simulations are with 100 agents for 1000 time steps.

analysis, hence a similar explanation could apply, i.e. that the opinions tend to contract and move towards consensus.

Looking now at the agent parameters of row 2, the two groups start to make more sense. From the agent parameter histogram it is clear that, in these particular societies, some agents have a very low entrenchment, while others have a very high entrenchment. One possible explanation for this 'two group' behaviour is that, depending on where the initial opinions are located (in the $(0,0)$, $(1,-1)$, $(1,1)$ triangle), one of the two groups of agents has a predominant effect. When the initial opinions are somewhat polarised, the backfire effect has nowhere to move the opinions, and thus the predominant effect is due to less entrenched agents. On the other hand, when the initial opinions are near consensus or perfect consensus, the biased assimilation effect of the highly entrenched agents is predominant and moves the opinion distribution to consensus or perfect consensus, effectively moving them to the lines $y = \pm x$.

The analysis of Tables 6.9 to 6.14 has provided insight into the opinion evolution capabilities of the BEBA model for a variety of initial opinions, agent parameters, and underlying digraphs. This model presents a rich and varied behaviour, resulting from the inclusion of both the backfire effect and the biased assimilation mechanism in the opinion formation process. In particular, the backfire effect allows for opinion distributions to achieve extreme opinion values, resulting in perfect consensus and polarisation. On the other hand, the biased assimilation mechanism moves opinions towards consensus on opinions that can have any value in the interval $[-1,1]$. The underlying digraph plays a fundamental and indirect role in the model behaviour. Unlike the Friedkin-Johnsen model, not only the Scale-Free topology results in a peculiar behaviour, but also the other considered topologies.

Now we can proceed with the Probabilistic Analysis of the BEBA model.

Probabilistic Analysis. Figures 6.22, 6.23, and 6.24 are analogous to Figures 6.18, 6.19, and 6.20 obtained in the Probabilistic Analysis of the Friedkin-Johnsen model. The Agreement Plot ordering of Figure 6.22 is quite similar to the one in Figure 6.18, therefore, the interpretation is the same: the first rows in the **QOL Tables** will correspond to initial opinions forming perfect consensus. Moving to towards the last rows the initial opinion goes from perfect consensus, to consensus, dissensus and clustering, and finally, polarisation.

On the other hand, the agent parameters ordering is quite different from the one for the Friedkin-Johnsen model. It is highly dependent on both the mean and variance of the agents' entrenchment. The first columns correspond to low values of entrenchment and as the ordering increases two groups appear, one with high mean entrenchment and the other with intermediate-high mean entrenchment and high variance. Therefore, when analysing the **QOL Tables** for the BEBA model, we need to take into account that right columns may correspond to societies where the mean entrenchment is very high, or where only *some* agents have a very high entrenchment (this would explain the mean of around 4 and high variance).

Figure 6.23 shows a model behaviour consistent with what was observed in the Agreement Plot analysis: the BEBA model has a strong tendency to move opinion distributions to the lines $y = \pm x$. Opinion distributions near these lines are mostly classified as consensus or perfect consensus. Figure 6.23 indicates that indeed most of these opinion distributions become perfect consensus.

According to Figure 6.23, the BEBA model results almost always in perfect consensus, and in the rare cases this is not the outcome, then the predicted opinions form polarisation. Initially, this can be somewhat unexpected, as the backfire effect and biased assimilation traits are generally associated with polarisation. In order to explain these results it is important to keep in mind that polarisation not only requires the existence of two 'distant enough' subgroups in the society, but also that these subgroups must have a comparable population. If this last condition is not met (one group has significantly more individuals than the other), then the opinion distribution is classified as perfect consensus.

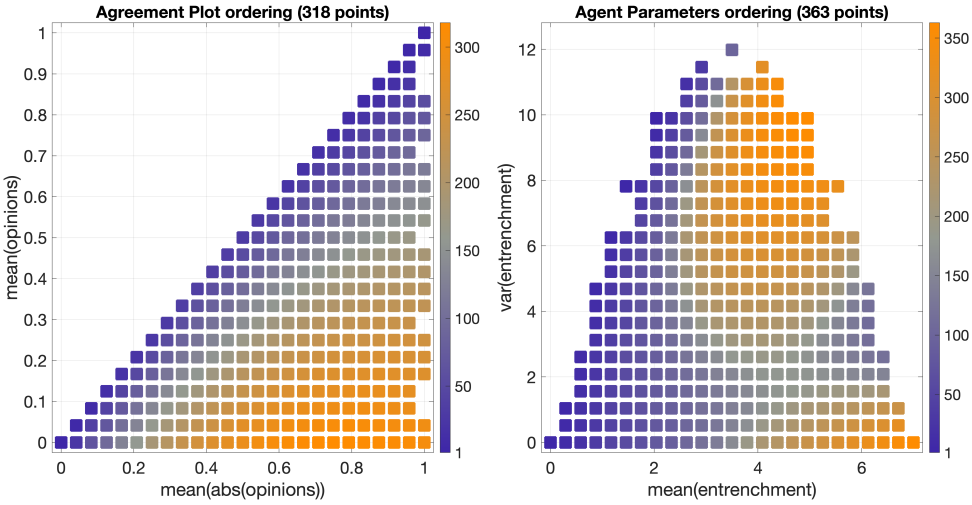


Figure 6.22: Ordering of the points in the Agreement Plot and Parameter Plane, for the Backfire Effect and Biased Assimilation model. Each point in the Agreement Plot corresponds to a pair $(\bar{x}, |\bar{x}|)$ that creates a non-empty set of opinion distributions $\mathcal{O}$. Each point in the Parameter Plane corresponds to a pair $(\beta, \sigma(\beta))$ that creates a non-empty set of agent parameter sets $\mathcal{P}$. The ordering is indicated by the colorbars and will be used in the plotting of the corresponding QOL Tables in Figure 6.23 as explained in Figure 6.17.

6

Although backfire effect and biased assimilation promote the existence of these two separate subgroups, there is no intrinsic mechanism to equalise the number of individuals in each one. And since random assignments most likely produce unequal subgroups, most of them are sorted as perfect consensus, as seen in Figure 6.23. The conditions for which the BEBA model most likely produces polarisation can be inferred from Figure 6.23. As evidenced by the middle image, polarisation mostly occurs in the lower right corner, which corresponds to initial opinions near the $(1, 0)$ point in the Agreement Plot, and agent parameters with either high mean entrenchment or intermediate-high mean entrenchment and high variance. This is not surprising: if the population is already polarised, adding high entrenchment ensures that the two separate subgroups forming the initial polarisation distribution are maintained.

Figure 6.24 (row 2, column 1) shows that perfect consensus is the most likely outcome even in the presence of high entrenchment. This further corroborates the conclusions drawn from Figure 6.23, that high entrenchment does not necessarily lead to polarisation. From row 2 in Figure 6.24 it is interesting to note that, when all initial opinions are considered, only intermediate-low mean values of entrenchment with high variance have a significant probability of producing opinions that are not sorted as perfect consensus.

After the Agreement Plot and Probabilistic Analysis, we can summarise the BEBA model behaviour and intrinsic characteristics in the following observations:

- The backfire effect and biased assimilation traits result in a model where agents have a tendency to (i) go to extreme opinions, or (ii) form perfect consensus. This behaviour is particularly evident in Table 6.14 (where the final opinions are located along the boundary of the $(0, 0)$, $(1, -1)$, $(1, 1)$ triangle), and Figure 6.23 (where perfect consensus is the most likely outcome).
- Although this tendency is intrinsic to the model, it can be affected by the underlying di-graph. The Scale-Free topology has a particularly noticeable effect on the model evolution,

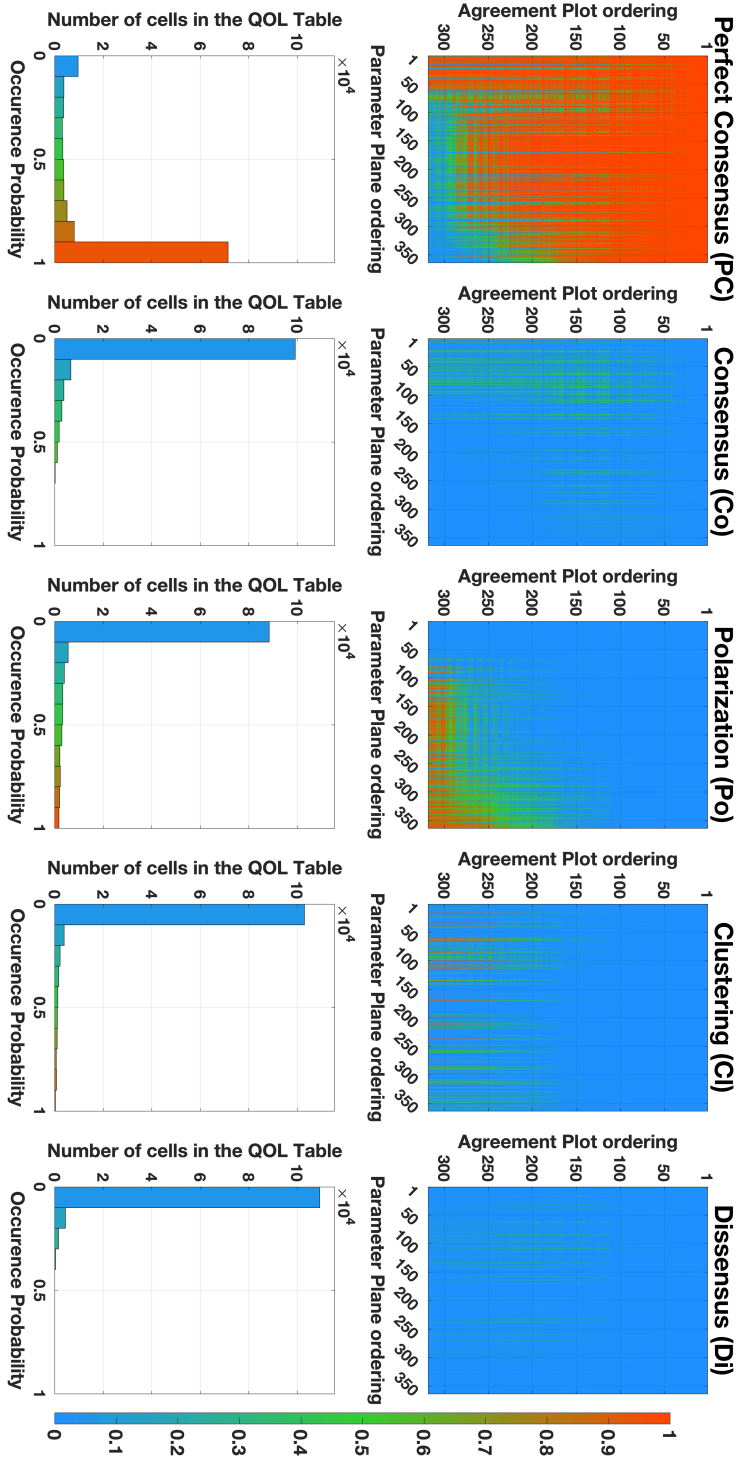


Figure 6.23: First row: **QOL Tables** for the Backfire Effect and Biased Assimilation model for $B \in \{PC, Co, Po, Ci, Di\}$. The rows corresponds to points $(\bar{x}, [x])$ in the Agreement Plot ordered according to Figure 6.22. The columns corresponds to points $(\bar{y}, \sigma(\bar{y}))$ in the Parameter Plane ordered according to Figure 6.22. Each cell is related to an initial opinion distribution and agent parameters. The cell colour represents the probability $\mathcal{P}_B$ that the resulting opinion distribution will be of qualitative type $B \in \{PC, Co, Po, Ci, Di\}$ as per Equation 6.2 (for details see Figure 6.17). Second row: the histograms corresponding to the **QOL Tables** presenting the number of cells in the table that have probability $\mathcal{P}_B$ in a given interval.

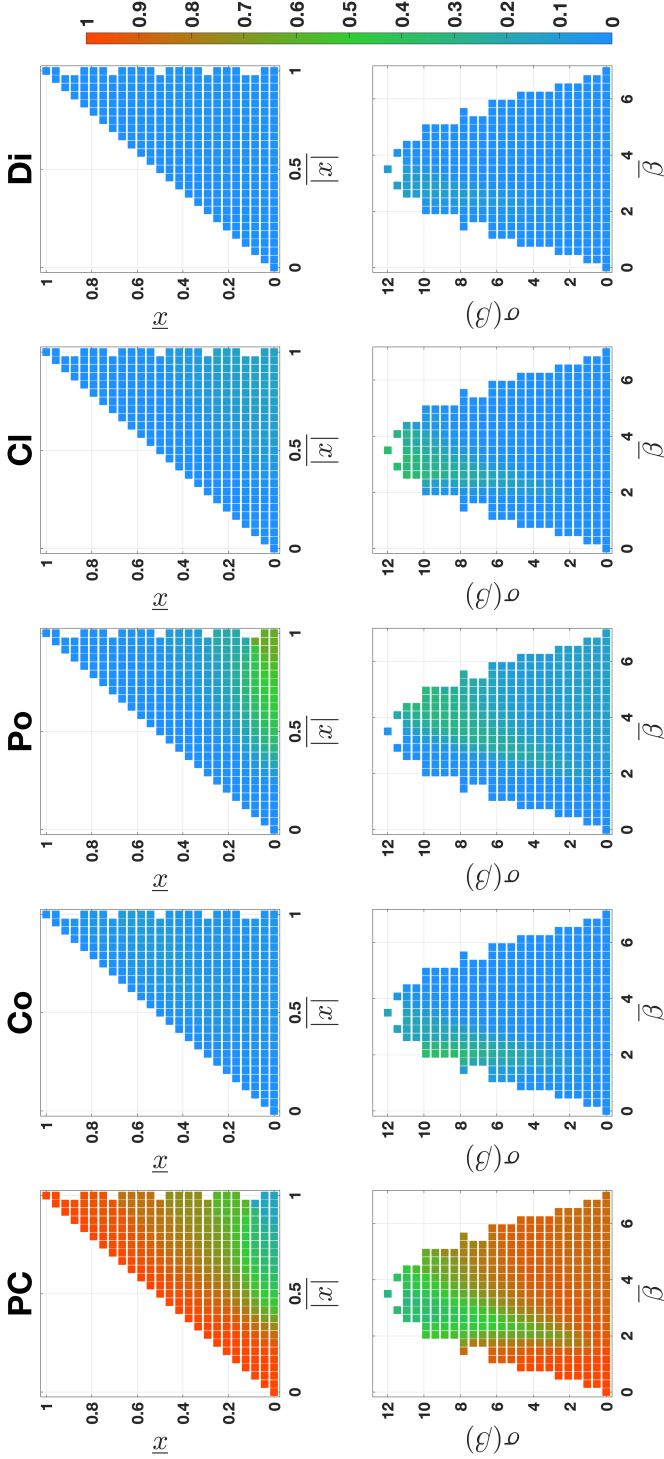


Figure 6.24: For the Backfire Effect and Biased Assimilation model, when agent parameter information (respectively, initial opinion) information is not available the probabilities $\mathcal{P}_B$ of the final opinions being of a qualitative type $B \in \{PC, Co, Po, CI, Di\}$ can be plotted in the Agreement Plot (resp. Parameter Plane) shown in Figure 6.22. These probability results are shown in the first (resp. second) row.

by allowing for more diverse opinion transformations.

- The underlying digraph has an indirect but significant effect on the overall model evolution. In a way, different digraph topologies enable the model to produce various outcomes via different agent parameters.
- Regarding the evolution of individual opinion distributions, depending on the agent parameters and the location of the initial opinion distribution in the Agreement Plot, the parametric curve corresponding to the opinion evolution moves to the left or right of the plot. This can be interpreted as opinions becoming more extreme, or going towards consensus.
- Overall the BEBA model has a great capacity of producing a wide range of quantitatively different opinions, evidenced for instance by Table 6.12, which shows that, starting from a single point in the Agreement Plot, the final opinions can be located almost everywhere in the $(0, 0)$, $(1, -1)$, $(1, 1)$ triangle.
- On the other hand, the qualitative range is way more limited, as evidenced by the **QOL Tables** in Figure 6.23. According to the tables and histograms in that figure, perfect consensus and polarisation are almost the two exclusive outcomes of the BEBA model.

6.5.2. BOUNDED CONFIDENCE MODEL

Unlike the Friedkin-Johnsen and BEBA models, for the Bounded Confidence model the digraph topology changes at each time-step, therefore there is no constant underlying digraph and, as a consequence, no **UDTE** or **UDSS** plots. Additionally, the **APTE**, **IOTE**, **APSS**, and **IOSS** plots will make no reference to a particular underlying digraph. Besides these changes, the analysis of this agent-based opinion formation model will follow the same structure as before. For the Bounded Confidence model, the agent parameter is the confidence radius, which is a number $r_i \in [0, 2]$. For the line and point colour coding, a confidence radius near 0 is represented by lavender colour, and a confidence radius near 2 by a pink colour.

Figure 6.25 shows the histograms corresponding to the agent parameters used for the **UDTE**, **IOTE**, **UDSS**, and **IOSS** plots for the Bounded Confidence model.

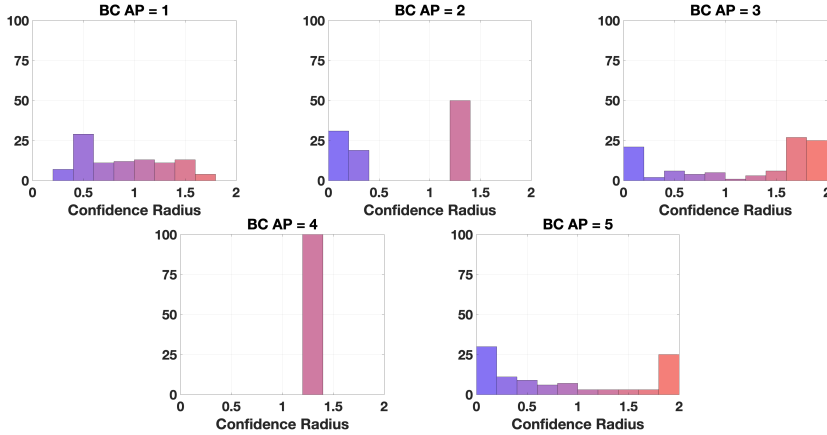


Figure 6.25: Histograms of the agent parameters used in the **IOTE**, and **IOSS** plots for the Bounded Confidence model (Tables 6.16, and 6.18 respectively).

Agreement Plot analysis. Starting with the Agreement Plot technique, Table 6.15 presents **APTE** plots for four different initial opinions. As with the Friedkin-Johnsen model, all the para-

metric curves move to the left, meaning that opinions get less extreme. This is due to the fact that the Bounded Confidence model has no mechanism to move opinions away from each other, so new opinions cannot exceed the initial maximum and initial opinions. In contrast to the Friedkin-Johnsen model, the parametric curves also have a significant component along the y -axis. A possible explanation for this is the following: if there are two subgroups of agents with distant opinions, one with agents having high confidence radius and the other with agents having low confidence radius, the first group will be influenced by the second group, but not vice-versa, resulting in a change in the mean of the opinion distribution.

Table 6.16 shows four **IOTE** plots for different choices of the agent parameters. The first notable observation is that, in some plots, some points seem not to have a parametric curve. This happens because for those opinions the agent parameters (confidence radius) are such that no agent influences each other, or if they do, their influence is minimal and therefore not noticeable. This explanation is further backed by the fact that the points without a visible trajectory are located to the right of the plots, where the mean of the opinions absolute values is relatively high, meaning that most agents have 'extreme' (and therefore, distant) opinions. Besides this observation, the plots in Figure 6.16 show a behaviour that is to be expected of a model that in essence averages opinions.

Table 6.17 presents four **APSS** plots for different initial opinions. The resulting plots are similar to the ones obtained for the Friedkin-Johnsen model in Table 6.5, with a key difference: some steady-state points can be located to the right of the initial opinion point and not near the $y = \pm x$ lines. Although not a common outcome, it is also significant. The same explanation as before applies: one subgroup with extreme opinions can influence the other agents without itself being influenced and thus, move them to opinions with higher value, without the need of achieving consensus. An additional clue is in the point's colour. Most of the points located at the right of the initial opinion point and not near the $y = \pm x$ lines are lavender, meaning a low mean confidence radius, therefore, if only some agents have a relatively high confidence radius they will move towards the other agents that could possibly have a more extreme opinion.

Regarding the colour coding interpretation, most of the lavender points are near the initial opinion point. This makes sense: if the mean confidence radius is low, then very few edges will be formed and the opinions will not change significantly. On the other hand, the more pink points are located near the $y = \pm x$ lines. Once again, this makes sense because, for those systems, the opinions evolve similarly to the French-DeGroot model and thus move towards perfect consensus.

Table 6.18 shows four **IOSS** plots for different choices of the agent parameters. As with the **IOSS** plots for the BEBA model, the points in these plots can be divided into two groups: a contracting group and an invariant group. The contracting group behaves as with the Friedkin-Johnsen model, except that points are more mixed: some blue points have a positive y -axis value, and some red points have a negative y -axis value. This has the same explanation as parametric curves having a significant component along the y -axis. On the other hand, the invariant group is self explanatory: for those initial opinions the agents have little influence among themselves (because the confidence radius is lower than the opinion differences) and thus the opinions remain almost the same.

These explanations are further backed by the agent parameters in each plot. For the first plot some agent parameters possibly have a confidence radius of 2 and therefore are affected by every agent and tend to consensus, resulting in a complete contraction (which can influence other agents with narrower confidence radius to also move) and empty invariant group. For the third plot the maximum confidence radius is probably around 1.9, therefore opinions with an absolute value mean above 1.9 tend to remain invariant, while the other opinions contract. For the second and fourth plots the maximum confidence radius is even lower, resulting in an invariant group consisting of more initial points.

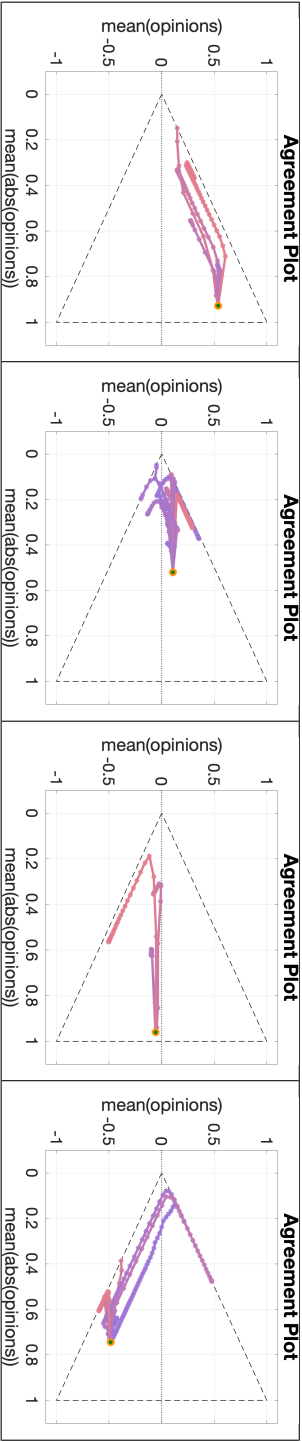


Table 6.15: **APTE** plots for the Bounded confidence model. Each of the 4 **APTE** plots includes 15 curves with different choices of agent parameters and constant initial opinion distributions. Since the Bounded Confidence model has no underlying digraph, this variable cannot be changed. Therefore, for this type of plot only the initial opinion can be changed. Hence, each column corresponds to a different initial opinion (shown in orange). All simulations are with 100 agents for 50 time steps.

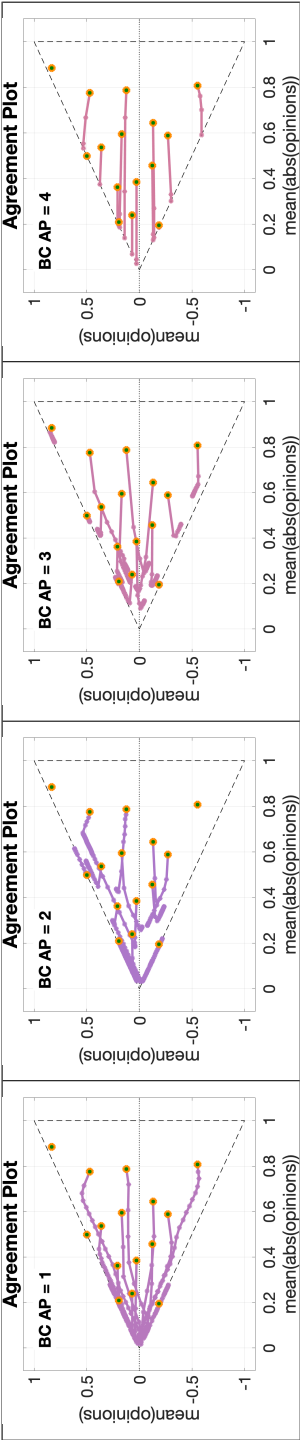


Table 6.16: **IOTE** plots for the Bounded confidence model. Each of the 4 **IOTE** plots includes 15 curves with different choices of initial opinion distributions and constant agent parameters. Since the Bounded Confidence model has no underlying digraph, this variable cannot be changed. Therefore, for this type of plot only the agent parameters can be changed. Hence, each column corresponds to a different agent parameter set. The agent parameter number can be seen in the upper left corner. Figure 6.25 shows a histogram of the corresponding agent parameters. All simulations are with 100 agents for 50 time steps.

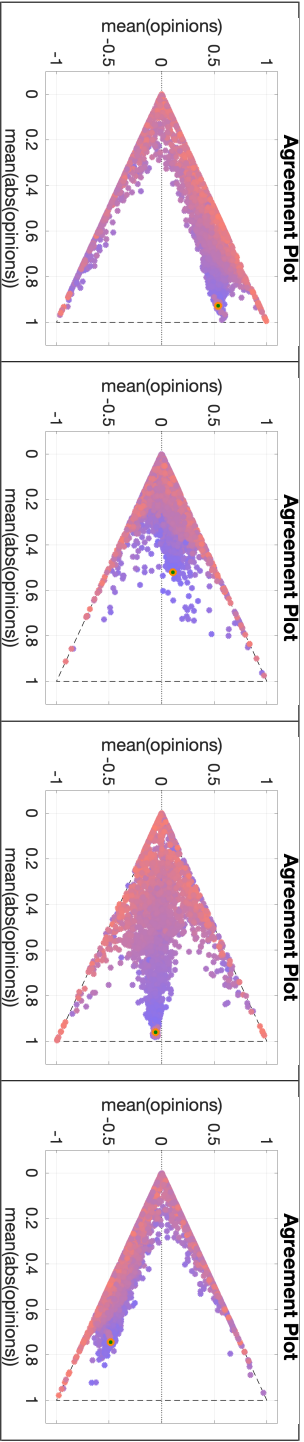


Table 6.1.7: **APSS** plots for the Bounded confidence model. Each of the 4 **APSS** plots includes 3528 points with different choices of agent parameters and constant initial opinion distributions and underlying digraphs. Since the Bounded Confidence model has no underlying digraph, this variable cannot be changed. Therefore, for this type of plot only the initial opinion can be changed. Hence, each column corresponds to a different initial opinion (shown in orange). All simulations are with 100 agents for 1000 time steps.

Unlike in the other plots, for the contracting group in the fourth plot, the opinions are exactly on the $y = \pm x$ lines. This happens because all agents have similar confidence radius and therefore the resulting digraphs are Complete and all agents converge to the same opinion. This does not happen in the other plots, where some agents have a very low confidence radius and therefore their opinion is almost invariant, preventing the steady state opinions from forming perfect consensus.

Probabilistic Analysis. Figures 6.26, 6.27, and 6.28 are analogous to Figures 6.18, 6.19, and 6.20 obtained in the Probabilistic Analysis of the Friedkin-Johnsen model.

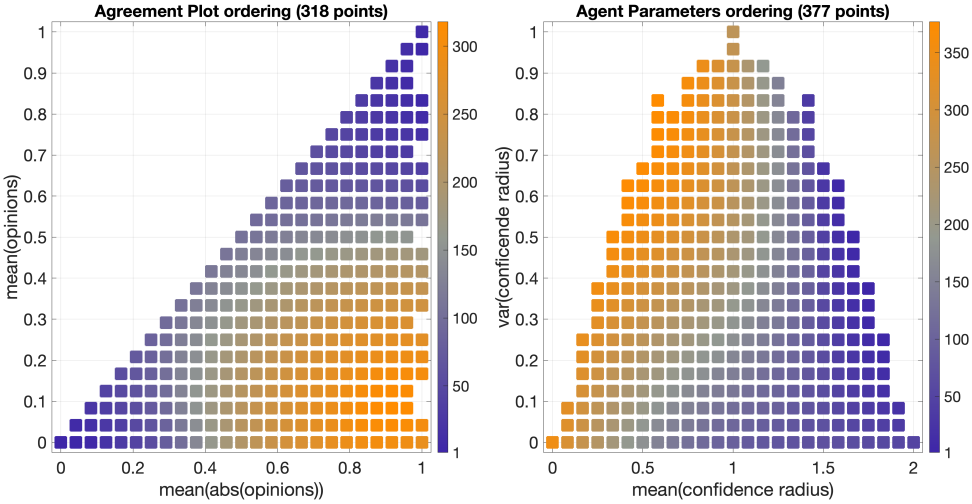


Figure 6.26: Ordering of the points in the Agreement Plot and Parameter Plane, for the Bounded Confidence model. Each point in the Agreement Plot corresponds to a pair $(\bar{x}, |\bar{x}|)$ that creates a non-empty set of opinion distributions $\mathcal{O}$. Each point in the Parameter Plane corresponds to a pair $(\bar{r}, \sigma(r))$ that creates a non-empty set of agent parameter sets $\mathcal{P}$. The ordering is indicated by the colorbars and will be used in the plotting of the corresponding **QOL Tables** in Figure 6.27 as explained in Figure 6.17.

The Agreement Plot ordering of Figure 6.26 means that the first rows will correspond to perfect consensus and will move through consensus, clustering, and dissensus until reaching polarisation for the last rows, which correspond to initial opinions near the point $(1, 0)$. The agent parameter ordering shows a significant dependence on the confidence radius variance: the first rows correspond to high mean radius with low variance and the last rows to intermediate-low mean confidence radius and high variance. This means that in the societies represented in the left columns most agents have the same confidence radius, around 2, and in societies to the right some agents have a high and others have a low confidence radius.

The **QOL Tables** in Figure 6.27 show that the Bounded Confidence model predominantly produces perfect consensus and that the second most likely outcome is clustering. Perfect consensus is almost guaranteed for the left part of the **QOL Tables**, that is, approximately when the mean confidence radius is above 1. This makes sense: with such a high mean confidence radius, even with high variance some agents will be influenced by almost all the other agents and cause global convergence to a single value, producing perfect consensus.

The **QOL Tables** show that clustering is almost guaranteed for the bottom right corner, that is, when the initial opinions are near the point $(1, 0)$ (most likely polarised), and the confidence radius has an intermediate-low mean and high variance. This happens because, when the initial opinions start polarised, the agents with intermediate confidence radius move towards the agents

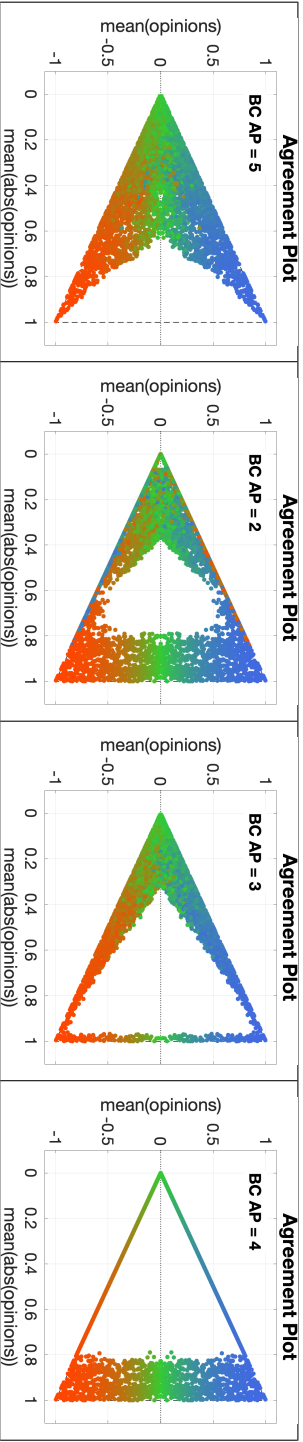


Table 6.18: **IOSS** plots for the Bounded confidence model. Each of the 4 **IOSS** plots includes 5314 points with different choices of initial opinion distributions and constant agent parameters and underlying digraphs. Since the Bounded Confidence model has no underlying digraph, this variable cannot be changed. Therefore, for this type of plot only the agent parameters can be changed. Hence, each column corresponds to a different agent parameter set. The agent parameter number can be seen in the upper left corner. Figure 6.25 shows a histogram of the corresponding agent parameters. All simulations are with 100 agents for 1000 time steps.

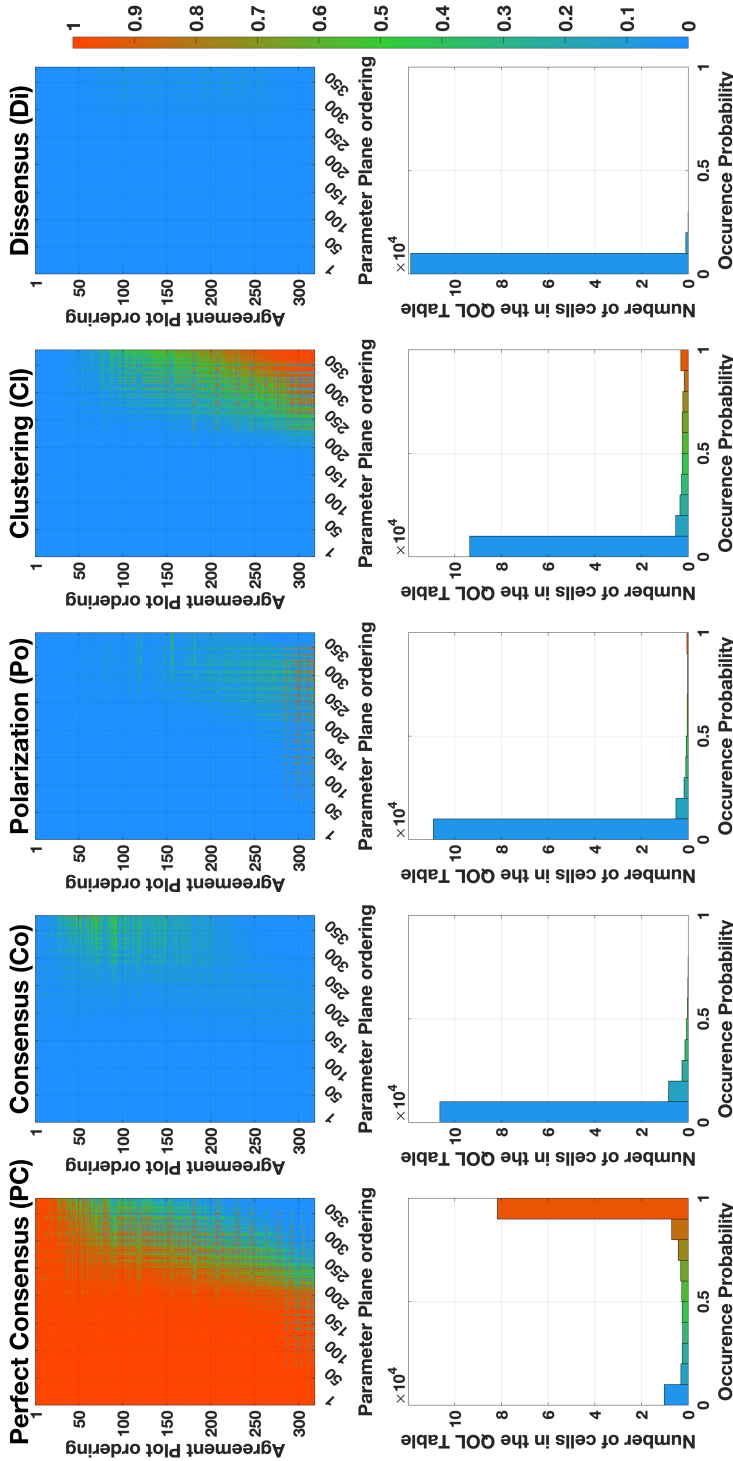


Figure 6.27: **QOL Tables** for the Bounded Confidence model for $B \in \{PC, Co, Po, Cl, Di\}$. The rows corresponds to points $(\bar{x}, |\bar{x}|)$ in the Agreement Plot ordered according to Figure 6.26. The columns corresponds to points $(\bar{r}, \sigma(\bar{r}))$ in the Parameter Plane ordered according to Figure 6.26. Each cell is related to an initial opinion distribution and agent parameters. The cell colour represents the probability $\mathcal{P}_B$ that the resulting opinion distribution will be of qualitative type $B \in \{PC, Co, Po, Cl, Di\}$ as per Equation 6.2 (for details see Figure 6.17). Second row: the histograms corresponding to the **QOL Tables** presenting the number of cells in the table that have probability $\mathcal{P}_B$ in a given interval.

with low confidence radius, which remain in place, as they are influenced by very few agents. This behaviour has the potential of moving agents away from the two initial polarised groups and thus lead to clustering.

From Figure 6.27 it is also clear that the other qualitative outcomes are extremely unlikely, because, in order to exist, they would need to start from very specific initial opinions and the initial opinions and agent parameters would need to be precisely allocated, which is not likely to happen at random.

Figure 6.28 provides more specific information about the conditions for which perfect consensus is not the almost sure outcome. This happens when the variance is maximal and the mean confidence radius is below 1. In that region there is also a significant probability of obtaining clustering, and when the mean confidence radius is below 0.5 there is a slight probability that the opinions form polarisation. This, however, is likely not because opinion distributions from different categories become polarisation, but because polarised opinions remain polarised.

To summarise, the conclusions drawn from the Bounded Confidence's behaviour and intrinsic properties using the Agreement Plot and Probabilistic Analysis techniques are:

- Although the Bounded Confidence model uses the same weighted sum law to evolve opinions (representing, thus, the trend towards agreement), as the French-DeGroot model, the possibility of some agents having a low confidence radius and others a high confidence radius allows the model to escape the 'averaging tendency' (the tendency of evolving opinions to have a near constant mean, represented by an almost horizontal movement along the Agreement Plot), and instead have a significant displacement along the y -axis of the Agreement Plot, i.e., significantly change the opinion's mean. This is because agents with low confidence radius can influence all the other agents without necessarily being influenced themselves. Hence, if there exists a group of agents with similar opinions and low confidence radius, that influence the other agents of the population, unavoidably, the mean of the population opinion will move towards the opinions of these group of low confidence radius agents.
- The lack of an underlying digraph both simplifies and limits the model. As noted for the Friedkin-Johnson and BEBA models, the digraph had a significant indirect effect on the model dynamics by 'modulating' the model response to different agent parameters and initial opinions. This cannot happen for the Bounded Confidence model, and thus one could say that it has only one response to agent parameters and initial opinions.
- Depending on the agent parameters, a significant percentage of possible initial opinion distributions may remain constant for the Bounded Confidence model. The lower the maximum confidence radius, the more likely societies with extreme opinions remain invariant as agents from opposite groups do not influence each other.
- A uniformly distributed agent parameter set creates more variety in the predicted opinions, because heterogeneous confidence radii result in a network that not necessarily is strongly connected and therefore allows agents to have different opinions. On the other hand, if all the confidence radii are similar, then the agents are either all isolated or strongly connected.

6.6. CONCLUSIONS

This chapter introduced and explained four agent-formation model analysis techniques: the Histogram-based Sorting Algorithm, the Transition Tables, the Agreement Plot, and the Probabilistic Analysis. These four techniques collectively form a framework that can investigate the intrinsic properties of agent-based opinion formation models. They rely on numerical simulations to examine the models from different perspectives.

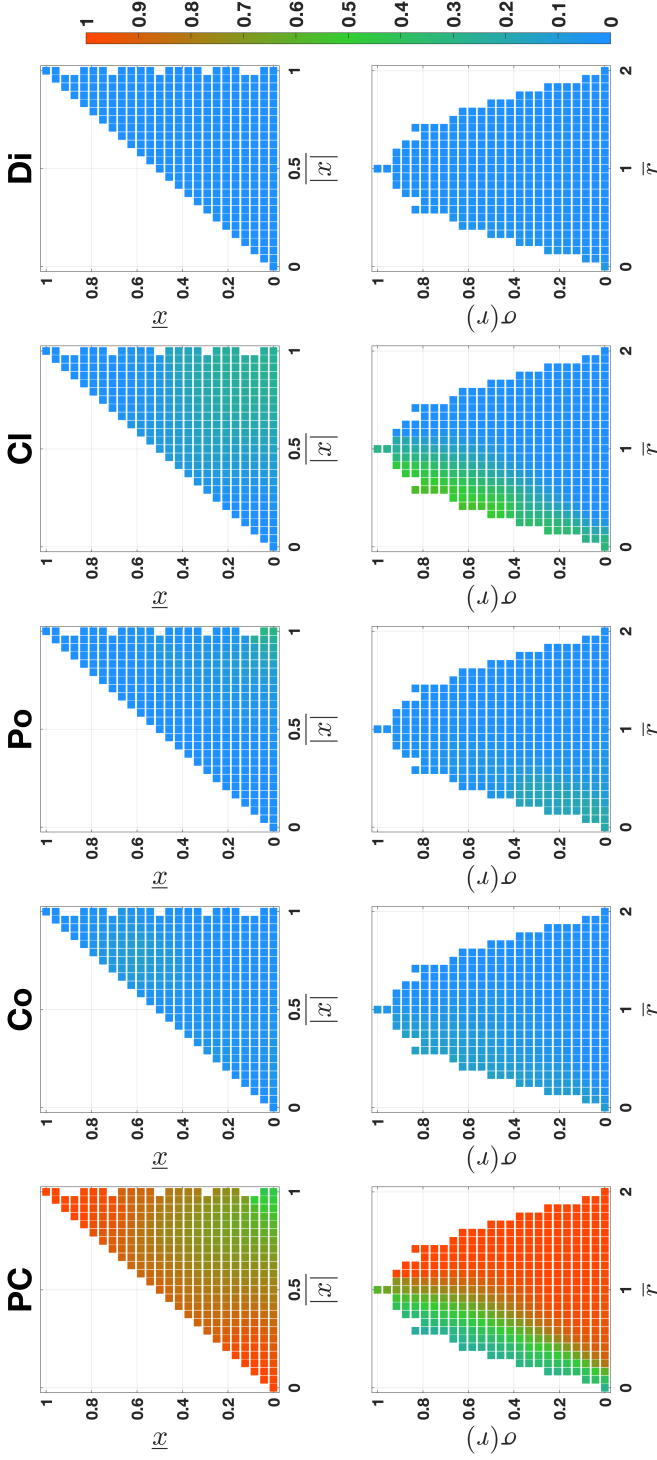


Figure 6.28: For the Bounded Confidence model, when agent parameter information (respectively, initial opinion) information is not available the probabilities $\mathcal{P}_B$ of the final opinions being of a qualitative type $B \in \{PC, Co, Po, Cl, Di\}$ can be plotted in the Agreement Plot (resp, Parameter Plane) shown in Figure 6.26. These probability results are shown in the first (resp, second) row.

We illustrated all the techniques with existing models in the literature, and we gained insight into the models and real opinion formation.

The first avenue of future work is to apply these techniques to other models and implement the required adaptations to put them into use with different opinion formation models (for which agents could have more than one parameter, for example). Another possibility is to export these ideas to other fields that use agent-based models, for instance, in disease spreading and epidemic modelling.

Chapter 7 will use the framework described in this chapter to analyse a newly proposed opinion formation model: the Classification-based model.

7

CLASSIFICATION-BASED OPINION FORMATION MODEL EMBEDDING AGENTS' PSYCHOLOGICAL TRAITS

One often meets his destiny
on the road he takes to avoid it.

Master Oogway, Kung Fu Panda

We propose an agent-based opinion formation model characterised by a two-fold novelty. First, we realistically assume that each agent cannot measure the opinion of its neighbours with infinite resolution and accuracy, and hence it can only classify the opinion of others as agreeing much more, or more, or comparably, or less, or much less (than itself) with a given statement. This leads to a classification-based rule for opinion update. Second, we consider three complementary agent traits suggested by significant sociological and psychological research: conformism, radicalism and stubbornness. We rely on World Values Survey data to show that the proposed model has the potential to predict the evolution of opinions in real life: the classification-based approach and complementary agent traits produce rich collective behaviours, such as polarisation, consensus, and clustering, which can yield predicted opinions similar to survey results.

This chapter is based on the manuscript "Classification-Based Opinion Formation Model Embedding Agents' Psychological Traits" [58] by Carlos Andrés Devia and Giulia Giordano, and on the work-in-progress paper "Graphical and Probabilistic Analysis of Agent-Based Opinion Formation Models" [59], by Carlos Andrés Devia and Giulia Giordano.

This Chapter is structured in the following way. In Section 7.1 the model is described and introduced. Section 7.2 presents the simulation results. These results consist of five parts: first, simulations over simple cases to develop an intuition on how the model behaves in simplified cases; next, a parameter sensitivity analysis that studies how different parameter changes affect the model behaviour; then, model validation with real data, where survey results from the World Value Survey data are used to determine the potential that the Classification-based model has of mimicking opinion evolutions seen in real societies; after this, simulations comparing the Classification-based model with the Friedkin-Johnsen model; and finally, results showing the model outcome capabilities. Section 7.3 discusses summary and conclusions. Throughout the simulation results the different analysis techniques described in Chapter 6 are used.

7.1. THE CLASSIFICATION-BASED MODEL

In our proposed classification-based (CB) model, the set $\mathcal{V} = \{1, 2, \dots, N\}$ indexes the agents. The opinion of agent $i \in \mathcal{V}$ at time k , representing its level of agreement with a statement, is denoted by $x_i[k] \in [-1, 1]$. The opinions $x_i = 1$, $x_i = 0$, and $x_i = -1$ represent complete agreement, indifference, and complete disagreement respectively. The vector of all opinions at time k is denoted by $x[k]$.

The agent opinions evolve in time due to opinion exchanges occurring over a signed digraph, represented by the matrix $W \in \{-1, 0, 1\}^{N \times N}$, whose entries are constant and, in particular, not opinion-dependent. The self-confidence of each agent is expressed by $w_{ii} = 1$ for all i . The coefficient w_{ij} represents the influence of agent j over agent i . If $w_{ij} = 0$, then agent i is not influenced by agent j . If $w_{ij} \neq 0$, then agent j is a neighbour of agent i : $w_{ij} = 1$ means that agent i approves, trusts, or follows agent j , while $w_{ij} = -1$ means that agent i disapproves, mistrusts, or antagonises agent j . Signed edges have been interpreted in the opinion formation literature in terms of either cooperative/antagonistic interactions [5], trust/mistrust [218], or approval/disapproval [36]. In our model, if $w_{ij} = 1$ (respectively $w_{ij} = -1$), then agent i perceives the opinion of agent j as x_j (resp. $-x_j$). The set of neighbours of agent $i \in \mathcal{V}$ is

$$\mathcal{N}_i = \{j \in \mathcal{V} \mid w_{ij} \neq 0\}. \quad (7.1)$$

The agent opinions evolve in discrete time and the opinion update relies on the assumption that agents cannot determine their neighbours' opinions precisely. Instead, each agent can classify its neighbours according to how close their *perceived* opinion is to its own opinion. For instance, if agent j influences agent i , and $x_i = 0.61$ and $x_j = 0.34$, then it is unrealistic to expect agent i to know *exactly* the opinion of agent j , or to assume that agent i knows that the opinion difference is *exactly* 0.27. However, agent i can perceive that agent j agrees less than itself. On the contrary, if $x_j = 0.89$, agent i can perceive that agent j agrees more than itself.

Therefore, agent i can at most classify agent j according to an estimation of Δ_{ij} , which is the weighted difference between its opinion x_i and the opinion of agent j , x_j : $\Delta_{ij} = x_i - w_{ij}x_j \in [-2, 2]$. Let us divide the interval $[-2, 2]$ in five equal subintervals. Then, depending on the subinterval to which Δ_{ij} belongs, agent i can *perceive* that agent j : (1) agrees much more, (2) agrees more, (3) agrees comparably, (4) agrees less, or (5) agrees much less with the statement; see Figure 7.1. If $w_{ij} = -1$, then agent i disapproves/mistrusts/antagonises agent j , therefore the weighted opinion difference is $\Delta_{ij} = x_i - (-x_j) \in [-2, 2]$. If $w_{ij} = 1$, then agent i approves/trusts/follows agent j and the weighted opinion difference is $\Delta_{ij} = x_i - x_j \in [-2, 2]$. The combined effect of signed edges and neighbour classification leads to a three-step process: first, agent i perceives the opinions of its neighbours; then, the opinions of neighbours that agent i disapproves, mistrusts, or antagonises have the sign reversed; finally, the neighbours are classified according to the adjusted perceived opinion distance.

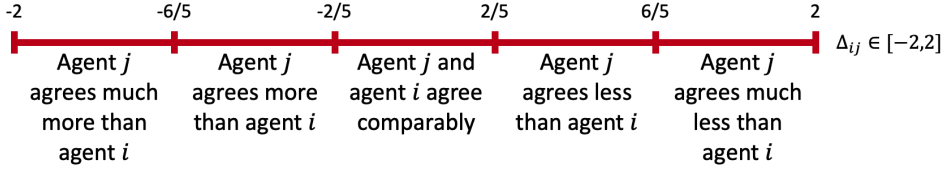


Figure 7.1: Partition of the interval $[-2, 2]$ in five equal subintervals. Depending on the interval to which the weighted opinion difference $\Delta_{ij} = x_i - w_{ij}x_j$ belongs, agent i will *perceive* that agent j agrees either: *much more*; or *more*; or *comparably*; or *less*; or *much less*.

The set $\mathcal{N}_i$ of all the neighbours of agent i is thus partitioned into five time-dependent subsets: $D_i^+[k]$, $D_i[k]$, $N_i[k]$, $A_i[k]$, and $A_i^+[k]$, which contain the neighbours that agree much less, less, comparably, more, and much more, respectively. Mathematically these subsets are defined as

$$\begin{aligned}
 D_i^+[k] &= \{j \in \mathcal{N}_i \mid 6/5 \leq \Delta_{ij}[k] \leq 2\} \\
 D_i[k] &= \{j \in \mathcal{N}_i \mid 2/5 \leq \Delta_{ij}[k] < 6/5\} \\
 N_i[k] &= \{j \in \mathcal{N}_i \mid -2/5 < \Delta_{ij}[k] < 2/5\} \\
 A_i[k] &= \{j \in \mathcal{N}_i \mid -6/5 < \Delta_{ij}[k] \leq -2/5\} \\
 A_i^+[k] &= \{j \in \mathcal{N}_i \mid -2 \leq \Delta_{ij}[k] \leq -6/5\}
 \end{aligned} \tag{7.2}$$

where $\Delta_{ij}[k] = x_i[k] - w_{ij}x_j[k]$. The cardinality of these sets has the following interpretation:

- $|D_i^+[k]|$ = number of neighbours that agent i *perceives* as agreeing much less than itself at time k
- $|D_i[k]|$ = number of neighbours that agent i *perceives* as agreeing less than itself at time k
- $|N_i[k]|$ = number of neighbours that agent i *perceives* as agreeing the same as itself at time k
- $|A_i[k]|$ = number of neighbours that agent i *perceives* as agreeing more than itself at time k
- $|A_i^+[k]|$ = number of neighbours that agent i *perceives* as agreeing much more than itself at time k

The overall behaviour of each agent results from the combination of three complementary inner traits: *conformism*, leading the agent to agree with its neighbours; *radicalism*, driving the agent to reinforce its opinion; and *stubbornness*, anchoring the agent to its current opinion. The conformism, radicalism and stubbornness degree of agent i is respectively denoted by α_i , β_i and γ_i . The parameters $\psi_i = (\alpha_i, \beta_i, \gamma_i)$, quantifying the *inner traits* of agent i , satisfy $\alpha_i, \beta_i, \gamma_i \in [0, 1]$ and $\alpha_i + \beta_i + \gamma_i = 1$ for all i . We call *inner traits assignation* the collection of inner traits of all agents, $\psi := (\psi_i)_{i \in \mathcal{V}}$. The model features are summarised in Figure 7.2.

The opinion change $\Delta x_i[k]$ of agent i at time k is thus the convex combination of the behaviour of a purely conformist, purely radical, and purely stubborn agent,

$$\Delta x_i[k] = \alpha_i f_i^{\text{con}} + \beta_i f_i^{\text{rad}} + \gamma_i f_i^{\text{stb}}, \tag{7.3}$$

with f_i^{con} , f_i^{rad} , and f_i^{stb} taken as

$$f_i^{\text{con}} = \frac{\lambda}{|\mathcal{N}_i|} \left(\xi |A_i^+| + |A_i| - |D_i| - \xi |D_i^+| \right), \quad f_i^{\text{rad}} = \frac{\lambda}{|\mathcal{N}_i|} \mu |N_i| x_i[k], \quad f_i^{\text{stb}} = 0, \tag{7.4}$$

where λ , ξ , and μ are positive parameters: λ weighs the overall opinion change magnitude, ξ weighs the increased influence that neighbours with distant opinions have over conformist traits, and μ weighs the influence of the agent's own opinion in radical traits. We call these *opinion evolution parameters*: $\Omega = (\lambda, \xi, \mu)$.

To better understand Equations (7.4) and choose reasonable values for the parameters, one can think of how an extreme agent ($\alpha_i = 1$, or $\beta_i = 1$, or $\gamma_i = 1$) behaves.

- A purely conformist agent ($\alpha_i = 1$, $\beta_i = 0$, $\gamma_i = 0$) evolves towards an opinion comparable to that of its neighbours. For instance, if $N_i = \mathcal{N}_i$ (all the neighbours of agent i agree comparably), then agent i does not change its opinion. If $A_i = \mathcal{N}_i$ (all the neighbours of agent i agree more), agent i increases its opinion x_i by λ ; given that all the neighbours of agent i are in the set A_i , a value $\lambda = 0.4$ guarantees that, if all the neighbour opinions remain unchanged, then at the next time step all the neighbours of agent i will be in the set N_i , hence perceived as having a comparable opinion. Instead, if $A_i^+ = \mathcal{N}_i$, then the opinion of agent i needs to increase $0.8 = 2\lambda$ in order to be perceived as comparable to its neighbours' at the next time step, and therefore a natural choice is $\xi = 2$. The same reasoning can be applied to the sets D_i and D_i^+ .
- A purely radical agent ($\alpha_i = 0$, $\beta_i = 1$, $\gamma_i = 0$) ignores neighbours with a different opinion and only cares about agents that think comparably to itself, hence it reinforces its current opinion $x_i[k]$ depending on the magnitude of its own opinion and on the fraction of its neighbours in the set N_i . To make sure that radical traits can affect the opinion change more strongly than conformist traits, we need $\mu > 1$. In fact, if $\mu = 1$, then $|f_i^{\text{rad}}| < |f_i^{\text{con}}|$ in general: the opinion change caused by the radical trait (which is proportional to $x_i[k]$, and $|x_i[k]| \leq 1$) is smaller in magnitude than the one caused by the conformist trait. In our simulations, we set $\mu = 5$. The effect of different values of μ can be seen in Table 7.7.
- A purely stubborn agent ($\alpha_i = 0$, $\beta_i = 0$, $\gamma_i = 1$) does not change its opinion under any circumstance.

The new opinion of agent i at time $k + 1$ is the sum of the previous opinion $x_i[k]$ and the opinion change $\Delta x_i[k]$, modulated by the saturation function σ

$$\sigma(x) = \begin{cases} x & \text{if } |x| \leq 1 \\ \text{sign}(x) & \text{if } |x| > 1 \end{cases} \quad (7.5)$$

so as to guarantee that the opinions remain in the interval $[-1, 1]$. The complete opinion update law is therefore

$$x_i[k+1] = \sigma \left(x_i[k] + \frac{\lambda}{|N_i|} \left(\alpha_i \xi (|A_i^+| - |D_i^+|) + \alpha_i (|A_i| - |D_i|) + \beta_i \mu |N_i| x_i[k] \right) \right), \quad \forall i \in \mathcal{V}. \quad (7.6)$$

7.1.1. MODEL PARAMETERS

The Classification-Based (CB) model has three types of parameters: the signed digraph weights w_{ij} ; the inner traits assignation $\psi_i = (\alpha_i, \beta_i, \gamma_i)$; and the opinion evolution parameters $\Omega = (\lambda, \xi, \mu) = (0.4, 2, 5)$ whose values are fixed, and chosen based on the model interpretation. Later, a parameter sensitivity analysis explores how the model evolution is affected by changes in opinion evolution parameters.

If the model has N agents, then:

- The signed digraph has weighted adjacency matrix (or weight matrix) $W \in \mathcal{W}_N$. In general, $\mathcal{W}_N = \{-1, 0, 1\}^{N \times N}$, but we can focus for instance on Small-World, or strongly connected, networks.

- The inner traits assignation is $\psi \in \mathcal{A}_N$, where

$$\mathcal{A}_N = \left\{ \psi = (\psi_i)_{i \in \mathcal{V}} = ((\alpha_i, \beta_i, \gamma_i))_{i=1}^N \mid \alpha_i, \beta_i, \gamma_i \in [0, 1] \text{ and } \alpha_i + \beta_i + \gamma_i = 1, \forall i \in \mathcal{V} \right\}. \quad (7.7)$$

We omit the subscript N from the sets $\mathcal{W}$ and $\mathcal{A}$ for simplicity. Given N agents, a signed digraph $W \in \mathcal{W}$, an inner traits assignation $\psi \in \mathcal{A}$, and a vector of initial opinions $x[0]$, the opinion formation model evolves according to Equation (7.6). The vector $x[K]$ of opinions after K iterations can be explicitly represented as a function of W , ψ , and $x[0]$ by the map $\mathcal{F}_\Omega(x[K])$ also depends on Ω , whose value, given by the model interpretation, is fixed) as

$$x[K] = \mathcal{F}_\Omega(x[0], W, \psi, K) \quad (7.8)$$

The value of K depends on the type of statements and the prediction horizon. For statements related to core values or beliefs, opinions are not expected to change very fast and one could consider roughly 10 changes per year (as done in the analysis of the Transition Tables, in Section 6.1.4). Therefore, if the model is used to predict the opinions after 5 years, $K = 50$. On the other hand, the opinions on more superficial topics could change faster and, over the same 5-year timespan, it could be $K = 500$. See Figure 7.2 for a summary of the model parameters and features.

To validate the model – namely, assess its potential to closely reproduce the evolution of opinions in real life with suitably chosen parameters – we consider *real* initial and final opinions, denoted by x and y respectively, taken from survey data. Assuming that y are the real opinions K iterations after the real initial opinions x , these data can be used to find values of the model parameters (edge weights W and inner traits ψ) that match as closely as possible the real opinion evolution, through the minimisation problem

$$(\widehat{W}, \widehat{\psi}) = \underset{\substack{W \in \mathcal{W} \\ \psi \in \mathcal{A}}}{\operatorname{argmin}} J(y, \tilde{y}) \quad \text{such that} \quad \tilde{y} = \mathcal{F}_\Omega(x, W, \psi, K), \quad (7.9)$$

where the cost function $J(y, \tilde{y}) = \sum_{i=1}^N |y_i - \tilde{y}_i|$ quantifies the mismatch between opinion vectors y and $\tilde{y}$.

If the same population is asked to quantify the agreement with Q different statements, the signed digraph cannot change. However, the inner traits assignation can vary depending on the statement, since each individual may have different attitudes towards different topics. Therefore, if $\psi^{(l)}$ represents the inner traits assignation associated with statement l , values for the parameters W and $(\psi^{(l)})_{l=1}^Q$ that produce predicted opinions as similar as possible to the real ones can be found through the *free optimisation problem*

$$(\widehat{W}, (\widehat{\psi}^{(l)})_{l=1}^Q) = \underset{\substack{W \in \mathcal{W} \\ \psi^{(l)} \in \mathcal{A}}}{\operatorname{argmin}} \sum_{l=1}^Q J(y_l, \tilde{y}_l) \quad \tilde{y}_l = \mathcal{F}_\Omega(x_l, W, \psi^{(l)}, K) \quad (7.10)$$

where x_l and y_l are the known initial and final opinions related to statement l .

If instead all the inner traits assignments are constrained to be the same for every question, we consider the *constrained optimisation problem*

$$(\widehat{W}, \widehat{\psi}) = \underset{\substack{W \in \mathcal{W} \\ \psi \in \mathcal{A}}}{\operatorname{argmin}} \sum_{l=1}^Q J(y_l, \tilde{y}_l) \quad \tilde{y}_l = \mathcal{F}_\Omega(x_l, W, \psi, K) \quad (7.11)$$

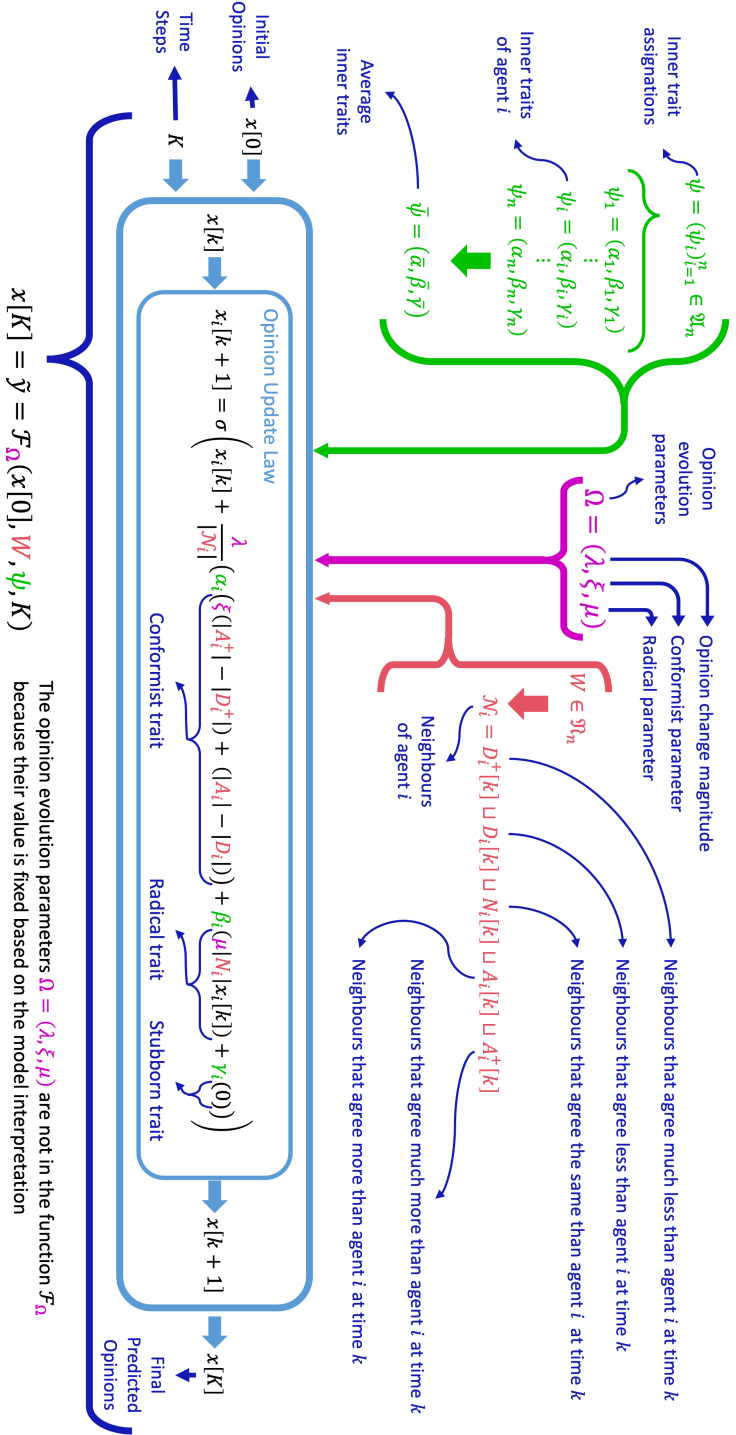


Figure 7.2: Visualisation of the model features and parameters. The model has three parameter types: inner traits assignment ψ (in green), opinion evolution parameters Ω (in magenta), and signed digraph weights W (in red). These parameters appear in the Opinion Update Law of Equation (7.6), which is a convex combination of contributions by conformist, radical, and stubborn traits, with the opinions of neighbouring agents evaluated through a classification-based approach. Given the initial conditions $x[0]$ and the number K of iterations over the prediction horizon, the Opinion Update Law produces the final predicted opinions $\hat{y} = x[K]$. The opinion evolution parameters Ω can be fixed based on the model interpretation. Then, the final predicted opinions in each particular case are a function $x[K] = \mathcal{F}_\Omega(x[0], W, \psi, K)$ of the chosen initial opinions, signed digraph weights, inner traits assignment, and number of iterations in the prediction horizon.

The free optimisation problem, where the inner assignments can change, allows for a more thorough study of the behaviour of a population, while the constrained optimisation problem allows for a more rigorous testing of the prediction capabilities of the model in the form of cross-validation: the answers to some questions can be used as training datasets to choose the model parameters, and the model performance can then be tested on the remaining questions.

7.2. SIMULATION RESULTS

To gain insight into the classification-based (CB) model, this section presents five different types of simulation results: 1) **Simulations in Simple Cases** (Section 7.2.1) evolve the model in simple, special cases to gain intuition into its behaviour; 2) **Parameter Sensitivity Analysis** (Section 7.2.2) studies how changes in each of the model parameters (inner traits assignment, signed digraph, opinion evolution parameters) affect the model behaviour; 3) **Model Validation with Real Data** (Section 7.2.3) leverages real data from the WVS to show that the CB model has the potential to reproduce the time evolution of real opinions in society (with parameters chosen through the *free* and the *constrained* optimisation problems of Equations (7.10) and (7.11) respectively) and presents the Transition Tables between different qualitative types of opinion distributions, as described in Section 6.1; 4) **Comparison with the Friedkin-Johnsen (FJ) Model** (Section 7.2.4) investigates the relation between the two models and their predictive capabilities (first, evolving equivalent populations; second, solving the optimisation problems (7.10) and (7.11); and third, computing the corresponding Transition Tables); 5) **Agreement Plot and Probabilistic Analysis** (Section 7.2.5) studies the model from the perspective presented in Sections 6.3 and 6.4. It is worth noting that the Agreement Plot is also used in other parts of the simulation results, but the six different plots described in Table 6.8 are found in Section 7.2.5.

To facilitate the interpretation of simulation results, we introduce some definitions. Given the inner traits assignment $\psi = (\psi_i)_{i \in \mathcal{V}} = ((\alpha_i, \beta_i, \gamma_i))_{i \in \mathcal{V}}$, the associated *average inner traits*

$$\bar{\psi} = (\bar{\alpha}, \bar{\beta}, \bar{\gamma}) \quad \text{where} \quad \bar{\alpha} = \frac{1}{N} \sum_{i \in \mathcal{V}} \alpha_i \quad \bar{\beta} = \frac{1}{N} \sum_{i \in \mathcal{V}} \beta_i \quad \bar{\gamma} = \frac{1}{N} \sum_{i \in \mathcal{V}} \gamma_i, \quad (7.12)$$

represent the traits of an average agent in the considered society or population. The inner traits assignment ψ and the corresponding average inner traits $\bar{\psi}$ can be plotted in a ternary diagram as shown in Figure 7.3a. Figure 7.3b explains how to interpret a point in the ternary diagram.

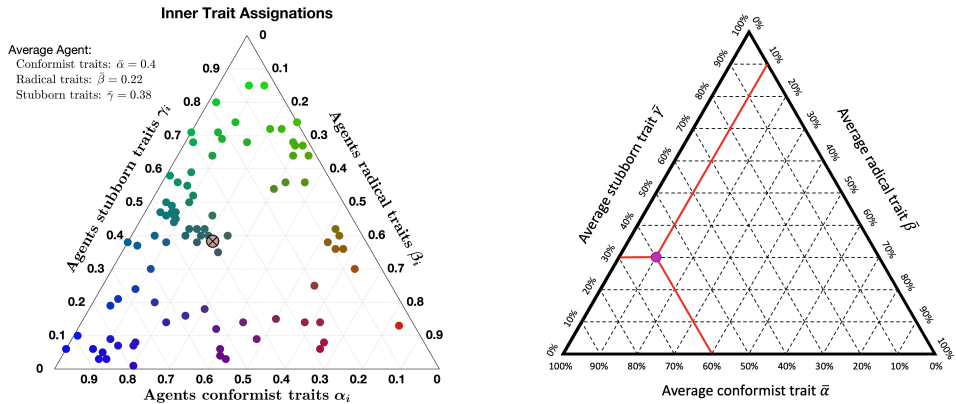
The *signed agreement* of an opinion vector $x = (x_i)_{i=1}^N$, quantified by the pair (θ_+, θ_-) where

$$\theta_- = \sum_{x_i < 0} x_i \quad \text{and} \quad \theta_+ = \sum_{x_i > 0} x_i, \quad (7.13)$$

is the overall level of agreement and disagreement in the whole society. The *signed agreement* $\theta = (\theta_+, \theta_-)$ and the *general agreement* $\pi = (\bar{x}, |\bar{x}|)$ are related as follows:

$$\bar{x} = \frac{1}{N}(\theta_+ + \theta_-) \quad |\bar{x}| = \frac{1}{N}(\theta_+ - \theta_-) \quad (7.14)$$

All the digraphs used in both Parameter Sensitivity Analysis and Model Validation with Real Data have a Small-World network topology, with an assigned probability for positive and negative edges, and are strongly connected. We consider Small-World networks because they have a high clustering coefficient (neighbours of neighbours of agent i are likely also neighbours of agent i) and low diameter (maximum distance between two agents of the network), which are believed to be characteristics of real-life social networks [68, 214]. The directed Small-World networks were built based on the Watts-Strogatz algorithm. Appendix A (Section 7.4) describes the computation of network metrics. The signed digraphs are not restricted to be structurally balanced, to account



(a) Each dot represents the inner traits of an agent; its RGB colour reflects the weight of each trait (blue: conformist; red: radical; green: stubborn). The crossed dot represents the average inner traits.

(b) Example of average inner traits in the ternary diagram: 60% conformist, 10% radical, 30% stubborn.

Figure 7.3: Ternary diagrams visualising inner traits assignments ψ and average inner traits $\bar{\psi}$. Panel 7.3a shows the whole inner traits assignment (along with its average), while panel 7.3b only shows the average inner traits. The difference between these diagrams is key: each non-crossed dot in Figure 7.3a represents the traits of a single agent (see the diagram labels), the crossed dot represents the average traits. The dot in Figure 7.3b represents the average traits of a complete population of (possibly) many agents (see the diagram labels). The ternary diagram in Figure 7.3b could also have multiple dots, and each would correspond to a different society.

for the fact that also non-structurally-balanced networks have been considered in the literature when modelling social dynamics [70, 140, 173].

In all the considered simulations, the initial opinions, traits and networks are assigned independently. A different approach could be to assign them in some correlated way: e.g., initial opinions and network could be correlated by assigning the initial opinions such that two vertices connected by an edge have a very similar (or very distant) initial opinion; traits and network could be correlated by assigning the agent parameters with a probability that depends on the corresponding vertex characteristics, for example assuming that vertices with higher out-degree have a higher probability of being completely conformist, or radical. Correlations between initial opinions, traits, and network characteristics can reproduce different types of societies present in real life (for instance, in a society that values tradition, highly stubborn agents may be more influential than others, and hence the corresponding vertices may have a higher out-degree).

7.2.1. SIMULATIONS IN SIMPLE CASES

To better understand the model behaviour, we simulate the model evolution over a digraph with a Lattice topology with varying inner traits assignments (Figure 7.4). We consider a signed Lattice digraph, where each agent has 4 in-neighbours and the edges are positive with probability 0.77. All the agents have the same inner traits, combining only two inner traits: stubbornness and radicalism; radicalism and conformism; conformism and stubbornness. Starting from the same initial opinions, Figure 7.4 shows the opinion evolution over 30 time steps. Radicalism tends to form polarisation by driving the agents to extreme opposite views. Conformism tends to create consensus; however, because of the classification approach, the agents do not converge to the very same opinion (close enough agents are unable to perceive their opinion difference, because opinions are assessed with finite resolution). Stubbornness slows down the effect of the other two traits; only in

a fully stubborn population everyone keeps its initial opinion. Among the three traits, radicalism appears to have the greatest effect: even a small amount of radicalism can prevent conformism from forming consensus, and can yield polarisation in a very stubborn society.

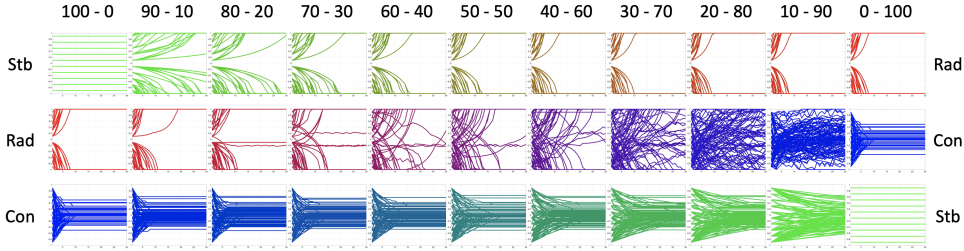


Figure 7.4: Evolution of the CB model, starting from the same initial opinions, over a signed Lattice digraph where the edges are positive with probability 0.77. In each simulation, all 100 agents have the same inner traits that are a combination of only two of the possible traits: stubbornness (Stb), radicalism (Rad), and conformism (Con). The top labels show the proportion of each trait: the upper left (respectively, right) graph corresponds to a simulation where all agents are purely stubborn (resp. radical). The colour of the lines is the RGB representation of the inner traits assignments (blue: conformist; red: radical; green: stubborn).

The opinion evolutions shown in Figure 7.4 can be interpreted as the model behaviour when the inner traits are at the boundary of the ternary diagram. In a way, these evolutions capture the dynamics that the model is capable of producing; since, as it can be seen in the next section, opinion evolutions when the inner traits are located inside the ternary diagram follow similar patterns. For instance, compare the plots in Figure 7.4 and Figures 7.6a, 7.7b, and 7.7e.

7.2.2. PARAMETER SENSITIVITY ANALYSIS

We select a set of *nominal parameters* (which, for given initial conditions, produce *nominal simulation results*) as a baseline with which other parameter choices can be compared. We choose a nominal inner traits assignment that leads to model outcomes that closely reproduce real data from the World Values Survey (in fact, it is close to some of the inner traits assignments resulting from the *Free* optimisation problem (7.9), see Figure 7.15a), and therefore has the potential to represent a realistic society; moreover, it allows us to showcase the wide range of different opinion evolutions that the model can produce. Then, we vary inner traits assignments, signed digraph and opinion evolution parameters, one by one, and study their effect on the simulated behaviour.

NOMINAL PARAMETERS AND NOMINAL RESULTS

We consider the initial opinions shown in Figure 7.5a, which evolve according to the model with the nominal parameters: $\lambda = 0.4$, $\xi = 2$, $\mu = 5$, inner traits assignments in Figure 7.5b, and signed digraph in Figure 7.5c.

The initial opinions shown in Figure 7.5a have $\theta_- = -19.3$ and $\theta_+ = 41.5$, indicating a strong general agreement since $\theta_+ > -\theta_-$. Figure 7.5b shows that most agents have very strong conformist traits, with a notable percentage of radicalism, resulting in an average agent (crossed dot) with 60% conformist traits, 30% radical traits, and 10% stubborn traits. The nominal signed digraph in Figure 7.5c is highly connected, with average path length 2.12, clustering coefficient 0.38, diameter 4. It has 834 positive edges and 767 negative edges.

The nominal results are shown in Figure 7.6. Figure 7.6a shows the opinion evolution of every agent. The line colour represents the percentage of conformist, radical, and stubborn agent traits (blue for conformist, red for radical, and green for stubborn). The purple colour of most lines

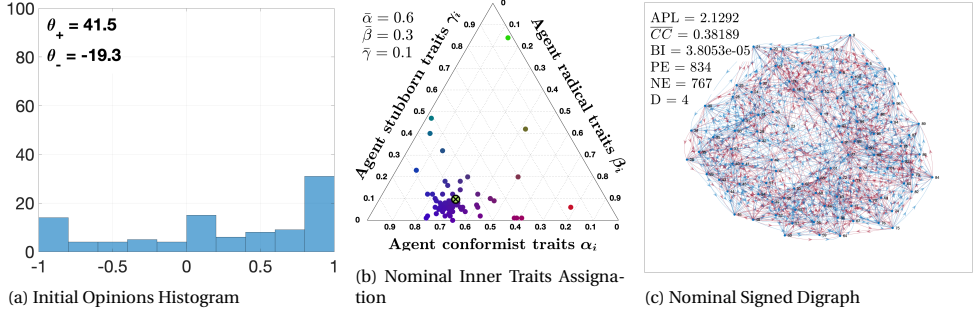


Figure 7.5: Initial opinions, nominal parameters, and nominal signed digraph.

corresponds to a combination of conformist and radical traits. The discontinuity in the opinion change is due to the classification process leading to a discontinuous opinion update law. The opinion evolution of the various agents shows a great variability in opinion changes, without a clear global tendency.

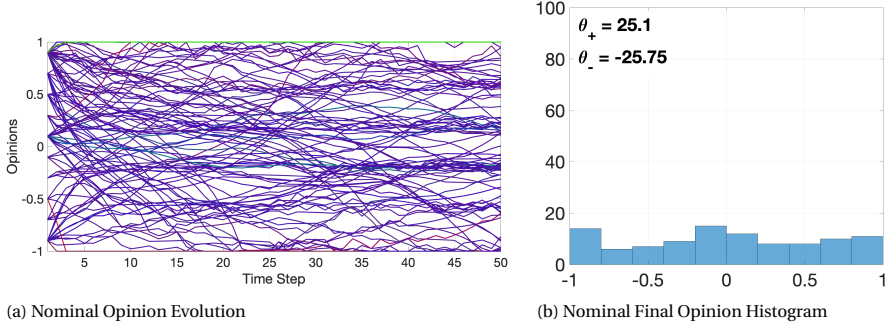


Figure 7.6: Opinion evolution with nominal parameter values ($\lambda = 0.4$, $\xi = 2$, $\mu = 5$) and final opinion histogram (100 agents).

Figure 7.6b shows the histogram of the nominal final opinions predicted by the model after 50 time steps. Compared with the initial opinions, the final opinions appear to have a more uniform distribution: in fact, for the nominal final opinions, $\theta_- = -25.75$ and $\theta_+ = 25.1$, hence $\theta_+ \approx -\theta_-$. The behaviour of the opinion evolution and the distribution of the final opinions is explained by the presence of two opposing forces that drive the opinion of all the agents: on one hand, the tendency to achieve consensus, due to the conformist traits, drives the agents towards the centre; on the other hand, the radical traits move the opinions towards extreme values.

VARYING THE INNER TRAITS ASSIGNATIONS

To evaluate the effect of different inner traits assignments, we change the nominal inner traits assignments of Figure 7.5b and simulate the opinion evolution, keeping all the other parameters unchanged. The two new inner traits assignments, shown in Figures 7.7a and 7.7d, are simply rotations of the nominal inner traits assignments. The corresponding opinion evolutions are shown in Figures 7.7b and 7.7e, while the final opinion histograms are presented in Figures 7.7c and 7.7f.

Comparing the opinion evolutions of Figures 7.6a, 7.7b, and 7.7e and the final opinion histograms of Figures 7.6b, 7.7c, and 7.7f reveals the profound effect of different inner traits assigna-

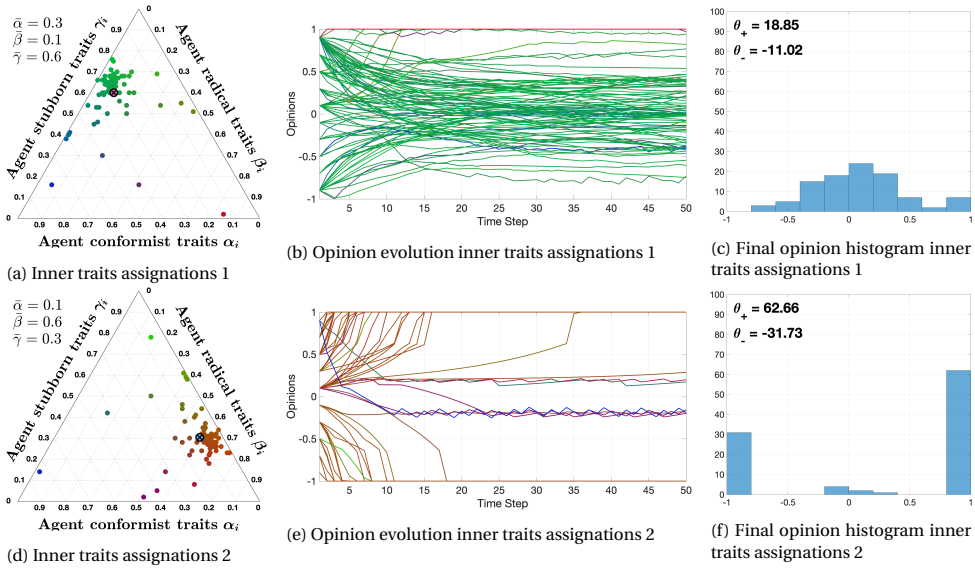


Figure 7.7: Effect of changing the inner traits assignments (evolving 100 agents).

tions on the opinion evolution. In the inner traits assignment of Figure 7.7a, the agents are mostly stubborn and conformist. This results in a very slow convergence towards the mean, spurred by conformist traits and slowed down by stubborn traits. Because of the neighbour classification, even completely conformist agents would not reach perfect consensus, but would rather converge to an opinion subinterval where all the agents perceive that the others have a comparable opinion. This tendency towards the mean can be seen in the final opinion histogram of Figure 7.7c, where both $\theta_- = -11.02$ and $\theta_+ = 18.85$ are much closer to 0.

On the other hand, the inner traits assignment of Figure 7.7d gives agents pronounced radical traits. Both the opinion evolution in Figure 7.7e and the final opinion histogram in Figure 7.7f show that agents lean towards extreme opinions. A bunch of agents keeps its opinion closer to zero. The line colours (closer to blue and green) show that these agents do not have very strong radical traits, and instead they are more conformist and stubborn: such traits allow these agents to avoid extreme opinions.

VARYING THE SIGNED DIGRAPH

To study the effect of changing the signs of the weights of the signed digraph, the nominal signed digraph of Figure 7.5c is modified into the signed digraphs shown in Figures 7.8a and 7.8d. The topology is unchanged, but the number of positive and negative edges is changed. The resulting opinion evolution and final opinion histograms are shown in Figures 7.8b and 7.8c, and in Figures 7.8e and 7.8f respectively.

Compared with the nominal results in Figures 7.6a and 7.6b, the most different outcome occurs when most edges are positive (digraph in Figure 7.8d). In this case, the end result is almost perfect consensus for the +1 opinion, because the initial opinion, with $\theta_- = -19.3$ and $\theta_+ = 41.5$, is more skewed towards +1. The presence of negative edges is crucial to avoid trivial consensus outcomes even when the agents are not completely conformist. The opinion evolution in Figure 7.8e shows that, initially, conformist traits pull the opinions towards positive values, and then radical

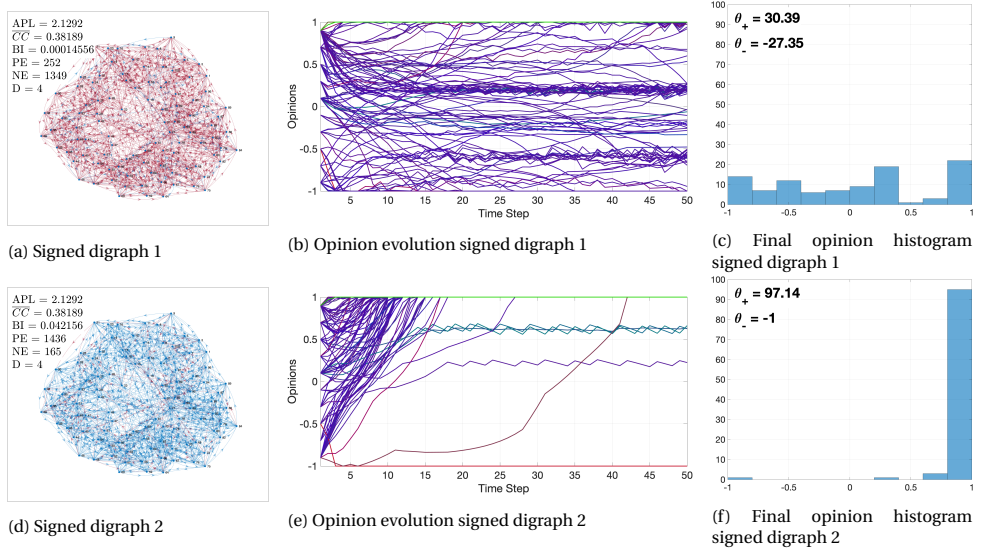


Figure 7.8: Effect of changing the signed digraph (evolving 100 agents).

traits make them increase in value until they reach +1. Purely radical agents would have produced polarisation instead of consensus.

When increasing the number of negative edges (digraph in Figure 7.8a), the final opinions in Figure 7.8c are different from the nominal ones, but the qualitative behaviour is comparable.

7

VARYING THE OPINION EVOLUTION PARAMETERS

We study the sensitivity with respect to the opinion evolution parameters $\Omega = (\lambda, \xi, \mu)$, where: λ is the overall opinion change magnitude, and can also be thought of as a time scaling parameter; ξ gives more weight to distant opinions for conformist traits; μ increases the opinion change for radical traits. We change these parameters one at the time, with respect to the nominal parameters, and compare the results with the nominal results in Figure 7.6.

Figure 7.9 shows the opinion evolution and final histogram for $\lambda = 0.2$ and $\lambda = 0.8$. The final histograms in Figures 7.9b and 7.9d do not change much with respect to the nominal. The most significant change can be noticed in Figures 7.9a and 7.9c, showing that indeed a higher value of λ produces larger changes in the opinions. Overall, however, the effect of varying λ is very limited.

The effect of varying ξ is shown in Figure 7.10. The changes in both the opinion evolution and the final opinion histogram are quite noticeable. A value of $\xi = 1$ means that distant opinions have the same attracting power as closer opinions for the conformist traits, hence in general the conformist trait has less influence over the whole opinion change, which is instead dominated by the radical traits. The result is visible in the opinion evolution in Figure 7.10a and the final opinion histogram in Figure 7.10b. On the contrary, increasing the value to $\xi = 4$ yields a stronger conformist tendency towards consensus, evident when comparing the nominal final opinions in Figure 7.6b with the final opinions with $\xi = 4$ in Figure 7.10d, and the respective θ_- and θ_+ .

Parameter μ modulates the effect of radical traits on the opinion evolution. Comparing Figure 7.11b with Figure 7.11d shows that a larger μ increases radicalism in the population, which leads to polarisation for the given initial opinions. A similar effect is achieved by varying ξ : in fact, both ξ and μ affect the balance between the conformist tendency towards consensus and the radical

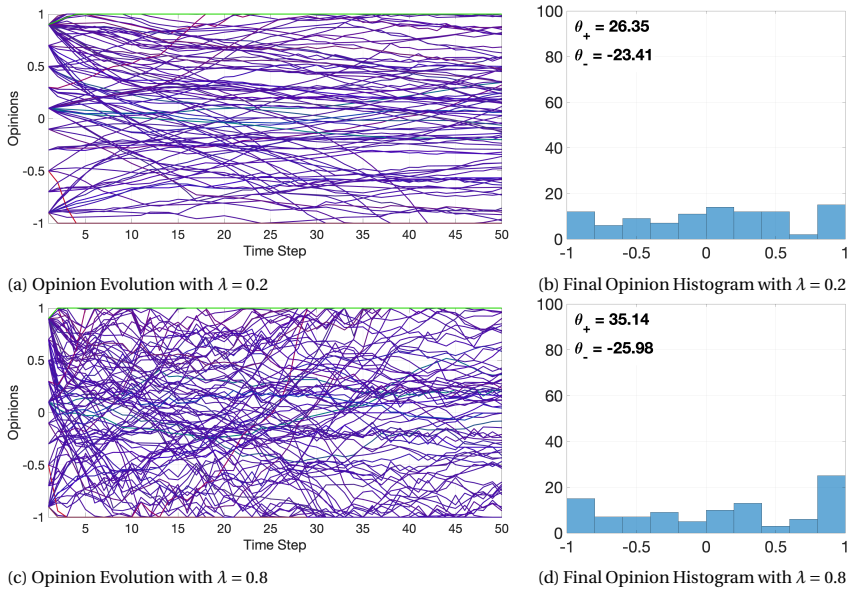


Figure 7.9: Effect of changing λ from the nominal value $\lambda = 0.4$ to $\lambda = 0.2$ (Figures 7.9a and 7.9b) and $\lambda = 0.8$ (Figures 7.9c and 7.9d) evolving 100 agents. The other model values are $\xi = 2$, and $\mu = 5$

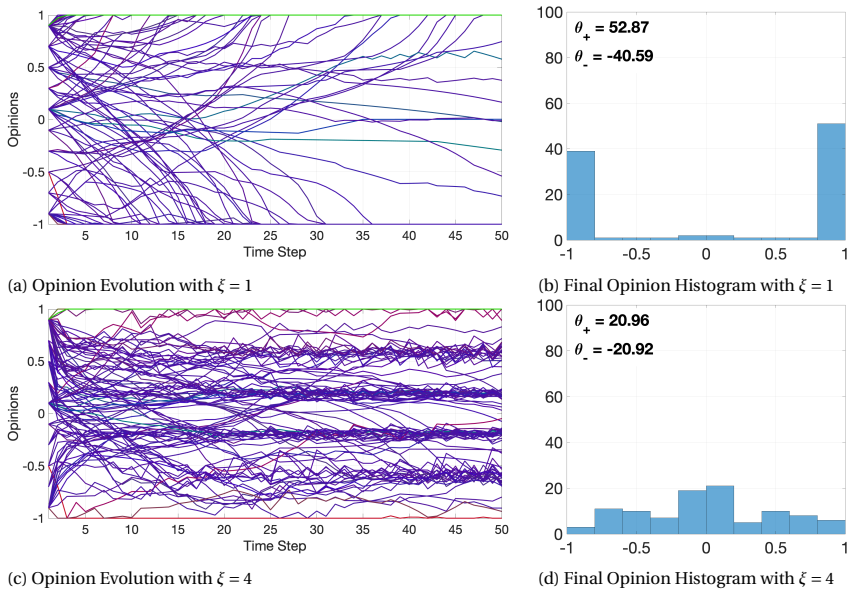


Figure 7.10: Effect of changing ξ from the nominal value $\xi = 2$ to $\xi = 1$ (Figures 7.10a and 7.10b) and $\xi = 4$ (Figures 7.10c and 7.10d) evolving 100 agents. The other model values are $\lambda = 0.4$, and $\mu = 5$

tendency towards polarisation. Although both ξ and μ play a role in the conformist-radical bal-

ance, they are not completely complementary: an increase in ξ is not the same as a decrease in μ . This can be seen by comparing Figures 7.10d and 7.11b: increasing ξ produces final opinions that are more evenly distributed than those obtained by decreasing μ . Moreover, increasing radicalism does not always lead to polarisation: this happens only when the opinions have both positive and negative values. If the opinions have only positive values or only negative values, then radicalism will move all of them to a single extreme, resulting in consensus. Therefore, it is not possible to generalise the idea that more radicalism always leads to polarisation, regardless of the initial opinions.

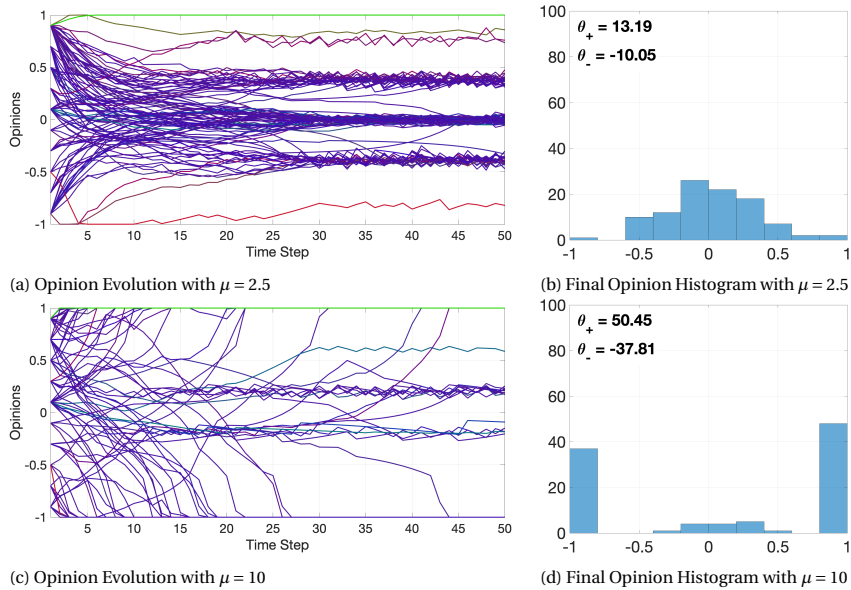


Figure 7.11: Effect of changing μ from the nominal value $\mu = 5$ to $\mu = 2.5$ (Figures 7.11a and 7.11b) and $\mu = 10$ (Figures 7.11c and 7.11d) evolving 100 agents. The other model values are $\lambda = 0.4$, and $\xi = 2$

7.2.3. MODEL VALIDATION WITH REAL DATA

Data from the World Values Survey are used to validate the CB model, namely, show that a suitable choice of the parameters allows the model to produce predicted opinions similar to the real opinions in a society. The World Values Survey is an international organisation that conducts surveys about ethics and values in different countries around the globe. These surveys are repeated every 5 years. We considered the answers to 30 questions, shown in Table 7.19, in 26 countries, shown in Table 7.18. In each question, the respondents are asked to state the extent to which they agree with a statement in a Likert scale 10. The answers given in the wave 5 surveys are taken as initial opinions, while the answers of wave 6 surveys are taken as final opinions.

Two minimisation problems are stated to find model parameters that produce predicted opinions similar to the ones found in the survey answers. The **Free Optimisation Problem** allows the inner traits assignment to change with questions; in the **Constrained Optimisation Problem**, the inner traits are fixed for all questions. The Transition Tables for the model with parameters provided by both optimisation problems are also computed.

Given real and model-generated opinion vectors r and y , the cost function J used in the minimisation problems (7.9), (7.10), and (7.11) is defined as

$$J(r, y) = \sum_{i=1}^N |\hat{r}_i - \tilde{y}_i|, \quad (7.15)$$

where $\tilde{r} = (\tilde{r}_i)_{i=1}^N$ is the vector $\hat{r} = (\hat{r}_i)_{i=1}^N$ sorted in descending order, and $\cdot^\wedge$ is the quantisation function

$$\hat{r}_i = \underset{\zeta \in \mathcal{R}}{\operatorname{argmin}} \{|\zeta - r_i|\} \quad \forall i = 1, \dots, N,$$

with $\mathcal{R}$ defined as $\mathcal{R} = \left\{ \frac{1}{2}(u_k + u_{k+1}) \mid u_k = -1 + k \frac{2}{10} \quad k = 1, \dots, 9 \right\}$. Quantisation is needed because the World Values Survey answers we consider as real opinions use a Likert scale 10: participants could choose their opinion from 10 different options. These opinions rescaled to be between -1 and 1 produce the set $\mathcal{R}$ and, therefore, the predicted opinions also need to be quantified in the same way. Both opinion vectors (real and predicted) are sorted in descending order, so that equal opinions add a zero to the total cost.

Even for a relatively small population $N = 100$, the size of the sets $\mathcal{W}$ (underlying signed digraph structures) and $\mathcal{A}$ (inner traits assignments) is enormous. Given the tremendous size of the parameter space $\mathcal{W} \times \mathcal{A}$, performing the minimisation over all possible signed digraphs and agent inner traits would be computationally intractable. Therefore, the minimisation occurs over small subsets $\tilde{\mathcal{W}} \subset \mathcal{W}$, $\tilde{\mathcal{A}} \subset \mathcal{A}$ of the whole parameter space. As a consequence, there is no guarantee that we are estimating the *real* parameter values or making the absolute best parameter choice: with other parameter choices, not included in $\tilde{\mathcal{W}} \times \tilde{\mathcal{A}}$, the model could reproduce the data with even better accuracy.

ID	1	2	3	4	5	6	7	8	9	10	11	12
APL	2.13	2.13	2.13	2.13	2.13	1.95	1.95	1.95	1.95	1.95	2.04	2.04
CC	0.38	0.38	0.38	0.38	0.38	0.18	0.18	0.18	0.18	0.18	0.16	0.16
PE	252	558	834	1115	1436	258	566	848	1145	1438	222	533
NE	1349	1043	767	486	165	1326	1018	736	439	146	1194	883
D	4	4	4	4	4	3	3	3	3	3	3	3
BI	0.00015	4.4e-05	3.8e-05	0.00013	0.042	0.00023	8.1e-05	4.8e-05	0.00027	0.049	0.00099	0.00025
ID	13	14	15	16	17	18	19	20	21	22	23	24
APL	2.04	2.04	2.04	1.75	1.75	1.75	1.75	1.75	1.68	1.68	1.68	1.68
CC	0.16	0.16	0.16	0.25	0.25	0.25	0.25	0.25	0.35	0.35	0.35	0.35
PE	746	1020	1259	362	864	1351	1813	2344	418	1079	1683	2372
NE	670	396	157	2227	1725	1238	776	245	2891	2230	1626	937
D	3	3	3	3	3	3	3	3	2	2	2	2
BI	0.00021	0.00056	0.047	2e-08	6.1e-09	4.1e-09	1e-07	0.0071	3.4e-11	5.8e-12	7.5e-13	6.7e-09
ID	25	26	27	28	29	30	31	32	33	34	35	
APL	1.68	1.68	1.68	1.68	1.68	1.68	1.62	1.62	1.62	1.62	1.62	
CC	0.35	0.32	0.32	0.32	0.32	0.32	0.39	0.39	0.39	0.39	0.39	
PE	2947	456	1063	1667	2329	2972	457	1259	1998	2717	3506	
NE	362	2823	2216	1612	950	307	3440	2638	1899	1180	391	
D	2	2	2	2	2	2	2	2	2	2	2	
BI	0.00074	4.8e-11	8.6e-12	1.3e-12	3.7e-09	0.0021	4.7e-14	3.6e-14	3.2e-14	4.3e-11	0.00033	

Table 7.1: Signed Digraph Information: Average Path Length (APL), Clustering Coefficient (CC), Positive Edges (PE), Negative Edges (NE), Diameter (D), and Balance Index (BI)

The subset $\tilde{\mathcal{W}}$ contains 35 different Small-World signed strongly connected digraphs. Table 7.1 shows the main characteristics of the networks. The subset $\tilde{\mathcal{A}}$ contains 3528 randomly generated inner traits assignments $\psi = (\psi_i)_{i=1}^N$. To avoid bias towards societies with average inner traits that are more conformist, radical, or stubborn, the set $\tilde{\mathcal{A}}$ satisfies the following property: for every inner traits assignment ψ , with corresponding average inner trait $\tilde{\psi} = (\tilde{\alpha}, \tilde{\beta}, \tilde{\gamma}) = (a_1, b_1, c_1)$, there are two inner traits assignments $\psi', \psi'' \in \tilde{\mathcal{A}}$ that satisfy $\psi' = (b_1, c_1, a_1)$, and $\psi'' = (c_1, a_1, b_1)$. In

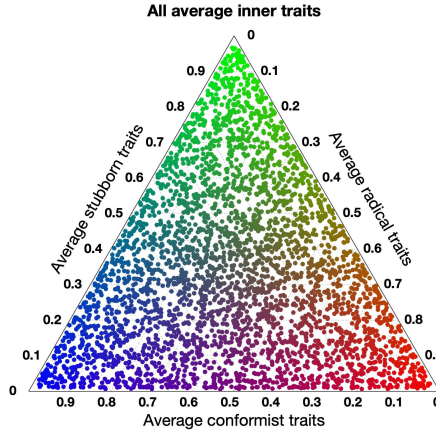


Figure 7.12: All the average inner traits $\bar{\psi}$ corresponding to inner traits assignments ψ in $\tilde{\mathcal{A}}$.

other words, the parameter space $\tilde{\mathcal{A}}$ is symmetric with respect to permutations of agent traits. Besides this property, the elements of this set were chosen at random. All the average inner traits $\bar{\psi}$ corresponding to inner traits assignments ψ in $\tilde{\mathcal{A}}$ are shown in Figure 7.12.

Due to the anonymity of the surveys, it is not possible to guarantee that the same people answered the survey in subsequent waves of the WVS. However, if the surveys are done correctly to represent society overall, the results can be anyway assumed to reflect the global opinion distribution of the general population about a given topic at a specific time, and this allows us to use the survey results in different waves in our minimisation problem, *as if* the very same people had answered.

Free OPTIMISATION PROBLEM

Assuming that the agents can have different inner traits for each question, Equation (7.10) was used to find model parameters for each country that yield opinions similar to the real ones. Once the parameters that solve the minimisation problem (7.10) were found for each country, the cost associated with the prediction discrepancy for each question-country pair was computed as in Equation (7.15) (see Figure 7.13). Due to its complexity and the huge size of the feasibility set, the minimisation problem is solved approximately: hence, a possibly suboptimal solution is found. By solving the optimisation problem more accurately, over a longer computation time (which we could not afford, due to the very large number of question-country pairs that we consider), even smaller costs could be achieved, and hence even better fits of the real data.

Figure 7.14 shows the model predictions for some question-country pairs. The original opinion is shown in blue, the real final opinion in orange, and the predicted final opinion in green; the corresponding cost (discrepancy) is reported. For costs less than 7, the model produces predicted final opinions that accurately represent the real final opinions, hence, we label these question-country pairs as ‘accurate’.

In total 780 question-country pairs were considered (30 questions and 26 countries). The individual cost of each question-country pair is reported in tables located in Appendix E. In order to evaluate the models, the most important metrics are: the average cost along all the countries, the number of ‘accurate’ question-country pairs, and their average cost. These values, for the *Free* optimisation problem are reported in Table 7.2.

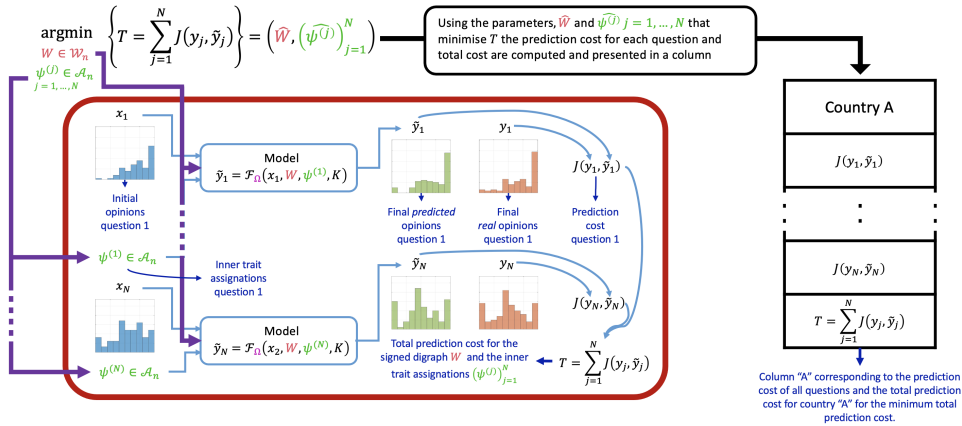


Figure 7.13: Visualisation of the procedure to solve minimisation problem in Equation (7.10). Assume that a survey conducted in two separate occasions in country A had Q questions. Given a signed digraph $W \in \mathcal{W}$ and Q inner traits assignments $(\psi^{(l)} \in \mathcal{A})_{l=1}^Q$ (one for each question), the model predicts a final opinion $\tilde{y}_l$. The cost function $J(y_l, \tilde{y}_l)$ measures how close the predicted final opinion is to the real final opinion y_l . The sum of all these costs gives the total cost T of the country. Minimising the value of T over the all the signed digraphs in $\mathcal{W}$ and inner traits assignments in $\mathcal{A}$ gives parameters that approximate the real opinion evolution, $\hat{W}$, and $(\hat{\psi}^{(l)})_{l=1}^Q$. The cost for each question and the average and total cost obtained using these optimal parameters are reported in the column of Table 7.24 corresponding to the considered country. Table 7.2 reports the global results, namely, the average cost along all countries, the number of 'accurate' question-country pairs (a question-country pair is labeled 'accurate' if the cost $J(y_l, \tilde{y}_l)$ is less than 7), and the average cost of 'accurate' question-country pairs. All simulations evolved 100 agents.

Average cost along all the countries	3.2815
Number of 'accurate' question-country pairs	755 from 780 (97% accuracy)
Average cost of 'accurate' question-country pairs	2.97

Table 7.2: Global results of the *Free* optimization problem (Equation (7.10)) using the Classification-based model. The individual costs for each question-country pair can be seen in Table 7.24. The corresponding results for the Friedkin-Johnsen model are shown in Table 7.9

To carry out a thorough comparison with standard models of opinion formation, an analysis equivalent to the one reported in Table 7.2 is performed also for the Null model (the model that assumes that the opinions do not change over time) and the French-DeGroot (FG) model [77, 105, 106, 54]. The results are reported in Tables 7.3 and 7.4, respectively.

Average cost along all the countries	10.5567
Number of 'accurate' question-country pairs	332 from 780 (43% accuracy)
Average cost of 'accurate' question-country pairs	4.17

Table 7.3: Global results using the Null model. The individual costs for each question-country pair can be seen in Table 7.25

To make Tables 7.2 and 7.4 comparable, for the FG model the digraphs used in each country are selected following the same minimisation problem as the one solved for the CB model. Since

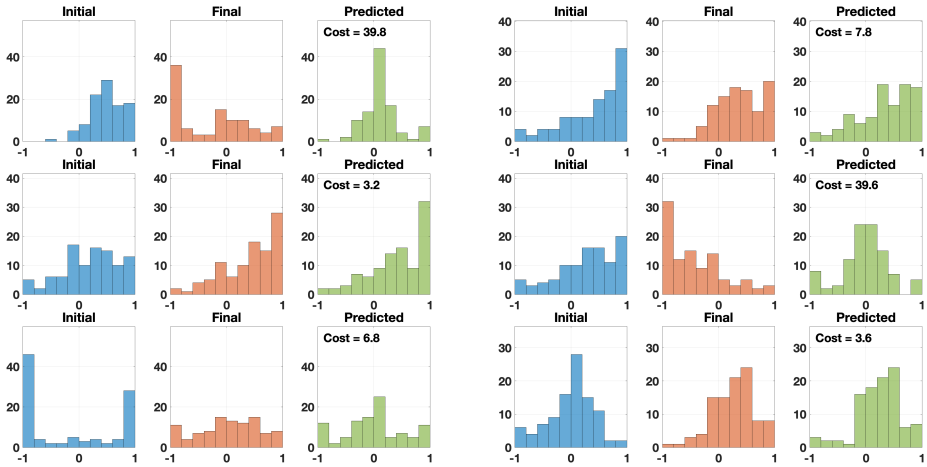


Figure 7.14: Examples of the opinion predictions achieved with the CB model, with varying resulting costs. For each trio of histograms, the initial opinions are in blue, the real final opinions in orange, and the predicted final opinions in green. The value of the cost J (in Equation (7.15)) is shown in the predicted histogram: cost values within 7 are shown to correspond to an accurate reproduction of the real opinion distribution.

Average cost along all the countries	38.4323
Number of 'accurate' question-country pairs	13 from 780 (2% accuracy)
Average cost of 'accurate' question-country pairs	5.43

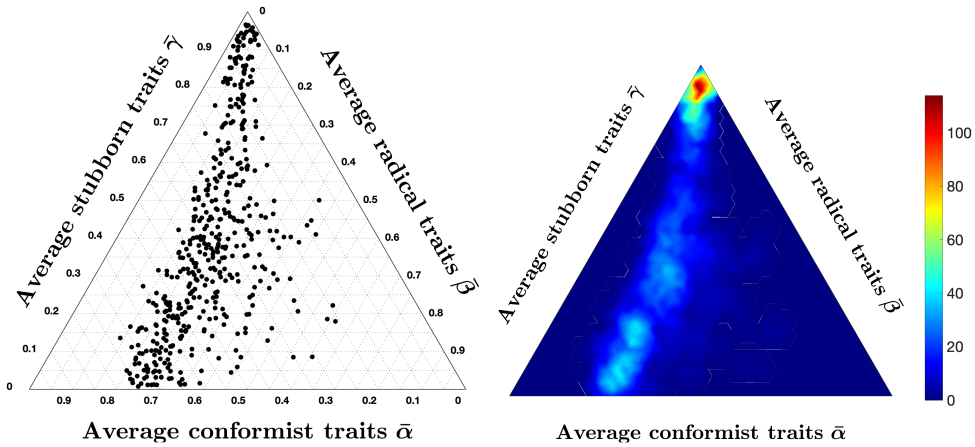
Table 7.4: Global results of the optimization problem using the French-DeGroot model. The individual costs for each question-country pair can be seen in Table 7.26

7

the FG model does not involve agent parameters, we only minimise over the set of digraphs $\mathcal{W}_{FG}$. Both the set of digraphs for the CB model, $\mathcal{W}$, and for the FG model, $\mathcal{W}_{FG}$, have the same number of elements and there is a one-to-one topology correspondence; the digraphs in $\mathcal{W}$ are signed and unweighted, while those in $\mathcal{W}_{FG}$ are unsigned and row-stochastic (as required by the different nature of the two models).

Comparing Table 7.2 with Tables 7.4 and 7.3 shows that the CB model performs remarkably well, yielding a 97% accuracy in contrast to the 43% accuracy of the Null model and the 2% accuracy of the French-DeGroot model. In fact, from Table 7.3 it is clear that, although there is a strong tendency towards stubbornness and opinion distribution tend to change only slightly over time, keeping the opinions exactly constant does not lead to good predictions. As shown in Table 7.4, the predictions of the French-DeGroot model are also not accurate, consistently with the evidence that perfect consensus is uncommon in real life.

Plotting the average inner traits $\bar{\psi}$ for all question-country pairs for which the cost is less than 7 provides possible hints on how these societies could potentially be formed. However, because of the large parameter space and relatively small data set, we cannot make conclusive statements on actual societies just based on the optimisation results, as very different inner traits assignments may produce similarly low costs: we just propose a *possible* explanation. The resulting ternary diagram is presented in Figure 7.15. Figure 7.15a shows the position of each question-country pair. Figure 7.15b shows a density plot over the ternary diagram indicating the regions where most question-country pairs are found.



(a) Ternary diagram plotting the average inner traits for the question-country pairs with cost < 7 according to Table 7.24. (b) Density plot of the ternary diagram: most question-country pairs have an average agent with more than 95% stubborn traits.

Figure 7.15: Analysis of the location of the average agents for all the question-country pairs with cost less than 7.

Despite the small data set and possible multiple local minima with similar low cost, fitting real data to gain an insight into the composition of actual societies reveals a clear trend: most average inner traits include a strong stubborn component, as shown by the high density in the stubbornness corner in Figure 7.15b. Also, the non-stubborn part can be roughly divided into 70% conformist and 30% radical, as shown by the trend in Figure 7.15a. This distribution is almost constant across all question-country pairs. Again, this is a possible explanation, and more data and more thorough explorations of the parameter space (extremely challenging from a computational standpoint) would be needed to make more conclusive statements. Hence, this is not conclusive evidence that most people are stubborn. There may be other explanations, for instance that not too many opinion exchange events take place in an average person's life. Graph-theoretically speaking, *isolation* due to the lack of outgoing edges from a node (i.e., lack of interactions) is associated with the concept of *stubbornness*. However, from a mathematical model it is impossible to draw conclusions on whether the opinion of an agent remains unchanged because the agent refuses to consider the different opinions it is exposed to, or because the agent intentionally avoids exposure to different opinions, or because the agent simply lacks the opportunity to come into contact with different opinions. Furthermore, the traits themselves can be interpreted in different ways: for instance, a lower value of stubbornness can be regarded as a greater openness to change.

Parameter Variation:

The results presented in Table 7.2 and Figures 7.14 and 7.15 are obtained by solving the minimisation problem (7.10) with nominal opinion evolution parameters $\lambda = 0.4$, $\xi = 2$, and $\mu = 5$. We now analyse the results of the minimisation problem when these parameters are changed. Tables 7.5 to 7.7 present how this variation affects the percentage of accurate question-country pairs (namely, those associated with a cost smaller than 7), the average cost of accurate question-country pairs, and the ternary diagram plot.

Tables 7.5 to 7.7 show that, even after varying the values of λ , ξ , and μ , the percentage of accurate question-country pairs remains around 96%, and the average cost of accurate question-country pairs is between 2.79 and 3.45 which is quite remarkable since it means that the high

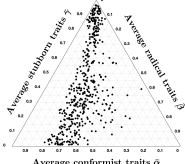
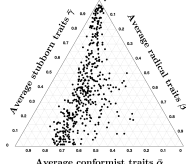
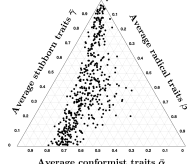
	$\lambda = 0.2$	$\lambda = 0.4$	$\lambda = 0.8$
% of 'accurate' country-question pairs	93.7	96.8	97.8
Average cost of 'accurate' country-question pairs	2.79	2.97	3.02
Ternary Diagram Plot			

Table 7.5: Effects of varying λ while keeping the nominal values $\xi = 2$, and $\mu = 5$.

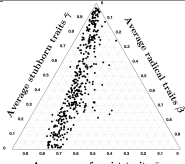
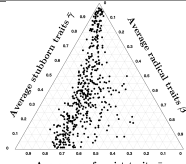
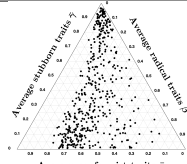
	$\xi = 1$	$\xi = 2$	$\xi = 4$
% of 'accurate' country-question pairs	96	96.8	95.8
Average cost of 'accurate' country-question pairs	2.84	2.97	3.45
Ternary Diagram Plot			

Table 7.6: Effects of varying ξ while keeping the nominal values $\mu = 5$, and $\lambda = 0.4$.

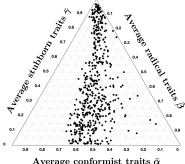
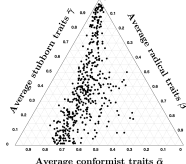
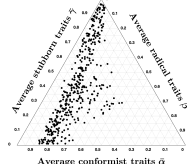
	$\mu = 2.5$	$\mu = 5$	$\mu = 10$
% of 'accurate' country-question pairs	96.8	96.8	97.2
Average cost of 'accurate' country-question pairs	2.84	2.97	3.15
Ternary Diagram Plot			

Table 7.7: Effects of varying μ while keeping the nominal values $\lambda = 0.4$, and $\xi = 2$.

accuracy achieved with the CB model is very robust to parameter variations.

Comparing the ternary diagrams shows the persistent tendency of question-country pairs to lie along a line where the proportion between conformist and radical traits is constant. For most simulation results, this proportion is still 70% conformist and 30% radical, as in the nominal case (Figure 7.15a). The proportion only changes when varying μ : for $\mu = 2.5$, we have 60% conformist and 40% radical agents, while for $\mu = 10$ we have 80% conformist and 20% radical agents. Therefore, it appears that μ can be tuned to regulate this proportion.

Constrained OPTIMISATION PROBLEM

If the agents are assumed to have the same inner traits for every question, then the model parameters can be found using the *constrained* optimisation problem in Equation (7.11). One advantage of using this approach is that, since each country has the same topology and inner traits assignation for all the questions, these parameters can be identified by solving the *constrained* optimisation problem (7.11) for a subset of all available questions (training dataset), and then tested on the remaining questions (test dataset). This was not possible previously, when assuming a different inner traits assignation associated with each question.

This procedure is commonly known as cross-validation. Generally, a subset of available data is used to train an algorithm (in this case, to identify the model parameters $\widehat{W}$ and $\widehat{\psi}$) and the remaining data is used to test the trained algorithm (in this case, the model with identified parameters $\widehat{W}$ and $\widehat{\psi}$). To eliminate result biases due to the selected training datasets and test datasets, cross-validation is performed multiple times for different partitions of the data. A common approach is to divide the data in K subsets and validate the model K times so that, at each iteration, only one subset is taken as the test dataset. This is known as K -fold cross-validation.

In the case of the *constrained* optimisation problem, a sixfold cross-validation was done on the available data (the questions are divided in six subsets of five questions each: $\{1, \dots, 5\}$, $\{6, \dots, 10\}$, $\dots$, $\{26, \dots, 30\}$). The results, reported in Table 7.8, show that the *free* optimisation problem yields better results, this is to be expected, as it has more degrees of freedom. Therefore, when looking at the results in Table 7.8 it is important to note that these predictions are done based on the assumption that the inner traits are the same for every question, while in reality the inner traits of the agents may change when considering their attitude towards different types of questions (which is taken into account by the free optimisation approach).

Average cost along all the countries	11.6746
Number of ‘accurate’ question-country pairs	220 from 780 (28% accuracy)
Average cost of ‘accurate’ question-country pairs	5.16

Table 7.8: Global results of the *Constrained* optimization problem (Equation (7.11)) using the Classification-based model. The individual costs for each question-country pair can be seen in Table 7.27. The corresponding results for the Friedkin-Johnsen model are shown in Table 7.10

TRANSITION TABLES

Here we briefly recall the definition of Transition Tables, for details see Section 6.1. Let x_o be an initial opinion vector, and x_f the final opinion vector predicted by the model. Both x_o and x_f can be sorted into one of five possible opinion distribution categories shown in Figure 7.16: 1) *perfect consensus*, PC; 2) *consensus*, Co; 3) *polarisation*, Po; 4) *clustering*, Cl; and 5) *dissensus*, Di.

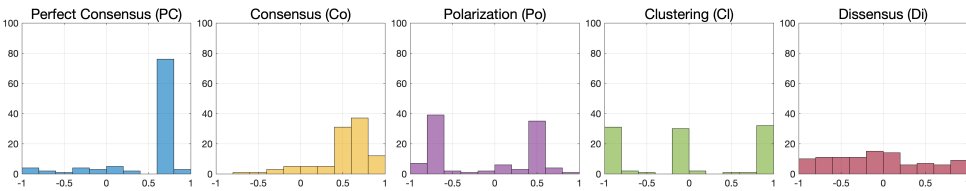


Figure 7.16: Histograms of opinion vectors that are representative of the five possible qualitative opinion distribution categories: Perfect Consensus, Consensus, Polarisation, Clustering, and Dissensus.

Now, let X_o be a set of initial opinions $X_o = \{x_o\}$ and X_f the corresponding set of predicted

opinions $X_f = \{x_f\}$. A *transition table* T , with 5 rows and 5 columns corresponding to the five possible opinion categories, is computed so that the coefficient in cell (a, b) is the number of initial opinion vectors $x_o \in X_o$ belonging to category a for which the corresponding predicted opinion x_f belongs to the category b . The table shows whether the model can evolve initial opinions belonging to any category into predicted opinions belonging to any other category.

Figure 7.17 shows three Transition Tables, where the set X_o represents the set of all World Values Survey answers to wave 5 for all questions and countries. For the real transition table A, the set X_f represents all the corresponding survey answers in wave 6, which are the true final opinions. For the transition table B (respectively, C) the set X_f contains all the corresponding predicted opinions produced by the CB model with the parameters obtained through the *Free* (respectively, *Constrained*) optimisation, namely, all the predicted final opinions $\hat{y}$ used to compute Table 7.2 (respectively, Table 7.8).

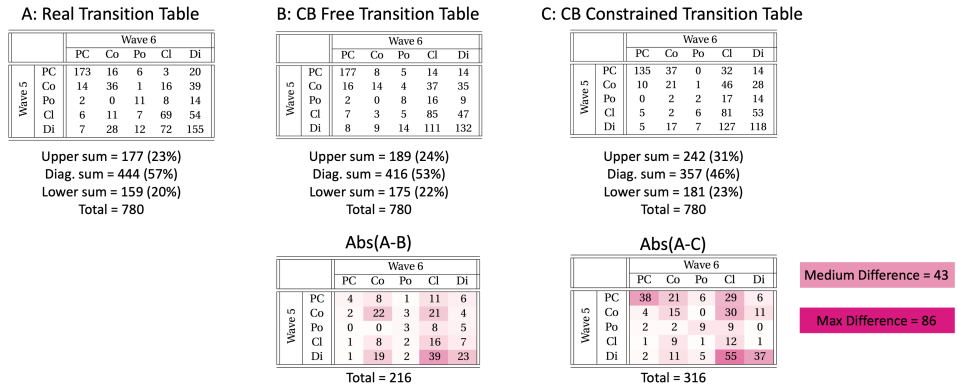


Figure 7.17: Real transition table and Transition Tables produced by the CB model using the digraphs and inner traits assignments given by the *Free* and *Constrained* optimisation problems, along with the sum and percentage of cells above / on / below the diagonal. The two tables below show the absolute value of the difference between predicted and real Transition Tables.

Interestingly, with the suitable choice of parameters, the CB model is capable of evolving opinions from any initial category into any other final category, as it happens with real opinion distributions. As expected, better results are achieved with the *Free* optimisation, in line with the results from Tables 7.2 and 7.8; still, both the free and the constrained Transition Tables show the versatility of the CB model, which can yield all transitions between opinion distribution categories that are seen in real life.

7.2.4. COMPARISON WITH THE FRIEDKIN-JOHNSEN MODEL

The classification-based model can be seen as an extension of the Friedkin-Johnsen (FJ) model [78, 79], in the sense that both models include in the agents' behaviour inner traits described by tuples. In the FJ model, each agent i is characterised by two parameters: *susceptibility* $a_i \in [0, 1]$, determining how strongly the agent is affected by its neighbours' opinions and forgets the initial opinion [79], and *prejudice* $b_i = 1 - a_i \in [0, 1]$. A value of $a_i = 1$ means that the agent has complete susceptibility to interpersonal influence (similar to complete conformism), while a value of $a_i = 0$ means that the opinion remains the same for all times (similar to complete stubbornness). When all agents are completely susceptible, the FJ model becomes the classic French-DeGroot (FG) model [77, 105, 106, 54]; when all agents are completely prejudgemental, the FJ model becomes the Null model (opinions do not change over time).

In the CB model, each agent i is associated with three parameters: *conformism* $\alpha_i \in [0, 1]$, *radicalism* $\beta_i \in [0, 1]$, and *stubbornness* $\gamma_i \in [0, 1]$, such that $\alpha_i + \beta_i + \gamma_i = 1$ for all $i \in \mathcal{V}$. Therefore, a FJ model where all agents have a susceptibility of a is similar to a CB model where all agents have inner traits $\alpha = a$ and $\gamma = 1 - a$ (hence $\beta = 0$). Still, the interpretations of stubbornness and prejudice are slightly different: prejudice in the FJ model means that agents tend to remain with their *initial opinion*, while stubbornness in the CB model means that agents tend to remain with their *current opinion*, which leads to the same outcome only when all the agents are completely stubborn. Apart from the outlined similarity, the FJ and CB models are different: crucially, the FJ model is linear, while the CB model is highly non-linear, which severely limits the applicability of closed-form analysis tools.

Figure 7.18 shows the evolution of the same initial opinions according to the two models, for different values of a , α , and γ . The digraphs had the same topology; randomly generated weights are considered for the FJ model, while for the CB model all the edge signs are taken positive to match the absence of antagonism in the FJ model.

The FJ model exhibits a slower change as a increases, while with the CB model, as soon as conformism is introduced, the opinions converge to an interval where all the agents perceive that their neighbours' opinion is similar enough to theirs. This difference is caused by the two different interpretations and implementations of stubbornness and prejudice. Another important difference is that, as the susceptibility value increases, the final opinions of the FJ model tend to converge to a single opinion, and yield perfect consensus when $a = 1$. Conversely, the final opinions of the CB model never converge to perfect consensus, even when $\alpha = 1$ and $\gamma = 0$, as a consequence of the classification-based approach.

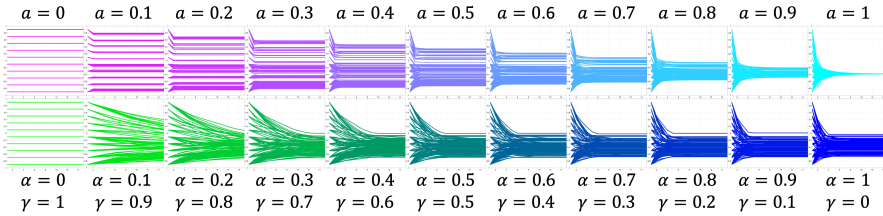


Figure 7.18: Comparison between the Friedkin-Johnsen (FJ) model, for different values of susceptibility a , and the Classification-Based (CB) model, for corresponding values of conformist (α) and stubborn (γ) weights. All the 100 agents have the same values of a (FJ) and of α and γ (CB). The simulations start from the same initial opinions and evolve over digraphs with the same topology. The degree of susceptibility, prejudice, conformism, and stubbornness is represented by the colours cyan, magenta, blue, and green respectively.

The differences between CB and FJ model help visualise the strong implications of the classification-based mechanism for assessing the opinion of others, which captures the fact that opinions cannot be perceived with perfect resolution and accuracy, and hence changes the model behaviour significantly: it grants the model new properties, such as the existence of multiple equilibria that can span the complete spectrum of opinions. For instance, in Figure 7.18, the CB model with $\alpha = 1$ and $\gamma = 0$ generates equilibrium opinions that span almost 40% of the opinion interval $[-1, 1]$ (a wider span can be achieved with different topologies), while the FJ model with $a = 1$ leads to identical equilibrium opinions.

The non-linearity introduced by the classification-based assessment of the opinion of others can completely change the resulting dynamics and lead to the emergence of peculiar features, which would not emerge from models where the agents have perfect access to the opinion of others. This is highlighted, for instance, by the comparison with the Friedkin-Johnsen model and with the French-DeGroot model (corresponding to the FJ model with $a = 1$).

The results of an analysis for the FJ model, equivalent to the one reported in Tables 7.2 and 7.8 for the CB model, are reported in Tables 7.9 and 7.10. To make the results comparable, when solving the optimisation problems (7.10) and (7.11) for the FJ model the sets $\mathcal{A}$ and $\mathcal{W}$ are modified as follows: the set of digraphs is the same used for the French-DeGroot model, since both models require row-stochastic adjacency matrices; the inner traits assignments in $\tilde{\mathcal{A}}$ are transformed into parameters of the FJ model using the mapping $a_i = \alpha_i / (\alpha_i + \gamma_i)$; if $\alpha_i + \gamma_i = 0$, then $a_i = 0.5$.

Average cost along all the countries	7.4897
Number of 'accurate' question-country pairs	460 from 780 (59% accuracy)
Average cost of 'accurate' question-country pairs	3.3

Table 7.9: Global results of the *Free* optimization problem (Equation (7.10)) using the Friedkin-Johnsen model. The individual costs for each question-country pair can be seen in Table 7.28. The corresponding results for the Classification-based model are shown in Table 7.2

Average cost along all the countries	10.3918
Number of 'accurate' question-country pairs	330 from 780 (42% accuracy)
Average cost of 'accurate' question-country pairs	4.1

Table 7.10: Global results of the *Constrained* optimization problem (Equation (7.11)) using the Friedkin-Johnsen model. The individual costs for each question-country pair can be seen in Table 7.29. The corresponding results for the Classification-based model are shown in Table 7.8

Comparing Tables 7.2 and 7.9 shows that the CB model outperforms the FJ model, yielding a 97% accuracy in contrast to 59%. Also, the average cost of 'accurate' country-question pairs is lower for the CB model (2.97) compared with the one produced by the FJ model (3.3), indicating that not only more question-country pairs are predicted satisfactorily, but also the predictions are more accurate.

Since the French-DeGroot (FG) and the Null model can be seen as extreme cases of the FJ model, as expected, the FJ model produces better results than the FG model and the Null model, as can be seen by comparing Tables 7.9 and 7.10 with Tables 7.4 and 7.3 (note that the optimisation was done for the total cost, not the number of 'accurate' question-country pairs). The fact that the Null model yields better predictions than the FG model suggests that opinions do not change much from one wave to the other, an observation that is also confirmed by the Transition Tables in Figure 7.17, where the sum (and percentage) of the values on the diagonal cells (associated with cases where initial and final opinion distributions both belong to the same qualitative category) is always the largest.

Figure 7.19 shows the Transition Tables for the FJ model (computed as for the CB model in Figure 7.17). In the CB Transition Tables, the diagonal sums (416 and 357) are always smaller than the real diagonal sum (444), while in the FJ Transition Tables the diagonal sums (544 and 630) are always larger, probably due to prejudice having a stronger effect than stubbornness in preserving the initial opinion. The lack of a "radical" trait makes the FJ model less versatile, which is reflected in the lower off-diagonal sum. According to Tables 7.2, 7.8, 7.9 and 7.10 the models that mimic the real opinion evolution at best are (in order): *Free* CB, *Free* FJ, *Constrained* FJ, and *Constrained* CB. Instead, according to the qualitative evaluation emerging from the Transition Tables, the models that are more faithful to the real opinion evolution are: *Free* CB, *Free* FJ, *Constrained* CB, and *Constrained* FJ.

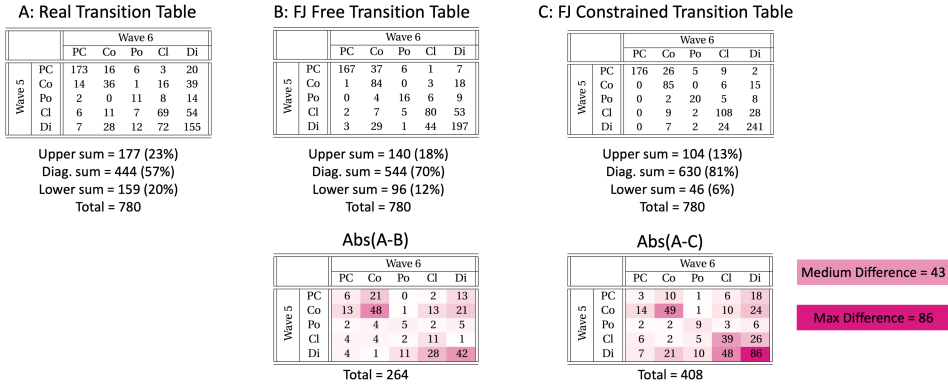


Figure 7.19: Real transition table and Transition Tables produced by the FJ model using the digraphs and inner traits assignments from the *Free* and *Constrained* optimisation problems. The sum and percentage of cells above, on, and below the diagonal is shown. The two bottom tables correspond to the absolute value of the difference between the predicted Transition Tables and the real one.

7.2.5. AGREEMENT PLOT AND PROBABILISTIC ANALYSIS

This final part of the simulation results for the Classification-based model investigates the model using the Agreement Plot and Probabilistic Analysis described in Sections 6.3 and 6.4, respectively.

AGREEMENT PLOT ANALYSIS

Starting with the Agreement Plot technique, Tables 7.11 to 7.16 present multiple plots in a way analogous to Tables 6.2 to 6.7 in Section 6.3. For Tables 7.11, 7.13, 7.14, and 7.16, the underlying digraphs have the same topology as the digraphs used in the Agreement Plot analysis of the Friedkin-Johnsen model. Therefore, the topology metrics are shown in Table 6.1.

Figure 7.20 shows the agent parameters used for the **UDTE**, **IOTE**, **UDSS**, and **IOSS** plots for the Classification-based model.

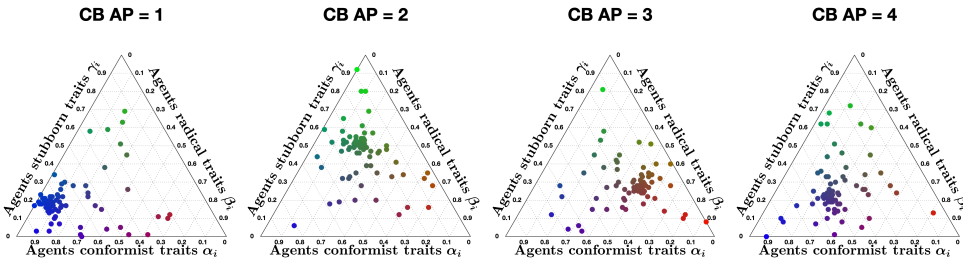


Figure 7.20: Ternary plots of the agent parameters used in the **UDTE**, **IOTE**, **UDSS**, and **IOSS** plots for the Classification-based model (Tables 7.12, 7.13, 7.15, and 7.16 respectively).

Unlike in Sections 6.3 and 6.5, the line and point colours in the plots on Tables 7.11 to 7.16 do not indicate a single mean. This is because, unlike the previous models, for the Classification-based model each agent has three parameters: the conformist, radical, and stubborn trait weights α_i , β_i , and γ_i , respectively. For each agent, these three weights take a value from the interval $[0, 1]$ and their sum is always 1. The average agent of a society will have conformist, radical, and

stubborn trait weights that are the corresponding averages among all the population agents, these average weights are denoted $\bar{\alpha}$, $\bar{\beta}$, and $\bar{\gamma}$ respectively; see Equation (7.12).

We will associate with each trait a colour, for conformism ($\bar{\alpha}$) blue, for radicalism ($\bar{\beta}$) red, and for stubbornness ($\bar{\gamma}$) green. Thus the line or point colour associated with a given society is the colour resulting from mixing the amount of blue, red, and green corresponding to the average weights in that particular society. So, for instance, a teal colour may indicate an equal amount of conformism and stubbornness and very low radicalism, while a red line indicates a highly radical society or a blue line a high level of conformism.

Table 7.11 shows 12 **APTE** plots for 3 different initial opinions and 4 different underlying digraphs. These **APTE** plots have a similar behaviour to the ones produced by the BEBA model; the parametric curves not only move towards the y -axis but also move away from it. There are also some curves that appear to move towards the centre of the Agreement Plot, but this behaviour can only be observed for some initial opinions. Looking at the line colours it is clear that the lines moving towards the y -axis are mainly blue, representing highly conformist societies. This makes sense, because in the Classification-based model the conformist trait produces consensus, and thus it is expected that the blue lines behave like the Friedkin-Johnsen and Bounded Confidence lines.

On the other hand, the curves moving away from the y -axis are mostly red and green. For the red lines, clearly radicalism tends to move the agent opinions to the extremes, resulting in a higher mean of the absolute values. For the green lines, the behaviour is unexpected, however upon closer inspection it is clear that these green lines also have a significant red component and this mild radical trait may be responsible for this change.

Regarding noticeable differences due to the underlying digraph, societies evolving over the Scale-Free digraph tend to move more towards the centre of the Agreement Plot, and in contrast the other societies (especially the ones with Complete digraph) move more to the right or left of the plot. One possible explanation could be that, since the Classification-based model allows for differences between the expressed opinion and how it is perceived by influenced neighbours, in networks with longer average path length these 'errors' accumulate, thus diluting the trait's effects and for instance preventing a highly conformist society from reaching consensus (blue curves not going completely to the $y = \pm x$ lines).

Table 7.12 shows 12 **UDTE** plots for 3 different initial opinions and 4 different choices of the agent parameters. These plots suggest that the agent parameters have a definitive effect on the overall opinion evolution: regardless of the underlying digraph topology, the curves in each plot appear to be moving towards the same regions in the Agreement Plot. The directions of these movements are consistent with the observations made for the **APTE** plots: Blue curves move generally towards the left of the Agreement Plot, red and green curves move to the right, and a more equal combination of the three colours moves to the centre-right. Taking into account that all the curves in a given plot move in the same direction, it is possible to conclude that the underlying digraph has an indirect influence on the opinion evolutions that cannot prevail over the agent parameters' effect.

Table 7.13 shows 12 **IOTE** plots for 3 different agent parameters and 4 different underlying digraphs. These plots suggest that, given a set of agent parameters and underlying digraph, all the initial opinions will move towards the same region in the Agreement Plot: this region just depends on the agent parameters and underlying digraph. This is a new behaviour. For the previous models, all the curves either moved to the same direction (Friedkin-Johnsen), or there were two qualitative different (and to some degree opposite) behaviours (for the BEBA model, moving either to the left or to the right of the Agreement Plot, and for the Bounded Confidence model moving to the left or not moving at all). In contrast, for the Classification-based model the general direction the curves follow depends on the initial opinion location in the Agreement Plot and the region towards all curves converge to.

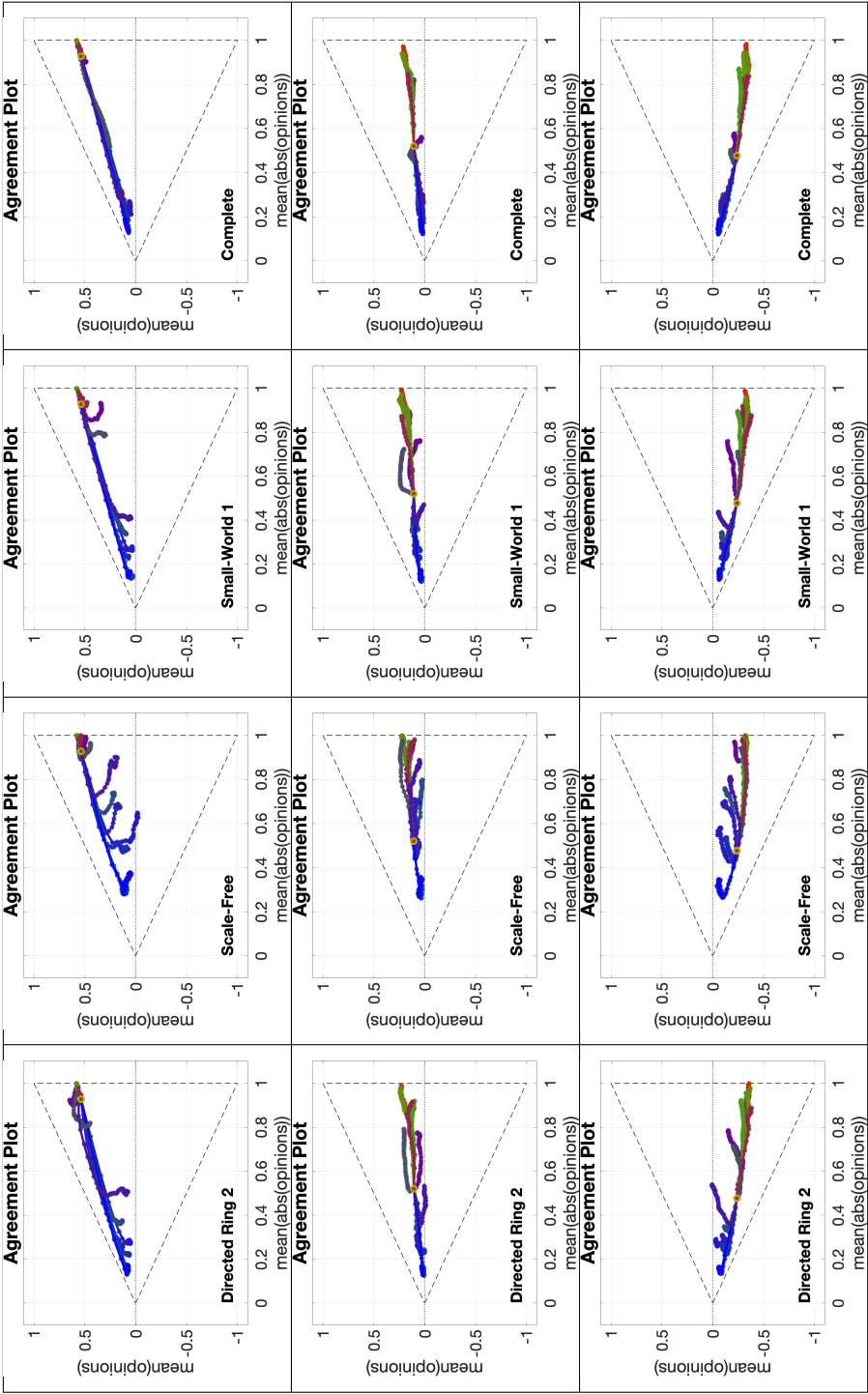


Table 7.11: APTE plots for the Classification-based model. Each of the 12 APTE plots includes 15 curves with different choices of agent parameters and constant initial opinion distributions and underlying digraphs. Plots in the same row start from the same initial opinion distribution (represented by the orange dot). Plots along the same column have the same underlying digraph, (from left to right, Directed Ring 2, Scale-Free, Small-World 1, and Complete, for metrics see Table 6.1). The digraph name can be seen in the bottom left corner, the corresponding topology is shown in Figure 5.1, and some metrics are presented in Table 6.1. All simulations are with 100 agents for 50 time steps.

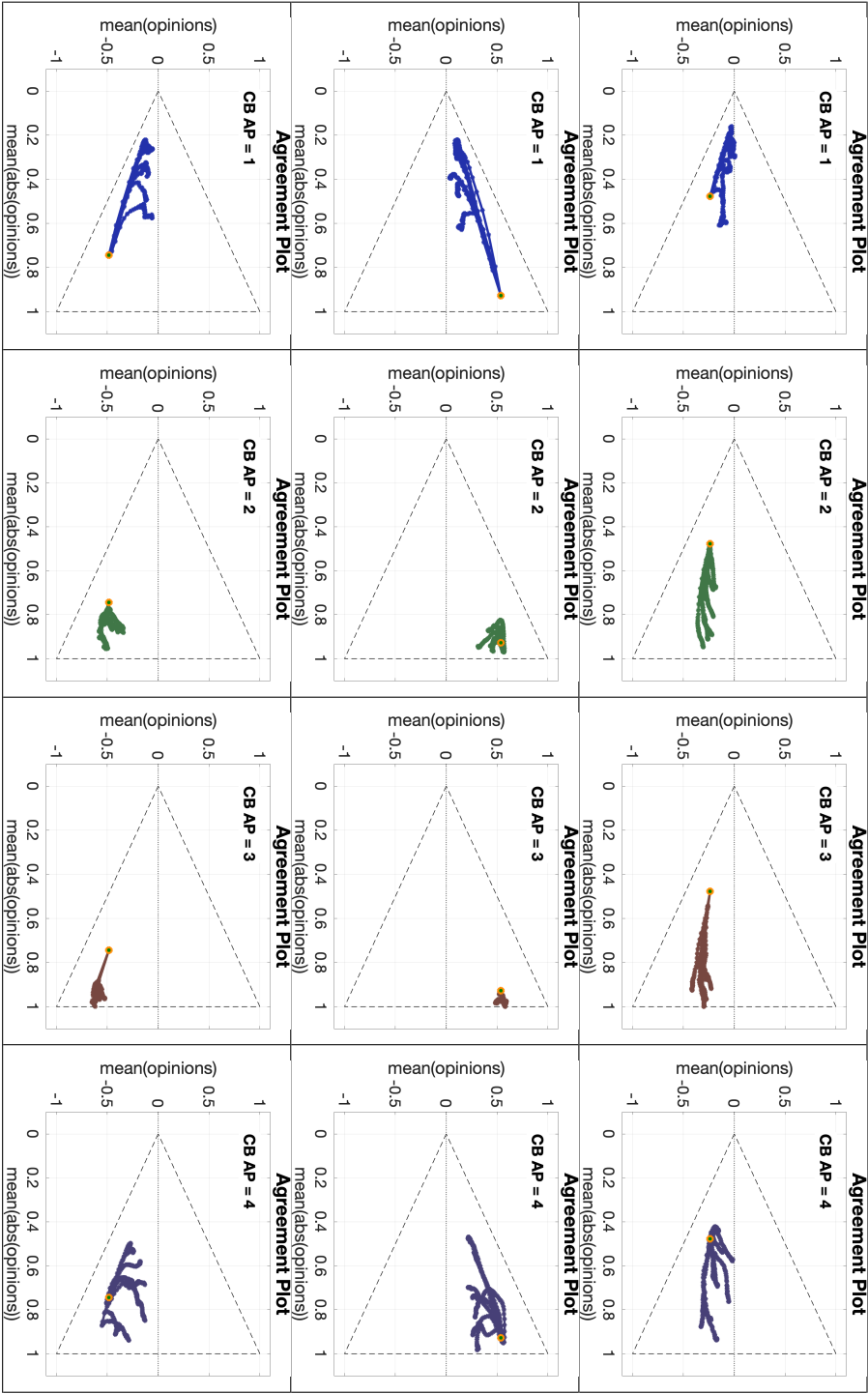


Table 7.12: UDTE plots for the Classification-based model. Each of the 12 UDTE plots includes 45 curves with different choices of underlying digraphs and constant initial opinion distributions and agent parameters. Plots in the same row start from the same initial opinion distribution (represented by the orange dot). Plots along the same column have the same agent parameters. The agent parameter number can be seen in the upper left corner. Figure 7.20 shows a histogram of the corresponding agent parameters. All simulations are with 100 agents for 50 time steps.

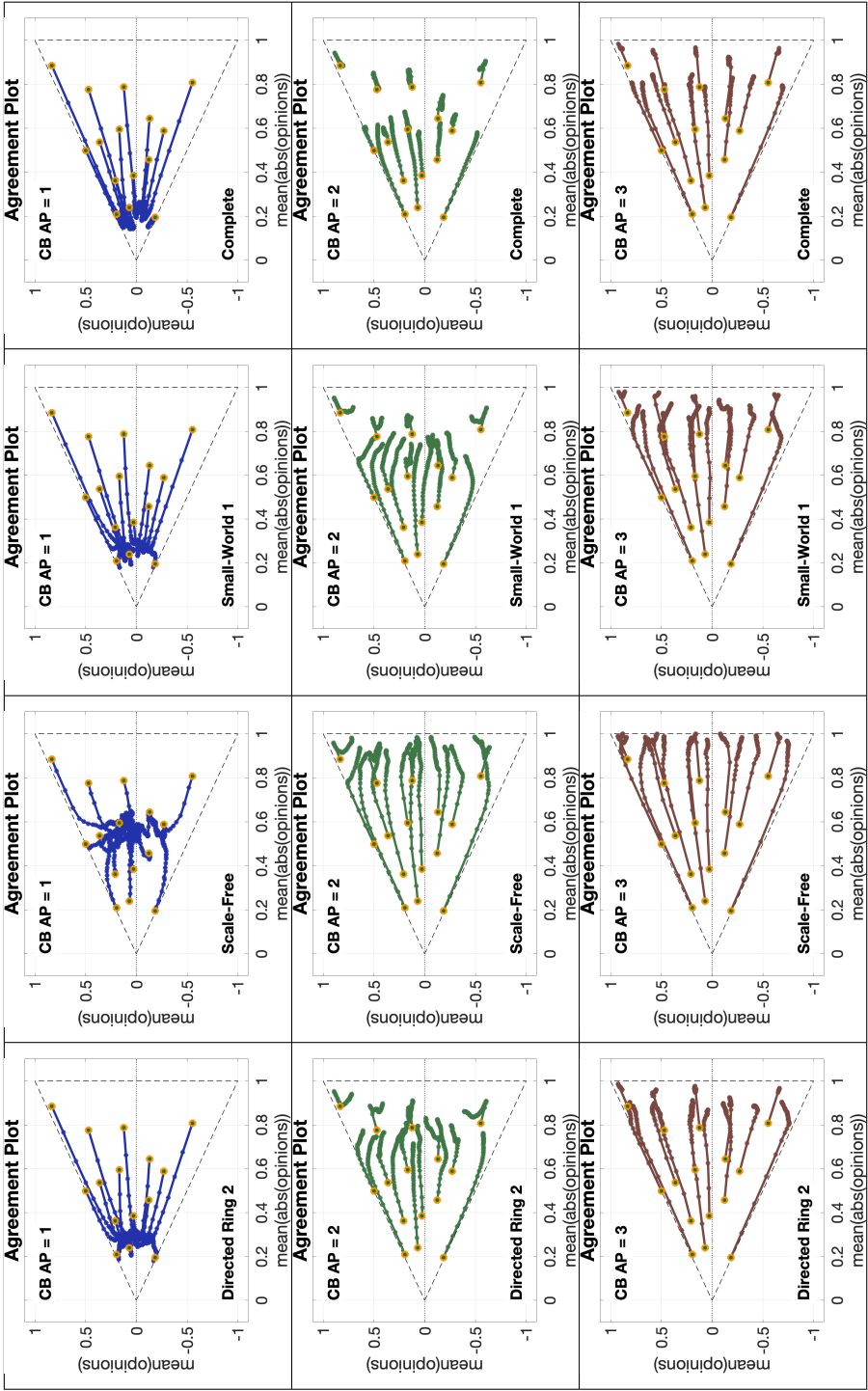


Table 7.13: IOTE plots for the Classification-based model. Each of the 12 IOTE plots includes 15 curves with different choices of initial opinion distributions and constant agent parameters and underlying digraphs. Plots in the same row have the same agent parameters. Plots along the same column have the same underlying digraph, (from left to right, Directed Ring 2, Scale-Free, Small-World 1, and Complete, for metrics see Table 6.1). The agent parameter number can be seen in the upper left corner. Figure 7.20 shows a histogram of the corresponding agent parameters. The digraph name can be seen in the bottom left corner, the corresponding topology is shown in Figure 5.1, and some metrics are presented in Table 6.1. All simulations are with 100 agents for 50 time steps.

The convergence speed (inferred from the distance between points in the parametric curves) seems to depend only on the agent parameters and not the underlying digraph. The distance between points in the green curves is noticeable less than in the other curves, indicating that societies with mostly stubborn traits change more slowly, which is coherent, given that the stubborn trait has the tendency of keeping the opinions unchanged and, as a side effect, decreases the opinion change caused by the other agents' traits.

Table 7.14 shows 12 **APSS** plots for 3 different initial opinions and 4 underlying digraphs. These **APSS** plots look remarkably different from the same type of plots for the other models. The first difference is that, unlike all the other **APSS** plots, the plots in Table 7.14 have few points located along the lines $y = \pm x$ and $x = 1$. The points mostly appear to converge to a connected (and in most cases convex) subset of the Agreement Plot. In some cases the initial point is clearly located in the interior of this region, while in others it is at the boundary. Unlike the other models (especially the BEBA model) the range of opinions that can be achieved is not too wide. This is particularly true when looking at the y -axis components of the points: most predicted points in the Classification-based model have a y component not too far from the initial opinions. This means that the Classification-based model tends to leave the mean of the opinions relatively unchanged.

Looking at the dot colours, we can see the right dots are mostly red, the left dots are mostly blue, and the green dots can be found in the vicinity of the initial opinion distribution point in either direction. This indicates that mostly conformist societies tend to less extreme opinions, mostly radical societies tend to more extreme opinions, and mostly stubborn societies don't change much and the change direction is determined by the trait with the second highest weight.

Table 7.15 shows 12 **UDSS** plots for 3 different initial opinions and 4 different choices of the agent parameters. Like the **APSS** plots, the **UDSS** plots for the Classification-based model are very contrasting with the ones obtained for the other models. The plots in Table 7.15 indicate that, for the Classification-based model, the digraph topology has a more direct and clear effect on the agents' opinion evolution. For instance, looking at column 1, we can see that, even for a highly conformist society, for some underlying topologies extreme opinions can appear (represented by dots near the $x = 1$ line), thus conformism alone is not enough to prevent extreme opinions from forming. Dots in columns 2 and 4 are also very scattered and in many cases can be found around the initial opinion point, further implying that even for a fixed agent parameter set, the digraph topology has a significant effect on the opinion evolution.

The only parameters for which this effect is somewhat less evident are the ones in column 3, where most of the dots are located to the right of the initial opinions distribution dot. This may suggest that the radical trait in the Classification-based model has a more powerful effect than the other traits, overshadowing the influence of the underlying digraph topology.

Table 7.16 shows 12 **IOSS** plots for 3 different agent parameters and 4 different underlying digraphs. These plots are also very interesting. They confirm the existence of 'convergence regions' towards which all the opinions move. Clearly these 'convergence regions' depend on the agent parameters and underlying digraph, as hinted by Table 7.13. These regions are symmetric with respect to reflections along the x -axis, given that the model has no preference for agreement or disagreement. Besides this observation there is no clear relation between the 'convergence region' shape or location and the agent parameters and underlying digraph.

An additional observation is that there are no blue dots with negative opinion mean, and there are no red dots with positive opinion mean. This indicates that the opinion's general agreement does not travel significantly along the y -axis, or in other words that the mean of the opinions does not change much, as previously noted. Finally, we can see that the 'convergence region' is sometimes located away from the $(0,0)$, $(1,-1)$, $(1,1)$ triangle's boundary, which is consistent with the behaviour noticed in Table 7.14.

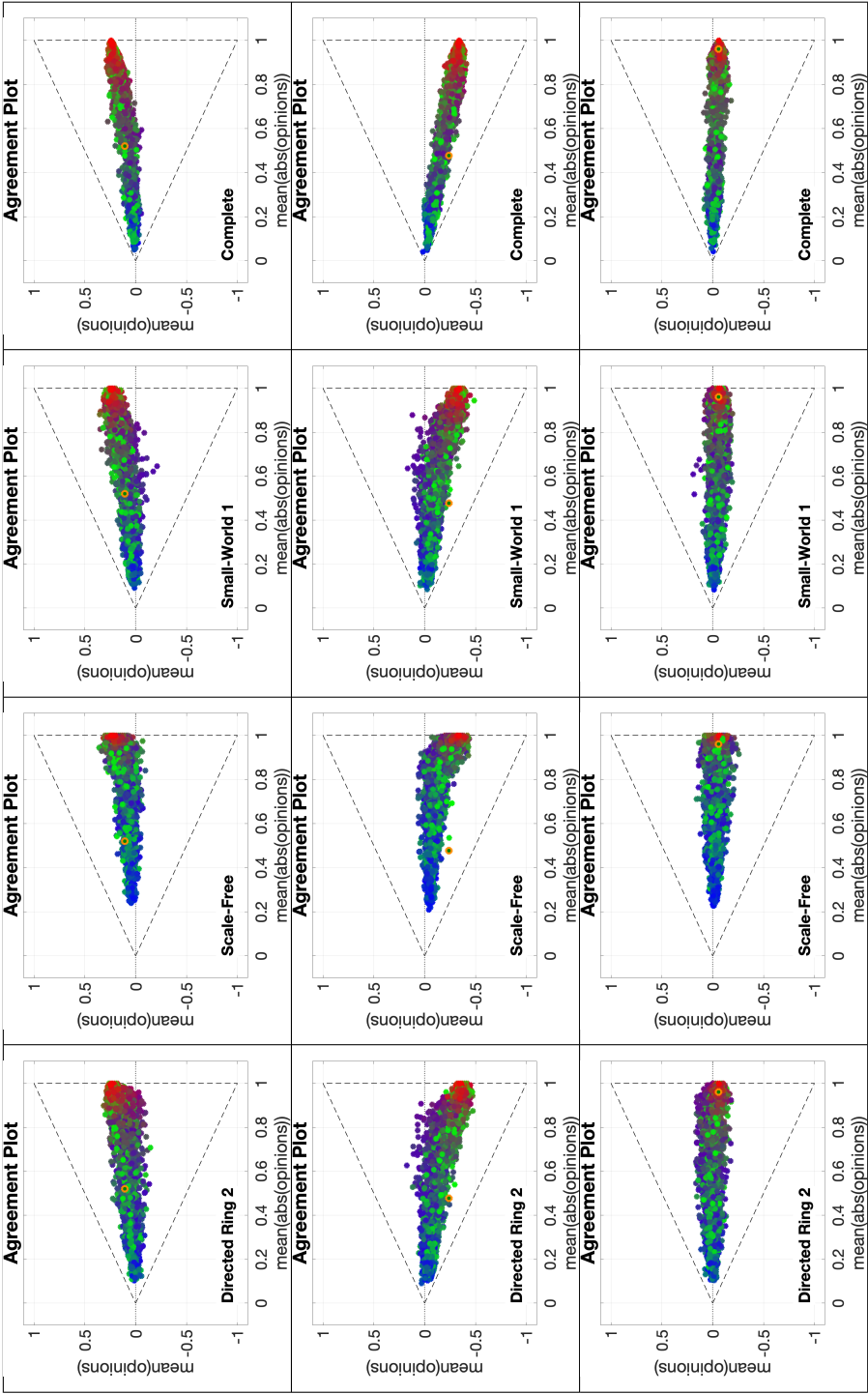


Table 7.14: APSS plots for the Classification-based model. Each of the 12 APSS plots includes 3528 points with different choices of agent parameters and constant initial opinion distributions and underlying digraphs. Plots in the same row start from the same initial opinion distribution (represented by the orange dot). Plots along the same column have the same underlying digraph, (from left to right, Directed Ring 2, Scale-Free, Small-World 1, and Complete, for metrics see Table 6.1). The digraph name can be seen in the bottom left corner, the corresponding topology is shown in Figure 5.1, and some metrics are presented in Table 6.1. All simulations are with 100 agents for 1000 time steps.

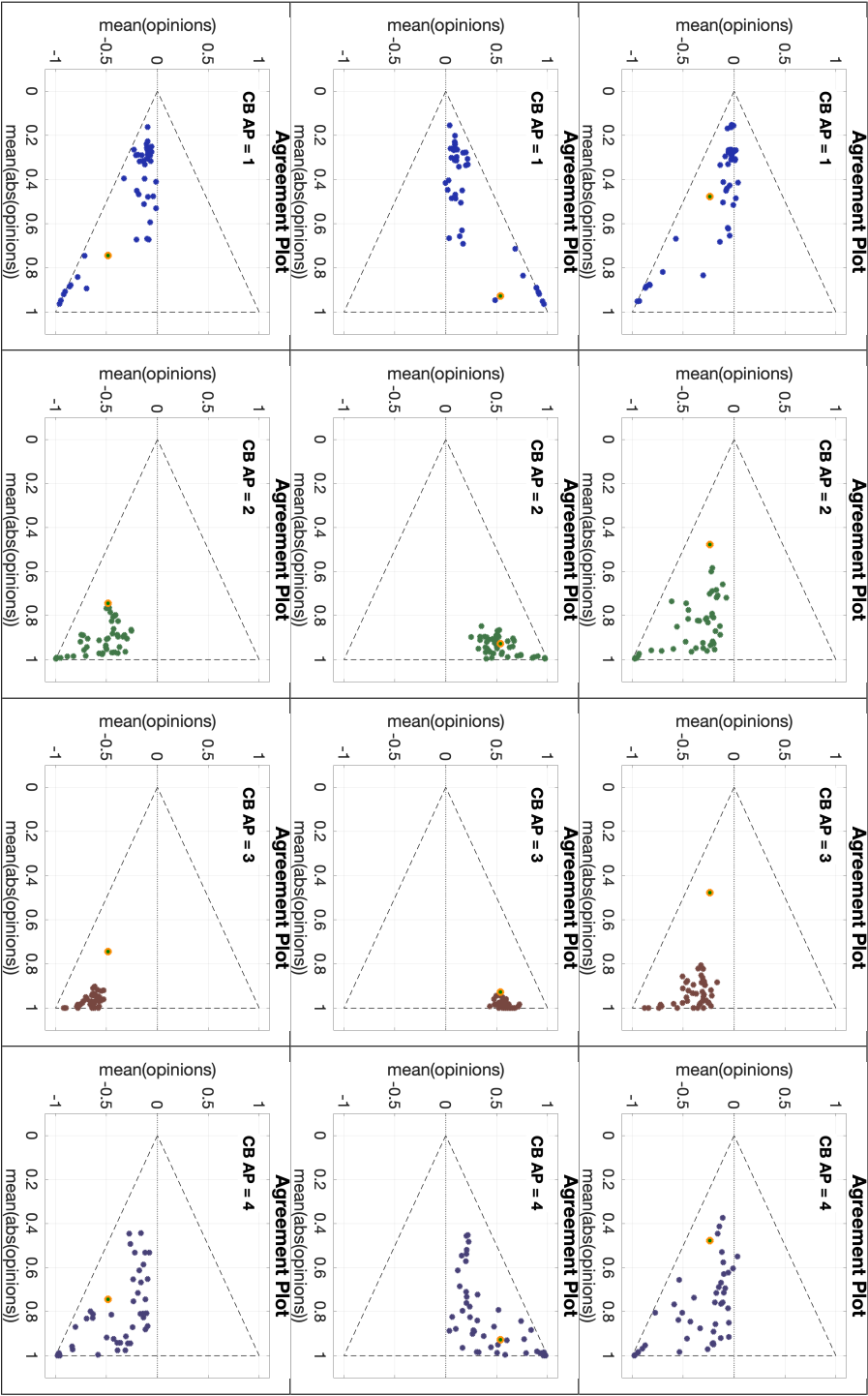


Table 7.15: **UDSS** plots for the Classification-based model. Each of the 12 **UDSS** plots includes 45 points with different choices of underlying digraphs and constant initial opinion distributions and agent parameters. Plots in the same row start from the same initial opinion distribution (represented by the orange dot). Plots along the same column have the same agent parameters. The agent parameter number can be seen in the upper left corner. Figure 7.20 shows a histogram of the corresponding agent parameters. All simulations are with 100 agents for 1000 time steps.

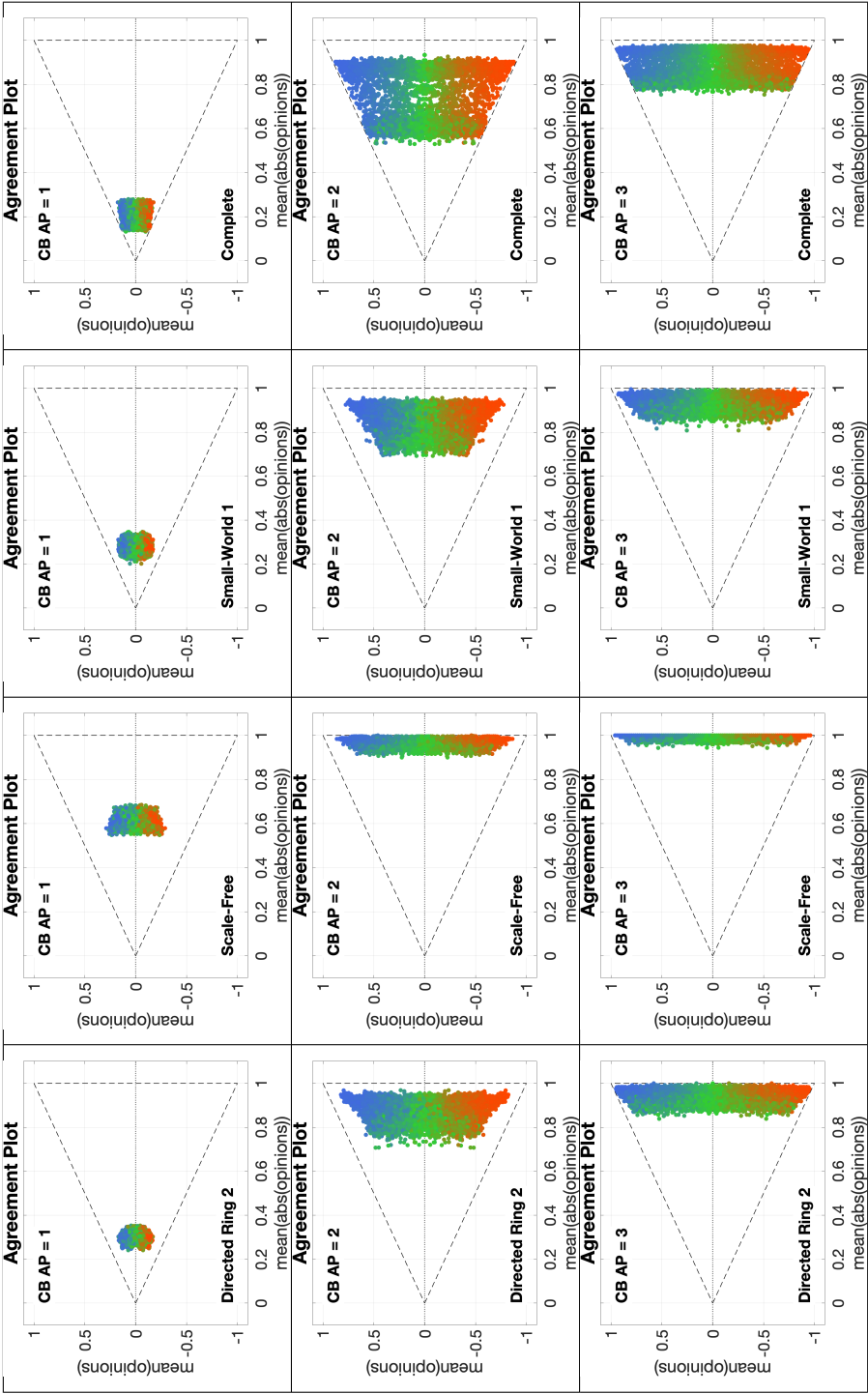


Table 7.16: IOSS plots for the Classification-based model. Each of the 12 IOSS plots includes 5314 points with different choices of initial opinion distributions and constant agent parameters and underlying digraphs. Plots in the same row have the same agent parameters. Plots along the same column have the same underlying digraph, (from left to right, Directed Ring 2, Scale-Free, Small-World 1, and Complete, for metrics see Table 6.1). The agent parameter number can be seen in the upper left corner. Figure 7.20 shows a histogram of the corresponding agent parameters. The digraph name can be seen in the bottom left corner, the corresponding topology is shown in Figure 5.1, and some metrics are presented in Table 6.1. All simulations are with 100 agents for 1000 time steps.

PROBABILISTIC ANALYSIS

Figures 7.21, 7.22, and 7.23 are analogous to Figures 6.18, 6.19, and 6.20 obtained in the Probabilistic Analysis of the Friedkin-Johnsen model.

First, it is important to point out that the agent parameter ordering in Figure 7.21 is different from the other agent orderings. Previously the x -axis (respectively, y -axis) corresponded to the mean (resp. variance) of the agent parameters, whether that was susceptibility, entrenchment, or confidence radius. For the Classification-based model, since each agent has three parameters (which, because they add up to 1, are in reality two *independent* parameters), the axes are changed to be: in the x -axis the mean radical trait weight ($\bar{\beta}$) and in the y -axis the mean stubborn trait ($\bar{\gamma}$). This choice of axis is convenient, because it is intuitive and, as seen in Figure 7.21, the ordering depends almost completely on the mean radical weight. The Agreement Plot ordering has the same axes as before, and the ordering is almost only dependent on the mean of the opinion absolute values.

Therefore, the first rows of the **QOL Tables** will correspond to societies with very extreme opinions, while the last rows are for almost completely indifferent societies. For the first columns the society is highly radical, and in contrast, the last columns have on average a low radical weight and any proportion of conformist and stubborn weights.

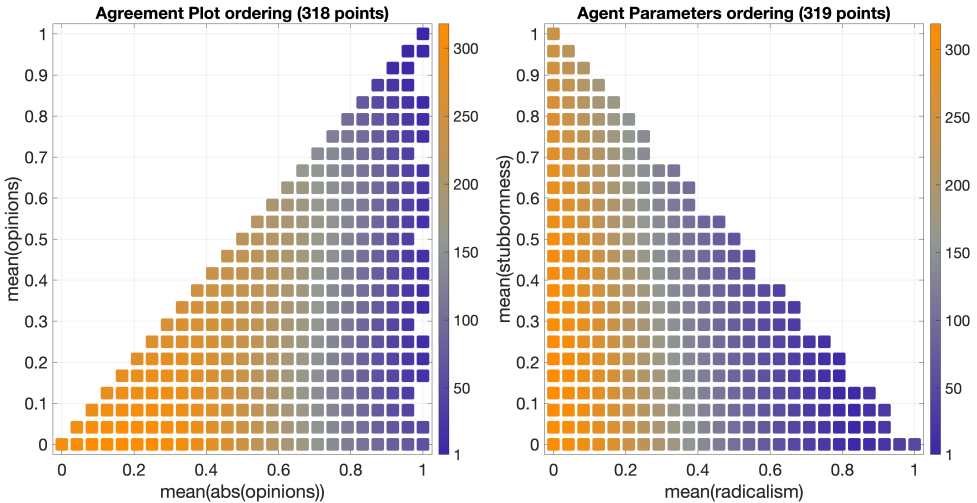


Figure 7.21: Ordering of the points in the Agreement Plot and Parameter Plane, for the Classification-based model. Each point in the Agreement Plot corresponds to a pair $(\bar{x}, |\bar{x}|)$ that creates a non-empty set of opinion distributions $\mathcal{O}$. Each point in the Parameter Plane corresponds to a pair $(\bar{\beta}, \bar{\gamma})$ that creates a non-empty set of agent parameter sets $\mathcal{P}$. The ordering is indicated by the colorbars and will be used in the plotting of the corresponding **QOL Tables** in Figure 7.22 as explained in Figure 6.17.

Looking at Figure 7.22 it can be seen that only perfect consensus and consensus have somewhat well-defined regions in the **QOL Tables**, which suggests that this model's behaviour is highly intricate. Here it is important to clarify that there could exist Agreement Plot and agent parameter orderings that create more well defined regions, however, these orderings probably won't follow an easy-to-interpret pattern and, as such, would not provide relevant information about the model behaviour.

Looking at the **QOL Tables'** histograms, it is possible to see that except for dissensus all qualitative outcomes have significant probability of being achieved. It is also interesting to note that

perfect consensus and consensus are likely outcomes in opposite scenarios. Perfect consensus occurs in the left columns corresponding with highly radical societies, while consensus can be found in scarcely radical societies. This happens because radicalism moves opinions to extremes: as with entrenchment, this often results in perfect consensus. On the other hand consensus requires a clear but not extreme tendency of all agents to have similar opinions, thus going to extremes cannot produce consensus.

Clustering is partially present in the left side of the last rows, which correspond to significantly radical societies starting from almost indifferent opinions. This can be explained in that individuals with radical traits may move from indifference towards a more well defined opinion, but not to an extreme that probably would produce perfect consensus, forming two or three not very distant subgroups that produce clustering. Interestingly, the histograms corresponding to the **QOL Tables** for the Classification-based model are similar to the histograms of the Friedkin-Johnsen model in two ways: (i) in both collections of histograms, there is no single qualitative outcome that is almost guaranteed for the majority of the **QOL Tables** (unlike with the BEBA and Bounded Confidence models, where perfect consensus is the almost sure result for the majority of cells); and (ii) in every table, a relatively significant number of cells has a probability that is not zero or one.

Unlike the **QOL Tables** in Figure 7.22, the plots in Figure 7.23 show clear regions in the Agreement Plot and parameter plane where qualitative outcomes are likely to be found. The first row indicates that only perfect consensus, polarisation, and clustering have clear regions in the Agreement Plot where they have significant probability of appearing: these regions are the vicinity of the points (1, 1), (1, 0), and (0, 0) for perfect consensus, polarisation, and clustering respectively. In the case of perfect consensus and polarisation, a possible explanation is that opinions located in those areas already belong to that qualitative category, therefore if they do not change much (for example by agent parameters with high stubbornness or radicalism) their qualitative category will not change. The possible explanation for the clustering case was previously discussed.

Looking at the second row, perfect consensus, consensus, polarisation, and clustering have regions in the parameter plane where they have significant probability of appearing. These regions seem to be dependent mostly on the average radical weight. From minimum to maximum radical weight, the order in which the qualitative outcomes are more probable is consensus, clustering, polarisation, and perfect consensus. A possible explanation follows the same reasoning previously presented: the existence of these categories requires a degree of intermediate and not extreme opinions: consensus needs multiple similar but not identical opinions, clustering two or more subgroups with distinguishable opinions, polarisation two subgroups with relatively equal number of agents having distant opinions, and finally in perfect consensus, all agents have very similar, and possibly extreme, opinions. The average radical weight directly relates to how extreme the population's opinions are.

Overall, the behaviour and intrinsic properties of the Classification-based model can be summarised as follows:

- Regardless of where the initial opinions are located in the Agreement Plot, they will move towards a 'convergence region' which depends on the agent parameters and the underlying digraph. It is possible for the convergence region to be located near the boundaries of the (0, 0), (1, -1), (1, 1) triangle, but that is not necessary.
- Although the agent parameters have a significant effect in the opinion evolution, the predicted opinion distribution location is also highly dependent on the underlying digraph, much more than in the Friedkin-Johnsen and BEBA models.
- For a given underlying digraph and initial opinion, the collection of possible opinion outcomes for a wide variety of agent parameters forms a connected and possibly convex region in the Agreement Plot that contains the initial point.

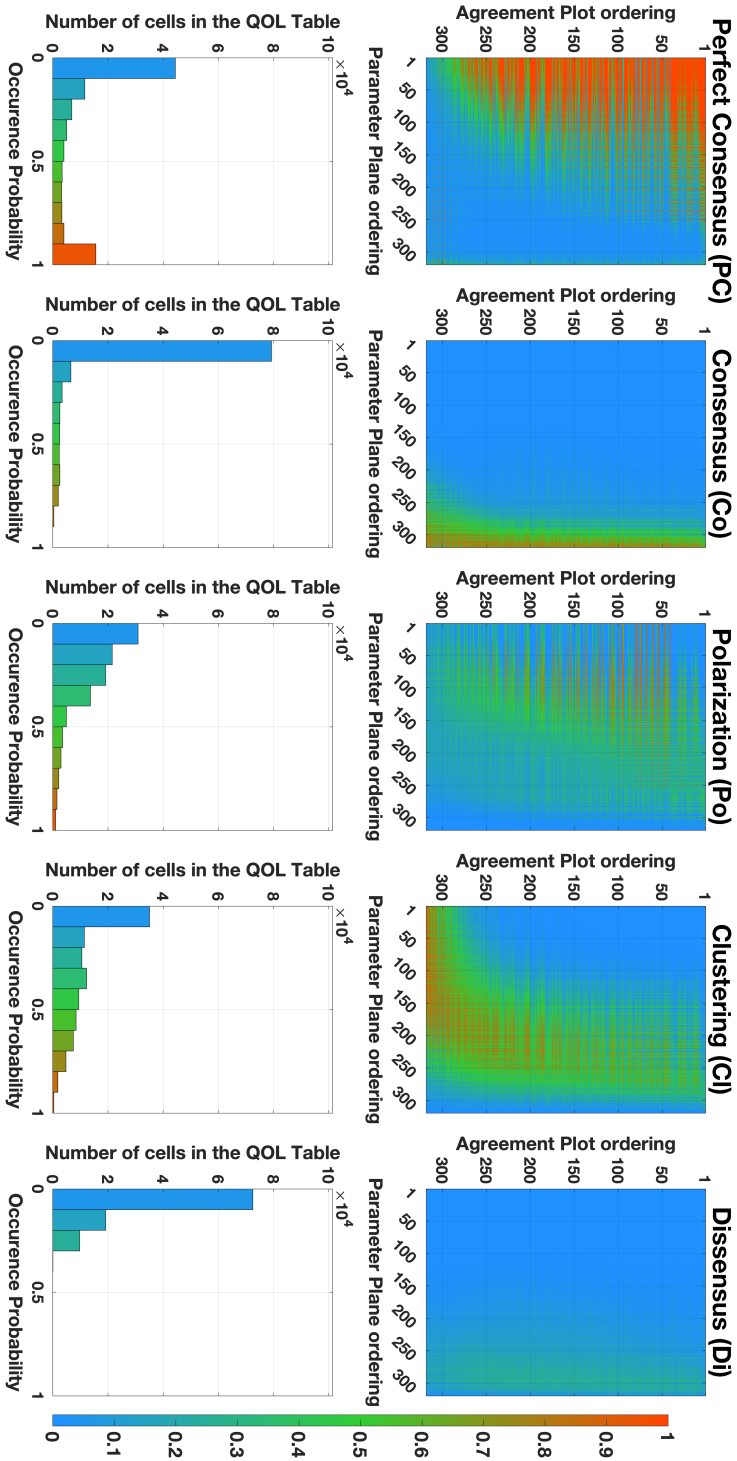


Figure 7.22: First row: **QOL Tables** for the Classification-based model for $B \in \{PC, Co, Po, Ci, Di\}$. The rows corresponds to points $(\bar{\beta}, \bar{\gamma})$ in the Parameter Plane ordered according to Figure 7.21. The columns corresponds to points $(\bar{\beta}, \bar{\gamma})$ in the Parameter Plane ordered according to Figure 7.21. Each cell is related to an initial opinion distribution and agent parameters. The cell colour represents the probability $\mathcal{P}_B$ that the resulting opinion distribution will be of qualitative type $B \in \{PC, Co, Po, Ci, Di\}$ as per Equation 6.2 (for details see Figure 6.17). Second row: the histograms corresponding to the **QOL Tables** presenting the number of cells in the table that have probability $\mathcal{P}_B$ in a given interval.

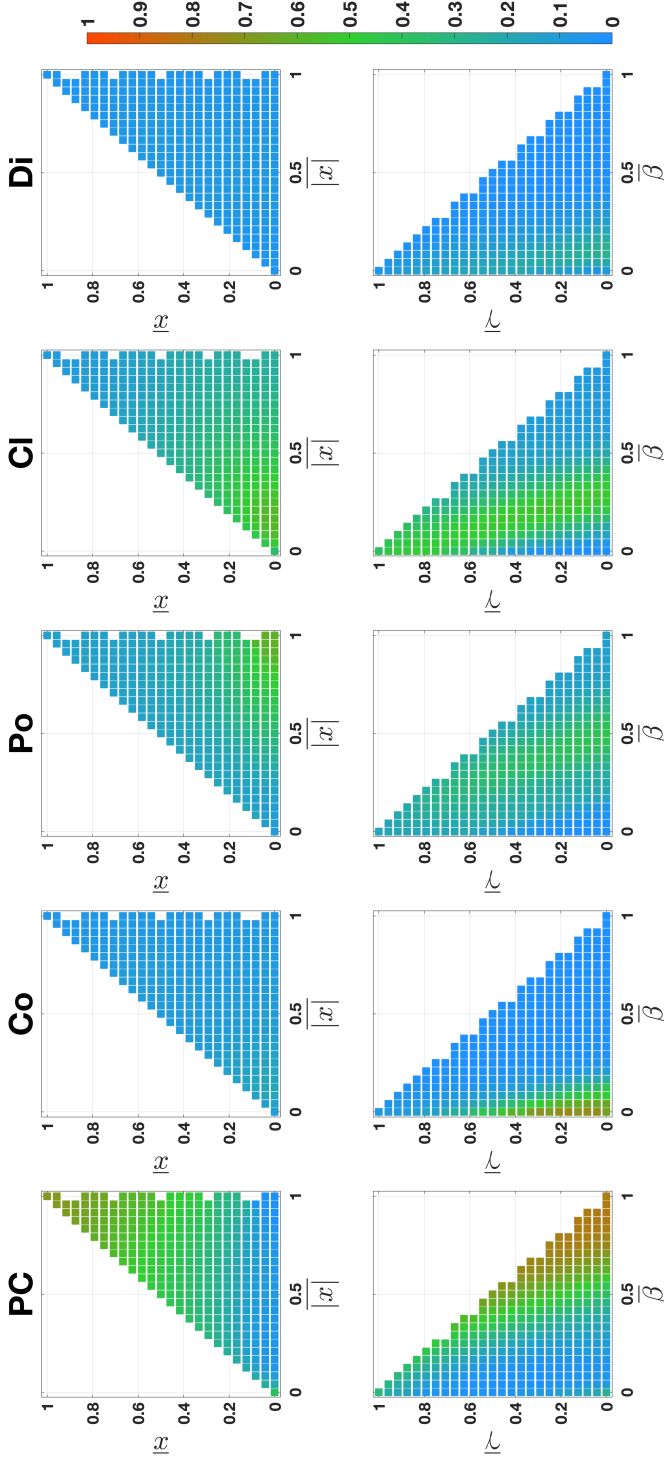


Figure 7.23: For the Classification-based model, when agent parameter information (respectively, initial opinion) information is not available the probabilities $\mathcal{P}_B$ of the final opinions being of a qualitative type $B \in \{PC, Co, Po, CI, Di\}$ can be plotted in the Agreement Plot (resp, Parameter Plane) shown in Figure 7.21. These probability results are shown in the first (resp, second) row.

- It is extremely rare for predicted opinions to be located near the $y = \pm x$ and $x = 1$, meaning that almost always there are some agents that agree and others that disagree, and that outcomes where all agents have extreme opinions are quite rare.
- The relation between qualitative outcomes and initial opinions and agent parameters is remarkably intricate. However, in general, the average radical weight is a good indication of which qualitative outcome is more likely.
- The only qualitative outcome that is very unlikely is dissensus. All the other categories have a significant number of cases where the probability of occurring is around 50%. Also, no qualitative category is almost guaranteed to occur in a majority of the considered cases.

7.3. SUMMARY AND CONCLUSIONS

We have proposed a novel agent-based opinion formation model that has two fundamental distinctive features. First, the model drops the unrealistic assumption that agents can measure the opinion of their neighbours with infinite precision, which drastically affects the opinion evolution, and introduces a novel classification-based approach that more realistically replicates the way individuals assess and evaluate the opinions of their neighbours, by classifying them as agreeing much less, less, comparably, more or much more. Second, the model captures the complexity of the behaviour of individuals by introducing three different internal traits, associated with conformism, radicalism, and stubbornness. Instead of considering agents of different types, the model allows all these tendencies to coexist in each agent, thus representing multifaceted psychological and sociological phenomena in action within each individual.

Five types of simulation analyses were carried out: (i) simulations over simple digraphs and agent parameters to gain insight into the model behaviour; (ii) simulations with varying model parameters to perform a parameter sensitivity analysis; (iii) simulations with parameters chosen through the approximate solution of two optimisation problems to assess the model's potential to predict opinions similar to those seen in real life; (iv) comparison with the Friedkin-Johnsen model; and (v) Agreement Plot and Probabilistic Analysis.

We used real data from the World Values Survey to assess the capability of our classification-based model to mimic actual opinion evolutions seen in real life. Despite its simplicity, the model can yield opinions similar to the ones in survey results and can also produce a rich and wide variety of collective behaviours, comparable to the ones obtained when other mechanisms, such as bounded confidence, randomness, biased assimilation, or backfire effect are used.

Possible further directions for future work include a more detailed study of the effects of different network topologies on the opinion evolution, and an investigation of what happens if the opinion evolution parameters Ω are agent-dependent.

7.4. APPENDIX A: NETWORK METRICS

The signed digraph is represented by the weight matrix $W \in \{-1, 0, 1\}^{N \times N}$, where w_{ij} is associated with the edge going from vertex j to vertex i . We consider the following network metrics: average path length (APL), clustering coefficient (CC), clustering variance ($\sigma(CC)$), mean in- and out-degree (δ^{in} , and δ^{out}), in- and out-degree variance ($\sigma(\delta^{in})$ and $\sigma(\delta^{out})$), positive edges (PE), negative edges (NE), diameter (D), bidirectional coefficient (BC), and balance index (BI). This appendix explains how these metrics are computed.

A directed path is a K -tuple of vertices $(p_1, p_2, \dots, p_i, p_{i+1}, \dots, p_K)$ such that there is an edge from vertex p_i to vertex p_{i+1} for $i = 1, \dots, K - 1$. The length $|p|$ of a directed path p is the number of edges that it crosses. Let $P(i, j)$ be the set of all directed paths from vertex i to vertex j (if there are none, then $P(i, j) = \emptyset$). Denote by $d(i, j)$ the length of the shortest directed path from i to j ,

i.e., $d(i, j) := \min_{p \in P(i, j)} |p|$. Let $C(W)$ be the set of vertex pairs (i, j) such that there exists a direct path from i to j and $i \neq j$, i.e. $C(W) = \{(i, j) \mid P(i, j) \neq \emptyset \text{ and } i \neq j\}$. Then the average path length and diameter of the digraph W are:

$$APL = \frac{1}{|C(W)|} \sum_{(i, j) \in C(W)} d(i, j) \quad \text{and} \quad D = \max_{(i, j) \in C(W)} d(i, j) \quad (7.16)$$

Note that, because all the networks are strongly connected, $|C(W)| = N(N - 1)$.

To compute the mean and variance clustering, consider agent i , with k_i in-neighbours excluding itself: $k_i = |\tilde{\mathcal{N}}_i|$, where $\tilde{\mathcal{N}}_i := \{j \in \mathcal{V} \mid w_{ij} \neq 0, i \neq j\}$. Then there are at most $k_i(k_i - 1)$ directed edges between these neighbours. The fraction c_i of these edges that is actually present is the clustering coefficient of agent i . If agent i has only one in-neighbour, then its clustering coefficient is 1, and if it has no in-neighbour but itself c_i is not defined:

$$c_i = \begin{cases} \frac{|\{(j, k) \mid j \neq k \text{ and } i, k \in \tilde{\mathcal{N}}_i\}|}{k_i(k_i - 1)} & \text{if } k_i > 1 \\ 1 & \text{if } k_i = 1 \\ \text{nan} & \text{if } k_i = 0 \end{cases} \quad (7.17)$$

The mean and variance clustering of the network are thus the average and variance of the clustering coefficients of all agents with at least one in-neighbour excluding themselves. The mean clustering is also sometimes called the clustering coefficient and simply denoted CC (defined by extending to digraphs the definition for undirected graphs by [214]).

For the connectivity degree measures, consider a vertex i . The in-degree (respectively, out-degree) of vertex i is denoted δ^{in} (resp. δ^{out}) and is the number of edges that enter (resp. exit) vertex i . Since each vertex has an individual in- and out-degree it is possible to compute the mean and variance of this collection of numbers.

The number of positive and negative edges are computed as

$$PE = \sum_{i, j \in \mathcal{V}: w_{ij} > 0} 1 \quad \text{and} \quad NE = \sum_{i, j \in \mathcal{V}: w_{ij} < 0} 1 \quad (7.18)$$

The bidirectional coefficient is computed as the ratio between edges for which an edge connecting the same vertices exists in the opposite direction, and the total number of edges. Mathematically it can be computed by:

$$BC = \frac{\sum_{i \neq j \in \mathcal{V}} |w_{ij} w_{ji}|}{\sum_{i \neq j \in \mathcal{V}} |w_{ij}|} \quad (7.19)$$

The numerator of Equation (7.19) counts the number of edges for which there is an edge connecting the same vertices in the opposite direction, and the denominator the total number of edges, excluding self-loops.

Finally, the balance index is computed as

$$BI = \frac{\text{tr}(\exp(W))}{\text{tr}(\exp(D))} \quad (7.20)$$

where $\text{tr}(\cdot)$ is the trace operator, $\exp(\cdot)$ is the matrix exponential, and $D = |W|$ component-wise. This formula is a direct extension of the balance index for undirected graphs proposed by [70, 71].

7.5. APPENDIX B: SIMULATION PROCESS

The **free optimisation problem** in Equation (7.10) with sets $\mathcal{W} = \tilde{\mathcal{W}}$ and $\mathcal{A} = \tilde{\mathcal{A}}$ was solved using the algorithm:

1. **Input:** survey answers for waves 5 and 6 for a given country.
2. Set $w_0 = \infty$; this will be the minimum cost across all networks
3. For network $W \in \tilde{\mathcal{W}}$
 - For question $q \in \{1, 2, \dots, 30\}$
 - Set $\nu_q = \infty$ to be the minimum cost for question q
 - For inner traits assignment $\psi^{(l)} \in \tilde{\mathcal{A}}$
 - ◊ Compute the predicted opinions $\tilde{y}_q$ after K iterations evolving over the network W with inner traits assignment $\psi^{(l)}$ starting with initial opinions x_q . These initial opinions are the survey results to question q in wave 5.

$$\tilde{y}_q = \mathcal{F}_\Omega(x_q, W, \psi^{(l)}, K)$$

- ◊ Compute the mismatch J (Equation (7.15)) between these predicted opinions $\tilde{y}_l$ and the real opinions y_l given by survey results of question q in wave 6.
- ◊ if $J(\tilde{y}_q, y_q) < \nu_q$
 - Set $\nu_q = J(\tilde{y}_q, y_q)$ as the current minimum cost across all inner traits assignments.
 - Set $\widehat{\psi^{(q)}} = \psi^{(l)}$ as the inner traits assignment that gives the lowest cost for question q .
- Add all the minimum costs to obtain the minimum cost for the network W

$$J_{\text{Total}} = \sum_{q=1}^{30} \nu_q$$

- if $J_{\text{Total}} < w_0$
 - Set $w_0 = J_{\text{Total}}$ as the current minimum cost across all networks for this country.
 - Set $\widehat{W} = W$ as the network that produces the minimum cost for this country.
- 4. **Output:** network $\widehat{W}$ and set of inner traits assignments $(\widehat{\psi^{(l)}})_{l=1}^{30}$ that give the minimum total cost across all questions.

In the algorithm used to solve the **constrained optimisation problem** in Equation (7.11), both the network and the inner traits assignments are the same for each question:

1. **Input:** survey answers for waves 5 and 6 for a given country.
2. Set $w_0 = \infty$; this will be the minimum cost across all networks and inner traits assignments
3. For network $W \in \tilde{\mathcal{W}}$
 - For inner traits assignment $\psi \in \tilde{\mathcal{A}}$
 - For question $q \in \{1, 2, \dots, 30\}$
 - ◊ Compute the predicted opinions $\tilde{y}_q$ after K iterations evolving over the network W with inner traits assignment ψ starting with initial opinions x_q . These initial opinions are the survey results to question q in wave 5.

$$\tilde{y}_q = \mathcal{F}_\Omega(x_q, W, \psi, K)$$

- ◇ Compute the mismatch J (Equation (7.15)) between the predicted opinions $\hat{y}_q$ and the real opinions y_q given by survey results of question q in wave 6.
- Add all the costs to obtain the cost for the network W and the inner traits assignment ψ .

$$J_{\text{Total}} = \sum_{q=1}^{30} v_q$$

- if $J_{\text{Total}} < w_0$
 - ◇ Set $w_0 = J_{\text{Total}}$ as the current minimum cost across all networks and inner traits assignments for this country.
 - ◇ Set $\hat{W} = W$ as the network that produces the minimum cost for this country.
 - ◇ Set $\hat{\psi} = \psi$ as the inner traits assignment that gives the minimum cost for this country.
- 4. **Output:** network $\hat{W}$ and inner traits assignments $\hat{\psi}$ that give the minimum total cost across all questions.

7.6. APPENDIX D: COUNTRIES AND QUESTIONS

We report here the list of countries and the list of questions we considered, from the real data collected by the World Values Survey. Table 7.17 reports the countries used in the Transition Table analysis of Section 6.2. Table 7.18 reports the countries used in the Optimisation problems of Section 7.2.3. Table 7.19 reports the questions from the World Values Survey used in Section 6.2 and Section 7.2.3.

C1	Argentina	C2	Australia	C3	Brazil
C4	Chile	C5	China	C6	Colombia
C7	Cyprus	C8	Egypt	C9	Germany
C10	Hong kong	C11	Iraq	C12	Jordan
C13	Japan	C14	Malaysia	C15	Mexico
C16	New Zealand	C17	Peru	C18	Romania
C19	Russia	C20	South Korea	C21	Thailand
C22	Taiwan	C23	Turkey	C24	Ukraine
C25	United States				

Table 7.17: Countries used in the Transition Table analysis of Section 6.2

C1	Australia	C2	Brazil	C3	Chile
C4	China	C5	Cyprus	C6	Georgia
C7	Ghana	C8	India	C9	Jordan
C10	Japan	C11	Malaysia	C12	Mexico
C13	Poland	C14	Romania	C15	Slovenia
C16	South Africa	C17	Spain	C18	Sweden
C19	South Korea	C20	Thailand	C21	Trinidad
C22	Taiwan	C23	Turkey	C24	Ukraine
C25	United States	C26	Uruguay		

Table 7.18: Countries used in the Optimisation problems of Section 7.2.3

Q1	Some people feel they have completely free choice and control over their lives, while other people feel that what they do has no real effect on what happens to them. Please use this scale where 1 means no choice at all and 10 means a great deal of choice to indicate how much freedom of choice and control you feel you have over the way your life turns out
Q2	All things considered, how satisfied are you with your life as a whole these days? Using this card on which 1 means you are completely dissatisfied and 10 means you are completely satisfied where would you put your satisfaction with your life as a whole?
Q3	How satisfied are you with the financial situation of your household? Please use this card again to help with your answer (1 is completely dissatisfied, 10 is completely satisfied)
Q4	How would you place your views on this scale? 1 means you completely agree with the statement Incomes should be made more equal ; 10 means you completely agree with the statement We need larger income differences as incentives for individual effort . And if your views fall somewhere in between, you can choose any number in between.
Q5	How would you place your views on this scale? 1 means you completely agree with the statement Private ownership of business and industry should be increased ; 10 means you completely agree with the statement Government ownership of business and industry should be increased . And if your views fall somewhere in between, you can choose any number in between.
Q6	How would you place your views on this scale? 1 means you completely agree with the statement The government should take more responsibility to ensure that everyone is provided for ; 10 means you completely agree with the statement People should take more responsibility to provide for themselves . And if your views fall somewhere in between, you can choose any number in between.
Q7	How would you place your views on this scale? 1 means you completely agree with the statement Competition is good. It stimulates people to work hard and develop new ideas ; 10 means you completely agree with the statement Competition is harmful. It brings out the worst in people . And if your views fall somewhere in between, you can choose any number in between.
Q8	How would you place your views on this scale? 1 means you completely agree with the statement In the long run, hard work usually brings a better life ; 10 means you completely agree with the statement Hard work doesn't generally bring success—it's more a matter of luck and connections . And if your views fall somewhere in between, you can choose any number in between.
Q9	How much you agree or disagree with the statement Science and technology are making our lives healthier, easier, and more comfortable.. For this questions, a 1 means that you "completely disagree" and a 10 means that you "completely agree."
Q10	How much you agree or disagree with the statement Because of science and technology, there will be more opportunities for the next generation. For this questions, a 1 means that you "completely disagree" and a 10 means that you "completely agree."
Q11	How much you agree or disagree with the statement We depend too much on science and not enough on faith.. For this questions, a 1 means that you "completely disagree" and a 10 means that you "completely agree."
Q12	All things considered, would you say that the world is better off, or worse off, because of science and technology? 1 means that "the world is a lot worse off," and 10 means that "the world is a lot better off."
Q13	How important is God in your life? 10 means "very important" and 1 means "not at all important."
Q14	Indicate if the action of Claiming government benefits to which you are not entitled can be never justified (1); always justified (10); or something in between in a scale from 1 to 10.
Q15	Indicate if the action of Cheating on taxes if you have a chance can be never justified (1); always justified (10); or something in between in a scale from 1 to 10.
Q16	Indicate if the action of Someone accepting a bribe in the course of their duties can be never justified (1); always justified (10); or something in between in a scale from 1 to 10.
Q17	Indicate if the action of Homosexuality can be never justified (1); always justified (10); or something in between in a scale from 1 to 10.
Q18	Indicate if the action of Abortion can be never justified (1); always justified (10); or something in between in a scale from 1 to 10.
Q19	Indicate if the action of Divorce can be never justified (1); always justified (10); or something in between in a scale from 1 to 10.
Q20	Indicate if the action of Suicide can be never justified (1); always justified (10); or something in between in a scale from 1 to 10.
Q21	Indicate if the action of For a man to beat his wife can be never justified (1); always justified (10); or something in between in a scale from 1 to 10.
Q22	Governments tax the rich and subsidize the poor. an essential characteristic of democracy? Use this scale where 1 means "not at all an essential characteristic of democracy" and 10 means it definitely is "an essential characteristic of democracy"
Q23	Religious authorities interpret the laws. an essential characteristic of democracy? Use this scale where 1 means "not at all an essential characteristic of democracy" and 10 means it definitely is "an essential characteristic of democracy"
Q24	Is People choose their leaders in free elections. an essential characteristic of democracy? Use this scale where 1 means "not at all an essential characteristic of democracy" and 10 means it definitely is "an essential characteristic of democracy"
Q25	Is People receive state aid for unemployment. an essential characteristic of democracy? Use this scale where 1 means "not at all an essential characteristic of democracy" and 10 means it definitely is "an essential characteristic of democracy"
Q26	Is The army takes over when government is incompetent. an essential characteristic of democracy? Use this scale where 1 means "not at all an essential characteristic of democracy" and 10 means it definitely is "an essential characteristic of democracy"
Q27	Is Civil rights protect people's liberty against oppression. an essential characteristic of democracy? Use this scale where 1 means "not at all an essential characteristic of democracy" and 10 means it definitely is "an essential characteristic of democracy"
Q28	Is Women have the same rights as men. an essential characteristic of democracy? Use this scale where 1 means "not at all an essential characteristic of democracy" and 10 means it definitely is "an essential characteristic of democracy"
Q29	How important is it for you to live in a country that is governed democratically? On this scale where 1 means it is "not at all important" and 10 means "absolutely important" what position would you choose?
Q30	And how democratically is this country being governed today? Again using a scale from 1 to 10, where 1 means that it is "not at all democratic" and 10 means that it is "completely democratic," what position would you choose?

Table 7.19: Questions from the World Values Survey used in Section 6.2 and Section 7.2.3.

7.7. APPENDIX E: TABLES

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	C16	C17	C18	C19	C20	C21	C22	C23	C24	C25	C26
Q1	4.2	3	3	3.6	3	2.8	2.6	2.6	2.4	2.4	1.6	3.2	3.2	2.2	2.8	3.6	2.2	3	2.2	2.4	2.8	3.6	3.2	2	3.2	3.6
Q2	5.2	0.8	2.4	3.6	4	4.4	4	3.6	2.2	2.8	2.2	2.2	4.2	2.6	3.4	3.4	3.6	4.4	2	2.2	2.4	2.4	2.6	2.8	6.2	3.4
Q3	4	3.8	3.4	3.2	2.4	4.6	2.4	4.8	2.6	5	2.6	4	4	6.4	4	3	3.2	3.8	2.8	2.2	2.2	2.4	2.2	2.4	4	3
Q4	5	2.4	3.2	5.2	4.6	9	6.4	6.6	2.4	4	6	5.2	6.4	6.2	2.2	2.6	4.4	6.4	4.6	22.6	3	4.8	3.4	39.6	4.4	3.8
Q5	3.4	2.8	2.6	9.6	3.8	4.4	3.2	9.8	6.4	2	4	2.6	3.6	11.4	3.4	3.6	3.6	4.6	2.6	3.8	4.6	2.6	3.8	3	2.2	3.8
Q6	4.4	2.2	2.2	2.2	2.4	2.2	4.6	6.8	3	1.8	3	2.6	2	4	3.2	1.8	1.6	3.4	4	12.4	1.8	3.4	2.6	3.2	1.6	5.2
Q7	1.4	2.4	3.2	3.2	5.2	2.4	3.4	4	2.8	4.8	2.8	2.6	3.6	4.8	3.8	2.8	2.8	2.2	2.4	2.4	1.8	2.2	4.8	2.2	3	5.8
Q8	4.2	2	2.6	3.8	4.2	6	4	5.2	2.2	3	3.2	2	3.4	3	3.8	2.2	3.8	1.8	5.4	3.2	2.4	2.6	3.2	2.4	2	3
Q9	3.6	3.4	2.8	4.2	2.4	2.4	2.8	2.6	3.8	4.6	2.8	3.6	3.4	4.2	2.2	4.4	3.6	3	2.8	3.8	3.4	2.4	3.8	2.6	3.8	1.6
Q10	3.2	2.4	2.8	3.8	3	2.4	2.4	2.2	5	3.2	2.4	2.4	2.4	4.6	3	7.8	4.2	3.8	3	3.6	3.8	2.2	2.4	2.4	2.4	3
Q11	3	2.4	5.2	2.6	4.2	3.4	3.4	2.6	6.8	3.2	4	3	4.2	4.2	3.4	4	2.8	3.6	2.6	3.6	5	6.6	3	4	2.4	3.4
Q12	4	2.6	5.4	5.8	2.6	13.4	5	3.2	3.8	3.6	2.4	4.2	3.2	1.4	2.8	3.8	2.2	2.6	2.2	2.2	2.8	2.2	2.6	3.8	2.2	2.8
Q13	6	0.8	2.2	4.6	1.8	1.8	1	5	0.8	4	1.6	1.8	4.2	2.4	4	2.4	3.2	3.2	2.2	39.8	1	2.8	0.8	3	4	7.6
Q14	0.8	1.4	2.8	3.8	1.6	1.8	4.8	1.8	2.8	2.4	2.2	3.2	2.6	1	3	3.8	1.6	1.8	2.2	4.2	2.8	2.8	4.2	1.8	2.4	2.4
Q15	1.4	2.2	1.8	3	2	1.4	1.8	3.2	3.6	1.2	2.8	1.4	1.8	1.8	2.4	2.4	1.4	2	1.4	2.6	1.8	3	1.4	2.4	1.6	1.8
Q16	1.2	0.8	2.4	1.6	0.6	0.8	3.2	4.8	1.6	1.4	1.8	1.4	1	0.6	2.2	3	1.4	1.4	2.6	3.2	1.2	1.8	1	2.6	0.6	1.4
Q17	3.2	4.2	4.4	3	2	0.4	1.8	1.8	1.2	3.6	2	2.6	3.2	1.8	3.8	3.2	2.2	2	2.6	2	0.8	4.2	2	2	8.4	2.6
Q18	4	1	2.2	4.6	3.2	2	2.2	4.4	1.4	3.8	2.4	2	3.4	2.8	3.4	3.6	2.4	3	3	2.4	1.4	2.4	1.8	2.2	4.2	2.8
Q19	3.4	2.6	3.6	6	3.2	4	3.2	4.2	2.4	4.6	2.4	3.2	4	2.4	2.8	2.4	2	2.2	3.8	4.6	8.2	3.2	1.6	3.8	5	2.6
Q20	4	0.2	1.2	1	2	1	0.4	2.6	0.6	2.2	2.6	1.4	2.2	0.4	3.2	4.2	2	3.6	2	3	0.8	2.2	1.4	0.8	3.6	3
Q21	0.2	1.2	1	2	1	0.4	2.6	2.6	1.6	1.6	2.4	1	0.6	0.4	0.6	3.6	1.4	0.4	1.6	3.4	1.4	1.4	1.2	1.2	0.4	1
Q22	3.2	3.2	2.8	7.6	3.2	2.8	4.6	4.8	7.8	2.6	4.4	2.8	3	3	3.2	3.6	2.4	4	2.4	2.4	3.6	3.2	2.2	3.8	3	5.2
Q23	2	4.2	2.8	4	2.2	3.2	3.4	3.4	10.4	3.6	4.2	2.6	2.4	4	2.6	10	4.4	3.8	2.4	2.2	2.4	1.8	3.4	2.8	1.8	2.8
Q24	1.6	2.6	2.2	6.4	3.2	3.6	2.4	4.8	4	2.4	4	4	2	6	3.4	3.4	2.2	1.8	3.2	2.2	2	1.6	2.4	3.6	2.8	2.4
Q25	2.2	3.4	2.4	3.2	5	9	7	4	6	3	3	3	3	3.4	4.6	3	2.8	1	2.8	3	2.4	2.2	2	2.6	2.2	2.4
Q26	2.6	3.2	3	5.8	3	3.2	2.6	3.6	17.4	1.4	3.2	2.6	3.2	2.4	3	5.2	2.2	3.2	3.2	6.8	5.4	2.2	3.4	4.4	3.4	4.4
Q27	3	2.8	6.2	3	6.6	10.4	10.8	3.4	6.2	2.8	2	3.8	2.8	7.6	2.2	4.2	1	4.6	3	3	3.8	2.6	3.6	2	3.8	5
Q28	2.4	1.6	1.4	2.6	3.6	5.8	4.6	2.8	2.8	4	2.4	1.6	2	5.4	2.2	2.2	1.2	0.8	4.4	2.6	4.8	2.8	2.8	2.8	3.2	2.8
Q29	2.4	3.4	1.6	2.4	1.6	2.4	2.6	3	2.4	3.6	2.2	1.6	1.8	1.4	2.6	3.6	1.8	1.2	2.4	1.6	1.4	2.4	4.2	3.2	1.8	1.8
Q30	2.8	2.6	5.2	3.8	9	3.2	4.4	1.6	3.4	5	3	3.4	2	4	3.4	3.2	4.2	4.2	6.4	2.2	1.2	2.4	4.2	4.4	1.6	3.6
Average	3.1	2.4	2.9	4	3.2	3.8	3.7	4.1	4	3.1	2.9	2.7	3	3.6	3	3.7	2.5	2.9	3	5.2	2.9	2.7	2.7	4.1	3	3.3
Total	92	71.6	88.4	121.4	96.4	114.4	109.6	122	119.8	93.6	85.6	82	89.2	107	89	109.8	76	88.2	88.8	154.6	87.2	82.4	79.8	122.2	90.8	97.8

Figure 7.24: Average cost along all the countries = 3.2815. Results of the *Free* optimization problem using the Classification-Based model. From 780 possible question-country pairs 755 have a cost less than 7 (an accuracy of 97% in total). The average cost of accepted question-country pairs is 2.97.

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	C16	C17	C18	C19	C20	C21	C22	C23	C24	C25	C26
Q1	2	8	5.4	7.4	3.4	6	8.4	13.4	12.8	7.6	3	1.2	3.4	6	9	12.8	4	3.8	2.8	11.6	9.6	2.2	6.2	7.8	3.4	5.6
Q2	6	4.6	5.2	7.8	10.4	9.8	7	16.8	16.6	3.8	5.2	5	3.4	16	4.4	12.4	11	2	4	6.6	4	5.2	5.2	4.6	5.8	4.8
Q3	3.6	7.6	7.6	10.8	9.2	4.2	6.4	16.2	19.8	5	4.8	3	13.6	21.2	2.4	6.6	10.6	2.2	2.2	11.2	5.2	3.2	4	4.6	5.8	9.8
Q4	19	13	23	26.2	37	13.4	14.6	38	23.2	19.2	9.6	13.2	12.2	26	26.8	11.6	9.4	21	5.2	39.2	7	17.4	10.6	69.2	14.8	8.4
Q5	4.4	6.8	12	20	7.8	5	14.2	16	29.2	4.2	10.8	5.6	9.8	25	7.4	18.2	1.8	11	3	12.2	10	4.8	5.4	5.8	6.8	5.2
Q6	5.2	8.4	10.4	14.2	17	7.4	9.2	21.6	13.2	6.8	11.8	15.8	5.6	10.2	6	17.2	2.6	18.4	4	25.6	18.6	3.8	8	18.8	5.8	7.6
Q7	10.8	8.4	16.6	6.8	5	5	6.4	7.2	7.4	12.2	5.8	14	8.4	2.6	12	11.4	28.4	6.2	5.4	5	6.4	3.8	9.8	3.8	4.8	9.2
Q8	12.4	10	6.2	6.8	12	8.4	13.2	10.8	15	6.2	16.8	2	4.4	9.4	7.2	26.4	5.4	7	6.2	17	2.2	3.8	4.8	8.4	5.4	4.8
Q9	14.2	7	5.8	5.8	30	6.8	15.2	13.6	22.4	17.8	16.2	3.8	16.8	9	10.8	9.6	9.2	17.2	10	10.2	7.6	6	12.2	8.6	9	20
Q10	21.6	12.8	11.2	4.8	32.4	6	10	14.4	25.4	13.8	13	8.4	10.8	7.8	10.8	14.2	6.2	12.6	6.2	9.6	7	4	13.2	12.2	3.8	27
Q11	22	4.2	10.2	7.4	21.2	14.6	15.8	17.2	36.6	18.4	13.4	4.4	23.6	13.4	23.2	12.6	19	14.4	5.4	8.8	17.8	23.2	10.4	10.4	7	4
Q12	17	10.2	7.8	6.8	2.8	46.4	9	31.6	20	26.8	20.8	9.8	17	11.4	8.2	2.8	15.8	9.6	16.2	10.4	9.2	3.8	4.8	4.4	7.2	16.6
Q13	15.2	3.4	9.8	5.2	4.4	2.2	1.8	21.8	5.4	2.8	19.4	2.2	14.8	3.4	10	20.6	5	6.2	2.6	73.4	1.8	3.6	4.8	3.4	8.2	14
Q14	2.2	10	12.6	8.6	9.6	3.4	12.4	10.4	12.8	2	26.2	9	7.4	4.6	8.6	39.4	8.2	6	7.6	26.6	19	13.8	3.2	14.4	2.2	6.8
Q15	4.6	24.8	5.6	7.2	10.4	3.8	12.2	27	8.4	3.2	20.4	8.6	1.6	5.2	12.8	29.8	6.6	6.6	2	19.4	16	3	2.4	12.2	3.2	9.2
Q16	2.4	8.4	4.4	6.8	3.6	2.4	12	27.4	8.4	3.4	16.6	7.2	3	2.4	10.4	37.8	5	2.6	2.8	21.2	7	4	0.8	8.2	1.2	7.6
Q17	18.6	10.4	10.4	9	4.4	2.2	5.6	29.6	5	9	13.4	4	11	3.4	7.8	26.8	10.8	3.6	11.4	10.4	5.8	14.4	4	4.2	17.6	13.2
Q18	2.8	4.8	6	18.4	3.2	14.4	8.6	32	13.8	4.6	17.6	4	4	6.2	15.6	26.8	2.4	6.6	2.6	11.2	2	2.4	2.4	9.4	8.6	14
Q19	6	16.4	14	20	8.4	3.2	6.8	42.4	16.6	5	16.8	8.2	8.4	7.2	9.4	12.4	9.4	4.8	3	14.6	9.4	8.8	5	6.4	9.2	8
Q20	6.4	3.6	8.6	12.6	2.4	2.4	6.8	37.4	9.2	12.4	20.8	6.2	3.8	3.2	6.6	35.2	4.8	6.6	9	12.6	9	4	3	6.6	7	13
Q21	3.6	5.6	2.4	12.2	2.2	2.2	4.6	28.4	4.6	3.6	15.6	4.8	1.4	5.4	2.4	33	6.2	2.8	2	18.2	12.2	2.2	1.4	5.4	3	9.2
Q22	4.4	9.4	10.6	7.2	7.6	16	23	20.6	19.4	5.2	11.2	7.8	14	13.4	8	4.2	7.4	3.4	4.2	25.8	33.4	11.8	6.6	9.2	2.4	8.2
Q23	8.8	10.4	9.6	19.4	3.4	15.6	8.2	13.8	36.4	16	16	4.4	11.2	20	9.6	16.8	6.4	8.8	3.4	18.8	19.2	9.6	3.4	9.6	2.8	4.2
Q24	4.2	6	6.4	24.2	2.6	19.8	7.6	18.6	25.6	3.4	15.6	7.8	3.8	18.8	4.4	24.4	4	5.2	4.4	20.6	4.2	2.8	6.2	3.2	2	4
Q25	5.4	8.8	4.6	5.8	6.2	24.4	23	19.2	20	4.2	14	4.8	6.4	18.8	1.8	21.6	5.2	6.8	3.2	22	17.4	5.4	4	7	4.8	4.4
Q26	4	4.2	3.8	22.2	3.8	8.6	3	17.8	38	6.4	12.4	9.2	3.2	3.6	4.6	27.2	4	16.6	3.2	18	23.4	9.2	10.4	12.4	4	2.8
Q27	2	11.8	10.6	9	19.6	39	38	17.4	35.6	4	21.8	5.8	3.2	3.6	1.6	16.4	4	15.2	11.4	22.4	7	6.8	15.2	4.2	8.4	12
Q28	4.6	3.6	1.2	9.2	11.6	23	11.6	11.8	24.4	8.6	20.8	2.4	4.4	11.8	2.6	20	3.8	4.8	12.4	10.8	9	8.4	10.8	2	3.2	6
Q29	1.6	8.2	5.4	5.6	1.2	7.6	13.8	22.4	22.2	5	14.2	5.4	1.6	4	3	21.4	1.6	4.6	4.6	2.2	10	2	2.6	12	2.8	2.2
Q30	2.6	12.6	15.4	6.2	23.4	10.8	22.8	25.2	22.8	7.4	4.2	6.4	3.6	19	24.2	11.4	14.2	3.4	11	9.8	11.2	2.4	9.8	5.4	2	2
Average	7.9	8.8	8.8	11.1	10.5	11.1	11.7	21.3	19.2	8.1	14.5	6.3	7.7	11.2	9	19.9	7	8	5.6	17.8	10.4	6.7	6.5	10.8	5.6	8.8
Total	237.6	283.4	294.8	338.2	316.2	332.6	352	640.2	573	243.6	436	188.2	230	335.6	271	596	210.2	239.2	168.6	534.6	311	202.2	194	323.8	169.4	263.8

Figure 7.25: Average cost along all the countries = 10.5567. Results using the Null model. From 780 possible question-country pairs 332 have a cost less than 7 (an accuracy of 43% in total). The average cost of accepted question-country pairs is 4.17.

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	C16	C17	C18	C19	C20	C21	C22	C23	C24	C25	C26
Q1	29	41	32	32	35	38	33	34	35	32	27	32	34	36	36	33	29	27	31	37	32	32	33	39	26	35
Q2	32	35	29	31	34	38	39	38	33	31	28	32	32	41	32	37	27	23	33	33	35	36	32	40	30	29
Q3	39	45	33	33	39	37	43	39	38	37	34	33	39	46	38	39	33	38	32	36	41	36	32	41	40	41
Q4	49	62	51	55	53	48	43	57	38	36	43	59	48	64	44	44	46	44	36	56	46	44	47	78	43	49
Q5	38	57	41	46	41	49	52	51	49	25	40	54	50	61	33	45	31	34	35	46	46	35	45	50	39	40
Q6	47	55	44	46	48	34	50	53	46	38	42	57	47	60	42	47	37	40	36	52	57	45	43	48	48	50
Q7	36	53	42	32	46	35	35	40	37	32	41	49	46	46	36	49	31	33	31	44	42	35	44	34	44	48
Q8	41	61	41	41	52	47	37	45	51	40	42	49	49	48	43	50	38	34	42	43	41	43	44	51	41	47
Q9	41	48	40	25	52	35	29	33	35	31	37	47	33	45	42	39	33	35	37	32	38	47	34	41	30	44
Q10	42	46	40	27	54	33	28	35	35	31	37	47	33	40	52	33	36	39	34	37	36	35	34	30	33	46
Q11	53	51	41	44	50	43	40	41	46	38	40	49	52	52	52	33	40	50	34	41	46	41	44	43	47	46
Q12	38	52	40	23	37	49	39	38	41	33	39	55	40	44	39	34	33	36	29	37	42	34	27	46	35	40
Q13	68	9	33	30	23	7	34	7	45	37	25	39	26	51	34	50	54	50	52	82	8	41	23	46	47	54
Q14	26	42	50	41	24	22	22	45	40	26	47	58	34	28	36	59	26	27	32	53	38	39	26	37	29	33
Q15	24	57	21	26	21	21	12	39	23	8	40	32	29	33	22	54	22	26	22	22	22	25	7	36	26	21
Q16	23	23	24	25	22	4	24	38	10	6	41	33	12	7	22	57	24	25	20	36	23	14	6	26	25	24
Q17	59	60	49	28	51	6	22	40	5	52	40	56	50	34	69	54	48	47	44	39	27	50	27	37	57	58
Q18	50	33	44	35	53	37	24	40	18	39	38	51	43	47	54	53	50	42	40	24	31	38	32	42	48	58
Q19	43	60	53	46	52	42	42	53	39	44	47	59	42	56	47	51	45	32	42	44	46	38	48	43	41	55
Q20	47	25	26	32	33	6	23	39	13	39	37	44	32	7	56	56	42	54	37	21	23	31	9	27	43	43
Q21	4	22	5	39	26	3	30	39	11	22	37	29	7	24	12	54	5	6	21	20	20	22	9	24	9	23
Q22	46	59	47	39	61	53	57	50	41	37	47	61	46	59	49	44	36	44	37	55	57	38	45	47	43	53
Q23	39	54	48	47	43	41	48	52	51	37	50	53	45	58	35	42	48	28	35	50	47	33	52	45	39	45
Q24	29	43	28	37	34	32	29	41	41	35	44	49	28	35	29	39	29	14	35	43	42	26	32	32	34	30
Q25	49	49	42	29	45	45	51	45	46	39	46	58	43	41	43	45	31	43	36	38	53	33	35	36	44	46
Q26	46	60	46	49	48	47	43	58	53	27	48	62	46	65	42	48	41	40	42	53	64	36	59	50	50	48
Q27	45	50	44	27	49	49	51	46	51	40	44	51	32	34	37	42	32	32	39	49	42	31	39	36	42	43
Q28	29	45	31	25	36	36	37	42	48	41	42	44	31	33	27	40	25	9	40	38	34	29	33	37	34	32
Q29	26	43	35	25	24	30	25	37	32	31	31	32	27	29	35	38	25	15	25	35	29	25	35	36	29	24
Q30	36	46	39	31	46	35	35	34	40	31	30	51	35	48	44	35	35	28	31	34	40	36	39	40	37	34
Average	39.1	46.2	38	35	41.3	33.9	35	42.6	35.2	33.4	39.8	47.3	37.9	41.5	39.4	44.1	34.3	33.4	34.4	41.8	38.6	34.5	33.5	40.9	37.5	41.3
Total	1174	1366	1139	1049	1239	1018	1050	1278	1056	1002	1195	1418	1136	1244	1181	1322	1028	1003	1033	1253	1157	1035	1004	1228	1124	1238

Figure 7.26: Average cost along all the countries = 38.4487. Results of the optimization problem using the French-DeGroot model. From 780 possible question-country pairs 11 have a cost less than 7 (an accuracy of 1% in total). The average cost of accepted question-country pairs is 5.09.

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	C16	C17	C18	C19	C20	C21	C22	C23	C24	C25	C26
Q1	9.4	11.8	18.6	5.2	10.2	12.2	9.2	22	6	15.8	26.8	5.8	8.4	7.2	7	4.8	12.2	15.6	6.6	4.6	5.4	5.8	12.6	8	7.6	7
Q2	11.6	10.4	13	22.2	14.2	10.6	10	13.6	8.8	14	20.4	3.4	8.4	25	6.4	6.2	26.4	15.4	7.4	12.2	8	6.8	12.4	6.8	6.6	8.2
Q3	7.8	9.4	13	9.4	14.8	10.6	8	26	17.8	12.6	15.2	14.4	11	23.2	6.6	5.6	18.4	8	7.4	22.8	10.8	7.2	11.8	7	7.3	10.4
Q4	22.6	9.4	22.2	25.8	42.8	14.2	16.8	14.8	16.8	29.6	20.4	14.6	17.8	23	24.8	12.8	14.2	25.8	9.2	51.2	9.8	17.2	15	7.3	22.8	8.6
Q5	4.6	4.2	7.6	20	11.4	5.6	13.2	20.2	27.2	11.2	8.2	5.2	14	25.2	11.2	15.4	9.8	16.2	5	8.2	8.8	7.8	6.8	10	12.6	10
Q6	6.2	3.2	7.2	16	15	4.4	12.4	9.6	4.6	10	11.4	14	6.2	5.8	5.2	12.2	5	23.4	4	29.8	20	5.2	7.6	15.4	5	4.8
Q7	9.8	8.2	15	5.2	7.2	3	6	7.8	5.4	5.2	4.4	9.8	6.4	7.2	25	17.8	4.8	16.6	7	10.2	6	6.8	7.6	3.6	13.2	6.6
Q8	7.8	7.2	8	4.6	12.2	12.8	12.8	8.6	7.6	6.8	7	3.8	8.6	5.2	16.4	11.4	7.4	14.6	7.2	18.6	1.4	2.8	7.8	7.4	8.2	6.6
Q9	8.4	6.2	5.4	4.6	34.2	8.6	14.6	21.8	8.8	14.8	6.2	7.8	16.2	8	5.6	12.8	12.8	15.2	10.8	6	7.4	6	15	2.4	9.6	21.8
Q10	16.2	9.4	14	2.6	33.8	9.8	5.8	20.4	13	7	13	5.6	7.2	11.4	4.6	10.6	10.6	6.4	8.4	11.6	6	5.8	15.4	5.2	7.4	29.6
Q11	17.4	6.6	18.6	6.6	20.6	13	12.4	40.8	23.8	16.2	23	5.2	23.4	12	22.4	14.2	18.8	13.4	7.8	7	23.4	27.4	7.2	15.6	8.8	4.4
Q12	4.4	9.4	16.2	5.8	9.8	49.2	15	18.2	10.2	25.2	12.4	10.2	8.8	11	7.8	10	11.8	6.6	14.2	13	6	5.6	6.2	4.7	3	16.8
Q13	17	5.2	13.4	7.4	1.8	11.2	7.6	14	20.6	5.6	7	8.2	17.4	3.2	6.4	12.8	8.4	8	5.6	79.4	8	6	8.6	6.6	8.2	13.6
Q14	6.4	8.8	6.2	4.4	8.4	7.8	13.6	8.2	23.2	12.6	17.8	6.2	11.2	8.8	6.8	19.4	11.6	5.4	4.2	19.6	16.6	12.6	14	14.6	8.8	7.8
Q15	5.4	23.2	3.4	4.2	14	14.4	4.2	17.2	24.2	12.4	9.2	15.4	3	11.8	14	13.8	9.8	5.2	6.6	14.8	19	7.6	13.2	6.8	9	11.8
Q16	1.2	12	4.8	6.4	6.8	13.4	22.8	28.4	18	8.4	18.4	15.6	7.2	8.4	13.8	16.8	11	6.4	8.8	21	11	7.6	13.6	8.6	8.2	10.8
Q17	16.8	10	8.4	8.4	9.2	9	20.6	32.4	21.8	14	6	6	13.6	3.8	15	10.4	16.6	13	14	13.2	8.8	13.2	12.8	5.6	17.8	16
Q18	10	7.8	9.4	13	7	20.6	20.6	29.6	10.2	9	20	8.4	6.4	9	13.4	11.2	6	6	8	18.6	6.6	5.6	11	7.8	9	13.8
Q19	5.2	17.2	14.8	16.2	9.4	8.4	15.2	48.4	8.4	6.4	12.6	11	10.4	7.4	7.6	8.2	10.8	6.2	7.8	9	11.2	10	11.6	5.4	9.4	8.2
Q20	9.4	6.6	15	8.8	6.2	11.2	21.2	34.2	15	14	15.6	10.4	6.4	11.8	9.6	16.6	10.8	10	12.6	19.8	12.4	7.4	15.4	6.4	7.4	14.6
Q21	8	6.2	9.6	6.4	8	9.4	12	4.8	15.8	9.2	10	9	6.6	8	4.6	13.4	13.6	5.8	9.8	18.6	11.2	9.2	13.2	7.8	11.2	13.2
Q22	4.8	11	12	8.2	8.8	17.4	26.8	21.4	6.6	17.6	11.4	10.2	13.6	16	7.2	5.2	6	7	4.2	24.8	32.4	10.4	8.4	6.2	5.2	11.6
Q23	7.6	12.6	11.8	19	4.2	21.4	10	26.4	25.8	18.2	11	6.2	10.8	19.2	7	22.8	10.4	10.2	6.8	9.2	17.6	12.6	5	10.6	4.2	6.8
Q24	7.4	3.2	8.8	20.4	3.8	10.8	5.8	19.8	17	8	3.4	8.4	4.2	17.2	3.2	4.2	5.4	7.4	5.4	17.2	3.6	4.6	5.2	5.4	6	4.4
Q25	3.6	9	4.4	3.2	8.4	25	23.4	25.8	6.4	9.8	12.6	4.4	7.8	17.2	6.6	7.8	3.8	7.2	3.8	22.8	19.6	6	5.6	5.4	5	7
Q26	6.2	7.2	7.8	21.6	8.2	6	9.4	22	30.2	7.6	10	6.6	10.4	2.4	3.4	25.4	8.8	14	6.8	23.8	20.8	14.8	12.6	16.6	5.6	5
Q27	5.6	6.8	16.4	3	18.6	43.2	32.6	7.2	30.2	7.6	8.8	12	6.4	8.4	16	8	3.6	14	10.2	16.6	10.4	4.6	13.2	3.8	6.2	18
Q28	6.6	3.2	7.2	2.6	7.2	23.8	11	6	23.2	12	7.8	6.4	6.8	8.4	7.4	7	5.8	7	11	7.2	12.4	3.4	9.2	4.6	3.8	6.6
Q29	4.2	8.4	6	2.4	1.8	10.6	11.2	14.4	15.2	12.4	12.2	3	2.8	4.8	6.8	10.8	2.6	6	2.6	7.2	5.6	6.2	10.2	12.4	4	4.2
Q30	14.4	17	25.6	7.6	33.4	22.8	28.4	19.6	20.4	11.8	16.8	9.4	9	20.8	30.8	6.2	18.2	6	14.8	17.4	10.6	7.4	8.8	12	9.6	13.4
Average	8.9	9	11.5	9.2	13.3	14.8	14.4	20.7	16.1	11.9	12.5	8.7	9.5	12	10.5	11.7	10.5	10.7	7.9	18.4	11.7	8.4	10.8	11.7	7.7	11
Total	266	270.8	343.8	277	399.4	444	431.4	619.6	484.4	357.4	375.6	260.4	286.4	359.6	315.4	351.6	314.4	322	236.6	551.4	350.8	253	322.6	350	231.8	330.8

Figure 7.27: Average cost along all the countries = 11.6746. Results of the *Constrained* optimization problem using the Classification-Based model. From 780 possible question-country pairs 220 have a cost less than 7 (an accuracy of 28% in total). The average cost of accepted question-country pairs is 5.16.

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	C16	C17	C18	C19	C20	C21	C22	C23	C24	C25	C26
Q1	1.8	8	2.4	1.2	1.8	2.2	2.4	6	7.2	6.2	2.6	1	1.4	4.8	5.4	9.4	4	2.6	1.6	10.6	4.2	1.8	1.2	5.6	2.4	5.4
Q2	5	4.2	1	1.4	8.8	7.4	2.6	14	4	3.6	3.8	3.6	1	13.8	2.4	8.6	9	1.4	2.8	6.4	3.2	3	1.6	3	4	3.8
Q3	3	6	2.2	2.2	7	3.2	3.2	13.4	11	5	4.8	2	10	18	2	1.4	9.4	1.8	1.4	9.4	2	2.2	2.8	4.4	5.4	8.6
Q4	15.6	12.2	18.6	19	29.4	9.8	10.8	25	3	15.6	9.6	9.8	10.8	24.4	17.8	5.6	6.2	17.2	3	3.8	2.4	15.8	5.4	5.2	14.2	8
Q5	4	6.6	8.4	13.2	4.6	4.8	8.2	10.4	18.4	3.2	9.2	3.2	7.6	23.6	1.6	11.2	0.8	8.8	0.4	12	1.6	4	3.6	5.6	6.6	5
Q6	5	7.8	6.4	5.8	13.8	4	4	14.6	7.6	4	8	11.8	3.8	10.2	3.6	13	1.8	14.8	0.4	23.8	11.8	3.2	1.8	15.8	4.6	7.4
Q7	7.2	7	11.6	1.6	4.8	2.2	3.8	3.2	8.4	5.2	11.8	5.4	1	11.8	7.8	25.2	5	4	1.6	6.2	0.6	6.8	2.8	3.8	4	8
Q8	10.6	8.6	1.8	1	11.2	5.4	7.2	5.2	4.6	5.2	12.6	1	3	9.2	4	21.8	5.2	3.6	5	14.2	0.4	3.2	1.6	8	5	4.6
Q9	13	6.6	3.2	2.2	20.6	1.8	9	6.6	13.4	15.2	11.2	3.6	15.4	8.8	8.4	3.8	7	15.6	6.6	10	5.6	4.2	9.8	7.2	6.6	14
Q10	17	9.8	7.2	1	20	4.4	5.2	8.4	11.8	11.8	10	5.2	9.2	7.4	6.8	5	5.4	11.2	3.2	9.4	3	3.2	10.8	8.6	1.8	18
Q11	18.2	4.2	6.2	2.8	12.8	9.6	3.2	10.4	12	13.6	10.4	1.2	19.8	10.2	17.6	2.6	14.2	12.2	1	8.4	13.8	17.4	5.8	7.4	4.6	3.8
Q12	13.4	9.6	6	2.4	2	31.8	6.4	11.6	3.8	19.8	16.2	6.4	13.8	9.6	4.8	0.6	8.6	9.2	11.6	8.4	5	2.8	1	32.2	6	13
Q13	13.8	3	6.8	1.6	1.2	1.4	1.2	8	4.2	2.6	15	1.6	12	3.4	5.2	17.4	3.6	4.8	2.2	64.8	1	3.6	2.8	3.2	6.2	9.6
Q14	1.8	7.4	7.4	3.8	6.8	0.2	6.4	8	2.2	1	23.6	7	6.4	3.6	6	35.2	3.6	5.4	2.6	25	12	10	2.8	10.2	1	5.4
Q15	2.8	18.4	2.6	4.2	4.4	0.8	12.2	17.4	0.6	2	17.6	6	0.4	3	7.6	26	2.2	4.4	0.8	17	8.4	2.6	1.6	8.6	1.8	3.4
Q16	0.6	4.8	3.8	4.4	2	0.4	5	14.2	5.6	0.6	15	4.4	2.6	1.2	6	35.4	3.2	1.6	0.4	16.4	2.2	3	0	4.6	0.6	4.8
Q17	13.8	10.2	5.8	6.6	3.2	1.2	2	17.6	2.2	7.6	9.2	1.8	8.6	3	7.8	22.4	7.4	3.4	9.2	10	4	10.6	3.8	1	14.4	12.6
Q18	1.6	3.6	3.8	16.2	2	10	4	21.2	10.4	2.6	14	3.6	1.6	4.8	11.8	22.4	2.4	6.4	1.8	8.8	1.6	1.2	2.2	5.6	5.2	11.2
Q19	4.6	14.4	9.2	16.8	4.2	2.4	4	26.6	13.8	4.6	11.6	7.2	5.2	6.6	5.2	9	6.4	4.6	2.6	14.4	7	4.4	4.4	2.8	7.2	7.4
Q20	5.4	2.4	4.6	10.8	2.4	1.8	4.8	20.6	6	8.6	12.2	3.6	3	2.2	5	32.2	3.2	5.2	6.6	9.4	4	2.4	1.8	4.2	5.8	10
Q21	1.4	1.8	0	9	1.8	0	1.8	15.6	0.8	4.2	11	2.6	1.2	4	1.4	28.2	0.6	2.4	0.2	12.4	2.4	0.8	0.6	1.8	2.8	4.6
Q22	4.2	9.2	8.6	3.2	7.2	11	19.2	11.8	13.8	4.8	10.8	5.2	10.6	12.2	5	0.6	2.2	2.6	1	24.6	26.6	10	4.4	7.8	2	8
Q23	5.6	10.2	7.2	13	1.6	14.2	3.4	6.2	18.6	11.8	16	1	5.8	15.8	5	11.4	5	8.6	1.6	18.6	13.6	6.2	1.6	7.2	1.6	3.4
Q24	3.8	5.8	3.8	17.8	2.4	12.8	2.2	12	20.6	3	13.2	5	3.2	17.2	2.6	15.4	0.8	5.2	4.2	17.6	3	2.8	4.6	3	1	3.2
Q25	4.8	8.8	1.2	1.8	5.6	20.2	16.4	10.6	8	3.8	12.4	2.2	3.8	16.6	1.6	18.4	1.4	2.8	0.8	19.2	13	4.2	3.4	4.6	3.4	3
Q26	3	3.8	1.8	14.8	3	1.6	1	6	25.2	4	11.8	5.2	2.4	3.2	1.6	17.6	0.8	7.2	2.6	17.4	18.8	6	7.4	10	3	2.4
Q27	2	9	7.2	6.2	16.6	35.2	23.4	8.4	26.2	3.8	20.8	2.8	3	15.8	0.6	12.6	1.4	14	5.4	19	4	6.4	12.6	2.8	6.2	8.2
Q28	4.4	3	0.4	6.2	9.2	19.6	9.2	3.8	15.2	8.4	17.4	1.4	4.2	10.8	1.8	14.2	1.2	4.8	9.4	10.4	7.2	6.4	8.2	1.6	2.2	4
Q29	1	8	4.2	1.6	1.2	6	9.8	7.6	16.4	4.8	12.6	3	1.4	3.8	2	17.6	0.2	4	1.4	9.4	0.8	2.4	9.6	2.4	1	0.4
Q30	2.2	10.4	11.8	3	21	7.4	20	3	16.6	7.2	3.4	5.8	3.4	18	20.6	8	11.2	3	8.4	9.4	7.8	1.6	3.8	0.8	1.8	1.8
Average	6.4	7.5	5.5	6.5	7.8	7.8	7.1	11.6	10.4	6.5	11.9	4.1	5.9	9.9	6	15.1	4.4	6.4	3.3	16	6.4	5.1	4.2	8	4.4	6.8
Total	190.6	224.8	165.2	194.8	232.6	232.8	212	347.4	311.6	193.8	357.8	124.2	176.4	297	179	452.2	133.4	192.8	99.8	480.6	191	152.2	126.8	238.8	131.4	203

Figure 7.28: Average cost along all the countries = 7.4897. Results of the *Free* optimization problem using the *Friedkin-Johnsen* model. From 780 possible question-country pairs 460 have a cost less than 7 (an accuracy of 59% in total). The average cost of accepted question-country pairs is 3.3.

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	C16	C17	C18	C19	C20	C21	C22	C23	C24	C25	C26
Q1	2.6	8.2	6.6	3	3.6	3.6	7	9.2	9	7.6	3.2	1.6	2.8	6	7.4	10.8	4.6	4	2.8	11.8	9.4	2.2	5.2	7.6	4	5.6
Q2	6.2	4.8	7.2	5.6	10.4	10.2	7.2	15.6	7.6	4.4	5.2	4.6	3.4	13	4.2	1.4	10.8	2	4.6	7.2	4.4	3.4	3.8	3.2	4.8	5.6
Q3	4	7.6	8.4	7.2	9.6	6	5.6	15.4	19.2	5.8	4.4	3.4	3.4	21.2	3.2	1.4	10.2	2.4	4.6	7.2	5	3.4	3.8	5.4	5.8	10.6
Q4	19.2	1.3	23.4	25	36.8	8.8	15	42.2	8	19.4	10.4	12.4	12.8	26.2	28.4	7.2	8.8	21.8	4.4	39	6	17.2	10.2	68.6	15.8	8.6
Q5	5.6	7.2	11.2	19.4	7.2	8.2	11.6	14	31.4	4.2	11	6.4	9.8	25.2	5.6	20.6	1.6	11.8	1.8	12.4	4.6	4.8	5.2	5.6	7	5.4
Q6	5.6	8.4	10	14	17	6	15.2	23.2	11.8	6.8	11.4	16.6	5.6	12	6.8	19.8	2.2	18.2	2.2	26	18	3.8	7.6	19.4	5.6	7.8
Q7	10.8	8.2	18.4	2.6	5.6	4.2	4.8	6.8	16.6	5.6	13.8	8.2	2.8	13.6	11	29.2	5.8	5.4	3	7	1.2	9.6	5.2	4.4	4	9.2
Q8	12.6	10	5	2.6	12.6	10	9.6	10.4	6.4	17	1.8	1.8	4	10.8	8.4	27.8	7.6	7.4	9	17.4	3	4.6	4.4	8.4	5.6	5
Q9	14	7	4.6	4.6	29.6	2.2	11.8	8.2	17.8	17.8	16.4	3.8	17.2	10.2	11.4	7.4	17.4	10.4	10.2	10.2	8.2	6.2	12.2	8.8	8.6	20.4
Q10	21.8	12.8	11.4	3.2	33	5.8	7.8	11	23.6	13.8	13	8.6	10.6	8.6	9.2	10.6	8.8	13	6.4	9.8	8.6	4.6	12.8	11.2	3.8	27
Q11	21.8	4.2	9.6	4.4	20.6	13	11.6	14.4	15.6	18.2	13.4	3.6	23.4	13.6	23.6	9	16.4	14.2	5.4	9	17.2	23	12.8	9.2	7.4	4.2
Q12	17	10.2	9.2	5.8	2.6	44.8	11.2	23.6	14.2	26	20.6	10.2	17	11.4	8	3.2	14.4	9.6	16.4	10.6	8	3.6	2.6	4.4	7.4	16.6
Q13	15.2	3.4	9.8	5.6	4	3.8	2	18.2	5.8	3.4	19.4	2	15	3.4	10	21.4	5.2	6.4	5	73.4	2	4.8	5	4.6	8.2	14.2
Q14	2.6	10.4	13.4	12.6	9.4	1.4	12.2	15.8	9.2	2	26	10	7.8	4.6	9.2	38.4	8.4	6.2	5.8	27.2	18.8	13.6	3.6	1.4	2.2	7
Q15	4.6	24.8	4.8	5.6	10.2	2.6	12.2	12.2	2.2	3	20.6	8.2	1.8	5.2	12.2	30	7.2	6.4	2.2	18.8	15.2	2.6	2.6	11	3.2	9.2
Q16	2.4	8.4	5.4	6.8	3.6	1.8	9.6	29	9.4	3.4	16.2	7.2	3	2.8	13	36.4	4.8	3	2.2	21.6	7.6	4	1.2	8.6	1.2	7.6
Q17	18.6	10.6	12.2	10.8	5.2	2.6	4.6	28.4	4.4	9	13.4	4	11.2	4.4	11	25.4	9	5.2	13.6	10.8	5.6	15.8	4.4	3	18.8	13.4
Q18	2.8	4.8	8	18.2	4.6	12.4	8.2	32.2	14	4.6	18	4.6	4	6	15.8	28.8	4.6	7.4	3.4	11.4	3	2.6	4.6	9.8	8.2	14.4
Q19	6	16.2	14	21	7.6	5.2	9	43	16	5.2	16.8	9.2	8.4	8.4	11.2	11	8.2	5.4	3.8	15	12.4	9	8.8	5.8	9.2	8.6
Q20	6.6	4.2	7.2	13.6	3.4	2.2	7	34.4	8.8	12.4	20.6	6.2	3.6	3.8	6.4	34.6	4.4	7.8	7.8	12.6	9.4	4	2.4	7.2	7.6	13.2
Q21	3.8	5.6	1.6	13.4	2.8	1.2	6	25.2	2.6	3.6	16.8	4	1.4	5.4	2.6	32.8	3.8	2.8	2	18.8	10.6	2.2	1.8	4.6	3.2	9.4
Q22	4.4	9.4	12	6.4	9.6	13	26.6	20.4	16.2	5.2	12	8.8	14.4	13	8	3.2	5.8	3.2	2.8	25.8	32.6	11.8	9.8	9	2.4	8.8
Q23	8.6	10.4	7	22.4	3.2	17.4	9.2	6.6	37.6	15.8	16.4	3	11	19.4	8.6	19.2	7	9	3.8	19.2	20.8	9.4	5.4	9.8	2.2	4.4
Q24	4.2	6	6.8	26	3	15.6	4.6	18.6	23	3.6	15.8	7.8	4	18.8	5	23.6	4.6	6.8	1.6	22.6	17.8	5.4	4.4	6.8	5.8	4.4
Q25	5.6	8.8	1.8	2.2	6	23.8	25.6	19.6	21	4.4	15.6	5.2	7	20.4	3.8	20.6	2.8	6.8	1.6	9.6	17.8	5.4	4.4	6.8	5.8	4.4
Q26	4	4.2	4	21	3.6	3.8	4.8	7.8	35.2	7	15.4	9.8	2.8	3.8	5.4	27.6	1.8	15.8	4.8	18	22	9.2	13.4	12	4.2	3.2
Q27	2.8	11.8	11.4	9.8	20.2	39.6	35.8	14	35.6	4.6	23	5.6	5.4	17.8	2.2	16.6	2.4	15.2	9.6	22.8	9.6	9.2	13.8	3.6	8.4	12
Q28	4.8	3.4	3	7.8	12	23.4	12.6	5.4	51.4	8.6	23	2.4	6	11.8	3.2	18.4	3	4.8	13.2	10.4	8.8	8.4	9.6	2.6	3	6
Q29	1.6	8.2	6.2	5	1.4	7.2	12.8	16.6	26.2	5.4	14.8	5.4	4.4	4	3	21.6	1.4	4.6	2.4	10.2	4	1.8	11.4	2.8	2	2.2
Q30	2.6	12.6	14.2	4.8	23.6	11.2	24	15.4	22.4	7.6	5	6.8	5.6	19.2	25.2	12.4	13.6	3.6	9.4	10	11	2.8	10.4	4.6	2.4	2
Average	8.1	8.8	8.8	10.3	10.7	10.4	11.5	19.1	16.5	8.2	15	6.4	8	11.6	9.4	19.5	6.6	8.2	5.6	18	10.3	6.8	6.9	10.6	5.8	9
Total	242.4	284.8	283.8	310.4	322	311	345.2	572.2	486.2	246.8	450	191.4	239	347	283.2	584.6	136.6	246	167.6	541.4	309.2	204.8	205.8	316.6	174	289.6

Figure 7.29: Average cost along all the countries = 10.3918. Results of the *Constrained* optimization problem using the Friedkin-Johnsen model. From 780 possible question-country pairs 330 have a cost less than 7 (an accuracy of 42% in total). The average cost of accepted question-country pairs is 4.1.

8

CONCLUSION

Following the self-contained nature of this dissertation's different parts, the conclusion chapter will present a brief overview of the results achieved in each part and their value and implications for society. We refer the reader to the specific chapters for more detailed, particular conclusions.

In the previous six chapters, we have studied and analysed different types of interconnected systems from a variety of perspectives and with various objectives. We provided topology-independent necessary conditions and sufficient conditions for the robust stability of LTI interconnected MIMO systems in Part 1, in the frequency (Chapter 3) and in the time (Chapter 4) domain. These conditions do not require complete information about the network topology and, in most cases, can be checked locally. They provide stability guarantees even when the interconnected systems have additive or dynamical uncertainties. In addition to stability certificates, a sufficient condition for convergence in a bounded time interval was also presented. Because of their generality, the provided conditions can be applied in various fields and contexts. Some examples were detailed in the corresponding chapters (a suspension bridge in Chapter 3 and a compartmental tumour system in Chapter 4), and others were mentioned.

Part 1 of the dissertation was rooted at the core of classical control theory. The studied systems were represented by transfer function matrices in the frequency domain and by ordinary differential equations in the state space in the time domain. The main goal was to find robust asymptotic stability conditions, which is a fundamental property of interest for dynamical systems. The tools used to prove the results are also classical in the context of system analysis.

On the one hand, the topology independence and robust properties of the provided conditions make them very general. However, on the other hand, their theoretical nature limits their immediate applications to real-world problems. Although very general and open from a theoretical point of view, some of the assumptions we make may be too constraining or difficult to guarantee in a real application. However, this is the case with most theoretical results, and transforming them into applicable tools is a challenge on its own.

For this reason, the actual value of the results and analysis presented in Part 1 of the dissertation is that of compiling and slightly advancing the tools and knowledge in the field of stability of interconnected LTI systems. The information, views, techniques, and results presented in this part are not intended to be applied directly to real applications but to be built upon.

There are many relevant directions to go from Part 1. These include exploring the case where both dynamical and additive uncertainties are present. Trying to prove similar results when the subsystems are slightly nonlinear, for instance, with Lur'e or linearised systems. Pursuing different

objectives besides asymptotic stability, such as coordination or synchronisation. Doing a similar analysis for discrete-time systems.

In many ways, Part 2 is very contrasting with Part 1. Part 2 studies opinion formation, which is a completely applied and ‘down to earth’ subject: everyone has an opinion, and we hear about opinions all the time. The subsystems that make the interaction network now are agents or people, ‘thinking’, ‘forming’, and ‘exchanging’ opinions. These subsystems are far from simple or well-defined as a LTI MIMO system is. In many ways, one could argue that we don’t know what opinions are and whether modelling opinion formation mathematically makes any sense.

Because of their very nature, large-scale population opinions have a more complex behaviour than interconnected LTI MIMO subsystems we have considered in Part 1. The latter can be unstable, marginally stable, and asymptotically stable, whereas opinions (represented by real numbers) can have all sorts of chaotic behaviours. Because of this, asymptotic stability is no longer of interest in Part 2. Instead, we are interested in understanding and analysing agent-based opinion formation models.

To do that, we first need to develop concepts and a complete framework to meaningfully and systematically evaluate opinion formation models. This includes introducing new definitions, asking *relevant* questions to direct our attention towards, defining qualitative behaviours to simplify and streamline the analysis of multiple opinion distributions, and defining different network topologies. The context and foundations required to build the framework are presented in Chapter 5.

Continuing from the bases in Chapter 5, we introduce a complete methodology to analyse agent-based opinion formation models in Chapter 6. This methodology comprises four techniques: the *Histogram-based Sorting Algorithm*, *Transition Tables*, *Agreement Plot*, and *Probabilistic Analysis*. Collectively, these techniques use numerical simulations to investigate (i) the intrinsic properties and patterns in opinion formation models; (ii) how these models change depending on agent parameters, initial opinions, and underlying interaction networks; and (iii) the potential the models have to mimic opinion transitions observed in real life.

We exemplified the techniques presented in Chapter 6 with several opinion formation models found in the literature: the French-DeGroot, Weighted-Median, Bounded Confidence, Quantum Game, Friedkin-Johnsen, and Backfire Effect and Biased Assimilation models. The results demonstrated the application of the proposed framework and provided an interesting insight into the behaviour of these models and real-life opinions.

To study real-life opinions, we used data from the World Values Survey (the WVS is a research project that explores people’s values and beliefs by conducting global surveys approximately every five years). Using the available data, we identified five qualitative categories to which any opinion distribution belongs. From this data, we could see that real-life transitions between all qualitative categories are possible and that the proportions of opinion distributions belonging to each category remain approximately constant.

In Chapter 7, we proposed the Classification-based model, an opinion formation model with two novel characteristics: first, it drops the assumption that agents have perfect access to the opinions their neighbours are communicating (whether or not those are their real opinions) and instead introduces a classification-based mechanism to represent how agents may interpret and process the opinions of others. Second, it combines three psychological traits well-known and studied in opinion formation literature: conformism, radicalism, and stubbornness. By assigning different weights to these traits, the model can produce agents that have diverse individual behaviours when forming their own opinion. Combined, these two characteristics embed the Classification-based model with a flexible and rich capacity for creating opinion transitions seen in real-life. After introducing the proposed model, Chapter 7 uses the framework presented in Chapter 6 to analyse the model’s intrinsic characteristics and behaviour.

By its very nature, the study of opinion formation processes and models has a much more direct impact on society than the theoretical work in Part 1. Regarding the ethical implications related to the study of this subject, we believe that the study of agent-based opinion formation models with the objective of *understanding* the processes behind opinion formation is valuable because it provides a complementary approach to that of sociology and psychology. In addition to this, it requires the analysis of complex interconnected systems and the techniques developed with this purpose can potentially be used in other applications, for instance, epidemic modelling. In this sense, opinion formation research is valuable and can positively impact society.

However, when the study of these models intends to *control* or *influence* populations' opinions, we believe that is unethical and, even with 'good intentions', should not be done. This belief is not restricted to the cases when the opinions are influenced by using systems and control approaches. Large-scale opinion manipulation by exploiting opinion formation mechanisms, such as biases and conformist traits, is unethical independent of whether it is done from a systems and control perspective, fake news, or targeted advertisement. People should be able to make their own judgements in an unbiased environment free of tactics or control laws that exploit our intrinsic opinion formation mechanism, bypassing our judgement.

We are aware that already for a long time internal biases and psychological mechanisms have been used to steer public opinion, whether it is to buy something, to act in a certain way, or to support a given cause. Also, one could argue that every time someone presents information in a certain way to convince a given individual of something, it is probably using information on the individual to convince her/him of something. There are many ambiguous cases in which it is not clear whether biases and psychological mechanisms are exploited, or information is simply arranged in a smart way to convince someone. Nevertheless, there are cases in which psychological traits, and biases in particular, are clearly exploited in an attempt to systematically control people's opinion beyond their judgment at large-scale. Those cases are unethical.

With this in mind, all the techniques presented in Part 2 of the dissertation are with the intent of analysing models and are on purpose designed so that it is easy to export them to other research fields. In this sense, we encourage adapting the knowledge presented in Part 2 to other research fields that use agent-based models, especially when models are complex and have different intrinsic properties that are challenging to grasp or understand. Similarly, we deter future research from using our results to design control algorithms to steer a population's opinion to a desired state, even if the intentions are judged to be 'good' by the researchers. Please, let people be, and let them make their own opinions.

Ultimately, the value of this dissertation can be found in the individual contributions to the collective knowledge of interconnected systems, a field so varied and diverse that it is only possible to investigate local problems, as reflected in the two parts of the dissertation. In this sense, the present dissertation can be seen as a compilation of ideas, approaches, techniques, and results related to the study of interconnected systems.

ACKNOWLEDGEMENTS

Although this is my PhD dissertation, it would be incorrect to think that I did all of this on my own. On the contrary, this dissertation would not have been possible without the contribution of many people whom I would like to thank and acknowledge now.

My partner Hanneke Hoogland not only created an environment where I could thrive and be happy but also encouraged me when I felt frustrated and overwhelmed when doing research and writing papers. Every day I am thankful for being able to share my life with you, and I hope I can do that for many, many more days. My friends Siert Sebus and Nejc Maček provided almost countless opportunities to laugh about silly things that would take my mind off work. Just thinking about the list of sponsors makes me smile. And Siert, please take good care of your plant Ricardo. My friend Suad Krilašević with whom I can laugh but also talk about more serious things, sometimes work-related, sometimes while drinking tea. I wish you the best with DragonX, and I hope to beat you at connect four one day. My friends Karthik Venkatesan, Jan Pruszyński, and Irene Cantoni, I enjoy a lot playing boardgames with you; we should play Kwakzalvers van Kakelenburg one day. And thank you, Irene and Jan, for your valuable advice when I needed it. My friend Pascal de Koster, I appreciate our interesting conversations, mostly about different ways of thinking about boardgames. We should play Gizmos one day. My friends Carolina Jorge and Larisa Gomaz, thanks for listening to my complaints when the PhD was not advancing as I would have liked. I wish you the best after the completion of your PhD. And Gabriel de Albuquerque Gleizer and Giannis Delimpaltadakis for your help with the defence preparations.

On the academic side, I want to thank and acknowledge my daily supervisor and promotor, Giulia Giordano, for her continued advice and guidance on technical and theoretical aspects. I am sure it was not easy. Also, my promotor Tamas Keviczky for his help navigating the university's bureaucracy. And all the support staff for planning events in the department. Most of them were fun.

I also want to acknowledge the indirect but fundamental contribution of my former university professors, including but not limited to Diego Patiño, Julian Colorado, Andrés Vargas, and Mario Velasquez, who taught me the tools and mindset I needed to undertake this quest. And Fredy Ruiz for sending me the information about the PhD vacancy.

Last but not least, I want to thank my mother, brother and father, who, in many ways, helped me be where I am now.

*Carlos Andrés
Amsterdam, May 2023*

BIBLIOGRAPHY

- [1] R.P. Abelson. “Mathematical models in social psychology”. In: *Advances in Experimental Social Psychology*. Vol. 3. Elsevier, 1967, pp. 1–54.
- [2] R.P. Abelson. “Mathematical models of the distribution of attitudes under controversy”. In: *Contributions to mathematical psychology* (1964).
- [3] M. Afshar and M. Asadpour. “Opinion formation by informed agents”. In: *Journal of artificial societies and social simulation* 13.4 (2010).
- [4] L.J.S. Allen and P. Van Den Driessche. “Stochastic epidemic models with a backward bifurcation”. In: *Mathematical Biosciences and Engineering* 3.3 (2006), pp. 445–458.
- [5] C. Altafini. “Consensus problems on networks with antagonistic interactions”. In: *IEEE Transactions on Automatic Control* 58.4 (2013), pp. 935–946.
- [6] C. Altafini. “Dynamics of opinion forming in structurally balanced social networks”. In: *PLoS ONE* 7.6 (2012).
- [7] V. Amelkin, F. Bullo, and A. K. Singh. “Polar Opinion Dynamics in Social Networks”. In: *IEEE Transactions on Automatic Control* 62.11 (2017), pp. 5650–5665.
- [8] M.S. Andersen, S.K. Pakazad, A. Hansson, and A. Rantzer. “Robust stability analysis of sparsely interconnected uncertain systems”. In: *IEEE Transactions on Automatic Control* 59.8 (2014), pp. 2151–2156.
- [9] B. Anderson, F. Dabbene, A.V. Proskurnikov, C. Ravazzi, and M. Ye. “Dynamical Networks of Social Influence: Modern Trends and Perspectives”. In: *IFAC-PapersOnLine* 53.2 (2020), pp. 17616–17627.
- [10] B. Anderson and M. Ye. “Recent advances in the modelling and analysis of opinion dynamics on influence networks”. In: *International Journal of Automation and Computing* 16.2 (2019), pp. 129–149.
- [11] H. Andersson. “Limit theorems for a random graph epidemic model”. In: *Annals of Applied Probability* 8.4 (1998), pp. 1331–1349.
- [12] O. Angulo, F. Milner, and L. Sega. “A SIR epidemic model structured by immunological variables”. In: *Journal of Biological Systems* 21.4 (2013).
- [13] S. Arimoto. “Passivity-based control”. In: vol. 1. 2000, pp. 227–232.
- [14] S.E. Asch. “Effects of group pressure upon the modification and distortion of judgments”. In: *Documents of gestalt psychology*. University of California Press, 1961, pp. 222–236.
- [15] S.E. Asch. “Opinions and social pressure”. In: *Scientific American* 193.5 (1955), pp. 31–35.

- [16] S.E. Asch. "Studies of independence and conformity: I. A minority of one against a unanimous majority." In: *Psychological monographs: General and applied* 70.9 (1956), p. 1.
- [17] S. Banisch and E. Olbrich. "Opinion polarization by learning from social feedback". In: *The Journal of Mathematical Sociology* 43.2 (2019), pp. 76–103.
- [18] S. Banisch and H. Shamon. "Biased processing and opinion polarisation: experimental refinement of argument communication theory in the context of the energy debate". In: *Available at SSRN 3895117* (2021).
- [19] B.R. Barmish and E.I. Jury. "New tools for robustness of linear systems". In: *IEEE Transactions on Automatic Control* 39.12 (1994), pp. 2525–2525.
- [20] P. Barooah and J. P. Hespanha. "Graph Effective Resistance and Distributed Control: Spectral Properties and Applications". In: *Proceedings of the 45th IEEE Conference on Decision and Control*. 2006, pp. 3479–3485.
- [21] P. Barooah and J.P. Hespanha. "Estimation from relative measurements: Electrical analogy and large graphs". In: *IEEE Transactions on Signal Processing* 56.6 (2008), pp. 2181–2193.
- [22] F.L. Bauer and C.T. Fike. "Norms and exclusion theorems". In: *Numerische Mathematik* 2.1 (1960), pp. 137–141.
- [23] F. Baumann, P. Lorenz-Spreen, I.M. Sokolov, and M. Starnini. "Modeling echo chambers and polarization dynamics in social networks". In: *Physical Review Letters* 124.4 (2020), p. 048301.
- [24] N. Bellomo et al. "What is life? A perspective of the mathematical kinetic theory of active particles". In: *Mathematical Models and Methods in Applied Sciences* 31.09 (2021), pp. 1821–1866.
- [25] D.A. Belsley, E. Kuh, and R.E. Welsch. *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*. Wiley Series in Probability and Statistics. Wiley, 2005, pp. 100–104. ISBN: 9780471725145.
- [26] J. Bendtsen, K. Trangbaek, and J. Stoustrup. "Plug-and-play control—Modifying control systems online". In: *IEEE Transactions on Control Systems Technology* 21.1 (2011), pp. 79–93.
- [27] B. Benigni, R. Gallotti, and M. De Domenico. "Potential-driven random walks on interconnected systems". In: *Physical Review E* 104.2 (2021).
- [28] C. Beumier and M. Idrissa. "Building change detection by histogram classification". In: *2011 Seventh International Conference on Signal Image Technology & Internet-Based Systems*. IEEE. 2011, pp. 409–415.
- [29] J. Biazar. "Solution of the epidemic model by Adomian decomposition method". In: *Applied Mathematics and Computation* 173.2 (2006), pp. 1101–1106.
- [30] F. Blanchini, D. Casagrande, G. Giordano, and U. Viaro. "A bounded complementary sensitivity function ensures topology-independent stability of homogeneous dynamical networks". In: *IEEE Transactions on Automatic Control* 63.4 (2017), pp. 1140–1146.

- [31] F. Blanchini, D. Casagrande, G. Giordano, and U. Viaro. "Topology-Independent Robust Stability of Homogeneous Dynamic Networks". In: *IFAC-PapersOnLine* 50.1 (2017). 20th IFAC World Congress, pp. 1736–1741.
- [32] F. Blanchini, E. Franco, and G. Giordano. "Network-decentralized control strategies for stabilization". In: *IEEE Transactions on Automatic Control* 60.2 (2014), pp. 491–496.
- [33] Y. Bouteraa and I.B. Abdallah. "Distributed control and speed sensorless for the synchronisation of multi-robot systems". In: *Automatic Control and Computer Sciences* 50.5 (2016), pp. 306–317.
- [34] F. Brauer, C. Castillo-Chavez, and C. Castillo-Chavez. *Mathematical models in population biology and epidemiology*. Vol. 2. Springer, 2012.
- [35] F.S. Burrhus. "Two Types of Conditioned Reflex and a Pseudo Type". In: *Journal of General Psychology* 12 (1935), pp. 66–77.
- [36] Dorwin C. and Frank H. "Structural balance: a generalization of Heider's theory." In: *Psychological review* 63 5 (1956), pp. 277–93.
- [37] J. Cai, X. Song, and S.-Y. Lee. "Bayesian analysis of nonlinear structural equation models with mixed continuous, ordered and unordered categorical, and nonignorable missing data". In: *Statistics and Its Interface Volume 1* (Jan. 2008), pp. 99–114.
- [38] X. Cao and Z. Jin. "Epidemic threshold and ergodicity of an SIS model in switched networks". In: *Journal of Mathematical Analysis and Applications* 479.1 (2019), pp. 1182–1194.
- [39] X. Cao and Z. Jin. "N-intertwined SIS epidemic model with Markovian switching". In: *Stochastics and Dynamics* 19.4 (2019).
- [40] F. Ceragioli and P. Frasca. "Consensus and Disagreement: The Role of Quantized Behaviors in Opinion Dynamics". In: *SIAM Journal on Control and Optimization* 56.2 (2018), pp. 1058–1080.
- [41] K. M. D. Chan, R. Duivenvoorden, A. Flache, and M. Mandjes. "A relative approach to opinion formation". In: *The Journal of Mathematical Sociology* 0.0 (2022), pp. 1–41.
- [42] E. Chattoe-Brown. "Using agent based modelling to integrate data on attitude change". In: *Sociological Research Online* 19.1 (2014), pp. 159–174.
- [43] C. Chen, H. Zhang, Y.-H. Wu, W.-Q. Feng, and J. Zhang. "Analysis and optimization of cross-immunity epidemic model on complex networks". In: *International Journal of Modern Physics C* 26.4 (2015).
- [44] L. Chen and J. Sun. "Global stability and optimal control of an SIRS epidemic model on heterogeneous networks". In: *Physica A: Statistical Mechanics and its Applications* 410 (2014), pp. 196–204.
- [45] X. Chen, P. Tsaparas, J. Lijffijt, and T. De Bie. "Opinion dynamics with backfire effect and biased assimilation". In: *PloS one* 16.9 (2021), e0256922.

- [46] G. Cheng. “Note on some upper bounds for the condition number”. In: *J. Math. Inequalities* 8.2 (2014), p. 6.
- [47] A. Chmiel et al. “Negative emotions boost user activity at BBC forum”. In: *Physica A: Statistical Mechanics and its Applications* 390.16 (2011), pp. 2936–2944.
- [48] B. Dadashova, B. Arenas-Ramírez, J. Mira-Mcwilliams, C. González-Fernández, and F. Aparicio-Izquierdo. “Simulation-based model comparison methodology with application to road accident models”. In: *Communications in Statistics - Simulation and Computation* 46.7 (2017), pp. 5340–5366.
- [49] P. Dandekar, A. Goel, and D.T. Lee. “Biased assimilation, homophily, and the dynamics of polarization”. In: *Proceedings of the National Academy of Sciences of the United States of America* 110.15 (2013), pp. 5791–5796.
- [50] F. Darabi Sahneh and C. Scoglio. “Epidemic spread in human networks”. In: 2011, pp. 3008–3013.
- [51] G. Deffuant. “Comparing extremism propagation patterns in continuous opinion models”. In: *Journal of artificial societies and social simulation* 9.3 (2006), pp. 29–52.
- [52] G. Deffuant, F. Amblard, G. Weisbuch, and T. Faure. “How can extremism prevail? A study based on the relative agreement interaction model”. In: *Journal of artificial societies and social simulation* 5.4 (2002), pp. 1–26.
- [53] G. Deffuant, D. Neau, F. Amblard, and G. Weisbuch. “Mixing beliefs among interacting agents”. In: *Advances in Complex Systems* 03.01n04 (2000), pp. 87–98.
- [54] M.H. DeGroot. “Reaching a consensus”. In: *Journal of the American Statistical Association* 69.345 (1974), pp. 118–121.
- [55] X. Deng, Y. Deng, Q. Liu, L. Shi, and Z. Wang. “Quantum games of opinion formation based on the Marinatto-Weber quantum game scheme”. In: *EPL (Europhysics Letters)* 114.5 (2016), p. 50012.
- [56] C. Desoer and Y.-T. Wang. “On the generalized Nyquist stability criterion”. In: *IEEE Transactions on Automatic Control* 25.2 (1980), pp. 187–196.
- [57] Carlos Andres Devia and Giulia Giordano. “A framework to analyze opinion formation models”. In: *Scientific Reports* 12 (2022). ISSN: 2045-2322.
- [58] Carlos Andres Devia and Giulia Giordano. “Classification-Based Opinion Formation Model Embedding Agents’ Psychological Traits”. In: *Journal of Artificial Societies and Social Simulation* (2023).
- [59] Carlos Andres Devia and Giulia Giordano. “Graphical and Probabilistic Analysis of Agent-Based Opinion Formation Models”. Working paper. 2023.
- [60] Carlos Andres Devia and Giulia Giordano. “MIMO Networks with Heterogeneous Uncertainties: Topology-Independent Robust Stability and α -Convergence”. In: *2021 European Control Conference (ECC)*. IEEE. 2021, pp. 884–889.
- [61] Carlos Andres Devia and Giulia Giordano. “Topology-independent robust stability conditions for uncertain MIMO networks”. In: *IEEE Control Systems Letters* 5.1 (2021), pp. 325–330.

- [62] Carlos Andres Devia and Giulia Giordano. "Topology-Independent Robust Stability for Networks of Homogeneous MIMO Systems". In: *IFAC-PapersOnLine* 53.2 (2020). 21st IFAC World Congress, pp. 3379–3384.
- [63] T. Dingzhong, Q. Ying, and W. Qiming. "Research on time synchronization in cluster robots based on wireless network". In: 2010, pp. 1807–1810.
- [64] B.M. Duarte, J.W. Fowler, K. Knutson, E. Gel, and D. Shunk. "A compact abstraction of manufacturing nodes in a supply network". In: *International Journal of Simulation and Process Modelling* 3.3 (2007), pp. 115–126.
- [65] P. Duggins. "A Psychologically-Motivated Model of Opinion Change with Applications to American Politics". In: *Journal of Artificial Societies and Social Simulation* 20.1 (2017), p. 13. ISSN: 1460-7425.
- [66] N.I. Eisenberger, M.D. Lieberman, and K.D. Williams. "Does rejection hurt? An fMRI study of social exclusion". In: *Science* 302.5643 (2003), pp. 290–292.
- [67] Y. Ekalo and M. Vukobratović. "Stabilization of robot motion and contact force interaction for third-order motor dynamics". In: *Journal of Intelligent & Robotic Systems* 10.3 (1994), pp. 257–282.
- [68] A.S. Elgazzar. "Applications of small-world networks to some socio-economic systems". In: *Physica A: Statistical Mechanics and its Applications* 324.1 (2003). Proceedings of the International Econophysics Conference, pp. 402–407. ISSN: 0378-4371.
- [69] J.M. Epstein. "Why model?" In: *Journal of artificial societies and social simulation* 11.4 (2008).
- [70] E. Estrada. "Rethinking structural balance in signed social networks". In: *Discrete Applied Mathematics* 268 (2019), pp. 70–90.
- [71] E. Estrada and M. Benzi. "Walk-based measure of balance in signed networks: Detecting lack of balance in social networks". In: *Physical Review E* 90.4 (2014), p. 042802.
- [72] M.-C. Fan, H.-T. Zhang, and M. Wang. "Bipartite flocking for multi-agent systems". In: *Communications in Nonlinear Science and Numerical Simulation* 19.9 (2014), pp. 3313–3322.
- [73] L. Festinger. *A theory of cognitive dissonance*. Stanford Univ Press, 1957.
- [74] M. Fiacchini and I.-C. Morarescu. "Convex conditions on decentralized control for graph topology preservation". In: *IEEE Transactions on Automatic Control* 59.6 (2014), pp. 1640–1645.
- [75] A. Flache et al. "Models of social influence: Towards the next frontiers". In: *Journal of Artificial Societies and Social Simulation* 20.4 (2017).
- [76] J.E. Franke and A.-A. Yakubu. "Discrete-time SIS epidemic model in a seasonal environment". In: *SIAM Journal on Applied Mathematics* 66.5 (2006), pp. 1563–1587.
- [77] J. R. P. French Jr. "A formal theory of social power". In: *Psychological review* 63.3 (1956), pp. 181–194.

- [78] N. Friedkin. "A formal theory of social power". In: *Journal of Mathematical Sociology* (1986).
- [79] N. Friedkin and E. Johnsen. "Social Influence Networks and Opinion Change". In: *Advances in Group Processes* 16 (Jan. 1999).
- [80] N.E. Friedkin and F. Bullo. "How truth wins in opinion dynamics along issue sequences". In: *Proceedings of the National Academy of Sciences* 114.43 (2017), pp. 11380–11385.
- [81] N.E. Friedkin and E.C. Johnsen. *Social Influence Network Theory: A Sociological Examination of Small Group Dynamics*. Structural Analysis in the Social Sciences. Cambridge University Press, 2011. ISBN: 9781107002463.
- [82] N.E. Friedkin, A.V. Proskurnikov, W. Mei, and F. Bullo. "Mathematical structures in group decision-making on resource allocation distributions". In: *Scientific reports* 9.1 (2019), pp. 1–9.
- [83] G. Fu and W. Zhang. "Opinion formation and bi-polarization with biased assimilation and homophily". In: *Physica A: Statistical Mechanics and its Applications* 444 (2016), pp. 700–712.
- [84] A. Gaget, J.-M. Montanier, and R. Doursat. "Branched Structure Formation in a Decentralized Flock of Wheeled Robots". In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 12421 LNCS (2020), pp. 42–54.
- [85] Z. Gajic and H. Khalil. "Multimodel strategies under random disturbances and imperfect partial observations". In: *Automatica* 22.1 (1986), pp. 121–125.
- [86] Y. Gandica, M. del Castillo-Mussot, G.J. Vázquez, and S. Rojas. "Continuous opinion model in small-world directed networks". In: *Physica A: Statistical Mechanics and its Applications* 389.24 (2010), pp. 5864–5870.
- [87] A. Gattami and R. Murray. "A frequency domain condition for stability of interconnected MIMO systems". In: *Proceedings of the 2004 American control conference*. Vol. 4. IEEE. 2004, pp. 3723–3728.
- [88] E. Genc, N. Duffie, and G. Reinhart. "Event-based Supply Chain Early Warning System for an adaptive production control". In: vol. 19. C. 2014, pp. 39–44.
- [89] S. A. Gershgorin. "Über die abgrenzung der eigenwerte einer matrix". In: *kad. Nauk SSSR* 6 (1931), pp. 749–754.
- [90] D. Geschke, J. Lorenz, and P. Holtz. "The triple-filter bubble: Using agent-based modelling to test a meta-theoretical framework for the emergence of filter bubbles and echo chambers". In: *British Journal of Social Psychology* 58.1 (2019), pp. 129–149.
- [91] G. Giordano, F. Blanchini, E. Franco, V. Mardanlou, and P.L. Montessoro. "The smallest eigenvalue of the generalized Laplacian matrix, with application to network-decentralized estimation for homogeneous systems". In: *IEEE Transactions on Network Science and Engineering* 3.4 (2016), pp. 312–324.

- [92] G. Giordano, A. Rantzer, and V.D. Jonsson. "A convex optimization approach to cancer treatment to address tumor heterogeneity and imperfect drug penetration in physiological compartments". In: *2016 IEEE 55th Conference on Decision and Control (CDC)*. IEEE, 2016, pp. 2494–2500.
- [93] J. Graham et al. "Moral Foundations Theory: The Pragmatic Validity of Moral Pluralism". In: *Advances in Experimental Social Psychology* 47 (2013), pp. 55–130.
- [94] M. Granovetter. "The strength of weak ties". In: *American journal of sociology* 78.6 (1973), pp. 1360–1380.
- [95] M. Granovetter. "Threshold Models of Collective Behavior". In: *American Journal of Sociology* 83 (1978), pp. 1420–1443.
- [96] M. Granovetter and S. Soong. "Threshold models of interpersonal effects in consumer demand". In: *Journal of Economic Behavior & Organization* 7.1 (1986), pp. 83–99. ISSN: 0167-2681.
- [97] M. Gravenotter. "The strength of weak ties: A network theory revisited". In: *Sociological Theory* 1 (1983), pp. 201–233.
- [98] R.M. Gray. "Toeplitz and circulant matrices: A review". In: *Foundations and Trends® in Communications and Information Theory* 2.3 (2006), pp. 155–239.
- [99] C. Haerpfer et al. *World Values Survey: Round Five - Country-Pooled Datafile*. Available at <https://www.worldvaluessurvey.org/WVSDocumentationWV5.jsp> (2021/12/15). 2010.
- [100] C. Haerpfer et al. *World Values Survey: Round Seven - Country-Pooled Datafile*. Available at <https://www.worldvaluessurvey.org/WVSDocumentationWV7.jsp> (2021/12/15). 2020.
- [101] C. Haerpfer et al. *World Values Survey: Round Six - Country-Pooled Datafile*. Available at <https://www.worldvaluessurvey.org/WVSDocumentationWV6.jsp> (2021/12/15). 2015.
- [102] S. Hara, T. Hayakawa, and H. Sugata. "LTI systems with generalized frequency variables: A unified framework for homogeneous multi-agent dynamical systems". In: *SICE Journal of Control, Measurement, and System Integration* 2.5 (2009), pp. 299–306.
- [103] S. Hara, T. Iwasaki, and Y. Hori. "Robust stability analysis for LTI systems with generalized frequency variables and its application to gene regulatory networks". In: *Automatica* 105 (2019), pp. 96–106.
- [104] S. Hara, H. Tanaka, and T. Iwasaki. "Stability analysis of systems with generalized frequency variables". In: *IEEE Transactions on Automatic Control* 59.2 (2013), pp. 313–326.
- [105] F. Harary. "A criterion for unanimity in French's theory of social power". In: *Studies in social power* (1959).
- [106] F. Harary, A. Harary, R.Z. Norman, D. Cartwright, and K. Esau. *Structural Models: An Introduction to the Theory of Directed Graphs*. vol. 82. Wiley, 1965. ISBN: 9780471351306.

- [107] S. Harat, T. Hayakawa, and H. Sugatata. "Stability analysis of linear systems with generalized frequency variables and its applications to formation control". In: *2007 46th IEEE Conference on Decision and Control*. IEEE. 2007, pp. 1459–1466.
- [108] F. Hassanibesheli and R.V. Donner. "Network inference from the timing of events in coupled dynamical systems". In: *Chaos: An Interdisciplinary Journal of Non-linear Science* 29.8 (2019), p. 083125.
- [109] G. He, J. Liu, Y. Wu, and J. Fang. "On Bipartite Consensus of Bounded Confidence Models for Opinion Dynamics". In: *International Journal of Control, Automation and Systems* 18.2 (2020), pp. 303–312.
- [110] R. Hegselmann and U. Krause. "Opinion dynamics and bounded confidence models, analysis, and simulation". In: *Journal of artificial societies and social simulation* 5.3 (2002).
- [111] R. Hegselmann and U. Krause. "Opinion dynamics under the influence of radical groups, charismatic leaders, and other constant signals: A simple unifying model". In: *Networks and Heterogeneous Media* 10.3 (2015), pp. 477–509.
- [112] R. Hegselmann and U. Krause. "Truth and cognitive division of labour first steps towards a computer aided social epistemology". In: *Journal of artificial societies and social simulation* 9.3 (2006), pp. 1–28.
- [113] P. Herman and K. Kozłowski. "A survey of equations of motion in terms of inertial quasi-velocities for serial manipulators". In: *Archive of Applied Mechanics* 76.9-10 (2006), pp. 579–614.
- [114] David Holdaway. *Integer partition generator*. Available at <https://www.mathworks.com/matlabcentral/fileexchange/36437-integer-partition-generator>. 2012.
- [115] Y. Hori, H. Miyazako, S. Kumagai, and S. Hara. "Coordinated spatial pattern formation in biomolecular communication networks". In: *IEEE Transactions on Molecular, Biological and Multi-Scale Communications* 1.2 (2015), pp. 111–121.
- [116] C.I. Hovland, I.L. Janis, and H.H. Kelley. *Communication and persuasion; psychological studies of opinion change*. Yale University Press, 1953.
- [117] J. Ito and K. Kaneko. "Spontaneous structure formation in a network of dynamic elements". In: *Physical Review E - Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics* 67.4 (2003), p. 14.
- [118] J.A. Jacquez. *Compartmental Analysis in Biology and Medicine*. BioMedware, 1996.
- [119] I. Jacyna-Golda, M. Izdebski, and E. Szczepański. "Assessment of the method effectiveness for choosing the location of warehouses in the supply network". In: *Communications in Computer and Information Science* 640 (2016), pp. 84–97.
- [120] L.-L. Jiang, D.-Y. Hua, J.-F. Zhu, B.-H. Wang, and T. Zhou. "Opinion dynamics on directed small-world networks". In: *The European Physical Journal B* 65.2 (2008), pp. 251–255.

- [121] Z. Jiang, B. Cui, and X. Lou. “Distributed consensus estimation for diffusion systems with missing measurements over sensor networks”. In: *International Journal of Systems Science* 47.12 (2016), pp. 2753–2761.
- [122] B. Joachimi, F. Köhlinger, W. Handley, and P. Lemos. “When tension is just a fluctuation - How noisy data affect model comparison”. In: *A&A* 647 (2021), p. L5.
- [123] U.T. Jönsson and C.-Y. Kao. “A scalable robust stability criterion for systems with heterogeneous LTI components”. In: *IEEE Transactions on Automatic Control* 55.10 (2010), pp. 2219–2234.
- [124] K. Kacperski and J.A. Holyst. “Opinion formation model with strong leader and external impact: A mean field approach”. In: *Physica A: Statistical Mechanics and its Applications* 269.2 (1999), pp. 511–526.
- [125] K. Kacperski and J.A. Holyst. “Phase transitions as a persistent feature of groups with leaders in models of opinion formation”. In: *Physica A: Statistical Mechanics and its Applications* 287.3-4 (2000), pp. 631–643.
- [126] A. Kandler and A. Powell. “Generative inference for cultural evolution”. In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 373.1743 (2018), p. 20170056.
- [127] C.-Y. Kao, U. Jönsson, and H. Fujioka. “Characterization of robust stability of a class of interconnected systems”. In: *Automatica* 45.1 (2009), pp. 217–224.
- [128] R.E. Kass and A.E. Raftery. “Bayes Factors”. In: *Journal of the American Statistical Association* 90.430 (1995), pp. 773–795.
- [129] S.Z. Khong and A. Rantzer. “Scalable stability conditions for heterogeneous networks via integral quadratic constraints”. In: *2014 European Control Conference (ECC)*. IEEE. 2014, pp. 2863–2867.
- [130] S. Kikuchi, T.H. Ngoc, M. Kawanishi, and T. Narikiyo. “A synchronizing approach for non-uniform time-varying communication delays on the average consensus of multi agent system”. In: 2019.
- [131] A. Kornilov. “Investigation of the properties of the model of a branched power supply network”. In: 2021.
- [132] I.V. Kozitsin. “A general framework to link theory and empirics in opinion formation models”. In: *Scientific Reports* 12.1 (2022), p. 5543.
- [133] M.J. Krawczyk, K. Malarz, R. Korff, and K. Kułakowski. “Communication and trust in the bounded confidence model”. In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 6421 LNAI.PART 1 (2010), pp. 90–99.
- [134] M. Kuperman and D. Zanette. “Stochastic resonance in a model of opinion formation on small-world networks”. In: *The European Physical Journal B-Condensed Matter and Complex Systems* 26.3 (2002), pp. 387–391.
- [135] C.E. La Rocca, L.A. Braunstein, and F. Vazquez. “The influence of persuasion in opinion formation and polarization”. In: *EPL* 106.4 (2014).

- [136] P. Lancaster and H.K. Farahat. "Norms on direct sums and tensor products". In: *mathematics of computation* 26.118 (1972), pp. 401–414.
- [137] K.S. Larsen. "Conformity in the Asch experiment". In: *Journal of Social Psychology* 94.2 (1974), pp. 303–304.
- [138] M. Laumanns and E. Lefeber. "Robust optimal control of material flows in demand-driven supply networks". In: *Physica A: Statistical Mechanics and its Applications* 363.1 (2006), pp. 24–31.
- [139] S.-Y. Lee and N.-S. Tang. "Analysis of Nonlinear Structural Equation Models with Nonignorable Missing Covariates and Ordered Categorical Data". In: *Statistica Sinica* 16.4 (2006), pp. 1117–1141.
- [140] S. Leinhardt. *Social networks : a developing paradigm / edited by Samuel Leinhardt*. English. Academic Press New York, 1977, xxxiv, 465 p. : ISBN: 0124424503.
- [141] M. Lemon, T. Anglea, and Y. Wang. "Experimental Study of Decentralized Robot Network Coordination". In: 2019, pp. 4863–4867.
- [142] N.E. Leonard. "Multi-agent system dynamics: Bifurcation and behavior of animal groups". In: *Annual Reviews in Control* 38.2 (2014), pp. 171–183.
- [143] I. Lestas and G. Vinnicombe. "Scalable decentralized robust stability certificates for networks of interconnected heterogeneous dynamical systems". In: *IEEE transactions on automatic control* 51.10 (2006), pp. 1613–1625.
- [144] S. Lewandowsky, U.K.H. Ecker, C.M. Seifert, N. Schwarz, and J. Cook. "Misinformation and Its Correction: Continued Influence and Successful Debiasing". In: *Psychological Science in the Public Interest, Supplement* 13.3 (2012), pp. 106–131.
- [145] C.-H. Li, C.-C. Tsai, and S.-Y. Yang. "Analysis of epidemic spreading of an SIRS model in complex heterogeneous networks". In: *Communications in Nonlinear Science and Numerical Simulation* 19.4 (2014), pp. 1042–1054.
- [146] J. Li, Y. Yang, and Y. Zhou. "Global stability of an epidemic model with latent stage and vaccination". In: *Nonlinear Analysis: Real World Applications* 12.4 (2011), pp. 2163–2173.
- [147] Y. Li and A. Ansari. "A Bayesian semiparametric approach for endogeneity and heterogeneity in choice models". In: *Management Science* 60.5 (2014), pp. 1161–1179.
- [148] Y.-X. Li, Y. Kano, J.-H. Pan, and X.-Y. Song. "A criterion-based model comparison statistic for structural equation models with heterogeneous data". In: *Journal of Multivariate Analysis* 112 (2012), pp. 92–107.
- [149] J. Liang, B. Shen, H. Dong, and J. Lam. "Robust distributed state estimation for sensor networks with multiple stochastic communication delays". In: *International Journal of Systems Science* 42.9 (2011), pp. 1459–1471.
- [150] Z. Lin, L. Wang, Z. Han, and M. Fu. "A graph Laplacian approach to coordinate-free formation stabilization for directed networks". In: *IEEE Transactions on Automatic Control* 61.5 (2015), pp. 1269–1280.

- [151] C. Liu and Z.-K. Zhang. "Information spreading on dynamic social networks". In: *Communications in Nonlinear Science and Numerical Simulation* 19.4 (2014), pp. 896–904.
- [152] Q. Liu and X. Wang. "Opinion dynamics with similarity-based random neighbors". In: *Scientific Reports* 3.1 (2013), p. 2968.
- [153] Q. Liu, J. Zhao, and X. Wang. "Multi-agent model of group polarisation with biased assimilation of arguments". In: *IET Control Theory and Applications* 9.3 (2015), pp. 485–492.
- [154] J. Lorenz, M/ Neumann, and T. Schröder. "Individual attitude change and societal dynamics: Computational experiments with psychological theories." In: *Psychological Review* 128.4 (2021), p. 623.
- [155] F. Lu, M. Maggioni, and S. Tang. "Learning interaction kernels in heterogeneous systems of agents from multiple trajectories". In: *Journal of machine learning research* 22 (2021).
- [156] A.CR Martins. "Continuous opinions and discrete actions in opinion dynamics problems". In: *International Journal of Modern Physics C* 19.04 (2008), pp. 617–624.
- [157] M. Mäs and A. Flache. "Differentiation without distancing. Explaining bipolarization of opinions without negative influence". In: *PloS one* 8.11 (2013), e74516.
- [158] P. Massioni and M. Verhaegen. "Distributed control for identical dynamically coupled systems: A decomposition approach". In: *IEEE Transactions on Automatic Control* 54.1 (2009), pp. 124–135.
- [159] L. Mastroeni, P. Vellucci, and M. Naldi. "Agent-Based Models for Opinion Formation: A Bibliographic Survey". In: *IEEE Access* 7 (2019), pp. 58836–58848.
- [160] N. Masuda. "Opinion control in complex networks". In: *New Journal of Physics* 17 (2015), pp. 1–11.
- [161] D.C. Matz and W. Wood. "Cognitive dissonance in groups: the consequences of disagreement." In: *Journal of personality and social psychology* 88.1 (2005), p. 22.
- [162] G. Mckeown and N. Sheehy. "Mass media and polarisation processes in the bounded confidence model of opinion dynamics". In: *Journal of artificial societies and social simulation* 9.1 (2006), pp. 33–63.
- [163] W. Mei, F. Bullo, G. Chen, J. Hendrickx, and F. Dörfler. *Rethinking the Micro-Foundation of Opinion Dynamics: Rich Consequences of the Weighted-Median Mechanism*. 2020.
- [164] D. Meng. "Convergence analysis of directed signed networks via an M-matrix approach". In: *International Journal of Control* 91.4 (2018), pp. 827–847.
- [165] D. Meng, M. Du, and Y. Jia. "Interval Bipartite Consensus of Networked Agents Associated With Signed Digraphs". In: *IEEE Transactions on Automatic Control* 61.12 (2016), pp. 3755–3770.

- [166] D. Meng, M. Du, and Y. Wu. "Extended Structural Balance Theory and Method for Cooperative-Antagonistic Networks". In: *IEEE Transactions on Automatic Control* 65.5 (2020), pp. 2147–2154.
- [167] G. Meng and V.D. D. "Consensus with quantized relative state measurements". In: *Automatica* 49.8 (2013), pp. 2531–2537. ISSN: 0005-1098.
- [168] Y. Mengbin, Q. Yuzhen, G. Alain, A. Brian, and C. Ming. "An influence network model to study discrepancies in expressed and private opinions". In: *Automatica* 107 (2019), pp. 371–381. ISSN: 0005-1098.
- [169] A.K. Misra, A. Sharma, and J.B. Shukla. "Stability analysis and optimal control of an epidemic model with awareness programs by media". In: *BioSystems* 138 (2015), pp. 53–62.
- [170] T. Nepusz and T. Vicsek. "Controlling edge dynamics in complex networks". In: *Nature Physics* 8.7 (2012), pp. 568–573.
- [171] B. Nyhan and J. Reifler. "When corrections fail: The persistence of political misperceptions". In: *Political Behavior* 32.2 (2010), pp. 303–330.
- [172] R. Olfati-Saber, J.A. Fax, and R.M. Murray. "Consensus and cooperation in networked multi-agent systems". In: *Proceedings of the IEEE* 95.1 (2007), pp. 215–233.
- [173] K.-D. Opp. "Balance theory: Progress and stagnation of a social psychological theory". In: *Philosophy of the Social Sciences* 14.1 (1984), pp. 27–49.
- [174] Z. Pan. "Trust, influence, and convergence of behavior in social networks". In: *Mathematical Social Sciences* 60.1 (2010), pp. 69–78.
- [175] R. Pates and G. Vinnicombe. "Scalable design of heterogeneous networks". In: *IEEE Transactions on Automatic Control* 62.5 (2016), pp. 2318–2333.
- [176] R. Pates and G. Vinnicombe. "Stability certificates for networks of heterogeneous linear systems". In: *2012 IEEE 51st IEEE Conference on Decision and Control (CDC)*. IEEE. 2012, pp. 6915–6920.
- [177] N. Perra, B. Gonçalves, R. Pastor-Satorras, and A. Vespignani. "Activity driven modeling of time varying networks". In: *Scientific Reports* 2 (2012).
- [178] J.P. Pinasco, V. Semeshenko, and P. Balenzuela. "Modeling opinion dynamics: Theoretical analysis and continuous approximation". In: *Chaos, Solitons and Fractals* 98 (2017), pp. 210–215.
- [179] A.V. Proskurnikov and R. Tempo. "A tutorial on modeling and analysis of dynamic social networks. Part I". In: *Annual Reviews in Control* 43 (2017), pp. 65–79.
- [180] A.V. Proskurnikov and R. Tempo. "A tutorial on modeling and analysis of dynamic social networks. Part II". In: *Annual Reviews in Control* 45 (2018), pp. 166–190.
- [181] M. Pruefer, C. Schmidt, and F. Wahl. "Identification of robot dynamics with differential and integral models: A comparison". In: pt 1. 1994, pp. 340–345.
- [182] M. Rafei, D.D. Ganji, and H. Daniali. "Solution of the epidemic model by homotopy perturbation method". In: *Applied Mathematics and Computation* 187.2 (2007), pp. 1056–1062.

- [183] C. Ravazzi, F. Dabbene, C. Lagoa, and A.V. Proskurnikov. “Learning Hidden Influences in Large-Scale Dynamical Social Networks: A Data-Driven Sparsity-Based Approach, in Memory of Roberto Tempo”. In: *IEEE Control Systems Magazine* 41.5 (2021), pp. 61–103.
- [184] S. Rivero, M. Farina, and G. Ferrari-Trecate. “Plug-and-play decentralized model predictive control for linear systems”. In: *IEEE Transactions on Automatic Control* 58.10 (2013), pp. 2608–2614.
- [185] W. S. Rossi, J. W. Polderman, and P. Frasca. “The Closed Loop Between Opinion Formation and Personalized Recommendations”. In: *IEEE Transactions on Control of Network Systems* 9.3 (2022), pp. 1092–1103.
- [186] D. Sabbagh, P. Ablin, G. Varoquaux, A. Gramfort, and D.A. Engemann. “Predictive regression modeling with MEG/EEG: from source power to signals and cognitive states”. In: *NeuroImage* 222 (2020), p. 116893. ISSN: 1053-8119.
- [187] H. Sadeghi and A.-A. Raie. “Histogram distance metric learning for facial expression recognition”. In: *Journal of Visual Communication and Image Representation* 62 (2019), pp. 152–165.
- [188] L. Salzarulo. “A continuous opinion dynamics model based on the principle of meta-contrast”. In: *Journal of artificial societies and social simulation* 9.1 (2006), pp. 65–93.
- [189] C. Sandrock. *alchemyst/ternplot*. Available at <https://www.mathworks.com/matlabcentral/fileexchange/2299-alchemyst-ternplot> (2022/04/26). 2022.
- [190] R. Sarraj, E. Ballot, S. Pan, D. Hakimi, and B. Montreuil. “Interconnected logistic networks and protocols: Simulation-based efficiency assessment”. In: *International Journal of Production Research* 52.11 (2014), pp. 3185–3208.
- [191] S. Schweighofer, F. Schweitzer, and D. Garcia. “A weighted balance model of opinion hyperpolarization”. In: *Journal of Artificial Societies and Social Simulation* 23.3 (2020), p. 5.
- [192] Y. Shang. “Resilient Consensus for Expressed and Private Opinions”. In: *IEEE Transactions on Cybernetics* 51.1 (2021), pp. 318–331.
- [193] J. Sijs and M. Lazar. “On event based state estimation”. In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 5469 (2009), pp. 336–350.
- [194] S. Silvestri, R. Uргаonkar, M. Zafer, and B. Jun Ko. “A framework for the inference of sensing measurements based on correlation”. In: *ACM Transactions on Sensor Networks* 15.1 (2018).
- [195] A. Sirbu, V. Loreto, V.D.P. Servedio, and F. Tria. *Opinion dynamics: Models, extensions and external effects*. Understanding Complex Systems 9783319256566. 2017, pp. 363–401.
- [196] P. Sobkowicz. “Opinion dynamics model based on cognitive biases of complex agents”. In: *Journal of artificial societies and social simulation* 21.4 (2018).

- [197] P. Sobkowicz. "Studies of opinion stability for small dynamic networks with opportunistic agents". In: *International Journal of Modern Physics C* 20.10 (2009), pp. 1645–1662.
- [198] P. Sobkowicz. "Whither Now, Opinion Modelers?" In: *Frontiers in Physics* 8 (2020).
- [199] P. Sobkowicz and A. Sobkowicz. "Dynamics of hate based Internet user networks". In: *European Physical Journal B* 73.4 (2010), pp. 633–643.
- [200] D. Spiegelhalter, N.G. Best, B.P. Carlin, and A. van der Linde. "Bayesian measures of model complexity and fit". In: *Quality control and applied statistics* 48.4 (2003), pp. 431–432.
- [201] J. Su, B. Liu, Q. Li, and H. Ma. "Coevolution of opinions and directed adaptive networks in a social group". In: *Journal of artificial societies and social simulation* 17.2 (2014).
- [202] K. Sznajd-Weron and J. Sznajd. "Opinion evolution in closed community". In: *International Journal of Modern Physics C* 11.6 (2000), pp. 1157–1165.
- [203] N.S. Thompson and P. Derr. "Contra Epstein, good explanations predict". In: *Journal of artificial societies and social simulation* 12.1 (2009).
- [204] H.L. Trentelman, K. Takaba, and N. Monshizadeh. "Robust synchronization of uncertain linear multi-agent systems". In: *IEEE Transactions on Automatic Control* 58.6 (2013), pp. 1511–1523.
- [205] K.G. Troitzsch. "Not all explanations predict satisfactorily, and not all good predictions explain". In: *Journal of artificial societies and social simulation* 12.1 (2009).
- [206] H.C. Tuckwell and R.J. Williams. "Some properties of a simple stochastic epidemic model of SIR type". In: *Mathematical Biosciences* 208.1 (2007), pp. 76–97.
- [207] V. Ugrinovskii. "Distributed robust estimation over randomly switching networks using H_∞ consensus". In: *Automatica* 49.1 (2013), pp. 160–168.
- [208] D. Urbig. "Attitude dynamics with limited verbalisation capabilities". In: *Journal of artificial societies and social simulation* 6.1 (2003), p. 2.
- [209] D. Urbig, J. Lorenz, and H. Herzberg. "Opinion dynamics: The effect of the number of peers met at once". In: *Journal of artificial societies and social simulation* 11.2 (2008).
- [210] E. Volz and L.A. Meyers. "Epidemic thresholds in dynamic contact networks". In: *Journal of the Royal Society Interface* 6.32 (2009), pp. 233–241.
- [211] W. Wang and X.-Q. Zhao. "An epidemic model in a patchy environment". In: *Mathematical Biosciences* 190.1 (2004), pp. 97–112.
- [212] S. Watanabe and M. Opper. "Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory." In: *Journal of machine learning research* 11.12 (2010).
- [213] J.B. Watson. "Psychology as the behaviourist views it". In: *Psychological Review* 20.2 (1913), pp. 158–177.

- [214] D.J. Watts and S.H. Strogatz. "Collective dynamics of 'small-world' networks". In: *nature* 393.6684 (1998), pp. 440–442.
- [215] G. Weisbuch. "Bounded confidence and social networks". In: *European Physical Journal B* 38.2 (2004), pp. 339–343.
- [216] G. Weisbuch, G. Deffuant, F. Amblard, and J.-P. Nadal. "Interacting agents and continuous opinions dynamics". In: *Heterogenous agents, interactions and economic performance*. Springer, 2003, pp. 225–242.
- [217] E.B. Wilson. "Probable inference, the law of succession, and statistical inference". In: *Journal of the American Statistical Association* 22.158 (1927), pp. 209–212.
- [218] Weiguo X., Ming C., and K. H. Johansson. "Structural balance and opinion separation in trust–mistrust social networks". English. In: *IEEE Transactions on Control of Network Systems* 3.1 (Mar. 2016), pp. 46–56. ISSN: 2325-5870.
- [219] J. Xu and J.X. Li. "State estimation with quantised sensor information in wireless sensor networks". In: *IET Signal Processing* 5.1 (2011), pp. 16–26.
- [220] L. Yan, F. Xu, J. Liu, K.L. Teo, and M. Lai. "Stability strategies of demand-driven supply networks with transportation delay". In: *Applied Mathematical Modelling* 76 (2019), pp. 109–121.
- [221] M. Yildirim Imamoğlu. "Supply chain management at real time customer demand". In: 2009.
- [222] X. Yin, H. Wang, P. Yin, and H. Zhu. "Agent-based opinion formation modeling in social network: A perspective of social psychology". In: *Physica A: Statistical Mechanics and its Applications* 532 (2019).
- [223] X. Yu and Y. Xue. "Smart grids: A cyber–physical systems perspective". In: *Proceedings of the IEEE* 104.5 (2016), pp. 1058–1070.
- [224] J.-P. Zhang and Z. Jin. "The analysis of an epidemic model on networks". In: *Applied Mathematics and Computation* 217.17 (2011), pp. 7053–7064.
- [225] X. Zhang, Q. Lu, and T. Wu. "Petri-net based applications for supply chain management: An overview". In: *International Journal of Production Research* 49.13 (2011), pp. 3939–3961.
- [226] Y. Zhao, G. Kou, Y. Peng, and Y. Chen. "Understanding influence power of opinion leaders in e-commerce networks: An opinion dynamics theory perspective". In: *Information Sciences* 426 (2018), pp. 131–147.
- [227] X. Zhou, X. Li, and W.-S. Wang. "Bifurcations for a deterministic SIR epidemic model in discrete time". In: *Advances in Difference Equations* 2014.1 (2014).
- [228] G. Zhu, G. Chen, and X. Fu. "Effects of active links on epidemic transmission over social networks". In: *Physica A: Statistical Mechanics and its Applications* 468 (2017), pp. 614–621.
- [229] G. Zhu, X. Fu, and G. Chen. "Spreading dynamics and global stability of a generalized epidemic model on complex heterogeneous networks". In: *Applied Mathematical Modelling* 36.12 (2012), pp. 5808–5817.

- [230] L. Zhu, G. Guan, and Y. Li. “Nonlinear dynamical analysis and control strategies of a network-based SIS epidemic model with time delay”. In: *Applied Mathematical Modelling* 70 (2019), pp. 512–531.

CURRICULUM VITÆ

Carlos Andrés DEVIA PINZON

03-12-1993 Born in Bogotá, Colombia.

EDUCATION

- 2011–2015 **B.S. in Electronic Engineering** (5-year diploma)
Pontificia Universidad Javeriana - Bogotá, Colombia
Project: 'Fusión Sensorial GPS/IMU usando un Filtro de Kalman en el Quadrotor AR Drone 2.0'. Overall rank 4.6/5.0
- 2013–2018 **B.S. in Mathematics** (5-year diploma)
Pontificia Universidad Javeriana - Bogotá, Colombia
Project: 'Classification of rank 1 and 2 affine homogeneous distributions on 3-manifolds'. Overall rank 4.8/5.0
- 2016–2018 **M.S. in Electronic Engineering** (2-year diploma)
Pontificia Universidad Javeriana - Bogotá, Colombia
Project: 'Nonlinear Model Predictive Control for Aggressive Maneuvers in a Variable Pitch Quadrotor based on the Extended Modal Series Method'.
- **Honour Mention**. Overall rank 4.7/5.0
- 2019 **PhD. Network Systems and Opinion Formation**
Delft University of Technology - Delft, The Netherlands
Thesis: Behaviour, Stability, and Control of Interconnected Systems: From Biological Applications to Opinion Dynamics
Promotor: Prof. dr. ir. G. Giordano

AWARDS

- 2011 Academic Scholarship Undergraduate Program (Electronic Engineering)
2013 Academic Scholarship Undergraduate Program (Electronic Engineering)
2018 Honour Mention for Master Graduation Project

WORK EXPERIENCE

2016–2017	Research Project: Aerial robot applied for autonomous detection of landmines using vision and ground-penetrating radar
2017–2018	Research Project: Aerial sensing and monitoring of rice crop fields applying precision agriculture techniques
2018–2019	University instructor in courses of Calculus, Linear Algebra, Differential Equations, Dynamical Systems, Electronic Circuits, and Control Theory

LIST OF PUBLICATIONS

Journal Papers

1. **C. A. Devia**, G. Giordano, “Classification-Based Opinion Formation Model Embedding Agents’ Psychological Traits”, *Journal of Artificial Societies and Social Simulation*, **2023**.
2. **C. A. Devia**, G. Giordano, “A framework to analyze opinion formation models”, *Scientific Reports*, 12:13441, **2022**.
3. F. Blanchini, **C. A. Devia**, G. Giordano, R. Pesenti, F. Rosset, “Fair and sparse solutions in network-decentralised flow control”, *IEEE Control Systems Letters*, 6:2984-2989, **2022**. (Also selected for presentation at the *61st IEEE Conference on Decision and Control*, Cancún, Mexico, December 2022.)
4. **C. A. Devia**, G. Giordano, “Topology-independent robust stability conditions for uncertain MIMO networks”, *IEEE Control Systems Letters*, 5(1), 325-330, **2021**. (Also selected for presentation as an invited paper at the *59th IEEE Conference on Decision and Control*, Jeju Island, Republic of Korea, December 2020.)
5. **C. Devia**, J. Rojas, E. Petro, C. Martinez, I. Mondragon, D. Patino, C. Rebollo, J. Colorado, “High-throughput biomass estimation in rice crops using UAV multispectral imagery”, *Journal of Intelligent and Robotic Systems*, 96(3–4), 573–589, **2019**.
6. J. Colorado, M. Perez, I. Mondragon, D. Mendez, C. Parra, **C. Devia**, J. Martinez, L. Neira, “An Integrated Aerial System for Landmine Detection: SDR-based Ground Penetrating Radar onboard an Autonomous Drone”, *Advanced Robotics*, 31(15), 791-808. **2017**

Conference Papers

1. **C. A. Devia**, G. Giordano, “MIMO networks with heterogeneous uncertainties: topology-independent robust stability and alpha-convergence”, *Proc. 2021 European Control Conference (ECC)*, Rotterdam, The Netherlands, June-July **2021**.
2. **C. A. Devia**, G. Giordano, “Topology-independent robust stability for networks of homogeneous MIMO systems”, *Proceedings of the 21st IFAC World Congress*, Berlin, Germany, July **2020**.
3. **C. A. Devia**, G. Giordano, “Optimal duration and planning of switching treatments taking drug toxicity into account: a convex optimisation approach”, *Proceedings of the 58th IEEE Conference on Decision and Control*, Nice, France, December **2019**.
4. **C. Devia**, D. Patino, and J. Colorado, “Gradient-Descent based Nonlinear Model Predictive Control for Input-Affine Systems”, *The 6th International Conference on Control, Decision and Information Technologies (CODIT’19)* Paris, France, April **2019**, .
5. **C. Devia**, J. Rojas, E. Petro, C. Martinez, I. Mondragon, D. Patino, C. Rebollo, J. Colorado, “Aerial monitoring of rice crop variables using an UAV robotic system”, *The 16th International Conference on Informatics in Control, Automation and Robotics -ICINCO*, Prague, Czech Republic, July **2019**, ISBN 978-989-758-380-3, pages 97-103

6. J. Colorado, C. Rossi, A. Barrientos, **C. Devia**, D. Patino, "The role of massive morphing wings for maneuvering a bio-inspired bat-like robot", *IEEE International Conference on Robotics and Automation -ICRA*, Brisbane, Australia, May **2018**.
7. **C. Devia**, M. Roa, D. Patino, and J. Colorado, "Towards a Nonlinear Model Predictive Control using the Extended Modal Series Method", *European Control Conference -ECC 2018*, Limassol, Cyprus, June **2018**.
8. J. Rojas, **C. Devia**, E. Petro, C. Martinez, I. Mondragon, D. Patino, C. Rebolledo, J. Colorado, "Aerial mapping of rice crops using mosaicing techniques for vegetative index monitoring", *International Conference on Unmanned Aircraft Systems -ICUAS'18*, Dallas, TX, USA, Jun , **2018**.
9. **C. Devia**, J. Peña, L. Roldán, D. Patino, "Comparison of two Model Free Optimization Techniques for Maximum Power Point Tracking", *IEEE Colombian Conference on Automatic Control (CCAC)*, Cartagena, Colombia, **2017**.
10. **C. Devia**, D. Patino, "Multivariable optimal Point Tracking algorithm", *IEEE Conference on Control Applications (CCA)*, Buenos Aires, Argentina, **2016**.
11. **C. Devia**, D. Patino, "Tracking of Optimal Local Points", *Colombian Conference on Automatic Control (CCAC)*, Manizales, Colombia **2015**.
12. **C. Devia**, N. Alvarez, J. Hurtado, "Reproduction of the Music Contained in a Music Sheet Using MATLAB", *Simposio de Tratamiento de Señales e Imágenes (STSIVA)*, Bogotá, Colombia **2015**.

Preprints

1. **C. A. Devia**, G. Giordano, "Classification-Based Opinion Formation Model Embedding Agents' Psychological Traits", October **2022**.
2. **C. A. Devia**, G. Giordano, "A framework to analyze opinion formation models", May **2022**.
3. F. Blanchini, **C. A. Devia**, G. Giordano, "Structural polyhedral stability of a biochemical network is equivalent to finiteness of the associated generalised Petri net", arXiv:2109.01709, September **2021**.

Working Papers

1. **C. A. Devia**, G. Giordano, "Graphical and Probabilistic Analysis of Agent-Based Opinion Formation Models".