

Collective decision-making on networked systems in presence of antagonistic interactions

Angela Fontan

Linköping studies in science and technology. Dissertations.
No. 2166

**Collective decision-making on networked systems in presence of antagonistic
interactions**

Angela Fontan

*angela.fontan@liu.se
www.control.isy.liu.se
Division of Automatic Control
Department of Electrical Engineering
Linköping University
SE-581 83 Linköping
Sweden*

ISBN 978-91-7929-017-7 ISSN 0345-7524

Copyright © 2021 Angela Fontan

Printed by LiU-Tryck, Linköping, Sweden 2021

Ad Anna, Gianna e Luigi

Abstract

Collective decision-making refers to a process in which the agents of a community exchange opinions with the objective of reaching a common decision. It is often assumed that a collective decision is reached through collaboration among the individuals. However in many contexts, concerning for instance collective human behavior, it is more realistic to assume that the agents can collaborate or compete with each other. In this case, different types of collective behavior can be observed. This thesis investigates collective decision-making problems in multiagent systems, both in the case of collaborative and of antagonistic interactions.

The first problem studied in the thesis is a special instance of the consensus problem, denoted “interval consensus” in this work. It consists in letting the agents impose constraints on the possible common consensus value. It is shown that introducing saturated nonlinearities in the decision-making dynamics to describe how the agents express their opinions effectively allows the agents to influence the achievable consensus value and steer it to the intersection of all the intervals imposed by the agents.

A second class of collective decision-making models discussed in the thesis is obtained by replacing the saturations with sigmoidal nonlinearities. This nonlinear interconnected model is first investigated in the collaborative case and then in the antagonistic case, represented as a signed graph of interactions. In both cases, it is shown that the behavior of the model can be described by means of bifurcation analysis, with the equilibria of the system encoding the possible decisions for the community. A scalar positive parameter, denoted “social effort”, is added to the model to represent the strength of commitment between the agents, and plays the role of bifurcation parameter in the analysis. It is shown that if the social effort is small, then the community is in a deadlock situation (i.e., no decision is taken), while if the agents have the “right” amount of commitment two alternative consensus decision states for the community are achieved. However, by further increasing the social effort, the agents may fall in a situation of “overcommitment” where multiple (more than 2) decisions are possible. When antagonistic interactions between the agents are taken into account, they may lead to conflicts or social tensions during the decision-making process, which can be quantified by the notion of “frustration” of the signed network representing the community. The aim is to understand how the presence of antagonism (represented by the amount of frustration of the signed network) influences the collective decision-making process. It is shown that, while the qualitative behavior of the system does not change, the value of social effort required from the agents to break the deadlock (i.e., the value for which the bifurcation is crossed) increases with the frustration of the signed network: the higher the frustration, the higher the required social commitment.

A natural context to apply these results is that of political decision-making. In particular it is shown in the thesis how the government formation process in parliamentary democracies can be modeled as a collective decision-making system, where the agents are the parliamentary members, the decision is the vote of confidence they cast to a candidate cabinet coalition, and the social effort parameter is a proxy for the duration of the government negotiation talks. A signed network

captures the alliances/rivalries between the political parties in the parliament. The idea is that the frustration of the parliamentary networks should correlate well with the duration of the government negotiation, and it is supported by the analysis of the legislative elections in 29 European countries in the last 40 years.

The final contribution of this thesis is an analysis of the structure of (signed) Laplacian matrices and of their pseudoinverses. It is shown that the pseudoinverse of a Laplacian is in general a signed Laplacian, and in particular that the set of eventually exponentially positive Laplacian matrices (i.e., matrices whose exponential is a matrix with negative entries which becomes and stays positive at a certain power) is closed under stability and matrix pseudoinversion.

Populärvetenskaplig sammanfattning

Kollektivt beslutsfattande kan definieras som en process där agenter i en grupp utbyter åsikter med målet att fatta ett beslut. Det antas ofta att ett kollektivt beslut nås genom samarbete mellan individerna. I många fall, till exempel mänskligt kollektivt beteende, är det däremot mer realistiskt att anta att agenterna kan både samarbeta och motverka varandra. I dessa fall kan olika typer av kollektivt beteende observeras. Denna avhandling undersöker kollektivt beslutsfattande i fleragentssystem, både med kollaborativa och antagonistiska interaktioner.

Först studeras ”intervallkonsensus”, ett särskilt fall av konsensusproblemet där agenterna tillåts begränsa det möjliga konsensusvärdet. Det visas att genom att introducera begränsningar i hur agenterna uttrycker sina åsikter i beslutfattningsprocessen blir det möjligt för agenterna att påverka det möjliga konsensusvärdet och styra det till ett värde som accepteras av alla agenter.

En andra klass av kollektiva beslutfattningsmodeller som diskuteras i avhandlingen fås genom att ersätta de hårda begränsningarna med mjukare (S-formade olinjäriteter), och utöka modellen med en parameter kallad ”social ansträngning” som representerar styrkan hos engagemanget mellan agenterna. Den sammankopplade modellen undersöks först i det kollaborativa fallet och sedan i det antagonistiska fallet. I grafen som beskriver gruppen representeras kollaborativa och antagonistiska interaktioner av bågar med positiva respektive negativa vikter. Målet är att förstå hur systemets beteende förändras när parametern för social ansträngning varieras. I båda fallen visas det att om den sociala ansträngningen är liten så hamnar gruppen i ett dödläge, men att om agenterna har lagom mängd engagemang så kan två alternativa konsensusbeslutstillstånd för gruppen uppnås. Genom att ytterligare öka den sociala ansträngningen kan agenterna däremot bli ”överengagerade” vilket leder till fler än två möjliga beslut. När antagonistiska kopplingar mellan agenterna beaktas kan dessa leda till social spänning under beslutfattningsprocessen, vilket kan kvantifieras av ”frustrationen” i det nätverk som representerar gruppen. Målet är att förstå hur närvaren av antagonism (representerad av mängden frustration) influerar den kollektiva beslutfattningsprocessen. Det visas att även om systemets kvalitativa beteende inte förändras, så ökar den mängd social ansträngning som krävs av agenterna för att bryta dödläget med frustrationen i nätnätverket: ju större frustration desto större socialt engagemang krävs.

Politiskt beslutsfattande är ett exempel där resultaten i avhandling kan tillämpas. Det visas hur regeringsbildningsprocessen i parlamentariska demokratier kan modelleras som ett kollektivt beslutfattningssystem där agenterna är parlamentarikerna, beslutet är förtroenerösten de ger till en kandiderande regeringskoalition, och den sociala ansträngningen representeras av regeringsförhandlingarnas längd. Parlamentet kan modelleras som en graf där noderna representerar parlamentarikerna och tecknet på bågarnas vikter representerar allianser och rivaliteter mellan deras respektive partier. Hypotesen är att frustrationen i det parlamentariska nätnätverket korrelerar väl med regeringsförhandlingarnas längd, vilket stöds av en analys av valen i 29 europeiska länder under de senaste 40 åren.

Det sista bidraget i denna avhandling är en analys av strukturen hos teckengrafer (d.v.s. grafer som kan ha både positiva och negativa bågvikter), och av de

associerade laplacematrider och deras pseudoinverser. Pseudoinversen av en laplacematrik har många tillämpningar vilket gör det viktigt att förstå dess egenskaper. Det visas att pseudomatrizen generellt är en laplacematrik svarandes mot teckengraf, och villkor som garanterar stabilitet föreslås.

Acknowledgments

The first person I need to thank is definitely my supervisor, Claudio Altafini. I am incredibly grateful for the opportunity of pursuing a PhD with you, your curiosity and passion for research are truly inspiring. Thank you for all your guidance and support, and for being available every time I needed help or got stuck. Grazie!

I would also like to thank my co-supervisor, Anders Hansson, for his support and for always having kind words towards me, and for the many discussions on our shared interests, such as music, opera, gardening, and much more.

I am grateful to Svante Gunnarsson, Martin Enqvist and Ninna Stensgård for providing a motivating, functioning and constructive work environment. To Alberto, Erik, Martin L., Gustaf, Jonatan and Anna, and all my colleagues at RT, thank you. Doing a PhD might feel lonely sometimes, and I am glad I have met so many nice, fun, and supportive colleagues and friends. Lunch walks, fika discussions, exchange of ideas at the reading group, workouts, dinners, BBQs, parties; in many occasions I have felt lucky to be part of the RT group. I treasure these memories. Kristoffer, Per, and Gustaf, I really appreciate all the help you have given me with thesis related issues in the last period.

I would also like to thank my friends outside RT. Elisabeth, I so appreciate our discussions and dinners (with cake and wine of course); Oskar, thank you for the lunches and for teaching me Swedish customs. Many thanks also to France (Alessandro) and Giulio, for all the great times we have shared together since university, and to Arianna, Mattia, and la compagnia di Breda, for welcoming me so warmly every time I return to Italy.

Finally, I want to thank my family. Thank you Olle, for being so supportive, caring and patient. I can't help but smile when I think about you. Un grazie speciale va ad Anna, Gianna e Luigi, per avermi sempre supportato (e sopportato direi) nei miei studi e sostenuto nelle mie scelte. Papà, per avermi insegnato a guardare il mondo con curiosità, ti penso spesso; mamma, per essere un ottimo esempio di donna forte e determinata; Anna, perchè so che posso sempre contare su di te (qui ci starebbe bene uno sticker).

Thank you all!

*Linköping, August 2021
Angela Fontan*

Contents

I Background

1	Introduction	3
1.1	Contributions	7
1.2	Thesis outline	8
1.2.1	Part I: Background	8
1.2.2	Part II: Publications	8
2	Concepts from Matrix Theory and Perron-Frobenius Theory	13
2.1	Elements of matrix theory	13
2.2	Perron-Frobenius Theory	15
3	Nonlinear systems	19
3.1	Preliminary definitions	19
3.1.1	Cooperativity and Monotonicity	23
3.2	Bifurcation analysis	24
3.2.1	Continuous-time systems: pitchfork bifurcation	25
3.2.2	Discrete-time systems: pitchfork and period-doubling bifurcations	27
4	Theory of graphs and social networks	31
4.1	Graphs	31
4.1.1	Laplacian of a graph and its properties	33
4.2	Signed graphs	34
4.2.1	Signed Laplacian matrices	34
4.3	Social networks as signed graphs	37
4.3.1	Structurally balanced graphs and frustration index	37
5	Concluding remarks	41
	Bibliography	45

II Publications

A Interval Consensus for Multiagent Networks	53
1 Introduction	55
2 Problem Definition	58
2.1 The Model	58
2.2 Examples	59
2.3 Paper Outline	59
3 Background Material	60
3.1 Cooperativity	60
3.2 Limit Set, Dini Derivatives, and Invariance Principle	61
3.3 Robust Consensus	62
4 Nonempty Interval Intersection: Interval Consensus	62
4.1 Proof of Theorem 1	63
5 Empty Interval Intersection: Existence and Stability of Equilibria	65
5.1 Proof of Theorem 2	66
5.2 Proof of Theorem 3	67
5.3 Proof of Theorem 4	69
5.4 Proof of Theorem 5	72
6 Nonempty Interval Intersection in Discrete-Time	73
7 Numerical Examples	77
8 Conclusion	80
Bibliography	81
B Multiequilibria analysis for a class of collective decision-making networked systems	85
1 Introduction	87
2 Preliminaries	89
2.1 Concave and convex functions	89
2.2 Nonnegative matrices and Perron-Frobenius	90
2.3 Symmetric, symmetrizable and congruent matrices	90
2.4 Cooperative systems	91
3 Multiple equilibria in collective decision-making systems	91
3.1 Existence of multiple equilibria: a necessary condition	92
3.2 A geometric necessary and sufficient condition	96
3.3 Stability properties of multiple equilibria	99
4 Location of the mixed-sign equilibria	100
5 Numerical Analysis	101
6 Final considerations and conclusions	104
Bibliography	106
C The role of frustration in collective decision-making dynamical processes on multiagent signed networks	109
1 Introduction	111
2 Preliminaries	114
2.1 Notation and linear algebra	114

2.2	Signed graphs	114
2.3	Monotone systems	115
3	Decision-making in antagonistic multiagent systems in continuous-time	116
3.1	Problem formulation	116
3.2	Structurally balanced case	117
3.3	Structurally unbalanced case	117
4	Discrete-time	119
5	Discussion and interpretation of the results	122
6	Numerical Examples	124
7	Conclusions	130
A	Technical preliminaries	130
B	Proof of Theorem 2	131
B.1	Proof of Theorem 2(i)	132
B.2	Proof of Theorem 2(ii.1): existence	132
B.3	Proof of Theorem 2(ii.2): stability	134
B.4	Proof of Theorem 2(ii.3): uniqueness	137
C	Proof of Lemma 2	137
D	Proof of Theorem 3	138
D.1	Proof of Theorem 3(i)	139
D.2	Proof of Theorem 3(ii)	141
E	Proof of Proposition 1	142
F	Proof of Proposition 2	142
G	Proof of Lemma 3 and Lemma 4	143
H	Proof of Theorem 4	144
H.1	Proof of Theorem 4(i)	144
H.2	Proof of Theorem 4(ii)	146
	Bibliography	148
D	A signed network perspective on the government formation process in parliamentary democracies	151
1	Introduction	153
2	Results	155
2.1	Correlation between frustration and duration of the government negotiation talks	155
2.2	Prediction of the successful cabinet coalition	158
2.3	Interpretation as collective decision dynamics	159
3	Discussion	163
4	Materials and Methods	170
4.1	Data description	170
4.2	Construction of parliamentary networks	171
4.3	Measuring the frustration of signed parliamentary networks	172
4.4	Frustration vs fractionalization index	172
4.5	Minimum energy government coalition	173
4.6	Dynamical model of government formation	174

4.7	Rationale behind the correlation between frustration and duration of the government negotiation phase	175
4.8	Leave-one-out analysis for scenario III	175
4.9	Yearly trends	176
A	Methods	178
A.1	Parliamentary network construction	178
A.2	Structural balance and frustration for signed networks . . .	181
A.3	Dynamical model of decision-making in presence of frustration	184
B	Application: from parliamentary networks to government formation	187
B.1	Frustration vs government negotiation days	187
B.2	Minimum energy government coalition	187
B.3	Frustration vs smallest eigenvalue of the normalized signed Laplacian	189
B.4	Frustration vs fractionalization index	190
B.5	Description of the results	192
B.6	Influential points of the regression are important	192
B.7	A brief discussion on national rules and traditions influencing the duration of the government negotiation talks	193
B.8	A brief discussion on elections resulting in a hung parliament	194
B.9	A brief discussion on elections where the government negotiations failed	195
B.10	Analysis of the type of governments formed after the elections	196
B.11	Analysis of the Italian bicameral parliamentary system . .	197
	Bibliography	208
E	On the properties of Laplacian pseudoinverses	215
1	Introduction	217
2	Preliminaries	218
2.1	Linear algebraic preliminaries	218
2.2	Signed graphs	219
2.3	Eventual exponential positivity	220
2.4	Kron reduction for undirected networks	220
3	Pseudoinverse of eventually exponentially positive Laplacians . .	221
3.1	Directed signed network case	221
3.2	Undirected signed network case	227
4	Electrical networks and effective resistance	228
5	Conclusions and future work	231
	Bibliography	232

Part I

Background

1

Introduction

The idea of networks of interacting agents exchanging opinions or preferences in order to achieve a common decision as a community is not specific of a particular field or discipline, and examples of collective behaviors can be observed in everyday life. Typical examples range from collective behaviors in animal groups, such as selection of nest location, food source or migration mechanisms [1–4], to coordination and formation control or task-allocation problems in robotics [5–7], to opinion formation [8, 9]. In many of these cases, an agreement is reached through collaboration between the agents. However, there are other scenarios, such as e.g. social networks, trade markets, sport games, or parliamentary systems, where competition among the agents is intrinsic and unavoidable [10], hence the outcome of the interaction is less clear and the fact of achieving some form of agreement is not guaranteed.

Focusing on opinion dynamics models, i.e., on exchange of opinions among agents on a “social network”, in order to describe a collective decision-making process two elements are required: a framework able to capture how the individuals represent their opinions and communicate with each other, and a dynamical model describing how the opinions of the agents evolve in time. The natural choice for the framework is a state space model with state variables representing the opinions and a graph representing the interactions among the agents. As for the dynamics, many choices have been proposed in the literature in the last fifty years, as reviewed for instance in [11, 12], ranging from linear to nonlinear, from continuous- to discrete-time, from deterministic to probabilistic, etc.

Given a community of n agents, represented as a network \mathcal{G} , a state variable $x_i(t) \in \mathbb{R}$ (to ease the notation, the time-dependence is omitted hereafter) is assigned to each node i of the network to represent the opinion of agent i (at time t). In continuous-time, the evolution of the opinions of the agents in the network

\mathcal{G} can be described by a dynamical model in the generic form

$$\dot{x} = f(x) = [f_1(x) \cdots f_n(x)]^T \quad (1.1)$$

where $x = [x_1 \cdots x_n]^T \in \mathbb{R}^n$ is the vector of opinions, and $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is a vector field which depends on the structure of the network \mathcal{G} . The possible decisions for the community can be encoded in the equilibria of the system. When $f(x) = -Lx$, where L is the Laplacian matrix of the network \mathcal{G} , then the equilibria of the system (1.1) have all equal components. These types of equilibria are called consensus, and have been extensively studied in the literature, see e.g. [6, 11, 13].

Linear models like the one leading to consensus may fail to capture more complex behaviors, and nonlinearities are sometimes included in an attempt to make the models more realistic. However, the adoption of nonlinear models means that the dynamics may exhibit new phenomena, such as multiple equilibria, periodic orbits, and bifurcations. For instance, a class of nonlinear interconnected models that combines consensus with saturated and sigmoidal nonlinearities is proposed in [4]. It has the following structure:

$$\dot{x} = -\Delta x + \pi A \psi(x) \quad (1.2)$$

where $A = [a_{ij}] \in \mathbb{R}^{n \times n}$ is the adjacency matrix of the graph \mathcal{G} , $\Delta = \text{diag}\{\delta_1, \dots, \delta_n\}$, with $\delta_i = \sum_{j=1}^n a_{ij}$ for all i , is the diagonal term representing the weighted indegree, $\pi > 0$ is a positive scalar parameter, and $\psi(x) = [\psi_1(x_1) \cdots \psi_n(x_n)]^T$ is a vector of sigmoidal and saturated nonlinearities, see Fig. 1.1. By construction, in this model the linearization at the origin (for $\pi = 1$) corresponds to the graph Laplacian $L = \Delta - A$. The model (1.2) is considered in paper A of part II of this thesis for the case of saturation nonlinearities, and in paper B for the case of sigmoidal nonlinearities.

The interaction graph \mathcal{G} considered in the model (1.1) (and (1.2)) typically has edges with nonnegative weights, which capture the fact that the agents are collaborating. A natural way to extend this model to the case of agents that compete is to consider signed graphs, i.e., graphs whose edges have positive or negative weight [14]: a positive edge indicates cooperation (or friendship, trust) while a negative edge competition (or antagonism, distrust). A notion that is often used when dealing with signed graphs is that of structural balance, which captures the idea that no social tension is present at a network level even if antagonism is present (at a local level) between the agents. In particular, a signed graph is structurally balanced if its agents can be divided into two subgroups such that the agents in each subgroup are mutual friends (meaning that they are linked by edges with positive weight), and the agents belonging to different subgroups are rivals (linked by edges with negative weight) [15, 16]. A closely related notion is that of frustration of a signed graph, a concept introduced to measure its distance from a structurally balanced state [17, 18] (see also [19] for an overview of the different measures proposed in the literature to estimate the level of imbalance a of signed graph): a signed graph is structurally balanced if and only if it has zero frustration.

In the context of multiagent dynamics, the notion of structural balance is strictly related to that of monotone systems (or order preserving systems) [20], and

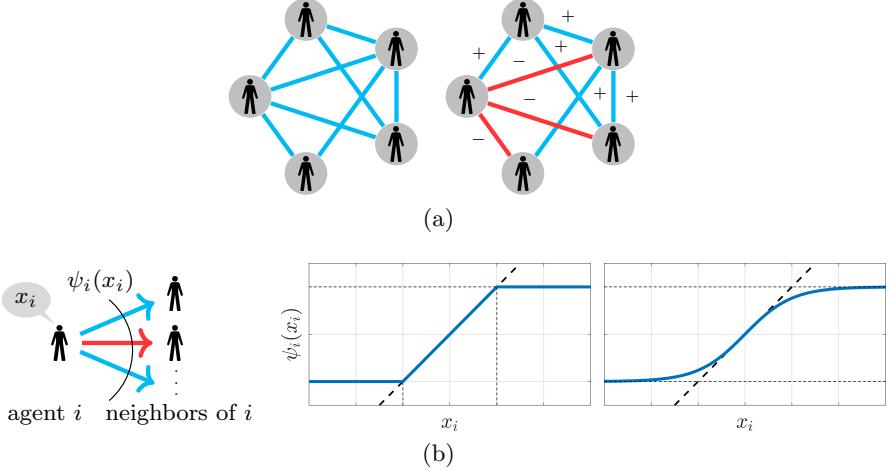


Figure 1.1: Representation of a collective decision-making process, described as a nonlinear system over (signed) networks, see (1.2). (a): A community of agents is modeled as a graph of interactions, which may be unsigned/cooperative (left) or signed (right). (b): Nonlinear functions $\psi_i(\cdot)$, $i = 1, \dots, n$, represent how each agent i transmits its opinion x_i to its neighbors in the network. The nonlinearities considered in this work are saturated and monotonically increasing.

the notion of frustration to that of distance to monotonicity [21]. In particular, for nonnegative graphs \mathcal{G} with saturated/sigmoidal nonlinearities, the model (1.2) is a so-called cooperative system, while when \mathcal{G} is signed and structurally balanced then (1.2) is a monotone system.

Real-world signed networks, from e.g., biological networks [18, 21, 22] to social networks [23, 24] and (multi-party) parliamentary networks [25], are in general not structurally balanced. For this reason, it is relevant to understand how the antagonism present in the signed networks affects the collective decision-making process. In paper C of part II of this thesis, the model (1.2) is extended to include the case in which the graph \mathcal{G} is signed. The difference is that the adjacency matrix A of \mathcal{G} is now a signed matrix, with the sign of each element (+1 or -1) representing the type of interaction (friendly/trust/alliance or unfriendly/distrust/rivalry), and the absolute value the amount of trust/distrust between the corresponding agents. Similarly, $\Delta = \text{diag}\{\delta_1, \dots, \delta_n\}$ is still a weighted in-degree matrix, but now with $\delta_i = \sum_{j=1}^n |a_{ij}|$ for all i . Under these assumptions, the linearization of the model (1.2) at the origin corresponds to a signed version of the Laplacian [16]. In particular, the system (1.2) is monotone if and only if the graph \mathcal{G} is structurally balanced. Consequently, when \mathcal{G} is not structurally balanced, its dynamics is not monotone and hence in general more complicated to understand. In papers C and D it is shown that the notion of frustration can be used to get some insight into the behavior of the system (1.2).

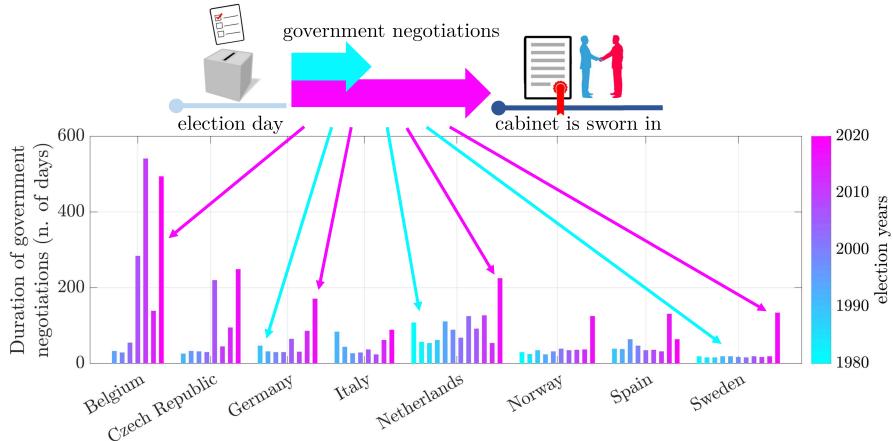


Figure 1.2: Duration of government negotiations (calculated as number of days between the election date and the date the government is sworn in) in selected European countries in the last 40 years (see paper D).

Example: Political Decision-Making

The government formation process in countries with a parliamentary system is a good example of collective decision-making applied to the field of Political Sciences. Briefly, the process starts with a legislative election after which each political party receives a certain number of seats in the parliament, depending on the number of votes gained at the election. In case of no clear winner, i.e., when no party or coalition has managed to secure a majority in the parliament, a negotiation starts between the political parties. The process concludes with the swearing-in of a new government, typically after winning a confidence vote in the parliament. The government negotiation phase might be complex and long. For instance, the past ten years have seen an increase in the duration of government negotiations, defined as the number of days between the election date and the date the government is sworn in. This has happened not only in countries famous for long periods of cabinet negotiations, such as Belgium, but also, e.g., in Germany, UK, and Sweden, see Fig. 1.2. The research question considered in paper D of part II of this thesis work is whether it is possible to use models for collective decision-making to explain this behavior. The idea is to describe the government formation process as a collective decision-making system on a signed graph having the parliamentary members as agents and their alliances/rivalries as edges. The resulting graphs are not structurally balanced, and the hypothesis is that their frustration correlates with the duration of the government negotiation phase.

1.1 Contributions

The contributions of this thesis work are within the area of complex (in general, signed) networks and nonlinear dynamics over networks, with applications to collective decision-making processes over social networks.

Starting from a problem of consensus proposed in Paper A, Papers B and C formulate a collective decision-making problem over cooperative and antagonistic networks, respectively, as a bifurcation problem, where the crossing of a pitchfork bifurcation corresponds to the achievement of a common decision. The results of Paper C admit a natural interpretation in the context of social networks, and in Paper D a concrete example of political decision-making is proposed. In paper E, departing from the study of nonlinear dynamics over networks, the focus is only on signed graphs and in particular on the structure of the signed Laplacian matrix and its pseudoinverse.

More specifically, the main contributions of this thesis work are:

- Formulation of a nonlinear interconnected cooperative model for “interval consensus”, where nonlinearities are introduced to allow the agents to impose constraints on the achievable consensus value. If the intersection of the constraints is nonempty, convergence to interval consensus is proven. (Paper A)
- Analysis of a nonlinear interconnected cooperative model for collective decision-making with saturating and sigmoidal nonlinearities, and scalar bifurcation parameter. Necessary and sufficient conditions for existence and stability of equilibria are proposed, and an exact upper bound for the norm of equilibria is derived. (Paper B)
- Extension of the nonlinear model of Paper B to antagonistic networks. Analysis of the qualitative behavior of the system, both in continuous- and discrete-time. (Paper C)
- Interpretation of the results of Paper C in terms of frustration of the signed network representing the community: the bifurcation value at which the first bifurcation is crossed is shown to depend on the frustration, and an upper bound for the 1-norm of the equilibria depending on the frustration is derived. (Paper C)
- Introduction of a concrete example of the theoretical work of paper C. The process of government formation in parliamentary democracies is explained as a collective decision-making process over signed “parliamentary” networks, where the crossing of a bifurcation corresponds to success/failure of a confidence vote. Through the collection and analysis of data of legislative elections in the last 40 years for 29 European countries, the frustration of the parliamentary networks is shown to be a good proxy for the complexity of the government negotiation process. (Paper D)
- Study of the properties of “repelling” signed Laplacian matrices and their pseudoinverses. The set of repelling signed Laplacian matrices which are eventually exponentially positive is shown to be closed under pseudoinversion and stability. (Paper E)

1.2 Thesis outline

The thesis is divided into two parts, with background material in the first part, and edited versions of published papers in the second part.

1.2.1 Part I: Background

The first part introduces theoretical background for the publications in part II. The preliminary material includes concepts and notions from matrix and graph theory, and nonlinear systems.

1.2.2 Part II: Publications

The second part of this thesis contains edited versions of the five papers listed below.

Paper A: Interval Consensus for Multiagent Networks

Paper A is an edited version of

A. Fontan, G. Shi, X. Hu, and C. Altafini, “Interval Consensus for Multiagent Networks,” *IEEE Transactions on Automatic Control*, vol. 65, no. 5, pp. 1855–1869, may 2020.

Summary The constrained consensus problem considered in paper A, denoted interval consensus, is characterized by the fact that each agent can impose a lower and upper bound on the achievable consensus value. Such constraints can be encoded in the consensus dynamics by saturating the values that an agent transmits to its neighboring nodes. We show in the paper that when the intersection of the intervals imposed by the agents is nonempty, the resulting constrained consensus problem must converge to a common value inside that intersection. In our algorithm, convergence happens in a fully distributed manner, and without need of sharing any information on the individual constraining intervals. When the intersection of the intervals is an empty set, the intrinsic nonlinearity of the network dynamics raises new challenges in understanding the node state evolution. Using Brouwer fixed-point theorem we prove that in that case there exists at least one equilibrium, and in fact the possible equilibria are locally stable if the constraints are satisfied or dissatisfied at the same time among all nodes. For graphs with sufficient sparsity it is further proven that there is a unique equilibrium that is globally attractive if the constraint intervals are pairwise disjoint.

Contribution and background The author of this thesis contributed with implementations, analysis and reviewing the manuscript.

Paper B: Multiequilibria Analysis for a Class of Collective Decision-Making Networked Systems

Paper B is an edited version of

A. Fontan and C. Altafini, “Multiequilibria Analysis for a Class of Collective Decision-Making Networked Systems,” *IEEE Transactions on Control of Network Systems*, vol. 5, no. 4, pp. 1931–1940, dec 2018.

Summary The models of collective decision-making considered in paper B are nonlinear interconnected cooperative systems with saturating interactions. These systems encode the possible outcomes of a decision process into different steady states of the dynamics. In particular, they are characterized by two main attractors in the positive and negative orthant, representing two choices of agreement among the agents, associated to the Perron-Frobenius eigenvector of the system. In this paper we give conditions for the appearance of other equilibria of mixed sign. The conditions are inspired by Perron-Frobenius theory and are related to the algebraic connectivity of the network. We also show how all these equilibria must be contained in a solid disk of radius given by the norm of the equilibrium point which is located in the positive orthant.

Contribution and background The author of this thesis contributed with the majority of the work including theoretical derivations, implementations, numerical calculations and the written manuscript.

Paper C: The role of frustration in collective decision-making dynamical processes on multiagent signed networks

Paper C is an edited version of

A. Fontan and C. Altafini, “The role of frustration in collective decision-making dynamical processes on multiagent signed networks,” *arXiv:2105.11396*, pp. 1–18, may 2021.

Summary In paper C we consider a collective decision-making process in a network of agents described by a nonlinear interconnected dynamical model with sigmoidal nonlinearities and signed interaction graph. The decisions are encoded in the equilibria of the system. The aim is to investigate this multiagent system when the signed graph representing the community is not structurally balanced and in particular as we vary its frustration, i.e., its distance to structural balance. The model exhibits bifurcations, and a “social effort” parameter, added to the model to represent the strength of the interactions between the agents, plays the role of bifurcation parameter in our analysis. We show that, as the social effort increases, the decision-making dynamics exhibits a pitchfork bifurcation behavior where, from a deadlock situation of “no decision” (i.e., the origin is the only globally stable equilibrium point), two possible (alternative) decision states for the community are achieved (corresponding to two nonzero locally stable equilibria).

The value of social effort for which the bifurcation is crossed (and a decision is reached) increases with the frustration of the signed network.

Contribution and background The author of this thesis contributed with the majority of the work including theoretical derivations, implementations, numerical calculations and the written manuscript.

Paper D: A signed network perspective on the government formation process in parliamentary democracies

Paper D is an edited version of

A. Fontan and C. Altafini, “A signed network perspective on the government formation process in parliamentary democracies,” *Scientific Reports*, vol. 11, no. 5134, dec 2021.

Summary In parliamentary democracies, government negotiations talks following a general election can sometimes be a long and laborious process. In order to explain this phenomenon, in this paper we use structural balance theory to represent a multiparty parliament as a signed network, with edge signs representing alliances and rivalries among parties. We show that the notion of frustration, which quantifies the amount of “disorder” encoded in the signed graph, correlates very well with the duration of the government negotiation talks. For the 29 European countries considered in this study, the average correlation between frustration and government negotiation talks ranges between 0.42 to 0.69, depending on what information is included in the edges of the signed network. Dynamical models of collective decision-making over signed networks with varying frustration are proposed to explain this correlation.

Contribution and background The author of this thesis contributed with the majority of the work including theoretical derivations, implementations, numerical calculations and the written manuscript.

Paper E: On the properties of Laplacian pseudoinverses

Paper E is an edited version of

A. Fontan and C. Altafini, “On the properties of Laplacian pseudoinverses,” in *60th IEEE Conference on Decision and Control*. Austin, TX, USA: IEEE, 2021.

Summary The pseudoinverse of a graph Laplacian is used in many applications and fields, such as for instance in the computation of the effective resistance in electrical networks, in the calculation of the hitting/commuting times for a Markov chain and in continuous-time distributed averaging problems. In this paper we show that the Laplacian pseudoinverse is in general not a Laplacian matrix but rather a signed Laplacian with the property of being an eventually exponentially

positive matrix, i.e., of obeying a strong Perron-Frobenius property. We show further that the set of signed Laplacians with this structure (i.e., eventual exponential positivity) is closed with respect to matrix pseudoinversion. This is true even for signed digraphs, and provided that we restrict to Laplacians that are weight balanced also stability is guaranteed.

Contribution and background The author of this thesis contributed with the majority of the work including theoretical derivations, implementations, numerical calculations and the written manuscript.

Publications

Works published by the author of this thesis are listed below in chronological order¹. Publications indicated by a \star are included in the thesis.

- A. Fontan, G. Shi, X. Hu, and C. Altafini, “Interval consensus: a novel class of constrained consensus problems for multiagent networks,” in *56th IEEE Conference on Decision and Control*. Melbourne, Australia: IEEE, dec 2017, pp. 4155–4160.
- A. Fontan and C. Altafini, “Investigating mixed-sign equilibria for nonlinear collective decision-making systems,” in *56th IEEE Conference on Decision and Control*. Melbourne, Australia: IEEE, dec 2017, pp. 781–786.
- \star A. Fontan and C. Altafini, “Multiequilibria Analysis for a Class of Collective Decision-Making Networked Systems,” *IEEE Transactions on Control of Network Systems*, vol. 5, no. 4, pp. 1931–1940, dec 2018.
- A. Fontan and C. Altafini, “Modeling wireless power transfer in a network of smart devices: A compartmental system approach,” in *2018 17th European Control Conference (ECC)*. Limassol, Cyprus: European Control Association (EUCA), 2018, pp. 1468–1473.
- A. Fontan and C. Altafini, “Achieving a decision in antagonistic multi agent networks: frustration determines commitment strength,” in *57th IEEE Conference on Decision and Control*. Miami Beach, FL, USA: IEEE, dec 2018, pp. 109–114.
- \star A. Fontan, G. Shi, X. Hu, and C. Altafini, “Interval Consensus for Multi-agent Networks,” *IEEE Transactions on Automatic Control*, vol. 65, no. 5, pp. 1855–1869, may 2020.
- \star A. Fontan and C. Altafini, “A signed network perspective on the government formation process in parliamentary democracies,” *Scientific Reports*, vol. 11, no. 5134, dec 2021.

¹Extended abstracts are not included in the list.

★A. Fontan and C. Altafini, “The role of frustration in collective decision-making dynamical processes on multiagent signed networks,” *arXiv:2105.11396*, pp. 1–18, may 2021.

★A. Fontan and C. Altafini, “On the properties of Laplacian pseudoinverses,” in *60th IEEE Conference on Decision and Control*. Austin, TX, USA: IEEE, 2021.

2

Concepts from Matrix Theory and Perron-Frobenius Theory

The aim of this chapter is to introduce notation and preliminary material from matrix theory, and it is mostly based on [34]. The second part of this chapter is dedicated to Perron-Frobenius theory. This topic is relevant for this thesis work due to its applications in dynamical systems (e.g., the study of cooperative/monotone systems) and graph theory (e.g., the study of the properties of Laplacian matrices).

2.1 Elements of matrix theory

This section presents a collection of definitions and results from matrix theory, with particular focus on the characterization of eigenvalues¹.

Let $\mathbb{N}, \mathbb{R}, \mathbb{C}$ indicate the set of natural, real, and complex numbers, respectively. The spectrum of a matrix $A \in \mathbb{R}^{n \times n}$ is denoted $\Lambda(A) = \{\lambda_1(A), \dots, \lambda_n(A)\}$, where $\lambda_i(A)$, $i = 1, \dots, n$, are the eigenvalues of A . The eigenvalues are in general assumed to be arranged in the following (nondecreasing) order: $\text{Re}[\lambda_1(A)] \leq \text{Re}[\lambda_2(A)] \leq \dots \leq \text{Re}[\lambda_n(A)]$, where $\text{Re}[\cdot]$ denotes the real part (similarly, $\text{Im}[\cdot]$ is used to denote the imaginary part). The *spectral radius* of A is $\rho(A) = \max\{|\lambda| : \lambda \in \Lambda(A)\}$, and its *spectral abscissa* is $\mu(A) = \max\{\text{Re}[\lambda] : \lambda \in \Lambda(A)\}$.

The *kernel* (or, null space) of A is $\ker(A) = \{x \in \mathbb{R}^n : Ax = 0\}$, and its *range* (or, column space) is $\text{range}(A) = \{y \in \mathbb{R}^n : y = Ax\}$. The *rank* of A is $\text{rank}(A) = \text{rank}(A^T) = n - \dim(\ker(A))$, and the *corank* of A is $\text{corank}(A) = \dim(\ker(A))$.

Characterization of symmetric and symmetrizable matrices A matrix $A \in \mathbb{R}^{n \times n}$ is called *symmetric* if $A = A^T$, and it is called (*diagonally*, omitted hereafter) *symmetrizable* if DA is symmetric for some diagonal matrix D with positive

¹Basically, the aim is to answer the questions: “if a matrix satisfies a certain property, what conclusions can be drawn on its eigenvalues?”, or, similarly: “if two matrices satisfy a certain relationship, what conclusions can be drawn on their eigenvalues?”

diagonal entries. Symmetric or symmetrizable matrices have real eigenvalues (see [35, 36] for a more detailed characterization of symmetrizable matrices).

Characterization of similar and congruent matrices Given two matrices $A, B \in \mathbb{R}^{n \times n}$, B is *similar* to A (denoted $A \sim B$) if there exists a nonsingular matrix $S \in \mathbb{R}^{n \times n}$ such that $B = S^{-1}AS$, while it is *congruent* to A if there exists a nonsingular matrix $S \in \mathbb{R}^{n \times n}$ such that $B = SAS^T$. Similar matrices have the same eigenvalues. Congruent matrices that are also symmetric have the same *inertia*, i.e., the same number (counting multiplicity) of positive, negative and zero eigenvalues, also known as Sylvester's law of inertia [34, Thm 4.5.8]. Moreover, a bound between the eigenvalues of congruent (and symmetric) matrices is provided by the Ostrowski's Theorem.

Theorem 2.1 (Ostrowski, 4.5.9 in [34]). Let $A, S \in \mathbb{R}^{n \times n}$ with A symmetric and S nonsingular. Let the eigenvalues of A , SAS^T and SS^T be arranged in nondecreasing order. For each $k = 1, \dots, n$, there exists a positive real number θ_k such that $\lambda_1(SS^T) \leq \theta_k \leq \lambda_n(SS^T)$ and $\lambda_k(SAS^T) = \theta_k \lambda_k(A)$.

Characterization of diagonally dominant matrices A matrix $A = [a_{ij}] \in \mathbb{R}^{n \times n}$ is called *diagonally dominant (by rows)*, omitted hereafter) if

$$|a_{ii}| \geq \sum_{j=1}^n |a_{ij}|, \quad i = 1, \dots, n, \tag{2.1}$$

diagonally equipotent (nomenclature from [37]) if (2.1) holds with equality for all i , *weakly diagonally dominant* if (2.1) holds and at least one (but not all) of the inequalities is strict, and *strictly diagonally dominant* if (2.1) holds with strict inequality for all i . A matrix $A \in \mathbb{R}^{n \times n}$ is called *irreducible* if there does not exist a permutation matrix P s.t. $P^T AP$ is block triangular.

Strictly diagonally dominant matrices are nonsingular, and, if all their diagonal elements are strictly positive, then their eigenvalues have positive real part, see [34, Thm 6.1.10 (Levy-Desplanques)]. The same characterization holds for matrices that are *irreducibly diagonally dominant*, i.e., matrices that are irreducible and weakly diagonally dominant, see [34, Corollary 6.2.27 (Taussky)]).

Diagonally equipotent matrices that are irreducible and have nonnegative diagonal elements, have eigenvalues with nonnegative real part, and their corank is at most 1 [37, Proposition 1].

The Geršgorin's Theorem is an useful tool to prove (or give the intuition behind) the aforementioned results.

Theorem 2.2 (Geršgorin, 6.1.1 in [34]). Let $A = [a_{ij}] \in \mathbb{R}^{n \times n}$. The eigenvalues of A are in the union of the n Geršgorin's disks, defined as

$$\left\{ \lambda \in \mathbb{C} : |\lambda - a_{ii}| \leq \sum_{j=1}^n |a_{ij}| \right\}, \quad i = 1, \dots, n. \tag{2.2}$$

2.2 Perron-Frobenius Theory

In 1907 Perron showed that the class of positive matrices (i.e., matrices whose elements are strictly positive) is characterized by an important property, namely, that the spectral radius of a positive matrix is an eigenvalue and that the corresponding (left and right) eigenvectors are strictly positive. Later, thanks to Frobenius, this result was extended to the class of irreducible and nonnegative matrices (i.e., matrices with nonnegative elements), and recently to the class of eventually positive matrices (i.e., matrices which become and stay positive after a certain power). These findings, as well as related results, are reported in this section.

Characterization of positive and nonnegative matrices A matrix $A = [a_{ij}] \in \mathbb{R}^{n \times n}$ is *positive* (denoted $A > 0$) if $a_{ij} > 0$ for all $i, j = 1, \dots, n$, and is *nonnegative* (denoted $A \geq 0$) if $a_{ij} \geq 0$ for all $i, j = 1, \dots, n$. Similarly, a vector $x \in \mathbb{R}^n$ is *positive* (denoted $x > 0$) if $x_i > 0$ for all $i = 1, \dots, n$, and is *nonnegative* (denoted $x \geq 0$) if $x_i \geq 0$ for all $i = 1, \dots, n$.

A useful property is that the largest and smallest row sum of a nonnegative matrix provide an upper and lower bound, respectively, for its spectral radius.

Lemma 2.1 (Thm 8.1.22 in [34]). *Let $A \in \mathbb{R}^{n \times n}$ be nonnegative. Then*

$$0 \leq \min_{i=1, \dots, n} \left\{ \sum_{j=1}^n a_{ij} \right\} \leq \rho(A) \leq \max_{i=1, \dots, n} \left\{ \sum_{j=1}^n a_{ij} \right\}$$

Perron-Frobenius theory can be used to understand: (i) when the spectral radius is strictly positive, and (ii) when it is an eigenvalue.

The following theorem, called Perron's Theorem, characterizes the spectral properties of positive matrices.

Theorem 2.3 (Perron, 8.2.8 in [34]). *Let $A \in \mathbb{R}^{n \times n}$ be positive. Then*

- (i) $\rho(A) > 0$ is an algebraically simple eigenvalue of A ;
- (ii) there is a unique real vector $\chi \in \mathbb{R}^n$ such that $A\chi = \rho(A)\chi$, and $\chi > 0$;
- (iii) there is a unique real vector $\xi \in \mathbb{R}^n$ such that $\xi^T A = \rho(A)\xi^T$, and $\xi > 0$;
- (iv) $\rho(A) > |\lambda|$ for all $\lambda \in \Lambda(A)$;
- (v) $\lim_{k \rightarrow \infty} (\rho(A)^{-1} A)^k = \chi \xi^T$.

In general, these properties do not hold for the class of nonnegative matrices. In this case, however, it is known that the properties (i)÷(iii) of Theorem 2.3 admit the following generalization: the spectral radius is always a nonnegative eigenvalue, and its corresponding (left and right) eigenvectors are nonnegative, see [34, Thm 8.3.1]. The Perron-Frobenius Theorem shows that this result can be further extended, and in particular that matrices that are nonnegative and irreducible satisfy the first three properties of Theorem 2.3.

Theorem 2.4 (Perron-Frobenius, 8.4.4 in [34]). *Let $A \in \mathbb{R}^{n \times n}$ ($n \geq 2$) be nonnegative and irreducible. Then*

- (i) $\rho(A) > 0$ is an algebraically simple eigenvalue of A ;

- (ii) there is a unique real vector $\chi \in \mathbb{R}^n$ such that $A\chi = \rho(A)\chi$, and $\chi > 0$;
- (iii) there is a unique real vector $\xi \in \mathbb{R}^n$ such that $\xi^T A = \rho(A)\xi^T$, and $\xi > 0$.

Remark 2.1. When A is a positive, or a nonnegative and irreducible, matrix (as in Theorems 2.3 or 2.4, respectively), its eigenvalue $\rho(A)$ is called the *Perron eigenvalue*, and the corresponding left and right eigenvectors $\chi > 0$ and $\xi > 0$ are called the *Perron eigenvectors*. ———

Characterization of matrices that possess the strong Perron-Frobenius property

An important difference between positive and nonnegative and irreducible matrices, is that for the latter the Perron eigenvalue may not be the unique eigenvalue with largest modulus.

If a nonnegative matrix A is irreducible and $\rho(A) > \lambda(A)$ for all $\lambda(A) \in \Lambda(A)$, then it is called *primitive*. The class of nonnegative and primitive matrices is particularly relevant due to their asymptotic properties: indeed, if $A \geq 0$ is primitive, then $\lim_{k \rightarrow \infty} (\rho(A)^{-1}A)^k = \chi\xi^T$, where χ, ξ are the right and left Perron eigenvectors, respectively. Essentially, Perron's Theorem can be generalized to the class of nonnegative and primitive matrices.

Recently (see [38]), it has been shown that Perron's Theorem generalizes also to the class of eventually positive matrices, which are matrices that may have negative or zero elements, but become and stay positive after a certain power. To adequately state this result (see Theorem 2.5), the following two definitions are needed. A matrix $A \in \mathbb{R}^{n \times n}$ is called *eventually positive* (denoted $A \stackrel{\vee}{>} 0$) if there exists an integer $k_0 \in \mathbb{N}_+$ such that $A^k > 0$ for all $k \geq k_0$. A matrix A possesses the *strong Perron-Frobenius property* (denoted $A \in \mathcal{PF}$) if its spectral radius is a simple positive eigenvalue and no other eigenvalue has the same modulus, and the corresponding right eigenvector is positive.

Theorem 2.5 (2.2 in [38]). Let $A \in \mathbb{R}^{n \times n}$. Then the following statements are equivalent:

- (i) $A \in \mathcal{PF}$ and $A^T \in \mathcal{PF}$;
- (ii) $A \stackrel{\vee}{>} 0$;
- (iii) $A^T \stackrel{\vee}{>} 0$.

Remark 2.2. Observe that if $A > 0$ or, if $A \geq 0$ and primitive, then $A \stackrel{\vee}{>} 0$. ———

Characterization of Metzler matrices and M-matrices Finally, this section concludes by introducing Metzler matrices and M-matrices, for which interesting properties can be derived using Perron-Frobenius theory. These notions will turn useful in the next chapters: for instance, Chapter 3 will show that Metzler matrices have a close relationship with cooperative systems, while the properties of M-matrices will be used in Chapter 4 to examine the structure of Laplacian matrices.

A matrix $A \in \mathbb{R}^{n \times n}$ is called a *Metzler matrix* if it can be written as $A = B - sI$, where $B \geq 0$ and $s \in \mathbb{R}$, i.e., it if all its off-diagonal elements are nonnegative. The following lemma follows from the Perron-Frobenius Theorem.

Lemma 2.2. Let $A = B - sI \in \mathbb{R}^{n \times n}$ be an irreducible Metzler matrix, i.e., $B \geq 0$ is irreducible and $s \in \mathbb{R}$. Then

- (i) $\mu(A)$ is a simple eigenvalue of A , and $\mu(A) = \rho(B) - s$;
- (ii) there is a unique real vector $\chi > 0$ such that $A\chi = \mu(A)\chi$;
- (iii) there is a unique real vector $\xi > 0$ such that $\xi^T A = \xi^T \mu(A)$

A matrix A is called an *M-matrix* if it can be written as $A = sI - B$, where $B \geq 0$ and $s \geq \rho(B)$. Perron-Frobenius theory has important consequences for the spectrum of M-matrices, as the following lemma shows.

Lemma 2.3. Let $A = sI - B \in \mathbb{R}^{n \times n}$ be an irreducible M-matrix, i.e., $B \geq 0$ is irreducible and $s \geq \rho(B)$. Then

- $\lambda_1(A)$ is real and simple, and $\lambda_1(A) = s - \rho(B) \geq 0$;
- there is a unique real vector $\chi > 0$ such that $A\chi = \lambda_1(A)\chi$;
- there is a unique real vector $\xi > 0$ such that $\xi^T A = \lambda_1(A)\xi^T$.

Therefore, a matrix A is an M-matrix if and only if $-A$ is Metzler and all its eigenvalues have nonnegative real part. If $s = \rho(B)$ then A is called a singular M-matrix. The properties of M-matrices have been extensively studied, see e.g. [39]. A noteworthy property is that the inverse of a (nonsingular) M-matrix is nonnegative.

3

Nonlinear systems

This chapter provides background theory on nonlinear systems and on some of the tools (e.g., monotonicity, bifurcation theory) on which the analysis presented in part II relies on.

Differently from linear systems, new phenomena arise when nonlinear dynamics is involved, such as e.g., multiple isolated equilibrium points, bifurcations, and limit cycles, see [40]. Investigating the behavior of a nonlinear system means to study these phenomena, in terms of existence and stability properties. An important class of nonlinear systems, whose asymptotic behavior has been extensively studied in the literature (see for instance [41–43]), is that of monotone systems, i.e., systems which generate an order preserving flow. The first section of this chapter gives a brief overview of nonlinear systems and their analysis, and introduces the class of monotone systems.

The second section of this chapter instead focuses on bifurcation theory, which studies how the qualitative behavior of a system changes under (small) variations of its parameters.

3.1 Preliminary definitions

Consider a nonlinear autonomous system

$$\dot{x}(t) = f(x(t)), \quad x(0) = x_0, \tag{3.1}$$

where $f : \mathcal{U} \rightarrow \mathbb{R}^n$ is a Lipschitz continuous function on an open convex subset $\mathcal{U} \subseteq \mathbb{R}^n$, that is, $\exists K \geq 0$ such that $\|f(x) - f(y)\| \leq K\|x - y\|$ for all $x, y \in \mathcal{U}$. Let $\varphi(t, x_0)$ be the solution $x(t)$ of (3.1) that satisfies $x(0) = x_0$; observe that existence and uniqueness of a solution of (3.1) are guaranteed by the Lipschitz condition imposed on f [40, Thm 3.2].

A point x^* is an *equilibrium point* of (3.1) if $f(x^*) = 0$. A nontrivial solution $x(t)$ of (3.1) is *periodic* of period T if there exists a $T > 0$ such that $x(t+T) = x(t)$ for all $t > 0$. A *periodic orbit* is the image of a periodic solution in the phase portrait. An isolated periodic orbit is called *limit cycle*.

Stability of an equilibrium point is characterized in Definition 3.1, and indicates whether solutions starting close to an equilibrium point will remain close to it (stable eq. point), converge to it (asymptotically stable eq. point), or get further away from it (unstable eq. point), as $t \rightarrow \infty$, see [40, Chapter 4]¹.

Definition 3.1. An equilibrium point x^* of (3.1) is:

- *stable* if, for each $\varepsilon > 0$, there is a $\delta > 0$ such that $\|x(0) - x^*\| < \delta$ implies that $\|x(t) - x^*\| < \varepsilon$ for all $t \geq 0$;
- *asymptotically stable* if it is stable and δ can be chosen such that $\|x(0) - x^*\| < \delta$ implies that $\lim_{t \rightarrow \infty} x(t) = x^*$;
- *unstable* if it is not stable.

Without loss of generality, hereafter it will be assumed that the system (3.1) has an equilibrium point in the origin ($x^* = 0 \in \mathcal{U}$), and all the definitions will be stated accordingly.

Lyapunov's theory provides sufficient conditions to determine the stability properties of an equilibrium point. The well-known Lyapunov's stability theorem states that if there exists a continuously differentiable *Lyapunov function*, i.e., a function that is positive definite in the domain and whose derivative along the trajectories of the system is negative semidefinite, then the equilibrium point is stable. Moreover, if the derivative is negative definite, then the equilibrium point is asymptotically stable.

Theorem 3.1 (4.1 in [40]). Let the origin be an equilibrium point for (3.1). Let $V : \mathcal{U} \rightarrow \mathbb{R}_+$ be a continuously differentiable function such that

- $V(0) = 0$ and $V(x) > 0$ for all $x \in \mathcal{U} \setminus \{0\}$;
- $\dot{V}(x) \leq 0$ for all $x \in \mathcal{U}$.

Then the origin is stable. Moreover, if $\dot{V}(x) < 0$ for all $x \in \mathcal{U} \setminus \{0\}$ then the origin is asymptotically stable.

In addition to a sufficient condition for asymptotic stability of the origin as an equilibrium point for (3.1), Lyapunov's theory can provide also an estimate (which is often, however, conservative) for its *region of attraction* (or, basin of attraction), defined as the set of points x_0 such that $\lim_{t \rightarrow \infty} \varphi(t, x_0) = 0$. Intuitively, the region of attraction indicates "how big" δ in Definition 3.1 can be for the origin to remain an attractive equilibrium point. In the particular case where the region of attraction is the whole \mathbb{R}^n , the origin is called *globally asymptotically stable*

¹Similarly, a limit cycle is stable (resp., unstable) if all the trajectories starting arbitrarily close to it will tend toward to (resp., away from) it as $t \rightarrow \infty$ [40, Chapter 8]. This section covers key results on stability of equilibrium points only. Stability of periodic orbits is not discussed, but an example of stable limit cycle will be shown in Section 3.2.

(GAS). A sufficient condition for the origin to be GAS is given by the Barbashin-Krasovskii theorem (see Theorem 3.2), namely, there exists a Lyapunov function which is radially unbounded.

Theorem 3.2 (4.2 in [40]). *Let the origin be an equilibrium point for (3.1). Let $V : \mathbb{R}^n \rightarrow \mathbb{R}_+$ be a continuously differentiable function such that*

- $V(0) = 0$ and $V(x) > 0$ for all $x \neq 0$;
- $\lim_{\|x\| \rightarrow +\infty} V(x) = +\infty$, i.e., V is radially unbounded;
- $\dot{V}(x) < 0$ for all $x \neq 0$.

Then the origin is globally asymptotically stable.

An important extension of these results is given by LaSalle's Invariant Set theory; a typical application is to establish the asymptotical stability of an equilibrium point when Lyapunov's theory fails (for instance, when asymptotic stability is expected but the derivative of the candidate Lyapunov function is only negative semidefinite). LaSalle's theory is built on the concept of invariant sets for the dynamical system. A set \mathcal{M} is called (*positively*, omitted hereafter) *invariant* w.r.t. (3.1) if every trajectory starting in \mathcal{M} remains in \mathcal{M} for all time, i.e., $x(0) \in \mathcal{M}$ implies $x(t) \in \mathcal{M}$ for all $t \geq 0$.

Theorem 3.3 (LaSalle, 4.4 in [40]). *Let $\Omega \subset \mathcal{U}$ be a compact set, which is positively invariant w.r.t. (3.1). Let $V : \mathcal{U} \rightarrow \mathbb{R}$ be a continuously differentiable function such that*

- $\dot{V}(x) \leq 0$ for all $x \in \Omega$;
- $\mathcal{Z} = \{x \in \Omega : \dot{V}(x) = 0\}$.

Then $\varphi(t, x_0)$ converges to \mathcal{M} for all $x_0 \in \Omega$, where \mathcal{M} is the largest invariant set in \mathcal{Z} .

A consequence of LaSalle's invariance principle is that the origin is an asymptotically stable equilibrium point of (3.1) if there exists a Lyapunov function V , and no solution (except for the trivial solution) can stay identically in the set $\mathcal{Z} = \{x \in \Omega : \dot{V}(x) = 0\}$ [40, Corollary 4.1].

Remark 3.1. In Theorems 3.1, 3.2 and 3.3, the hypothesis that the Lyapunov function V is continuously differentiable is not essential. However, if this assumption is removed, the derivative needs to be substituted by the upper Dini derivative of V in the theorems.

The following material on the upper Dini derivative is from [44]. The upper Dini derivative of a continuous function $r : (a, b) \rightarrow \mathbb{R}$, with $a, b \in \mathbb{R}$, at $t \in (a, b)$ is defined as

$$d^+ r(t) = \limsup_{s \rightarrow 0^+} \frac{r(s+t) - r(t)}{s}.$$

If r is Lipschitz on some neighborhood of t then the derivative is finite. Key result is that r is decreasing on (a, b) if and only if $d^+ r(t) \leq 0$ for all $t \in (a, b)$.

Now, let $V : \mathcal{U} \rightarrow \mathbb{R}$ be a continuous and locally Lipschitz function, and $x(t)$ be a solution of (3.1). The upper Dini derivative of $V(x(t))$ is defined accordingly to (3.1), $d^+ V(x(t)) = \limsup_{s \rightarrow 0^+} \frac{V(x(s+t)) - V(x(t))}{s}$. The upper Dini derivative of V along the

vector field f of (3.1), denoted d^+V or d_f^+V , is given by

$$d^+V(x) = \limsup_{s \rightarrow 0^+} \frac{V(x + sf(x)) - V(x)}{s},$$

and it holds that $d^+V(x)|_{x=x^*} = d^+V(x(t))|_{t=t^*}$ with $x(t^*) = x^*$. If $d^+V(x) \leq 0$ on \mathcal{U} , then V is decreasing along the solutions of (3.1). _____

The theorems discussed so far are main results of the *Lyapunov's direct method*, which focuses on finding a Lyapunov function V . The next theorem, denoted *Lyapunov's indirect method*, offers an alternative way to determine the stability of an equilibrium point, by means of linearization.

Theorem 3.4 (4.7 in [40]). *Let the origin be an equilibrium point for (3.1), and assume that f is continuously differentiable in \mathcal{U} . Let $f_x(x) := \frac{\partial f}{\partial x}(x)$ indicate the Jacobian of (3.1) evaluated at x . Then the origin is:*

- *asymptotically stable if $f_x(0)$ is Hurwitz, i.e., $\text{Re}[\lambda] < 0$ for all $\lambda \in \Lambda(f_x(0))$;*
- *unstable if there exists at least one eigenvalue of $f_x(0)$ with positive real part.*

This method fails when $f_x(0)$ has eigenvalues with nonpositive real part and at least one eigenvalue with zero real part. In this case, more advanced tools from *center manifold* theory can be used to study the stability properties of the origin [40, Chapter 8]. It is out of the scope of this thesis work to detail center manifold theory, however a brief overview is given in what follows, which will be useful in Section 3.2. Assume that the origin is an equilibrium point of the system (3.1). The generalized eigenspaces of $f_x(0)$ relative to eigenvalues with strictly negative, zero, and strictly positive real part are denoted stable, center, and unstable generalized eigenspaces, respectively. Center manifold theory states that it is possible to define three manifolds \mathcal{W}_i , $i = s, c, u$, denoted the *stable*, *center* and *unstable* manifolds, respectively, which are invariant w.r.t. (3.1) and tangent to the generalized stable, center and unstable eigenspaces eigenspaces of $f_x(0)$ at 0, respectively. When there is no unstable manifold, the stability properties of the origin as an equilibrium point of (3.1) can be investigated by studying the dynamics of (3.1) restricted to the center manifold \mathcal{W}_c , referred to as the *reduced system*. If the origin is an asymptotically stable (resp., unstable) equilibrium point of the reduced dynamics then it is also an asymptotically stable (resp., unstable) equilibrium point of the original (higher-order) system [40, 45].

Discrete-time systems The notions and results introduced in Section 3.1 for continuous-time systems admit a natural extension and can be restated for discrete-time nonlinear systems, represented as

$$x(k+1) = f(x(k)), \quad x(0) = x_0. \tag{3.2}$$

For instance, a point x^* is an *equilibrium point* of (3.2) if $f(x^*) = x^*$. Provided that the derivative of V is substituted with the increment of V along the trajectories, $V_\Delta(x) := V(f(x)) - V(x)$, Theorems 3.1 and 3.2 can be used to provide sufficient

conditions for stability and asymptotical stability, and global asymptotic stability, respectively, of an equilibrium point. The discrete-time version of Lyapunov's indirect method can be stated as follows.

Theorem 3.5. *Let the origin be an equilibrium point for (3.2), and assume that f is continuously differentiable in \mathcal{U} . Let $f_x(x) := \frac{\partial f}{\partial x}(x)$ indicate the Jacobian of (3.2) evaluated at x . Then the origin is:*

- *asymptotically stable if all the eigenvalues of $f_x(0)$ are inside the unit circle, i.e., $|\lambda| < 1$ for all $\lambda \in \Lambda(f_x(0))$;*
- *unstable if there exists at least one eigenvalue of $f_x(0)$ outside the unit circle, i.e., $\exists \lambda \in \Lambda(f_x(0))$ such that $|\lambda| > 1$.*

When $f_x(0)$ has eigenvalues inside the unit circle and at least one eigenvalue on the unit circle, tools from center manifold theory (as in the continuous-time case) can be used to study the stability properties of the origin as an equilibrium point of the system (3.2), see for instance [46].

3.1.1 Cooperativity and Monotonicity

An interesting class of nonlinear systems is that of monotone systems, whose characteristic asymptotic properties have motivated their comprehensive analysis in the literature.

Consider a nonlinear system, described as (3.1). To introduce the concept of monotonicity, the notions of orthant of \mathbb{R}^n and of partial ordering generated by an orthant are needed. Let $S \in \mathbb{R}^{n \times n}$ be a signature matrix, i.e., $S = \text{diag}\{s_1, \dots, s_n\}$ with $s_i = \pm 1$ for all $i = 1, \dots, n$, and let $S\mathbb{R}^n$ indicate an orthant of \mathbb{R}^n , $S\mathbb{R}^n = \{x \in \mathbb{R}^n : s_i x_i \geq 0, i = 1, \dots, n\}$. Let $\leq_{S\mathbb{R}^n}$ indicate the partial ordering generated by $S\mathbb{R}^n$, i.e., $x \leq_{S\mathbb{R}^n} y$ if and only if $y - x \in S\mathbb{R}^n$.

Definition 3.2. The partial ordering $\leq_{S\mathbb{R}^n}$ is preserved by the solution operator $\varphi(t, \cdot)$ and the system (3.1) is *type $S\mathbb{R}^n$ monotone* if whenever $x, y \in \mathcal{U}$ and $x \leq_{S\mathbb{R}^n} y$ then $\varphi(t, x) \leq_{S\mathbb{R}^n} \varphi(t, y)$ for all $t \geq 0$. ———

Lemma 3.1 (2.1 in [42]). *If $f \in C^1(\mathcal{U})$ where \mathcal{U} is open and convex in \mathbb{R}^n then $\varphi(t, \cdot)$ preserves the partial ordering $\leq_{S\mathbb{R}^n}$ for $t \geq 0$ if and only if $S \frac{\partial f}{\partial x}(x)S$ has nonnegative off-diagonal elements for every $x \in \mathcal{U}$.*

As previously stated, the importance behind the analysis of monotone systems lies in their asymptotic behavior. In particular, it is known that if all the trajectories of a monotone system are bounded, then they generically converge to an equilibrium point, and that there are no attracting periodic orbits other than equilibria [43].

Cooperativity is a special case of monotonicity corresponding to $S\mathbb{R}^n = \mathbb{R}_+^n$, where \mathbb{R}_+^n is the nonnegative orthant. The partial ordering generated by \mathbb{R}_+^n is denoted \leq .

Definition 3.3. The partial ordering \leq is preserved by the solution operator $\varphi(t, \cdot)$ and the system (3.1) is *cooperative* if whenever $x, y \in \mathcal{U}$ and $x \leq y$ then $\varphi(t, x) \leq \varphi(t, y)$ for all $t \geq 0$. ———

The next theorem provides a necessary and sufficient condition for a vector field f to determine a cooperative system.

Theorem 3.6 (12.11 in [47]). *The system (3.1) is cooperative if and only if it satisfies the Kamke-Muller condition:*

$$x \leq y \Rightarrow f_i(x) \leq f_i(y) \quad \forall i \text{ such that } x_i = y_i$$

If $f \in C^1(\mathcal{U})$, a necessary and sufficient condition is that $\frac{\partial f}{\partial x}(x)$ has nonnegative off-diagonal elements for every $x \in \mathcal{U}$.

Remark 3.2. Observe that in Theorem 3.6, the necessary and sufficient condition for a differentiable vector field f to determine a cooperative system is that $\frac{\partial f}{\partial x}(x)$ is Metzler (in the domain). This implies that, if the Jacobian of (3.1) is Metzler, then \mathbb{R}_+^n is positively invariant w.r.t. (3.1). ———

3.2 Bifurcation analysis

In this section a brief overview of bifurcation analysis for nonlinear autonomous systems is presented, both in continuous-time and discrete-time. The main idea is that when a nonlinear system is dependent on a parameter, its qualitative behavior may change as the value of the parameter varies. A *bifurcation* is defined as a change in the number of equilibrium points or periodic orbits, or in their stability properties, as a parameter is varied. The parameter is called *bifurcation parameter*, and the points at which changes occur are called *bifurcation points*. A visual representation of a bifurcation is given by a *bifurcation diagram* [40].

A bifurcation can occur only when, for a certain value of the bifurcation parameter, the Jacobian of the system at an equilibrium point has (at least) one eigenvalue on the imaginary axis (in continuous-time systems), or has (at least) one eigenvalue of the unit circle (in discrete-time systems), see Fig. 3.1, which implies that the equilibrium point has a center manifold²; for simplicity, this thesis work considers only the class of steady-state bifurcations for which the center manifold has dimension one.

From center manifold theory, in order to investigate the qualitative behavior of a nonlinear (in general, n -dimensional) system near a bifurcation point it is sufficient to study the center manifold (scalar) dynamics. A key concept of bifurcation theory is that each steady-state bifurcation can be represented by a canonical scalar system, called *normal form*, and that every problem can be reduced to a specific normal form exhibiting the same qualitative behavior under variation of a parameter.

This chapter provides a brief introduction on the normal forms of (steady-state) bifurcations that are of interest for this thesis work, and, later, on the procedure to reduce general (higher-dimensional) problems to these canonical forms. It focuses specifically on systems that have an odd symmetry, denoted \mathbf{Z}_2 -equivariant

²In this case, the equilibrium point is called *nonhyperbolic* [45, 46]. However, this notation is not used in the papers of Part II.

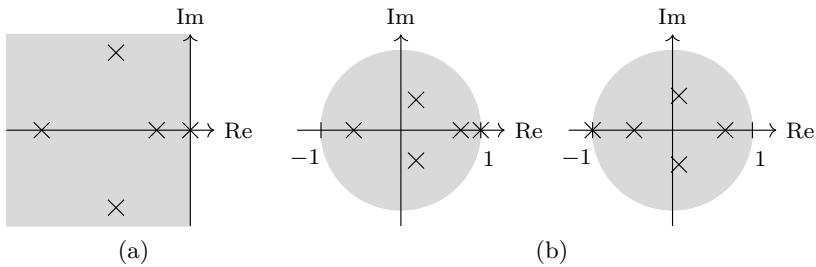


Figure 3.1: Necessary condition for occurrence of a bifurcation at an equilibrium point. (a): The Jacobian at the equilibrium point has eigenvalues with zero real part (continuous-time systems). (b): The Jacobian at the equilibrium point has eigenvalues on the unit circle (discrete-time systems). Legend: crosses indicate eigenvalues of the Jacobian at an equilibrium point.

systems [48, 49], which exhibit a specific type of steady-state bifurcation, denoted *pitchfork bifurcation*.

Since the (stability) characterization of equilibrium points for continuous-time and discrete-time systems is different, these cases will be treated separately. The material presented in this chapter is inspired mostly by [45, 46, 48].

3.2.1 Continuous-time systems: pitchfork bifurcation

Consider the nonlinear system

$$\dot{x} = f(x, \pi), \quad x(0) = x_0, \quad (3.3)$$

where $x \in \mathbb{R}^n$, $\pi \in \mathbb{R}$ is a scalar parameter, and $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is infinitely differentiable everywhere. Let $f_x(x, \pi) := \frac{\partial f}{\partial x}(x, \pi)$ indicate the Jacobian of (3.3) evaluated at (x, π) . A point x^* is an equilibrium point of (3.3) at $\pi = \pi^*$ if $f(x^*, \pi^*) = 0$.

Definition 3.4. The parameter π is called *bifurcation parameter*, and the set of (x, π) satisfying $f(x, \pi) = 0$ *bifurcation diagram*. A point (x^*, π^*) is called *bifurcation point* if the number of solutions of $f(x, \pi) = 0$ changes as π varies in a neighborhood of π^* .

Necessary conditions for a bifurcation to occur at a point (x^*, π^*) are that $f(x^*, \pi^*) = 0$, and $f_x(x^*, \pi^*)$ has an eigenvalue on the imaginary axis. Without loss of generality, in what follows it is assumed that $(x^*, \pi^*) = (0, 0)$, and that the zero eigenvalue of $f_x(0, 0)$ is real and *simple*³.

This section focuses on a specific type of (steady-state) bifurcation, called *pitchfork bifurcation*, characterized by the fact that at the bifurcation point the number of solutions of $f(x, \pi) = 0$ jumps from one to three. This type of bifurcation is ubiquitous of systems that have *odd symmetry*, i.e., systems for which $f(x, \pi) = -f(-x, \pi)$ for all $x \in \mathbb{R}^n$.

³The case in which $f_x(0, 0)$ has complex conjugate eigenvalues on the imaginary axis is not considered.

Normal form of a pitchfork bifurcation

The normal form of a pitchfork bifurcation is described by the scalar nonlinear system

$$\dot{x} = \pi x - x^3, \quad x \in \mathbb{R}, \pi \in \mathbb{R}. \quad (3.4)$$

Let $g(x, \pi) = -x^3 + \pi x$ and $g_x(x, \pi) = -3x^2 + \pi$. Observe that the system (3.4) has odd symmetry since g is odd in x .

The origin is an equilibrium point of the system (3.4) for all values of $\pi \in \mathbb{R}$. It is globally asymptotically stable when $\pi \leq 0$, and unstable when $\pi > 0$. When $\pi > 0$ the system admits two new equilibrium points in $\pm\sqrt{\pi}$, which are locally asymptotically stable since $g_x(\pm\sqrt{\pi}, \pi) < 0$ for all $\pi > 0$. This means that the system (3.4) undergoes a pitchfork bifurcation at $\pi = 0$, and the corresponding bifurcation diagram is depicted in Figure 3.2⁴.

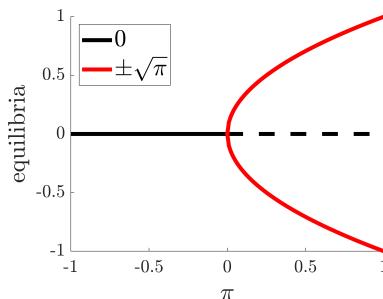


Figure 3.2: Bifurcation diagram of a pitchfork bifurcation, whose normal form is shown in (3.4). Solid lines indicate stable equilibrium points, while a dashed line unstable equilibrium points.

How to recognize a pitchfork bifurcation

In the previous section, the normal form of a pitchfork bifurcation has been introduced. In general if a bifurcation problem, represented by a nonlinear system

$$\dot{x} = g(x, \pi), \quad x \in \mathbb{R}, \pi \in \mathbb{R}, \quad (3.5)$$

at a point $(x, \pi) = (0, 0)$ satisfies the conditions

$$g = g_x = g_{xx} = g_\pi = 0, \quad g_{xxx} < 0, \quad g_{\pi x} > 0, \quad (3.6)$$

then it is said to be *equivalent* to the normal form (3.4), and (3.6) is said to solve the recognition problem for the pitchfork bifurcation⁵. Intuitively, this means that there exists an adequate change of coordinates that transforms the problem

⁴Figure 3.2 shows a *supercritical* pitchfork bifurcation. A *subcritical* pitchfork bifurcation instead is characterized by the number of equilibria jumping from 3 to 1, and is described by the normal form $g(x, \pi) = \pi x + x^3 = 0$. This thesis work focuses only on the supercritical pitchfork bifurcation, therefore the term “supercritical” will be dropped from now on.

⁵The subscripts in (3.6) indicate partial derivatives.

$g(x, \pi) = 0$ into the standard form of a pitchfork, $-x^3 + \pi x = 0$, see [48, §1] for details. Therefore, the number of solutions of $g(x, \pi) = 0$ jumps from one to three as π crosses 0.

Similarly, every nonlinear model that exhibits a pitchfork bifurcation behavior can be reduced to a scalar differential equation $\dot{x} = g(x, \pi)$ that satisfies (3.6). A method to obtain such scalar equation from a general n -dimensional system is the *Lyapunov-Schmidt reduction*, explained in [48, §3] and used in Papers B and C.

3.2.2 Discrete-time systems: pitchfork and period-doubling bifurcations

Consider the nonlinear system

$$x(k+1) = f(x(k), \pi), \quad x(0) = x_0, \quad (3.7)$$

where $x \in \mathbb{R}^n$, $\pi \in \mathbb{R}$ is a scalar parameter, and $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is infinitely differentiable everywhere. Let $f_x(x, \pi) := \frac{\partial f}{\partial x}(x, \pi)$ indicate the Jacobian of (3.7) evaluated at (x, π) . A point x^* is an equilibrium point of (3.7) at $\pi = \pi^*$ if $f(x^*, \pi^*) = x^*$.

Definition 3.5. The parameter π is called *bifurcation parameter*, and the set of (x, π) satisfying $f(x, \pi) = x$ *bifurcation diagram*. A point (x^*, π^*) is called *bifurcation point* if the number of solutions of $f(x, \pi) = x$ changes as π varies in a neighborhood of π^* .

Necessary conditions for a bifurcation to occur at a point (x^*, π^*) are that $f(x^*, \pi^*) = x^*$, and $f_x(x^*, \pi^*)$ has an eigenvalue on the unit circle. Again, without loss of generality, in what follows it is assumed that $(x^*, \pi^*) = (0, 0)$, and that the eigenvalue of $f_x(0, 0)$ on the unit circle is real and *simple*, that is, $\exists \lambda \in \Lambda(f_x(0, 0))$ such that either $\lambda = +1$ or $\lambda = -1$.

The type of bifurcation the system (3.7) undergoes at $(0, 0)$ depends on the sign of the eigenvalue λ . The next paragraphs introduce the normal forms of a pitchfork bifurcation ($\lambda = 1$, assuming that the system has odd symmetry), and of a period-doubling bifurcation ($\lambda = -1$).

Normal form of a pitchfork bifurcation

The normal form of a pitchfork bifurcation is described by the scalar nonlinear system

$$x(k+1) = (1 + \pi)x(k) - x(k)^3, \quad x \in \mathbb{R}, \pi \in \mathbb{R}. \quad (3.8)$$

Let $g(x, \pi) = (1 + \pi)x - x^3$ and $g_x(x, \pi) = (1 + \pi) - 3x^2$. The system (3.8) has odd symmetry since g is odd in x .

The analysis is similar to the continuous-time case. The origin is an equilibrium point of the system (3.4) for all values of $\pi \in \mathbb{R}$, and it is stable when $\pi \leq 0$ (and $|\pi|$ small), and unstable when $\pi > 0$. When $\pi > 0$ the system admits two new equilibrium points in $\pm\sqrt{\pi}$, which are locally asymptotically stable since $g_x(\pm\sqrt{\pi}, \pi) < 1$ for all $\pi > 0$ (and π small).

This means that the system (3.4) undergoes a pitchfork bifurcation at $\pi = 0$, and the corresponding bifurcation diagram is the same as Fig. 3.2.

Normal form of a period-doubling bifurcation

The normal form of a period-doubling bifurcation is described by the scalar nonlinear system

$$x(k+1) = -(1+\pi)x(k) + x(k)^3, \quad x \in \mathbb{R}, \pi \in \mathbb{R}. \quad (3.9)$$

Let $g(x, \pi) = -(1+\pi)x + x^3$ and $g_x(x, \pi) = -(1+\pi) + 3x^2$. The origin is an equilibrium point of the system (3.9) for all values of $\pi \in \mathbb{R}$. It is asymptotically stable when $\pi < 0$ (and $|\pi|$ small), and unstable when $\pi > 0$.

For values of π near 0 the system (3.9) does not admit any other equilibrium point. Instead, as the name suggests, the characteristic of a period-doubling bifurcation is that it produces an oscillation of period 2. To understand, one needs to consider the second iterate of the system (3.9),

$$x(k+2) = g^2(x(k), \pi) \quad (3.10)$$

described by

$$g^2(x, \pi) := g(g(x, \pi), \pi) = (1+\pi)^2 x - (1+\pi)(2+2\pi+\pi^2)x^3 + O(x^5),$$

and observe that a nontrivial equilibrium point of (3.10) corresponds to a periodic oscillation of (3.9) of period 2.

The bifurcation analysis of the system (3.10), see Fig. 3.3, shows that (3.10) undergoes a pitchfork bifurcation from the origin at $\pi = 0$, which implies that (3.9) admits a stable limit cycle of period 2 when π crosses 0.

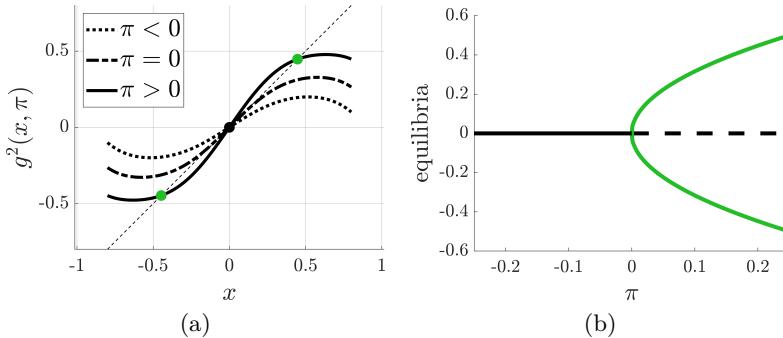


Figure 3.3: Analysis of the second iterate (3.10) near $(0,0)$. (a): $g^2(x, \pi)$ near $(0,0)$ for different values of π . (b): Bifurcation diagram of (3.10).

How to recognize a pitchfork or a period-doubling bifurcation

In the previous section, the normal forms of a pitchfork and a period-doubling bifurcation have been introduced. In general if a bifurcation problem, represented by a scalar nonlinear system

$$x(k+1) = g(x(k), \pi), \quad x \in \mathbb{R}, \pi \in \mathbb{R}, \quad (3.11)$$

at a point $(x, \pi) = (0, 0)$ satisfies the conditions

$$g = 0, \quad g_x = 1, \quad g_{xx} = 0, \quad g_\pi = 0, \quad g_{xxx} < 0, \quad g_{\pi x} > 0, \quad (3.12)$$

then it is equivalent to the normal form of a pitchfork bifurcation (3.8). If, instead, it satisfies the conditions

$$g = 0, \quad g_x = -1, \quad g_{xx} = 0, \quad g_\pi = 0, \quad g_{xxx} > 0, \quad g_{\pi x} < 0, \quad (3.13)$$

then it is equivalent to the normal form of a period-doubling bifurcation (3.9).

In n -dimensional systems (3.7), the bifurcation analysis reduces to that of the dynamics along the center manifold of the equilibrium point. The details can be found in [46, Chapter 5].

4

Theory of graphs and social networks

Graphs are the natural mathematical structure to model networks of interacting agents. For instance, a social network can be modeled as a graph where the nodes represent individuals, and the edges the social interactions between them [10]. The general aim of this chapter is to introduce key notions and results from *graph theory*, i.e., the study of networks structure, and the interpretation will be given in terms of social networks. In particular, the second part of this chapter is dedicated to the theory of *signed graphs*, which are used to model networks where the interactions between the agents are not restricted to be friendly or cooperative, but can be hostile or antagonistic.

The main references for the material of this chapter are [50–52] and [14, 16, 53, 54], for graph and signed graph theory, respectively. In the rest of this thesis work the terms graphs and networks will be used interchangeably.

4.1 Graphs

A graph \mathcal{G} is defined as a triple $\mathcal{G} = (\mathcal{V}, \mathcal{E}, A)$ where $\mathcal{V} = \{1, \dots, n\}$ is a finite set of nodes, $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ is a set of edges, and $A = [a_{ij}] \in \mathbb{R}_+^{n \times n}$ is the adjacency matrix of \mathcal{G} , i.e., $a_{ij} > 0$ if and only if $(j, i) \in \mathcal{E}$. A graph is *unweighted* if $a_{ij} \in \{0, 1\}$ for all $i, j \in \mathcal{V}$, and it is *weighted* otherwise.

A *self-loop* is an edge from a node to itself.

A graph \mathcal{G} is called *undirected* if \mathcal{E} is a set of unordered edges, i.e., if $(i, j) \in \mathcal{E}$ implies that $(j, i) \in \mathcal{E}$, and it is called *directed* (or, a *digraph*) if \mathcal{E} is a set of ordered edges. In this case, $(i, j) \in \mathcal{E}$ denotes an edge from i to j .

A (*directed*) *path* in a graph is an ordered sequence of nodes such that any pair of consecutive nodes in the sequence is a (directed) edge of the graph. It is *simple* if all nodes in the (directed) path are distinct. A graph is (*strongly*) *connected* if there is a (directed) path between any two distinct nodes. A *cycle* is a simple

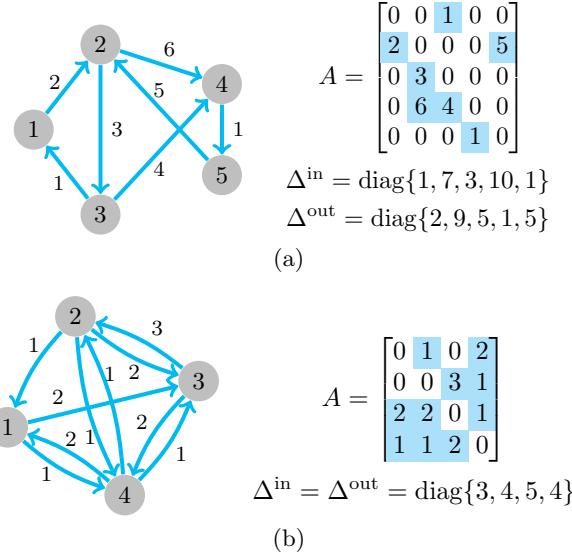


Figure 4.1: Examples of simple, weighted graphs. (a): Strongly connected digraph with $\text{card}(\mathcal{V}) = 5$ nodes and 7 directed edges, and corresponding adjacency, in-degree and out-degree matrices. (b): Weight balanced graph with $\text{card}(\mathcal{V}) = 4$ nodes and 10 directed edges, and corresponding adjacency, in-degree and out-degree matrices.

(directed) path that starts and ends at the same node.

A node j is a *in-neighbor* of i (or, j is *adjacent* to i) if $(j, i) \in \mathcal{E}$, and is a *out-neighbor* of i if $(i, j) \in \mathcal{E}$, and the in-neighborhood $\mathcal{N}_i^{\text{in}}$ and out-neighborhood $\mathcal{N}_i^{\text{out}}$ of a node i indicate the set of all in- and out-neighbors of i , respectively. A node is called *isolated* if its in-neighborhood and out-neighborhood are both empty sets. A graph is *complete* if every pair of nodes is adjacent, i.e., $\mathcal{N}_i^{\text{in}} = \mathcal{N}_i^{\text{out}} = \mathcal{V} \setminus \{i\}$ for all $i \in \mathcal{V}$.

The *in-degree* and *out-degree* of a node $i \in \mathcal{V}$ are defined by $\delta_i^{\text{in}} = \sum_{j=1}^n a_{ij}$ and $\delta_i^{\text{out}} = \sum_{j=1}^n a_{ji}$, respectively. A node is called *isolated* if its in-degree and out-degree are both zero (i.e., $\delta_i^{\text{in}} = \delta_i^{\text{out}} = 0$). A graph is called *weight-balanced* if in-degree and out-degree coincide for each node. If a graph is undirected then it is also weight-balanced and in this case the superscripts “in” and “out” are dropped ($\delta_i^{\text{in}} = \delta_i^{\text{out}} =: \delta_i$ for all i). The in-degree matrix and out-degree matrix of a graph are defined by $\Delta^{\text{in}} = \text{diag}\{\delta_1^{\text{in}}, \dots, \delta_n^{\text{in}}\}$ and $\Delta^{\text{out}} = \text{diag}\{\delta_1^{\text{out}}, \dots, \delta_n^{\text{out}}\}$, respectively. This thesis work mostly deals with the notion of in-degrees and, when there is no ambiguity, the notation δ_i and Δ is used to indicate the in-degree of a node i and the in-degree matrix, respectively.

These concepts are illustrated in the examples of Figure 4.1.

Remark 4.1. The adjacency matrix of a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, A)$ has the following properties:

- \mathcal{G} is without self-loops if and only if A has zero diagonal elements;
- \mathcal{G} has an isolated node i if and only if the i -th row and i -th column of A are zero;
- \mathcal{G} is undirected if and only if $A = A^T$;

- \mathcal{G} is connected/strongly connected if and only if A is irreducible, i.e., if there does not exist a permutation matrix P such that $P^T A P$ is block triangular;
- \mathcal{G} is weight-balanced if and only if $A\mathbf{1} = A^T \mathbf{1}$, where $\mathbf{1}$ is the vector of all 1.

Remark 4.2. This thesis work considers graphs that are (strongly) connected and without self-loops.

4.1.1 Laplacian of a graph and its properties

This section introduces the definition of Laplacian matrix of a graph. The definitions and lemmas are stated for the general case of directed graphs; however, they hold also for undirected graphs.

The *Laplacian* of a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, A)$, with $\text{card}(\mathcal{V}) = n$, is the matrix $L \in \mathbb{R}^{n \times n}$ defined by

$$[L]_{ij} = \begin{cases} \sum_{k=1}^n a_{ik}, & j = i \\ -a_{ij}, & j \neq i. \end{cases} \quad (4.1)$$

In compact form, $L = \Delta^{\text{in}} - A$.

The following lemma is a collection of fundamental and well-known properties of the Laplacian matrix, see for instance [51, 52, 55–57]. Intuitively, (i)÷(vi) follow directly from the definition of Laplacian, the Geršgorin's Theorem, and the Perron-Frobenius Theorem, and (vii) is shown in [57, Theorem 2].

Lemma 4.1. *The Laplacian L of a digraph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, A)$, with $\text{card}(\mathcal{V}) = n$, has the following properties:*

- (i) $L\mathbf{1} = 0$, i.e., 0 is always an eigenvalue of L with right eigenvector $\mathbf{1}$;
- (ii) all the eigenvalues of L have nonnegative real part, i.e., $\text{Re}[\lambda(L)] \geq 0$ for all $\lambda(L) \in \Lambda(L)$;
- (iii) L is a singular M -matrix.

Moreover, if the digraph \mathcal{G} is strongly connected, then:

- (iv) L is a singular and irreducible M -matrix;
- (v) the zero eigenvalue of L is simple, $\text{rank}(L) = n - 1$ and $\ker(L) = \text{span}\{\mathbf{1}\}$;
- (vi) the left eigenvector of L associated with the zero eigenvalue is positive, i.e., $\xi > 0$ with $L^T \xi = 0$ and $\ker(L^T) = \text{span}\{\xi\}$;
- (vii) $\xi > 0$ is the unique vector (up to a scaling factor) such that $\Xi L + L^T \Xi$ is positive semidefinite, with $\Xi := \text{diag}\{\xi\}$, and $\text{rank}(\Xi L + L^T \Xi) = n - 1$.

Something more can be added when a digraph is not only strongly connected but also weight-balanced, as shown for instance in the recent work [57].

Lemma 4.2. *Let $\mathcal{G} = (\mathcal{V}, \mathcal{E}, A)$, with $\text{card}(\mathcal{V}) = n$, be a strongly connected digraph with Laplacian L . The following statements are equivalent:*

- (i) \mathcal{G} is weight-balanced;
- (ii) $L^T \mathbf{1} = 0$ and, therefore, $\ker(L) = \ker(L^T) = \text{span}\{\mathbf{1}\}$;

- (iii) the symmetric part of L , $L_s := \frac{L+L^T}{2}$, is positive semidefinite and $\text{rank}(L_s) = n - 1$.

When a digraph \mathcal{G} is strongly connected then it does not have isolated nodes, which means that the in-degree matrix Δ^{in} is positive definite. Therefore, its inverse $(\Delta^{\text{in}})^{-1}$ is well-defined and positive definite. The *normalized Laplacian* of a strongly connected digraph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, A)$ is the matrix $\mathcal{L} \in \mathbb{R}^{n \times n}$ defined by

$$[\mathcal{L}]_{ij} = \begin{cases} 1, & j = i \\ -\frac{a_{ij}}{\sum_{k=1}^n a_{ik}}, & j \neq i. \end{cases} \quad (4.2)$$

In compact form, $\mathcal{L} = I - (\Delta^{\text{in}})^{-1}A$. The normalized Laplacian L is a Laplacian matrix, therefore it satisfies the same properties of L listed in Lemma 4.1. For completeness, its key properties are summarized in the following lemma: in this case, a bound for the eigenvalues of \mathcal{L} can be provided using the Geršgorin's Theorem.

Lemma 4.3. *The normalized Laplacian \mathcal{L} of a strongly connected digraph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, A)$, with $\text{card}(\mathcal{V}) = n$, has the following properties:*

- (i) $\mathcal{L}\mathbf{1} = 0$, i.e., 0 is always an eigenvalue of \mathcal{L} with right eigenvector $\mathbf{1}$;
- (ii) the zero eigenvalue of \mathcal{L} is simple, $\text{rank}(\mathcal{L}) = n - 1$ and $\ker(\mathcal{L}) = \text{span}\{\mathbf{1}\}$;
- (iii) every eigenvalue of \mathcal{L} lies in the closed bounded disk of center 1 and radius 1 (i.e., $|\lambda - 1| \leq 1$ for all $\lambda \in \Lambda(\mathcal{L})$);
- (iv) \mathcal{L} is a singular and irreducible M-matrix.

4.2 Signed graphs

A signed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, A)$ is a graph where each edge in \mathcal{E} is associated with a positive or negative sign. The corresponding adjacency matrix is a signed matrix, i.e., $\text{sign}(a_{ij}) = \pm 1$ for all $(j, i) \in \mathcal{E}$, see Fig. 4.2a. A signed graph is *unweighted* if $a_{ij} \in \{-1, 0, 1\}$ for all $i, j \in \mathcal{V}$, and it is *weighted* otherwise. In this setting the graphs described in Section 4.1, i.e., graphs whose edges have all positive weights or, equivalently, whose adjacency matrix is nonnegative, are also denoted *unsigned graphs*.

Except for the notions of in-degrees and out-degrees, all the definitions introduced in Section 4.1 hold also for graphs that are signed. The notions of signed (in- and out-)degrees will be discussed in Remark 4.3 of Section 4.2.1.

4.2.1 Signed Laplacian matrices

In the literature of signed graphs different definitions of signed Laplacian are proposed, see [16, 54]. Using the notation introduced in [54], the *opposing signed Laplacian* L^o is defined as

$$[L^o]_{ij} = \begin{cases} \sum_{k=1}^n |a_{ik}|, & j = i \\ -a_{ij}, & j \neq i. \end{cases} \quad (4.3)$$

while the *repelling signed Laplacian* L^r is defined as

$$[L^r]_{ij} = \begin{cases} \sum_{k=1}^n a_{ik}, & j = i \\ -a_{ij}, & j \neq i, \end{cases} \quad (4.4)$$

When it is clear from context, the superscripts “ o ” and “ r ” will be dropped.¹

Compared with the Laplacian of an unsigned graph introduced in Section 4.1.1, here denoted *unsigned Laplacian*, the signed Laplacian matrices may or may not satisfy the properties listed in Lemma 4.1, as summarized and illustrated in Table 4.1. Counterexamples, see for instance Figure 4.2, can be found to show that a certain property does not hold in the signed case.

It is possible to define a signed Laplacian matrix, i.e., a matrix that induces a simple weighted signed digraph.

Definition 4.1. A matrix $L \in \mathbb{R}^{n \times n}$ ($n \geq 2$) is a:

- *opposing signed Laplacian* matrix if
 - (i) L has nonnegative diagonal elements;
 - (ii) L is diagonally equipotent.
- *repelling signed Laplacian* matrix if
 - (i) $L\mathbf{1} = 0$.

A signed Laplacian matrix L (repelling or opposing) induces a simple weighted signed digraph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, A)$ with $\text{card}(\mathcal{V}) = n$, where $(i, j) \in \mathcal{E}$ if and only if $[L]_{ji} \neq 0$, and $[A]_{ij} = -[L]_{ij}$ if $j \neq i$ or $[A]_{ij} = 0$ otherwise. ———

Remark 4.3. As for the notion of signed Laplacian, in this thesis work two alternative definitions for the signed in-degrees and out-degrees of a signed graph \mathcal{G} are proposed. The *in-degree* and *out-degree* of a node $i \in \mathcal{V}$ are defined either by

$$\delta_i^{\text{in}} = \sum_{j=1}^n |a_{ij}|, \quad \delta_i^{\text{out}} = \sum_{j=1}^n |a_{ji}|, \quad (4.5)$$

or

$$\delta_i^{\text{in}} = \sum_{j=1}^n a_{ij}, \quad \delta_i^{\text{out}} = \sum_{j=1}^n a_{ji} \quad (4.6)$$

in the context of the opposing or repelling signed Laplacian, respectively. Therefore, the signed Laplacian of a graph, given by (4.3) or (4.4), can be written in compact form as $L = \Delta^{\text{in}} - A$, where Δ^{in} is the signed in-degree matrix with diagonal elements given by (4.5) or (4.6), respectively.

¹To avoid confusion in the papers of part II, the following legend holds:

- Paper A and Paper B consider unsigned graphs, and use the notion of unsigned Laplacian (4.1) and its normalization (4.2);
- Paper C and Paper D consider signed graphs, and use the notion of opposing signed Laplacian (4.3) and its normalization (4.7), defined in Section 4.3.1;
- Paper E considers signed graphs, and uses the notion of repelling signed Laplacian (4.4).

Notice that once a definition for the signed in-degrees and out-degrees is adopted, the notion of weight balanced signed graph is equivalent to the one introduced in Section 4.1, namely, if signed in-degree and out-degree coincide for each node.

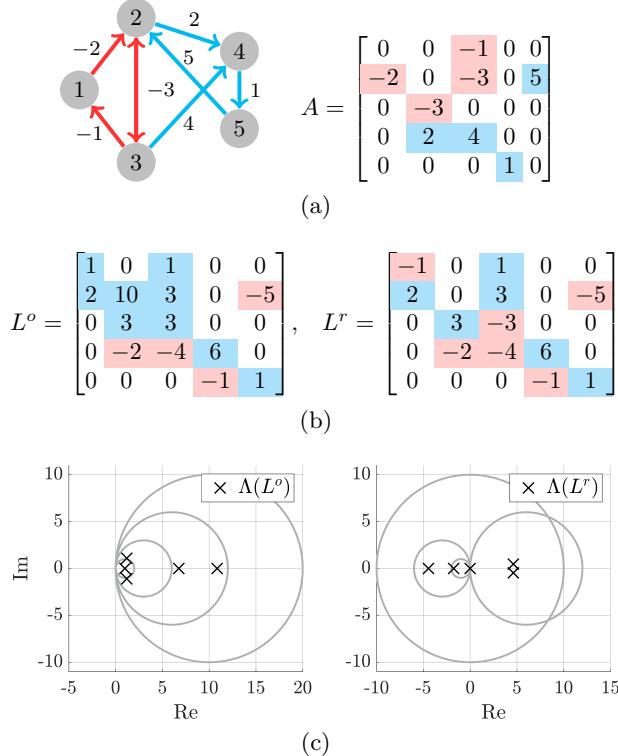


Figure 4.2: Example of a weighted signed digraph. (a): Strongly connected digraph with $\text{card}(\mathcal{V}) = 5$ nodes and 8 directed edges, and associated adjacency matrix. A red color is used to indicate edges with negative weights, and a blue color to indicate edges with positive weights. (b): Repelling and opposing signed Laplacian. Observe that only L^o is diagonally dominant. (c): Gershgorin's disks and eigenvalues of L^o (left) and L^r (right). Observe that $0 \in \Lambda(L^r)$ but $0 \notin \Lambda(L^o)$, and that while all the eigenvalues of L^o have positive real part, L^r has two eigenvalues with negative real part.

L	L^o	L^r	Property
■	■	■	0 is an eigenvalue
■	■	■	all the eigenvalues have nonnegative real part
■	■	■	it is diagonally dominant

Table 4.1: Properties of the opposing and repelling signed Laplacian compared with the unsigned Laplacian. Legend: ■ = the property is always satisfied, ■ = the property may or may not be satisfied.

4.3 Social networks as signed graphs

In the context of social networks, edges with positive weights represent friendship or cooperation, while edges with negative weights represent hostility or competition. The notion of structurally balanced network, first introduced in the field of social psychology by Heider [58] and then translated into the setting of signed graph theory by Harary and Cartwright [15, 59], represents a situation where, even if there are antagonistic interactions between the agents, at a network level there is no social tension. There are only two ways a social network can be structurally balanced: if, trivially, there is no antagonism between the individuals; or, if it is possible to divide the group of agents into two groups of mutual friends linked only by hostile interactions [10]. When a network is not structurally balanced, measures are proposed to characterize its distance from a structurally balanced state, see e.g., *line index of balance* introduced by Harary [17]. In this section these concepts are formulated in terms of (undirected) signed graphs.

4.3.1 Structurally balanced graphs and frustration index

Let $\mathcal{G} = (\mathcal{V}, \mathcal{E}, A)$ be a signed, undirected, and simple graph.

Definition 4.2. A signed graph is *structurally balanced* if there exists a partition of the node set $\mathcal{V} = \mathcal{V}_1 \cup \mathcal{V}_2$, with $\mathcal{V}_1 \cap \mathcal{V}_2 = \emptyset$, such that every edge between \mathcal{V}_1 and \mathcal{V}_2 is negative, and every edge within \mathcal{V}_1 or \mathcal{V}_2 is positive.

Figure 4.3a illustrates a structurally balanced signed graph. An unsigned graph is a special case of structurally balanced graph, in which for instance $\mathcal{V}_1 = \mathcal{V}$ and $\mathcal{V}_2 = \emptyset$.

The opposing signed Laplacian (4.3)² has been used in the literature of signed graphs to characterize the structural balance of a signed graph \mathcal{G} . In compact form, $L = \Delta - A$, where Δ is the (signed) degree matrix (4.5), $\Delta := \text{diag}\{\delta_1, \dots, \delta_n\}$ with $\delta_i = \sum_{j=1}^n |a_{ij}| \geq 0$ for all i . When the signed graph \mathcal{G} is connected, then it does not have isolated vertices, which means that the degree matrix Δ is positive definite, and that its inverse Δ^{-1} is well-defined and positive definite. As for unsigned graphs, the normalized signed Laplacian can be defined as

$$[\mathcal{L}]_{ij} = \begin{cases} 1, & j = i \\ -\frac{a_{ij}}{\sum_{k=1}^n |a_{ik}|}, & j \neq i. \end{cases} \quad (4.7)$$

which in compact form reads $\mathcal{L} = \Delta^{-1}L = I - \Delta^{-1}A$. Equivalent conditions for \mathcal{G} to be structurally balanced can be formulated in terms of \mathcal{L} , see for instance [16].

Theorem 4.1. Let \mathcal{G} be a signed connected graph ($\text{card}(\mathcal{V}) = n$) with normalized signed Laplacian \mathcal{L} . Then \mathcal{G} is structurally balanced if and only if any of the following conditions are satisfied:

- (i) all the cycles in \mathcal{G} are positive, i.e., they all have an even number of edges with negative weights;

²Notice that in this section the superscript “o” will be dropped, and the opposing signed Laplacian L^o will be referred to just as signed Laplacian L .

- (ii) there exists a signature matrix $S = \text{diag}\{s_1, \dots, s_n\}$, with $s_i = \pm 1$ for all i , such that $S\mathcal{L}S$ has all nonpositive off-diagonal entries;
- (iii) the smallest eigenvalue of \mathcal{L} is zero, i.e., $\lambda_1(\mathcal{L}) = 0$.

The first condition of Theorem 4.1 corresponds to the original definition of structurally balanced network introduced by Harary and Cartwright [15].

The second condition, illustrated in Fig. 4.3b and Fig. 4.3c, means that applying special transformations, called *switching equivalence* or *signature similarity transformations* in the signed graph theory literature [14], it is possible to obtain a graph whose edges have all positive weights. In practice, these transformations consist of changing the signs of all the edges adjacent to a set of nodes, see Fig. 4.3c, which can be identified by the signature matrix S and is given either by $\{i \in \mathcal{V} : s_i = +1\}$ or $\{i \in \mathcal{V} : s_i = -1\}$.

Finally, the third condition can be used to clarify one of the properties of the signed (opposing) Laplacian, see the first line of Table 4.1. Using the Geršgorin's Theorem, it is clear that all the eigenvalues of \mathcal{L} have nonnegative real part. Thanks to Theorem 4.1 it is possible to add that, in particular, 0 is an eigenvalue of \mathcal{L} (with eigenvector $S\mathbf{1}$) if and only if the graph is structurally balanced.

When a graph is structurally unbalanced, the question naturally arises as to how far the signed graph is from a structurally balanced state. The standard “measure” of such a distance is the *frustration index*, hereafter simply denoted *frustration*, originally introduced by Harary [17, 60] for unweighted signed graphs under the name *line index of balance*. It indicates the minimum (weighted) sum of the negative edges that have to be removed in order to obtain a graph that is structurally balanced.

Definition 4.3. The frustration $\epsilon(\mathcal{G})$ of a signed graph \mathcal{G} is defined as the minimum (weighted) sum of the positive edges over all signature similarity transformations of \mathcal{L} :

$$\epsilon(\mathcal{G}) = \min_{\substack{S=\text{diag}\{s_1, \dots, s_n\} \\ s_i = \pm 1 \forall i}} \frac{1}{2} \sum_{\substack{i,j=1 \\ j \neq i}}^n [|S\mathcal{L}S|_{ij}] \quad (4.8)$$

Observe that if a signed graph \mathcal{G} is structurally balanced, then $\epsilon(\mathcal{G}) = 0$ (see e.g., Theorem 4.1(ii)).

Another (maybe more intuitive) interpretation of the notion of frustration, which comes from the statistical physics literature³, is that the frustration is the minimum over all possible signature matrices S of an energy functional, defined as

$$e(S) = \frac{1}{2} \sum_{\substack{i,j=1 \\ j \neq i}}^n [|S\mathcal{L}S|_{ij}] \quad (4.9)$$

Using this analogy, the frustration quantifies the amount of “disorder” in a signed network.

³In particular from the Ising spin glass literature, see [18, 61–63], where the frustration is defined as the energy in the ground state of an Ising spin glass.

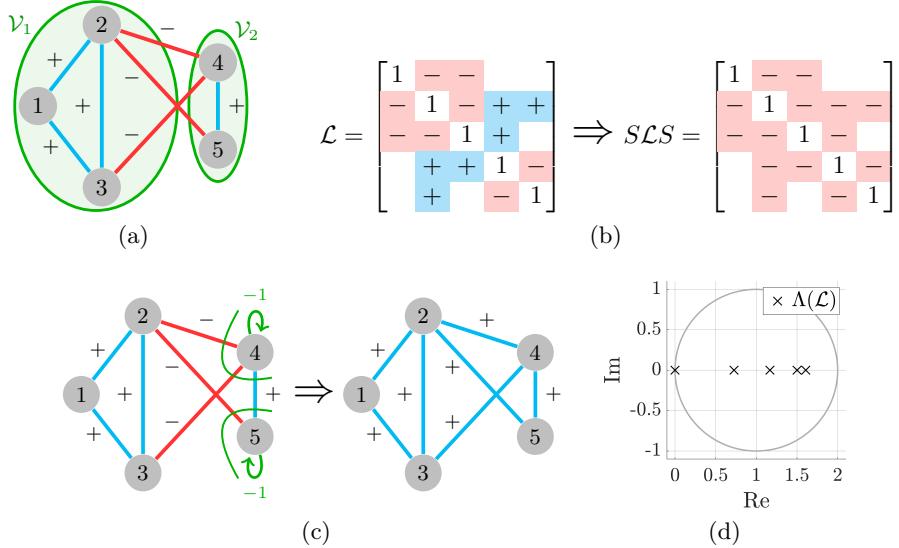


Figure 4.3: Illustration of the notion of structural balance, see Definition 4.2 and Theorem 4.1. (a): Structurally balanced signed graph. (b): Applying the signature matrix $S = \text{diag}\{1, 1, 1, -1, -1\}$ to the signed graph in (a) shows that $S\mathcal{L}S$ is a M-matrix. (c): Equivalently, changing the signs of the edges adjacent to nodes 4 and 5 (with $\{4, 5\} = \{i \in \mathcal{V} : s_i = -1\}$) yields a graph with positive edges. (d) Geršgorin's disks and eigenvalues of \mathcal{L} (computed for the signed graph in (a) with weights $a_{ij} \in \{-1, 0, 1\}$).

In practice, the computation of the frustration (4.8) is an NP-hard problem. Several algorithms are proposed in the literature, see e.g., [18, 21, 64], and in particular the papers in Part II use the algorithm introduced in [21] to compute numerically the frustration of a signed graph.

Another way to measure the level of imbalance of a signed graph \mathcal{G} is the *algebraic conflict*, defined simply as the smallest eigenvalue of the normalized signed Laplacian \mathcal{L} of \mathcal{G} , $\lambda_1(\mathcal{L})$, which is known to be strictly positive for structurally unbalanced signed graphs (see Theorem 4.1(iii)). In addition, $\lambda_1(\mathcal{L})$ is known to approximate well the frustration of a signed network [19, 24, 33], as shown in the following example (inspired by [28, 33]).

Example 4.1

In this example a sequence of signed graphs with increasing frustration is considered and, for each graph of the sequence, the values of frustration and of the smallest eigenvalue of the normalized signed Laplacian are computed and compared. The sequence is constructed as follows. Each graph of the sequence $\mathcal{G} = (\mathcal{V}, \mathcal{E}, A)$ has $\text{card}(\mathcal{V}) = n = 500$ nodes, and the edge weights are drawn from a uniform distribution (where $p = 0.8$ is the edge probability). The signature of \mathcal{G} is dependent on a parameter $\beta \in [0, 1]$: in particular, if $(j, i) \in \mathcal{E}$ then $P[a_{ij} < 0] = \beta$. In practice, β represents the percentage of edges with negative weight in the signed graph \mathcal{G} .

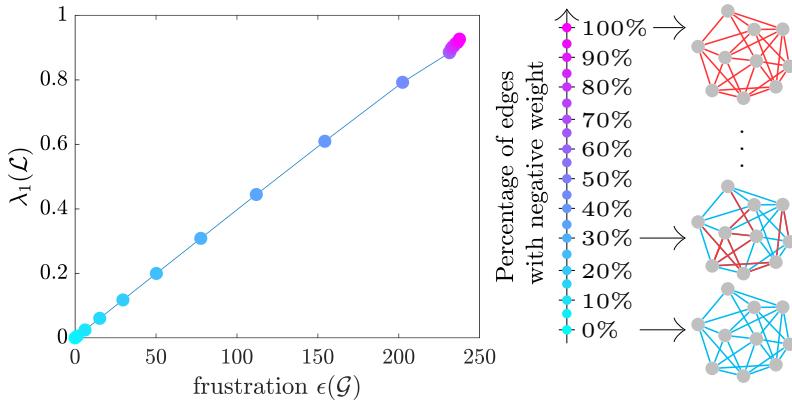


Figure 4.4: Example 4.1. Comparison between smallest eigenvalue of the normalized signed Laplacian $\lambda_1(\mathcal{L})$ (i.e., algebraic conflict in the structural balance literature [24]) and frustration $\epsilon(\mathcal{G})$, for a sequence of signed graphs with increasing frustration.

The sequence is considered for increasing values of β , $\beta \in \{0, 0.05, \dots, 1\}$, and the frustration $\epsilon(\mathcal{G})$ of each signed graph in the sequence, computed numerically using the algorithm proposed in [21], is shown to increase. Figure 4.4 plots the smallest eigenvalue of the normalized signed Laplacian \mathcal{L} vs the frustration, for each signed graph of the sequence.

5

Concluding remarks

The focus of this thesis is on the study of the collective behavior in a community of agents exchanging opinions, and on understanding how it could be influenced by the presence of antagonistic interactions among the agents.

Several classes of models are considered for this scope. Even though different protocols for consensus have been proposed in the literature, there are cases in which these protocols may prove limiting/restrictive: for instance they fail to represent situations in which the agents may want to rule out particular values of consensus, or to impose that the consensus value respects their own “comfort interval”.

For this reason, the first paper of this thesis proposes a new protocol for consensus, denoted interval consensus, which allows the agents to impose constraints on the achievable consensus value. It is a nonlinear interconnected model with a Laplacian-like scheme, where saturations are introduced to describe how the agents transmit their opinions to their neighbors in the network. It is shown that, while during the transient the opinions need not satisfy the constraints imposed by the agents through the saturations, asymptotic convergence to interval consensus is achieved: the opinions converge to a common value belonging to the intersection of the constraints (if nonempty). Formally, existence and asymptotic stability of the equilibrium point of consensus is proven, in case of strongly connected networks. Existence of (at least) an equilibrium point (which, in general, is not a consensus value) is also shown when the intersection is empty, a case which is perhaps less relevant in practical applications.

The second class of models considered in this thesis work, and introduced in paper B, presents a similar design, in that it is characterized by a Laplacian-like structure at the origin, and saturated nonlinearities of sigmoidal type. Differently from the first class of models, the amplitude of the interaction part of the dynamics is modulated by a scalar positive parameter, which plays the role of bifurcation parameter in the analysis. As the parameter is varied, the system exhibits a pitchfork

bifurcation behavior: for small values of the parameter the origin is the unique equilibrium point of the system and it is globally asymptotically stable. When the parameter increases and crosses a first threshold value, the system undergoes a pitchfork bifurcation and two new alternative (one positive, one negative) equilibria appear, which are locally asymptotically stable, and, in the particular case of identical nonlinearities, are consensus values. The uniqueness of these three equilibria (the origin, and the two alternative consensus equilibria) is proven for all values of the bifurcation parameter less than a second threshold value, after which the system undergoes a second pitchfork bifurcation and new mixed-sign equilibria appear. The equilibrium points that the system admits for high values of the bifurcation parameter may be stable or unstable (as shown by means of numerical simulations), and, while an estimate of their total number is not provided, their norm is shown to be upper bounded by the norm of the (stable) consensus equilibria. Finally, it is shown in paper B that, while the first threshold value for the bifurcation parameter is fixed/constant, the second threshold value depends on the algebraic connectivity of the network, meaning that the “robustness” of the consensus problem (in the sense of size of the interval of values of the bifurcation parameter for which the system admits only three equilibria) is influenced by the network’s topology.

The behavior of the system can be interpreted in terms of collective decision-making, where the bifurcation parameter represents the “social effort” or “strength of commitment” among the agents. In order to escape from a deadlock situation (where the origin is the only equilibrium point, i.e., no decision is taken) and achieve a decision (i.e., convergence to one of the two alternative nonzero equilibrium points is achieved), the agents need to have the “right” amount of commitment; while this decision represents typically an agreement, disagreement situations are always possible if the agents become “overcommitted” (i.e., the second bifurcation threshold is crossed).

The nonlinear models considered in papers A and B are cooperative, meaning that a decision (in general, of consensus) is reached through collaboration among the agents. This assumption is however restrictive in many real-world scenarios, where rivalry among the agents is inevitable. For this reason, in paper C the nonlinear model of paper B is extended to the case of antagonistic networks, by representing the competitive interactions among the agents through a signed network. When the network is structurally balanced, the system is monotone and its behavior is identical to the model of paper B, modulo a change of orthant: instead of consensus, the agents’ opinions reach bipartite consensus (i.e., they reach consensus in absolute value). The qualitative behavior of the system does not change also when the network is structurally unbalanced, however the value of the bifurcation parameter for which the first pitchfork bifurcation is crossed is shown to increase with the frustration of the signed network, which represents the amount of social tension present in the community. The interpretation is that the higher is the frustration, the higher is the social commitment required from the agents in order to reach a decision.

This idea is then used in paper D, whose aim is to investigate the complexity of government formation processes in countries with a parliamentary system,

through an approach based on decision dynamics over signed networks. In particular, by mapping the (post-election) parliament into a signed network, and treating the government formation process as a collective decision-making system over a network (with parliamentary members as agents, and vote of confidence given to the post-election new government as decision), the hypothesis, whose intuition comes from the theoretical analysis of paper C, is that the frustration of the parliamentary network should correlate with the duration of the government negotiation talks: a long period of negotiations is to be expected when the social tension in the parliament is high. This can indeed be seen in the real electoral data in 29 European countries analyzed in paper D. Moreover, it is shown that the concept of frustration of the parliamentary networks can be used also to predict the composition of the successful post-election cabinet coalition.

The work of papers C and D (focused on social/parliamentary networks), where structural balance theory proves significant in order to investigate and interpret the collective behavior of multiagent systems, is one example of the relevance of the study of signed graphs. The focus of paper E is on the structure of signed digraphs, and in particular of a signed version of Laplacian, denoted “repelling” signed Laplacian, and its pseudoinverse, used for instance in the context of electric networks. Passing from unsigned to signed graphs, important properties of this Laplacian matrix, such as diagonal dominance and marginal stability, are in general no longer guaranteed. Moreover, even in the unsigned case, the set of Laplacian matrices is not closed under pseudoinversion. In paper E, the set of signed Laplacian matrices that are weight balanced and eventually exponentially positive (i.e., matrices whose exponential satisfies the strong Perron-Frobenius property) is shown to be closed not only w.r.t. marginal stability but also to matrix pseudoinversion.

Directions of future research

This section discusses possible directions in which the results presented in this thesis work could be extended.

Regarding the nonlinear (possibly, antagonistic) model over networks presented in papers B to D, it is important to observe that the pitchfork bifurcation behavior the multiagent system exhibits is due to symmetry (in the dynamics and in the graph), and an interesting development would be to understand the effect of perturbations or symmetry-breaking phenomena. One direction would be to remove the assumption that the (signed) network representing the community of agents is undirected. In the cooperative or, in general, monotone case (corresponding to a nonnegative or structurally balanced graph) the intuition is that the general behavior of the system would not change: for instance, it would still not admit stable limit cycles. However, this would be no longer true when monotonicity does not hold (i.e., when the signed graph is not structurally balanced): in this case, a Hopf bifurcation might occur and periodic solutions arise. Perhaps more interesting would be to study the effect of adding a constant input to the dynamics, representing for instance external factors, or preference of the agents for a certain alternative. This has been investigated in [4] for the cooperative case, where the

authors use unfolding theory of a pitchfork bifurcation (see e.g. [48]) to discuss how the preference of the agents for one of the two possible alternatives (represented by two consensus equilibria) affects the choice they make as a group.

Concerning instead the results presented in paper E, while from one side efforts should be spent in trying to characterize better the structure of the set of eventually exponentially positive and weight balanced Laplacian matrices, on the other side it would be interesting to consider the possible applications of the Laplacian pseudoinverse to signed networks: for instance, in the context of electrical networks, by exploring the connection between the signed Laplacian pseudoinverse and the concept of effective resistance, which has been only briefly mentioned in paper E.

The study of theory of eventually (exponentially) positive matrices presented in paper E (and [57]) could also be beneficial in the context of nonlinear dynamics over signed networks of paper C. Indeed, the way paper C extends the nonlinear co-operative model presented in paper B to the antagonistic case is to assume that its linearization at the origin corresponds to the version of signed Laplacian denoted “opposing” signed Laplacian. It would be interesting instead to consider the version denoted “repelling” signed Laplacian, and use the work of paper E (and [57]) to investigate the behavior of the corresponding nonlinear model and characterize its equilibria (e.g., if they represent agreement or disagreement); additionally, the results could be compared with the ones obtained for linear systems in [65].

Bibliography

- [1] N. E. Leonard, “Multi-agent system dynamics: Bifurcation and behavior of animal groups,” *Annual Reviews in Control*, vol. 38, no. 2, pp. 171–183, 2014.
- [2] I. D. Couzin, C. C. Ioannou, G. Demirel, T. Gross, C. J. Torney, A. Hartnett, L. Conradt, S. A. Levin, and N. E. Leonard, “Uninformed individuals promote democratic consensus in animal groups,” *Science*, vol. 334, no. 6062, pp. 1578–1580, 2011.
- [3] D. T. Swain, I. D. Couzin, and N. Ehrich Leonard, “Real-Time Feedback-Controlled Robotic Fish for Behavioral Experiments With Fish Schools,” *Proceedings of the IEEE*, vol. 100, no. 1, pp. 150–163, jan 2012.
- [4] R. Gray, A. Franci, V. Srivastava, and N. E. Leonard, “Multiagent Decision-Making Dynamics Inspired by Honeybees,” *IEEE Transactions on Control of Network Systems*, vol. 5, no. 2, pp. 793–806, jun 2018.
- [5] J. A. Fax and R. M. Murray, “Information Flow and Cooperative Control of Vehicle Formations,” *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1465–1476, sep 2004.
- [6] W. Ren and R. W. Beard, *Distributed Consensus in Multi-vehicle Cooperative Control*, ser. Communications and Control Engineering. London: Springer London, 2008.
- [7] W. Ren and Y. Cao, *Distributed Coordination of Multi-agent Networks*, ser. Communications and Control Engineering. London: Springer London, 2011.
- [8] R. Hegselmann and U. Krause, “Opinion Dynamics and Bounded Confidence,” *Journal of Artificial Societies and Social Simulation*, vol. 5, no. 3, p. 2, 2002.
- [9] P. Jia, A. MirTabatabaei, N. E. Friedkin, and F. Bullo, “Opinion Dynamics and the Evolution of Social Power in Influence Networks,” *SIAM Review*, vol. 57, no. 3, pp. 367–397, 2015.
- [10] D. Easley and J. Kleinberg, *Networks, Crowds, and Markets: Reasoning about a Highly Connected World*. Cambridge: Cambridge University Press, 2010.

- [11] A. V. Proskurnikov and R. Tempo, "A tutorial on modeling and analysis of dynamic social networks. Part I," *Annual Reviews in Control*, vol. 43, pp. 65–79, 2017.
- [12] A. V. Proskurnikov and R. Tempo, "A tutorial on modeling and analysis of dynamic social networks. Part II," *Annual Reviews in Control*, vol. 45, pp. 166–190, jan 2018.
- [13] W. Ren, R. W. Beard, and E. M. Atkins, "A survey of consensus problems in multi-agent coordination," in *2005 American Control Conference (ACC)*, vol. 3. IEEE, 2005, pp. 1859–1864.
- [14] T. Zaslavsky, "Signed graphs," *Discrete Applied Mathematics*, vol. 4, no. 1, pp. 47–74, 1982.
- [15] D. Cartwright and F. Harary, "Structural balance: a generalization of Heider's theory." *Psychological Review*, vol. 63, no. 5, pp. 277–293, 1956.
- [16] C. Altafini, "Consensus Problems on Networks With Antagonistic Interactions," *IEEE Transactions on Automatic Control*, vol. 58, no. 4, pp. 935–946, apr 2013.
- [17] F. Harary, "On the measurement of structural balance," *Behavioral Science*, vol. 4, no. 4, pp. 316–323, jan 1959.
- [18] G. Facchetti, G. Iacono, and C. Altafini, "Computing global structural balance in large-scale signed social networks," *Proceedings of the National Academy of Sciences*, vol. 108, no. 52, pp. 20 953–20 958, 2011.
- [19] S. Aref and M. C. Wilson, "Measuring partial balance in signed networks," *Journal of Complex Networks*, vol. 6, no. 4, pp. 566–595, aug 2018.
- [20] M. W. Hirsch, S. Smale, and R. L. Devaney, *Differential Equations, Dynamical Systems, and an Introduction to Chaos*, 2nd ed., A. Press, Ed. Elsevier, 2004.
- [21] G. Iacono, F. Ramezani, N. Soranzo, and C. Altafini, "Determining the distance to monotonicity of a biological network: a graph-theoretical approach," *IET Systems Biology*, vol. 4, no. 3, pp. 223–235, may 2010.
- [22] E. D. Sontag, "Monotone and near-monotone biochemical networks," *Systems and Synthetic Biology*, vol. 1, no. 2, pp. 59–87, 2007.
- [23] J. Kunegis, S. Schmidt, A. Lommatzsch, J. Lerner, E. W. De Luca, and S. Albayrak, "Spectral Analysis of Signed Graphs for Clustering, Prediction and Visualization," in *2010 SIAM International Conference on Data Mining*. Philadelphia, PA: Society for Industrial and Applied Mathematics, apr 2010, pp. 559–570.
- [24] J. Kunegis, "Applications of Structural Balance in Signed Social Networks," *arXiv:1402.6865v1*, pp. 1–37, 2014.

- [25] A. Fontan and C. Altafini, “A signed network perspective on the government formation process in parliamentary democracies,” *Scientific Reports*, vol. 11, no. 5134, dec 2021.
- [26] A. Fontan, G. Shi, X. Hu, and C. Altafini, “Interval Consensus for Multiagent Networks,” *IEEE Transactions on Automatic Control*, vol. 65, no. 5, pp. 1855–1869, may 2020.
- [27] A. Fontan and C. Altafini, “Multiequilibria Analysis for a Class of Collective Decision-Making Networked Systems,” *IEEE Transactions on Control of Network Systems*, vol. 5, no. 4, pp. 1931–1940, dec 2018.
- [28] A. Fontan and C. Altafini, “The role of frustration in collective decision-making dynamical processes on multiagent signed networks,” *arXiv:2105.11396*, pp. 1–18, may 2021.
- [29] A. Fontan and C. Altafini, “On the properties of Laplacian pseudoinverses,” in *60th IEEE Conference on Decision and Control*. Austin, TX, USA: IEEE, 2021.
- [30] A. Fontan, G. Shi, X. Hu, and C. Altafini, “Interval consensus: a novel class of constrained consensus problems for multiagent networks,” in *56th IEEE Conference on Decision and Control*. Melbourne, Australia: IEEE, dec 2017, pp. 4155–4160.
- [31] A. Fontan and C. Altafini, “Investigating mixed-sign equilibria for nonlinear collective decision-making systems,” in *56th IEEE Conference on Decision and Control*. Melbourne, Australia: IEEE, dec 2017, pp. 781–786.
- [32] A. Fontan and C. Altafini, “Modeling wireless power transfer in a network of smart devices: A compartmental system approach,” in *2018 17th European Control Conference (ECC)*. Limassol, Cyprus: European Control Association (EUCA), 2018, pp. 1468–1473.
- [33] A. Fontan and C. Altafini, “Achieving a decision in antagonistic multi agent networks: frustration determines commitment strength,” in *57th IEEE Conference on Decision and Control*. Miami Beach, FL, USA: IEEE, dec 2018, pp. 109–114.
- [34] R. A. Horn and C. R. Johnson, *Matrix analysis*, 2nd ed. Cambridge University Press, 2013.
- [35] E. Kaszkurewicz and A. Bhaya, *Matrix Diagonal Stability in Systems and Computation*. Boston: Birkhäuser, 2000.
- [36] S. Camiz and S. Stefani, *Matrices and Graphs*. World Scientific, 1996.
- [37] C. Altafini, “Stability analysis of diagonally equipotent matrices,” *Automatica*, vol. 49, no. 9, pp. 2780–2785, sep 2013.

- [38] D. Noutsos, “On Perron–Frobenius property of matrices having some negative entries,” *Linear Algebra and its Applications*, vol. 412, no. 2–3, pp. 132–153, jan 2006.
- [39] A. Berman and R. J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*, ser. Classic in applied mathematics. Society for Industrial and Applied Mathematics, 1994.
- [40] H. K. Khalil, *Nonlinear Systems: Third Edition*. Pearson, 2002.
- [41] M. W. Hirsch, “Systems of Differential Equations that are Competitive or Cooperative II: Convergence Almost Everywhere,” *SIAM Journal on Mathematical Analysis*, vol. 16, no. 3, pp. 423–439, may 1985.
- [42] H. L. Smith, “Systems of Ordinary Differential Equations Which Generate an Order Preserving Flow. A Survey of Results,” *SIAM Review*, vol. 30, no. 1, pp. 87–113, mar 1988.
- [43] M. W. Hirsch and H. L. Smith, “Monotone Dynamical Systems,” in *Handbook of Differential Equations: Ordinary Differential Equations*. North-Holland, 2006, vol. 2, ch. 4, pp. 239–357.
- [44] N. Rouche, P. Habets, and M. Laloy, *Stability Theory by Liapunov’s Direct Method*. Springer New York, 1977.
- [45] L. Perko, *Differential Equations and Dynamical Systems*, ser. Texts in Applied Mathematics. New York, NY: Springer New York, 2001, vol. 7.
- [46] Y. A. Kuznetsov, *Elements of Applied Bifurcation Theory, Second Edition*. Springer, 1998.
- [47] F. Blanchini and S. Miani, *Set-Theoretic Methods in Control*, ser. Systems & Control: Foundations & Applications. Boston, MA: Birkhäuser Boston, 2008.
- [48] M. Golubitsky and D. G. Schaeffer, *Singularities and Groups in Bifurcation Theory. Volume I*, ser. Applied Mathematical Sciences. Springer New York, 1985, vol. 51.
- [49] M. Golubitsky and I. Stewart, *The Symmetry Perspective*, ser. Progress in Mathematics (Volume 200), H. Bass, J. Oesterle, and A. Weinstein, Eds. Basel: Birkhäuser Basel, 2002.
- [50] C. Godsil and G. Royle, *Algebraic Graph Theory*, ser. Graduate Texts in Mathematics. New York, NY: Springer New York, 2001, vol. 207.
- [51] F. R. K. Chung, *Spectral Graph Theory*, ser. CBMS Number 92. American Mathematical Society, 1997.
- [52] F. Bullo, *Lectures on Nonlinear Network Systems* (ed. 1.4). Kindle Direct Publishing, 2020. <http://motion.me.ucsb.edu/book-lns>

- [53] T. Zaslavsky, “Matrices in the Theory of Signed Simple Graphs,” *arXiv:1303.3083v1*, pp. 1–20, mar 2013.
- [54] G. Shi, C. Altafini, and J. S. Baras, “Dynamics over Signed Networks,” *SIAM Review*, vol. 61, no. 2, pp. 229–257, jan 2019.
- [55] R. Olfati-Saber and R. Murray, “Consensus Problems in Networks of Agents With Switching Topology and Time-Delays,” *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1520–1533, sep 2004.
- [56] R. Agaev and P. Chebotarev, “On the spectra of nonsymmetric Laplacian matrices,” *Linear Algebra and Its Applications*, vol. 399, no. 1-3, pp. 157–168, 2005.
- [57] C. Altafini, “Investigating stability of Laplacians on signed digraphs via eventual positivity,” in *58th IEEE Conference on Decision and Control*. Nice, France: IEEE, dec 2019, pp. 5044–5049.
- [58] F. Heider, “Attitudes and Cognitive Organization,” *The Journal of Psychology*, vol. 21, no. 1, pp. 107–112, jan 1946.
- [59] F. Harary, “On the notion of balance of a signed graph.” *The Michigan Mathematical Journal*, vol. 2, no. 2, pp. 143–146, 1953.
- [60] F. Harary and J. A. Kabell, “A simple algorithm to detect balance in signed graphs,” *Mathematical Social Sciences*, vol. 1, no. 1, pp. 131–136, sep 1980.
- [61] F. Barahona, “On the computational complexity of ising spin glass models,” *Journal of Physics A: Mathematical and General*, vol. 15, no. 10, pp. 3241–3253, 1982.
- [62] K. Binder and A. P. Young, “Spin glasses: Experimental facts, theoretical concepts, and open questions,” *Reviews of Modern Physics*, vol. 58, no. 4, pp. 801–976, oct 1986.
- [63] M. Mézard and G. Parisi, “The Bethe lattice spin glass revisited,” *The European Physical Journal B*, vol. 20, no. 2, pp. 217–233, mar 2001.
- [64] S. Aref, A. J. Mason, and M. C. Wilson, “An exact method for computing the frustration index in signed networks using binary programming,” *arXiv:1611.09030v2*, nov 2016.
- [65] C. Altafini and G. Lini, “Predictable dynamics of opinion forming for networks with antagonistic interactions,” *IEEE Transactions on Automatic Control*, vol. 60, no. 2, pp. 342–357, feb 2015.

Part II

Publications

Paper A

Interval Consensus for Multiagent Networks

Authors: Angela Fontan, Guodong Shi, Xiaoming Hu and Claudio Altafini

© [2020] IEEE. *Edited version of the paper:*

A. Fontan, G. Shi, X. Hu, and C. Altafini, “Interval Consensus for Multiagent Networks,” *IEEE Transactions on Automatic Control*, vol. 65, no. 5, pp. 1855–1869, may 2020.

Early versions of this work were submitted in:

A. Fontan, G. Shi, X. Hu, and C. Altafini, “Interval consensus: a novel class of constrained consensus problems for multiagent networks,” in *56th IEEE Conference on Decision and Control*. Melbourne, Australia: IEEE, dec 2017, pp. 4155–4160.

Interval Consensus for Multiagent Networks

Angela Fontan^{*}, Guodong Shi[†], Xiaoming Hu[‡] and Claudio Altafini^{*}

^{*}Dept. of Electrical Engineering
Linköping University
SE-581 83 Linköping, Sweden
angela.fontan@liu.se,
claudio.altafini@liu.se

[†]Australian Center for Field Robotics,
School of Aerospace, Mechanical
and Mechatronic Engineering
The University of Sydney
NSW 2008 Sydney, Australia
guodong.shi@sydney.edu.au

[‡]Dept. of Optimization and Systems Theory
Royal Institute of Technology
SE-100 44 Stockholm, Sweden
hu@kth.se

Abstract

The constrained consensus problem considered in this paper, denoted interval consensus, is characterized by the fact that each agent can impose a lower and upper bound on the achievable consensus value. Such constraints can be encoded in the consensus dynamics by saturating the values that an agent transmits to its neighboring nodes. We show in the paper that when the intersection of the intervals imposed by the agents is nonempty, the resulting constrained consensus problem must converge to a common value inside that intersection. In our algorithm, convergence happens in a fully distributed manner, and without need of sharing any information on the individual constraining intervals. When the intersection of the intervals is an empty set, the intrinsic nonlinearity of the network dynamics raises new challenges in understanding the node state evolution. Using Brouwer fixed-point theorem we prove that in that case there exists at least one equilibrium, and in fact the possible equilibria are locally stable if the constraints are satisfied or dissatisfied at the same time among all nodes. For graphs with sufficient sparsity it is further proven that there is a unique equilibrium that is globally attractive if the constraint intervals are pairwise disjoint.

1 Introduction

The basic idea of a consensus problem is to achieve an agreement among a group of agents through a distributed dynamical system, encoding the values that the agents want to contribute as initial conditions of a Laplacian-like system which

represents the exchanges of information among the first neighbors of a communication graph. Owing to the Laplacian structure of the dynamics, each agent is driven only by relative states, i.e., differences between its own state and that of its neighbors. Various algorithms have been developed using this scheme. For instance, the average consensus problem consists of computing the average of such initial conditions, see [3]. In a leader-follower scenario, instead, only the initial conditions of the leaders matter, and provide the values to which the followers converge, see [4]. In a max consensus problem, the agents determine the max of their initial conditions, and all settle to that value, see [5]. When cooperation and competition among the agents coexist, a bipartite consensus can be achieved, provided that the graph is structurally balanced, see [6].

In all these protocols, an agent has no authority to veto certain values of consensus, or to impose that the consensus is restricted within an admissible region. This is a drastic limitation in certain contexts. For instance, in a network of processors trying to agree on sharing a computational load, each processor might have constraints on the computational resources allocable to the shared task, and accept only consensus values which are within that range. In an opinion dynamics context, an agent might agree on a common opinion only if this is not too extreme. In a formation docking problem, a robot might be able to achieve alignment with the rest of the formation only if the consensus position is within a certain region. In all these cases, what one would like to add is a state constraint to the consensus problem.

Consensus problems with constraints have been studied from different perspectives in the literature. A significant group of papers deals with the use of state projections on convex sets, mostly in discrete-time consensus problems and motivated by optimization algorithms [7]. Projection-based methods for state constraint satisfaction have been introduced also for continuous-time consensus problems, using projection operators inspired by the adaptive control literature [8], or logarithmic barrier functions [9]. Continuous flows can be used to solve convex intersection computation problems when the states of the nodes are not necessarily satisfying the constraints for all time [10]. In [11] a discontinuous vector field is used to describe the state saturation. Alternative approaches for imposing state constraints on consensus problems are proposed in e.g., [12, 13]. A different situation of consensus with state constraints is the positive consensus problem studied in [14]. In this case, the aim is to achieve consensus while respecting the positivity of the state variables, representing e.g., quantities that are intrinsically nonnegative (masses, concentrations, etc.).

Other types of constrained consensus problems that have been considered in the literature include for instance the discarded consensus algorithm of [15], that discards the state of neighbors if they are outside of certain bounds, or the distributed averaging with flow constraints considered in [16]. Sometimes instead of state constraints one is interested in models with inputs constraints, representing e.g., actuator saturations, see e.g., [17, 18]. The opinion dynamics literature offers several other contexts in which models are endowed with state constraints in order to better represent a phenomenon. In [19] for instance, interactions are unilateral, i.e., are considered only if the state of the neighboring nodes is higher than the

agent's state for optimistic models, or lower for pessimistic models. A different approach, used in opinion dynamics, is proposed in the so-called bounded confidence models [20, 21], in which states that are more distant than a certain threshold ignore each other. The result is that these models produce clusters of opinions, and a local consensus value within each cluster. Various variants of this opinion dynamics problem have been proposed, to accommodate other constraints in addition to bounded confidence. For instance in [22] the sign of the initial conditions is maintained throughout the opinion clustering process.

The problem we intend to study in this paper is different from all the aforementioned state-constrained consensus problems. The main idea we want to introduce in a consensus problem is that we want to give to each agent the possibility of limiting the interval of values in which a consensus value can be accepted, and therefore force the agreed consensus value to belong to the intersection of all such intervals, if such intersection is nonempty. The constraints we want to impose are however not classical hard constraints on the state variables. Rather, they should only condition the range in which the steady state consensus value belongs to, but should be trespassable during the transient evolution. To distinguish our problem from these other forms of consensus with hard-wired constraints, we call it *interval consensus*.

It is worth observing that our interval consensus problem is not related to the notion of “bipartite interval consensus” introduced in [23]. In that paper, in fact, lack of strong connectivity of the graph is used to achieve some form of containment control (or leader-follower scheme [4]), but no common value (monopartite or bipartite) is achieved. In our problem, instead, the objective of the agents is to achieve a common consensus value, in spite of the interval constraints imposed by each of them.

Technically an agent implements an interval consensus by transmitting a value of its state which is saturated between an upper and a lower bound. By limiting the transmitted state we can skip the projection step, and obtain the same result of imposing constraints on the consensus value although only asymptotically. Practically it means that the agents keep seeking a compromise value fitting all constraints, and it is only through “stubbornly” transmitting a saturated value to its neighbors that an agent manages to carry the common consensus value within the interval imposed by its constraints. Clearly, properties like the presence of a conserved quantity in average consensus, or the “diffusion-like” structure of any linear consensus algorithm are lost when the constraints become active. In particular, when this happens the terms in the vector field that drive the consensus may no longer represent relative distances between agents states, meaning that the overall dynamical system behaves like a (Lipschitz continuous) switching system. Nevertheless, in the paper we show that when the intersection of the intervals admissible by the agents is nonempty, a consensus is always achieved, and convergence must necessarily be to a value in the intersection.

In our model each agent decides independently what saturation values to choose for its interval. Consequently, there is no guarantee that the intervals have a nonzero intersection. When the intersection is nonempty (the most interesting case from an application perspective) our results provide a complete and global

description of the behavior of the system. The system is marginally stable inside the intersection of the allowed consensus intervals, but is asymptotically stable outside it, because of the saturations. When the intersection of the admissible intervals is instead empty, the analysis of the model turns out to be much more challenging, and we could obtain only partial results on the uniqueness and stability character of the equilibrium points.

It is worth mentioning that the case of nonempty intersection of the admissible intervals is the only one treated in the papers dealing with state constraints. Furthermore, in this literature invariance of the dynamics to the interval intersection is typically imposed. In continuous time, saturation of the dynamics by itself is not enough to guarantee forward invariance of the interval intersection. In fact, in order to avoid excursions of the dynamics outside the intersection, one needs to resort to vector fields with special structure, like projection operators or discontinuities, which however render the dynamics significantly more complicated and add additional burden to the problem. In [9] for instance the logarithmic barrier approach requires the agent to make use of auxiliary variables that must be transmitted alongside the state variables. The model [11], which imposes forward invariance of the interval intersection by means of a discontinuous vector field, is problematic to deal with because uniqueness of the solutions might be lost at the saturation points. Also the projection-based approach of [8] relies intrinsically on rendering the interval intersection invariant, and can only be applied under that assumption. In general, to the best of our knowledge, none of the available methods deals with the case of empty interval intersection and even for the case of nonempty intersection the analysis is restricted to initial conditions already inside the interval intersection (i.e., global attractivity of the intersection is never shown). For this last case our analysis is instead global.

In the paper we treat both the continuous-time and discrete-time interval consensus problems. In both cases we normally assume that the graph of interactions is directed and strongly connected. Needless to say, our interval consensus protocol respects the fully distributed nature of the problem, including for what concerns the individual upper and lower bounds, which are unknown to the other agents.

2 Problem Definition

2.1 The Model

We consider a network with n nodes indexed in the set $\mathcal{V} = \{1, \dots, n\}$. The structure of node interconnections is described by a simple directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where each element in \mathcal{E} is an ordered pair of two distinct nodes in the set \mathcal{V} . The neighbor set of node i in the graph \mathcal{G} is denoted $\mathcal{N}_i := \{j : (j, i) \in \mathcal{E}\}$. Each edge $(j, i) \in \mathcal{E}$ is associated with a weight $a_{ij} > 0$. Each node m holds a state $x_m(t) \in \mathbb{R}$ at time $t \geq 0$. Instead of $x_m(t)$, the node transmits to its neighbors in \mathcal{V} a value

$\psi_m(x_m(t))$ lying within an interval $\mathcal{I}_m := [p_m, q_m]$, where

$$\psi_m(z) = \begin{cases} p_m, & \text{if } z < p_m; \\ z, & \text{if } p_m \leq z \leq q_m; \\ q_m, & \text{if } z > q_m. \end{cases} \quad (1)$$

The evolution of $x_i(t) \in \mathbb{R}$ is therefore described by

$$\frac{d}{dt}x_i(t) = \sum_{j \in \mathcal{N}_i} a_{ij} (\psi_j(x_j(t)) - x_i(t)), \quad i \in \mathcal{V}. \quad (2)$$

The nonlinear consensus system (2) will be studied in this paper.

2.2 Examples

A few more specific examples in which our notion of interval consensus is of interest are the following.

- *Achieving a price agreement among shareholders.* Assume the board members of a company are negotiating a buy or sell order, and have to find an agreement among themselves on a price, price for which each of them is imposing boundaries. If unanimity of the board is required, then the request of a consensus value that respects everybody's constraints has priority over for instance a consensus value which preserves the average of the initial bids.
- *Load sharing under load assignment constraints.* A network of computational units must share in equal parts a certain workload, under the constraint that each unit can allocate to the workload only a certain amount of resources, not known a priori to the other units. When is it possible for the units to agree on an equal load sharing policy and how?
- *Social interactions under observer effect.* The observer effect is a generalization of the DeGroot type social interaction rule [24], accounting for the fact that in face-to-face interactions opinions exchanged tend to be more "moderate" than they are in reality [25, 26]. In particular, an agent tends to avoid assuming extremist opinions in a debate, but instead let them fall in a "comfort interval" shared with the other agents. Seeking a consensus under such observer effect can be modeled as a saturation in the values of the transmitted opinions, as we do here.

In each of these cases, constraints are part of the problem, and if a consensus solution exists, then it has to respect them. There is however no need to impose that the transient dynamics respects them, i.e., the constraints are soft, not hard, as captured by the model (2).

2.3 Paper Outline

The behavior of (2) depends crucially on the intersection of intervals $\bigcap_{m=1}^n \mathcal{I}_m$:

- (i) When the intersection is nonempty, $\bigcap_{m=1}^n \mathcal{I}_m \neq \emptyset$, then the system (2) always achieves a consensus value belonging to that intersection. This case is the most interesting from an application point of view. A complete analysis of its behavior is provided in both continuous-time (Section 4) and discrete-time (Section 6).
- (ii) When instead the intersection is empty, $\bigcap_{m=1}^n \mathcal{I}_m = \emptyset$, then (at least) an equilibrium is always present, but it is typically not a consensus value. As shown in Section 5, only in some special cases uniqueness and asymptotic stability can be proven explicitly, although numerical simulations (Section 7) suggest that a unique global attractor should be present in all cases.

3 Background Material

Due to the nonlinearity in the network dynamics (2), our work relies heavily on tools from nonlinear systems, non-smooth analysis, and robust consensus which are now briefly reviewed.

3.1 Cooperativity

Let $y = [y_1 \dots y_n]^T, z = [z_1 \dots z_n]^T \in \mathbb{R}^n$. We say $y \leq z$ if $y_i \leq z_i$ for all i . We next consider an autonomous dynamical system described by

$$\frac{d}{dt}x(t) = f(x(t)) = [f_1(x(t)) \dots f_n(x(t))]^T, \quad (3)$$

where $f : \mathbb{R}^n \mapsto \mathbb{R}^n$ is Lipschitz continuous everywhere. Let $x(t, y)$ be the solution of the system (3) with $x(0) = y$. We recall the following definition.

Definition 1. The system (3) is cooperative if $y \leq z$ implies $x(t, y) \leq x(t, z)$ for all $y, z \in \mathbb{R}^n$.

Cooperativity is a special case of monotonicity [27], in correspondence of a Jacobian matrix which is Metzler. An effective test for cooperativity of the dynamical systems from properties of the vector field relies on the so-called Kamke condition [28, Theorem 12.11, p. 581]. The system (3) is cooperative if and only if

$$y \leq z \text{ and } y_i = z_i \implies f_i(y) \leq f_i(z)$$

holds for any $i = 1, \dots, n$. It is easy to verify this condition for the network dynamics (2). In fact, let

$$f_i(x) = \sum_{j \in \mathcal{N}_i} a_{ij} (\psi_j(x_j) - x_i), \quad i \in \mathcal{V}.$$

Then $y \leq z$ and $y_i = z_i$ implies $\psi_j(y_j) \leq \psi_j(z_j)$ for all $j \in \mathcal{N}_i$ as it is straightforward to show. Hence, since $a_{ij} > 0$, $\psi_j(y_j) \leq \psi_j(z_j)$ for all $j \in \mathcal{N}_i$ and $y_i = z_i$ imply

$$f_i(y) = \sum_{j \in \mathcal{N}_i} a_{ij} (\psi_j(y_j) - z_i) \leq \sum_{j \in \mathcal{N}_i} a_{ij} (\psi_j(z_j) - z_i) = f_i(z).$$

Therefore, (2) is a cooperative dynamical system.

3.2 Limit Set, Dini Derivatives, and Invariance Principle

Consider the autonomous system (3), where $f : \mathbb{R}^d \rightarrow \mathbb{R}^d$ is a continuous function. Then $\Omega_0 \subset \mathbb{R}^d$ is called a *positively invariant set* of (3) if, for any $t_0 \in \mathbb{R}$ and any $x(t_0) \in \Omega_0$, we have $x(t) \in \Omega_0$, $t \geq t_0$, along every solution $x(t)$ of (3).

Let $x : (a, b) \rightarrow \mathbb{R}^d$ be a non-continuable solution of (3) with initial condition $x(0) = x_0$, where the time interval is such that $-\infty \leq a < b \leq \infty$. We call y an ω -limit point of $x(t)$ if there exists a sequence $\{t_k\}$ with $\lim_{k \rightarrow \infty} t_k = \omega$ such that $\lim_{k \rightarrow \infty} x(t_k) = y$. The set of all ω -limit points of $x(t)$ is called the ω -limit set of $x(t)$, and is denoted $\Lambda^+(x_0)$. The following lemma is given in [29, Appendix III (Theorems at pp. 364-365)].

Lemma 1. *Let $x(t)$ be a solution of (3). If $x(t)$ is bounded, then $\Lambda^+(x_0)$ is nonempty, compact, connected, and positively invariant. Moreover, there holds $x(t) \rightarrow \Lambda^+(x_0)$ as $t \rightarrow \omega$ with $\omega = \infty$.*

The upper *Dini derivative* of a continuous function $r : (a, b) \rightarrow \mathbb{R}$ at t is defined as

$$d^+r(t) = \limsup_{s \rightarrow 0^+} \frac{r(t+s) - r(t)}{s}.$$

When r is continuous on (a, b) , r is nonincreasing on (a, b) if and only if $d^+r(t) \leq 0$ for any $t \in (a, b)$.

Now let $x(t)$ be a solution of (3) and let $V : \mathbb{R}^d \rightarrow \mathbb{R}$ be a continuous, locally Lipschitz function. The Dini derivative of $V(x(t))$, $d^+V(x(t))$, thereby follows the above definition. On the other hand, one can also define

$$d_f^+V(x) = \limsup_{s \rightarrow 0^+} \frac{V(x + sf(x)) - V(x)}{s}, \quad (4)$$

namely the upper Dini derivative of V along the vector field (3). There holds that

$$d_f^+V(x)|_{x^*} = d^+V(x(t))|_{t_*} \quad (5)$$

when putting $x(t_*) = x^*$ [29]. The next result is convenient for the calculation of the Dini derivative [30, 31].

Lemma 2. *Let $V_i(x) : \mathbb{R}^d \rightarrow \mathbb{R}$ ($i = 1, \dots, n$) be C^1 and $V(x) = \max_{i=1, \dots, n} V_i(x)$. Let $x(t) \in \mathbb{R}^d$ be an absolutely continuous function over an interval (a, b) . If $I(t) = \{i \in \{1, 2, \dots, n\} : V(x(t)) = V_i(x(t))\}$ is the set of indices where the maximum is reached at t , then $d^+V(x(t)) = \max_{i \in I(t)} V_i(x(t))$, $t \in (a, b)$.*

The following is the well-known LaSalle invariance principle.

Lemma 3 (LaSalle (1968), Thm 3.2 in [29]). *Let $x(t)$ be a solution of (3). Let $V : \mathbb{R}^d \rightarrow \mathbb{R}$ be a continuous, locally Lipschitz function with $d^+V(x(t)) \leq 0$ on $[0, \omega)$. Then $\Lambda^+(x_0)$ is contained in the union of all solutions that remain in $\mathcal{Z} := \{x : d_f^+V(x) = 0\}$ on their maximal intervals of definition.*

3.3 Robust Consensus

The following lemma deals with a robust version of the usual consensus problem, and it is a special case of Theorem 4.1 and Proposition 4.10 in [32].

Lemma 4. *Consider the following network dynamics defined over the digraph \mathcal{G} :*

$$\frac{d}{dt}x_i(t) = \sum_{j \in \mathcal{N}_i} a_{ij}(x_j(t) - x_i(t)) + w_i(t), \quad i = 1, \dots, n \quad (6)$$

where $w_i(t)$ is a piecewise continuous function. Let the initial time be $t = t_*$ and the initial condition be $x(t_*) = x_*$. Let \mathcal{G} contain a directed spanning tree. Denote

$$\|w(t)\|_{[t_*, \infty)} := \max_{i \in \mathcal{V}} \sup_{t \in [t_*, \infty)} |w_i(t)|.$$

Then for any $\epsilon > 0$, there exists $\delta > 0$ such that

$$\|w(t)\|_{[t_*, \infty)} \leq \delta \implies \limsup_{t \rightarrow +\infty} \max_{i, j \in \mathcal{V}} |x_i(t) - x_j(t)| \leq \epsilon$$

for any initial value x_* .

4 Nonempty Interval Intersection: Interval Consensus

Denote $x(t) = [x_1(t) \cdots x_n(t)]^T$ the network state. Let $x_0 = [x_1(0) \cdots x_n(0)]^T$ be the network initial value. The following theorem says that node state consensus can be enforced by the interval constraints node dynamics if the intervals admit some nonempty intersection.

Theorem 1. *Suppose $\bigcap_{m=1}^n \mathcal{I}_m \neq \emptyset$ and let the underlying graph \mathcal{G} be strongly connected. Then along the system (2), for any initial value x_0 , there is a $c^*(x_0) \in \bigcap_{m=1}^n \mathcal{I}_m$ such that $\lim_{t \rightarrow \infty} x_i(t) = c^*$, $i \in \mathcal{V}$.*

Remark 1. It is worth observing that Theorem 1 is not valid if we replace the strongly connectivity of \mathcal{G} with a weaker condition, like \mathcal{G} containing a directed spanning tree. In fact in the latter case only the state of the root nodes matters when achieving an (unconstrained) consensus value, and such states may not be compatible with the saturations imposed e.g., on the leaf nodes, meaning that consensus may not be achieved even when $\bigcap_{m=1}^n \mathcal{I}_m \neq \emptyset$. —————

The condition $\bigcap_{m=1}^n \mathcal{I}_m \neq \emptyset$ is equivalent to $p_* \leq q_*$ with

$$p_* = \max_{i \in \mathcal{V}} p_i, \quad q_* = \min_{i \in \mathcal{V}} q_i.$$

When such condition holds we have $\bigcap_{m=1}^n \mathcal{I}_m = [p_*, q_*]$.

4.1 Proof of Theorem 1

We proceed in steps.

Step 1. Introduce $H(x(t)) = \max \{ \max_{i \in \mathcal{V}} x_i(t), q_* \}$. Clearly H is continuous and locally Lipschitz. If $H(x(t)) > q_*$, then $\max_{i \in \mathcal{V}} x_i(t) > q_*$ for $[t, t + \epsilon)$ for some sufficiently small ϵ . Let $I_0(t) := \{j : x_j(t) = \max_{i \in \mathcal{V}} x_i(t)\}$. As a result, from Lemma 2,

$$d^+ H(x(t)) = d^+ \max_{i \in \mathcal{V}} x_i(t) = \max_{i \in I_0(t)} \dot{x}_i(t) = \max_{i \in I_0(t)} \left(- \sum_{j \in \mathcal{N}_i} a_{ij}(x_i(t) - \psi_j(x_j(t))) \right). \quad (7)$$

Let $i_0 \in I_0(t)$. Then $x_{i_0}(t) \geq x_j(t)$ for all j . Moreover, by definition we have $q_j \geq q_*$, which implies:

- (i) $\psi_j(x_j(t)) \leq x_j(t)$ if $x_j(t) > q_*$;
- (ii) $\psi_j(x_j(t)) \leq q_*$ if $x_j(t) \leq q_*$.

Combining the two cases we can conclude that $x_{i_0}(t) - \psi_j(x_j(t)) \geq 0$ since $x_{i_0}(t) > q_*$. From (7) we further know that $d^+ H(x(t)) \leq 0$ if $H(x(t)) > q_*$. This in fact further assures that if $H(x(t_*)) = q_*$, then $H(x(t)) = q_*$ for all $t \geq t_*$. We have proved that $H(x(t))$ is a nonincreasing function for all t .

Also introduce $h(x(t)) = \min \{ \min_{i \in \mathcal{V}} x_i(t), p_* \}$. The same argument leads to $d^+ h(x(t)) \geq 0$, i.e., $h(x(t))$ is a nondecreasing function for all t . Consequently, for $V(x(t)) = H(x(t)) - h(x(t))$, there holds $d^+ V(x(t)) \leq 0$.

Step 2. Denote¹ $\mathcal{Z} := \{x : d^+ V(x) = 0\}$. In this step, we show $\mathcal{Z} \subseteq [p_*, q_*]^n$ when \mathcal{G} is strongly connected.

We use a contradiction argument. Let $x^* = [x_1^* \cdots x_n^*]^T \in \mathcal{Z}$ with $x^* \notin [p_*, q_*]^n$. Then there must be a node i satisfying $x_i^* \notin [p_*, q_*]$. By symmetry we assume $x_i^* > q_*$, and without loss of generality we let $x_i^* = \max_{j \in \mathcal{V}} x_j^*$. Let us consider a solution $x(t)$ of (2) with $x(0) = x^*$.

Denote $I_* := \{j : x_j^* = x_i^* = \max_{k \in \mathcal{V}} x_k^*\}$. Because \mathcal{G} is strongly connected, along the system (2), nodes in I_* will either be attracted by other nodes (if any) in $\mathcal{V} \setminus I_*$ which hold values strictly smaller than x_i^* , or simply by q_* . Therefore, there is $\epsilon > 0$ such that $x_j(\epsilon) < x_i^*$ for all j . This is to say, $H(x(\epsilon)) < H(x(0))$ and therefore the trajectory cannot be within \mathcal{Z} . We have proved $\mathcal{Z} \subseteq [p_*, q_*]^n$.

Now by Lemma 3, $\Lambda^+(x_0)$ is always contained in \mathcal{Z} , and therefore $\Lambda^+(x_0) \subseteq [p_*, q_*]^n$. Further by Lemma 1, there holds²

$$x(t) \rightarrow [p_*, q_*]^n \quad (8)$$

as $t \rightarrow \infty$.

Step 3. By (8), for any $\delta > 0$, there is a finite $T(\delta) > 0$ such that along (2), there holds

$$|x_i(t) - \psi_i(x_i(t))| \leq \frac{\delta}{\alpha}$$

¹More precisely, it is the Dini derivative of V along system (2). But by (5), there is no harm writing it in this way.

²From the properties of V , each trajectory is obviously contained in a compact set with $\omega = \infty$.

for all $t \geq T(\delta)$ and $i \in \mathcal{V}$, with $\alpha = \max\{|a_{ij}| \cdot \text{card}(\mathcal{N}_i) : (j, i) \in \mathcal{E}\}$. We can therefore rewrite (2) as

$$\frac{d}{dt}x_i(t) = - \sum_{j \in \mathcal{N}_i} a_{ij}(x_i(t) - x_j(t)) + w_i(t), \quad (9)$$

with

$$w_i(t) := \sum_{j \in \mathcal{N}_i} a_{ij}(\psi_j(x_j(t)) - x_j(t)),$$

and conclude that the following claim holds true.

Claim. For any $\delta > 0$, there is a $T(\delta) > 0$ such that $|w_i(t)| \leq \delta$ for all $t \geq T(\delta)$ and $i \in \mathcal{V}$, i.e., $\|w(t)\|_{[T(\delta), \infty)} \leq \delta$.

Consider the sequence $\epsilon_k = \frac{1}{k}$ for $k = 1, 2, \dots$. For any fixed k , Lemma 4 can be invoked to conclude that we can find a δ_k such that if $\|w(t)\|_{[t_*, \infty)} \leq \delta_k$ for some $t_* > 0$, then

$$\limsup_{t \rightarrow +\infty} \max_{i, j \in \mathcal{V}} |x_i(t) - x_j(t)| \leq \epsilon_k. \quad (10)$$

Then, from the above claim, for such δ_k there exists indeed a $T(\delta_k)$ for which $|w_i(t)| \leq \delta_k$ for all $t \geq T(\delta_k)$ and $i \in \mathcal{V}$, i.e., $\|w(t)\|_{[t_*, \infty)} \leq \delta_k$ for $t_* = T(\delta_k)$. As a result, for any fixed k , (10) holds. In other words,

$$\ell(x_0) := \limsup_{t \rightarrow +\infty} \max_{i, j \in \mathcal{V}} |x_i(t) - x_j(t)|$$

is a well-defined constant for which it holds $0 \leq \ell(x_0) \leq \frac{1}{k}$ for all k . As a result, $\ell(x_0) = 0$, hence

$$\lim_{t \rightarrow +\infty} |x_i(t) - x_j(t)| = 0$$

for all $i, j \in \mathcal{V}$.

Step 4. In this step, we finally show that each $x_i(t)$ admits a finite limit. Let c^* be a limit point of $x_j(t)$ for a fixed j . Based on the fact that $\mathcal{Z} \subseteq [p_*, q_*]^n$, there must hold $c^* \in [p_*, q_*]$. If $p_* = q_*$, the result already holds. We assume $p_* < q_*$ in the following.

According to (10), for any $\epsilon > 0$, there exists $t_* > 0$ such that

$$|x_i(t_*) - c^*| \leq \epsilon. \quad (11)$$

for all i . There are three cases.

(i) Let $c^* \in (p_*, q_*)$. We let ϵ be sufficiently small so that

$$p_* < c^* - \epsilon \leq x_i(t_*) \leq c^* + \epsilon < q_*$$

for all i . This means the system (2) is a standard consensus dynamics for $t \geq t_*$ because $\psi_i(x_i(t)) = x_i(t)$ for all $t \geq t_*$. Of course all $x_i(t)$ converge to the same limit, which must be c^* .

- (ii) Let $c^* = q_*$. We let ϵ be sufficiently small so that

$$p_* < q_* - \epsilon < q_*.$$

As a result, $p_* < x_i(t_*) \leq q_* + \epsilon$ for all i . Repeating the argument we used in Step 1, it is easy to see that $\max_{i \in \mathcal{V}} x_i(t)$ is nonincreasing for $t \geq t_*$, and therefore it converges to a finite limit, say M_* . While from (10), $\min_{i \in \mathcal{V}} x_i(t)$ must converge to the same limit M_* . This leaves $c^* = q_* = M_*$ to be the only possibility, and all $x_i(t)$ converge to q_* .

- (iii) Let $c^* = p_*$. The argument is symmetric to Case (ii). By first showing that $\min_{i \in \mathcal{V}} x_i(t)$ converges with $p_* < p_* + \epsilon < q_*$, we know all $x_i(t)$ converge to p_* .

We have now proved that all $x_i(t)$ will converge to a common limit within $[p_*, q_*]$. The proof is complete. \square

5 Empty Interval Intersection: Existence and Stability of Equilibria

In this section, we study the network dynamics (2) when the intervals \mathcal{I}_m admit an empty intersection. To this end, we denote $x(t, y)$ the solution of (2) with the initial condition $x(0) = y$. Denote $\underline{p} = \min_{i \in \mathcal{V}} p_i$ and $\bar{q} = \max_{i \in \mathcal{V}} q_i$. It is obvious that $\text{conv}(\bigcup_{m \in \mathcal{V}} \mathcal{I}_m) = [\underline{p}, \bar{q}]$, where conv denotes the convex hull of a set. It turns out that, regardless of the network topology \mathcal{G} and the intervals \mathcal{I}_m , the nonlinearity of (2) always defines equilibria dynamics.

Theorem 2. *The system (2) has at least one equilibrium. In fact, all equilibria of the system (2) lie within $[\underline{p}, \bar{q}]^n$ if \mathcal{G} is strongly connected.*

Naturally we are interested in the stability of the equilibria. We introduce the following definitions.

Definition 2. An equilibrium $e = [e_1 \cdots e_n]^T$ is an equi-unconstrained equilibrium if e_m is an interior point of $[p_m, q_m]$ for all $m \in \mathcal{V}$; an equi-constrained equilibrium if e_m is an interior point of $\mathbb{R} \setminus [p_m, q_m]$ for all $m \in \mathcal{V}$. \square

Definition 3. An equilibrium e is:

- (i) locally stable if for any $\epsilon > 0$, there exists $\delta > 0$ such that $\|x(t, y)\| \leq \epsilon$ for all $t \geq 0$ and all $\|y - e\| \leq \delta$;
- (ii) locally asymptotically stable if for any $\epsilon > 0$, there exists $\delta > 0$ such that $\|x(t, y)\| \leq \epsilon$ for all $t \geq 0$ and all $\|y - e\| \leq \delta$, and $\lim_{t \rightarrow \infty} x(t, y) = e$ for all $\|y - e\| \leq \delta$.

We present the following result for the stability of equi-unconstrained or equi-constrained equilibria.

Theorem 3. *Suppose $\bigcap_{m=1}^n \mathcal{I}_m = \emptyset$. The following statements hold for the system (2).*

- (i) Every equi-unconstrained equilibrium is locally stable;
- (ii) Every equi-constrained equilibrium $e = [e_1 \cdots e_n]^T$ is locally asymptotically stable if $\mathcal{N}_i \neq \emptyset$ for all $i \in \mathcal{V}$.

Apparently the classes of equi-unconstrained and equi-constrained equilibria only cover a fraction of possible equilibria of the network dynamics. With pairwise disjoint constraint intervals, i.e., $\mathcal{I}_{m_1} \cap \mathcal{I}_{m_2} = \emptyset \forall m_1, m_2 \in \mathcal{V}$, we can establish a full picture regarding the stability of the equilibria.

Theorem 4. Let the graph \mathcal{G} be strongly connected and suppose the \mathcal{I}_m are pairwise disjoint. Then for the system (2) the following statements hold.

- (i) There cannot exist equi-unconstrained equilibria;
- (ii) Every equilibrium is locally asymptotically stable.

We conjecture that the system (2) should have a unique equilibrium which is globally attractive when the interaction graph \mathcal{G} is strongly connected and the \mathcal{I}_m are pairwise disjoint. It seems that there are some major difficulties in establishing such an assertion due to the nonlinear node dynamics. Nonetheless, we manage to prove the following result for directed graphs with the in-degree no more than one at the majority of the nodes.

Theorem 5. Let the graph \mathcal{G} be strongly connected and assume that $\text{card}(\mathcal{N}_m) \leq 2$ for all $m \in \mathcal{V}$ with the equality holding at most for exactly one node. Suppose the \mathcal{I}_m are pairwise disjoint. Then along the system (2), there exists $d^* = [d_1^* \cdots d_n^*]^T \in \mathbb{R}^n$ such that

$$\lim_{t \rightarrow \infty} x_i(t, x_0) = d_i^*, \quad i \in \mathcal{V}$$

for all initial values x_0 .

Remark 2. The value of d_i^* depends on the network topology. For example, following the proof of Theorem 5, assume without loss of generality that $p_1 < p_2 < \cdots < p_n$ and let $i_0 \neq \{1, n\}$ be a node satisfying $\mathcal{N}_{i_0} = \{n\}$; then $d_{i_0}^* = p_n$. If instead $\mathcal{N}_{i_0} = \{1\}$, then $d_{i_0}^* = q_1$. If $\mathcal{N}_{i_0} = \{1, n\}$, then $d_{i_0}^* = \frac{p_n a_{i_0 n} + q_1 a_{i_0 1}}{a_{i_0 n} + a_{i_0 1}}$. ———

Remark 3. The underlying graph \mathcal{G} is termed a symmetric undirected graph if $(i, j) \in \mathcal{V}$ if and only if $(j, i) \in \mathcal{V}$, and $a_{ij} = a_{ji}$ for all $(i, j) \in \mathcal{V}$. Undirected graphs would not help too much to simplify the stability analysis because there can be the case with $\psi_i(x_i) = x_i$ while $\psi_j(x_j) = p_j$. Therefore locally the node interactions could be essentially directional even with bidirectional interactions. ———

5.1 Proof of Theorem 2

We rewrite the system (2) as

$$\frac{d}{dt} x(t) = g(x(t)) = [g_1(x(t)) \cdots g_n(x(t))]^T$$

with $g_i(x(t)) = \sum_{j \in \mathcal{N}_i} a_{ij} (\psi_j(x_j(t)) - x_i(t))$. Now let $x_0 \in [\underline{p}, \bar{q}]^n$. Then it is straightforward to verify that $x(t, x_0) \in [\underline{p}, \bar{q}]^n$ for all $t \geq 0$ because the vector field g is pointing inwards the n-dimensional cube $[\underline{p}, \bar{q}]^n$. This leads to the following lemma.

Lemma 5. *The set $[\underline{p}, \bar{q}]^n$ is positively invariant along the system (2).*

Therefore, $x(t, \cdot)$ defines a continuous mapping from $[\underline{p}, \bar{q}]^n$ to itself. By the famous Brouwer fixed-point theorem, there is at least one point $e \in [\underline{p}, \bar{q}]^n$ satisfying $x(t, e) = e$, i.e., e is a fixed point. We have proved existence of equilibria of the network dynamics within the set $[\underline{p}, \bar{q}]^n$. In order to further prove that there can be no equilibrium outside the set $[\underline{p}, \bar{q}]^n$ when \mathcal{G} is strongly connected, we need the following lemma. We introduce the notation $\text{dist}(x(t), [\underline{p}, \bar{q}]^n) = \|x(t)\|_{[\underline{p}, \bar{q}]^n} = \min_{y \in [\underline{p}, \bar{q}]^n} \|x(t) - y\|$ to indicate the distance between $x(t) \in \mathbb{R}^n$ and the set $[\underline{p}, \bar{q}]^n$.

Lemma 6. *Let the graph \mathcal{G} be strongly connected. Then along the system (2) there holds for any $x_0 \in \mathbb{R}^n$ that*

$$\lim_{t \rightarrow \infty} \text{dist}(x(t), [\underline{p}, \bar{q}]^n) = 0.$$

Proof. Define $\beta(t) = \max_{m \in \mathcal{V}} x_m(t)$ and again let $I_0(t) := \{j : x_j(t) = \max_{i \in \mathcal{V}} x_i(t)\}$. Note that, $\psi_m(x_m(t)) \leq \bar{q}$ for all $m \in \mathcal{V}$ and for all $t \geq 0$. From Lemma 2 and noticing the structure of the node dynamics, there holds that if $\beta(t) \geq \bar{q}$, then

$$\begin{aligned} d^+ \beta(t) &= \max_{i \in I_0(t)} \frac{d}{dt} x_i(t) \\ &= \max_{i \in I_0(t)} \sum_{j \in \mathcal{N}_i} a_{ij} (\psi_j(x_j(t)) - x_i(t)) \\ &\leq \max_{i \in I_0(t)} \sum_{j \in \mathcal{N}_i} a_{ij} (\bar{q} - x_i(t)) \\ &= \max_{i \in I_0(t)} \sum_{j \in \mathcal{N}_i} a_{ij} (\bar{q} - \beta(t)) \\ &\leq \min \{a_{ij} > 0 : (j, i) \in \mathcal{E}\} (\bar{q} - \beta(t)) \end{aligned} \tag{12}$$

when \mathcal{G} is strongly connected. Similarly, $d^+ \beta(t) < 0$ if $\beta(t) > \bar{q}$. As a result, we can obtain $\limsup_{t \rightarrow \infty} \beta(t) \leq \bar{q}$. A symmetric argument leads to the fact that $\liminf_{t \rightarrow \infty} \beta(t) \geq \underline{p}$. We have now proved the desired lemma. \square

Based on Lemma 6, obviously every equilibrium must be within the set $[\underline{p}, \bar{q}]^n$ when \mathcal{G} is strongly connected. This proves Theorem 2.

5.2 Proof of Theorem 3

(i) Let the equilibrium e be an equi-unconstrained equilibrium, i.e., $e_m \in [p_m, q_m]$ for all $m \in \mathcal{V}$.

Assume first that $e_m \in (p_m, q_m)$ for all $m \in \mathcal{V}$. Under this condition on e there exists $\epsilon > 0$ such that for $\mathbb{B}_\epsilon(e) := \{y \in \mathbb{R}^n : \|y - e\| < \epsilon\}$, there holds

$$\frac{d}{dt} x(t) = -Lx(t), \quad x(t) \in \mathbb{B}_\epsilon(e) \tag{13}$$

where L is the network Laplacian. Clearly (13) is standard consensus dynamics. Therefore e is locally stable.

Assume now that $e_m \in [p_m, q_m]$ for all $m \in \mathcal{V}$ and that there exists (at least) an $m \in \mathcal{V}$ such that $e_m \in \{p_m, q_m\}$. Denote \mathcal{V}_q , \mathcal{V}_p and $\bar{\mathcal{V}}^\dagger$ the node sets with $\mathcal{V}_q := \{m \in \mathcal{V} : e_m = q_m\}$, $\mathcal{V}_p := \{m \in \mathcal{V} : e_m = p_m\}$, $\bar{\mathcal{V}}^\dagger := \{m \in \mathcal{V} : e_m \in (p_m, q_m)\}$. Then

$$\delta_i e_i = \sum_{j \in \mathcal{N}_i \cap \mathcal{V}_p} a_{ij} p_j + \sum_{j \in \mathcal{N}_i \cap \mathcal{V}_q} a_{ij} q_j + \sum_{j \in \mathcal{N}_i \cap \bar{\mathcal{V}}^\dagger} a_{ij} e_j, \quad i \in \mathcal{V},$$

where $\delta_i = \sum_{j \in \mathcal{N}_i} a_{ij}$, $i \in \mathcal{V}$.

We divide the analysis considering each of the following 2^n orthants of $\mathbb{B}_\epsilon(e)$ around the equilibrium,

$$\mathbb{N}_\epsilon^s(e) := \left\{ y = [y_1 \cdots y_n]^T : y_i \in \mathcal{J}_i^{s_i}(e, \epsilon) \quad i \in \mathcal{V} \right\} \quad (14)$$

with $s_i \in \{1, 2\}$ and

$$\mathcal{J}_i^{s_i}(e, \epsilon) = \begin{cases} [e_i, e_i + \epsilon], & s_i = 1 \\ (e_i - \epsilon, e_i], & s_i = 2 \end{cases}, \quad i \in \mathcal{V}.$$

We first consider the orthant described by $s_i = 1$ for all $i \in \mathcal{V}$, which we denote $\mathbb{N}_\epsilon^+(e)$,

$$\mathbb{N}_\epsilon^+(e) := \left\{ y = [y_1 \cdots y_n]^T : y_i \in [e_i, e_i + \epsilon] \quad \forall i \in \mathcal{V} \right\}.$$

For sufficiently small ϵ , $x(t) \in \mathbb{N}_\epsilon^+(e)$ implies that $\psi(x_m(t)) = x_m(t)$ if $m \in \bar{\mathcal{V}}^\dagger \cup \mathcal{V}_p$, while $\psi(x_m(t)) = \psi(e_m) = q_m$ if $m \in \mathcal{V}_q$. Hence, when $x(t) \in \mathbb{N}_\epsilon^+(e)$, it follows that

$$\begin{aligned} \frac{d}{dt}(x_i(t) - e_i) &= -\delta_i x_i(t) + \sum_{j \in \mathcal{N}_i \cap (\bar{\mathcal{V}}^\dagger \cup \mathcal{V}_p)} a_{ij} x_j(t) + \sum_{j \in \mathcal{N}_i \cap \mathcal{V}_q} a_{ij} q_j \\ &= -\delta_i(x_i(t) - e_i) + \sum_{j \in \mathcal{N}_i \cap (\bar{\mathcal{V}}^\dagger \cup \mathcal{V}_p)} a_{ij}(x_j(t) - e_j), \quad i \in \mathcal{V}. \end{aligned}$$

The network dynamics can be rewritten as

$$\frac{d}{dt}(x(t) - e) = -\bar{L}(x(t) - e), \quad x(t) \in \mathbb{N}_\epsilon^+(e) \quad (15)$$

where $\bar{L} = [\bar{l}_{ij}]$ is given by

$$\bar{l}_{ij} = \begin{cases} \delta_i, & j = i, \\ -a_{ij}, & j \in \mathcal{N}_i \cap (\bar{\mathcal{V}}^\dagger \cup \mathcal{V}_p), \\ 0, & \text{otherwise.} \end{cases}$$

If $\mathcal{V}_q = \emptyset$, $\bar{\mathcal{V}}^\dagger \cup \mathcal{V}_p \supseteq \mathcal{N}_i$ and $\bar{L} = L$. Otherwise, if $\mathcal{V}_q \neq \emptyset$, then $\bar{L} = \Delta - \bar{A}$ where $\Delta = \text{diag}\{\delta_1, \dots, \delta_n\}$ is the degree matrix, and each element of \bar{A} is given by $\bar{a}_{ij} = a_{ij}$ if $j \in \mathcal{N}_i \cap (\bar{\mathcal{V}}^\dagger \cup \mathcal{V}_p)$ and $\bar{a}_{ij} = 0$ otherwise, for all $i \in \mathcal{V}$. Hence \bar{L} has nonnegative eigenvalues and the zero eigenvalue has equal algebraic and geometric multiplicities.

Moreover, the set $\mathbb{N}^+(e) = \{y = [y_1 \cdots y_n]^T : y_i \geq e_i \ \forall i \in \mathcal{V}\}$ is positively invariant along the network dynamics (2). This is an immediate conclusion from the cooperativity of the system (2) established in Section 3.1. We then conclude that there exists $\delta > 0$ such that for any $x_0 \in \mathbb{N}_\delta^+(e)$, there holds $x(t, x_0) \in \mathbb{N}_\epsilon^+(e)$ for all $t \geq 0$.

To complete the proof we repeat the same reasoning for the other $2^n - 1$ orthants around the equilibrium described by (14). Let $\mathcal{V}^+ = \{m \in \mathcal{V} : x_m \geq e_m\}$ and $\mathcal{V}^- = \{m \in \mathcal{V} : x_m \leq e_m\}$. For sufficiently small ϵ , $x(t) \in \mathbb{N}_\epsilon^s(e)$ implies that

$$\psi(x_m(t)) = \begin{cases} q_m, & m \in \mathcal{V}^+ \cap \mathcal{V}_q =: \mathcal{V}_q^+ \\ p_m, & m \in \mathcal{V}^- \cap \mathcal{V}_p =: \mathcal{V}_p^- \\ x_m(t), & m \in \mathcal{V} \setminus (\mathcal{V}_q^+ \cup \mathcal{V}_p^-). \end{cases}$$

Hence, as before, the network dynamics can be rewritten as

$$\frac{d}{dt}(x(t) - e) = -\bar{L}^s(x(t) - e), \quad x(t) \in \mathbb{N}_\epsilon^s(e) \quad (16)$$

where $\bar{L}^s = [\bar{l}_{ij}^s]$ is given by

$$\bar{l}_{ij}^s = \begin{cases} \delta_i, & j = i, \\ -a_{ij}, & j \in \mathcal{N}_i \cap (\mathcal{V} \setminus (\mathcal{V}_q^+ \cup \mathcal{V}_p^-)), \\ 0, & \text{otherwise.} \end{cases}$$

The problem is then identical to the case $\mathbb{N}_\epsilon^+(e)$.

(ii) Let the equilibrium e be an equi-constrained equilibrium, i.e., $A\psi(e) = \Delta e$ where $\Delta = \text{diag}\{\delta_1, \dots, \delta_n\}$ is the degree matrix with $\delta_i = \sum_{j \in \mathcal{N}_i} a_{ij}$, and e_m is an interior point of $\mathbb{R} \setminus [p_m, q_m]$ for all $m \in \mathcal{V}$. When $x(t) \in \mathbb{B}_\epsilon(e)$, for a sufficiently small ϵ , $\psi(x(t)) = \psi(e)$ and hence

$$\frac{d}{dt}(x(t) - e) = -\Delta x(t) + A\psi(e) = -\Delta(x(t) - e), \quad x(t) \in \mathbb{B}_\epsilon(e). \quad (17)$$

Now let $x_0 \in \mathbb{B}_\epsilon(e)$. Then there exists $\mu > 0$ such that $x(t, x_0) \in \mathbb{B}_\epsilon(e)$ for $t \in [0, \mu]$ simply by continuity of the trajectory. However, along the interval $t \in [0, \mu]$ for the system (17) there holds $\|x(t, x_0) - e\| \leq \|x_0 - e\|$. Therefore again we have shown that $\mathbb{B}_\epsilon(e)$ is an invariant set in such case along the network system (2). The local asymptotical stability of e is then straightforward to verify.

We have now proved Theorem 3.

5.3 Proof of Theorem 4

Without loss of generality we assume $p_1 < p_2 < \dots < p_n$ and therefore $p_1 \leq q_1 < p_2 \leq q_2 < \dots < p_n \leq q_n$. We first establish a technical lemma strengthening the statement of Lemma 6.

Lemma 7. *Let the graph \mathcal{G} be strongly connected. Suppose the \mathcal{I}_m are pairwise disjoint with $p_1 < p_2 < \dots < p_n$. Then along the system (2) there holds that for any $x_0 \in \mathbb{R}^n$, $\lim_{t \rightarrow \infty} \|x(t)\|_{[q_1, p_n]^n} = 0$.*

Proof. Since the \mathcal{I}_m are pairwise disjoint with $p_1 < p_2 < \dots < p_n$, we have $\psi_j(x_j(t)) \leq q_{n-1} < p_n$ for all $j = 1, \dots, n-1$. Thus,

$$\frac{d}{dt}x_n(t) = \sum_{j \in \mathcal{N}_n} a_{nj}(\psi_j(x_j(t)) - x_n(t)) \leq \sum_{j \in \mathcal{N}_n} a_{nj}(q_{n-1} - x_n(t)),$$

which implies that $\limsup_{t \rightarrow \infty} x_n(t) \leq q_{n-1}$. Therefore, for any x_0 , there is $T_1(x_0) > 0$ such that

$$\psi_n(x_n(t)) = p_n, \quad t \geq T_1.$$

Let $\beta(t)$ and $\mathbf{l}_0(t)$ be defined as in the proof of Lemma 6. For $t \geq T_1$, if $\beta(t) \geq p_n$, then

$$\begin{aligned} d^+ \beta(t) &= \max_{i \in \mathbf{l}_0(t)} \frac{d}{dt} x_i(t) = \max_{i \in \mathbf{l}_0(t)} \sum_{j \in \mathcal{N}_i} a_{ij}(\psi_j(x_j(t)) - x_i(t)) \\ &\leq \max_{i \in \mathbf{l}_0(t)} \sum_{j \in \mathcal{N}_i} a_{ij}(p_n - x_i(t)) = \max_{i \in \mathbf{l}_0(t)} \sum_{j \in \mathcal{N}_i} a_{ij}(p_n - \beta(t)) \\ &\leq \min\{a_{ij} > 0 : (j, i) \in \mathcal{E}\}(p_n - \beta(t)) \end{aligned} \quad (18)$$

when \mathcal{G} is strongly connected. This in turn leads to the fact that $\limsup_{t \rightarrow \infty} \beta(t) \leq p_n$. A symmetric argument will give us $\liminf_{t \rightarrow \infty} \beta(t) \geq q_1$ based on the fact that there is $T_2(x_0) > 0$ that $\psi_1(x_1(t)) = q_1$, $t \geq T_2$. We have now completed the proof of the desired lemma. \square

We are now ready to prove Theorem 4, following a similar reasoning to the proof of Theorem 3. Let $e = [e_1 \dots e_n]^T$ be an equilibrium. From Lemma 7 and its proof there must hold $\psi_1(e_1) = q_1$ and $\psi_n(e_n) = p_n$ with $e_1 > q_1$ and $e_n < p_n$. We denote \mathcal{V}^\dagger the node set with $\mathcal{V}^\dagger := \{m \in \mathcal{V} : e_m \in \mathbb{R} \setminus (p_m, q_m)\}$ and $\bar{\mathcal{V}}^\dagger$ the node set with $\bar{\mathcal{V}}^\dagger = \mathcal{V} \setminus \mathcal{V}^\dagger$. Then

$$\delta_i e_i = \sum_{j \in \mathcal{N}_i \cap \mathcal{V}^\dagger} a_{ij} \kappa_j + \sum_{j \in \mathcal{N}_i \cap \bar{\mathcal{V}}^\dagger} a_{ij} e_j, \quad i \in \mathcal{V},$$

where $\delta_i = \sum_{j \in \mathcal{N}_i} a_{ij}$ and $\kappa_j = p_j$ or q_j for all $j \in \mathcal{N}_i \cap \mathcal{V}^\dagger$. There holds $\mathcal{V}^\dagger \neq \emptyset$ since $\{1, n\} \in \mathcal{V}^\dagger$. Let us introduce

$$\mathbb{N}_\epsilon^+(e) := \{y = [y_1 \dots y_n]^T : y_i \in [e_i, e_i + \epsilon] \quad \forall i \in \mathcal{V}\}.$$

For sufficiently small ϵ , $x(t) \in \mathbb{N}_\epsilon^+(e)$ implies that $\psi(x_m(t)) = x_m(t)$ if $m \in \bar{\mathcal{V}}^\dagger$, while $\psi(x_m(t)) = \psi(e_m) = \kappa_m$ if $m \in \mathcal{V}^\dagger$. Hence, when $x(t) \in \mathbb{N}_\epsilon^+(e)$, it follows that

$$\begin{aligned} \frac{d}{dt}(x_i(t) - e_i) &= -\delta_i x_i(t) + \sum_{j \in \mathcal{N}_i \cap \mathcal{V}^\dagger} a_{ij} \kappa_j + \sum_{j \in \mathcal{N}_i \cap \bar{\mathcal{V}}^\dagger} a_{ij} x_j(t) \\ &= -\delta_i (x_i(t) - e_i) + \sum_{j \in \mathcal{N}_i \cap \bar{\mathcal{V}}^\dagger} a_{ij} (x_j(t) - e_j), \quad i \in \mathcal{V}. \end{aligned}$$

The network dynamics can be rewritten as

$$\frac{d}{dt}(x(t) - e) = -\bar{L}(x(t) - e), \quad x(t) \in \mathbb{N}_\epsilon^+(e) \quad (19)$$

where $\bar{L} = [\bar{l}_{ij}]$ depends on the structure of \mathcal{V}^\dagger , and is given by

$$\bar{l}_{ij} = \begin{cases} \delta_i, & j = i, \\ -a_{ij}, & j \in \mathcal{N}_i \cap \bar{\mathcal{V}}^\dagger, \\ 0, & \text{otherwise.} \end{cases}$$

Now that \mathcal{G} is strongly connected and $\{1, n\} \in \mathcal{V}^\dagger$, the matrix $-\bar{L}$ is Hurwitz since it is weakly diagonally dominant and irreducible with negative diagonal entries [33]. Therefore, given I_n as the n dimensional identity matrix, there exists a unique symmetric positive definite matrix P such that

$$P\bar{L} + \bar{L}^T P = I_n. \quad (20)$$

We establish two facts.

- F1. $\mathbb{N}^+(e) = \{y = [y_1 \dots y_n]^T : y_i \geq e_i \ \forall i \in \mathcal{V}\}$ is positively invariant along the network dynamics (2). This is an immediate conclusion from the cooperativity of the system (2) established in Section 3.1.
- F2. From the Lyapunov equation (20) we can routinely obtain

$$\|x(t) - e\|^2 \leq e^{-t/\lambda_{\max}(P)} \frac{\lambda_{\max}(P)}{\lambda_{\min}(P)} \|x(0) - e\|^2$$

along the linear dynamics $\frac{d}{dt}(x(t) - e) = -\bar{L}(x(t) - e)$.

Combining the two facts, we conclude that there exists $\delta > 0$ such that for any $x_0 \in \mathbb{N}_\delta^+(e)$, there holds $x(t, x_0) \in \mathbb{N}_\epsilon^+(e)$ for all $t \geq 0$. Further, it is straightforward to verify that $\lim_{t \rightarrow \infty} x(t, x_0) = e$ if $x_0 \in \mathbb{N}_\delta^+(e)$.

In order to complete the proof we also need to consider the other $2^n - 1$ orthants around the equilibrium

$$\mathbb{N}_\epsilon^s(e) := \left\{ y = [y_1 \dots y_n]^T : y_i \in \mathcal{J}_i^{s_i}(e, \epsilon) \ i \in \mathcal{V} \right\}$$

with $s_i \in \{1, 2\}$ and

$$\mathcal{J}_i^{s_i}(e, \epsilon) = \begin{cases} [e_i, e_i + \epsilon], & s_i = 1 \\ (e_i - \epsilon, e_i], & s_i = 2 \end{cases}, \quad i \in \mathcal{V}.$$

For each $\mathbb{N}_\epsilon^s(e)$ we can use the transform

$$\begin{cases} y_i = -x_i + 2e_i, & \text{if } \mathcal{J}_i^{s_i}(e, \epsilon) = (e_i - \epsilon, e_i]; \\ y_i = x_i, & \text{if } \mathcal{J}_i^{s_i}(e, \epsilon) = [e_i, e_i + \epsilon]. \end{cases}$$

Then the problem will become identical to the case of $\mathbb{N}_\epsilon^+(e)$. This concluded the proof of Theorem 4.

5.4 Proof of Theorem 5

The proof relies on the intermediate statements in the proof of Lemma 7 that

$$\psi_1(x_1(t)) = q_1, \quad \psi_n(x_n(t)) = p_n$$

for all $t \geq \max\{T_1, T_2\}$.

Since $\text{card}(\mathcal{N}_m) \leq 2$ for all $m \in \mathcal{V}$ with the equality holding at most for exactly one node and since \mathcal{G} is strongly connected, there must be a node $i_0 \notin \{1, n\}$ satisfying $\mathcal{N}_{i_0} = \{1\}$, $\mathcal{N}_{i_0} = \{n\}$, or $\mathcal{N}_{i_0} = \{1, n\}$. Consequently, from the network dynamics (2) it is obvious that there is a $d_{i_0}^*$ such that $\lim_{t \rightarrow \infty} x_{i_0}(t) = d_{i_0}^*$. The continuity of $\psi_{i_0}(\cdot)$ in turn implies

$$\lim_{t \rightarrow \infty} \psi_{i_0}(x_{i_0}(t)) = \psi_{i_0}(d_{i_0}^*).$$

Next, there exists a node $i_1 \neq i_0$ with $\mathcal{N}_{i_1} \in \{i_0, 1, n\}$ since \mathcal{G} is strongly connected. From the fact that $\psi_m(x_m(t))$ converges to finite limits for $m \in \{i_0, 1, n\}$, we further know that there is a $d_{i_1}^*$ such that $\lim_{t \rightarrow \infty} x_{i_1}(t) = d_{i_1}^*$. This can be shown using the following argument: let $\lim_{t \rightarrow \infty} b(t) = b^*$ and consider

$$\dot{a}(t) = -(a(t) - b(t)) = -(a(t) - b^*) + \xi(t) \quad (21)$$

with $\xi(t) = b(t) - b^*$. Since $\xi(t) \rightarrow 0$, then $a(t) \rightarrow b^*$. We will now apply this argument to our problem.

If $\mathcal{N}_{i_1} = \{i_0\}$, then the evolution of $x_{i_1}(t)$ is described by

$$\dot{x}_{i_1}(t) = a_{i_1 i_0}(-(x_{i_1}(t) - \psi_{i_0}(x_{i_0}(t))). \quad (22)$$

Let $a(t) := x_{i_1}(t)$, $b(t) := \psi_{i_0}(x_{i_0}(t))$, $b^* := \psi_{i_0}(d_{i_0}^*)$, and $\xi(t) := \psi_{i_0}(x_{i_0}(t)) - \psi_{i_0}(d_{i_0}^*)$. Then (22) can be rewritten in a similar form as (21),

$$\dot{a}(t) = a_{i_1 i_0}(-(a(t) - b^*) + \xi(t)).$$

Since $\xi(t) \rightarrow 0$ as $t \rightarrow \infty$, then $a(t) \rightarrow b^*$, i.e., $x_{i_1}(t) \rightarrow d_{i_1}^* := \psi_{i_0}(d_{i_0}^*)$. Instead, if $\mathcal{N}_{i_1} = \{i_0, 1\}$, then the evolution of $x_{i_1}(t)$ is described by

$$\dot{x}_{i_1}(t) = - \sum_{j=i_0,1} a_{i_1 j} x_{i_1}(t) + \sum_{j=i_0,1} a_{i_1 j} \psi_j(x_j(t)) \quad (23)$$

Let $a(t) := x_{i_1}(t)$, $b(t) := (a_{i_1 i_0} + a_{i_1 1})^{-1}(a_{i_1 i_0} \psi_{i_0}(x_{i_0}(t)) + a_{i_1 1} \psi_1(x_1(t)))$, $b^* := (a_{i_1 i_0} + a_{i_1 1})^{-1}(a_{i_1 i_0} \psi_{i_0}(d_{i_0}^*) + a_{i_1 1} q_1)$ and $\xi(t) := b(t) - b^*$. Then (23) can be rewritten as

$$\dot{a}(t) = (a_{i_1 i_0} + a_{i_1 1}) \cdot (-(a(t) - b^*) + \xi(t)).$$

Since $\xi(t) \rightarrow 0$ as $t \rightarrow \infty$, then $a(t) \rightarrow b^*$, i.e., $x_{i_1}(t) \rightarrow d_{i_1}^* := \frac{a_{i_1 i_0} \psi_{i_0}(d_{i_0}^*) + a_{i_1 1} q_1}{a_{i_1 i_0} + a_{i_1 1}}$.

Similarly, if $\mathcal{N}_{i_1} = \{i_0, n\}$, then $x_{i_1}(t) \rightarrow d_{i_1}^* := \frac{a_{i_1 i_0} \psi_{i_0}(d_{i_0}^*) + a_{i_1 n} p_n}{a_{i_1 i_0} + a_{i_1 n}}$. This recursion can be repeated until all nodes in the set \mathcal{V} have been visited, which implies the conclusion of Theorem 5.

6 Nonempty Interval Intersection in Discrete-Time

Let us consider the discrete-time network dynamics analogous to (2) as below:

$$\begin{aligned} x_i(t+1) &= x_i(t) + \varepsilon \sum_{j \in \mathcal{N}_i} a_{ij} (\psi_j(x_j(t)) - x_i(t)) \\ &= \left(1 - \varepsilon \sum_{j \in \mathcal{N}_i} a_{ij}\right) x_i(t) + \varepsilon \sum_{j \in \mathcal{N}_i} a_{ij} \psi_j(x_j(t)) \end{aligned} \quad (24)$$

for all $i \in \mathcal{V}$. Clearly (24) is the Euler approximation of (2) with ε a small step size.

Theorem 6. Suppose $\bigcap_{m=1}^n \mathcal{I}_m \neq \emptyset$ and let the underlying graph \mathcal{G} be strongly connected. Suppose $\varepsilon < 1 / \max_{i \in \mathcal{V}} \sum_{j \in \mathcal{N}_i} a_{ij}$. Then along the system (24), for any initial value x_0 , there is $c^*(x_0) \in \bigcap_{m=1}^n \mathcal{I}_m$ such that

$$\lim_{t \rightarrow \infty} x_i(t) = c^*, \quad i \in \mathcal{V}.$$

Proof. The proof has to rely on some new development from the proof of Theorem 1 since we cannot use LaSalle invariance principle. We continue to use the definitions of $H(x(t))$, $h(x(t))$, and $V(x(t))$, but defined over the discrete-time system (24). Again we proceed in steps.

Step 1. In this step, let us establish the monotonicity of the functions $H(x(t))$ and $h(x(t))$. We introduce a function $\mathbf{I}_a^+(\cdot)$ by $\mathbf{I}_a^+(y) = y, y > a$ and $\mathbf{I}_a^+(y) = a, y \leq a$. Therefore,

$$\begin{aligned} H(x(t+1)) &= \mathbf{I}_{q_*}^+ \left(\max_{i \in \mathcal{V}} x_i(t+1) \right) \\ &= \mathbf{I}_{q_*}^+ \left(\max_{i \in \mathcal{V}} \left((1 - \varepsilon \sum_{j \in \mathcal{N}_i} a_{ij}) x_i(t) + \varepsilon \sum_{j \in \mathcal{N}_i} a_{ij} \psi_j(x_j(t)) \right) \right) \\ &\leq \mathbf{I}_{q_*}^+ \left(\max_{i \in \mathcal{V}} \left((1 - \varepsilon \sum_{j \in \mathcal{N}_i} a_{ij}) \mathbf{I}_{q_*}^+(x_i(t)) + \varepsilon \sum_{j \in \mathcal{N}_i} a_{ij} \mathbf{I}_{q_*}^+(x_j(t)) \right) \right) \\ &\leq \mathbf{I}_{q_*}^+ \left(\max_{i \in \mathcal{V}} \left((1 - \varepsilon \sum_{j \in \mathcal{N}_i} a_{ij}) \mathbf{I}_{q_*}^+ \left(\max_{j \in \mathcal{V}} x_j(t) \right) + \varepsilon \sum_{j \in \mathcal{N}_i} a_{ij} \mathbf{I}_{q_*}^+ \left(\max_{j \in \mathcal{V}} x_j(t) \right) \right) \right) \\ &= \mathbf{I}_{q_*}^+ \left(\mathbf{I}_{q_*}^+ \left(\max_{j \in \mathcal{V}} x_j(t) \right) \right) = H(x(t)), \end{aligned} \quad (25)$$

where the first inequality holds due to the definition of $\mathbf{I}_{q_*}^+(\cdot)$, and the second inequality is based on the monotonicity of $\mathbf{I}_{q_*}^+(\cdot)$ as well as the assumption that $\varepsilon < 1 / \max_{i \in \mathcal{V}} \sum_{j \in \mathcal{N}_i} a_{ij}$. We can use a symmetric argument to establish $h(x(t+1)) \geq h(x(t))$.

Step 2. From the conclusion of the previous analysis, there are two constants H_* and h_* such that

$$\lim_{t \rightarrow \infty} H(x(t+1)) = H_*, \quad \lim_{t \rightarrow \infty} h(x(t+1)) = h_*.$$

Note that there always holds $H_* \geq q_* \geq p_* \geq h_*$. In this step, we prove $q_* = H_*$ and $p_* = h_*$.

We use a contradiction argument. Let us assume for the moment $p_* > h_*$ in order to eventually build a contradiction.

Fix a time s and take a node i_0 with $x_{i_0}(s) = \min_{j \in \mathcal{V}} x_j(s) \leq h_* < p_*$. Such a node always exists in view of the fact that $p_* > h_* = \lim_{t \rightarrow \infty} \min\{\min_{j \in \mathcal{V}} x_j(t), p_*\}$. The graph is strongly connected, therefore there must exist a node $i_1 \neq i_0$ that is influenced by node i_0 in the interaction graph, resulting in

$$\begin{aligned} x_{i_1}(s+1) &= (1 - \varepsilon \sum_{j \in \mathcal{N}_{i_1}} a_{i_1 j}) x_{i_1}(s) + \varepsilon \sum_{j \in \mathcal{N}_{i_1}} a_{i_1 j} \psi_j(x_j(s)) \\ &= (1 - \varepsilon \sum_{j \in \mathcal{N}_{i_1}} a_{i_1 j}) x_{i_1}(s) + \varepsilon \sum_{j \neq i_0 \in \mathcal{N}_{i_1}} a_{i_1 j} \psi_j(x_j(s)) + \varepsilon a_{i_1 i_0} \psi_{i_0}(x_{i_0}(s)) \\ &\leq (1 - \varepsilon \sum_{j \in \mathcal{N}_{i_1}} a_{i_1 j}) \mathbf{I}_{q_*}^+ (\max_{j \in \mathcal{V}} x_j(s)) + \varepsilon \sum_{j \neq i_0 \in \mathcal{N}_{i_1}} a_{i_1 j} \mathbf{I}_{q_*}^+ (\max_{j \in \mathcal{V}} x_j(s)) + \varepsilon a_{i_1 i_0} p_* \\ &= (1 - \varepsilon a_{i_1 i_0}) \mathbf{I}_{q_*}^+ (\max_{j \in \mathcal{V}} x_j(s)) + \varepsilon a_{i_1 i_0} p_* \\ &\leq (1 - \theta) \mathbf{I}_{q_*}^+ (\max_{j \in \mathcal{V}} x_j(s)) + \theta p_* \end{aligned} \quad (26)$$

for

$$\theta = \min \left\{ \varepsilon \min_{(i,j) \in \mathcal{E}} \{a_{ij} : i \neq j, a_{ij} \neq 0\}, \min_{i \in \mathcal{V}} \left\{ 1 - \varepsilon \sum_{j \in \mathcal{N}_i} a_{ij} \right\} \right\},$$

where in the first inequality we have used

$$x_{i_1}(s) \leq \mathbf{I}_{q_*}^+ (\max_{j \in \mathcal{V}} x_j(s)), \quad \psi_j(x_j(s)) \leq \mathbf{I}_{q_*}^+ (\max_{j \in \mathcal{V}} x_j(s)), \quad \psi_{i_0}(x_{i_0}(t)) \leq \psi_{i_0}(h_*) \leq p_*,$$

and the second inequality is due to the fact that $\mathbf{I}_{q_*}^+ (\max_{j \in \mathcal{V}} x_j(s)) \geq p_*$ and $\varepsilon a_{i_1 i_0} \geq \theta$.

On the other hand, for node i_0 , we have

$$\begin{aligned} x_{i_0}(s+1) &= (1 - \varepsilon \sum_{j \in \mathcal{N}_{i_0}} a_{i_0 j}) x_{i_0}(s) + \varepsilon \sum_{j \in \mathcal{N}_{i_0}} a_{i_0 j} \psi_j(x_j(s)) \\ &\leq (1 - \varepsilon \sum_{j \in \mathcal{N}_{i_0}} a_{i_0 j}) p_* + (\varepsilon \sum_{j \in \mathcal{N}_{i_0}} a_{i_0 j}) \mathbf{I}_{q_*}^+ (\max_{j \in \mathcal{V}} x_j(s)) \\ &\leq \theta p_* + (1 - \theta) \mathbf{I}_{q_*}^+ (\max_{j \in \mathcal{V}} x_j(s)) \end{aligned}$$

Therefore, for $k = i_0, i_1$, we have

$$x_k(s+1) \leq \theta p_* + (1 - \theta) \mathbf{I}_{q_*}^+ (\max_{j \in \mathcal{V}} x_j(s))$$

Continuing to investigate time instant $s + 2$, we have

$$\begin{aligned}
x_k(s+2) &= (1 - \varepsilon \sum_{j \in \mathcal{N}_k} a_{kj}) x_k(s+1) + \varepsilon \sum_{j \in \mathcal{N}_k} a_{kj} \psi_j(x_j(s+1)) \\
&\leq (1 - \varepsilon \sum_{j \in \mathcal{N}_k} a_{kj}) (\theta p_* + (1 - \theta) \mathbf{I}_{q_*}^+ (\max_{j \in \mathcal{V}} x_j(s))) + (\varepsilon \sum_{j \in \mathcal{N}_k} a_{kj}) \mathbf{I}_{q_*}^+ (\max_{j \in \mathcal{V}} x_j(s)) \\
&= \theta p_* + \theta (\varepsilon \sum_{j \in \mathcal{N}_k} a_{kj}) (\mathbf{I}_{q_*}^+ (\max_{j \in \mathcal{V}} x_j(s)) - p_*) + (1 - \theta) \mathbf{I}_{q_*}^+ (\max_{j \in \mathcal{V}} x_j(s)) \\
&\leq \theta p_* + \theta (1 - \theta) (\mathbf{I}_{q_*}^+ (\max_{j \in \mathcal{V}} x_j(s)) - p_*) + (1 - \theta) \mathbf{I}_{q_*}^+ (\max_{j \in \mathcal{V}} x_j(s)) \\
&= \theta^2 p_* + (1 - \theta^2) \mathbf{I}_{q_*}^+ (\max_{j \in \mathcal{V}} x_j(s)), \quad k = i_0, i_1.
\end{aligned}$$

This recursion gives us

$$x_k(s+\tau) = \theta^\tau p_* + (1 - \theta^\tau) \mathbf{I}_{q_*}^+ (\max_{j \in \mathcal{V}} x_j(s))$$

for $k = i_0, i_1$, $\tau = 1, \dots, n - 1$. Note that i_2 is influenced by either i_0 or i_1 , and without loss of generality we assume it is i_1 that is affecting i_2 . Then

$$\begin{aligned}
x_{i_2}(s+2) &= (1 - \varepsilon \sum_{j \in \mathcal{N}_{i_2}} a_{i_2 j}) x_{i_2}(s+1) + \varepsilon \sum_{j \neq i_1 \in \mathcal{N}_{i_2}} a_{i_2 j} \psi_j(x_j(s+1)) + \varepsilon a_{i_2 i_1} \psi_{i_1}(x_{i_1}(s+1)) \\
&\leq (1 - \varepsilon a_{i_2 i_1}) \mathbf{I}_{q_*}^+ (\max_{j \in \mathcal{V}} x_j(s)) + \varepsilon a_{i_2 i_1} ((1 - \theta) \mathbf{I}_{q_*}^+ (\max_{j \in \mathcal{V}} x_j(s)) + \theta p_*) \\
&\leq (1 - \theta) \mathbf{I}_{q_*}^+ (\max_{j \in \mathcal{V}} x_j(s)) + \theta ((1 - \theta) \mathbf{I}_{q_*}^+ (\max_{j \in \mathcal{V}} x_j(s)) + \theta p_*) \\
&= (1 - \theta^2) \mathbf{I}_{q_*}^+ (\max_{j \in \mathcal{V}} x_j(s)) + \theta^2 p_*. \tag{27}
\end{aligned}$$

A similar recursion leads to

$$x_{i_2}(s+\tau) \leq (1 - \theta^\tau) \mathbf{I}_{q_*}^+ (\max_{j \in \mathcal{V}} x_j(s)) + \theta^\tau p_*$$

for $\tau = 2, \dots, n - 1$. The strong connectivity of the graph allows us to continue the process until all nodes are visited, leading to

$$x_k(s+n-1) \leq (1 - \theta^{n-1}) \mathbf{I}_{q_*}^+ (\max_{j \in \mathcal{V}} x_j(s)) + \theta^{n-1} p_* \tag{28}$$

for $k = i_0, \dots, i_{n-1}$, and thus

$$\max_{j \in \mathcal{V}} x_j(s+n-1) \leq (1 - \theta^{n-1}) \mathbf{I}_{q_*}^+ (\max_{j \in \mathcal{V}} x_j(s)) + \theta^{n-1} p_* \tag{29}$$

with $k = i_0, \dots, i_{n-1}$.

At this point we investigate two cases.

- (1) Let $p_* < H_*$. In this case, for sufficiently large s , $\mathbf{I}_{q_*}^+(\max_{j \in \mathcal{V}} x_j(s))$ will be so close to H_* that

$$(1 - \theta^{n-1})\mathbf{I}_{q_*}^+(\max_{j \in \mathcal{V}} x_j(s)) + \theta^{n-1}p_* < H_*.$$

Therefore, (29) implies that $\max_{j \in \mathcal{V}} x_j(s+n-1) < H_*$ for all s that are sufficiently large. From the definition of $H(x(t))$ and H_* , we can only conclude $H_* = q_*$. As a result, $\max_{j \in \mathcal{V}} x_j(s+n-1) < q_*$ for sufficiently large s , which implies that there exists $T > 0$ such that

$$\psi_j(x_j(t)) \leq \mathbf{I}_{p_*}^+(\max_{j \in \mathcal{V}} x_j(t))$$

for all j and all $t \geq T$.

This means, the term $\mathbf{I}_{q_*}^+(\max_{j \in \mathcal{V}} x_j(s))$ in eqs. (26), (27), (28) can be replaced by $\mathbf{I}_{p_*}^+(\max_{j \in \mathcal{V}} x_j(s))$ for $s > T$. In this case (29) becomes

$$\max_{j \in \mathcal{V}} x_j(s+n-1) \leq (1 - \theta^{n-1})\mathbf{I}_{p_*}^+(\max_{j \in \mathcal{V}} x_j(s)) + \theta^{n-1}p_*$$

for all $s \geq T$. Letting s tend to infinity from both sides of the inequality we know $\limsup_{t \rightarrow \infty} \max_{j \in \mathcal{V}} x_j(t) \leq p_*$.

- (2) Suppose $p_* = H_*$. Then, $\limsup_{t \rightarrow \infty} \max_{j \in \mathcal{V}} x_j(t) \leq p_* = H_*$.

Therefore, there must hold true that $\limsup_{t \rightarrow \infty} \max_{j \in \mathcal{V}} x_j(t) \leq p_*$. On the other hand, $p_* > h_*$ implies that there also holds true $\lim_{t \rightarrow \infty} \min_{j \in \mathcal{V}} x_j(t) = h_*$. An immediate conclusion we can draw from the structure of the algorithm is that it can only be the case $\limsup_{t \rightarrow \infty} \max_{j \in \mathcal{V}} x_j(t) = p_*$, because otherwise there is a node i_* with $\psi_{i_*}(x_{i_*}(t)) = p_*$ for all t that are large enough. However, even $\limsup_{t \rightarrow \infty} \max_{j \in \mathcal{V}} x_j(t) = p_*$ ensures that there must always be nodes whose states are arbitrarily close to p_* for an infinite amount of times, a similar contradiction argument would clarify that in that case $\lim_{t \rightarrow \infty} \min_{j \in \mathcal{V}} x_j(t) = p_*$ holds as well. This contradicts our standing assumption $p_* > h_*$.

We have now proved $p_* = h_*$. A symmetric argument leads to $q_* = H_*$ as well.

Step 3. We rewrite the update of node i as

$$x_i(t+1) = (1 - \varepsilon \sum_{j \in \mathcal{N}_i} a_{ij})x_i(t) + \varepsilon \sum_{j \in \mathcal{N}_i} a_{ij}x_j(t) + w_i(t) \quad (30)$$

with $w_i(t) := \varepsilon \sum_{j \in \mathcal{N}_i} a_{ij}(\psi_j(x_j(t)) - x_j(t))$. Then we can reach

$$\limsup_{t \rightarrow +\infty} \max_{i, j \in \mathcal{V}} |x_i(t) - x_j(t)| = 0. \quad (31)$$

by the robust consensus results for discrete-time dynamics [34]. The final piece of proof for node state convergence follows from the same argument as the proof for continuous-time dynamics, and then we finally have $\lim_{t \rightarrow \infty} x_i(t) = c^*$ for all i with $c^* \in [p_*, q_*]$. This completes the proof. \square

7 Numerical Examples

In this section we first consider a case in which the intervals \mathcal{I}_m have nonempty intersection, and then an empty intersection case. Our third example is a cycle graph also with empty interval intersection for which the equilibrium point can be computed explicitly.

Example 1

In Fig. 1 an example of interval consensus on strongly connected graph with $n = 5$ and adjacency matrix

$$A = \begin{bmatrix} 0 & 0 & 0.3360 & 0 & 0 \\ 0.0451 & 0 & 0.0465 & 0.0104 & 0.0641 \\ 0.2096 & 0 & 0 & 0 & 0.1768 \\ 0.0054 & 0.0012 & 0.0038 & 0 & 0 \\ 0.0759 & 0.1650 & 0 & 0 & 0 \end{bmatrix}$$

is shown in which $\bigcap_{m=1}^n \mathcal{I}_m = [p_*, q_*] \neq \emptyset$. In the left column the consensus value c^* is strictly inside the interval $[p_*, q_*]$. In the right column instead c^* is on the boundary of $[p_*, q_*]$ ($c^* = p_*$) and it is clearly driven there by the saturation on $\psi(x)$. Notice that, unlike for a standard consensus problem, in the process of converging $\max_i \{x_i(t)\} - \min_i \{x_i(t)\}$ is not monotonically decreasing, see Fig. 2. Notice further that x_0 need not belong to $\bigotimes_{i=1}^n [p_i, q_i] = [p_1, q_1] \times \cdots \times [p_n, q_n]$, i.e., convergence is for any $x_0 \in \mathbb{R}^n$. As Fig. 1a and Fig. 1c show, $x_0 \notin \bigotimes_{i=1}^n [p_i, q_i]$ does not necessarily lead to c^* on the boundary of $[p_*, q_*]$.

Example 2

In the $n = 5$ example of Fig. 3, the graph is strongly connected (the same adjacency matrix of Example 1 is used), but the intervals \mathcal{I}_m have empty intersection, i.e., $\bigcap_{m=1}^n \mathcal{I}_m = \emptyset$, and are not pairwise disjoint. Numerically the system (2) admits a unique equilibrium point which is not a consensus value, but which appears to be asymptotically stable in the entire \mathbb{R}^5 .

Example 3

In this example, we still consider empty intersection between the sets, that is, $\bigcap_{m=1}^n \mathcal{I}_m = \emptyset$ or, equivalently, $q_* < p_*$, and in addition we assume that $p_1 \leq p_2 \leq \cdots \leq p_n$ and $q_1 \leq \cdots \leq q_n$. The intervals need not be pairwise disjoint. We show that if the graph is a particular (strongly connected) cycle graph (which has $\text{card}(\mathcal{N}_m) = 1$ for all m , see Fig. 4a), then (2) admits a unique equilibrium point e which is in $[q_*, p_*]^n$. This special case is interesting because it is possible to compute e in an explicit way, directly from the p_m and q_m . The adjacency matrix $A = [a_{ij}]$ has the following cyclic structure:

$$a_{ij} = \begin{cases} a_{i,i+1} \neq 0, & j = i + 1 \\ 0, & j \neq i + 1 \end{cases} \quad \text{for } i = 1, \dots, n-1, \text{ and } a_{nj} = \begin{cases} a_{n1} \neq 0, & j = 1 \\ 0, & j \neq 1. \end{cases}$$

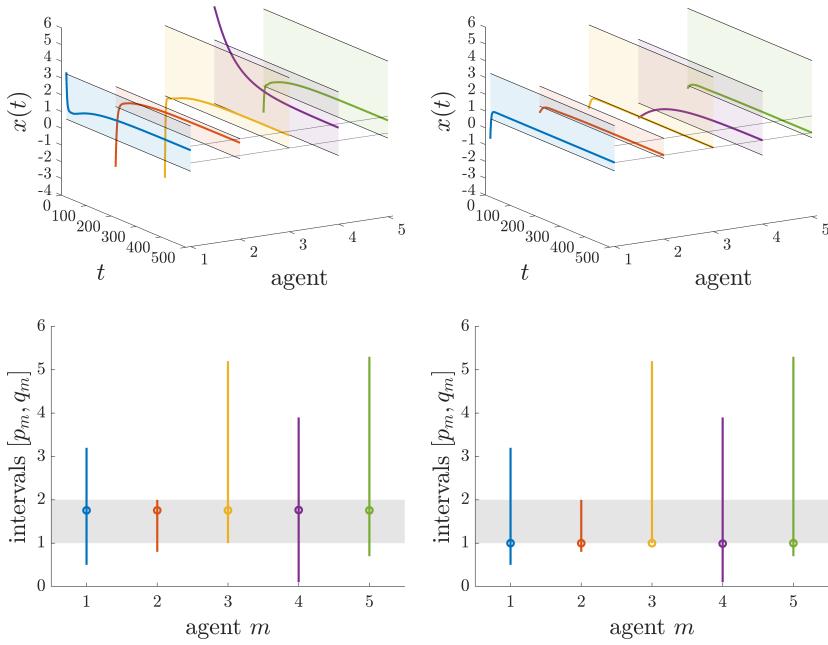


Figure 1: Example 1. Two simulations from different initial conditions. Top row: the trajectories $x(t)$ of the agents are shown in solid color lines. For each agent the shaded region represents the intervals $[p_m, q_m]$, while the transversal dotted lines are p_* and q_* . Bottom row: Intervals $[p_m, q_m]$ for each of the 5 agents, and consensus value c^* (circle) are shown in color, while the gray shaded region corresponds to $[p_*, q_*]$. In the left column we have $c^* \in (p_*, q_*)$, while in the right column $c^* = p_*$.

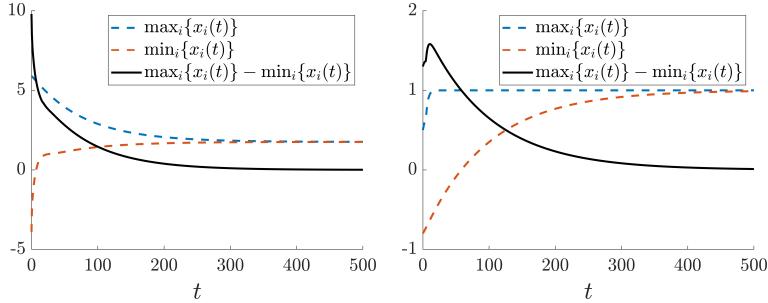


Figure 2: Example 1. Values of $\max_i\{x_i(t)\}$ (blue), $\min_i\{x_i(t)\}$ (red), and $\max_i\{x_i(t)\} - \min_i\{x_i(t)\}$ (black) for the simulations in Fig. 1. It can be seen that $\max_i\{x_i(t)\} - \min_i\{x_i(t)\}$ is not monotonically decreasing in the second case.

In this case (2) becomes

$$\begin{aligned} \frac{d}{dt}x_i(t) &= a_{i,i+1}(\psi_{i+1}(x_{i+1}(t)) - x_i(t)), \quad i = 1, \dots, n-1 \\ \frac{d}{dt}x_n(t) &= a_{n1}(\psi_1(x_1(t)) - x_n(t)) \end{aligned} \tag{32}$$

and, from Theorem 2, it admits at least one equilibrium point, which is in $[\underline{p}, \bar{q}]^n = [p_1, q_n]^n$. Let e be an equilibrium point of (32), that is

$$\begin{cases} e_i = \psi_{i+1}(e_{i+1}), & i = 1, \dots, n-1 \\ e_n = \psi_1(e_1). \end{cases}$$

From Theorem 2, we know that $e_1 \geq p_1$, which implies that $\psi_1(e_1) = e_1$ if $e_1 \leq q_1$ or $\psi_1(e_1) = q_1$ if $e_1 > q_1$. Then

$$e_n = \begin{cases} e_1, & \text{if } e_1 \leq q_1 \\ q_1, & \text{if } e_1 > q_1 \end{cases}, \quad e_{n-1} = \begin{cases} \psi_n(e_1), & \text{if } e_1 \leq q_1 \\ \psi_n(q_1), & \text{if } e_1 > q_1 \end{cases} = \begin{cases} p_n, & \text{if } e_1 \leq q_1 \\ p_n, & \text{if } e_1 > q_1 \end{cases} = p_n$$

because $q_1 = q_* < p_* = p_n$. Therefore

$$e_{n-2} = \psi_{n-1}(p_n) = \begin{cases} p_n, & \text{if } p_n \leq q_{n-1} \\ q_{n-1}, & \text{if } p_n > q_{n-1} \end{cases}$$

because $p_n \geq p_{n-1}$, and

$$\begin{aligned} e_{n-3} = \psi_{n-2}(e_{n-2}) &= \begin{cases} \psi_{n-2}(p_n), & \text{if } p_n \leq q_{n-1} \\ \psi_{n-2}(q_{n-1}), & \text{if } p_n > q_{n-1} \end{cases} = \begin{cases} q_{n-2}, & \text{if } p_n \in (q_{n-2}, q_{n-1}] \\ p_n, & \text{if } p_n \leq q_{n-2} \\ q_{n-2}, & \text{if } p_n > q_{n-1} \end{cases} \\ &= \begin{cases} q_{n-2}, & \text{if } p_n > q_{n-2} \\ p_n, & \text{if } p_n \leq q_{n-2}. \end{cases} \end{aligned}$$

Iterating yields

$$e_{n-i} = \begin{cases} q_{n-i+1}, & \text{if } p_n > q_{n-i+1} \\ p_n, & \text{if } p_n \leq q_{n-i+1} \end{cases} \quad i = 1, \dots, n-1$$

and in particular

$$e_1 = \begin{cases} q_2, & \text{if } p_n > q_2 \\ p_n, & \text{if } p_n \leq q_2. \end{cases}$$

Since $q_2 \geq q_1$ and $p_n > q_1$, it follows that $e_1 \geq q_1$ and hence that $e_n = q_1$. In conclusion, the system (2) admits a unique equilibrium point e such that

$$\begin{aligned} e_n &= q_1 \\ e_{n-1} &= p_n \\ e_{n-i} &= \begin{cases} q_{n-i+1}, & \text{if } p_n > q_{n-i+1} \\ p_n, & \text{if } p_n \leq q_{n-i+1} \end{cases} \quad i = 2, \dots, n-1 \end{aligned}$$

Following the same reasoning as the proof of Theorem 4, the equilibrium must be locally asymptotically stable. Moreover, it must be $e \in [q_1, p_n]^n = [q_*, p_*]^n$. Figure 4b and Fig. 4c shows the result for a cycle graph of size $n = 10$ nodes and edges weight drawn from a uniform distribution. The asymptotic stability character of the unique equilibrium point is confirmed, see Fig. 4b.

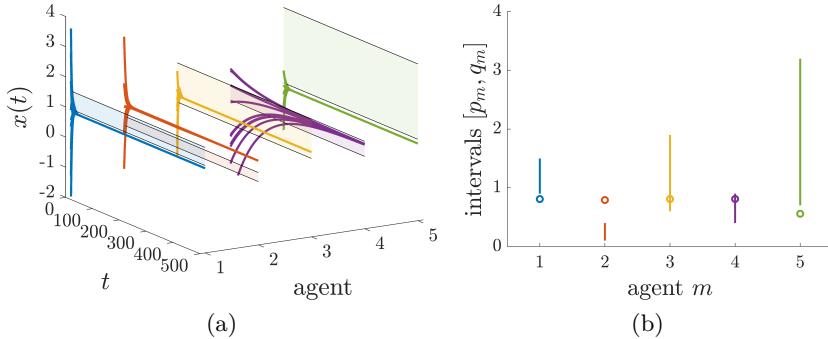


Figure 3: Example 2. (a): Trajectories $x(t)$ of the system (2) from 10 random initial conditions. Shaded region: p_m and q_m for each agent. (b): Intervals $[p_m, q_m]$ for each agent and equilibrium e (circle).

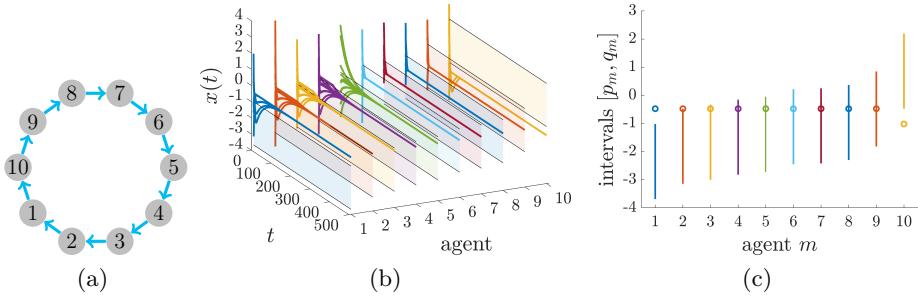


Figure 4: Example 3. (a): Cycle Graph. (b): Trajectories $x(t)$ of the system (2) from 10 random initial conditions. (c): Intervals $[p_m, q_m]$ for each agent and equilibrium e (circle).

8 Conclusion

If a group of agents seeking a consensus has non-dispensable requests on the range of values that such a consensus can achieve, then standard consensus algorithms must be replaced by something more sophisticated. The scheme proposed in this paper, interval consensus, allows to do this efficiently in both continuous and discrete-time with the only (unavoidable) prerequisite that the intersection of the agent intervals is nonempty.

To complete the understanding of our saturated dynamics (2) some work still need to be done for the cases with empty interval intersection. In particular, the mixed case of an equilibrium which is neither equi-constrained nor equi-unconstrained (see Example 2) is not treated at all in the paper. The conjecture which we could not fully prove is that the empty interval intersection case always leads to a single (asymptotically stable) equilibrium point.

Bibliography

- [1] A. Fontan, G. Shi, X. Hu, and C. Altafini, “Interval Consensus for Multiagent Networks,” *IEEE Transactions on Automatic Control*, vol. 65, no. 5, pp. 1855–1869, may 2020.
- [2] A. Fontan, G. Shi, X. Hu, and C. Altafini, “Interval consensus: a novel class of constrained consensus problems for multiagent networks,” in *56th IEEE Conference on Decision and Control*. Melbourne, Australia: IEEE, dec 2017, pp. 4155–4160.
- [3] R. Olfati-Saber and R. Murray, “Consensus Problems in Networks of Agents With Switching Topology and Time-Delays,” *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1520–1533, sep 2004.
- [4] Y. Hong, J. Hu, and L. Gao, “Tracking control for multi-agent consensus with an active leader and variable topology,” *Automatica*, vol. 42, no. 7, pp. 1177–1182, 2006.
- [5] G. Shi, W. Xia, and K. H. Johansson, “Convergence of max-min consensus algorithms,” *Automatica*, vol. 62, pp. 11–17, 2015.
- [6] C. Altafini, “Consensus Problems on Networks With Antagonistic Interactions,” *IEEE Transactions on Automatic Control*, vol. 58, no. 4, pp. 935–946, apr 2013.
- [7] A. Nedic, A. Ozdaglar, and P. A. Parrilo, “Constrained consensus and optimization in multi-agent networks,” *IEEE Transactions on Automatic Control*, vol. 55, no. 4, pp. 922–938, 2010.
- [8] X. Wang and Z. Zhou, “Projection-based consensus for continuous-time multi-agent systems with state constraints,” in *54th IEEE Conference on Decision and Control*. Osaka, Japan: IEEE, 2015, pp. 1060–1065.
- [9] U. Lee and M. Mesbahi, “Constrained consensus via logarithmic barrier functions,” in *50th IEEE Conference on Decision and Control and European control Conference*. Orlando, FL, USA: IEEE, 2011, pp. 3608–3613.
- [10] G. Shi, K. H. Johansson, and Y. Hong, “Reaching an optimal consensus: Dynamical systems that compute intersections of convex sets,” *IEEE Transactions on Automatic Control*, vol. 58, no. 3, pp. 610–622, 2013.
- [11] Y. Cao, “Consensus of multi-agent systems with state constraints: a unified view of opinion dynamics and containment control,” in *2015 American Control Conference (ACC)*. Chicago, IL, USA: IEEE, jul 2015, pp. 1439–1444.
- [12] W. Meng, Q. Yang, J. Si, and Y. Sun, “Consensus Control of Nonlinear Multiagent Systems with Time-Varying State Constraints,” *IEEE Transactions on Cybernetics*, vol. 47, no. 8, pp. 2110–2120, 2017.

- [13] C. Sun, C. J. Ong, and J. K. White, “Consensus control of multi-agent system with constraint - The scalar case,” in *52nd IEEE Conference on Decision and Control*. Florence, Italy: IEEE, 2013, pp. 7345–7350.
- [14] M. E. Valcher and I. Zorzan, “New results on the solution of the positive consensus problem,” in *55th IEEE Conference on Decision and Control*. Las Vegas, USA: IEEE, 2016, pp. 5251–5256.
- [15] Z. Liu and Z. Chen, “Discarded Consensus of Network of Agents With State Constraint,” *IEEE Transactions on Automatic Control*, vol. 57, no. 11, pp. 2869–2874, nov 2012.
- [16] M. Barić and F. Borrelli, “Distributed averaging with flow constraints,” in *2011 American Control Conference (ACC)*, San Francisco, CA, USA, 2011, pp. 4834–4839.
- [17] H. Su, M. Z. Chen, J. Lam, and Z. Lin, “Semi-global leader-following consensus of linear multi-agent systems with input saturation via low gain feedback,” *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 60, no. 7, pp. 1881–1889, 2013.
- [18] Y. Li, J. Xiang, and W. Wei, “Consensus problems for linear time-invariant multi-agent systems with saturation constraints,” *IET Control Theory and Applications*, vol. 5, no. 6, pp. 823–829, 2011.
- [19] S. Manfredi and D. Angeli, “Consensus for nonlinear monotone networks with unilateral interactions,” in *55th IEEE Conference on Decision and Control*. Las Vegas, USA: IEEE, dec 2016, pp. 2609–2614.
- [20] V. D. Blondel, J. M. Hendrickx, and J. N. Tsitsiklis, “Continuous-time average-preserving opinion dynamics with opinion-dependent communications,” *SIAM Journal on Control and Optimization*, vol. 48, no. 8, pp. 5214–5240, 2010.
- [21] V. D. Blondel, J. M. Hendrickx, and J. N. Tsitsiklis, “On Krause’s multi-agent consensus model with state-dependent connectivity,” *IEEE Transactions on Automatic Control*, vol. 54, no. 11, pp. 2586–2597, 2009.
- [22] C. Altafini and F. Ceragioli, “Signed bounded confidence models for opinion dynamics,” *Automatica*, vol. 93, pp. 114–125, jul 2018.
- [23] D. Meng, M. Du, and Y. Jia, “Interval Bipartite Consensus of Networked Agents Associated With Signed Digraphs,” *IEEE Transactions on Automatic Control*, vol. 61, no. 12, pp. 3755–3770, dec 2016.
- [24] M. H. Degroot, “Reaching a consensus,” *Journal of the American Statistical Association*, vol. 69, no. 345, pp. 118–121, 1974.
- [25] M. D. LeCompte and J. P. Goetz, “Problems of Reliability and Validity in Ethnographic Research,” *Review of Educational Research*, vol. 52, no. 1, p. 31, 1982.

- [26] R. Spano, “Potential sources of observer bias in police observational data,” *Social Science Research*, vol. 34, no. 3, pp. 591–617, 2005.
- [27] H. L. Smith, “Systems of Ordinary Differential Equations Which Generate an Order Preserving Flow. A Survey of Results,” *SIAM Review*, vol. 30, no. 1, pp. 87–113, mar 1988.
- [28] F. Blanchini and S. Miani, *Set-Theoretic Methods in Control*, ser. Systems & Control: Foundations & Applications. Boston, MA: Birkhäuser Boston, 2008.
- [29] N. Rouche, P. Habets, and M. Laloy, *Stability Theory by Liapunov’s Direct Method*. Springer New York, 1977.
- [30] J. M. Danskin, “The Theory of Max-Min, with Applications,” *SIAM Journal on Applied Mathematics*, vol. 14, no. 4, pp. 641–664, jul 1966.
- [31] Z. Lin, B. Francis, and M. Maggiore, “State Agreement for Continuous-Time Coupled Nonlinear Systems,” *SIAM Journal on Control and Optimization*, vol. 46, no. 1, pp. 288–307, jan 2007.
- [32] G. Shi and K. H. Johansson, “Robust Consensus for Continuous-Time Multi-agent Dynamics,” *SIAM Journal on Control and Optimization*, vol. 51, no. 5, pp. 3673–3691, jan 2013.
- [33] O. Taussky, “A Recurring Theorem on Determinants,” *The American Mathematical Monthly*, vol. 56, no. 10, p. 672, dec 1949.
- [34] L. Wang and Z. Liu, “Robust consensus of multi-agent systems with noise,” *Science in China Series F: Information Sciences*, vol. 52, no. 5, pp. 824–834, may 2009.

Paper B

Multiequilibria analysis for a class of collective decision-making networked systems

Authors: Angela Fontan and Claudio Altafini

© [2018] IEEE. *Edited version of the paper:*

A. Fontan and C. Altafini, “Multiequilibria Analysis for a Class of Collective Decision-Making Networked Systems,” *IEEE Transactions on Control of Network Systems*, vol. 5, no. 4, pp. 1931–1940, dec 2018.

Early versions of this work were submitted in:

A. Fontan and C. Altafini, “Investigating mixed-sign equilibria for non-linear collective decision-making systems,” in *56th IEEE Conference on Decision and Control*. Melbourne, Australia: IEEE, dec 2017, pp. 781–786.

Multiequilibria analysis for a class of collective decision-making networked systems

Angela Fontan^{*} and Claudio Altafini^{*}

^{*}Dept. of Electrical Engineering
Linköping University
SE-581 83 Linköping, Sweden
angela.fontan@liu.se, claudio.altafini@liu.se

Abstract

The models of collective decision-making considered in this paper are nonlinear interconnected cooperative systems with saturating interactions. These systems encode the possible outcomes of a decision process into different steady states of the dynamics. In particular, they are characterized by two main attractors in the positive and negative orthant, representing two choices of agreement among the agents, associated to the Perron-Frobenius eigenvector of the system. In this paper we give conditions for the appearance of other equilibria of mixed sign. The conditions are inspired by Perron-Frobenius theory and are related to the algebraic connectivity of the network. We also show how all these equilibria must be contained in a solid disk of radius given by the norm of the equilibrium point which is located in the positive orthant.

1 Introduction

Nonlinear interconnected systems are used in broadly different contexts to describe the collective dynamical behavior of an ensemble of “agents” interacting with each other in a non-centralized manner. They are used for instance to represent collective decision-making by animal groups [3–5], formation of opinion in social communities [6, 7], dynamics of a gene regulatory network [8], or neural networks [9]. Such models often share similar features, like the fact of using first order dynamics at a node and sigmoidal saturation-like functions to describe the interactions among the nodes [9–12]. The latter functional form is instrumental to avoid diverging dynamics. The price to pay for having an effectively “bounded” dynamics is however the appearance of complex dynamical phenomena such as periodic orbits or multiple equilibrium points, which complicate considerably the behavior of the system and its understanding. While (stable) periodic orbits can be ruled out by choosing functional forms which are, beside saturated, also monotone [13, 14], it is in general more difficult to deal with multiple equilibria. These

are often a necessary feature of a model, like for instance, when describing bistability in a biological system. In the literature on agent-based animal groups for instance, the collective decision is typically a selection between two different attractors [3–5]. A scenario that occurs often is that of a finite number of attractors, each with its own basin of attraction, encoding the possible outcomes of the collective decision-making process. This is for instance the model of choice of several classes of neural networks, like the so-called Hopfield [10] or Cohen-Grossberg [15] neural networks. When a neural network is interpreted as an associative memory storage device or is used for pattern recognition, the presence of a high number of stable equilibria increases the storage capacity of the network. The multistability of neural networks has in fact been extensively investigated in recent years [16–20], see also [21] for an overview.

The model adopted in this paper is described in [3, 22, 23]. All interactions are “activatory”, i.e., the adjacency matrix is nonnegative. In addition, it is symmetric or diagonally symmetrizable and irreducible. The model has a Laplacian-like structure at the origin and monotone saturating nonlinearities to represent the interaction terms. The amplitude of the interaction part is modulated by a scalar parameter, interpretable as the strength of the social commitment of the agents, and playing the role of bifurcation parameter. When the parameter is small, the origin is globally stable, as can be easily deduced by (global) diagonal dominance. The interesting dynamics happens when the parameter passes a bifurcation threshold: the origin becomes unstable, and two locally stable equilibria, one positive, the other negative, are created. This is the behavior described in [3]. The bifurcation analysis of [3], however, captures what happens only in a neighborhood of the bifurcation point. Moving away from the bifurcation, all is known is that the positive and negative orthants remain invariant sets for the dynamics, and each keeps having a single asymptotically stable equilibrium point, see [22, 23]. What happens in the remaining orthants is unknown and its investigation is the scope of this paper.

It is useful to look at the neural network literature (in our knowledge the only field that has studied the multiequilibria problem is a systematic way for such interconnected systems). It is known since [10] that for neural networks with connections that are symmetric, monotone increasing and sigmoidal, the equations of motion always lead to convergence to stable steady states. The number of such equilibria is shown to grow exponentially with the number n of “neurons” for various specific models [16–19]. In order to count equilibria, often in these papers one has to resort to nondecreasing piecewise linear functions to describe the saturations, and to obtain algebraic conditions on the equilibria by looking at the corners and at the constant slopes. The existence of the equilibria is also checked through Brouwer fixed-point arguments. None of these methods apply in our case.

The consequence is that in order to investigate the presence of multiple equilibria in our system we have to take a completely different approach. Since the adjacency matrix of our network is nonnegative, we can use Perron-Frobenius theorem and the geometrical considerations that follow from it. The main result of the paper is a necessary and sufficient condition for existence of equilibria outside

$\mathbb{R}_+^n/\mathbb{R}_-^n$, formulated in terms of the second largest positive eigenvalue of the adjacency matrix. Roughly speaking, this condition says that the interval of values of the bifurcation parameter in which no mixed-sign equilibrium can appear is determined by the algebraic connectivity of the Laplacian of the system [24–26].

While our necessary condition is always valid, our proof of sufficiency is based on singularity analysis of bifurcations [27], and is valid under the assumption that the algebraic connectivity has multiplicity one, condition which is generically true [28]. For consensus problems, the role of the algebraic connectivity is well-known: the bigger is the gap between 0 (the least eigenvalue of the Laplacian) and the algebraic connectivity, the more robust the consensus is to model uncertainties, parameter variations, node or link failures, etc. [25]. In the present context, the spectral gap in the Laplacian (or, more properly, in the adjacency matrix), represents the range of values of social commitment of the agents which leads to a choice between two alternative agreement solutions, with a guaranteed global convergence. Beyond the value represented by the algebraic connectivity, however, the system bifurcates again, and a number of (stable and unstable) mixed-sign equilibria appears quickly, destroying the global convergence to the agreement manifold in which the two alternative attractors live. Although we can investigate these extra equilibria only numerically, it is nevertheless possible to compute exactly the region in which they must be. The second analytical result of the paper is in fact that for all values of the bifurcation parameter the mixed-sign equilibria have to have norm less than the norm of the equilibrium in \mathbb{R}_+^n . In other words, all the equilibria of the system must be contained in a ball centered in the origin of \mathbb{R}^n and of radius equal to the norm of the positive equilibrium point. Our numerical analysis shows that stable equilibria tend to localize towards the boundary of this ball, while those near the origin tend to have Jacobians with several positive eigenvalues and a similar number of positive and negative components.

The rest of this work is organized as follows. Preliminary material is introduced in Section 2. The main results are stated in Sections 3 and 4; the first section provides the necessary and sufficient condition for the existence of mixed-sign equilibria, while the second describes the region in which they must be contained. Section 5 finally provides a numerical analysis of the equilibria.

2 Preliminaries

2.1 Concave and convex functions

Let $\mathcal{U} \in \mathbb{R}^n$ be a convex set. A function $f : \mathcal{U} \rightarrow \mathbb{R}$ is *convex* if for all $x_1, x_2 \in \mathcal{U}$ and θ with $0 \leq \theta \leq 1$, we have

$$f(\theta x_1 + (1 - \theta)x_2) \leq \theta f(x_1) + (1 - \theta)f(x_2). \quad (1)$$

It is *strictly convex* if strictly inequality holds in (1) whenever $x_1 \neq x_2$ and $0 < \theta < 1$. We say f is *concave* if $-f$ is convex, and *strictly concave* if $-f$ is strictly convex. Suppose $f : \mathcal{U} \rightarrow \mathbb{R}$ is differentiable. Then f is convex if and only if

$$f(x_1) \geq f(x_2) + \frac{\partial f}{\partial x}(x_2)(x_1 - x_2) \quad (2)$$

for all $x_1, x_2 \in \mathcal{U}$. It is strictly convex if and only if strict inequality holds in (2) for all $x_1 \neq x_2$ and $x_1, x_2 \in \mathcal{U}$.

2.2 Nonnegative matrices and Perron-Frobenius

The set of all $\lambda \in \mathbb{C}$ that are eigenvalues of $A \in \mathbb{R}^{n \times n}$ is called the *spectrum* of A and is denoted by $\Lambda(A)$. The *spectral radius* of A is the nonnegative real number $\rho(A) = \max\{|\lambda| : \lambda \in \Lambda(A)\}$. A matrix $B \in \mathbb{R}^{n \times n}$ is said to be *similar* to a matrix $A \in \mathbb{R}^{n \times n}$, abbreviated $B \sim A$, if there exists a nonsingular matrix $S \in \mathbb{R}^{n \times n}$ such that $B = S^{-1}AS$. If A and B are similar, then they have the same eigenvalues, counting multiplicity. A matrix $A \in \mathbb{R}^{n \times n}$ is said to be *reducible* if either $n = 1$ and $A = 0$ or if $n \geq 2$, there is a permutation matrix $P \in \mathbb{R}^{n \times n}$ and some integer r with $1 \leq r \leq n - 1$ such that $P^TAP = \begin{bmatrix} B & C \\ 0 & D \end{bmatrix}$ where $B \in \mathbb{R}^{r \times r}$, $C \in \mathbb{R}^{r \times (n-r)}$, $D \in \mathbb{R}^{(n-r) \times (n-r)}$. A matrix $A \in \mathbb{R}^{n \times n}$ is said to be *irreducible* if it is not reducible. The matrix $A = [a_{ij}] \in \mathbb{R}^{n \times n}$ is said to be *diagonally dominant* if $|a_{ii}| \geq \sum_{j \neq i} |a_{ij}|$ for all i . It is said to be *strictly diagonally dominant* if $|a_{ii}| > \sum_{j \neq i} |a_{ij}|$ for all i .

Theorem 1 (6.1.10 in [29]). Let $A \in \mathbb{R}^{n \times n}$ be strictly diagonally dominant. Then A is nonsingular. If $a_{ii} > 0$ for all $i = 1, \dots, n$, then every eigenvalue of A has positive real part. If A is symmetric and $a_{ii} > 0$ for all $i = 1, \dots, n$, then A is positive definite.

Theorem 2 (Perron-Frobenius, 1.4 in [30]). If $A \in \mathbb{R}^{n \times n}$ is irreducible and nonnegative then $\rho(A)$ is a real, positive, algebraically simple eigenvalue of A , of right (resp., left) eigenvector $v > 0$ (resp., $w > 0$). Furthermore, for every eigenvalue $\lambda \in \Lambda(A)$ such that $\lambda \neq \rho(A)$ it is $|\lambda| < \rho(A)$ and the corresponding eigenvector v_λ cannot be nonnegative.

Corollary 1 (Perron-Frobenius). If $A \in \mathbb{R}^{n \times n}$ is irreducible and nonnegative then either

$$\rho(A) = \sum_{j=1}^n a_{ij}, \quad \forall i = 1, \dots, n \quad (3)$$

or

$$\min_i \left\{ \sum_{j=1}^n a_{ij} \right\} \leq \rho(A) \leq \max_i \left\{ \sum_{j=1}^n a_{ij} \right\}. \quad (4)$$

2.3 Symmetric, symmetrizable and congruent matrices

Let $A \in \mathbb{R}^{n \times n}$ be symmetric. Then all the eigenvalues of A are real and S^TAS is symmetric for all $S \in \mathbb{R}^{n \times n}$. A matrix $A \in \mathbb{R}^{n \times n}$ is (*diagonally*) *symmetrizable* if DA is symmetric for some diagonal matrix D with positive diagonal entries. The matrix DA is called *symmetrization* of A and the matrix D is called *symmetrizer* of A [11, 31]. The eigenvalues of a symmetrizable matrix are real. $A \in \mathbb{R}^{n \times n}$ is symmetrizable if and only if it is sign symmetric, i.e., $a_{ij} = a_{ji} = 0$ or $a_{ij}a_{ji} > 0$, $\forall i \neq j$, and $a_{i_1 i_2}a_{i_2 i_3} \cdots a_{i_k i_1} = a_{i_2 i_1}a_{i_3 i_2} \cdots a_{i_1 i_k}$ for all i_1, \dots, i_k . A matrix $B \in$

$\mathbb{R}^{n \times n}$ is said to be *congruent* to the matrix $A \in \mathbb{R}^{n \times n}$ if there exists a nonsingular matrix S such that $B = SAS^T$.

Theorem 3 (Ostrowski, 4.5.9 in [29]). Let $A, S \in \mathbb{R}^{n \times n}$ with A symmetric and S nonsingular. Let the eigenvalues of A , SAS^T and SST^T be arranged in nondecreasing order. For each $k = 1, \dots, n$, there exists a positive real number θ_k such that $\lambda_1(SST^T) \leq \theta_k \leq \lambda_n(SST^T)$ and $\lambda_k(SAS^T) = \theta_k \lambda_k(A)$.

2.4 Cooperative systems

Consider the system

$$\dot{x} = f(x), \quad x(0) = x_0 \quad (5)$$

where f is a continuously differentiable function defined on a convex, open set $\mathcal{U} \subseteq \mathbb{R}^n$. We write $x(t, x_0)$ for the forward solution of (5) with initial condition $x_0 \in \mathbb{R}^n$ at $t = 0$. Let $S = \text{diag}\{s_1, \dots, s_n\} \in \mathbb{R}^{n \times n}$, where $s_i = \pm 1$ for all i . The set $S\mathbb{R}^n$ given by $S\mathbb{R}^n = \{x \in \mathbb{R}^n : s_i x_i \geq 0, s_i \in \{\pm 1\}, i = 1, \dots, n\}$ is an orthant in \mathbb{R}^n . $S\mathbb{R}^n$ is a cone in \mathbb{R}^n and it generates a partial ordering " $\leq_{S\mathbb{R}^n}$ ", i.e., $x \leq_{S\mathbb{R}^n} y$ iff $y - x \in S\mathbb{R}^n$. The subscript " $S\mathbb{R}^n$ " will be dropped in case $S\mathbb{R}^n = \mathbb{R}_+^n$, the nonnegative orthant.

The system (5) is said to be type- $S\mathbb{R}^n$ *monotone* [14] if whenever \bar{x} and \bar{y} lie in \mathcal{U} and if $\bar{x} \leq_{S\mathbb{R}^n} \bar{y}$ then $x(t, \bar{x}) \leq_{S\mathbb{R}^n} x(t, \bar{y})$ for all $t \geq 0$ for which both solutions are defined. In this case we say that the flow of (5) preserves the ordering $\leq_{S\mathbb{R}^n}$.

Lemma 1 ([14]). If $f \in C^1(\mathcal{U})$ where \mathcal{U} is open and convex in \mathbb{R}^n , then $x(t, x_0)$ preserves the partial ordering $\leq_{S\mathbb{R}^n}$ for $t \geq 0$ if and only if $S \frac{\partial f}{\partial x}(x)S$ has nonnegative off-diagonal elements for every $x \in \mathcal{U}$, where $S = \text{diag}\{s_1, \dots, s_n\}$, $s_i \in \{\pm 1\}$, is the "signature" of the orthant $S\mathbb{R}^n$.

If $S\mathbb{R}^n = \mathbb{R}_+^n$ then we have the class of cooperative systems. System (5) is said to be *cooperative* in $\mathcal{U} \subset \mathbb{R}^n$ if the differentiable vector field $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is such that the Jacobian matrix $\frac{\partial f}{\partial x}(x)$ is Metzler for all $x \in \mathcal{U}$, that is, $\left[\frac{\partial f}{\partial x}(x) \right]_{ij} \geq 0$ for all $i \neq j$.

3 Multiple equilibria in collective decision-making systems

The class of nonlinear systems considered in this work is the following [3, 22]

$$\dot{x} = -\Delta x + \pi A \psi(x) \quad (6)$$

where $x \in \mathbb{R}^n$, $\pi > 0$ is a scalar parameter, $A = [a_{ij}]$ is the weighted adjacency matrix of the network, $\Delta = \text{diag}\{\delta_1, \dots, \delta_n\}$, and $\psi(x) = [\psi_1(x_1) \cdots \psi_n(x_n)]^T$. The matrix A is assumed to be *nonnegative* with *null diagonal*, *irreducible* and *symmetrizable*. A Laplacian-like assumption links Δ and A : $\delta_i = \sum_j a_{ij}$. In the context of agent-based group decisions, δ_i represents the inertia of the i -th agent

to the development of an opinion, $\psi_i(x_i)$ the capacity of the i -th agent of transmitting its opinion to the other agents, mediated by the pairwise susceptibilities a_{ji} . From the irreducibility assumption on A it follows that each δ_i is strictly positive, $\delta_i > 0$. Denote $\delta_{\min} = \min_i \{\delta_i\}$ and $\delta_{\max} = \max_i \{\delta_i\}$. The parameter π represents a community social effort. See [3, 5] for more details.

The vector of functions $\psi(x)$ is such that each $\psi_i(x_i) : \mathbb{R} \rightarrow \mathbb{R}$ satisfies the following conditions

$$\psi_i(x_i) = -\psi_i(-x_i), \forall x_i \in \mathbb{R} \quad (\text{odd}) \quad (\text{A.1})$$

$$\frac{\partial \psi_i}{\partial x_i}(x_i) > 0, \forall x_i \in \mathbb{R} \text{ and } \frac{\partial \psi_i}{\partial x_i}(0) = 1 \quad (\text{monotone}) \quad (\text{A.2})$$

$$\lim_{x_i \rightarrow \pm\infty} \psi_i(x_i) = \pm 1 \quad (\text{saturated}). \quad (\text{A.3})$$

The assumption (A.3) guarantees the boundedness of the solutions, and together with (A.2) allows to exclude the presence of limit cycles. Typical choices for ψ_i are a (modified) Michaelis-Menten function $\psi_i(x_i) = \frac{x_i}{1+|x_i|}$, $x_i \in \mathbb{R}$ [7], or a (modified) Boltzmann function $\psi_i(x_i) = \frac{1-e^{-2x_i}}{1+e^{-2x_i}}$, $x_i \in \mathbb{R}$ [9] (that is, the hyperbolic tangent function). The versions here proposed satisfy the conditions (A.1)÷(A.3).

Additionally, if each nonlinear function $\psi_i(x_i)$ satisfies the following condition

$$\psi_i(x_i) \begin{cases} \text{strictly convex} & \forall x_i < 0 \\ \text{strictly concave} & \forall x_i > 0 \end{cases} \quad (\text{sigmoidal}) \quad (\text{A.4})$$

then, from (2) and (A.4), $|\psi_i(x_i)| < |x_i|$ for all i and $x_i \neq 0$. While all the examples mentioned above (Michaelis-Menten and Boltzmann functions) satisfy also (A.4), nonsigmoidal functions often satisfy (A.1)÷(A.3) but not (A.4). From these assumptions, it follows that the Jacobian matrix of (6), given by $-\Delta + \pi A \frac{\partial \psi}{\partial x}(x)$, is Metzler. Therefore, the system (6) is cooperative.

It is convenient to rewrite the system (6) in the following (“normalized”) form

$$\dot{x} = \Delta(-x + \pi H \psi(x)), \quad x \in \mathbb{R}^n. \quad (7)$$

where the matrix H is defined as $H := \Delta^{-1} A$. Denote also $H_\pi := \pi H$. Observe that it satisfies some useful properties:

- H_π is nonnegative and irreducible, so Theorem 2 applies.
- All the row sums of H_π are equal to π , that is, $H_\pi \mathbf{1}_n = \pi \mathbf{1}_n$. It follows that $(\pi, \mathbf{1}_n)$ is the Perron-Frobenius eigenpair of H_π .
- As the matrix A is symmetrizable, also H_π is symmetrizable, hence it has real eigenvalues.

We will see below that the existence (and the stability) of multiple equilibria is strictly related to the structure of the spectrum of the matrix H_π .

3.1 Existence of multiple equilibria: a necessary condition

Consider the system (7) (or (6)) where each nonlinear function $\psi_i(\cdot)$ satisfies the properties (A.1)÷(A.4). Let us start by recalling what is known for this system

when we vary the parameter π . By construction, the origin is always an equilibrium point for (7) (or (6)). When $\pi < 1$, $x = 0$ is the only equilibrium point, it is globally asymptotically stable and locally exponentially stable. This follows from diagonal dominance, and can be easily shown by a Lyapunov argument, see [3, 22, 23]. At $\pi = \pi_1 = 1$, the system undergoes a pitchfork bifurcation, the origin becomes a saddle point and two more equilibria emerge, $x^+ \in \mathbb{R}_+^n$ and $x^- \in \mathbb{R}_-^n$ [3]. It follows from the analysis of [22] that, for all $\pi > 1$, $x = 0$ is an unstable equilibrium point, while both x^+ and x^- are locally asymptotically stable with domain of attraction given by (at least) the entire orthant for any $\pi > 1$. \mathbb{R}_+^n (respectively \mathbb{R}_-^n) are in fact invariant for the system (7). What happens outside these two orthants is, however, unknown. When $\pi > 1$ and $\pi - 1$ is sufficiently small, the behavior of the system (7) outside \mathbb{R}_+^n and \mathbb{R}_-^n has been discussed in [3]. Only the three equilibrium points mentioned above are possible, two locally stable [22] and the origin as a saddle point. However, when $\pi > 1$ grows, the bifurcation analysis of [3] does not suffice anymore.

Our task is, therefore, to investigate the behavior of the system (7) when $\pi > 1$ grows and $x \in \mathbb{R}^n$ (case not described by [22] and [3]). In particular, we would like to understand for what interval $(1, \pi_2)$ of the bifurcation parameter π extra equilibria not contained in $\mathbb{R}_+^n/\mathbb{R}_-^n$ cannot appear, and what happens for $\pi > \pi_2$.

The following theorem introduces a necessary condition that has to be verified in order to have an equilibrium point \bar{x} in a generic orthant $S\mathbb{R}^n \neq \mathbb{R}_+^n, \mathbb{R}_-^n$ for the system (7).

Theorem 4. Consider the system (7) where each nonlinear function $\psi_i(\cdot)$ satisfies the properties (A.1)÷(A.4). If the system admits an equilibrium point $\bar{x} \in S\mathbb{R}^n$, where $S\mathbb{R}^n$ is an orthant in \mathbb{R}^n and $S\mathbb{R}^n \neq \mathbb{R}_+^n, \mathbb{R}_-^n$, then $\exists \lambda(H_\pi) \in \Lambda(H_\pi)$ such that $\lambda(H_\pi) > 1$ and $\lambda(H_\pi) \neq \rho(H_\pi)$.

Proof. Let $\bar{x} \in S\mathbb{R}^n$ be an equilibrium point for (7). Because Δ is diagonal and positive definite, from (7) it follows

$$\bar{x} = H_\pi \psi(\bar{x}). \quad (8)$$

First notice that if $\bar{x} \in S\mathbb{R}^n$ also $\psi(\bar{x}) \in S\mathbb{R}^n$, because from (A.1) and (A.2) $\psi_i(x_i)$ keeps the same sign of x_i for all $i = 1, \dots, n$. Introduce the diagonal matrix $M(\bar{x}) = \text{diag}\{m_1(\bar{x}_1), \dots, m_n(\bar{x}_n)\}$, where each element is given by

$$m_i(\bar{x}_i) = \frac{\psi_i(\bar{x}_i)}{\bar{x}_i}, \quad i = 1, \dots, n.$$

Since $\bar{x} >_{S\mathbb{R}^n} 0$, the ratio is well-posed. The dependence of $M(\bar{x})$ from \bar{x} will be omitted from now on. From (A.2) and (A.4) one gets $|\psi_i(x_i)| < |x_i| \quad \forall i, x_i \neq 0$, which leads to $m_i = \frac{\psi_i(\bar{x}_i)}{\bar{x}_i} \in (0, 1)$ for all i . Then $0 < \text{diag}\{M\} < \mathbf{1}_n$. Applying the change of coordinates $\bar{z} = M^{\frac{1}{2}}\bar{x}$, from $\psi(\bar{x}) = M\bar{x}$, we get

$$\bar{z} = M^{\frac{1}{2}}H_\pi M^{\frac{1}{2}}\bar{z}. \quad (9)$$

Eq. (9) represents the eigenvalue equation for the matrix $M^{\frac{1}{2}}H_\pi M^{\frac{1}{2}}$, that is, $(1, \bar{z})$ is an eigenpair of $M^{\frac{1}{2}}H_\pi M^{\frac{1}{2}}$. Furthermore, $\bar{z} \in S\mathbb{R}^n$ since, for each i , $\bar{z}_i = \sqrt{m_i}\bar{x}_i$

and m_i is strictly positive. Observe that $M^{\frac{1}{2}}H_\pi M^{\frac{1}{2}}$ is nonnegative and irreducible and let its eigenvalues be arranged in a nondecreasing order. Theorem 2 states that $\lambda_n(M^{\frac{1}{2}}H_\pi M^{\frac{1}{2}}) = \rho(M^{\frac{1}{2}}H_\pi M^{\frac{1}{2}})$ is real and positive and that its associated eigenvector is real and in the positive orthant of \mathbb{R}^n . Then $\rho(M^{\frac{1}{2}}H_\pi M^{\frac{1}{2}}) > 1$, since the eigenvector associated to 1 is $\bar{z} \in S\mathbb{R}^n \neq \mathbb{R}_+^n, \mathbb{R}_-^n$. To prove that $\exists \lambda(H_\pi) \in \Lambda(H_\pi)$ such that $\lambda(H_\pi) > 1$, we proceed in steps.

Step 1. Since the matrix A is symmetrizable, let $A = D_A S_A$ where D_A is a diagonal matrix with positive diagonal entries and S_A is a symmetric matrix. Notice that the matrix S_A is still irreducible, nonnegative and with null diagonal.

Step 2. The matrix H_π , which can be written as $H_\pi = \pi\Delta^{-1}D_A S_A$, is symmetrizable. Define the matrix $D_H := \pi\Delta^{-1}D_A$, diagonal with positive diagonal entries. Then, $S_A = D_H^{-1}H_\pi$ is the symmetrization of H_π while D_H^{-1} is the symmetrizer of H_π .

Step 3. Consider the matrix \tilde{H}_π defined as

$$\tilde{H}_\pi := D_H^{-\frac{1}{2}}H_\pi D_H^{\frac{1}{2}} = D_H^{\frac{1}{2}}S_A D_H^{\frac{1}{2}}.$$

By construction it is symmetric, nonnegative, irreducible, and similar to H_π . Because H_π and \tilde{H}_π have the same eigenvalues, it is just necessary to prove that $\exists \lambda(\tilde{H}_\pi) \in \Lambda(\tilde{H}_\pi)$ such that $\lambda(\tilde{H}_\pi) > 1$ and $\lambda(\tilde{H}_\pi) \neq \rho(\tilde{H}_\pi)$.

Step 4. The matrices $M^{\frac{1}{2}}H_\pi M^{\frac{1}{2}}$ and $M^{\frac{1}{2}}\tilde{H}_\pi M^{\frac{1}{2}}$ are similar. Indeed

$$M^{\frac{1}{2}}H_\pi M^{\frac{1}{2}} = D_H^{\frac{1}{2}}(M^{\frac{1}{2}}\tilde{H}_\pi M^{\frac{1}{2}})D_H^{-\frac{1}{2}} \sim M^{\frac{1}{2}}\tilde{H}_\pi M^{\frac{1}{2}}.$$

Then, they have the same eigenvalues and, in particular, from equation (9) it follows that $(1, D_H^{-\frac{1}{2}}\bar{z})$ is an eigenpair of the matrix $M^{\frac{1}{2}}\tilde{H}_\pi M^{\frac{1}{2}}$. Since $D_H^{-\frac{1}{2}}\bar{z} \in S\mathbb{R}^n$, it follows that $\exists k \neq n$ such that $\lambda_k(M^{\frac{1}{2}}\tilde{H}_\pi M^{\frac{1}{2}}) = 1$.

Step 5. The matrix $M^{\frac{1}{2}}\tilde{H}_\pi M^{\frac{1}{2}}$ is symmetric and $M^{\frac{1}{2}}$ is nonsingular so it is possible to apply Theorem 3. There exists a positive real number θ_k such that

$$\lambda_k(M^{\frac{1}{2}}\tilde{H}_\pi M^{\frac{1}{2}}) = \theta_k \lambda_k(\tilde{H}_\pi)$$

and

$$\lambda_1(M^{\frac{1}{2}}(M^{\frac{1}{2}})^T) \leq \theta_k \leq \lambda_n(M^{\frac{1}{2}}(M^{\frac{1}{2}})^T)$$

where $\lambda_1(M) = \min_i\{m_i\}$ and $\lambda_n(M) = \max_i\{m_i\}$. Then $\theta_k \leq \max_i\{m_i\} < 1$. Since $k \neq n$ is the index for which $\lambda_k(M^{\frac{1}{2}}\tilde{H}_\pi M^{\frac{1}{2}}) = 1$, it follows

$$1 = \lambda_k(M^{\frac{1}{2}}\tilde{H}_\pi M^{\frac{1}{2}}) = \theta_k \lambda_k(\tilde{H}_\pi) < \lambda_k(\tilde{H}_\pi).$$

Because $k \neq n$, this implies the existence of an eigenvalue $\lambda(\tilde{H}_\pi) \in \Lambda(\tilde{H}_\pi)$ such that $\lambda(\tilde{H}_\pi) > 1$ and $\lambda(\tilde{H}_\pi) \neq \rho(\tilde{H}_\pi)$. Consequently, since H_π and \tilde{H}_π are similar, there exists $\lambda(H_\pi) \in \Lambda(H_\pi)$ such that $\lambda(H_\pi) > 1$ and $\lambda(H_\pi) \neq \rho(H_\pi)$. \square

When instead of (7) the system (6) is considered, then the results are less sharp, since they depend on the diagonal terms, which are not all identical as in (7).

Corollary 2. Consider the system (6), where each nonlinear function $\psi_i(\cdot)$ satisfies the properties (A.1)÷(A.4). If the system admits an equilibrium point $\bar{x} \in S\mathbb{R}^n$, where $S\mathbb{R}^n$ is an orthant in \mathbb{R}^n and $S\mathbb{R}^n \neq \mathbb{R}_+^n, \mathbb{R}_-^n$, then $\exists \lambda(A) \in \Lambda(A)$ such that $\lambda(A) > 0$ and $\lambda(A) \neq \rho(A)$ for which $\pi \lambda(A) > \delta_{\min}$.

Proof. Let the symmetrizable matrix A be written as $A = DS_A$, where D is a diagonal matrix with positive diagonal entries and S_A is a symmetric matrix. Then $H_\pi = \pi \Delta^{-1} DS_A$. Define a new matrix

$$\tilde{H}_\pi := \Delta^{\frac{1}{2}} D^{-\frac{1}{2}} H_\pi D^{\frac{1}{2}} \Delta^{-\frac{1}{2}} = \pi \Delta^{-\frac{1}{2}} \tilde{A} \Delta^{-\frac{1}{2}}$$

where \tilde{A} is defined as $\tilde{A} := D^{-\frac{1}{2}} A D^{\frac{1}{2}} = D^{\frac{1}{2}} S_A D^{\frac{1}{2}}$. By construction, \tilde{H}_π is symmetric, similar to H_π and congruent to \tilde{A} , while \tilde{A} is symmetric, similar to A and congruent to S_A . Because \tilde{H}_π and \tilde{A} are both symmetric, it is possible to apply Theorem 3. To simplify the notation, let $\tilde{S} := \sqrt{\pi} \Delta^{-\frac{1}{2}}$ and $\tilde{H}_\pi = \tilde{S} \tilde{A} \tilde{S}^T$, and the eigenvalues be arranged in a nondecreasing order. Therefore, there exists a positive real number θ_k such that the following conditions hold

$$\lambda_k(\tilde{H}_\pi) = \theta_k \lambda_k(\tilde{A}) \quad (10)$$

$$\lambda_1(\tilde{S} \tilde{S}^T) \leq \theta_k \leq \lambda_n(\tilde{S} \tilde{S}^T) \quad (11)$$

From Theorem 4, $\exists k \neq n$ such that $\lambda_k(H_\pi) > 1$. It follows, by similarity, that $\lambda_k(\tilde{H}_\pi) > 1$. Then, the condition (10) where $\theta_k > 0$ and $k \neq n$ yields $\lambda_k(\tilde{A}) > 0$ and $\lambda_k(\tilde{A}) \neq \rho(\tilde{A})$. Moreover, since $\tilde{S} \tilde{S}^T = \pi \Delta^{-1}$, then

$$\lambda_1(\tilde{S} \tilde{S}^T) = \frac{\pi}{\delta_{\max}}, \quad \lambda_n(\tilde{S} \tilde{S}^T) = \frac{\pi}{\delta_{\min}}.$$

From (10), (11), and the result of Theorem 4, it follows that

$$1 < \frac{\pi}{\delta_{\min}} \lambda_k(\tilde{A}),$$

that is, $\pi \lambda_k(\tilde{A}) > \delta_{\min}$. But \tilde{A} and A are similar, that is, they have the same eigenvalues. Then $\pi \lambda_k(A) > \delta_{\min}$, which concludes the proof. \square

It is possible to relax the assumption (A.4). For each $i = 1, \dots, n$, define the coefficients $\mu_i := \max_{x_i \in \mathbb{R}} \left\{ \frac{\psi_i(x_i)}{x_i} \right\}$ (with $\left. \frac{\psi_i(x_i)}{x_i} \right|_{x_i=0} = 1$); then, let

$$\mu := \max_{i=1, \dots, n} \{\mu_i\}. \quad (12)$$

This means that the condition $|\psi_i(x_i)| \leq \mu |x_i|$ holds for each i and $x_i \in \mathbb{R}$.

Theorem 5. Consider the system (7) where each nonlinear function $\psi_i(\cdot)$ satisfies the properties (A.1)÷(A.3). If the system (7) admits an equilibrium point $\bar{x} >_{S\mathbb{R}^n} 0$, where $S\mathbb{R}^n$ is an orthant in \mathbb{R}^n and $S\mathbb{R}^n \neq \mathbb{R}_+^n, \mathbb{R}_-^n$, then $\exists \lambda(H_\pi) \in \Lambda(H_\pi)$ such that $\lambda(H_\pi) \geq 1/\mu$ and $\lambda(H_\pi) \neq \rho(H_\pi)$, where μ is given by (12).

The proofs of this theorem and of the following corollary are omitted as they are completely analogous to those of Theorem 4 and Corollary 2.

Corollary 3. Consider the system (6) where each nonlinear function $\psi_i(\cdot)$ satisfies the properties (A.1)–(A.3). If the system (6) admits an equilibrium point $\bar{x} \in S\mathbb{R}^n$, where $S\mathbb{R}^n$ is an orthant in \mathbb{R}^n and $S\mathbb{R}^n \neq \mathbb{R}_+^n, \mathbb{R}_-^n$, then $\exists \lambda(A) \in \Lambda(A)$ such that $\lambda(A) > 0$ and $\lambda(A) \neq \rho(A)$, for which $\pi \lambda(A) \geq \frac{\delta_{\min}}{\mu}$, where μ is given by (12).

Remark 1. Observe that if $\bar{x} \in \mathbb{R}^n$ is any equilibrium point for the system (7) (or (6)), also $-\bar{x}$ is an equilibrium point as well. Indeed, as by assumption (A.1) $\psi(-\bar{x}) = -\psi(\bar{x})$, it follows that $H_\pi \psi(-\bar{x}) = -H_\pi \psi(\bar{x}) = -\bar{x}$. _____

Remark 2. The necessary conditions given by Theorem 4 and Theorem 5 imply that in order to have an equilibrium point $\bar{x} \in S\mathbb{R}^n$ for the system (7), where $S\mathbb{R}^n$ is an orthant in \mathbb{R}^n and $S\mathbb{R}^n \neq \mathbb{R}_+^n, \mathbb{R}_-^n$, the number of nodes in the network should be strictly greater than three. The next proposition, in fact, shows that it is impossible for A (and hence for H_π) to have a second positive eigenvalue which differs from the Perron-Frobenius eigenvalue if $n \leq 3$, and is of independent interest. _____

Proposition 1. Let $n \leq 3$ and $A \in \mathbb{R}^{n \times n}$ be an irreducible, symmetrizable, non-negative matrix with null diagonal. Then, A cannot have two different real positive eigenvalues.

Proof. The matrix A is symmetrizable, which implies that its eigenvalues are real. Let them be arranged in a nondecreasing order, that is, $\lambda_n(A) > \lambda_{n-1}(A) \geq \dots \geq \lambda_1(A)$. For all n , it is always true that

$$\sum_{i=1}^n \lambda_i(A) = \text{Tr}(A) = 0, \quad \prod_{i=1}^n \lambda_i(A) = \det(A).$$

Since A is nonnegative and irreducible, it is possible to apply Theorem 2 from which it follows that $\lambda_n(A) = \rho(A) > 0$. Then

- If $n = 2$, $A = \begin{bmatrix} 0 & a_{12} \\ a_{21} & 0 \end{bmatrix}$ and the conditions on its trace and determinant become

$$\begin{cases} \lambda_1(A) + \lambda_2(A) = 0 \\ \lambda_1(A)\lambda_2(A) = -a_{12}a_{21} < 0 \end{cases}$$

which yields $\lambda_1(A) < 0$.

- If $n = 3$, $A = \begin{bmatrix} 0 & a_{12} & a_{13} \\ a_{21} & 0 & a_{23} \\ a_{31} & a_{32} & 0 \end{bmatrix}$ and the conditions on its trace and determinant become

$$\begin{cases} \lambda_1(A) + \lambda_2(A) + \lambda_3(A) = 0 \\ \lambda_1(A)\lambda_2(A)\lambda_3(A) = a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \geq 0 \end{cases}$$

which yields $\lambda_1(A) = -\lambda_3(A) < 0$ and $\lambda_2(A) = 0$ or $\lambda_1(A), \lambda_2(A) < 0$.

□

3.2 A geometric necessary and sufficient condition

The following lemma provides a geometric interpretation of the condition of Theorem 4. Consider the Laplacian $\mathcal{L} = I - H$. Since $\rho(H) = 1$, by construction,

the least eigenvalue $\lambda_1(\mathcal{L}) = 1 - \rho(H)$ is the origin. Recall that the second left-most eigenvalue of \mathcal{L} , $\lambda_2(\mathcal{L})$, is called the algebraic connectivity of \mathcal{L} [26]. If $\lambda_{n-1}(H)$ is the second largest eigenvalue of H , then the algebraic connectivity of \mathcal{L} is $\lambda_2(\mathcal{L}) = 1 - \lambda_{n-1}(H)$.

Lemma 2. *The range of values of π for which no extra equilibrium of the system (7) (other than 0 , x^+ and x^-) can appear is given by $(1, \pi_2)$, with $\pi_2 = \frac{1}{\lambda_{n-1}(H)} > 1$, and, hence, it is determined by the algebraic connectivity $\lambda_2(\mathcal{L})$.*

Proof. For (3), it holds that $\sum_j h_{ij} = \rho(H) = 1 \forall i$. This means that the Laplacian \mathcal{L} has all identical Geršgorin disks, all centered at 1 and passing through the origin:

$$\left\{ s \in \mathbb{C} \text{ s.t. } |s - 1| \leq \sum_{j=1}^n \frac{a_{ij}}{\delta_i} = 1 \right\}.$$

From the Geršgorin's Theorem [29], the eigenvalues of \mathcal{L} are located in the union of the n disks. The least eigenvalue $\lambda_1(\mathcal{L}) = 0$ lies on the boundary of the Geršgorin's disks. The other eigenvalues of \mathcal{L} are strictly inside the disks because of irreducibility and positive semidefiniteness of \mathcal{L} . When $\pi > 1$, all the eigenvalues of $H_\pi = \pi H$ are increased in modulus, and the Geršgorin disks of $I - H_\pi$ are still all centered in 1 but have radius π . The condition of Theorem 4, $\exists \lambda(H_\pi) > 1$, $\lambda(H_\pi) \neq \rho(H_\pi)$ corresponds to the existence of π s.t. $\pi\lambda(H) > 1$ for $\lambda(H) > 0$ and $\lambda(H) \neq 1$, that is, it corresponds to requiring that $\pi > \pi_2$, where $\pi_2 = \frac{1}{\lambda_{n-1}(H)} > 1$. When this happens, $1 - \pi\lambda_{n-1}(H)$ crosses into the left half of the complex plane. \square

The insight given by Lemma 2 allows to show that the condition of Theorem 4 is also sufficient. Since we use bifurcation theory and singularity analysis in the proof, we have however to restrict to the case of algebraic connectivity $\lambda_2(\mathcal{L})$ which is a simple eigenvalue (property which is generic for weighted graphs, see [28]).

Theorem 6. *Consider the system (7) where each nonlinear function $\psi_i(\cdot)$ satisfies the properties (A.1)÷(A.4). Assume further that the second largest eigenvalue of H , $\lambda_{n-1}(H)$ is simple. The system admits an equilibrium point $\bar{x} \in S\mathbb{R}^n$, where $S\mathbb{R}^n$ is an orthant in \mathbb{R}^n and $S\mathbb{R}^n \neq \mathbb{R}_+^n, \mathbb{R}_-^n$, if and only if $\pi > \pi_2 = \frac{1}{\lambda_{n-1}(H)} > 1$.*

Proof. Necessity was proven in Theorem 4. In fact, as shown in Lemma 2, since $\rho(H) = 1$, the condition of Theorem 4 corresponds to $\pi > \pi_2 = \frac{1}{\lambda_{n-1}(H)} > 1$. To show that mixed sign equilibria appear exactly at $\pi = \pi_2 = \frac{1}{\lambda_{n-1}(H)}$, we use bifurcation theory, in particular singularity theory and Lyapunov-Schmidt reduction for pitchfork bifurcations [27], see also [3]. Consider

$$\Phi(x, \pi) = -x + \pi H\psi(x) = 0. \quad (13)$$

Denote v_2 and w_2 , $w_2^T v_2 = 1$, the right and left Fiedler vectors, i.e., the eigenvectors relative to $\lambda_{n-1}(H)$. From Theorem 2, $v_2, w_2 \in S\mathbb{R}^n$ for some $S\mathbb{R}^n \neq \mathbb{R}_+^n, \mathbb{R}_-^n$. Let $J = \frac{\partial \Phi}{\partial x}(0, \pi_2) = -I + \pi_2 H$. Observe that $\text{range}(J) = (\ker(J^T))^\perp = (\text{span}\{w_2\})^\perp$ and that $\ker(J) = \text{span}\{v_2\}$. Let E denote the projection of \mathbb{R}^n onto $\text{range}(J) =$

$(\text{span}\{w_2\})^\perp$, $E = I - v_2 w_2^T$, and $I - E$ the projection onto $\ker(J) = \text{span}\{v_2\}$. Split x accordingly: $x = yv_2 + r$, where $y \in \mathbb{R}$ and $r = Ex \in (\text{span}\{w_2\})^\perp$.

Near $(0, \pi_2)$, (13) can be split into

$$E\Phi(x, \pi) = E(-x + \pi H\psi(x)) = 0 \quad (14a)$$

$$(I - E)\Phi(x, \pi) = (I - E)(-x + \pi H\psi(x)) = 0 \quad (14b)$$

Since $\lambda_{n-1}(H)$ is simple, at $\pi = \pi_2$ (13) has a simple singularity. Hence, for (14a) the implicit function theorem applies and it is possible to express $r = R(yv_2, \pi)$. The explicit expression of $R(\cdot)$ is not needed in what follows. Replacing r in (14b), we get the center manifold

$$\phi(y, \pi) = (I - E)(-yv_2 - R(yv_2, \pi) + \pi H\psi(yv_2 + R(yv_2, \pi))) = 0. \quad (15)$$

Define $g(y, \pi) = w_2^T \phi(y, \pi)$, where $w_2 \in (\text{range}(J))^\perp$. The recognition problem for a pitchfork bifurcation requires computing at $(0, \pi_2)$ the partial derivatives $g_y, g_{yy}, g_{yyy}, g_\pi, g_{\pi y}$. In this case, the calculation is simplified by the fact that $\Phi(x, \pi)$ is odd in x . For instance, since Φ has a singularity at $(0, \pi_2)$, it follows that the directional derivative along v_2 vanishes:

$$\begin{aligned} \Phi_y(0, \pi_2) &= \frac{\partial \Phi}{\partial x}(x, \pi) \Big|_{(0, \pi_2)} v_2 = \left(-I + \pi H \frac{\partial \psi(x)}{\partial x} \right) \Big|_{(0, \pi_2)} v_2 \\ &= (-I + \pi_2 H)v_2 = -v_2 + \pi_2 \lambda_{n-1}(H)v_2 = 0 \end{aligned}$$

where we have used $\frac{\partial \psi(x_i)}{\partial x_i} \Big|_0 = 1$. Similarly (see [27])

$$\Phi_{yy}(0, \pi_2) = \frac{\partial}{\partial x} \left(\frac{\partial \Phi}{\partial x}(x, \pi) \Big|_{(0, \pi_2)} v_2 \right) v_2 = \pi H \frac{\partial}{\partial x} \left(\frac{\partial \psi(x)}{\partial x} v_2 \right) \Big|_{(0, \pi_2)} v_2 = 0$$

because $\frac{\partial^2 \psi(x_i)}{\partial x_i^2} \Big|_0 = 0$, and

$$\Phi_\pi(0, \pi_2) = \frac{\partial \Phi}{\partial \pi}(x, \pi) \Big|_{(0, \pi_2)} = H\psi(x) \Big|_{(0, \pi_2)} = 0$$

since $\psi_i(0) = 0$. The two remaining partial derivatives are

$$\Phi_{\pi y}(0, \pi_2) = \frac{\partial}{\partial \pi} \left(\frac{\partial \Phi}{\partial x}(x, \pi) \Big|_{(0, \pi_2)} v_2 \right) = H \frac{\partial \psi(x)}{\partial x} \Big|_{(0, \pi_2)} v_2 = Hv_2 = \lambda_{n-1}(H)v_2$$

and (using a notation similar to [27, (3.16)])

$$\Phi_{yyy}(0, \pi_2) = \frac{\partial^3 \Phi}{\partial x^3}(x, \pi) \Big|_{(0, \pi_2)} (v_2, v_2, v_2) = \pi_2 H \begin{bmatrix} \beta_1 & & \\ & \ddots & \\ & & \beta_n \end{bmatrix} \begin{bmatrix} v_{2,1}^3 \\ \vdots \\ v_{2,n}^3 \end{bmatrix}$$

where, from (A.4), $\beta_i = \left. \frac{\partial^3 \psi_i(x_i)}{\partial x_i^3} \right|_0 < 0$. Consequently, for the projections along v_2 : $\phi_y(0, \pi_2) = (I - E)\Phi_y(0, \pi_2) = 0$, and, similarly, $\phi_{yy}(0, \pi_2) = 0$, $\phi_\pi(0, \pi_2) = 0$. As for the two nonvanishing partial derivatives: $\phi_{\pi y}(0, \pi_2) = (I - E)\Phi_{\pi y}(0, \pi_2) = \lambda_{n-1}(H)v_2$ and $\phi_{yyy}(0, \pi_2) = (I - E)\Phi_{yyy}(0, \pi_2) = v_2 \sum_{i=1}^n \beta_i w_{2,i} v_{2,i}^3$. Therefore $g_y(0, \pi_2) = g_{yy}(0, \pi_2) = g_\pi(0, \pi_2) = 0$ and $g_{\pi y}(0, \pi_2) = \lambda_{n-1}(H) > 0$, $g_{yyy}(0, \pi_2) = \sum_{i=1}^n \beta_i w_{2,i} v_{2,i}^3 < 0$ since $w_{2,i} v_{2,i}^3 \geq 0 \forall i$, which completes the recognition problem for a pitchfork bifurcation [27]. Hence, at $\pi = \pi_2$ the system crosses a second bifurcation through the origin and two new equilibria appear along $\text{span}\{v_2\}$. Since $v_2 \in S\mathbb{R}^n$, these equilibria must belong to $S\mathbb{R}^n$ and $-S\mathbb{R}^n$. \square

Notice that unlike most arguments based on singularity analysis of bifurcations, our result in Theorem 6 is not a local one, as our proof of necessity (Theorem 4) is nonlocal.

When $\lambda_{n-1}(H)$ has multiplicity higher than one, then singularity analysis based on the normal form of a pitchfork bifurcation does not apply, although we expect that similar sufficiency results can be obtained through more advanced bifurcation theory.

A pictorial view of the situation described in Theorem 6 will be given below in Example 1, see in particular Fig. 1, Fig. 2a and Fig. 2b.

If instead we look at system (6), and at the Laplacian $L = \Delta - A$, then, when $\pi = 1$, the Geršgorin disks are centered at δ_i and have different radii equal to δ_i . However, this cannot be straightforwardly reformulated in terms of $\rho(A)$, as (4) (instead of (3)) now holds: $\delta_{\min} \leq \rho(A) \leq \delta_{\max}$. When exploring the values $\pi > 1$, then the Geršgorin disks of $L_\pi = \Delta - \pi A$ are contained one in the other, according to the corresponding δ_i , and all have nonempty intersection with the left half of \mathbb{C} , see Fig. 2c and Fig. 2d. In this case, while the eigenvalue condition introduced in Theorem 4 represents the necessary and sufficient condition for the negativity of a second leftmost eigenvalue of L_π , the eigenvalue condition introduced in Corollary 2 is just necessary, due to the presence of the diagonal matrix Δ . Using a similar reasoning it is in fact possible to prove that $\lambda_{n-1}(H) = \pi \theta \lambda_{n-1}(A)$, where the value of the positive constant θ is not fixed but $\delta_{\min} \leq \theta \leq \delta_{\max}$.

3.3 Stability properties of multiple equilibria

Theorem 7. Suppose the system (7) admits an equilibrium point $\bar{x} \in S\mathbb{R}^n$, where $S\mathbb{R}^n$ is an orthant in \mathbb{R}^n and $S\mathbb{R}^n \neq \mathbb{R}_+^n, \mathbb{R}_-^n$. If

$$\pi \max_j \left\{ \frac{\partial \psi_j}{\partial x_j}(\bar{x}_j) \right\} < 1, \quad (16)$$

then \bar{x} is locally asymptotically stable. Instead, if

$$\pi \min_j \left\{ \frac{\partial \psi_j}{\partial x_j}(\bar{x}_j) \right\} > 1, \quad (17)$$

then the equilibrium point \bar{x} is unstable.

Proof. Let $\bar{x} \in S\mathbb{R}^n$ be an equilibrium point for (7). To study the behavior of the dynamical system (7) near the equilibrium point \bar{x} , consider the linearization around \bar{x}

$$\dot{x} = \Delta \left(-I + H_\pi \frac{\partial \psi}{\partial x}(\bar{x}) \right) (x - \bar{x}). \quad (18)$$

Under the condition (16), it can be proven that the matrix $I - H_\pi \frac{\partial \psi}{\partial x}(\bar{x})$ is strictly diagonally dominant, that is, $1 > \sum_{j \neq i} [H_\pi]_{ij} \frac{\partial \psi_j}{\partial x_j}(\bar{x}_j)$ for all $i = 1, \dots, n$. Indeed

$$\sum_{j \neq i} [H_\pi]_{ij} \frac{\partial \psi_j}{\partial x_j}(\bar{x}_j) \leq \max_j \left\{ \frac{\partial \psi_j}{\partial x_j}(\bar{x}_j) \right\} \sum_{j \neq i} [H_\pi]_{ij} = \pi \max_j \left\{ \frac{\partial \psi_j}{\partial x_j}(\bar{x}_j) \right\} < 1.$$

Then, it is possible to apply Theorem 1 and state that, under the condition (16), each eigenvalue of the matrix $I - H_\pi \frac{\partial \psi}{\partial x}(\bar{x})$ has strictly positive real part. Therefore, each eigenvalue of the matrix $\Delta \left(-I + H_\pi \frac{\partial \psi}{\partial x}(\bar{x}) \right)$ has strictly negative real part, that is, \bar{x} is locally asymptotically stable.

Now, suppose that (17) holds and consider the linearization around \bar{x} (18). Let $\bar{H}_\pi := H_\pi \frac{\partial \psi}{\partial x}(\bar{x})$ and notice that it is nonnegative and irreducible (therefore it is possible to apply Theorem 2). If the matrix $\bar{H}_\pi - I$ admits an eigenvalue with positive real part, it is possible to conclude that the equilibrium point \bar{x} is unstable. For each $i = 1, \dots, n$, it holds that

$$\sum_{j=1}^n [\bar{H}_\pi]_{ij} = \sum_{j=1}^n [H_\pi]_{ij} \frac{\partial \psi_j}{\partial x_j}(\bar{x}_j) \geq \min_j \left\{ \frac{\partial \psi_j}{\partial x_j}(\bar{x}_j) \right\} \sum_{j=1}^n [H_\pi]_{ij} = \min_j \left\{ \frac{\partial \psi_j}{\partial x_j}(\bar{x}_j) \right\} \pi > 1,$$

under the hypothesis (17). According to Corollary 1, $\rho(\bar{H}_\pi) \geq \min_i \left\{ \sum_{j=1}^n [\bar{H}_\pi]_{ij} \right\}$.

From the previous reasoning, it follows that $\rho(\bar{H}_\pi) > 1$. Therefore, the matrix $\bar{H}_\pi - I$ admits a real positive eigenvalue given by $\rho(\bar{H}_\pi) - 1$, which implies that the equilibrium point \bar{x} is unstable. \square

4 Location of the mixed-sign equilibria

In this section we restrict our analysis to the special case of all identical $\psi_i(x_i)$. In this case, the equilibrium point in the positive orthant has all identical components as shown in the following lemma.

Lemma 3. Consider the system (7) where each $\psi_i(\cdot)$ satisfies the properties (A.1)÷(A.4) and $\psi_i(\xi) = \psi_j(\xi) \forall i, j$ and $\forall \xi \in \mathbb{R}$. When $\pi > 1$, the positive equilibrium point $x^+ \in \mathbb{R}_+^n$ is such that $\frac{\psi_i(x_i^+)}{x_i^+} = \frac{1}{\pi}$. Furthermore x^+ is the Perron-Frobenius (right) eigenvector of H_π .

Proof. For the matrix H_π we have that the Perron-Frobenius eigenvalue is $\rho(H_\pi) = \pi > 1$ and the corresponding eigenvector is $\mathbf{1}_n$. At x^+ , it must be

$H_\pi \psi(x^+) = x^+$. If $\psi(x^+) = \frac{1}{\pi}x^+$ with $x^+ = \alpha \mathbf{1}_n$ for some $\alpha \in (0, \pi)$, then we have $H_\pi \psi(x^+) = \frac{1}{\pi} \alpha H_\pi \mathbf{1}_n = \alpha \mathbf{1}_n$ from which we get the eigenvalue/eigenvector equation $H_\pi \mathbf{1}_n = \pi \mathbf{1}_n$. It follows that $x^+ = \alpha \mathbf{1}_n$ is the Perron-Frobenius eigenvector of H_π . The specific value of α depends on the functional form of $\psi_i(x_i)$. \square

A more important consequence is that for each value of π the positive equilibrium point x^+ provides an upper bound on the norm that any mixed-sign equilibrium \bar{x} can assume.

Theorem 8. Consider the system (7), where each nonlinear function $\psi_i(\cdot)$ satisfies the properties (A.1)÷(A.4) and $\psi_i(\xi) = \psi_j(\xi) \forall i, j$ and $\forall \xi \in \mathbb{R}$. If the system admits an equilibrium point $\bar{x} \in S\mathbb{R}^n$, where $S\mathbb{R}^n$ is an orthant in \mathbb{R}^n and $S\mathbb{R}^n \neq \mathbb{R}_+^n, \mathbb{R}_-^n$, then $\|\bar{x}\|_2 \leq \|x^+\|_2$.

In order to prove the theorem we use the following lemma.

Lemma 4. Consider the system (7), where each nonlinear function $\psi_i(\cdot)$ satisfies the properties (A.1)÷(A.4) and $\psi_i(\xi) = \psi_j(\xi) \forall i, j$ and $\forall \xi \in \mathbb{R}$. If the system admits an equilibrium point $\bar{x} \in S\mathbb{R}^n$, where $S\mathbb{R}^n$ is an orthant in \mathbb{R}^n and $S\mathbb{R}^n \neq \mathbb{R}_+^n, \mathbb{R}_-^n$, then $\frac{\psi_i(\bar{x}_i)}{\bar{x}_i} \geq \frac{1}{\pi}$ for all i .

Proof. Introduce the diagonal matrix $M(\bar{x}) = \text{diag}\{m_1(\bar{x}_1), \dots, m_n(\bar{x}_n)\}$ as in Theorem 4, where each element is given by $m_i(\bar{x}_i) = \frac{\psi_i(\bar{x}_i)}{\bar{x}_i}$, $i = 1, \dots, n$. Since $\bar{x} >_{S\mathbb{R}^n} 0$, the ratio is well-posed. The dependence of $M(\bar{x})$ from \bar{x} will be omitted from now on. Let $m_{\min} = \min_i \left\{ \frac{\psi_i(\bar{x}_i)}{\bar{x}_i} \right\}$ for all i and $l \in \{1, \dots, n\}$ be the index such that $\frac{\psi_l(\bar{x}_l)}{\bar{x}_l} = m_{\min}$ and suppose without loss of generality that $\bar{x}_l > 0$. By definition of m_{\min} and from (A.2) and (A.4), $\psi_i(\bar{x}_i) \leq \psi_l(\bar{x}_l)$ and $\bar{x}_i \leq \bar{x}_l$ for all $i \neq l$. At \bar{x} , it must be $\bar{x} = H_\pi \psi(\bar{x})$, that is, $\bar{x}_i = \sum_{j \neq i} [H_\pi]_{ij} \psi_j(\bar{x}_j)$ for all $i = 1, \dots, n$. This yields

$$\bar{x}_l = \sum_{j \neq l} [H_\pi]_{lj} \psi_j(\bar{x}_j) \leq \left(\sum_{j \neq l} [H_\pi]_{lj} \right) \psi_l(\bar{x}_l) = \pi \psi_l(\bar{x}_l)$$

which implies that $m_{\min} = \frac{\psi_l(\bar{x}_l)}{\bar{x}_l} \geq \frac{1}{\pi}$. Then, since $m_i \geq m_{\min}$ for all $i \neq l$, it follows that $\frac{\psi_i(\bar{x}_i)}{\bar{x}_i} \geq \frac{1}{\pi}$ for all $i = 1, \dots, n$. \square

Proof. [Theorem 8] Let $x^+ \in \mathbb{R}_+^n$ and $\bar{x} \in S\mathbb{R}^n$ be equilibrium points for the system (7). From Lemma 3 and Lemma 4 it follows that $\frac{\psi_i(\bar{x}_i)}{\bar{x}_i} \geq \frac{1}{\pi} = \frac{\psi_i(x_i^+)}{x_i^+}$ for all $i = 1, \dots, n$. This implies that $|\bar{x}_i| \leq x_i^+$ for each i , which yields $|\bar{x}| \leq x^+$. Therefore, $\|\bar{x}\|_2 \leq \|x^+\|_2$. \square

5 Numerical Analysis

In this section we first look at the trajectories of a specific numerical example of size $n = 6$. Then we perform a computational analysis of the properties of the equilibria for a system of size $n = 20$.

Example 1

Consider a network of $n = 6$ nodes of weighted adjacency matrix

$$A = \begin{bmatrix} 0 & 0.1477 & 0 & 0.1698 & 0 & 0.0135 \\ 0.4242 & 0 & 0.2626 & 0.3621 & 0 & 0 \\ 0 & 0.1889 & 0 & 0 & 0.2502 & 0.4158 \\ 0.4036 & 0.2997 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.2427 & 0 & 0 & 0.2513 \\ 0.0301 & 0 & 0.4474 & 0 & 0.2787 & 0 \end{bmatrix}$$

which is irreducible, nonnegative, and symmetrizable. Consider the system (6) (and (7)) and assume that each nonlinear function $\psi_i(\cdot)$ is given by the hyperbolic tangent $\psi_i(x_i) = \tanh(x_i)$, which satisfies the assumptions (A.1)÷(A.4). The Perron-Frobenius eigenvalue of the matrix A is $\rho(A) = 0.706$ and its second largest positive eigenvalue is $\lambda_{n-1}(A) = 0.515$, meaning $\lambda_{n-1}(H) = 0.822$ or $\pi_2 = 1.216$. Hence we have:

- for $\pi < 1$, $\bar{x} = 0$ is the only equilibrium point;
- for $\pi \in (1, 1.216)$, the only equilibria are $\{0, x^+, x^-\}$, with $x^\pm = \pm\alpha\mathbf{1}$, see red branch in Fig. 1a;
- for $\pi > 1.216$, unstable equilibria $\bar{x}_2 \in S\mathbb{R}^n = \text{diag}\{-1, -1, 1, -1, 1, 1\}$ and $-\bar{x}_2 \in -S\mathbb{R}^n$ bifurcate from 0, see blue branch in Fig. 1a.

When π grows further, new bifurcation points soon appear. Fig. 1b shows that these bifurcations are not associated to singularities in the origin, but rather they branch out of $\pm\bar{x}_2$. For instance, choosing $\pi = 1.838$, also the condition $\pi\lambda_{n-1}(A) > \delta_{\min}$ is satisfied. Numerical computations confirm the existence of at least 3 equilibria in $S\mathbb{R}^n$, denoted $\bar{x}_3, \bar{x}_4, \bar{x}_5$, which shows that more than one equilibrium can exist per orthant. The condition presented in Theorem 7 is satisfied for \bar{x}_3 , and this assures its local asymptotic stability. Instead, \bar{x}_4 and \bar{x}_5 are unstable. This is shown in the simulation of Fig. 1c. Following the reasoning introduced in Section 3.2, it is possible to observe that at $\pi = 1.838$ the second leftmost eigenvalue of the matrix $L_\pi = \Delta - \pi A$, $\lambda_2(L_\pi) = -0.302$, is negative. However, notice that even if the necessary condition presented in Corollary 2 is satisfied, the condition $\pi\lambda_{n-1}(A) > \delta_{\max}$ does not hold. Fig. 2 shows the Geršgorin's disks and the eigenvalues of different matrices, respectively $I - H$ (when $\pi = 1$), $I - \pi H$ (when $\pi > 1$), $L = \Delta - A$ (when $\pi = 1$), $L_\pi = \Delta - \pi A$ (when $\pi > 1$). Notice that while it is possible to determine the exact value of $\lambda_2(I - \pi H) = 1 - \pi\lambda_{n-1}(H)$, it is only possible to give a bound for the second leftmost eigenvalue of L_π , i.e., $\delta_{\max}[1 - \lambda_{n-1}(H_\pi)] \leq \lambda_2(L_\pi) \leq \delta_{\min}[1 - \lambda_{n-1}(H_\pi)]$ or $\delta_{\min} - \pi\lambda_{n-1}(A) \leq \lambda_2(L_\pi) \leq \delta_{\max} - \pi\lambda_{n-1}(A)$. Finally, observe that this eigenvalue has to be in the union of all the Geršgorin's disks of L_π , but it is not possible to define a priori the smallest disk that contains $\lambda_2(L_\pi)$.

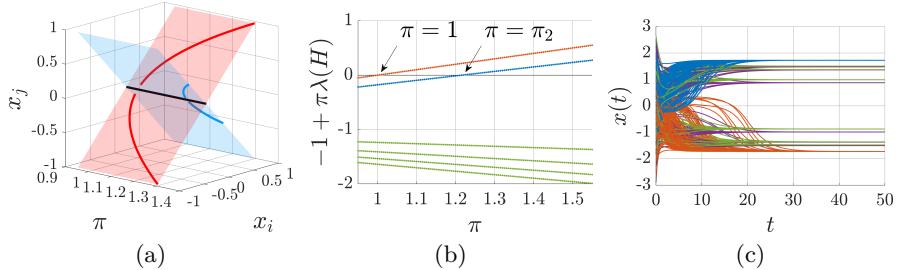


Figure 1: Example 1. (a): Bifurcation diagram near $x = 0$ for two components x_i and x_j . When $\pi = 1$, then the origin bifurcates a first time, and the two locally stable equilibria x^\pm (red branches) are created along the consensus manifold (red plane). When $\pi = \pi_2 = \frac{1}{\lambda_{n-1}(H)}$, a second bifurcation occurs in the origin, with two equilibria (blue branches) locally branching out along $\text{span}\{v_2\}$ (blue plane). (b): The corresponding set of eigenvalues of $-I + \pi H$ (linearization in the origin) as a function of π . The two crossings at $\pi = 1$ and $\pi = \pi_2$ are highlighted. (c): 100 random initial conditions converging to x^+ (blue), x^- (red), \bar{x}_3 (purple) and $-\bar{x}_3$ (green).

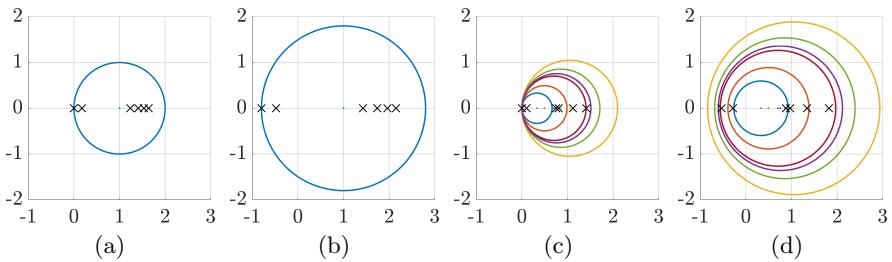


Figure 2: Example 1. Geršgorin's disks and eigenvalues of (a): $I - H$ (when $\pi = 1$). (b): $I - \pi H$ (when $\pi > \pi_2$). (c): $L = \Delta - A$ (when $\pi = 1$). (d): $L_\pi = \Delta - \pi A$ (when $\pi > \pi_2$). The eigenvalues are indicated by cross symbols.

Example 2

A network of $n = 20$ already has $> 10^6$ orthants, all potentially containing equilibria of the system (7). Since, as shown in Example 1, multiple equilibria can appear in the same orthant, this size is already by far out of reach of exhaustive analysis. The results shown in Fig. 3 are for a single (nonnegative, irreducible, symmetrizable) realization A , with edges chosen as in an Erdős-Rényi graph (edge probability $p = 0.1$) and weights drawn from a uniform distribution. All $\psi_i(x_i)$ are chosen equal (again hyperbolic tangent). The results appear to be robust across different realizations of A . Figure 3a shows the number of equilibria (and, in red, the number of orthants to which these equilibria belong) for 500 choices of π uniformly distributed between 1 and 20. For π very small no equilibrium appears, as expected. Equilibria start to appear for values of π that satisfy the condition of Theorem 6. When π is increased further, then the number of equilibria rapidly

grows. For each value of π , 10^4 different initial conditions were tested (we used the `fsoeve` function of Matlab to compute equilibria). The number of equilibria found in this way oscillated between 500 and 600 for a broad range of π values, belonging to 400-500 different orthants. As shown in Theorem 8, for each π all of the equilibria have a norm which is less than the norm of the corresponding positive/negative equilibrium, see Fig. 3b, where the ratio $\frac{\|\bar{x}\|_2}{\|x^+\|_2}$ is shown. These equilibria were tested for local stability. As shown in Fig. 3c, most but not all of them are unstable, with up to 7 unstable eigenvalues in the Jacobian linearization. It is also remarkable that all stable equilibria tend to have high norm $\|\bar{x}\|_2$, i.e., they tend to be near the boundary of the ball of radius $\|x^+\|_2$ to which they have to belong, see Fig. 3c. Notice from Fig. 3b and Fig. 3c that equilibria of small norm ratio $\frac{\|\bar{x}\|_2}{\|x^+\|_2}$, corresponding to small values of π , tend also to have an equal ratio of positive and negative components: in Fig. 3c the radial directions are determined by the fraction of + and - signs of an equilibrium, and the bisectrix of the second and fourth quadrant, corresponding to 50% of + and 50% of -, is where these equilibria tend to be localized.

6 Final considerations and conclusions

In this work we have investigated the presence of fixed points for a particular class of nonlinear interconnected cooperative systems, where the nonlinearities are (strictly) monotonically increasing and saturated. We have proposed necessary and sufficient conditions on the spectrum of the adjacency matrix of the network which are required for the existence of multiple equilibria not contained in $\mathbb{R}_+^n/\mathbb{R}_-^n$ when the nonlinearities assume both sigmoidal and nonsigmoidal shapes. The stability properties of these equilibria have also been investigated. Although we cannot analytically quantify the number of such equilibria, we can locate them in the solid disk whose radius is given by the norm of the positive equilibrium x^+ .

When interpreted in terms of collective decision-making of agent systems, our results can be recapitulated as follows:

- For a low value of the social effort parameter π (i.e., for $\pi < 1$), the agents are not committed enough to reach an agreement;
- For values of π between 1 and $\frac{1}{\lambda_{n-1}(H)}$, the agents have the right dose of commitment to achieve an agreement among two alternative options x^+, x^- ;
- For values of π bigger than $\frac{1}{\lambda_{n-1}(H)}$, the agents start to become overcommitted, which can lead to other possible decisions, depending on the initial conditions of the system. All these extra decisions represent disagreement situations, i.e., they do not belong to \mathbb{R}_+^n or \mathbb{R}_-^n .

Future work includes gaining a better understanding of the bifurcation pattern for $\pi > \frac{1}{\lambda_{n-1}(H)}$, and in presence of “informed agents” in the sense of [3]. It is well-known that the algebraic connectivity is strongly influenced by the topology of the network [25]. We expect that similar arguments apply *tamquam* to

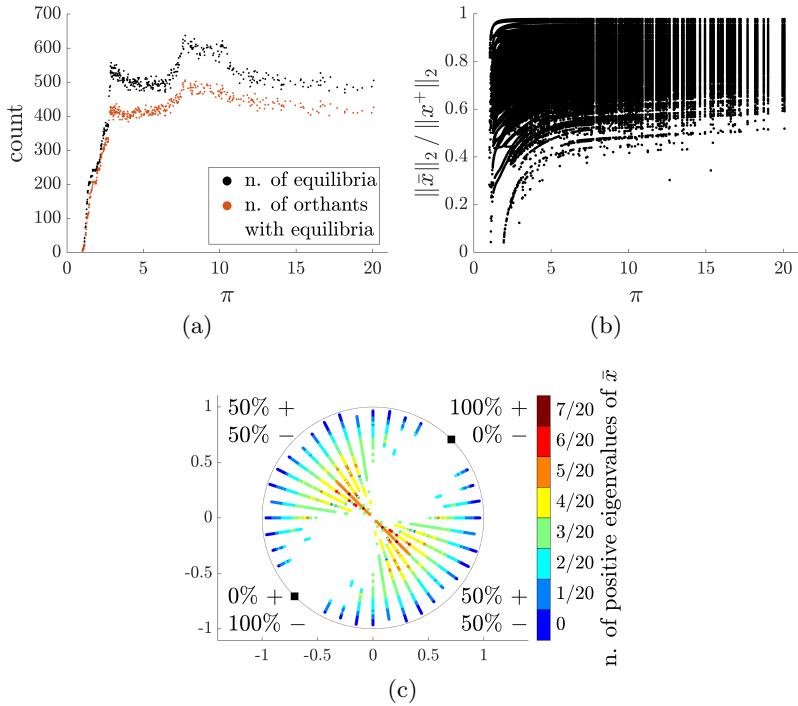


Figure 3: Example 2. Equilibria for a system (7) of size $n = 20$. (a): N. of equilibria (black) found on 10^4 trials for 500 different values of π , and n. of orthants to which these equilibria belong (red). (b): Norm ratio $\frac{\|\bar{x}\|_2}{\|x^+\|_2}$ of the equilibria \bar{x} with respect to the corresponding value of x^+ , as π changes. (c): Distribution of the equilibria according to the norm ratio $\frac{\|\bar{x}\|_2}{\|x^+\|_2}$ (radial distance from origin), to the number of positive eigenvalues of the Jacobian linearization (colormap), and to the n. of components of negative sign (radial angle). Black squares on the unit disk are x^+ and x^- .

$\lambda_{n-1}(H)$. What remains to be checked is whether these topological considerations are applicable to concrete examples of collective decision-making.

Bibliography

- [1] A. Fontan and C. Altafini, “Multiequilibria Analysis for a Class of Collective Decision-Making Networked Systems,” *IEEE Transactions on Control of Network Systems*, vol. 5, no. 4, pp. 1931–1940, dec 2018.
- [2] A. Fontan and C. Altafini, “Investigating mixed-sign equilibria for nonlinear collective decision-making systems,” in *56th IEEE Conference on Decision and Control*. Melbourne, Australia: IEEE, dec 2017, pp. 781–786.
- [3] A. Franci, V. Srivastava, and N. E. Leonard, “A Realization Theory for Bio-inspired Collective Decision-Making,” *arXiv:1503.08526v3*, mar 2017.
- [4] R. Gray, A. Franci, V. Srivastava, and N. Ehrich Leonard, “An agent-based framework for bio-inspired, value-sensitive decision-making,” *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 8238–8243, 2017.
- [5] N. E. Leonard, “Multi-agent system dynamics: Bifurcation and behavior of animal groups,” *Annual Reviews in Control*, vol. 38, no. 2, pp. 171–183, 2014.
- [6] C. Altafini and G. Lini, “Predictable dynamics of opinion forming for networks with antagonistic interactions,” *IEEE Transactions on Automatic Control*, vol. 60, no. 2, pp. 342–357, feb 2015.
- [7] C. Altafini, “Dynamics of Opinion Forming in Structurally Balanced Social Networks,” *PLoS ONE*, vol. 7, no. 6, p. e38135, jun 2012.
- [8] C. Li, L. Chen, and K. Aihara, “Stability of genetic networks with SUM regulatory logic: Lur'e system and LMI approach,” *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 53, no. 11, pp. 2451–2458, 2006.
- [9] M. Kubat and S. Haykin, “Neural networks: a comprehensive foundation,” *The Knowledge Engineering Review*, vol. 13, no. 4, p. 842, 1999.
- [10] J. J. Hopfield, “Neurons with graded response have collective computational properties like those of two-state neurons.” *Proceedings of the National Academy of Sciences*, vol. 81, no. 10, pp. 3088–3092, 1984.
- [11] E. Kaszkurewicz and A. Bhaya, *Matrix Diagonal Stability in Systems and Computation*. Boston: Birkhäuser, 2000.
- [12] D. Siljak, *Large-Scale Dynamic Systems: Stability and Structure*. North-Holland, 1978.
- [13] M. W. Hirsch and H. L. Smith, “Monotone Dynamical Systems,” in *Handbook of Differential Equations: Ordinary Differential Equations*. North-Holland, 2006, vol. 2, ch. 4, pp. 239–357.
- [14] H. L. Smith, “Systems of Ordinary Differential Equations Which Generate an Order Preserving Flow. A Survey of Results,” *SIAM Review*, vol. 30, no. 1, pp. 87–113, mar 1988.

- [15] M. A. Cohen and S. Grossberg, *Artificial Neural Networks: Theoretical Concepts*, V. Vemuri, Ed. IEEE Computer Society Press, 1988.
- [16] C.-Y. Cheng, K.-H. Lin, and C.-W. Shih, “Multistability in Recurrent Neural Networks,” *SIAM Journal on Applied Mathematics*, vol. 66, no. 4, pp. 1301–1320, 2006.
- [17] C.-Y. Cheng, K.-H. Lin, and C.-W. Shih, “Multistability and convergence in delayed neural networks,” *Physica D: Nonlinear Phenomena*, vol. 225, no. 1, pp. 61–74, jan 2007.
- [18] Z. Zeng and W. X. Zheng, “Multistability of neural networks with time-varying delays and concave-convex characteristics,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 23, no. 2, pp. 293–305, 2012.
- [19] Wenlian Lu, Lili Wang, and Tianping Chen, “On Attracting Basins of Multiple Equilibria of a Class of Cellular Neural Networks,” *IEEE Transactions on Neural Networks*, vol. 22, no. 3, pp. 381–394, mar 2011.
- [20] M. Forti and A. Tesi, “New conditions for global stability of neural networks with application to linear and quadratic programming problems,” *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, vol. 42, no. 7, pp. 354–366, jul 1995.
- [21] H. Zhang, Z. Wang, and D. Liu, “A Comprehensive Review of Stability Analysis of Continuous-Time Recurrent Neural Networks,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 25, no. 7, pp. 1229–1262, jul 2014.
- [22] P. U. Abara, F. Ticozzi, and C. Altafini, “Spectral Conditions for Stability and Stabilization of Positive Equilibria for a Class of Nonlinear Cooperative Systems,” *IEEE Transactions on Automatic Control*, vol. 63, no. 2, pp. 402–417, feb 2018.
- [23] P. U. Abara, F. Ticozzi, and C. Altafini, “Existence, uniqueness and stability properties of positive equilibria for a class of nonlinear cooperative systems,” in *54th IEEE Conference on Decision and Control*. Osaka, Japan: IEEE, dec 2015, pp. 4406–4411.
- [24] N. M. M. de Abreu, “Old and new results on algebraic connectivity of graphs,” *Linear Algebra and Its Applications*, vol. 423, no. 1, pp. 53–73, 2007.
- [25] L. Donetti, F. Neri, and M. A. Muñoz, “Optimal network topologies: expanders, cages, Ramanujan graphs, entangled networks and all that,” *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2006, no. 08, pp. P08007–P08007, aug 2006.
- [26] R. Olfati-Saber and R. Murray, “Consensus Problems in Networks of Agents With Switching Topology and Time-Delays,” *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1520–1533, sep 2004.

- [27] M. Golubitsky and D. G. Schaeffer, *Singularities and Groups in Bifurcation Theory. Volume I*, ser. Applied Mathematical Sciences. Springer New York, 1985, vol. 51.
- [28] C. Poignard, T. Pereira, and J. P. Pade, “Spectra of Laplacian Matrices of Weighted Graphs: Structural Genericity Properties,” *SIAM Journal on Applied Mathematics*, vol. 78, no. 1, pp. 372–394, 2018.
- [29] R. A. Horn and C. R. Johnson, *Matrix analysis*, 2nd ed. Cambridge University Press, 2013.
- [30] A. Berman and R. J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*, ser. Classic in applied mathematics. Society for Industrial and Applied Mathematics, 1994.
- [31] S. Camiz and S. Stefani, *Matrices and Graphs*. World Scientific, 1996.

Paper C

The role of frustration in collective decision-making dynamical processes on multiagent signed networks

Authors: Angela Fontan and Claudio Altafini

Edited version of the paper:

A. Fontan and C. Altafini, “The role of frustration in collective decision-making dynamical processes on multiagent signed networks,” *arXiv:2105.11396*, pp. 1–18, may 2021.

Early versions of this work were submitted in:

A. Fontan and C. Altafini, “Achieving a decision in antagonistic multi agent networks: frustration determines commitment strength,” in *57th IEEE Conference on Decision and Control*. Miami Beach, FL, USA: IEEE, dec 2018, pp. 109–114.

The role of frustration in collective decision-making dynamical processes on multiagent signed networks

Angela Fontan^{*} and Claudio Altafini^{*}

^{*}Dept. of Electrical Engineering
Linköping University
SE-581 83 Linköping, Sweden
angela.fontan@liu.se, claudio.altafini@liu.se

Abstract

In this work we consider a collective decision-making process in a network of agents described by a nonlinear interconnected dynamical model with sigmoidal nonlinearities and signed interaction graph. The decisions are encoded in the equilibria of the system. The aim is to investigate this multiagent system when the signed graph representing the community is not structurally balanced and in particular as we vary its frustration, i.e., its distance to structural balance. The model exhibits bifurcations, and a “social effort” parameter, added to the model to represent the strength of the interactions between the agents, plays the role of bifurcation parameter in our analysis. We show that, as the social effort increases, the decision-making dynamics exhibits a pitchfork bifurcation behavior where, from a deadlock situation of “no decision” (i.e., the origin is the only globally stable equilibrium point), two possible (alternative) decision states for the community are achieved (corresponding to two nonzero locally stable equilibria). The value of social effort for which the bifurcation is crossed (and a decision is reached) increases with the frustration of the signed network.

1 Introduction

In this paper we want to study a nonlinear model for decision-making in a community of agents where antagonistic interactions may exist between the agents. Indeed, while collaboration between agents is often assumed in order to reach a common decision (for instance in applications such as collective behavior in animal groups [3, 4], cooperative control in robotics [5, 6], or opinion forming [7, 8]), there are applications in which restricting to collaborative interactions means oversimplifying the relationship among the agents [9, 10]. Classes of multiagent systems in which the presence of antagonism is plausible include for instance “social networks”, i.e., groups of individuals interacting and exchanging opinions in a friendly/unfriendly manner or trusting/mistrusting each other. Other scenarios in

which antagonism is unavoidable are team games, where different teams have to compete against each other, or parliamentary democracies, where parties can be allied or rival.

Signed networks [11, 12] are a natural framework to model a community of agents where both cooperative and antagonistic interactions coexist: a positive sign labeling an edge between two agents represents a friendly (or cooperative) relationship, while a negative sign labeling an edge an unfriendly (or competitive) relationship. If the group of agents can be divided into two subgroups such that the agents inside each group are mutual friends (i.e., they are linked by edges with positive weight) while the agents across the two subgroups are enemies (i.e., they are linked by edges with negative weight), we say that the network is structurally balanced [13, 14]. If we assume that a network is undirected and connected, an equivalent condition to structural balance is that the smallest eigenvalue of the normalized signed Laplacian \mathcal{L} is zero, $\lambda_1(\mathcal{L}) = 0$ in the notation we introduce below. As for instance the works [15–18] show, real signed social networks are in general not structurally balanced.

To model the evolution of the opinions of the agents in a community represented as a signed social network we use the model of opinion forming previously introduced in [3, 19, 20]. This model is characterized by sigmoidal and saturated nonlinearities, describing how the agents transmit their opinion to their neighbors. It has a (signed) Laplacian-like structure at the origin and it is endowed with a social effort parameter π which in our analysis plays the role of bifurcation parameter. Our aim is to study how the strength of the commitment among the agents, represented by π , affects the presence and stability of the equilibrium points of the system, which represent the decision states for the community. Under our assumptions, the system is monotone [21] if and only if the corresponding signed social network is structurally balanced. In this case the behavior of the system can be easily deduced from [3, 19, 20], where the authors consider a cooperative system (i.e., only friendly interactions exist between the agents), which is a particular case of monotone system. In this case the analysis shows that for increasing values of the social effort parameter π , the system undergoes two sequential pitchfork bifurcations: after the first bifurcation the number of equilibria jumps from one to three, while after the second bifurcation multiple (more than three) equilibrium points arise. In particular when crossing the first bifurcation the system passes from having the origin as globally asymptotically stable equilibrium to a situation in which two nonzero locally stable equilibria exist while the origin becomes a saddle point. This situation is maintained up to the second bifurcation where novel equilibria, stable or unstable, appear. In the context of social interactions this behavior can be interpreted as follows: if the social effort between the agents is small then no decision is achieved (the origin is the only attractor), while two alternative decision states can be reached if the agents have the “right” amount of commitment. However, by further increasing the social effort, the agents may fall in a situation of overcommitment where multiple (more than 2) decisions are possible. For cooperative networks the first threshold value is fixed and constant, while the second threshold value depends on the algebraic connectivity of the network.

We show in this work that if we consider signed networks that are structurally

unbalanced then, while the qualitative behavior of the system does not change, the value of social effort parameter for which the first bifurcation is crossed is no longer constant but grows with the smallest eigenvalue of the normalized signed Laplacian of the network, which for structurally unbalanced networks is strictly positive ($\lambda_1(\mathcal{L}) > 0$). In particular, its value increases with the amount of “frustration” encoded in the signed network, i.e., with the amount of “disorder” that the negative edges introduce in a network, see [15] for a more thorough statistical physics interpretation. First introduced by Harary [11, 22] and denoted “line index of balance”, the frustration is a standard measure to express the distance of a signed network from a structurally balanced state and is defined as the minimum (weighted) sum of the negative edges that need to be removed in order to obtain a structurally balanced network, see [2] for details.

For our model of decision-making, this means that when we consider signed networks with higher frustration, the first bifurcation is crossed at higher values of the social effort parameter π , meaning that a higher commitment is required from the agents in order to converge to a nontrivial equilibrium point. From a sociological point of view, the result admits a fairly reasonable interpretation: the more in a community there are “unresolved tensions” among the agents (i.e., unbalanced interactions, as measured by the frustration), the more commitment is required by the agents to achieve a common nontrivial decision and to “escape” the (trivial) zero equilibrium point. On the other hand, the value of social effort for which the second bifurcation is crossed is independent of the frustration of the network [2], meaning that for highly frustrated graphs the range of social commitment values for which only two nontrivial equilibria are present shrinks. As a concrete application of these results, in a recent work (see [18]) we have described the process of government formation in parliamentary democracies as a collective decision-making process where the members of the parliament (the agents) are required to cast a vote of confidence (the decision) to a candidate cabinet coalition. In this context, the social effort parameter π is a proxy for the complexity of the government negotiations (measured as duration of the negotiation phase), while a signed network describes the composition of the parliament after each election with signs representing party alliances/rivalries. These “parliamentary networks” are in general not structurally balanced and their frustration correlates well with the duration of the government negotiation processes.

Because of the nonlinearities the behavior of our system is fundamentally different from that of [14]. In the case of [14] in fact, structural balance leads to bipartite consensus and structural unbalance to asymptotic stability. In our case, instead, balanced and unbalanced cases are qualitatively similar, with only the bifurcation point gradually moving to higher values of social commitment as the frustration grows. In this respect, the model we present here has a more reasonable behavior than the one in [14], at the cost of a higher complexity.

Even though the behavior of the system in the structurally unbalanced case is qualitatively similar to the structurally balanced case, the technical tools that must be used to show the results become much more challenging because the system is no longer monotone. An important technical contribution of this paper is in fact to develop methods able to perform a global state space analysis of a broad

class of nonlinear nonmonotone interconnected systems which are not diagonally dominant. Familiar examples of Hopfield-like neural networks fall in this category [23, 24]. Another noteworthy result we obtain is a description of the region in which all equilibria of the system must be contained. In particular also the upper bound to the 1-norm of the equilibria we provide depends on the frustration of the signed network.

The paper investigates also a discrete-time version of our multiagent decision-making system. Such extension is nontrivial in several directions: for instance new phenomena, like period-2 limit cycles, appear in the discrete-time case. Also the techniques that must be used to prove the results are largely different from those of the continuous-time case. In particular, we show that the first bifurcation occurring at the origin is either a pitchfork or a period-doubling bifurcation, depending on the relative positions of the corresponding threshold values for the social effort parameter π , where the value for which a pitchfork bifurcation is crossed is the same as in the continuous-time case. Interestingly, we show that if the signed network has zero or small frustration, the value of π for which a period-doubling bifurcation is crossed is always bigger than the usual bifurcation threshold.

The rest of the paper is organized as follows: in Section 2 we introduce preliminary material. In Sections 3 and 4 we present our results for collective decision-making over signed networks in continuous- and discrete-time, respectively. The results are discussed and interpreted in Section 5. Numerical simulations and examples are shown in Section 6. Technical preliminaries (useful Lemmas and Theorems) and most of the proofs are put in the Appendices at the end of the paper.

2 Preliminaries

2.1 Notation and linear algebra

Given a matrix $A = [a_{ij}] \in \mathbb{R}^{n \times n}$, $A \geq 0$ means element-wise nonnegative, i.e., $a_{ij} \geq 0$ for all $i, j = 1, \dots, n$, while $A > 0$ means element-wise positive, i.e., $a_{ij} > 0$ for all $i, j = 1, \dots, n$. The spectrum of A is denoted $\Lambda(A) = \{\lambda_1(A), \dots, \lambda_n(A)\}$, where $\lambda_i(A)$, $i = 1, \dots, n$, are the eigenvalues of A . A matrix A is called irreducible if there does not exist a permutation matrix P s.t. $P^T AP$ is block triangular. If $x, y \in \mathbb{R}^n$ then $x \geq y$ ($x > y$) means that $x_i \geq y_i$ (resp., $x_i > y_i$) for all $i = 1, \dots, n$. Given two matrices $A, B \in \mathbb{R}^{n \times n}$, the notation $A \sim B$ means that A and B are similar, and hence that they have the same eigenvalues. Given a diagonal positive definite matrix D , we denote the unique (diagonal) positive definite square root of D by $D^{\frac{1}{2}}$. The symbol $\mathbf{1}$ indicates the vector of ones ($\mathbf{1}_m$ is used when the dimension m is not clear from the context) and $0_{n,m}$ the $n \times m$ zero matrix (0 if it is clear from the context).

2.2 Signed graphs

Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a graph with vertex set \mathcal{V} (such that $\text{card}(\mathcal{V}) = n$) and edge set $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$. Let $A = [a_{ij}] \in \mathbb{R}^{n \times n}$ be the adjacency matrix of \mathcal{G} , i.e., $a_{ij} \neq 0$ if

and only if $(j, i) \in \mathcal{E}$. In this work we consider undirected and connected graphs without self-loops.

A graph \mathcal{G} is *signed* if each of its edges is labeled by a sign, that is, $\text{sign}(a_{ij}) = \text{sign}(a_{ji}) = \pm 1$ if $(i, j) \in \mathcal{E}$. The *signed Laplacian* of a graph \mathcal{G} is the symmetric matrix $L = \Delta - A$, where $\Delta = \text{diag}\{\delta_1, \dots, \delta_n\}$ and each diagonal element δ_i is given by $\delta_i = \sum_{j=1}^n |a_{ij}|$, $i = 1, \dots, n$ [14]. The *normalized signed Laplacian* of a graph \mathcal{G} , see [25, 26], is the non-symmetric (symmetrizable, see Appendix A for a definition) matrix defined as

$$\mathcal{L} = \Delta^{-1} L = I - \Delta^{-1} A. \quad (1)$$

Notice that since the graph \mathcal{G} is connected, it does not have isolated vertices, hence $\delta_i \neq 0$ for all i and the matrix Δ^{-1} is well-defined and positive definite.

All the matrices we consider in this work are either symmetric (e.g., A and L) or symmetrizable (e.g., \mathcal{L}), hence they have real eigenvalues, which we assume to be arranged in a nondecreasing order. Let $\lambda_i(A)$, $\lambda_i(L)$ and $\lambda_i(\mathcal{L})$, $i = 1, \dots, n$, be the eigenvalues of A , L and \mathcal{L} , respectively. By construction the eigenvalues of the signed Laplacian L and the normalized signed Laplacian \mathcal{L} are nonnegative, which can be easily shown using the Geršgorin's Theorem, see [27, Thm 6.1.1].

A cycle of a signed graph \mathcal{G} is said *positive* if it contains an even number of negative edges, *negative* otherwise. A graph \mathcal{G} is *structurally balanced* if all its cycles are positive. Equivalent conditions for \mathcal{G} (connected) to be structurally balanced are the following [14]: (i) there exists a partition of the node set $\mathcal{V} = \mathcal{V}_1 \cup \mathcal{V}_2$ such that every edge between \mathcal{V}_1 and \mathcal{V}_2 is negative and every edge within \mathcal{V}_1 or \mathcal{V}_2 is positive; (ii) there exists a signature matrix $S = \text{diag}\{s_1, \dots, s_n\}$ with diagonal entries $s_i = \pm 1$ ($i = 1, \dots, n$), such that $S\mathcal{L}S$ has all nonpositive off-diagonal entries; (iii) $\lambda_1(\mathcal{L}) = 0$. The *frustration index* of a signed graph \mathcal{G} is defined as

$$\epsilon(\mathcal{G}) = \min_{\substack{S = \text{diag}\{s_1, \dots, s_n\} \\ s_i = \pm 1 \forall i}} \frac{1}{2} \sum_{i \neq j} [|\mathcal{L}| + S\mathcal{L}S]_{ij}, \quad (2)$$

where $[\cdot]_{ij}$ indicates the i, j element and $|\cdot|$ the element-wise absolute value, and it provides a measure of the distance of \mathcal{G} from a structurally balanced state [2]. If \mathcal{G} is structurally balanced, $\epsilon(\mathcal{G}) = 0$.

2.3 Monotone systems

Consider the system

$$\dot{x} = f(x), \quad x(0) = x_0 \quad (3)$$

where f is a continuously differentiable function defined on a convex open set $\mathcal{U} \subseteq \mathbb{R}^n$. Let $x(t, \bar{x})$ be the solution $x(t)$ of (3) s.t. $x(0) = \bar{x}$.

Let S be a signature matrix, i.e., $S = \text{diag}\{s_1, \dots, s_n\}$ with $s_i = \pm 1 \forall i$, and let $S\mathbb{R}^n$ indicate an orthant of \mathbb{R}^n , $S\mathbb{R}^n = \{x \in \mathbb{R}^n : s_i x_i \geq 0, i = 1, \dots, n\}$. The partial ordering $\leq_{S\mathbb{R}^n}$ is preserved by the solution operator $x(t, \cdot)$ and the system (3) is type $S\mathbb{R}^n$ monotone if whenever $\bar{x}, \bar{y} \in \mathcal{U}$ with $\bar{x} \leq_{S\mathbb{R}^n} \bar{y}$ then $x(t, \bar{x}) \leq_{S\mathbb{R}^n} x(t, \bar{y})$ for all $t \geq 0$ [21].

Lemma 1 (2.1 in [21]). *If $f \in C^1(\mathcal{U})$ where \mathcal{U} is open and convex in \mathbb{R}^n then $x(t, \cdot)$ preserves the partial ordering $\leq_{S\mathbb{R}^n}$ for $t \geq 0$ if and only if $S \frac{\partial f}{\partial x}(x)S$ has nonnegative off-diagonal elements for every $x \in \mathcal{U}$.*

Therefore, a system (3) is monotone if and only if the graph described by the Jacobian $\frac{\partial f}{\partial x}$ as adjacency matrix is structurally balanced with fixed $S \forall x \in \mathcal{U}$.

3 Decision-making in antagonistic multiagent systems in continuous-time

3.1 Problem formulation

To model the process of decision-making in a community of n agents represented by a signed network \mathcal{G} , we consider the following class of nonlinear interconnected systems,

$$\dot{x} = -\Delta x + \pi A \psi(x), \quad x \in \mathbb{R}^n. \quad (4)$$

The state vector $x = [x_1 \cdots x_n]^T \in \mathbb{R}^n$ represents the agents' opinions, $A = [a_{ij}]$ is the adjacency matrix of the network \mathcal{G} and describes how the agents interact with each other, $\Delta = \text{diag}\{\delta_1, \dots, \delta_n\}$, $\pi > 0$ is a positive scalar parameter and $\psi(x) = [\psi_1(x_1) \cdots \psi_n(x_n)]^T$. Each nonlinear function $\psi_i(x_i)$ describes how an agent i transmits its opinion x_i to its neighbors in the network. This term is then weighted first by the element a_{ij} , describing the influence between agents i and j (positive/friendly if $a_{ij} > 0$ or negative/unfriendly if $a_{ij} < 0$), and then by the parameter π representing the global "social effort" or "strength of commitment" among the agents [3]. The equilibria of the system represent the decision states for the community.

We assume that the signed network \mathcal{G} is undirected (two agents able to influence each other's opinion share the same amount of trust/distrust in each other), connected (there are no isolated agents) and without self-loops, meaning that the signed adjacency matrix A is symmetric, irreducible and with null diagonal. We also assume that a Laplacian-like assumption relates Δ and A , $\delta_i = \sum_j |a_{ij}|$. Finally, we assume that each nonlinear function $\psi_i(x_i) : \mathbb{R} \rightarrow \mathbb{R}$ of the vector $\psi(x)$ satisfies the following conditions

$$\psi_i(x_i) = -\psi_i(-x_i), \forall x_i \in \mathbb{R} \quad (\text{odd}) \quad (A.1)$$

$$\frac{\partial \psi_i}{\partial x_i}(x_i) > 0, \forall x_i \in \mathbb{R} \text{ and } \frac{\partial \psi_i}{\partial x_i}(0) = 1 \quad (\text{monotone}) \quad (A.2)$$

$$\lim_{x_i \rightarrow \pm\infty} \psi_i(x_i) = \pm 1 \quad (\text{saturated}) \quad (A.3)$$

$$\psi_i(x_i) \begin{cases} \text{strictly convex} & \forall x_i < 0 \\ \text{strictly concave} & \forall x_i > 0 \end{cases} \quad (\text{sigmoidal}). \quad (A.4)$$

The system (4) can be rewritten in a "normalized" form,

$$\dot{x} = \Delta(-x + \pi H \psi(x)), \quad x \in \mathbb{R}^n, \quad (5)$$

where we consider the normalized interaction matrix $H := \Delta^{-1}A$. The Jacobian of (5) is $J(x) = -\Delta(I - \pi H \frac{\partial \psi}{\partial x}(x))$ which at the origin for $\pi = 1$ reduces to $J = -\Delta \mathcal{L}$ (where \mathcal{L} is the normalized signed Laplacian of the network); hence, under our assumptions and from Lemma 1, the system (5) is monotone if and only if the signed network \mathcal{G} is structurally balanced.

Our aim is use bifurcation analysis to investigate how the social effort parameter π (our bifurcation parameter) affects the presence of the equilibrium points of the system (5).

3.2 Structurally balanced case

Previous works, such as [19, 20], have studied the behaviour of the system (5) when the adjacency matrix A of the network is nonnegative, i.e., when the system is cooperative [21]. These results, summarized in the following theorem, still hold when the system is in general monotone, that is, when the network \mathcal{G} described by the matrix A is structurally balanced.

Theorem 1 ([20]). Consider the system (5) where each nonlinear function $\psi_i(\cdot)$, $i = 1, \dots, n$, satisfies the properties (A.1)÷(A.4). Assume that the signed graph \mathcal{G} is structurally balanced and let S be the signature matrix s.t. $S\mathcal{L}S$ has all nonpositive off-diagonal entries ($|A| = SAS$).

- (i) When $\pi < 1$, the origin is the unique equilibrium point and it is asymptotically stable.
- (ii) When $\pi = \pi_1 = 1$, the system undergoes a pitchfork bifurcation, the origin becomes unstable and two new equilibria appear, in the orthants described by S and $-S$, respectively, denoted $S\mathbb{R}_+^n$ and $S\mathbb{R}_-^n$. These equilibria are locally asymptotically stable for all values of $\pi > 1$, with domain of attraction at least equal to $S\mathbb{R}_+^n$ and $S\mathbb{R}_-^n$, respectively.
- (iii) If $\lambda_2(\mathcal{L}) < 1$ and simple, when $\pi = \pi_2 = \frac{1}{1-\lambda_2(\mathcal{L})}$, the system undergoes a second pitchfork bifurcation, and new equilibria in other orthants of \mathbb{R}^n appear. When $\pi > \pi_2$, these equilibria may be stable or unstable.

3.3 Structurally unbalanced case

In this section we want to introduce our novel results, i.e., the extension of Theorem 1 to signed networks which are structurally unbalanced: we show that, by redefining the threshold values π_1 and π_2 , the system (5) behaves in a similar manner as the one described in Theorem 1.

Theorem 2 summarizes our findings. We proceed as follows: first, in (i), we prove that the origin is the unique equilibrium point for the system when $\pi < \pi_1$ and it is globally asymptotically stable, where π_1 depends on the smallest eigenvalue of the normalized signed Laplacian \mathcal{L} . Then, in (ii) we show that when $\pi = \pi_1$ the system undergoes a pitchfork bifurcation and two new equilibria appears, which are locally asymptotically stable for all values of the bifurcation parameter in the interval (π_1, π_2) , where π_2 depends on the second smallest eigenvalue of the normalized signed Laplacian \mathcal{L} . Similarly to [3, 20], the proof relies on

bifurcation theory. Lack of monotonicity however implies that most of the proofs require different arguments than those used in [3, 20]. At π_2 the system bifurcates again and new equilibria appear, see (iii).

Theorem 2. Consider the system (5) where each nonlinear function $\psi_i(\cdot)$, $i = 1, \dots, n$, satisfies the properties (A.1)÷(A.4). Assume that the signed graph \mathcal{G} is structurally unbalanced with normalized signed Laplacian \mathcal{L} .

- (i) When $\pi \leq \pi_1 = \frac{1}{1-\lambda_1(\mathcal{L})}$, the origin is the unique equilibrium point of (5) and it is globally asymptotically stable.
- (ii) Let $\lambda_1(\mathcal{L})$ be simple,
 - (ii.1) (existence): when π crosses π_1 , the system undergoes a pitchfork bifurcation and two new equilibria (x^* and $-x^*$) appear;
 - (ii.2) (stability): when $\pi > \pi_1$, the origin is an unstable equilibrium point, while the equilibria $\pm x^* \neq 0$ are locally asymptotically stable for all values of $\pi \in (\pi_1, \pi_2)$, with $\pi_2 = \frac{1}{1-\lambda_2(\mathcal{L})}$;
 - (ii.3) (uniqueness): when $\pi \in (\pi_1, \pi_2)$, the system admits exactly three equilibria, the origin and the two nontrivial equilibrium points $\pm x^* \neq 0$.
- (iii) If $\lambda_2(\mathcal{L})$ is simple, when $\pi = \pi_2$, the system undergoes a second pitchfork bifurcation and new equilibria appear.

Proof in Appendix B.

Remark 1. It follows from the assumption (A.1) that if the system (5) admits an equilibrium point $x^* \neq 0$, then $-x^*$ is also an equilibrium point. _____

Remark 2. Differently from Theorem 1(iii), in Theorem 2(iii) the assumption $\lambda_2(\mathcal{L}) < 1$ is not needed: if the network \mathcal{G} is structurally unbalanced and connected it is always true that $\lambda_2(\mathcal{L}) < 1$, as shown in Lemma 2 below. Therefore, if $\lambda_2(\mathcal{L})$ is simple, $\pi_2 = \frac{1}{1-\lambda_2(\mathcal{L})}$ is always well-defined (i.e., strictly positive and greater than π_1). On the other hand, examples of structurally balanced graphs for which $\lambda_2(\mathcal{L}) > 1$ are complete graphs, whose adjacency matrix is a Euclidean distance matrix. This means that in the structurally balanced case when $\lambda_2(\mathcal{L}) > 1$ the system (5) admits only 3 equilibrium points $(0, \pm x^*)$ for all values of $\pi > \pi_1$ and that the trajectories converge either to x^* or $-x^*$. However, this situation can never happen in the structurally unbalanced case: if π is “large enough” (i.e., it is above the threshold π_2) the system (5) will always admit new equilibria (other than $0, \pm x^*$), which may be attractors. _____

Lemma 2. Let \mathcal{G} be a signed connected network with normalized signed Laplacian \mathcal{L} . If \mathcal{G} is structurally unbalanced, $\lambda_2(\mathcal{L}) < 1$.

Proof in Appendix C.

To conclude this part, we show that for $\pi > \pi_1$ the 1-norm of the equilibria of the system is upper bounded by $\pi(n - 2\epsilon(\mathcal{G}))$, where $\epsilon(\mathcal{G})$ is the frustration of the signed network. Moreover, if the matrix \mathcal{L} is symmetric (i.e., if $\Delta = \delta I$), we can show that the solutions of (5) are all bounded and converge to a set $\Omega_{\epsilon(\mathcal{G})}$, which implies that all the equilibria of the system (5) belong to $\Omega_{\epsilon(\mathcal{G})}$.

Theorem 3. Consider the system (5) where each nonlinear function $\psi_i(\cdot)$, $i = 1, \dots, n$, satisfies the properties (A.1)÷(A.4). Let $\epsilon(\mathcal{G})$ be the frustration of the signed network \mathcal{G} defined in (2), and \mathcal{L} its normalized signed Laplacian.

- (i) If x^* is an equilibrium point of (5), then $\|x^*\|_1 \leq \pi(n - 2\epsilon(\mathcal{G}))$.
- (ii) Let $\pi > \pi_1$. Under the assumption that \mathcal{L} is symmetric (i.e., $\Delta = \delta I$), the trajectories of (5) asymptotically converge to the set $\Omega_{\epsilon(\mathcal{G})}$, where

$$\Omega_{\epsilon(\mathcal{G})} = \{x \in \mathbb{R}^n : \|x\|_1 \leq \pi(n - 2\epsilon(\mathcal{G}))\}.$$

Proof in Appendix D.

Proposition 1. Let \mathcal{G} be a signed graph with normalized signed Laplacian \mathcal{L} , and assume that \mathcal{L} is symmetric (i.e., $\Delta = \delta I$). Then it is possible to derive an upper bound for the social effort at the first bifurcation point, π_1 , which depends on the frustration of the network $\epsilon(\mathcal{G})$:

$$1 \leq \pi_1 \leq \min\left\{\frac{n}{n - 2\epsilon(\mathcal{G})}, \pi_2\right\}. \quad (6)$$

Proof in Appendix E. Notice that if the frustration is zero (i.e., the network is structurally balanced) then $\pi_1 = 1$.

4 Discrete-time

The Euler approximation of system (5) with step ε is

$$x_i(k+1) = (1 - \varepsilon\delta_i)x_i(k) + \varepsilon\pi \sum_{j \neq i} a_{ij}\psi_j(x_j(k)), \quad i = 1, \dots, n. \quad (7)$$

Let $x_k := x(k) = [x_1(k) \cdots x_n(k)]^T$ and $\psi(x_k) = [\psi_1(x_1(k)) \cdots \psi_n(x_n(k))]^T$. Equation (7) can be rewritten in a more compact form as follows:

$$x_{k+1} = (I - \varepsilon\Delta)x_k + \varepsilon\pi A\psi(x_k). \quad (8)$$

The Jacobian at the origin is given by:

$$J_\pi = I - \varepsilon\Delta + \varepsilon\pi A = I - \varepsilon L_\pi \quad (9)$$

where

$$L_\pi = \Delta - \pi A = \Delta(I - \pi(I - \mathcal{L})) \quad (10)$$

and \mathcal{L} is the normalized signed Laplacian of the network. As in Section 3, we want to study how the social effort parameter π affects the existence of the equilibria of the system (8), again relying on tools from bifurcation analysis [28].

A local bifurcation occurs at the origin if the Jacobian J_π has an eigenvalue with absolute value equal to 1 (that is, equal to ± 1 since J_π is symmetric and has real eigenvalues). When π is small and in particular is such that all the eigenvalues of J_π have magnitude less than one, following the proof of [29, Thm 2] it is possible

to prove (under the additional condition $\varepsilon \max_i \delta_i < 1$) that the origin is globally asymptotically stable and hence the unique equilibrium point of the system (8). As π grows, the magnitude of the eigenvalues of J_π increases and for values of π such that J_π has a simple eigenvalue λ at ± 1 the system (8) can undergo either a pitchfork ($\lambda = +1$) or a period-doubling ($\lambda = -1$) bifurcation [28].

Let $\lambda_i(J_\pi)$ and $\lambda_i(L_\pi)$, $i = 1, \dots, n$, be the eigenvalues of J_π and L_π , respectively, which we assume to be arranged in a nondecreasing order. We denote π_1 the value of social effort for which the biggest eigenvalue of J_π crosses $+1$ and $\pi_{1,d}$ the value of social effort for which the smallest eigenvalue of J_π crosses -1 :

$$\pi_1 : \lambda_n(J_{\pi_1}) = 1, \quad \pi_{1,d} : \lambda_1(J_{\pi_{1,d}}) = -1. \quad (11)$$

Remark 3. From (9), the biggest and smallest eigenvalues of J_π are given by

$$\lambda_n(J_\pi) = 1 - \varepsilon \lambda_1(L_\pi), \quad \lambda_1(J_\pi) = 1 - \varepsilon \lambda_n(L_\pi),$$

which means that J_π is Schur stable (i.e., its eigenvalues have magnitude strictly less than one) if and only if the following two conditions hold:

- (i) $\lambda_1(L_\pi) > 0$, that is, L_π is positive definite.

From (10) and since Δ is positive definite, this condition is equivalent to $I - \pi(I - \mathcal{L})$ having (strictly) positive eigenvalues, i.e., $0 < 1 - \pi(1 - \lambda_1(\mathcal{L}))$;

- (ii) $\varepsilon \lambda_n(L_\pi) < 2$, that is, $\varepsilon L_\pi - 2I$ is negative definite.

Hence π_1 is the value of social effort for which the smallest eigenvalue of L_π crosses 0, while $\pi_{1,d}$ is the value of social effort for which the biggest eigenvalue of εL_π crosses 2:

$$\pi_1 : \lambda_1(L_{\pi_1}) = 0 \Rightarrow \pi_1 = \frac{1}{1 - \lambda_1(\mathcal{L})} \quad (12)$$

$$\pi_{1,d} : \lambda_n(L_{\pi_{1,d}}) = \frac{2}{\varepsilon}. \quad (13)$$

In the analysis of the discrete-time model (8) (see Theorem 4 below) it is relevant to know where $\pi_{1,d}$ lies compared with π_1 . The next proposition shows that $\pi_1 < \pi_{1,d}$ always holds if the network is structurally balanced ($\lambda_1(\mathcal{L}) = 0$) or if it is structurally unbalanced but $\lambda_1(\mathcal{L}) > 0$ is small.

Proposition 2. *Assume that $\varepsilon \max_i \delta_i < 1$. If any of the two following conditions on the signed graph \mathcal{G} with normalized signed Laplacian \mathcal{L} is satisfied:*

- (i) \mathcal{G} is structurally balanced (i.e., $\lambda_1(\mathcal{L}) = 0$), or
- (ii) \mathcal{G} is structurally unbalanced and $\lambda_1(\mathcal{L}) < 2 - \lambda_n(\mathcal{L})$,

then $\pi_1 < \pi_{1,d}$.

Proof in Appendix F.

The next two lemmas show that if the system (8) admits a nontrivial equilibrium point then $\pi > \pi_1$ (Lemma 3), while if it admits a period-2 orbit then $\pi > \pi_{1,d}$ (Lemma 4).

Lemma 3. *Consider the system (8) where each nonlinear function $\psi_i(\cdot)$, $i = 1, \dots, n$, satisfies the properties (A.1)÷(A.4). If $x^* \neq 0$ is an equilibrium point of the system (8) then $\pi > \pi_1$.*

Proof in Appendix G.

Lemma 4. Consider the system (8) where each nonlinear function $\psi_i(\cdot)$, $i = 1, \dots, n$, satisfies the properties (A.1)÷(A.4). If $\varepsilon < \frac{2}{\max_i \delta_i}$ and the system (8) admits a period-2 limit cycle ($\exists K > 0$ such that $x_{k+2} = x_k$ for all $k \geq K$) then $\pi > \pi_{1,d}$.

Proof in Appendix G.

[Remark 4.] The condition $\varepsilon < \frac{2}{\max_i \delta_i}$ imposed by Lemma 4 represents an upper bound on the step size ε in the Euler approximation. ———

We are now ready to state our results for the discrete-time system (8), summarized in Theorem 4. Similarly to Theorem 2, we show first that the origin is the unique equilibrium point for the system when $\pi < \min\{\pi_1, \pi_{1,d}\}$ and that it is globally asymptotically stable, see Theorem 4(i). However, differently from the continuous-time case, when π crosses $\min\{\pi_1, \pi_{1,d}\}$ two different behaviors can happen. If $\pi_1 < \pi_{1,d}$ we expect the system (8) to undergo a pitchfork bifurcation when $\pi = \pi_1$ while if $\pi_1 > \pi_{1,d}$ we expect a period-doubling bifurcation when $\pi = \pi_{1,d}$, see Theorem 4(ii). The special case where $\pi_1 = \pi_{1,d}$ is here not discussed, but the intuition is that a Neimark-Sacker bifurcation occurs at the origin when $\pi = \pi_{1,d} = \pi_1$ [28]. Observe also that $\pi > \pi_1$ and $\pi > \pi_{1,d}$ are necessary conditions (not only sufficient) in order for the system (8) to admit a nontrivial equilibrium or a periodic solution, respectively, as shown in Lemma 3 and Lemma 4.

Finally, notice that the following theorem holds also for structurally balanced networks. However, in that case the condition $\pi_1 < \pi_{1,d}$ would always be satisfied (see Proposition 2) meaning that the formulation of theorem could be simplified.

Theorem 4. Consider the system (8) where each nonlinear function $\psi_i(\cdot)$, $i = 1, \dots, n$, satisfies the properties (A.1)÷(A.4). Assume that the signed graph \mathcal{G} is structurally unbalanced with normalized signed Laplacian \mathcal{L} . Let J_π , L_π , π_1 and $\pi_{1,d}$ be as in (9), (10), (12) and (13), respectively. Assume that $1 - \varepsilon \max_i \delta_i \geq 0$.

- (i) If $\pi < \min\{\pi_1, \pi_{1,d}\}$ then the origin is the unique equilibrium point of the system (8) and it is globally asymptotically stable.
- (ii) If $\pi_1 < \pi_{1,d}$ and the biggest eigenvalue of J_{π_1} , $\lambda_n(J_{\pi_1}) = +1$, is simple, when $\pi = \pi_1$ the system (8) undergoes a pitchfork bifurcation;
If $\pi_1 > \pi_{1,d}$ and the smallest eigenvalue of $J_{\pi_{1,d}}$, $\lambda_1(J_{\pi_{1,d}}) = -1$, is simple, when $\pi = \pi_{1,d}$ the system (8) undergoes a period-doubling bifurcation.

Proof in Appendix H.

Notice that, compared with Theorems 1 and 2, Theorem 4 considers only the first bifurcation the system (8) undergoes at the origin, i.e., it does not consider for instance secondary bifurcations at the origin happening for values of π such that $\lambda_{n-1}(J_\pi) = +1$ or $\lambda_2(J_\pi) = -1$.

Corollary 1. Consider the system (8) where each nonlinear function $\psi_i(\cdot)$, $i = 1, \dots, n$, satisfies the properties (A.1)÷(A.4). Let $\epsilon(\mathcal{G})$ be the frustration of the signed network \mathcal{G} defined in (2).

- (i) If x^* is an equilibrium point of (8), then $\|x^*\|_1 \leq \pi(n - 2\epsilon(\mathcal{G}))$.
- (ii) If $\Delta = \delta I$ with $\delta\varepsilon < 1$, the trajectories of (8) asymptotically converge to the set $\{x \in \mathbb{R}^n : \|x\|_1 \leq \pi n\}$.

The proof is omitted since (i) follows from the observation that the discrete- and continuous-time models share the same equilibrium points, therefore the upper bound on the 1-norm of the equilibria found in Theorem 3(i) still holds, and (ii) follows from the fact that the nonlinearities are saturated (and can be shown for instance using the Lyapunov function $V(x_k) = \|x_k\|_1 - \pi n$ for all $\|x_k\|_1 > \pi n$ and $V(x_k) = 0$ otherwise).

5 Discussion and interpretation of the results

Comparing Theorem 1 with Theorem 2, the general behavior of the continuous-time system (5) (illustrated in Figure 1) does not change when, instead of a structurally balanced network, we assume that the network is structurally unbalanced. However, while in the structurally balanced case (see Figure 1a) the first threshold value for the social effort parameter π is constant ($\pi_1 = 1$), in the structurally unbalanced case (see Figure 1b) this value is strictly greater than 1 and increases with the smallest eigenvalue of the normalized signed Laplacian, $\lambda_1(\mathcal{L})$. In a recent work [2] we have shown that $\lambda_1(\mathcal{L})$ approximates well the frustration of a signed network \mathcal{G} (measured by $\epsilon(\mathcal{G})$ introduced in equation (2)), while the intuition is that $\lambda_2(\mathcal{L})$ is independent from the frustration $\epsilon(\mathcal{G})$. As a consequence, a higher frustration $\epsilon(\mathcal{G})$ (reflecting a situation where the system (5) is “far” from being monotone) implies (i) a higher value of π_1 and (ii) the shrinkage of the interval (π_1, π_2) for which only two alternative equilibria are admitted. These conclusions are illustrated in Example 2.

In the context of social networks the decision-making process (4) can be summarized as follows:

- $\pi < \pi_1$: No decision will be reached if the social effort among the agents is small.
- $\pi \in (\pi_1, \pi_2)$: The “right” level of commitment among the agents leads to two possible (alternative) decisions. If the signed social network is not structurally balanced, a higher frustration implies that a higher effort will be required from the agents in order to achieve this decision.
- $\pi > \pi_2$: An overcommitment between the agents (high value of social effort) leads to a situation where several alternative decisions are possible.

When we instead compare Theorems 1 and 2 with Theorem 4, we observe that the discrete-time system (8) exhibits a “richer” behavior, in that it admits (stable) periodic solutions, as illustrated in Figure 2 (which, for the sake of simplicity, does not consider secondary pitchfork or period-doubling bifurcations at the origin). This is related to the presence of a new threshold value for the parameter π , denoted $\pi_{1,d}$: understanding where $\pi_{1,d}$ lies compared with π_1 plays a key role when investigating the behavior of the system (8) over a signed network. In particular, Proposition 2 suggests that the condition $\pi_1 > \pi_{1,d}$ cannot hold unless a signed network is structurally unbalanced and has high frustration (i.e., $\lambda_1(\mathcal{L}) \gg 0$).

This implies first that, if we consider networks that are structurally balanced (for which $\pi_1 < \pi_{1,d}$ always holds) or that are structurally unbalanced for which

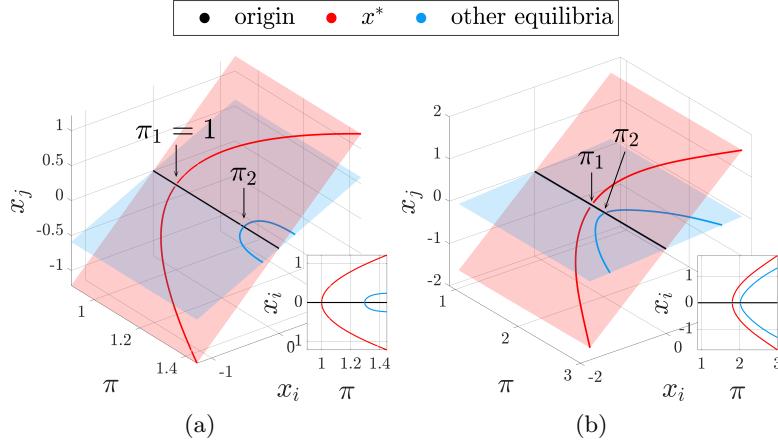


Figure 1: The system (5) undergoes two pitchfork bifurcations, respectively for $\pi = \pi_1$ and $\pi = \pi_2$. The bifurcation diagram for two components x_i and x_j is here shown for two different signed networks. (a): Structurally balanced network (monotone system). (b): Structurally unbalanced network.

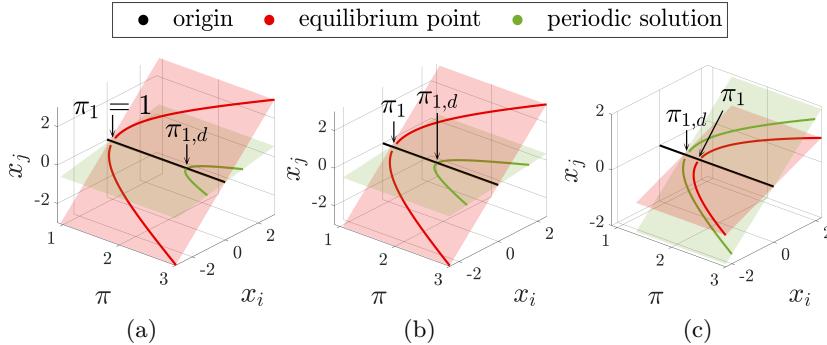


Figure 2: The system (8) undergoes a pitchfork bifurcations and a period-doubling bifurcation, respectively for $\pi = \pi_1$ and $\pi = \pi_{1,d}$. The bifurcation diagram for two components x_i and x_j is here shown for three different signed networks. (a): Structurally balanced network. (b): Structurally unbalanced network with $\pi_1 < \pi_{1,d}$. (c): Structurally unbalanced network with $\pi_1 > \pi_{1,d}$.

$\pi_1 < \pi_{1,d}$ (typically, with low frustration) the general behavior of the discrete-time system (8) resembles that of its continuous-time counterpart, see Fig. 2a and Fig. 2b: the crossing of a (pitchfork) bifurcation yields two (alternative) nontrivial equilibrium points representing two possible (alternative) decisions. Hence, the general idea that the higher is the frustration of the network the higher is the social effort needed to converge to a nontrivial equilibrium point still holds. Instead, if we consider networks that are structurally unbalanced for which $\pi_1 > \pi_{1,d}$ (typically, with high frustration), see Fig. 2c, then there exists an interval of values for the social effort parameter, $(\pi_{1,d}, \pi_1)$, for which the collective decision-making process

still ends in a deadlock situation where the opinions of the agents do not settle but keep fluctuating: only by further increasing the commitment among the agents the process can be settled and the community can reach a decision. In conclusion, in the discrete-time model the presence of high frustration in the graph leads to agents who will never cease to change opinion (a somewhat artificial behavior). Other recent works in the literature that propose models characterized by fluctuations of opinions of the agents are for instance [30, 31].

6 Numerical Examples

In this section we first illustrate the bound (6) in Proposition 1 (Example 1). As a byproduct we observe (numerically) that the bound is tight if the smallest and largest eigenvalues of \mathcal{L} satisfy the condition $\lambda_1(\mathcal{L}) < 2 - \lambda_n(\mathcal{L})$ (typically, if the network does not have high frustration). Then, we show the behavior of the system (5) over signed (structurally unbalanced) networks with increasing frustration (Example 2 and Example 3). In Example 4 we show that when the social effort parameter π crosses the second threshold π_2 the system admits multiple equilibria which are stable (i.e., several decision states for the community are possible). Example 5 is used to illustrate a case which has not been treated by our analysis. Indeed, in Example 5 we illustrate the behavior of the system (5) in presence of symmetries implying an algebraic multiplicity of $\lambda_1(\mathcal{L})$ higher than 1. The case where the smallest eigenvalue of \mathcal{L} is not simple is in fact not covered by Theorem 2. However the intuition, supported by the reading of [32, 33], is that when $\pi > \pi_1$ the system admits multiple (more than three) equilibria. Finally, in Example 6 we illustrate the behavior of the discrete-time system (8) and compare it with that of the continuous-time system (5).

If not specified otherwise, we assume that each nonlinear function $\psi_i(\cdot)$ ($i = 1, \dots, n$) is given by the hyperbolic tangent $\psi(\varepsilon) = \tanh(\varepsilon)$. Moreover, to compute numerically the frustration $\epsilon(\mathcal{G})$ of a signed network \mathcal{G} we use the algorithm proposed in [34].

Example 1

This example wants to illustrate the bound (6) in Proposition 1 and show that it holds also for graphs whose normalized signed Laplacian is not symmetric. In Fig. 3 we consider two sequences of signed networks \mathcal{G} with $n=500$ agents (in which the edge weights are drawn from a uniform distribution and $p=0.8$ is the edge probability) and with increasing frustration $\epsilon(\mathcal{G})$. In the first sequence (see Fig. 3a), each adjacency matrix A of the network is rescaled so that $|A|\mathbf{1} = \delta\mathbf{1}$, which implies that the normalized signed Laplacian \mathcal{L} is symmetric. Instead, in the second sequence (see Fig. 3b), each matrix \mathcal{L} is not symmetric (but is symmetrizable). As Fig. 3 illustrates, the bound (6) holds for both sequences; moreover, when the frustration is small (numerically, when the condition $\lambda_1(\mathcal{L}) < 2 - \lambda_n(\mathcal{L})$ is satisfied) the upper bound $\frac{n}{n-2\epsilon(\mathcal{G})}$ for π_1 is tight (this is not surprising since we know that $\lambda_1(\mathcal{L})$ approximates well the frustration).

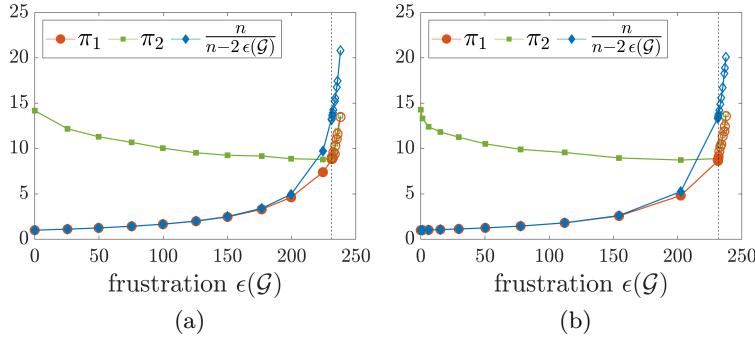


Figure 3: Example 1. Plot of π_1, π_2 and $\frac{n}{n-2\epsilon(\mathcal{G})}$, for two sequences of signed networks with increasing frustration $\epsilon(\mathcal{G})$. (a): Sequence 1: for each network, the normalized signed Laplacian \mathcal{L} is symmetric. (b): Sequence 2: for each network, the matrix \mathcal{L} is not symmetric. A full (resp., empty) symbol means that $\lambda_1(\mathcal{L}) < 2 - \lambda_n(\mathcal{L})$ (resp., $\lambda_1(\mathcal{L}) > 2 - \lambda_n(\mathcal{L})$); for clarity, a dashed line shows the maximum value of frustration above which the condition $\lambda_1(\mathcal{L}) < 2 - \lambda_n(\mathcal{L})$ does not hold.

Example 2

We consider three signed networks with $n = 20$ agents in which the edge weights are drawn from a uniform distribution and $p = 0.5$ is the edge probability. These networks are chosen to be structurally unbalanced and with increasing frustration. In Fig. 4 the euclidean norm of the equilibria of the system (5) for values of π in $\{0.005, 0.010, \dots, 4\}$ is depicted. As Table 1 shows, the smallest eigenvalue of the normalized signed Laplacian increases with the frustration of the network while the second smallest eigenvalue remains almost constant, hence the interval for π for which the system admits only two equilibria becomes smaller (compare Fig. 4a and Fig. 4c).

Example 3

Consider a network \mathcal{G} with $n = 100$ agents in which the edge weights are drawn from a uniform distribution and $p = 0.8$ is the edge probability. Let $A = [a_{ij}]$ be its weighted nonnegative adjacency matrix. Consider now a sequence of signed networks \mathcal{G}_β with weighted adjacency matrices $A_\beta = [a_{\beta ij}]$ constructed such that $|A_\beta| = A$ and their signature is dependent on a parameter $\beta \in \{0, 0.05, 0.1, \dots, 1\}$: if $a_{ij} \neq 0$ then $a_{\beta ij} \neq 0$ and $P[a_{\beta ij} < 0] = \beta$. When $\beta = 1$, $A_1 = -|A|$. As β increases also the frustration of the networks increases. For each network, we numerically compute the equilibria x^* of the system (5) for values of π in $\{1, 1.05, \dots, 9\}$ and their 1-norm $\|x^*\|_1$: let $\mathcal{X} = \{x^* \in \mathbb{R}^n : x^* \text{ is an equilibrium point of the system (5)}\}$ be the set of equilibria. In Fig. 5, for each network of the sequence we plot $\frac{1}{\pi} \max_{x^* \in \mathcal{X}} \|x^*\|_1$ (the maximum 1-norm of the equilibrium points divided by π) for each value of π ; the colormap illustrates the sequence of signed networks \mathcal{G}_β with increasing frustration. As Theorem 3 states, the maximum 1-norm of the equilibria

is upper bounded by $\pi(n - 2\epsilon(\mathcal{G}_\beta))$, where $\epsilon(\mathcal{G}_\beta)$ indicates the frustration of \mathcal{G}_β : as the frustration increases, the bound decreases.

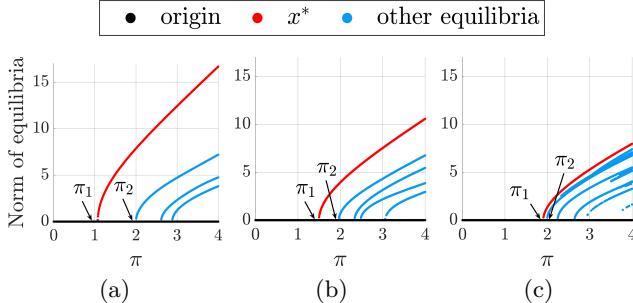


Figure 4: Example 2. Norm of the equilibrium points of the system (5) as a function of π . The networks \mathcal{G} we use in this example are structurally unbalanced, with increasing frustration $\epsilon(\mathcal{G})$, see Table 1. (a): $\epsilon(\mathcal{G}) = 0.677$. (b): $\epsilon(\mathcal{G}) = 4.285$. (c): $\epsilon(\mathcal{G}) = 5.536$.

	$\epsilon(\mathcal{G})$	λ_1	λ_2	π_1	π_2
(a)	0.677	0.065	0.500	1.069	2.000
(b)	4.285	0.332	0.491	1.496	1.966
(c)	5.536	0.475	0.499	1.905	1.995

Table 1: Example 2. Values of frustration, first two eigenvalues of the normalized signed Laplacian and bifurcation points for the three cases, (a), (b) and (c), depicted in Figure 4.

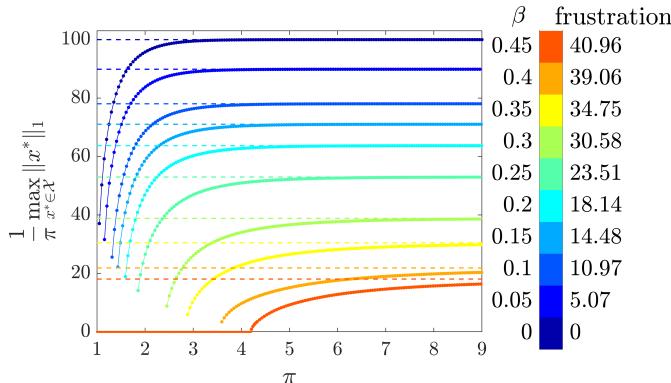


Figure 5: Example 3. Plot of the maximum 1-norm of x^* , where x^* is an equilibrium point of the system (5), for a sequence of signed networks with increasing frustration. The values of $n - 2\epsilon(\mathcal{G}_\beta)$ are shown as dashed lines.

Example 4

When $\pi > \pi_2$, Theorem 2(iii) proves that the system (5) admits multiple equilibrium points. Through numerical simulations it is possible to see that some of these equilibria may be stable. In Fig. 6 the multistability for the system (5) is highlighted as we depict the evolution of the first component of $x(t)$ for 50 random initial conditions and $\pi = 4 > \pi_2 = 1.995$. The same signed network as case (c) in Fig. 4 and Table 1 is used.

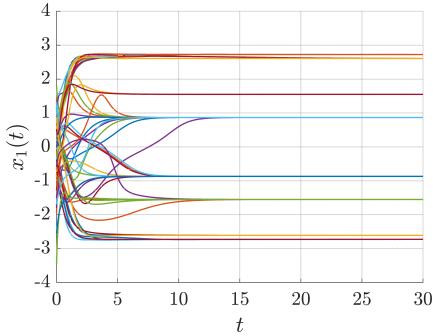


Figure 6: Example 4. Evolution of state variable $x_1(t)$ for 50 random initial conditions and $\pi = 4$. The signed network considered in this example corresponds to the one used to obtain Fig. 4c ($\pi_2 = 1.995$).

Example 5

Consider the system (5) where each nonlinear function $\psi_i(\cdot)$, $i = 1, \dots, n$, satisfies the properties (A.1)÷(A.4). Moreover, assume that

$$\psi_i(\varepsilon) = \psi_j(\varepsilon) =: \psi_u(\varepsilon), \forall i, j = 1, \dots, n, \varepsilon \in \mathbb{R} \quad (\text{identical nonlinearities}). \quad (\text{A.5})$$

Notice that under these assumptions $P\psi(x) = \psi(Px)$ for all signed permutation matrices P .

Let $n = 3$ and the adjacency matrix of the network be

$$A = \begin{bmatrix} 0 & -1 & -1 \\ -1 & 0 & -1 \\ -1 & -1 & 0 \end{bmatrix} = I - \mathbf{1}\mathbf{1}^T,$$

which implies that the signed graph described by A is structurally unbalanced and that the smallest eigenvalue of the normalized signed Laplacian \mathcal{L} , $\lambda_1(\mathcal{L})$, is not simple (the spectrum of \mathcal{L} is $\Lambda(\mathcal{L}) = \{\frac{1}{2}, \frac{1}{2}, 2\}$). This represents an interesting case for our analysis since the (algebraic and geometric) multiplicity of the smallest eigenvalue of \mathcal{L} is 2, hence we cannot straightforwardly apply Theorem 2(ii). However, in this case the equilibria of the system (5) for $\pi > \pi_1 = \frac{1}{1-\lambda_1(\mathcal{L})} = 2$ can be computed explicitly.

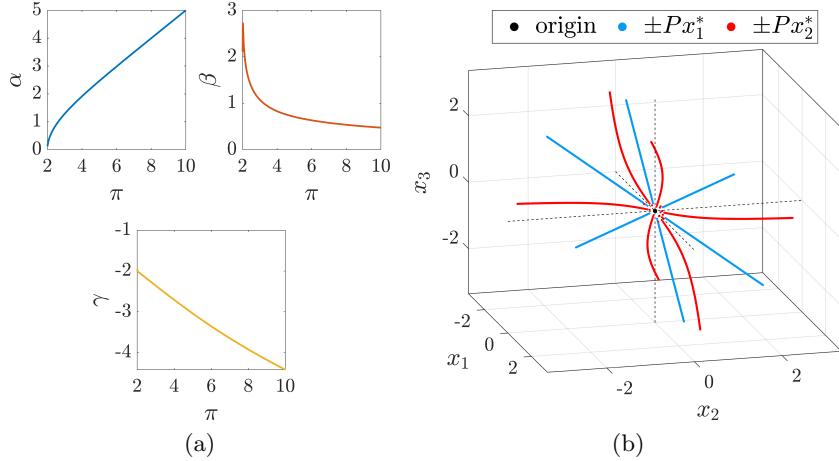


Figure 7: Example 5. (a): α, β, γ as functions of π . (b): Equilibria of the system (5) as described by (14), for $\pi = 2.001, 2.002, \dots, 4$: the origin (black dot), $\pm Px_1^*$ (blue branches) and $\pm Px_2^*$ (red branches).

Under assumption (A.5), let $\pi > 2$ and $\alpha(\pi), \beta(\pi) > 0, \gamma(\pi) < 0$ be such that

$$\alpha : \frac{\psi_u(\alpha)}{\alpha} = \frac{2}{\pi}, \quad \beta, \gamma : \begin{cases} \gamma = -\pi\psi_u(\beta) \\ \psi_u(\beta) + \frac{2}{\pi}\beta + \psi_u(\gamma) = 0, \end{cases}$$

see also Fig. 7a. Then $x_1^* = \alpha[1, -1, 0]^T$, $x_2^* = [\beta, \beta, \gamma]^T$ are equilibria of (5). Indeed

$$\begin{aligned} \pi H\psi(x_1^*) &= \pi\psi_u(\alpha)H\frac{x_1^*}{\alpha} = \frac{\pi}{2}\frac{\psi_u(\alpha)}{\alpha}x_1^* = x_1^*, \\ \pi H\psi(x_2^*) &= -\frac{\pi}{2} \begin{bmatrix} \psi_u(\beta) + \psi_u(\gamma) \\ \psi_u(\beta) + \psi_u(\gamma) \\ 2\psi_u(\beta) \end{bmatrix} = \begin{bmatrix} \beta \\ \beta \\ \gamma \end{bmatrix} = x_2^*. \end{aligned}$$

Let $\Phi(x, \pi) = -x + \pi H\psi(x)$. Under assumption (A.1), $\Phi(x, \pi)$ is odd. Moreover, since $PHP^T = H$ for all permutation matrices $P \in \mathbb{R}^{3 \times 3}$, it holds that

$$P\Phi(x, \pi) = \Phi(Px, \pi) \quad \forall P \in \mathbf{S}_3,$$

that is, $\Phi(x, \pi)$ is \mathbf{S}_3 -equivariant (\mathbf{S}_3 indicates the symmetric group of order 3, i.e., the group of all permutations of a three-element set). Hence if $x(t)$ is a solution of (5), then $\pm Px(t)$, $P \in \mathbf{S}_3$, is also a solution of (5) [32].

To conclude, the equilibria of (5) can be written as

$$\pm Px_1^*, \quad \pm Px_2^* \quad \forall P \in \mathbf{S}_3. \quad (14)$$

Figure 7b shows the equilibrium points of the system (5), where the nonlinear function ψ_u is the hyperbolic tangent $\psi_u(\varepsilon) = \tanh(\varepsilon)$, as π increases.

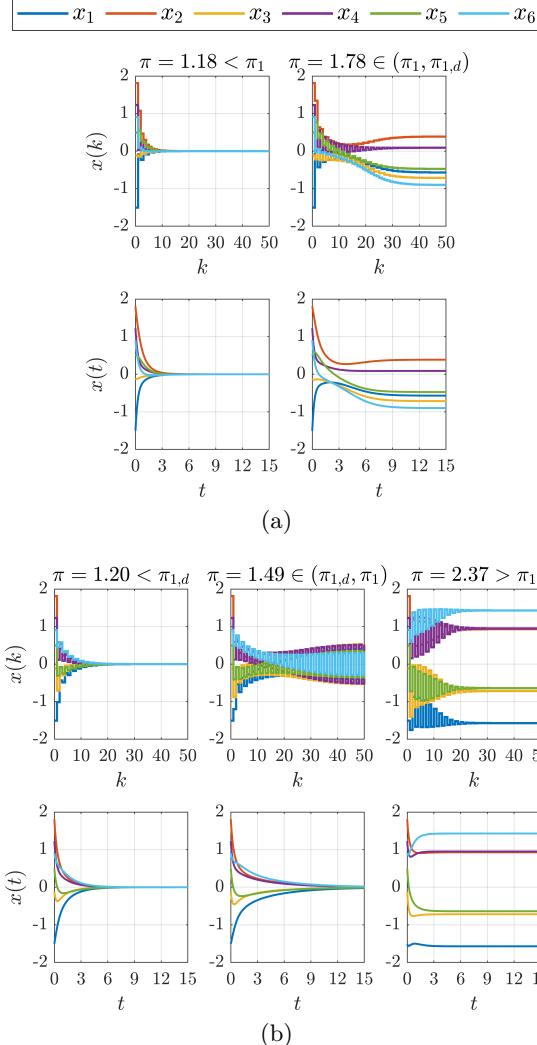


Figure 8: Example 6. Trajectories of the discrete-time system (8) with $\varepsilon = 0.3$ (top panels) vs trajectories of the continuous-time system (5) (bottom panels) for different values of π . (a): Network \mathcal{G}_1 , $\pi_1 < \pi_{1,d}$. (b): Network \mathcal{G}_2 , $\pi_1 > \pi_{1,d}$.

Example 6

This last example wants to illustrate the results of Theorem 4 for the discrete-time system (8) and compare them with the results of Theorem 2 for the continuous-time system (5). We consider two structurally unbalanced networks (\mathcal{G}_1 and \mathcal{G}_2) with $n = 6$ agents in which the edge weights are drawn from a uniform distribution and $p = 0.9$ is the edge probability. The network \mathcal{G}_1 is such that $1 < \pi_1 = 1.53 < 1.89 = \pi_{1,d}$ while the network \mathcal{G}_2 is such that $1 < \pi_{1,d} = 1.40 < 1.63 = \pi_1$, where π_1 and $\pi_{1,d}$ are defined in (12) and (13), respectively.

Figure 8 plots the trajectories of the discrete-time system (8) with $\varepsilon = 0.3$ (top panels) and the trajectories of the continuous-time system (5) (bottom panels) for different values of π and the same initial condition $x(0) = [-1.51, 1.81, -0.12, 1.23, 0.49, 0.91]^T$: in Fig. 8a we consider the network \mathcal{G}_1 , while in Fig. 8b the network \mathcal{G}_2 . When $\pi_1 < \pi_{1,d}$ (see Fig. 8a), we expect the trajectories of both the discrete- and continuous-time system to converge to the origin for all values of π less than π_1 (see left panels) and to converge to a nontrivial equilibrium point for values of π greater than (and in a neighborhood of) π_1 (see right panels). When $\pi_1 > \pi_{1,d}$ (see Fig. 8b), we expect the trajectories of both the discrete- and continuous-time system to converge to the origin for all values of π less than $\pi_{1,d}$ (see left panels). However, when $\pi \in (\pi_{1,d}, \pi_1)$ (see middle panels), while the trajectories of continuous-time system still converge to the origin, the discrete-time system admits a periodic solution. Finally, for both the discrete- and continuous-time system to admit a nontrivial equilibrium point π needs to be greater than π_1 (see right panels).

7 Conclusions

In this work we have extended the analysis of a decision-making process in a community of agents, described by the nonlinear interconnected model introduced in [3, 20], to the case in which the signed network representing the group of agents is not structurally balanced. We provided necessary and sufficient conditions for the existence (and stability) of equilibrium points of the system showing that, qualitatively, the bifurcation behavior of the system does not change when we assume that it is not monotone, i.e., that the signed social network is not structurally balanced. What changes, however, is the threshold at which the bifurcation occurs. In particular, we have shown in the paper that this bifurcation threshold grows with the frustration of the signed network.

Given the interpretation of the bifurcation parameter as “social effort” of the network of agents, from a sociological point of view, this behavior is reasonable and plausible: the more “disorder” (i.e., frustration) a social network contains, the more difficult it is for its actors to achieve a common decision.

Appendix

A Technical preliminaries

In this section we introduce definitions and technical theorems and lemmas from linear algebra that will be necessary in order to prove the main results of this work.

Definition 1 ([35, 36]). A matrix $A \in \mathbb{R}^{n \times n}$ is (diagonally) symmetrizable if DA is symmetric for some diagonal matrix D with positive diagonal entries. The matrices DA and D are called symmetrization and symmetrizer of A , respectively.

Theorem 5 (Ostrowski, 4.5.9 in [27]). Let $A, S \in \mathbb{R}^{n \times n}$ with A symmetric and S nonsingular. Let the eigenvalues of A , SAS^T and SS^T be arranged in nondecreasing order. For each $k = 1, \dots, n$, there exists a positive real number θ_k such that $\lambda_1(SS^T) \leq \theta_k \leq \lambda_n(SS^T)$ and $\lambda_k(SAS^T) = \theta_k \lambda_k(A)$.

The following lemma results from Theorem 5.

Lemma 5. Let $B \in \mathbb{R}^{n \times n}$ be symmetrizable, and $S = \text{diag}\{s_1, \dots, s_n\} \in \mathbb{R}^{n \times n}$ be a positive definite diagonal matrix. Let the eigenvalues of B , BS , SB and $S^{\frac{1}{2}}BS^{\frac{1}{2}}$ be arranged in nondecreasing order. Then, for all $k \in \{1, \dots, n\}$, it holds that $\exists \theta_k \in [\min_i\{s_i\}, \max_i\{s_i\}]$ such that

$$\lambda_k(BS) = \lambda_k(SB) = \lambda_k(S^{\frac{1}{2}}BS^{\frac{1}{2}}) = \theta_k \lambda_k(B).$$

Proof. The matrix B is symmetrizable, hence there exist a diagonal positive definite matrix $D \in \mathbb{R}^{n \times n}$ and a symmetric matrix $A \in \mathbb{R}^{n \times n}$ such that $B = DA$. Define the symmetric matrix

$$B_{\text{sym}} := D^{\frac{1}{2}}AD^{\frac{1}{2}} = D^{-\frac{1}{2}}BD^{\frac{1}{2}} \sim B$$

which, from similarity, has the same eigenvalues of B . Notice that, with S diagonal positive definite, the products BS , SB and $S^{\frac{1}{2}}BS^{\frac{1}{2}}$ are similar matrices,

$$BS = S^{-1}(SB)S \sim SB \quad \text{and} \quad BS = S^{-\frac{1}{2}}(S^{\frac{1}{2}}BS^{\frac{1}{2}})S^{\frac{1}{2}} \sim S^{\frac{1}{2}}BS^{\frac{1}{2}},$$

which implies that they have the same eigenvalues. Moreover,

$$S^{\frac{1}{2}}BS^{\frac{1}{2}} = S^{\frac{1}{2}}D^{\frac{1}{2}}B_{\text{sym}}D^{-\frac{1}{2}}S^{\frac{1}{2}} = D^{\frac{1}{2}}(S^{\frac{1}{2}}B_{\text{sym}}S^{\frac{1}{2}})D^{-\frac{1}{2}} \sim S^{\frac{1}{2}}B_{\text{sym}}S^{\frac{1}{2}},$$

which implies that $S^{\frac{1}{2}}BS^{\frac{1}{2}}$ and $S^{\frac{1}{2}}B_{\text{sym}}S^{\frac{1}{2}}$ have the same eigenvalues. Since $S^{\frac{1}{2}}B_{\text{sym}}S^{\frac{1}{2}}$ is a symmetric matrix, with $S^{\frac{1}{2}}$ diagonal and nonsingular, it is possible to apply Theorem 5. For all $k \in \{1, \dots, n\}$ there is a positive real number $\theta_k \in [\min_i\{s_i\}, \max_i\{s_i\}]$ such that

$$\lambda_k(S^{\frac{1}{2}}B_{\text{sym}}S^{\frac{1}{2}}) = \theta_k \lambda_k(B_{\text{sym}}).$$

Therefore, from similarity, it follows that for all $k \in \{1, \dots, n\}$ there exists a $\theta_k \in [\min_i\{s_i\}, \max_i\{s_i\}]$ such that

$$\lambda_k(S^{\frac{1}{2}}BS^{\frac{1}{2}}) = \lambda_k(SB) = \lambda_k(BS) = \theta_k \lambda_k(B).$$

□

B Proof of Theorem 2

To improve readability, the proof of Theorem 2 is divided as follows: in Section B.1 we prove (i); in Sections B.2 and B.3, B.4 we prove the existence (ii.1), stability (ii.2) and uniqueness (ii.3) part, respectively, of (ii). The proof of (iii) is omitted since it is identical to the proof of (ii.1).

B.1 Proof of Theorem 2(i)

The condition for the existence of a unique equilibrium point for the system (5) can be rewritten in terms of the biggest eigenvalue of the normalized interaction matrix $H = \Delta^{-1}A = I - \mathcal{L}$. Define the following symmetric matrix:

$$H_{\text{sym}} := \Delta^{-\frac{1}{2}} A \Delta^{-\frac{1}{2}} = \Delta^{\frac{1}{2}} H \Delta^{-\frac{1}{2}} \sim H, \quad (15)$$

By construction, H and H_{sym} have the same eigenvalues.

Proof. First, notice that since $\lambda_i(\mathcal{L}) = 1 - \lambda_{n-i+1}(H)$ for all $i = 1, \dots, n$, then $\pi_1 = \frac{1}{1-\lambda_1(\mathcal{L})} = \frac{1}{\lambda_n(H)}$.

Let $V : \mathbb{R}^n \rightarrow \mathbb{R}_+$ be the Lyapunov function described by

$$V(x) = \sum_{i=1}^n \int_0^{x_i} \psi_i(s) ds. \quad (16)$$

Since each function $\psi_i(\cdot)$ is monotonically increasing and $\psi_i(s) = 0$ if and only if $s = 0$, then $V(x) > 0$ for all $x \in \mathbb{R}^n \setminus \{0\}$ and $V(0) = 0$. Moreover, $V(x)$ is radially unbounded. From the assumptions (A.1), (A.2) and (A.4), we know that

$$x_i \begin{cases} > \psi_i(x_i), & \text{if } x_i > 0 \text{ (i.e., } \psi_i(x_i) > 0) \\ < \psi_i(x_i), & \text{if } x_i < 0 \text{ (i.e., } \psi_i(x_i) < 0) \\ = 0, & \text{if } x_i = 0, \end{cases}$$

i.e., $\psi(x)^T \Delta x > \psi(x)^T \Delta \psi(x) > 0$ since $x \neq 0$. Hence, computing the derivative of V along the trajectories gives

$$\begin{aligned} \dot{V}(x) &= \psi(x)^T \dot{x} = \psi(x)^T [-\Delta x + \pi A \psi(x)] \\ &= -\psi(x)^T \Delta x + \psi(x)^T \Delta^{\frac{1}{2}} (\pi H_{\text{sym}}) \Delta^{\frac{1}{2}} \psi(x) \\ &< -\psi(x)^T \Delta^{\frac{1}{2}} (I - \pi H_{\text{sym}}) \Delta^{\frac{1}{2}} \psi(x) \\ &\leq -\left(1 - \frac{\pi}{\pi_1}\right) \psi(x)^T \Delta \psi(x) \end{aligned}$$

Since $\pi \leq \pi_1$, then $\dot{V}(x) < 0$ for all $x \neq 0$, i.e., the origin is globally asymptotically stable, hence the unique equilibrium point for the system (5). \square

B.2 Proof of Theorem 2(ii.1): existence

Since $\lambda_1(\mathcal{L}) = 1 - \lambda_n(H)$, the condition on π can be rewritten as $\pi > \pi_1 = \frac{1}{1-\lambda_1(\mathcal{L})} = \frac{1}{\lambda_n(H)}$. The proof follows [37, Chapter I§3]. The equilibrium points of the system (5) are solution of

$$\Phi(x, \pi) = -x + \pi H \psi(x) = 0. \quad (17)$$

Let $J := \frac{\partial \Phi}{\partial x}(0, \pi_1) = -I + \pi_1 H$ be the Jacobian matrix at $(0, \pi_1)$, and let v and w be its left and right eigenvectors (such that $w^T v = 1$), respectively, associated with

the zero eigenvalue (notice that $\lambda_n(J) = -1 + \pi_1 \lambda_n(H) = 0$ by construction). By assumption $\lambda_n(H)$ is simple, which implies that $\dim(\ker(J)) = 1$ and that we can write $\ker(J) = \text{span}\{v\}$ and $\text{range}(J) = (\ker(J^T))^\perp = (\text{span}\{w\})^\perp$. Let E denote the projection of \mathbb{R}^n onto $\text{range}(J)$, $E = I - vw^T$, and $I - E = vw^T$ the projection onto $\ker(J)$. The system of equations (17) can be expanded as follows

$$E\Phi(x, \pi) = 0 \quad (18a)$$

$$(I - E)\Phi(x, \pi) = 0. \quad (18b)$$

Split the vector x accordingly, $x = yv + r$ with $y \in \mathbb{R}$, $yv \in \ker(J)$ and $r \in \text{range}(J)$. From the implicit function theorem, solving (18a) for r gives $r = R(yv, \pi)$, which we can substitute in (18b) obtaining

$$0 = (I - E)\Phi(yv + R(yv, \pi), \pi) = (I - E)(-yv - R(yv, \pi) + \pi H\psi(yv + R(yv, \pi))).$$

Defining the center manifold $g : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ by

$$g(y) = w^T(I - E)\Phi(yv + R(yv, \pi), \pi)$$

we obtain that the zeros of g are in one-to-one correspondence with the solutions of $\Phi(x, \pi) = 0$. We say that the system (5) undergoes a pitchfork bifurcation at $(0, \pi_1)$ if

$$g = g_y = g_{yy} = g_\pi = 0, \quad g_{yyy} < 0, \quad g_{\pi y} > 0,$$

where the subscript indicates partial derivative.

In what follows we will use the following notation: $\frac{\partial^2 \psi}{\partial x^2}(x) := \text{diag}\{\frac{\partial^2 \psi_1}{\partial x_1^2}(x_1), \dots, \frac{\partial^2 \psi_n}{\partial x_n^2}(x_n)\}$, $\frac{\partial^3 \psi}{\partial x^3}(x) := \text{diag}\{\frac{\partial^3 \psi_1}{\partial x_1^3}(x_1), \dots, \frac{\partial^3 \psi_n}{\partial x_n^3}(x_n)\}$. Moreover, as $\Phi(x, \pi)$ is an odd function of x , we can neglect $R(\cdot)$ [37, p. 33]. Calculations yield

$$\Phi(0, \pi_1) = 0 \quad (19a)$$

$$\Phi_y(0, \pi_1) = \left. \left(-I + \pi H \frac{\partial \psi}{\partial x}(x) \right) \right|_{(0, \pi_1)} v = (-I + \pi_1 H)v = 0 \quad (19b)$$

$$\Phi_{yy}(0, \pi_1) = \left. \left(\pi H \frac{\partial \psi}{\partial x} \left(\frac{\partial \psi}{\partial x}(x)v \right) \right) \right|_{(0, \pi_1)} v = \pi_1 H \frac{\partial^2 \psi}{\partial x^2}(0) \begin{bmatrix} v_1^2 \\ \vdots \\ v_n^2 \end{bmatrix} = 0 \quad (19c)$$

$$\Phi_\pi(0, \pi_1) = H\psi(0) = 0 \quad (19d)$$

$$\Phi_{yyy}(0, \pi_1) = \left. \left(\pi H \frac{\partial \psi}{\partial x} \left(\frac{\partial \psi}{\partial x} \left(\frac{\partial \psi}{\partial x}(x)v \right) v \right) \right) \right|_{(0, \pi_1)} v = \pi_1 H \frac{\partial^3 \psi}{\partial x^3}(0) \begin{bmatrix} v_1^3 \\ \vdots \\ v_n^3 \end{bmatrix} \quad (19e)$$

$$\Phi_{\pi y}(0, \pi_1) = H \frac{\partial \psi}{\partial x}(x) \Big|_{(0, \pi_1)} v = Hv = \lambda_n(H)v \quad (19f)$$

where $\psi(0) = 0$, $\frac{\partial \psi}{\partial x}(0) = I$ and $\frac{\partial^2 \psi}{\partial x^2}(0) = 0$ directly follow from assumptions (A.1), (A.2) and (A.4). Equations (19a), (19b), (19c) and (19d) yield respectively

$$g(0, \pi_1) = 0, \quad g_y(0, \pi_1) = 0, \quad g_{yy}(0, \pi_1) = 0, \quad g_\pi(0, \pi_1) = 0.$$

The last two derivatives are given by

$$g_{\pi y}(0, \pi_1) = \lambda_n(H) w^T v = \lambda_n(H) > 0$$

and

$$g_{yy}(0, \pi_1) = \pi_1 w^T H \frac{\partial^3 \psi}{\partial x^3}(0) \begin{bmatrix} v_1^3 \\ \vdots \\ v_n^3 \end{bmatrix} = \pi_1 \lambda_n(H) \sum_{i=1}^n w_i v_i^3 \frac{\partial^3 \psi_i}{\partial x_i^3}(0) < 0.$$

The last inequality holds since $\pi_1 \lambda_n(H) = 1$, $w_i v_i \geq 0 \forall i$ and $\frac{\partial^3 \psi_i}{\partial x_i^3}(0) < 0 \forall i$, which follows from assumption (A.4). We have shown that $(0, \pi_1)$ is a pitchfork bifurcation point for the system (5) and, as a consequence, that two new (nontrivial) equilibrium branches are created. \square

B.3 Proof of Theorem 2(ii.2): stability

Before stating the proof we show that, when $\pi \in (\pi_1, \pi_2)$, all the nontrivial equilibrium points $x^* \neq 0$ of the system (5) (if present) are locally asymptotically stable. This lemma will be used in the proof of both the stability and uniqueness part of Theorem 2(ii). Moreover, notice that the threshold values π_1 and π_2 can be rewritten as follows:

$$\pi_1 = \frac{1}{\lambda_n(H)}, \quad \pi_2 = \frac{1}{\lambda_{n-1}(H)}.$$

Lemma 6. *Under the assumptions of Theorem 2(ii), when $\pi \in (\pi_1, \pi_2)$, if $x^* \neq 0$ is a (nontrivial) equilibrium point of the system (5) then it is locally asymptotically stable.*

Proof. Let $x^* \neq 0$ be an equilibrium point for the system (5),

$$x^* = \pi H \psi(x^*). \tag{20}$$

To prove that x^* is locally asymptotically stable, consider the linearization around x^* ,

$$\dot{x} = \Delta \left(-I + \pi H \frac{\partial \psi}{\partial x}(x^*) \right) (x - x^*). \tag{21}$$

The equilibrium point x^* is asymptotically stable for the system (21), and consequently locally asymptotically stable for the system (5), if the matrix $\Delta(-I + \pi H \frac{\partial \psi}{\partial x}(x^*))$ is Hurwitz stable, i.e., its eigenvalues are strictly negative. Since Δ is diagonal and positive definite, this holds if and only if the matrix $-I + \pi H \frac{\partial \psi}{\partial x}(x^*)$ is Hurwitz stable. The following proof shows that the largest eigenvalue of $-I + \pi H \frac{\partial \psi}{\partial x}(x^*)$ is strictly smaller than 0, i.e., that $\pi \lambda_n(H \frac{\partial \psi}{\partial x}(x^*)) < 1$. It is a two-steps proof, showing first that $\pi \lambda_n(H \frac{\partial \psi}{\partial x}(x^*)) \leq 1$ and then by contradiction that $\pi \lambda_n(H \frac{\partial \psi}{\partial x}(x^*)) \neq 1$.

Step 1. From the assumptions (A.1), (A.2) and (A.4), for all $i = 1, \dots, n$ it holds that

$$\begin{cases} x_i^* > \psi(x_i^*) > \frac{\partial \psi_i}{\partial x_i}(x_i^*) x_i^*, & \text{if } x_i^* > 0 \\ x_i^* < \psi(x_i^*) < \frac{\partial \psi_i}{\partial x_i}(x_i^*) x_i^*, & \text{if } x_i^* < 0 \\ x_i^* = \psi(x_i^*) = 0, \frac{\partial \psi_i}{\partial x_i}(x_i^*) = 1, & \text{if } x_i^* = 0. \end{cases}$$

Therefore, $\exists \Xi(x^*) = \text{diag}\{\xi_1(x_1^*), \dots, \xi_n(x_n^*)\} \in \mathbb{R}^{n \times n}$ such that, $\forall i = 1, \dots, n$,

$$\begin{aligned} \psi(x_i^*) &= \xi_i(x_i^*) \cdot \frac{\partial \psi_i}{\partial x_i}(x_i^*) \cdot x_i^*, \\ \begin{cases} \xi_i(x_i^*) > 1, & 0 < \xi_i(x_i^*) \cdot \frac{\partial \psi_i}{\partial x_i}(x_i^*) < 1, \\ \xi_i(x_i^*) = 1, & \xi_i(x_i^*) \cdot \frac{\partial \psi_i}{\partial x_i}(x_i^*) = 1, \end{cases} &\begin{cases} \text{if } x_i^* \neq 0 \\ \text{if } x_i^* = 0. \end{cases} \end{aligned}$$

This can be rewritten in compact form as

$$\psi(x^*) = \Xi(x^*) \cdot \frac{\partial \psi}{\partial x}(x^*) \cdot x^*, \quad (22)$$

where

$$\text{diag}\{\Xi(x^*)\} \geq \mathbb{1}, \quad 0 < \text{diag}\{\Xi(x^*) \cdot \frac{\partial \psi}{\partial x}(x^*)\} \leq \mathbb{1}. \quad (23)$$

To simplify the notation, we will neglect the dependence from x^* in what follows; moreover, we define $\Psi_x := \frac{\partial \psi}{\partial x}(x^*)$. From (20) and (22), it follows that

$$x^* = (\pi H \cdot \Psi_x \cdot \Xi) x^*, \quad (24)$$

that is, $(1, x^*)$ is an eigenpair of $\pi H \Psi_x \Xi$. Therefore, by Lemma 5 with $B = \pi H$ and $S = \Psi_x \Xi$, and by (23), it follows that for all $k \in \{1, \dots, n\}$, $\exists \theta_k \in (0, 1]$ such that $\lambda_k(\pi H \Psi_x \Xi) = \theta_k \pi \lambda_k(H)$. In particular $\exists \theta_n, \theta_{n-1} \in (0, 1]$ such that

$$\lambda_n(\pi H \Psi_x \Xi) = \theta_n \pi \lambda_n(H), \quad (25)$$

$$\lambda_{n-1}(\pi H \Psi_x \Xi) = \theta_{n-1} \pi \lambda_{n-1}(H) < \theta_{n-1} \leq 1, \quad (26)$$

where $\pi \lambda_{n-1}(H) = \frac{\pi}{\pi_2} < 1$ under the assumption $\pi \in (\pi_1, \pi_2)$. Then, (24) and (26) yield $1 = \lambda_n(\pi H \Psi_x \Xi)$.

Applying again Lemma 5 with $B = \pi H \Psi_x$ and $S = \Xi$, there exists a $\theta \in [\min_i\{\xi_i\}, \max_i\{\xi_i\}]$ such that

$$1 = \lambda_n(\pi H \Psi_x \Xi) = \theta \pi \lambda_n(H \Psi_x) \geq \pi \lambda_n(H \Psi_x) > 0,$$

since $\xi_i \geq 1$ for all i implies $\theta \geq 1$. Hence, $\pi \lambda_n(H \Psi_x) \leq 1$.

Step 2. Suppose by contradiction that $\pi \lambda_n(H \Psi_x) = 1$. First, we need to define the “symmetric versions” of the matrices $H \Psi_x$ and $H \Psi_x \Xi$.

$$\begin{aligned} [H \Psi_x]_{\text{sym}} &:= (\Psi_x)^{\frac{1}{2}} H_{\text{sym}} (\Psi_x)^{\frac{1}{2}} \\ &\sim (\Delta \Psi_x)^{-\frac{1}{2}} \cdot \left((\Psi_x)^{\frac{1}{2}} H_{\text{sym}} (\Psi_x)^{\frac{1}{2}} \right) \cdot (\Delta \Psi_x)^{\frac{1}{2}} \\ &= (\Delta^{-\frac{1}{2}} H_{\text{sym}} \Delta^{\frac{1}{2}}) \Psi_x = H \Psi_x \end{aligned}$$

with H_{sym} defined in (15), and

$$\begin{aligned} [H\Psi_x \Xi]_{\text{sym}} &:= (\Psi_x \Xi)^{\frac{1}{2}} H_{\text{sym}} (\Psi_x \Xi)^{\frac{1}{2}} \\ &\sim (\Delta \Psi_x \Xi)^{-\frac{1}{2}} \cdot \left((\Psi_x \Xi)^{\frac{1}{2}} H_{\text{sym}} (\Psi_x \Xi)^{\frac{1}{2}} \right) \cdot (\Delta \Psi_x \Xi)^{\frac{1}{2}} \\ &= (\Delta^{-\frac{1}{2}} H_{\text{sym}} \Delta^{\frac{1}{2}}) \Psi_x \Xi = H \Psi_x \Xi. \end{aligned}$$

Moreover, $[H\Psi_x \Xi]_{\text{sym}} = \Xi^{\frac{1}{2}} [H\Psi_x]_{\text{sym}} \Xi^{\frac{1}{2}}$ by construction. From similarity, $1 = \pi \lambda_n([H\Psi_x]_{\text{sym}})$. Let v be the right eigenvector of $\pi[H\Psi_x]_{\text{sym}}$ associated with its largest eigenvalue $\pi \lambda_n([H\Psi_x]_{\text{sym}}) = 1$. From (24), since $(1, x^*)$ is an eigenpair of $\pi H \Psi_x \Xi$ (with 1 being the largest eigenvalue, as proven previously), it follows that $(1, (\Delta \Psi_x \Xi)^{\frac{1}{2}} x^*)$ is an eigenpair of $\pi[H\Psi_x \Xi]_{\text{sym}}$ (with 1 being the largest eigenvalue). To summarize

$$\begin{cases} [H\Psi_x \Xi]_{\text{sym}} = \Xi^{\frac{1}{2}} [H\Psi_x]_{\text{sym}} \Xi^{\frac{1}{2}}; \\ \pi \lambda_n([H\Psi_x]_{\text{sym}}) = 1 \quad \text{and} \quad \pi[H\Psi_x]_{\text{sym}} v = v; \\ \pi \lambda_n([H\Psi_x \Xi]_{\text{sym}}) = 1 \quad \text{and} \quad \pi[H\Psi_x \Xi]_{\text{sym}} (\Delta \Psi_x \Xi)^{\frac{1}{2}} x^* = (\Delta \Psi_x \Xi)^{\frac{1}{2}} x^*. \end{cases} \quad (27)$$

Applying Rayleigh's Theorem [27, Thm 4.2.2] with $\pi[H\Psi_x \Xi]_{\text{sym}}$, one obtains

$$\begin{aligned} 1 &= \pi \lambda_n([H\Psi_x \Xi]_{\text{sym}}) = \pi \max_{y \neq 0} \frac{y^T [H\Psi_x \Xi]_{\text{sym}} y}{y^T y} \\ &= \pi \max_{y \neq 0} \frac{(y^T \Xi^{\frac{1}{2}}) \cdot [H\Psi_x]_{\text{sym}} \cdot (\Xi^{\frac{1}{2}} y)}{y^T y} \\ &\Downarrow \quad \text{with} \quad y = \Xi^{-\frac{1}{2}} v \neq 0 \\ &\geq \pi \frac{v^T [H\Psi_x]_{\text{sym}} v}{v^T \Xi^{-1} v} = \frac{v^T v}{v^T \Xi^{-1} v}. \end{aligned} \quad (28)$$

The inequality $v^T v \leq v^T \Xi^{-1} v$, which can be rewritten as

$$\sum_{i=1}^n v_i^2 \leq \sum_{i=1}^n \frac{v_i^2}{\xi_i} = \sum_{i: \xi_i=1} v_i^2 + \sum_{i: \xi_i>1} \frac{v_i^2}{\xi_i},$$

holds, as equality, if and only if

$$v_i = 0 \quad \forall i \quad \text{s.t.} \quad \xi_i > 1 \quad (\Leftrightarrow x_i^* \neq 0). \quad (29)$$

Hence, (28) can only hold as equality if (29) holds, which further implies that $\Xi^{\frac{1}{2}} v = v$ and consequently that $(1, v)$ is an eigenpair of $\pi[H\Psi_x \Xi]_{\text{sym}}$ (with 1 being the largest eigenvalue). Indeed

$$\pi[H\Psi_x \Xi]_{\text{sym}} v = \pi \Xi^{\frac{1}{2}} [H\Psi_x]_{\text{sym}} \Xi^{\frac{1}{2}} v = \pi \Xi^{\frac{1}{2}} [H\Psi_x]_{\text{sym}} v = \Xi^{\frac{1}{2}} v = v.$$

Since $\lambda_n(\pi[H\Psi_x \Xi]_{\text{sym}}) = 1$ is simple, as shown in (25) and (26), it follows that v should be equivalent to the corresponding right eigenvector of $\pi[H\Psi_x \Xi]_{\text{sym}}$,

i.e., $(\Delta\Psi_x \Xi)^{\frac{1}{2}}x^*$. This however yields a contradiction, since by (29) $v^T x^* = 0$ (in particular $v_i = 0$ for all $x_i^* \neq 0$ and $v_i \neq 0$ for at least one i s.t. $x_i^* = 0$).

To conclude, $\pi\lambda_n(H\Psi_x) < 1$ and the matrix $-I + \pi H \frac{\partial\psi}{\partial x}(x^*)$ is Hurwitz stable, which implies that the nontrivial equilibrium point x^* is locally asymptotically stable. \square

Finally, we are ready to prove Theorem 2(ii.2): stability.

Proof of Theorem 2(ii.2). The linearized system around the origin is $\dot{x} = \Delta(-I + \pi H)x$ (given assumption (A.2)), where Δ is positive definite and $-I + \pi H$ has eigenvalues $\{\pi\lambda_1(H) - 1, \dots, \pi\lambda_n(H) - 1\}$. When $\pi > \pi_1$, the matrix $-I + \pi H$ has at least one positive eigenvalue, which proves the instability of the origin as equilibrium point of (5).

Instead, let $x^* \neq 0$ be an equilibrium point of the system 5, whose existence is shown in Theorem 2(ii.1). To prove its stability we can apply Lemma 6, which shows that if $x^* \neq 0$ is an equilibrium point of (5) then it must be locally asymptotically stable. \square

B.4 Proof of Theorem 2(ii.3): uniqueness

Let $f(x, \pi) = \Delta[-x + \pi H\psi(x)]$. We can divide the proof into two parts. (i) We prove that in a neighborhood of π_1 the system admits only 3 equilibria. (ii) We prove, by contradiction, that there are no bifurcation points for $\pi \in (\pi_1, \pi_2)$. We use Lemma 6 which proves that *all* the nontrivial equilibrium points \bar{x} of the system (5) are locally asymptotically stable, hence isolated and with each matrix $\frac{\partial f}{\partial x}(\bar{x}, \bar{\pi})$ Hurwitz.

(i). The existence is shown in Theorem 2(ii.1), where it is also proven that the bifurcation is a pitchfork. This means that in a neighborhood of π_1 the system (5) admits exactly three equilibrium points: the origin and two nontrivial equilibrium points, $\pm x^* \neq 0$.

(ii). The necessary condition for an equilibrium point $(\bar{x}, \bar{\pi})$ (where $\bar{\pi} \in (\pi_1, \pi_2)$) to be a bifurcation point is that the Jacobian $\frac{\partial f}{\partial x}(\bar{x}, \bar{\pi}) = \Delta[-I + \pi H \frac{\partial\psi}{\partial x}(\bar{x})]$ is not invertible (i.e., there is an $i \in \{1, \dots, n\}$ such that $\bar{\pi}\lambda_i(H \frac{\partial\psi}{\partial x}(\bar{x})) = 1$). Suppose by contradiction that, for $\bar{\pi} \in (\pi_1, \pi_2)$, \bar{x} is an equilibrium point of the system (5), i.e., $\bar{x} = \bar{\pi}H\psi(\bar{x})$, and a bifurcation point, i.e., $\exists i$ s.t. $\bar{\pi}\lambda_i(H \frac{\partial\psi}{\partial x}(\bar{x})) = 1$. However, Lemma 6 shows that if $\bar{x} \neq 0$ is an equilibrium point of (5) then it must be locally asymptotically stable., i.e., $\bar{\pi}\lambda_n(H \frac{\partial\psi}{\partial x}(\bar{x})) < 1$. Moreover, if $\bar{x} = 0$, $\bar{\pi}\lambda_{n-1}(H) < 1 < \bar{\pi}\lambda_n(H)$ for $\bar{\pi} \in (\pi_1, \pi_2)$. Hence, \bar{x} cannot be a bifurcation point.

To conclude, we know that the system (5) admits three equilibria $(0, x^*, -x^*)$ and that it cannot bifurcate further from them for values of $\pi \in (\pi_1, \pi_2)$. Hence, the only possible equilibrium points for the system are the origin and those originated from the first bifurcation at $\pi = \pi_1$. \square

C Proof of Lemma 2

It is useful to first reformulate structural balance of \mathcal{G} (see Section 2) in terms of the matrix H_{sym} defined in (15), which is congruent to A and similar to H : \mathcal{G}

(connected) is structurally balanced if and only if there exists a signature matrix $S = \text{diag}\{s_1, \dots, s_n\}$ with diagonal entries $s_i = \pm 1$ ($i = 1, \dots, n$) such that $SH_{\text{sym}}S$ is nonnegative. We are now ready to prove Lemma 2. In what follows we say that a symmetric matrix is “special” if it has zero diagonal entries, and “elliptic” if it has exactly one and simple positive eigenvalue. This notation is from [38], whose results are used in the proof. Notice that both A and H_{sym} are special matrices.

Proof. Let \mathcal{G} be structurally unbalanced.

Assume by contradiction that $\lambda_2(\mathcal{L}) > 1$, which is equivalent to $\lambda_{n-1}(H) = \lambda_{n-1}(H_{\text{sym}}) < 0$. This means that H_{sym} is special, nonsingular and elliptic. Then, all the off-diagonal entries of H_{sym} are different from zero (see [38, Corollary 2.7]) and hence there exists a diagonal (signature) matrix $S = \text{diag}\{s_1, \dots, s_n\}$ with $s_i = \pm 1$ s.t. $SH_{\text{sym}}S$ is nonnegative (see [38, Thm 2.5]). Therefore, \mathcal{G} is structurally balanced and we obtain a contradiction.

Assume by contradiction that $\lambda_2(\mathcal{L}) = 1$, which is equivalent to $\lambda_{n-1}(H) = \lambda_{n-1}(H_{\text{sym}}) = 0$. This means that H_{sym} is special, singular and elliptic. Let $r = \text{rank}(H_{\text{sym}}) < n$ and observe that H_{sym} cannot have zero rows or columns since \mathcal{G} is connected. From [38, Thm 2.9], there exist a permutation matrix $P \in \mathbb{R}^{n \times n}$ and an integer $t \in \{r, \dots, n\}$ s.t.

$$PH_{\text{sym}}P^T = DH_{\text{sym}_0}D^T,$$

where $H_{\text{sym}_0} \in \mathbb{R}^{t \times t}$ is elliptic, special and nonnegative with rank r and $D = d_1 \oplus d_2 \oplus \dots \oplus d_t \in \mathbb{R}^{n \times t}$ (here \oplus indicates the direct sum) is a block matrix where each $d_i \in \mathbb{R}^{n_i}$ (with $\sum_{i=1}^t n_i = n$) is a unit vector (i.e., $\|d_i\|_2 = 1$) with all elements different from zero (i.e., $|d_i| > 0$). Define the signature (block) matrix $S := S_1 \oplus \dots \oplus S_t \in \mathbb{R}^{n \times n}$, where each $S_i := \text{diag}\{\text{sign}(d_i)\} \in \mathbb{R}^{n_i \times n_i}$, $i = 1, \dots, t$, is a signature matrix. Then, $PH_{\text{sym}}P^T$ can be rewritten as

$$PH_{\text{sym}}P^T = S|D|H_{\text{sym}_0}|D^T|S.$$

It follows that $S(PH_{\text{sym}}P^T)S$ is nonnegative, and hence, that $\bar{S}H_{\text{sym}}\bar{S}$ is nonnegative, with $\bar{S} = P^TSP$ still a signature matrix. Therefore, \mathcal{G} is structurally balanced and we obtain a contradiction.

D Proof of Theorem 3

In the following proofs we use the notation $S_x := \text{diag}\{\text{sign}(x)\}$ where $x \in \mathbb{R}^n$ and the signum function is defined as $\text{sign}(y) = 1$ if $y \leq 0$ or $\text{sign}(y) = -1$ if $y < 0$, where $y \in \mathbb{R}$.

Remark 5. The frustration of the network \mathcal{G} is defined in equation (2), which can be

rewritten as follows:

$$\begin{aligned}
2\epsilon(\mathcal{G}) &= \min_{\substack{S=\text{diag}\{s_1, \dots, s_n\} \\ s_i=\pm 1 \forall i}} \mathbb{1}^T (|H| - SHS) \mathbb{1} \\
&= n - \max_{\substack{S=\text{diag}\{s_1, \dots, s_n\} \\ s_i=\pm 1 \forall i}} \mathbb{1}^T SHS \mathbb{1} \\
&= n - \max_{\substack{S=\text{diag}\{s_1, \dots, s_n\} \\ s_i=\pm 1 \forall i}} \mathbb{1}^T S \left(\frac{H + H^T}{2} \right) S \mathbb{1}.
\end{aligned}$$

[Remark 6.] Let

$$u^* = S_{u^*} \mathbb{1} \text{ where } S_{u^*} = \arg \max_{\substack{S=\text{diag}\{s_1, \dots, s_n\} \\ s_i=\pm 1 \forall i}} \mathbb{1}^T SHS \mathbb{1}, \quad (30)$$

that is, $2\epsilon(\mathcal{G}) = n - \mathbb{1}^T S_{u^*} HS_{u^*} \mathbb{1}$. From the results on (symmetric) Hopfield neural networks (see [23, 39]) we know that the vector u^* satisfies $u^* = \text{sign} \left(\left(\frac{H + H^T}{2} \right) u^* \right)$. If H is symmetric, then $u^* = \text{sign} \left(H^T u^* \right)$, meaning that the vector $S_{u^*} \text{sign} \left(H^T S_{u^*} \mathbb{1} \right) = S_{u^*} \text{sign} \left(HS_{u^*} \mathbb{1} \right)$ has all strictly positive components (equal to 1).

D.1 Proof of Theorem 3(i)

The proof can be divided into three steps. First, we show that if x^* is an equilibrium point of the system (5) and H is symmetric, then $\|x^*\|_1 \leq \pi(n - 2\epsilon(\mathcal{G}))$. Then, we show that if H is not symmetric and x^* is such that $x_i^* \neq 0$ for all i (or, $|x^*| > 0$) we can apply the same reasoning to prove that $\|x^*\|_1 \leq \pi(n - 2\epsilon(\mathcal{G}))$. Finally, we complete the proof and show that each equilibrium point x^* of the system (5) (without assuming that H is symmetric) satisfies the inequality $\|x^*\|_1 \leq \pi(n - 2\epsilon(\mathcal{G}))$.

Step 1. We first consider the particular case of the matrix $H = \Delta^{-1}A$ being symmetric, that is, $\Delta = \delta I$. Let x^* be an equilibrium point of the system (5), that is, $x^* = \pi H \psi(x^*)$ and let S_{x^*} be its signature, i.e., $|x^*| = S_{x^*} x^*$. It follows that

$$|x^*| = \pi S_{x^*} H \psi(x^*) = \pi |H \psi(x^*)| = \pi S_{x^*} H S_{x^*} |\psi(x^*)|.$$

Observe that $S_{H\psi(x^*)} = S_{x^*} = S_{\psi(x^*)}$. Then

$$\frac{\|x^*\|_1}{\pi} = \frac{\mathbb{1}^T |x^*|}{\pi} = \mathbb{1}^T |H \psi(x^*)| = \mathbb{1}^T S_{x^*} H S_{x^*} |\psi(x^*)|$$

and

$$\max_{\substack{x^* \in \mathbb{R}^n \text{ s.t.} \\ x^* = \pi H \psi(x^*)}} \frac{\|x^*\|_1}{\pi} = \max_{\substack{x^* \in \mathbb{R}^n \text{ s.t.} \\ x^* = \pi H \psi(x^*)}} \mathbb{1}^T |H \psi(x^*)| \leq \max_{\substack{u \in \mathbb{R}^n \text{ s.t.} \\ S_{Hu} = S_u, -\mathbb{1} \leq u \leq \mathbb{1}}} \mathbb{1}^T |Hu|.$$

because the constraint $x^* = \pi H\psi(x^*)$ (i.e., x^* is an equilibrium point) implies the constraint $S_{H\psi(x^*)} = S_{\psi(x^*)}$, and $|\psi_i(x_i)| \leq 1 \forall x_i \in \mathbb{R}$. Then

$$\begin{aligned} \max_{\substack{x^* \in \mathbb{R}^n \text{ s.t.} \\ x^* = \pi H\psi(x^*)}} \frac{\|x^*\|_1}{\pi} &\leq \max_{\substack{u \in \mathbb{R}^n \text{ s.t.} \\ S_{Hu} = S_u, -1 \leq u \leq 1}} \mathbf{1}^T |Hu| \\ &= \max_{\substack{u \in \mathbb{R}^n \text{ s.t.} \\ -1 \leq u \leq 1}} \mathbf{1}^T S_u H S_u |u| \end{aligned} \quad (31a)$$

$$= \max_{\substack{S_u = \text{diag}\{s_{u,1}, \dots, s_{u,n}\} \\ s_{u,i} = \pm 1 \forall i}} \mathbf{1}^T S_u H S_u \mathbf{1} \quad (31b)$$

$$= n - 2\epsilon(\mathcal{G}).$$

Notice that $\mathbf{1}^T S_u H S_u |u| \leq |\mathbf{1}^T S_u H S_u| \mathbf{1}$ for all $u \in \mathbb{R}^n$ s.t. $|u| \leq \mathbf{1}$. The equality between (31a) and (31b) means that the maxima are obtained when u lies in the corners of the hypercube $|u| \leq \mathbf{1}$ (i.e., $|u| = \mathbf{1}$). In particular, u^* defined in (30) is a solution of this maximization problem since it is feasible and $\text{sign}(\mathbf{1}^T S_{u^*} H S_{u^*}) = \text{sign}(\mathbf{1}^T S_{u^*} H) S_{u^*} = \mathbf{1}^T \geq 0$, meaning that $\mathbf{1}^T S_{u^*} H S_{u^*} = |\mathbf{1}^T S_{u^*} H S_{u^*}|$.

Step 2. Let x^* be an equilibrium point of the system (5) and assume that $x_i^* \neq 0$ for all i . In this step we do not assume the symmetry of the matrix H . Following the reasoning of Step 1, and by adding the additional constraint $|u| > 0$ (which comes from $x_i^* \neq 0$ for all i), it is still possible to prove (see below) that the maxima are obtained when u lies in the corners of the hypercube $|u| \leq \mathbf{1}$ (which yields the equivalence between (31a) and (31b)). This is equivalent to show that the maxima are obtained when u is s.t. $\mathbf{1}^T S_u H S_u \geq 0$. Let

$$\tilde{u} = \arg \max_{\substack{u \in \mathbb{R}^n \text{ s.t.} \\ -1 \leq u \leq \mathbf{1}, |u| > 0}} \mathbf{1}^T S_u H S_u |u|$$

and $v^T = \mathbf{1}^T S_{\tilde{u}} H S_{\tilde{u}}$. Suppose, by contradiction, that $\exists j$ s.t. $v_j < 0$ (and $\tilde{u}_j \neq 0$, since $|\tilde{u}| > 0$). Define \bar{u} s.t. $S_{\bar{u}} = S_{\tilde{u}}$ and

$$|\bar{u}_j| = \begin{cases} 1, & \text{if } v_j \geq 0 \\ \varepsilon |\tilde{u}_j|, & \text{if } v_j < 0 \end{cases}$$

with $0 < \varepsilon < 1$ (which means that $|\bar{u}| \leq \mathbf{1}$ and $|\bar{u}| > 0$, i.e., \bar{u} is a feasible point of the maximization problem). Then

$$\begin{aligned} \mathbf{1}^T S_{\bar{u}} H S_{\bar{u}} |\bar{u}| &= v^T |\bar{u}| = \sum_{j:v_j \geq 0} v_j |\bar{u}_j| + \sum_{i:v_i < 0} v_j |\bar{u}_j| \\ &= \sum_{j:v_j \geq 0} v_j + \sum_{j:v_j < 0} v_j \varepsilon |\tilde{u}_j| \\ &\geq \sum_{j:v_j \geq 0} v_j |\tilde{u}_j| + \sum_{j:v_j < 0} v_j \varepsilon |\tilde{u}_j| \\ &> \sum_{j:v_j \geq 0} v_j |\tilde{u}_j| + \sum_{i:v_i < 0} v_j |\tilde{u}_j| = \mathbf{1}^T S_{\tilde{u}} H S_{\tilde{u}} |\tilde{u}|, \end{aligned}$$

which implies a contradiction.

Step 3. Finally, we want to extend the idea presented in the previous steps to show, without imposing any constraint on H or x^* (except for being an equilibrium point of (5)), that $\|x^*\|_1 \leq \pi(n - 2\epsilon(\mathcal{G}))$. In this case, assume that $x_i^* = 0$ for some i : let n_0 (resp., $n - n_0$) be the number of zero (resp., nonzero) components of x_i^* and let P be a permutation matrix s.t. $Px^* = \begin{bmatrix} x_{nz}^* \\ 0 \end{bmatrix}$, where $x_{nz}^* \in \mathbb{R}^{n-n_0}$ and $|x_{nz}^*| > 0$. Let $P\psi(x^*) = \begin{bmatrix} \psi(x_{nz}^*) \\ 0 \end{bmatrix}$ and $PHP^T = \begin{bmatrix} H_{nz} & * \\ * & * \end{bmatrix}$. Then, $\|x^*\|_1 = \|x_{nz}^*\|_1$, $x_{nz}^* = \pi H_{nz}\psi(x_{nz}^*)$ and (following the reasoning of Step 2)

$$\frac{\|x^*\|_1}{\pi} \leq \max_{\substack{S_{nz} = \text{diag}\{s_1, \dots, s_{n-n_0}\} \\ s_i = \pm 1 \forall i}} \mathbf{1}_{n-n_0}^T S_{nz} H_{nz} S_{nz} \mathbf{1}_{n-n_0}.$$

Notice that for all signature matrices $S_{nz} = \text{diag}\{s_1, \dots, s_{n-n_0}\} \in \mathbb{R}^{n-n_0 \times n-n_0}$ with $s_i = \pm 1 \forall i$, the following inequality holds:

$$\begin{aligned} & \mathbf{1}_{n-n_0}^T S_{nz} H_{nz} S_{nz} \mathbf{1}_{n-n_0} \\ &= \left([\mathbf{1}_{n-n_0}^T \quad 0_{n_0}^T] P \right) \cdot \left(P^T \begin{bmatrix} S_{nz} & 0 \\ 0 & I_{n_0} \end{bmatrix} P \right) \cdot H \cdot \left(P^T \begin{bmatrix} S_{nz} & 0 \\ 0 & I_{n_0} \end{bmatrix} P \right) \cdot \left(P^T \begin{bmatrix} \mathbf{1}_{n-n_0} \\ 0_{n_0} \end{bmatrix} \right) \\ &\leq \max_{u \in \mathbb{R}^n \text{ s.t. } |u| \leq 1} u^T H u = n - 2\epsilon(\mathcal{G}). \end{aligned}$$

Summarizing, we have shown that $\|x^*\|_1 \leq \pi(n - 2\epsilon(\mathcal{G}))$. \square

D.2 Proof of Theorem 3(ii)

From Theorem 3(i) we know that the set

$$\Omega_{\epsilon(\mathcal{G})} = \{x \in \mathbb{R}^n : \|x\|_1 \leq \pi(n - 2\epsilon(\mathcal{G}))\}$$

contains all equilibria of the system (5). We now want to prove that $\Omega_{\epsilon(\mathcal{G})}$ is attractive when the signed normalized Laplacian \mathcal{L} or, equivalently, the normalized interaction matrix $H = \Delta^{-1}A = I - \mathcal{L}$, is symmetric (i.e., $\Delta = \delta I$).

Let $V : \mathbb{R}^n \rightarrow \mathbb{R}_+$ be the Lyapunov function described by

$$V(x) = \begin{cases} \frac{1}{\delta} (\|x\|_1 - \pi(n - 2\epsilon(\mathcal{G}))), & x \notin \Omega_{\epsilon(\mathcal{G})} \\ 0, & x \in \Omega_{\epsilon(\mathcal{G})} \end{cases}. \quad (32)$$

Since Δ is positive definite then $V(x) > 0$ for all $x \notin \Omega_{\epsilon(\mathcal{G})}$. Moreover, $V(x)$ is radially unbounded. Let $S_x := \text{diag}\{\text{sign}(x)\}$ (observe that $S_x = S_{\psi(x)}$). The

upper Dini derivative of V along the trajectories (5) (with $\Delta = \delta I$) gives

$$\begin{aligned}
d^+V(x) &= \limsup_{s \rightarrow 0^+} \frac{V(x + s\dot{x}) - V(x)}{s} \\
&= \frac{1}{\delta} \limsup_{s \rightarrow 0^+} \frac{\sum_i |x_i + s\dot{x}_i| - |x_i|}{s} \\
&= \frac{1}{\delta} \sum_i d^+ |x_i| = \frac{1}{\delta} \sum_{i=1}^n \text{sign}(x_i) \dot{x}_i \\
&= \mathbf{1}^T \Delta^{-1} S_x \dot{x} = \mathbf{1}^T S_x [-x + \pi H \psi(x)] \\
&= -\|x\|_1 + \pi \mathbf{1}^T S_x H \psi(x) \\
&= -\|x\|_1 + \pi \mathbf{1}^T S_x H S_x |\psi(x)|, \\
&\leq -\|x\|_1 + \pi \max_{u \in \mathbb{R}^n \text{ s.t. } |u| \leq \mathbf{1}} \mathbf{1}^T S_u H u
\end{aligned}$$

Again, the intuition is that $\mathbf{1}^T S_u H u = \mathbf{1}^T S_u S_{H u} |H u| \leq \mathbf{1}^T |H u|$, which means that the maxima of $\mathbf{1}^T S_u H u$ are obtained when $S_{H u} = S_u$. Hence, following the reasoning of the proof of Theorem 3, we conclude that for all $x \notin \Omega_{\epsilon(\mathcal{G})}$

$$d^+V(x) \leq -\|x\|_1 + \pi(n - 2\epsilon(\mathcal{G})) < 0.$$

□

E Proof of Proposition 1

To show that the bound (6) holds, observe first that $1 \leq \pi_1 \leq \pi_2$ holds trivially since $\frac{1}{\pi_1} = 1 - \lambda_1(\mathcal{L}) \leq 1$ and $\pi_1 = \frac{1}{1 - \lambda_1(\mathcal{L})} \leq \frac{1}{1 - \lambda_2(\mathcal{L})} = \pi_2$. Therefore, it remains to show that $\pi_1 \leq \frac{n}{n - 2\epsilon(\mathcal{G})}$. It is easier to write the rest of the proof in terms of the normalized interaction matrix $H = I - \mathcal{L} = \Delta^{-1} A$; then, $\pi_1 = \frac{1}{1 - \lambda_1(\mathcal{L})} = \frac{1}{\lambda_n(H)}$. Let S_{best} be the matrix yielding the minimum value of energy in (2), that is, $2\epsilon(\mathcal{G}) = n - \mathbf{1}^T S_{\text{best}} H S_{\text{best}} \mathbf{1}$ (Remark 5 in the proof of Theorem 3 shows how the frustration of the network \mathcal{G} can be rewritten in terms of H). Since \mathcal{L} is symmetric then H is also symmetric and using the Rayleigh's Theorem [27, Thm 4.2.2], it follows that

$$\frac{1}{\pi_1} = \lambda_n(H) = \max_{v \in \mathbb{R}^n} \frac{v^T H v}{v^T v} \geq \frac{\mathbf{1}^T S_{\text{best}} H S_{\text{best}} \mathbf{1}}{\mathbf{1}^T \mathbf{1}} = \frac{n - 2\epsilon(\mathcal{G})}{n},$$

which implies (6). □

F Proof of Proposition 2

The first part of the proof holds for both structurally balanced and structurally unbalanced (connected) graphs \mathcal{G} with normalized signed Laplacian \mathcal{L} .

The condition $\pi_{1,d} > \pi_1$ holds if at π_1 the biggest eigenvalue of L_{π_1} is smaller than $\frac{2}{\varepsilon}$, that is, $\lambda_n(L_{\pi_1}) < \frac{2}{\varepsilon} = \lambda_n(L_{\pi_{1,d}})$. Since $L_{\pi_1} = \Delta(I - \pi_1(I - \mathcal{L}))$, from

Lemma 5 it follows that $\exists \theta \in [\min_i \delta_i, \max_i \delta_i]$ such that

$$\lambda_n(L_{\pi_1}) = \theta \lambda_n(I - \pi_1(I - \mathcal{L})) = \theta(1 - \pi_1(1 - \lambda_n(\mathcal{L}))).$$

Let $\varepsilon \max_i \delta_i < 1$. Observe that the smallest and largest eigenvalues of \mathcal{L} always satisfy $\lambda_1(\mathcal{L}) < 1 < \lambda_n(\mathcal{L}) \leq 2$ (see, e.g., [25]). Then $1 - \lambda_n(\mathcal{L}) < 0$ and

$$\begin{aligned} \lambda_n(L_{\pi_1}) &= \theta(1 + \pi_1(\lambda_n(\mathcal{L}) - 1)) \\ &\leq \max_i \delta_i(1 + \pi_1(\lambda_n(\mathcal{L}) - 1)) < \frac{1}{\varepsilon}(1 + \pi_1(\lambda_n(\mathcal{L}) - 1)). \end{aligned}$$

Therefore, showing that $\pi_1(\lambda_n(\mathcal{L}) - 1) \leq 1$ is sufficient to conclude that $\lambda_n(L_{\pi_1}) < \frac{2}{\varepsilon}$ and hence that $\pi_{1,d} > \pi_1$. We now need to treat the cases (i) and (ii) separately.

Assume that (i) holds, i.e., $\lambda_1(\mathcal{L}) = 0$. Then, since $\lambda_n(\mathcal{L}) \leq 2$, $\pi_1(\lambda_n(\mathcal{L}) - 1) = \lambda_n(\mathcal{L}) - 1 \leq 1$.

Assume that (ii) holds, i.e., $\lambda_1(\mathcal{L}) < 2 - \lambda_n(\mathcal{L})$. Then, $\pi_1(\lambda_n(\mathcal{L}) - 1) < \pi_1(1 - \lambda_1(\mathcal{L})) = 1$.

Hence, the condition $\pi_{1,d} > \pi_1$ holds. \square

G Proof of Lemma 3 and Lemma 4

Proof of Lemma 3. Assume, by contradiction, that $\pi \leq \pi_1$ which implies that L_π is positive semidefinite (positive definite if $\pi < \pi_1$), see Remark 3. Let $x^* \neq 0$ be an equilibrium point of (8), that is, $\Delta x^* = \pi A\psi(x^*) = \Delta\psi(x^*) - L_\pi\psi(x^*)$. Then,

$$0 < \psi(x^*)^T \Delta(x^* - \psi(x^*)) = -\psi(x^*)^T L_\pi \psi(x^*) \leq 0,$$

which leads to a contradiction. Hence, $\pi > \pi_1$. \square

Proof of Lemma 4. Assume, by contradiction, that $\pi \leq \pi_{1,d}$ which implies that $\varepsilon L_\pi - 2I$ is negative semidefinite (negative definite if $\pi < \pi_{1,d}$), see Remark 3. Assume that the system (8) admits a period-2 limit cycle: $\exists K > 0$ such that $x_{k+2} = x_k \neq 0$ for all $k \geq K$, that is

$$\begin{cases} x_k = (I - \varepsilon\Delta)x_{k+1} + \varepsilon\pi A\psi(x_{k+1}) \\ x_{k+1} = (I - \varepsilon\Delta)x_k + \varepsilon\pi A\psi(x_k), \end{cases}$$

which implies that

$$0 = (2I - \varepsilon\Delta)(x_{k+1} - x_k) + \varepsilon\pi A(\psi(x_{k+1}) - \psi(x_k)).$$

Then,

$$\begin{aligned} 0 &= (\psi(x_{k+1}) - \psi(x_k))^T (2I - \varepsilon\Delta)(x_{k+1} - x_k) \\ &\quad + \varepsilon\pi(\psi(x_{k+1}) - \psi(x_k))^T A(\psi(x_{k+1}) - \psi(x_k)) \\ &> (\psi(x_{k+1}) - \psi(x_k))^T (2I - \varepsilon\Delta + \varepsilon\pi A)(\psi(x_{k+1}) - \psi(x_k)) \\ &= (\psi(x_{k+1}) - \psi(x_k))^T (2I - \varepsilon L_\pi)(\psi(x_{k+1}) - \psi(x_k)) \geq 0 \end{aligned}$$

which leads to a contradiction since $\psi(x_{k+1}) - \psi(x_k) \neq 0$. Hence, $\pi > \pi_{1,d}$. Observe that the first inequality holds since $\varepsilon \max_i \delta_i < 2$ and each nonlinearity is monotonically increasing (i.e., if $x_i(k+1) - x_i(k) \geq 0$ then $\psi_i(x_i(k+1)) - \psi_i(x_i(k)) \geq 0$ for all i) and Lipschitz with constant 1 (i.e., $|x_i(k+1) - x_i(k)| \geq |\psi_i(x_i(k+1)) - \psi_i(x_i(k))|$ for all i). \square

H Proof of Theorem 4

To improve readability, the proof of Theorem 4 is divided as follows: in Section H.1 we prove (i) and in Section H.2 we prove (ii).

H.1 Proof of Theorem 4(i)

It is useful to first introduce the following two lemmas.

Lemma 7. Consider the function $f : [-1, 1] \rightarrow \mathbb{R}$ defined as

$$f(y) = \int_0^y \psi^{-1}(s) ds - \frac{1}{2}y^2.$$

where $\psi : \mathbb{R} \rightarrow [-1, 1]$ is a nonlinear function satisfying the properties (A.1)÷(A.4) and ψ^{-1} indicates the inverse function. Then, $f(y) > 0$ for all $y \neq 0$ and $f(0) = 0$.

Proof. Form the definition of f it follows that $f(0) = 0$ and $f'(y) = \psi^{-1}(y) - y$. Since $f'(0) = 0$, a sufficient condition for $f(y) > 0$ to hold for all $y \neq 0$ is that f is a convex function. To prove that f is convex we compute the second derivative:

$$f''(y) = \frac{1}{\psi'(\psi^{-1}(y))} - 1 \begin{cases} > 1 - 1 = 0, & y \neq 0 \\ = 0, & y = 0 \end{cases}$$

since $0 < \psi'(y) < 1 \forall y \neq 0$ and $\psi'(0) = 1$. It follows that f is convex, which implies that $f(y) \geq 0 \forall y \in [-1, 1]$. \square

Lemma 8 (Taylor expansion). Consider the function $g : [-1, 1] \rightarrow \mathbb{R}_+$ given by $g(y) = \int_0^y \psi^{-1}(s) ds$, where $\psi : \mathbb{R} \rightarrow [-1, 1]$ is a nonlinear function satisfying the properties (A.1)÷(A.4). Expanding g around y_0 yields

$$\int_0^y \psi^{-1}(s) ds = \int_0^{y_0} \psi^{-1}(s) ds + (y - y_0)\psi^{-1}(y_0) + \frac{(y - y_0)^2}{2}g''(z)$$

where $z \in [y_0, y]$ and $g''(z) := \frac{1}{\psi'(\psi^{-1}(z))} \geq 1$.

Now we are ready to prove Theorem 4(i); the proof follows the work [29].

Proof. Let $\psi_{k,i} := \psi_i(x_i(k))$, $i = 1, \dots, n$, and $\psi_k = [\psi_{k,1} \dots \psi_{k,n}]^T$. Then system (8) can be rewritten (using ψ_k instead of x_k as state variable) as:

$$\psi^{-1}(\psi_{k+1}) = (I - \epsilon\Delta)\psi^{-1}(\psi_k) + \epsilon\pi A\psi_k. \quad (33)$$

Let $V: [-1, 1]^n \rightarrow \mathbb{R}_+$ be the Lyapunov function described by

$$V(\psi_k) = -\frac{1}{2}\pi\varepsilon\psi_k^T A\psi_k + \varepsilon \sum_i \delta_i \int_0^{\psi_{k,i}} \psi_i^{-1}(s) ds.$$

Observe that from Lemma 7, $V(\psi_k) > 0$ for all $\psi_k \in [-1, 1]^n \setminus \{0\}$ and that $V(0) = 0$. Indeed,

$$V(\psi_k) \geq -\frac{1}{2}\pi\varepsilon\psi_k^T A\psi_k + \varepsilon \sum_i \frac{\delta_i}{2} \psi_{k,i}^2 = \frac{\varepsilon}{2} \psi_k^T (\Delta - \pi A) \psi_k = \frac{\varepsilon}{2} \psi_k^T L_\pi \psi_k \geq 0.$$

Let $\psi_\Delta = \psi_{k+1} - \psi_k$. Computing the increment of V along the trajectories gives

$$\begin{aligned} V_\Delta &= V(\psi_{k+1}) - V(\psi_k) \\ &= -\frac{\pi\varepsilon}{2} (\psi_{k+1}^T A \psi_{k+1} - \psi_k^T A \psi_k) + \varepsilon \sum_i \delta_i \left(\int_0^{\psi_{k+1,i}} \psi_i^{-1}(s) ds - \int_0^{\psi_{k,i}} \psi_i^{-1}(s) ds \right) \\ &= -\frac{1}{2} \psi_\Delta^T (\pi\varepsilon A) \psi_\Delta - \sum_i \psi_{\Delta,i} (\pi\varepsilon \sum_j a_{ij} \psi_{k,j}) \\ &\quad - \sum_i (1 - \varepsilon\delta_i) \int_{\psi_{k,i}}^{\psi_{k+1,i}} \psi_i^{-1}(s) ds + \sum_i \int_{\psi_{k,i}}^{\psi_{k+1,i}} \psi_i^{-1}(s) ds. \end{aligned} \tag{34}$$

From (33),

$$\sum_i \psi_{\Delta,i} (\pi\varepsilon \sum_j a_{ij} \psi_{k,j}) = \sum_i \psi_{\Delta,i} \psi_i^{-1}(\psi_{k+1,i}) - \sum_i \psi_{\Delta,i} (1 - \varepsilon\delta_i) \psi_i^{-1}(\psi_{k,i}).$$

Lemma 8 with $y = \psi_{k+1,i}$ and $y_0 = \psi_{k,i}$ yields

$$\int_{\psi_{k,i}}^{\psi_{k+1,i}} \psi_i^{-1}(s) ds = \psi_{\Delta,i} \psi_i^{-1}(\psi_{k,i}) + \frac{\psi_{\Delta,i}^2}{2} d_2(z_i), \tag{35}$$

where $d_2(z_i) := \frac{1}{\psi_i'(\psi_i^{-1}(z_i))} \geq 1$ and $z_i \in [\psi_{k,i}, \psi_{k+1,i}]$. While, with $y_0 = \psi_{k+1,i}$ and $y = \psi_{k,i}$, it yields

$$\int_{\psi_{k,i}}^{\psi_{k+1,i}} \psi_i^{-1}(s) ds = \psi_{\Delta,i} \psi_i^{-1}(\psi_{k+1,i}) - \frac{\psi_{\Delta,i}^2}{2} d_2(y_i) \tag{36}$$

where $d_2(y_i) := \frac{1}{\psi_i'(\psi_i^{-1}(y_i))} \geq 1$ and $y_i \in [\psi_{k,i}, \psi_{k+1,i}]$. Substituting (35), (35) and

(36) in (34), one obtains

$$\begin{aligned}
V_\Delta &= -\frac{1}{2}\psi_\Delta^T(\pi\varepsilon A)\psi_\Delta - \sum_i \psi_{\Delta,i}\psi_i^{-1}(\psi_{k+1,i}) + \sum_i (1-\varepsilon\delta_i)\psi_{\Delta,i}\psi_i^{-1}(\psi_{k,i}) \\
&\quad - \sum_i (1-\varepsilon\delta_i)\psi_{\Delta,i}\psi_i^{-1}(\psi_{k,i}) - \sum_i (1-\varepsilon\delta_i)\frac{\psi_{\Delta,i}^2}{2} d_2(z_i) \\
&\quad + \sum_i \psi_{\Delta,i}\psi_i^{-1}(\psi_{k+1,i}) - \sum_i \frac{\psi_{\Delta,i}^2}{2} d_2(y_i) \\
&\leq -\frac{1}{2}\psi_\Delta^T(\pi\varepsilon A)\psi_\Delta - \sum_i (2-\varepsilon\delta_i)\frac{\psi_{\Delta,i}^2}{2} \\
&= \frac{1}{2}\psi_\Delta^T(-2I + \varepsilon\Delta - \pi\varepsilon A)\psi_\Delta = \frac{1}{2}\psi_\Delta^T(-2I + \varepsilon L_\pi)\psi_\Delta.
\end{aligned}$$

The inequality holds since $1-\varepsilon\delta_i \geq 0$, $d_2(z_i) \geq 1$, and $d_2(y_i) \geq 1$ for all i . Under the assumption that $-2I + \varepsilon L_\pi$ is negative definite, we obtain $V_\Delta < 0$ if $\psi_\Delta \neq 0$. Therefore the trajectories of the system (8) converge asymptotically to a fixed equilibrium point which must be the origin (see Lemma 3). \square

H.2 Proof of Theorem 4(ii)

In this proof we follow [28, Chapter 5] and [40]. The system (8) can be rewritten as

$$x_{k+1} = J_\pi x_k + F(x_k) \quad (37)$$

where $J_\pi = I - \varepsilon L_\pi$ and $F(x_k) = \pi\varepsilon A(-x_k + \psi(x_k))$. The proof is divided into two steps. First, we assume that $\pi_1 < \pi_{1,d}$ and prove that the system (37) undergoes a pitchfork bifurcation at $\pi = \pi_1$. Then, we assume that $\pi_{1,d} < \pi_1$ and prove that the system (37) undergoes a period-doubling bifurcation at $\pi = \pi_{1,d}$.

Step 1. Assume that $\pi_1 < \pi_{1,d}$. At $\pi = \pi_1$, J_π has a simple eigenvalue at $+1$ and the corresponding eigenspace $\text{span}\{q\}$ has dimension one, where q is the eigenvector of J_{π_1} associated with 1 (i.e., $J_{\pi_1}q = q$) normalized such that $\|q\|_2 = 1$. Since J_{π_1} is symmetric, the left eigenvector of J_{π_1} is also q (i.e., $q^T J_{\pi_1} = q^T$). We can decompose any vector $x \in \mathbb{R}^n$ as $x = uq + y$ where $u = q^T x$ and $y = x - (q^T x)q \in (\text{span}\{q\})^\perp$. The system of equations (37) in the coordinates (u_k, y_k) can be written as follows

$$u_{k+1} = u_k + q^T F(u_k q + y_k) \quad (38a)$$

$$y_{k+1} = J_\pi y_k + F(u_k q + y_k) - (q^T F(u_k q + y_k))q. \quad (38b)$$

Center manifold theory (see [28, Chapter 5.4.2]) demonstrates that the restriction of (38) to the center manifold takes the form

$$u_{k+1} = u_k + b u_k^2 + c u_k^3 + O(u_k^4), \quad (39)$$

where, under the assumption (A.1), the parameters b and c in (39) simplify to

$$b := \frac{1}{2}q^T F_{yy}(0, \pi_1), \quad c := \frac{1}{6}q^T F_{yyy}(0, \pi_1).$$

Let $\frac{\partial^2 \psi}{\partial x^2}(x) := \text{diag}\{\frac{\partial^2 \psi_1}{\partial x_1^2}(x_1), \dots, \frac{\partial^2 \psi_n}{\partial x_n^2}(x_n)\}$ and $\frac{\partial^3 \psi}{\partial x^3}(x) := \text{diag}\{\frac{\partial^3 \psi_1}{\partial x_1^3}(x_1), \dots, \frac{\partial^3 \psi_n}{\partial x_n^3}(x_n)\}$. Then

$$b = \frac{1}{2} q^T F_{yy}(0, \pi_1) = \frac{\pi_1 \varepsilon}{2} q^T A \frac{\partial^2 \psi}{\partial x^2}(0) \begin{bmatrix} q_1^2 \\ \vdots \\ q_n^2 \end{bmatrix} = 0$$

since $\frac{\partial^2 \psi_i}{\partial x_i^2}(0) = 0$ for all i . Moreover,

$$\begin{aligned} c &= \frac{1}{6} q^T F_{yyy}(0, \pi_1) = \frac{\pi \varepsilon}{6} q^T A \frac{\partial^3 \psi}{\partial x^3}(0) \begin{bmatrix} q_1^3 \\ \vdots \\ q_n^3 \end{bmatrix} \\ &= \frac{\varepsilon}{6} q^T \Delta \frac{\partial^3 \psi}{\partial x^3}(0) \begin{bmatrix} q_1^3 \\ \vdots \\ q_n^3 \end{bmatrix} = \frac{\varepsilon}{6} \sum_{i=1}^n \delta_i \frac{\partial^3 \psi_i}{\partial x_i^3}(0) q_i^4 < 0, \end{aligned}$$

since $\frac{\partial^3 \psi_i}{\partial x_i^3}(0) < 0$ for all i . Hence, (39) can be rewritten as

$$u_{k+1} = u_k - |c| u_k^3 + O(u_k^4) \quad (40)$$

which means that at $\pi = \pi_1$ the system (8) undergoes a pitchfork bifurcation.

Step 2. Assume that $\pi_1 > \pi_{1,d}$. At $\pi = \pi_{1,d}$, J_π has a simple eigenvalue at -1 and the corresponding eigenspace $\text{span}\{q\}$ has dimension one, where q is the eigenvector of $J_{\pi_{1,d}}$ associated with -1 (i.e., $J_{\pi_{1,d}} q = q$) normalized such that $\|q\|_2 = 1$. We can decompose any vector $x \in \mathbb{R}^n$ as $x = uq + y$ where $u = q^T x$ and $y = x - (q^T x)q \in (\text{span}\{q\})^\perp$. The system of equations (37) in the coordinates (u_k, y_k) can be written as follows

$$u_{k+1} = -u_k + q^T F(u_k q + y_k) \quad (41a)$$

$$y_{k+1} = J_\pi y_k + F(u_k q + y_k) - (q^T F(u_k q + y_k)) q. \quad (41b)$$

Center manifold theory (see [28, Chapter 5.4.2]) demonstrates that the restriction of (38) to the center manifold takes the form

$$u_{k+1} = -u_k + b u_k^2 + c u_k^3 + O(u_k^4), \quad (42)$$

where, as in Step 1, since each $\psi_i(\cdot)$ has odd symmetry (hence $\frac{\partial^2 \psi_i}{\partial x_i^2}(0) = 0$ for all i), the parameters b and c in (42) are given by

$$b = q^T F_{yy}(0, \pi_{1,d}) = 0, \quad c = \frac{1}{6} q^T F_{yyy}(0, \pi_{1,d}) > 0.$$

Hence, (42) can be rewritten as

$$u_{k+1} = -u_k + |c| u_k^3 + O(u_k^4) \quad (43)$$

which means that at $\pi = \pi_{1,d}$ a cycle of period 2 bifurcates from the origin. \square

Bibliography

- [1] A. Fontan and C. Altafini, “The role of frustration in collective decision-making dynamical processes on multiagent signed networks,” *arXiv:2105.11396*, pp. 1–18, may 2021.
- [2] A. Fontan and C. Altafini, “Achieving a decision in antagonistic multi agent networks: frustration determines commitment strength,” in *57th IEEE Conference on Decision and Control*. Miami Beach, FL, USA: IEEE, dec 2018, pp. 109–114.
- [3] R. Gray, A. Franci, V. Srivastava, and N. E. Leonard, “Multiagent Decision-Making Dynamics Inspired by Honeybees,” *IEEE Transactions on Control of Network Systems*, vol. 5, no. 2, pp. 793–806, jun 2018.
- [4] D. T. Swain, I. D. Couzin, and N. Ehrich Leonard, “Real-Time Feedback-Controlled Robotic Fish for Behavioral Experiments With Fish Schools,” *Proceedings of the IEEE*, vol. 100, no. 1, pp. 150–163, jan 2012.
- [5] W. Ren and R. W. Beard, *Distributed Consensus in Multi-vehicle Cooperative Control*, ser. Communications and Control Engineering. London: Springer London, 2008.
- [6] W. Ren and Y. Cao, *Distributed Coordination of Multi-agent Networks*, ser. Communications and Control Engineering. London: Springer London, 2011.
- [7] R. Hegselmann and U. Krause, “Opinion Dynamics and Bounded Confidence,” *Journal of Artificial Societies and Social Simulation*, vol. 5, no. 3, p. 2, 2002.
- [8] P. Jia, A. MirTabatabaei, N. E. Friedkin, and F. Bullo, “Opinion Dynamics and the Evolution of Social Power in Influence Networks,” *SIAM Review*, vol. 57, no. 3, pp. 367–397, 2015.
- [9] S. Wasserman and K. Faust, *Social Network Analysis: Methods and applications*, ser. Structural Analysis in the Social Sciences. Cambridge University Press, 1994.
- [10] D. Easley and J. Kleinberg, *Networks, Crowds, and Markets: Reasoning about a Highly Connected World*. Cambridge: Cambridge University Press, 2010.
- [11] F. Harary, “On the notion of balance of a signed graph.” *The Michigan Mathematical Journal*, vol. 2, no. 2, pp. 143–146, 1953.
- [12] T. Zaslavsky, “Signed graphs,” *Discrete Applied Mathematics*, vol. 4, no. 1, pp. 47–74, 1982.
- [13] D. Cartwright and F. Harary, “Structural balance: a generalization of Heider’s theory.” *Psychological Review*, vol. 63, no. 5, pp. 277–293, 1956.

- [14] C. Altafini, “Consensus Problems on Networks With Antagonistic Interactions,” *IEEE Transactions on Automatic Control*, vol. 58, no. 4, pp. 935–946, apr 2013.
- [15] G. Facchetti, G. Iacono, and C. Altafini, “Computing global structural balance in large-scale signed social networks,” *Proceedings of the National Academy of Sciences*, vol. 108, no. 52, pp. 20 953–20 958, 2011.
- [16] J. Kunegis, S. Schmidt, A. Lommatzsch, J. Lerner, E. W. De Luca, and S. Albayrak, “Spectral Analysis of Signed Graphs for Clustering, Prediction and Visualization,” in *2010 SIAM International Conference on Data Mining*. Philadelphia, PA: Society for Industrial and Applied Mathematics, apr 2010, pp. 559–570.
- [17] J. Kunegis, “Applications of Structural Balance in Signed Social Networks,” *arXiv:1402.6865v1*, pp. 1–37, 2014.
- [18] A. Fontan and C. Altafini, “A signed network perspective on the government formation process in parliamentary democracies,” *Scientific Reports*, vol. 11, no. 5134, dec 2021.
- [19] P. U. Abara, F. Ticozzi, and C. Altafini, “Spectral Conditions for Stability and Stabilization of Positive Equilibria for a Class of Nonlinear Cooperative Systems,” *IEEE Transactions on Automatic Control*, vol. 63, no. 2, pp. 402–417, feb 2018.
- [20] A. Fontan and C. Altafini, “Multiequilibria Analysis for a Class of Collective Decision-Making Networked Systems,” *IEEE Transactions on Control of Network Systems*, vol. 5, no. 4, pp. 1931–1940, dec 2018.
- [21] H. L. Smith, “Systems of Ordinary Differential Equations Which Generate an Order Preserving Flow. A Survey of Results,” *SIAM Review*, vol. 30, no. 1, pp. 87–113, mar 1988.
- [22] F. Harary, “On the measurement of structural balance,” *Behavioral Science*, vol. 4, no. 4, pp. 316–323, jan 1959.
- [23] J. J. Hopfield, “Neurons with graded response have collective computational properties like those of two-state neurons.” *Proceedings of the National Academy of Sciences*, vol. 81, no. 10, pp. 3088–3092, 1984.
- [24] H. Zhang, Z. Wang, and D. Liu, “A Comprehensive Review of Stability Analysis of Continuous-Time Recurrent Neural Networks,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 25, no. 7, pp. 1229–1262, jul 2014.
- [25] H.-H. Li and J.-S. Li, “Note on the normalized Laplacian eigenvalues of signed graphs,” *The Australasian Journal of Combinatorics*, vol. 44, pp. 153–162, 2009.

- [26] Y. Hou, J. Li, and Y. Pan, "On the Laplacian Eigenvalues of Signed Graphs," *Linear and Multilinear Algebra*, vol. 51, no. 1, pp. 21–30, 2003.
- [27] R. A. Horn and C. R. Johnson, *Matrix analysis*, 2nd ed. Cambridge University Press, 2013.
- [28] Y. A. Kuznetsov, *Elements of Applied Bifurcation Theory, Second Edition*. Springer, 1998.
- [29] W. Zhao, W. Lin, R. Liu, and J. Ruan, "Asymptotical Stability in Discrete-Time Neural Networks," *IEEE Transactions on Circuits and Systems*, vol. 49, no. 10, pp. 1516–1520, 2002.
- [30] D. Acemoğlu, G. Como, F. Fagnani, and A. Ozdaglar, "Opinion Fluctuations and Disagreement in Social Networks," *Mathematics of Operations Research*, vol. 38, no. 1, pp. 1–27, feb 2013.
- [31] P. Cisneros-Velarde, K. Chan, and F. Bullo, "Polarization and Fluctuations in Signed Social Networks," *IEEE Transactions on Automatic Control*, 2020.
- [32] M. Golubitsky and I. Stewart, *The Symmetry Perspective*, ser. Progress in Mathematics (Volume 200), H. Bass, J. Oesterle, and A. Weinstein, Eds. Basel: Birkhäuser Basel, 2002.
- [33] M. Golubitsky, I. Stewart, and D. Schaeffer, *Singularities and Groups in Bifurcation Theory: Volume II*, ser. Applied Mathematical Sciences 69. Springer-Verlag, 1988.
- [34] G. Iacono, F. Ramezani, N. Soranzo, and C. Altafini, "Determining the distance to monotonicity of a biological network: a graph-theoretical approach," *IET Systems Biology*, vol. 4, no. 3, pp. 223–235, may 2010.
- [35] E. Kaszkurewicz and A. Bhaya, *Matrix Diagonal Stability in Systems and Computation*. Boston: Birkhäuser, 2000.
- [36] S. Camiz and S. Stefani, *Matrices and Graphs*. World Scientific, 1996.
- [37] M. Golubitsky and D. G. Schaeffer, *Singularities and Groups in Bifurcation Theory. Volume I*, ser. Applied Mathematical Sciences. Springer New York, 1985, vol. 51.
- [38] M. Fiedler, "Elliptic matrices with zero diagonal," *Linear Algebra and Its Applications*, vol. 197-198, no. C, pp. 337–347, 1994.
- [39] J. J. Hopfield and D. W. Tank, ""Neural" Computation of Decisions in Optimization Problems," *Biological Cybernetics*, vol. 52, no. 3, pp. 141–152, 1985.
- [40] Y. A. Kuznetsov and H. G. Meijer, "Numerical normal forms for codim 2 bifurcations of fixed points with at most two critical eigenvalues," *SIAM Journal on Scientific Computing*, vol. 26, no. 6, pp. 1932–1954, 2005.

Paper D

A signed network perspective on the government formation process in parliamentary democracies

Authors: Angela Fontan and Claudio Altafini

Edited version of the paper:

A. Fontan and C. Altafini, “A signed network perspective on the government formation process in parliamentary democracies,” *Scientific Reports*, vol. 11, no. 5134, dec 2021.

A signed network perspective on the government formation process in parliamentary democracies

Angela Fontan^{*} and Claudio Altafini^{*}

^{*}Dept. of Electrical Engineering
Linköping University
SE-581 83 Linköping, Sweden
angela.fontan@liu.se, claudio.altafini@liu.se

Abstract

In parliamentary democracies, government negotiations talks following a general election can sometimes be a long and laborious process. In order to explain this phenomenon, in this paper we use structural balance theory to represent a multiparty parliament as a signed network, with edge signs representing alliances and rivalries among parties. We show that the notion of frustration, which quantifies the amount of “disorder” encoded in the signed graph, correlates very well with the duration of the government negotiation talks. For the 29 European countries considered in this study, the average correlation between frustration and government negotiation talks ranges between 0.42 to 0.69, depending on what information is included in the edges of the signed network. Dynamical models of collective decision-making over signed networks with varying frustration are proposed to explain this correlation.

1 Introduction

In a country adopting a multiparty parliamentary system, the process of forming a government after a general election can be quick, or can sometimes be impervious and characterized by lengthy negotiations. Indeed, especially in the last decade, many European countries have experienced unusually long cabinet bargaining phases, for instance Belgium in 2010 and 2019, Bosnia Herzegovina in 2010 and 2018, Croatia in 2015, Czech Republic in 2006, Germany in 2017, Ireland in 2020, Italy in 2018, North Macedonia in 2016, Netherlands in 2017, Norway in 2017, Spain in 2016 (and 2015), Sweden in 2018 and UK in 2018. Typically, long delays appear when a parliament is fragmented and a clear majority is missing: parties have to enter multiple bargainings with each other in order to form a coalition able to win a parliamentary confidence vote.

To model this government formation process, descriptive statistics and game-theoretical models of bargaining have often been used in the literature, see [2–10] and references therein. These models use the data to fit coefficients expressing

the relative importance of certain factors, like number of parties and Members of Parliament (MP), seats per party, ideological polarization, etc., in predicting the cabinet formation times.

One of the important determinants of a government formation timeline is the fragmentation of a legislature, measured for instance in terms of the Laakso-Taagepera effective number of parties [11] or of the related fractionalization index [12]. This and other factors, such as the ideological diversity of a legislature, the idea of minimum winning coalition [3], etc., can be used in descriptive statistical models to evaluate how different features influence the government formation process [5, 7, 9]. For instance, the Cox hazard model of [5] groups factors into “bargaining complexity” (n. of parties and seats per party, ideological polarization, requirements for the investiture, type of parliamentarism – “negative” or “positive”, see [2]) and “uncertainty” of the political landscape (previous history of collaborations or defeats among the parties, whether the government is a post-electoral or an inter-electoral one, etc.) and uses empirical data to estimate the importance of the different terms.

The literature on bargaining models applied to coalition formation is vast, see for instance [8, 10, 13–17] (see also [18] for an overview). A typical bargaining procedure is organized in two stages: first a player (often the formateur) proposes a coalition of parties and a certain allocation of government portfolios among them, then the members of the candidate coalition can decide to either accept or refuse the proposal. A government is formed only if all the candidate parties agree on the proposal, otherwise the bargaining process continues until a government is formed. This model is used for instance in [19] to perform an empirical analysis in 11 parliamentary democracies, using electoral data from 1945–1997. More complex bargaining procedures are discussed e.g. in [10].

Departing from the aforementioned literature, this paper aims to propose a novel approach to model the process of government formation, rooted in the social network theory of multiagent systems with antagonistic interactions [20, 21]. In particular we represent antagonism among parties as a signed graph, and consider the process of government formation as a collective decision-making over such signed network.

More specifically, our approach relies on the concept of structural balance [22–26] and of graph frustration, intended as distance from a structurally balanced situation. The idea of structural balance is well-known in social network theory [20, 21], and easily exemplified by notions like “the enemy of my enemy is my friend”: when a signed graph can be partitioned into two subgroups such that all nodes in each subgroup are connected by positive (i.e., “friendly”) edges and all edges across the two subgroups are instead negative (i.e., “unfriendly”) then the graph is structurally balanced and shows no frustration. A two-party parliament (e.g., Malta, if we restrict to European nations) is structurally balanced: all cycles on the signed graph are positive (i.e., have an even number of negative edges). For a structurally balanced parliament, the government formation process is typically straightforward: the party winning the elections is in charge of forming a cabinet. Signed graphs that are not structurally balanced have instead some amount of negative cycles (i.e., cycles with an odd number of negative edges). The notion

of frustration can be used to quantify the distance to structural balance induced by these negative cycles [26]. This notion, adopted from the statistical physics literature, is expressed as the energy in the ground state of an “Ising spin glass” [26–28] and quantifies the amount of “disorder” contained in the system.

Our hypothesis is that in parliamentary networks this disorder strongly influences the process of government formation. Namely, the higher the amount of frustration, the longer the government negotiation phase is expected to be. Indeed if there is no clear winner after the elections (i.e., no political party or alliance managed to secure a majority in the parliament) then the parties have to negotiate and overcome their ideological differences in order to form a coalition cabinet backed by the majority in the parliament. The data we collected from 29 European nations over the last 30–40 years show that indeed the frustration of the parliamentary networks correlates well with the duration of the government negotiations talks. This correlation is due to a large extent to the fragmentation of a legislature. In fact, interestingly enough, we show that in the simplest scenario we consider (referred to as scenario **I** below) frustration combines together the notions of fractionalization index and minimum winning coalition.

The correlation between frustration and duration of the government negotiations can be explained also at a deeper level, using dynamical models of collective decision-making in multiagent systems [29–31]. Such models represent a decision as the trespassing of a bifurcation threshold, and the corresponding bifurcation parameter has the interpretation of “social commitment” of the agents, i.e., intensity of the interactions among the agents. In presence of signed graphs, it is known that the bifurcation threshold can be pushed to higher values of the bifurcation parameter, and in particular that its value is proportional to the frustration encoded in the signed network [32]. If we consider as decision a confidence vote on a candidate cabinet, and as “social commitment” the intensity of the government negotiation talks (quantified as duration of the bargaining phase), then in a multiagent dynamics perspective the positive correlation we find between frustration and duration of the government negotiation phase is an expected and reasonable property: more “disordered” parliaments require longer negotiations to form a government. By capturing and quantifying this disorder, the frustration of a parliamentary network allows to predict the government formation timing and also, to some extent, the composition of parties that managed to form a successful cabinet coalition.

2 Results

2.1 Correlation between frustration and duration of the government negotiation talks

In the 29 European parliamentary democracies or constitutional monarchies considered in this study, see Table 1 (and Table 2 for more details), we computed the frustration associated to the parliamentary network resulting after each general election (see Section 4). Its value depends on what information is encoded in the weights of the parliamentary network: three different scenarios (denoted **I** to **III**,

Country	System of Government (PR or PCM)	Pre-electoral coalitions ✓ = yes or ✗ = no	N	(from-to)
Albania	PR	✓	8	(1992-2017)
Andorra	PCM	✓	8	(1993-2019)
Austria	PR	✗	13	(1979-2019)
Belgium	PCM	✓	7	(1995-2019)
Bosnia Herzegovina	PR	✗	8	(1996-2018)
Bulgaria	PR	✗	9	(1991-2017)
Croatia	PR	✓	9	(1992-2020)
Czech Republic	PR	✓	8	(1992-2017)
Denmark	PCM	✗	13	(1981-2019)
Estonia	PR	✗	8	(1992-2019)
Finland	PR	✓	8	(1991-2019)
Germany	PR	✓	8	(1990-2017)
Greece	PR	✗	11	(1990-2019)
Hungary	PR	✓	6	(1990-2010)
Iceland	PR	✗	8	(1995-2017)
Ireland	PR	✓	7	(1992-2020)
Italy	PR	✓	8	(1992-2018)
Latvia	PR	✗	9	(1993-2018)
Luxembourg	PCM	✗	8	(1984-2018)
North Macedonia	PR	✓	10	(1990-2020)
Moldova	PR	✗	8	(1994-2019)
Netherlands	PCM	✗	12	(1981-2017)
Norway	PCM	✓	10	(1981-2017)
Serbia	PR	✗	6	(2007-2020)
Slovakia	PR	✓	9	(1990-2020)
Slovenia	PR	✗	8	(1992-2018)
Spain	PCM	✓	9	(1989-2019)
Sweden	PCM	✓	11	(1982-2018)
United Kingdom	PCM	✓	10	(1983-2019)

Table 1: List of countries considered in this study. For each country the following are shown: the system of government (PR = Parliamentary Republic, PCM = Parliamentary Constitutional Monarchy), the existence of pre-electoral coalitions, the number of general elections considered and their time span.

see Fig. 1a) are considered in this study, from no a-priori information at all on the parties and their relationships, to ideological and electoral coalition information. In particular, in scenario **I** a parliamentary network is built using information on (1): n. of parties of a legislature, and (2): n. of MPs per party. In scenarios **II** and **III** also (3): pre-electoral party coalitions, and (4): ideological polarization of the parties are considered in forming the adjacency matrices of the parliamentary networks. More details on these four factors are provided below, in the Section 4 (see also Fig. 1b and Fig. 8).

Even for the most basic scenario (**I**: the parties are “all-against-all”, which leads to an unweighted fully-connected signed graph with on-diagonal blocks of +1 and off-diagonal blocks of -1, see Fig. 1b), our calculations show that the frustration of a parliamentary network is indeed a good indicator of the complexity of a post-electoral government formation process: the average correlation (defined by the Pearson correlation coefficient) between the frustration (ζ , computed according to

SCENARIO	I	II	III
Party grouping	All-against-all	Pre-electoral coalitions	Pre-electoral coalitions
Weight selection	Unweighted	RILE	(Optimized) Left-Right grid

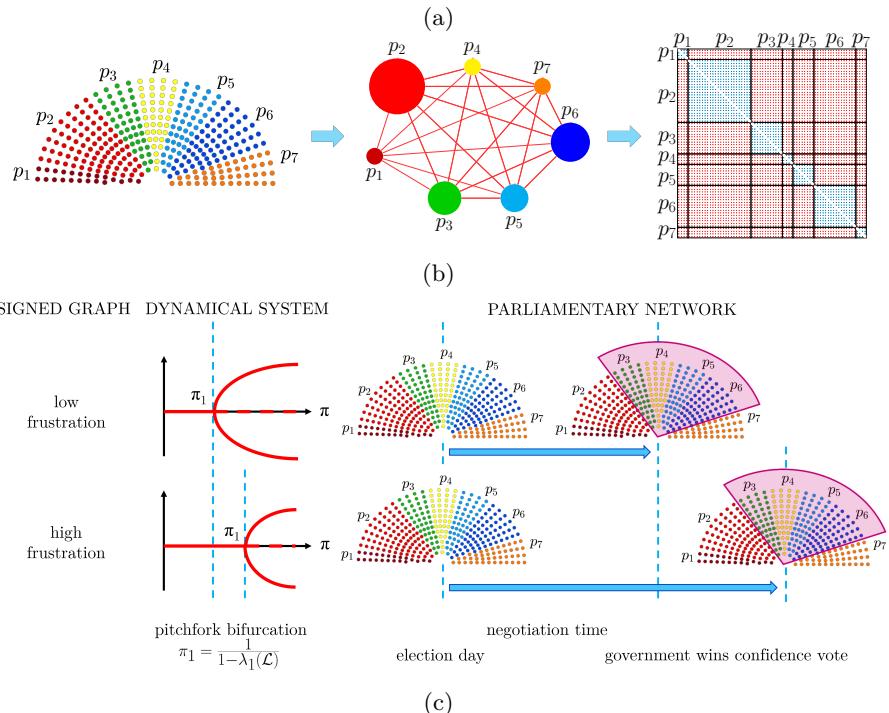


Figure 1: (a): The three different scenarios considered in this study. (b): Constructing a parliamentary network and the corresponding adjacency matrix in scenario I (all parties against all parties and weights equal to +1, in blue, and -1, in red). Each node p_i represents a political party. See Fig. 8a for scenario II and III. (c): Process of government formation described as a dynamical model of decision-making over signed networks: the higher is the frustration of the network, the higher is the value of the bifurcation point $\pi_1 = \frac{1}{1-\lambda_1(\mathcal{L})}$ ($\lambda_1(\mathcal{L})$ = least eigenvalue of the normalized Laplacian of the signed graph). In turn, the higher is π_1 , the longer is the expected government negotiation time between the parties.

the formula (2), see Fig. 10b) and the length of the government negotiation phase (computed as number of days between the general election and the day the new government is sworn in, see Fig. 10a) is 0.42, see Fig. 2a. In this scenario, what is modeled is essentially only the fragmentation of a legislature, i.e., n. of parties and n. of MPs per party, and in fact in this case the frustration is nearly identical to the fractionalization index F (correlation is 0.99, see Fig. 2d) with a very minor correction needed only when in a country no coalition exists able to achieve exactly

50% +1 of the seats. An analytical formula relating these quantities is provided in Section 4.4, see also Section 3 for more details.

Given the many unmodeled factors constraining and influencing the cabinet bargaining process (see e.g. [33, Chapter 9] for an overview) we find it remarkable that this level of correlation can be achieved by such simple models, especially since a clear reason behind the lack of correlation exists for several of the countries showing the worst fit (e.g., Czech Republic, Estonia and Greece), see Section 3.

More complex adjacency matrices for our parliamentary networks can be formed if we include information such as pre-electoral party coalitions, if available (explicit or implicit, depending on the country, see Table 1), and ideological classification of parties. When for the latter factor we use standard indexes such as *rile* (based on the electoral manifestos of the parties, see [33–35]) then the correlation between network frustration and negotiation days is slightly lower, with an average value of 0.32 (scenario **II** in Fig. 2a). When instead the edge weights are optimized based on a predetermined left-right grid (see Fig. 8c and Section A.1 for the details) then the correlation grows up to 0.69 in average (scenario **III** in Fig. 2a). In both scenarios **II** and **III**, frustration and fractionalization index no longer coincide, see Fig. 2d. Because of the weight optimization, scenario **III** cannot be considered parameter-free, unlike scenarios **I** and **II**. To motivate (and validate) the use of this scenario we have performed a “leave-one-out” analysis, where the optimization of the political positions of the parties in the left-right scale is performed only on $N - 1$ elections (denoted “training set”, where N is the number of elections for each country, see Section 4.8 for more details) and the remaining election (denoted “validation set”) is used to evaluate how well the model fits the data (in terms of correlation). As Fig. 4 shows, scenario **III** is in general performing well in the leave-one-out analysis. A special case this analysis, when the excluded election is the last one, can be interpreted as the capacity of the model to predict the length of a future negotiation process based only on the frustration of a newly elected parliament. Also in this case the predictive power of the model is in most cases reasonably good, see Fig. 17 and Section 3 for more details.

2.2 Prediction of the successful cabinet coalition

The notion of frustration of a parliamentary network can be used also to predict the successful cabinet composition ensuing from the negotiations. If in our signed graphs we assign a “spin” variable (polarized into ± 1 values, corresponding to “spin up” and “spin down”, and here interpretable as yes / no in a confidence vote) to all parties, then it is possible to compute an energy-like function for each spin configuration, as well as for the party coalition that succeeds in forming a post-election government. The frustration then corresponds to the global minimum of this energy functional (i.e., the energy in the ground state, denoted S_{best}). The index ρ_{gov} described in Section 4.5 represents the overlap between the party coalition that succeeded in forming a government, S_{gov} , and the majoritarian group of parties in the ground state S_{best} . In our data, ρ_{gov} varies between a 68% of scenario **I** and a 79% of scenario **III**, see Fig. 2b. It is worth remarking that for several countries (e.g., Czech Republic, Denmark, Norway, Sweden) the overlap

is reduced by the fact that minority governments are supported (actively or passively, e.g., through abstentions in confidence votes) by other parties which do not figure explicitly in the cabinet coalition (and hence do not appear in the spin configuration S_{gov} we use to quantify ρ_{gov}). The lower value of ρ_{gov} for scenario **I** is also due to the parliamentary network being an unweighted matrix, which leads to many states of near-identical energy, see Fig. 11. For weighted adjacency matrices (scenario **II** and **III**) such degenerate cases are much less frequent and in fact ρ_{gov} increases. Indeed, if we look at the energy gap between ground state and “government state” (i.e., S_{best} and S_{gov} , see Section 4.5), expressed by the index η_{gov} described in Section 4.5, then it is $\eta_{\text{gov}} > 90\%$ in scenario **I**, even better than for the remaining scenarios ($\eta_{\text{gov}} = 79.2\%$ in scenario **II** and $\eta_{\text{gov}} = 79.7\%$ in scenario **III**, see Fig. 2c), which confirms that even when non-optimal, the successful government coalition is always “energetically” close enough to the ground state coalition S_{best} identified by our method.

2.3 Interpretation as collective decision dynamics

The high correlation between frustration and government negotiation days admits an interpretation in terms of dynamical models of collective decision-making, see Section 4.6, Section 4.7, and Section A.3. These models, used for instance in [31, 32], are inspired by the behavior of animal groups [29, 30, 36] and in our case associate a state variable to each MP, variable that represents the decision in a confidence vote. In order for a government to win a confidence vote, an attractor corresponding to a majority of positive decisions must be present. The government formation process is represented in this model by means of a bifurcation occurring for a certain value of a scalar parameter π which we can refer to as strength of the “social commitment” among the parties, here intended as a proxy for the duration of the cabinet negotiations process. Increasing π , the system passes from having the origin as the only globally asymptotically stable equilibrium to a situation in which two extra nonzero states of decision, i.e., two locally asymptotically stable equilibria, are present, while the origin becomes unstable, see Fig. 1c. In the model, the origin represents a state of “no decision” (the MPs do not take any side), while the two stable equilibria appearing after the crossing of the bifurcation point correspond to success and failure of a confidence vote. Since the bifurcation is of pitchfork type and the network is symmetric, these two equilibria are identical up to a change of sign, and they represent a partition of the parliament into two factions, the majoritarian one being the winner of the confidence vote. The bifurcation point is a function of the least eigenvalue of the normalized Laplacian of the network, $\lambda_1(\mathcal{L})$, see (8). The latter is called in the signed graph literature the *algebraic conflict* [25], and it is well-known to be closely related to the frustration [25, 32, 37], see Table 4 for an analysis on our data. Similarly, the signature of the corresponding eigenvector overlaps substantially with the signature of the ground state spin configuration S_{best} , see Table 4. Combining the proportionality between $\lambda_1(\mathcal{L})$ and frustration with (8), we get a relationship between the value of π at the bifurcation point, π_1 , and the frustration of the network. Interpreted in the context of a government formation process, this relationship states that the duration of



Figure 2: (a): Correlation between frustration and duration of the government negotiation talks, quantified in number of days from the general election to the date the government is sworn in or to the date the government negotiations fail, for each country and scenarios **I**, **II** and **III**. Blue bars consider only the cases of success, yellow bars consider both successes and failures. (b): Index ρ_{gov} (overlap between party coalition entering in the government and party majority in the ground state). (c): Index η_{gov} (energy gap between ground state and “government state”). (d): Correlation between frustration and fractionalization index ($r_{\zeta,F}$), for each country and scenarios **I**, **II** and **III**. The average values (of correlation, in (a) and (d), or of the indexes, in (b) and (c)) are reported inside each plot.

the government negotiation talks (a proxy of π_1) is directly proportional to the frustration of the parliamentary network, see Section 4.7. When the frustration increases, the model predicts that the bifurcation threshold increases as well (see [32] and also Fig. 9), meaning that a higher commitment will be required from the agents in order to escape a state of no decision and to reach a collective nontrivial equilibrium point. This translates in our model into longer negotiation times for the government formation phase. The concept is illustrated in Fig. 1c.

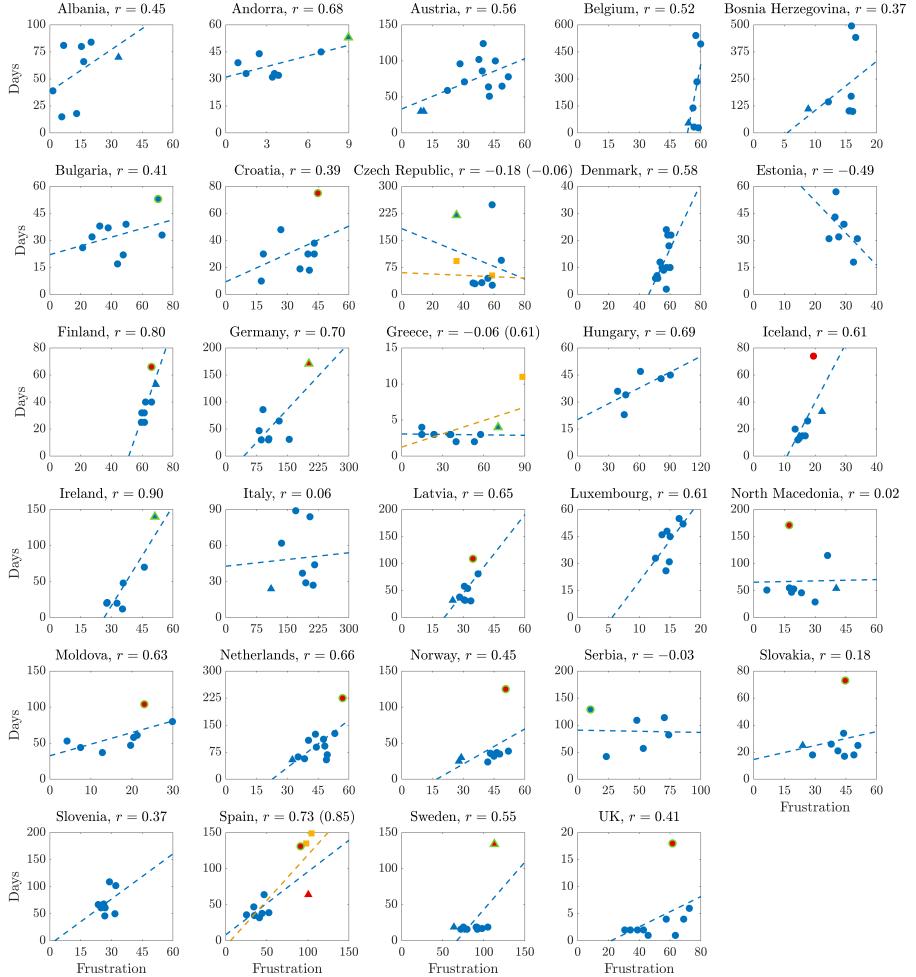


Figure 3: Scenario I. Frustration of the parliamentary networks v.s. duration of the government negotiation talks (days) and corresponding linear regression line, for all countries of Table 1. The value of Pearson correlation (r) for each country is reported in the plot heading. Legend: blue circles represent points that are neither outliers, nor high leverage nor influential. A red symbol indicates an outlier, a triangle a high leverage point and a symbol with green outline an influential point. Residual analysis, leverage statistic and delete-1 statistics are used to identify outliers, high leverage and influential points, respectively. Yellow square data points indicate elections corresponding to failure of government negotiations resulting in votes of no-confidence (Czech Republic in 2006 and 2017) and new elections (Spain in December 2015 and April 2019, Greece in May 2012). Blue regression lines include only the successful government formations. Including also the failure points we obtain the yellow regression lines. In all 3 cases, the correlation increases (values are reported between parentheses).

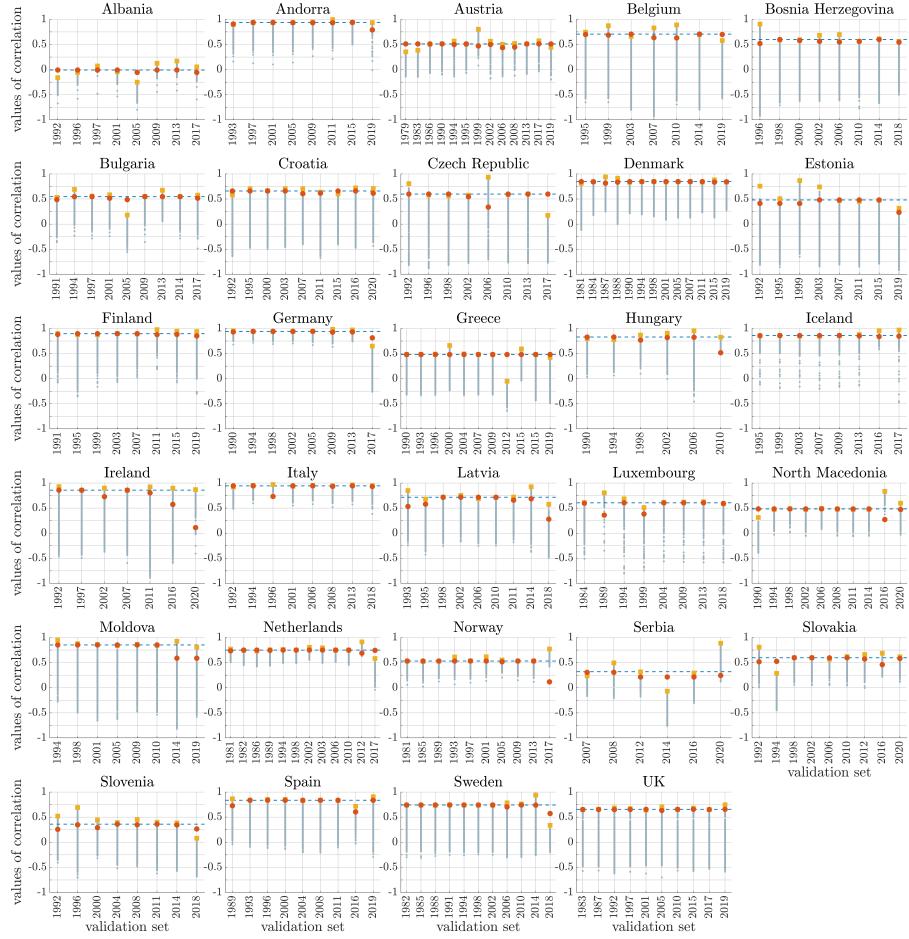


Figure 4: Leave-one-out analysis of the correlation for scenario **III**. A single election is used for the validation set, and the remaining elections (training set) are used to calculate the correlation (frustration v.s. duration of the government negotiation talks) for each choice of political positions in the left-right grid. 1000 sets of values for the left-right political positions of the parties were randomly selected on the preassigned left-right grid. Here the corresponding correlation values for each set are shown (gray dots), together with the overall country maximum value (yellow square) corresponding to the optimal choice for the weights. Red circles represent the values of correlation obtained when the excluded election is considered (the frustration is calculated using the optimal choice for the weights found using the training set). A blue dashed line represents the value of correlation obtained if the weights are tuned so as to maximize the correlation when all the elections are considered (i.e., scenario **III**). The red circles are normally very close to the yellow squares, meaning that the leave-one-out test is normally accurate.

3 Discussion

A cornerstone of parliamentary democracies is plurality of opinions, which often manifests itself as antagonism between parties of different factions. This is the rationale behind our use of the formalism of signed graphs for describing parliamentary networks and their collective decision processes. In particular, in the paper we investigate how our knowledge over signed parliamentary networks can be used in order to predict (i) the duration of the government negotiation talks and (ii) the composition of the successful coalition cabinet formed after the elections. We show that the frustration of the parliamentary networks can be considered a good proxy for the complexity of the negotiations among the parties and, in particular, that it correlates well with the duration of the government negotiation talks. Moreover, we show that it can be used to obtain an estimate of the post-election government composition.

In the literature on political networks, signed networks have been used for instance to investigate coalition formation from policy beliefs [38–40], to analyze the influence of political polarization on cabinet stability and duration [41], or for inferring political polarization from voting records [42]. The approach we take in this paper is different as we use signed networks to investigate the government negotiation process, in particular the one that immediately follows a general election. It is known in the literature [5, 7] that the dynamics of such post-electoral government formation is somewhat simpler to analyze than that of inter-election governments. In fact, the antagonism among parties likely tends to peak during the pre-electoral period, hence it is natural that also the early phase of a legislation keeps reflecting the pattern of alliances and rivalries that characterized an electoral ballot. Furthermore, it is also plausible to assume that in early post-election decisions, MPs tend to follow thoroughly the party lines, and therefore that parties behave as homogeneous entities, as we are doing here. Modeling inter-election government processes may require to take into account the behavior of individual MPs and their history of votes, which requires to collect different datasets, difficult to obtain and even more difficult to analyze [42].

If we zoom on the scatter plots of frustration vs negotiation days of the individual nations (see Fig. 3 for scenario I, Fig. 12 and Fig. 13 for scenarios II and III), we see that in several cases what determines the high correlation is a small fraction of the data, corresponding to elections with a “hung parliament”. In our framework, rather than being considered spurious points, these data carry a strong informative value, as they correspond to parliamentary networks having a frustration higher than usual for that nation (in Fig. 3: a right outlier on the horizontal axis, according to a leverage statistic test). In the vast majority of cases, they happen to correspond to long cabinet negotiation times, i.e., to outliers from above also in the vertical axis (residual analysis test), see for example the plots for Germany, Ireland, Spain and Sweden in Fig. 3.

The Czech Republic seems to be the only exception to this rule, with a high value of negotiation days for the 2006 election corresponding to the least frustration. As a matter of fact, this point represents a special situation that does not violate the general rule proposed in the paper. In fact, the 2006 Czech election

saw a parliament split into two exact halves, which also corresponds to our S_{best} having 50% of +1 and 50% of -1. Since a tie is not a majority, its frustration may fail to capture the true complexity of the negotiation, if the two halves correspond also to opposite ideological factions, as in the Czech case. More precise comments on this example are reported in Section B.8.

In the Czech Republic constitution, a cabinet is first sworn in and then confirmed by a confidence vote within 30 days. In two occasions (after the 2006 and 2017 elections) the candidate cabinets (Topolánek I and Babiš I) failed to obtain the confidence of the Parliament and new rounds of government negotiations started, after which the new candidate cabinets (Topolánek II and Babiš II) were sworn in. In our model, the 2006 and 2017 Czech elections give rise to 2 points, one corresponding to failure and the other to success, see Fig. 3 and Fig. 14a.

The Czech unsuccessful attempts are among the few cases of “failure” in our dataset. In fact, in the 29 countries we analyzed, restricting to the post-election phase, only in a few cases unsuccessful (“official”) confidence votes have occurred and are therefore available for our data analysis (the countless failed attempts of formateurs normally stop without a formal vote, and are impossible to document systematically). Other failure points can be obtained if we consider legislatures that ended without a government being sworn in. We could find such data for Spain and Greece, see Section B.9 and Fig. 14a for more details. Figure 3 shows the correlation between frustration and duration of the government negotiation talks with and without these failure points. It can be seen that our predictions improve when also the latter are included.

Alongside the Czech case, in Section B.8 we analyze in detail several cases in which a hung parliament led to long negotiation times (Ireland in 2016/2020, Spain in 2015/2016/2019 and Germany in 2017). Here it is instructive to give some detail for the Swedish elections of 2018. In Sweden after the 2018 elections it took a record four months to form a government, and in fact the 2018 point has the highest frustration and is highly influential in the regression line. This example is also emblematic of the difficulties encountered when taking into account a priori information like pre-electoral coalitions or ideological classifications of parties (i.e., rile). The government formed after the 2018 Swedish elections, Social Democrats (S) and Green Party (MP) with support in the form of abstention from Centre Party (CP), Liberals (L) and Left Party (V), broke the pre-electoral coalition alliances (S+MP+V, Moderate Party+CP+Christian Democrats+L) hence in scenario **II** the 2018 election point is still influential, but leads to an overall negative correlation, see Fig. 12. On the other hand, for certain countries adding coalition and ideological spectrum information increases significantly the correlation. For instance in Italy it increases from 0.06 to 0.45 (scenario **II**), which grows further to 0.95 in scenario **III**, see Fig. 12 and Fig. 13.

Under the assumption of homogeneous party behavior, the adjacency matrices describing the parliamentary networks have blockwise identical entries, see Section A.1 for the details. This means that the energy landscape determined by the functional (1) with spin-like (i.e., ± 1) states can be explored systematically. For scenario **I**, the values of energy corresponding to the various configurations are shown in Fig. 11 for all countries and all elections. Since the weights in this sce-

nario are restricted to $+1$ and -1 , there are normally several degenerate ground states (leftmost bin in the histograms of Fig. 11, see also Fig. 15), as well as many other low energy states. Each state corresponds to a partition of the parliament into two aggregations of parties. The index η_{gov} shown in Fig. 2c expresses how close the energy of the “true government” is to that of the ground state. As can be seen in Fig. 11, in the vast majority of cases, it is very close to the ground state energy ($\eta_{\text{gov}} = 0.95$ in scenario **I**). When the adjacency matrix is weighted (scenario **II** and **III**) degenerate states become less frequent. The value of η_{gov} decreases, although it remains always around 0.8, sign that the “true government” is never “energetically implausible” for our selected weights (i.e., formed by implausible alliances of ideologically distant parties, unless this is strictly necessary for a particular parliament to achieve a majority).

En passant it is worth observing that our candidate government coalition S_{best} subsumes also the notion of “minimal winning coalition” of [2], i.e., in the ground state a government majority is reached but not unnecessarily exceeded. This is always true for scenario **I**, while in scenarios **II** and **III** the government coalition can be minimal or surplus, depending on the party coalitions.

The fractionalization index F (see Section 4.4 for a definition) is a widely used measure of concentration [12]. In our context, it is linked to the probability that two randomly chosen MPs belong to different parties. The equally popular effective number of parties N_2 [11] can be derived straightforwardly from it: $N_2 = 1/(1-F)$. The “power 2” present in the formulas for F and N_2 (see Section 4.4) has been interpreted for instance in terms of mean and variance of the distribution of seats per party [43]. By looking from a signed network perspective, for scenario **I**, we can give another interpretation, namely in terms of number of positive/negative edges associated to the parliamentary network. In fact, F is the fraction of negative edges in the signed graph, and for scenario **I** the frustration follows very closely this index, see Fig. 2d. The only cases in which it deviates slightly (e.g., Albania, North Macedonia, UK, see Fig. 16) are associated to predicted minimum winning government coalitions P_{best} (corresponding to S_{best}) which consist of a number of seat in excess of the absolute minimum of 50% +1 of the seats. An exact formula expressing frustration in terms of fractionalization index and seat excess in a minimum winning coalition is given in Section 4.4 (see also Section B.4). The consequence is twofold: first, even though frustration (a measure on signed graphs) is conceptually different from standard measures of parliamentary fragmentation, the fact that it nearly perfectly correlates with the fractionalization index F means that it is consistent with the existing literature. Second, as opposed to the classical fragmentation indexes, frustration also encodes information on the minimum winning coalition principle.

If in scenario **I** frustration is a close proxy for classical measures of parliamentary fragmentation, in scenarios **II** and **III** the two quantities ζ and F no longer coincide, as the signed adjacency matrix of a parliamentary network now has become weighted, which alters the value of ζ , see Fig. 2d. In this respect, the method we are proposing is quite versatile, as it can easily incorporate both the basic fragmentation measure used in the literature as well as other factors, like ideological polarization and pre-existing party coalitions, as long as these factors can be ex-

pressed as sets of weights for the adjacency matrix of the signed parliamentary network.

In conjunction with the dynamical model shown in Fig. 1c, it is worth observing that crossing a bifurcation in our model (7) always leads to the onset of two equilibria with opposite signatures, say x^* and $-x^*$. These two equilibria are aligned with the direction of the eigenvector of the least eigenvalue of the normalized Laplacian, and are related to success (S_{best} , which has more positive than negative entries) and failure ($-S_{\text{best}}$) of a confidence vote. In the vast majority of cases, only the first equilibrium is seen in our data, although some sporadic cases of failure to win parliamentary confidence (with a formal vote) are reported for some countries and are mentioned above. Our model (7) is symmetric and cannot account for this symmetry breaking effect. However, more sophisticated models, such as the one considered in [29, 30], in which external influences and peer pressure lead to unfolding of the pitchfork bifurcation, can be used to capture this phenomenon. In particular, in the context of parliamentary systems, the pressure from the public opinion, from the media and from the parliament itself is to avoid a formal vote of confidence when the chances of passing it are null or even just slim, unless dictated by constitutional rules. How to modify the model (7) to account for these external factors will be the subject of future research.

In order to understand the confounding influence of unmodeled factors on our statistics, it is worth noticing that low values of ρ_{gov} or η_{gov} can often be observed in two special cases not explicitly included in our model: when a minority government is formed after the election (our government configuration S_{gov} does not include parties that support the government in the parliament but have no assigned ministers) or, in bicameral systems, when a candidate cabinet needs to achieve a majority in both chambers of the parliament (only the lower chamber is analyzed in this study). Consider for example Italy: low values of η_{gov} are obtained after the 1996, 2013 and 2018 elections, see scenario **II** and **III** in Fig. 5. The 1996 election sees the formation of a minority government, the Prodi I cabinet (a coalition of centre-left parties) which, while enjoying the confidence in the Senate of the Republic, did not manage to secure a majority in the Chamber of Deputies. This government had to be supported by the Communist Refoundation Party (PRC) and some other smaller parties in order to achieve majority in both chambers of the parliament. If we include these parties in the government configuration, now described by $S_{\text{gov+supp}}$ (see Section 4.5), we can observe that the energy gap between the ground state and the energy functional $e(S_{\text{gov+supp}})$ decreases, see the 1996 election in Fig. 5. The opposite situation happened instead at the 2013 election where, while a clear majority was reached in the Chamber of Deputies by the centre-left alliance (that won 345 of the 630 seats), none of the party alliances obtained a majority in the Senate, whose seats were won mostly by three party blocs: the centre-left alliance, the centre-right alliance and the Five Star Movement (M5S), holding respectively 123, 117 and 54 of the 315 total seats. However, neither the centre-left nor the centre-right coalition wanted to form a government with the M5S party. After 62 days of government negotiations, the Letta cabinet, a grand coalition comprising parties from both the left and right side of the political spectrum, was sworn in. This configuration of parties was far

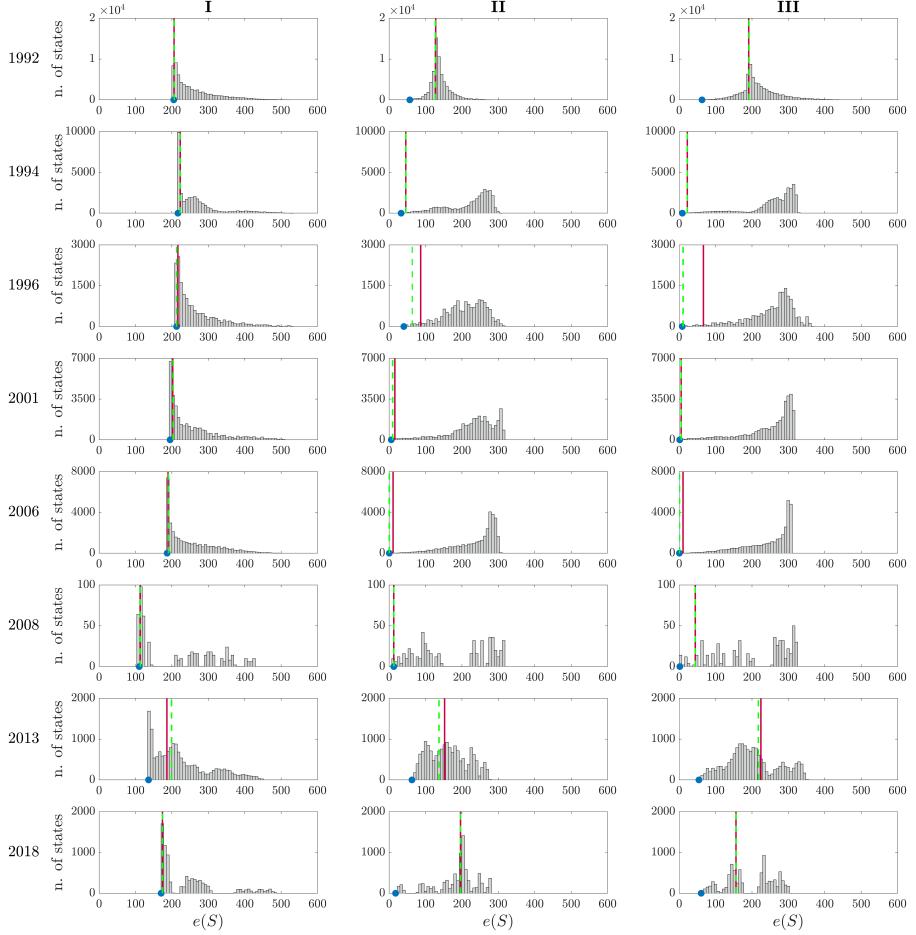


Figure 5: Energy landscape in Italy for each scenario (**I**, **II** and **III**) and election, i.e., values of the energy $e(S)$ calculated according to (1) for all the 2^{n_p} ($n_p = n$. of parties in the parliament) choices of $S = \text{diag}\{S_1, \dots, S_{n_p}\}$, $S_i = s_i I_{c_i}$, $s_i = \pm 1$. The blue dot indicates the minimum of such energy functionals, i.e., the frustration (2). The red line represents $e(S_{\text{gov}})$, i.e., the value of the energy functional corresponding to the government, while the dashed green line represents $e(S_{\text{gov+supp}})$, i.e., the value of the energy functional corresponding to the government plus the other parties providing support in parliament without formally entering into the government coalition (see Section 4.5).

from being a “minimum winning coalition” which, together with the fact that pre-electoral coalitions had to be broken to reach a cabinet coalition of ideologically distant parties, explains the low value of η_{gov} we obtain for the 2013 election in all three scenarios. For example, in scenario **III** where the pre-electoral coalitions and the ideological differences between parties are taken into consideration, our ground state corresponds exactly to the centre-left coalition, which is energetically

far from the Letta cabinet coalition, see Fig. 5 (2013 election). Moreover, both the 2013 and 2018 elections are examples on how the presence of the “big tent” M5S party (which does not have a clear classification on the left-right scale) further increases the complexity of the government negotiation process: in 2018 the Conte I cabinet was sworn in only after an agreement was reached between the far-right Lega and the M5S and, at 89 days, it is the longest government negotiation process since 1992 (where we start our analysis). The same conclusion applies if we consider the signed network corresponding to the Senate of the Republic. In all three scenarios the governing coalition (together with its support in the parliament) corresponds to the least energy achieving a majority in both upper and lower chamber, see Fig. 6 for scenario III, with the exception of the 1992, 2013 and 2018 elections.

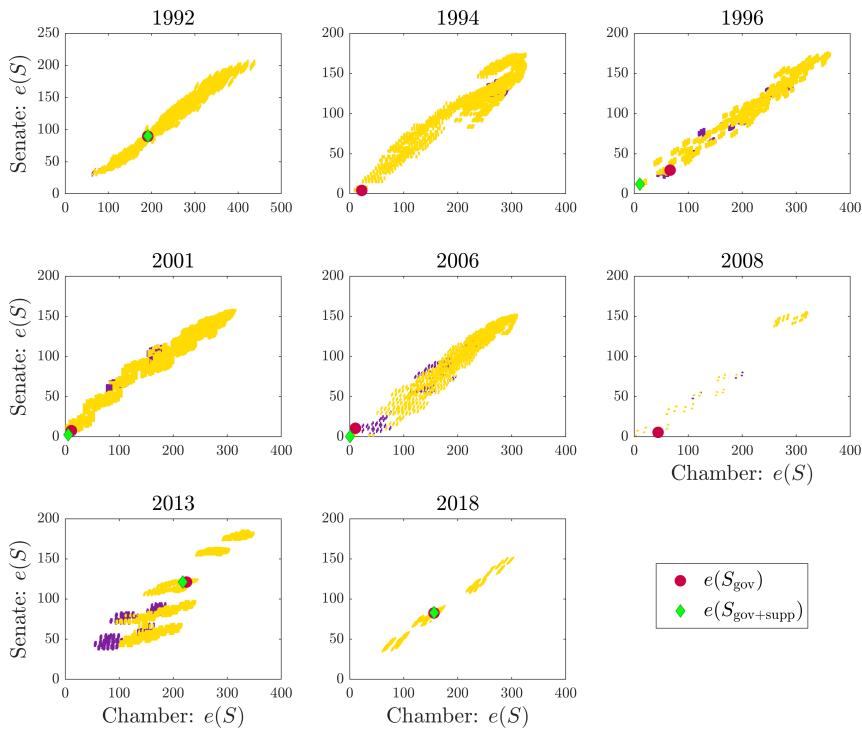


Figure 6: Comparison of the energy landscapes in the Chamber of Deputies and Senate of the Republic in Italy for scenario III. The red dot indicates the value of the energy functional corresponding to the government configuration S_{gov} , while the green diamond indicates the value of the energy functional corresponding to the party configuration $S_{\text{gov+supp}}$, which includes also the parties providing external support to the government in the parliament (see Section 4.5). The values of the energy functional corresponding to majority configurations (i.e., party coalitions holding majority of seats in both chambers of the parliament) are highlighted with a yellow color, while a purple color indicates party coalitions which have the majority of seats in only one chamber (see Section B.11 for details).

There is a significant difference between scenarios **I** and **II** on the one hand, and scenario **III** on the other. In the first two scenarios, in fact, it is possible to claim that the correlation observed is actually a form of *causation*: the frustration of a network is computed solely on the basis of the data of the problem available *ex ante* (n. of parties and n. of MPs per party for scenario **I**; n. of parties and MPs plus coalition and rile for scenario **II**) without making use of any tuning parameter to fit the data. Hence a claim that the level of frustration is a “cause” (or one of the causes) behind the duration of the government negotiation phase, seems a reasonable one, given our interpretation of frustration as amount of antagonism present in a parliament. In scenario **III**, instead, the edge weights of the parliamentary network are tuned so as to maximize the correlation between frustration (a function of the edge weights) and duration of the negotiations, hence this scenario corresponds more to an *a posteriori* party classification based on a party behavior over the years. The rationale behind this approach is that a party “realpolitik” is a combination of factors such as those composing indexes like rile (ideological, social, economical, military, etc) but also of other variables like pragmatism, opportunity, necessity, national interest, etc, whose relative importance is difficult to assess and quantify. Notice that unlike in scenario **II** where new rile values are available at each new general election, in scenario **III** we keep the weights fixed throughout each party history, meaning that this method can have predictive power for future elections.

To evaluate such predictive power, we can consider a “leave-the-last-one-out” validation, which is a particular case of the “leave-one-out” analysis we have performed (see Fig. 4 and Section 4.8). When we consider the last (most recent) election as the validation set, we can observe that our model (scenario **III**) is able to predict the duration of the negotiations in most countries, see e.g., Denmark, Finland, Iceland, Italy, Slovakia, Spain, etc., in the regression plots of Fig. 17. However, there are few cases in which our scenario **III** must necessarily fail, namely when a country experiences a hung parliament for the first time in the latest election, see Sweden or Germany in Fig. 17. These failures are of course expected given the nature of the model and of the test (the validation set proposes a novel situation to the model). This confirms that high leverage and influential points are crucial to our approach.

As can be seen in Fig. 3, the scales for both frustration and government negotiation days vary widely between states. For the frustration this is due to factors such as number of MPs and parties, while for the negotiation days it is due to different national constitutional rules and traditions. Nevertheless, after a suitable normalization (see Section 4.9), the data for the different nations can be assembled into pan-European time-series. As shown in Fig. 7, a few significant trends are clearly emerging: both frustration and government negotiation times have nearly doubled over the last 30 years. The increased frustration is due to an increase in parliamentary fragmentation (recall that in this scenario the frustration follows very closely the fractionalization index F), which is reflected by the increase in the number of parties represented in parliament, see Fig. 7c, and by parties of smaller size, see Fig. 7d. Likely factors behind these trends are the erosion of ideology-driven historical parties, the rise of populist and nationalistic parties and the appearance

of post-ideological movements such as the “big tent” parties (e.g., M5S in Italy and United We Can in Spain). Overall, these factors add “disorder” to our parliamentary networks. In spite of the political sphere being commonly referred to as the quintessential “art of compromise”, this increased disorder puts a strain on the functioning of our political systems. By representing disorder in terms of graph frustration, our models provide a natural explanation and a quantitative assessment of some of the observable consequences of this phenomenon.

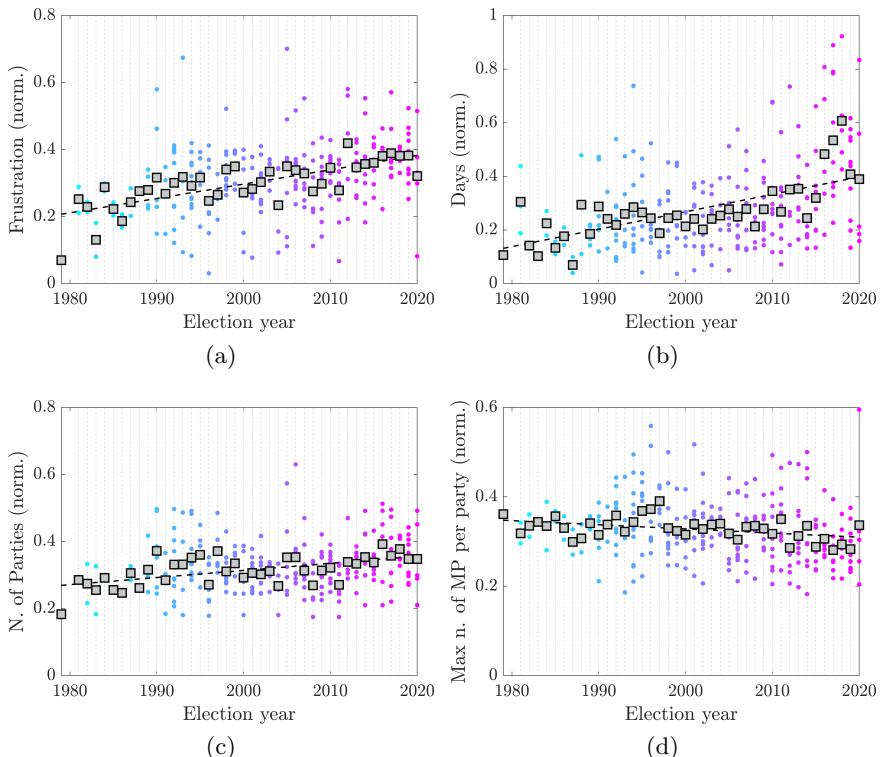


Figure 7: Yearly trends for the ensemble of 29 nations of Table 1 (normalized values). The gray boxes indicate the mean value for each year, the dashed black line the best linear fit for these mean values. (a): Frustration for scenario I; (b): Duration of the government negotiation talks (days); (c): Number of parties in parliament; (d): Maximum number of MPs per party in parliament.

4 Materials and Methods

4.1 Data description

A total of 29 European countries, whose system of government is a Parliamentary Republic or a Parliamentary Constitutional Monarchy, were considered in

the analysis and are listed in Table 1. Countries adopting a Presidential (e.g., Belarus, Cyprus) or Semi-presidential (e.g., France, Georgia, Lithuania, Poland, Portugal, Romania, Russia, Ukraine) system were not considered. Countries that have switched from Semi-presidential to Parliamentary system in recent times (e.g., Armenia) were also disregarded. Countries with a short history of political elections (less than five) were excluded from the analysis (e.g., Montenegro, Kosovo). The Republic of Malta has not been considered given that there are only two parties competing in the elections, hence its parliamentary graph is always structurally balanced, with zero frustration (see below). When a bicameral system is adopted, we only consider data for the lower house.

We are interested only in general elections called after the dissolution of the parliament and only in the cabinet formation process immediately following the elections.

Data were collected from various sources such as the Manifesto Project Database [44], the Parliaments and Governments Database [45], the new Parline (IPU’s Open Data Platform) [46], the Chapel Hill Expert Survey (CHES) [47, 48] and WIKIPEDIA. To review the information other references were consulted, such as [49–53] and [4] (in particular the last reference provided additional information regarding the existence of pre-electoral coalitions). The following data are considered: the election dates, the political parties winning seats at the elections and their position in the left-right political spectrum, the existing pre-electoral alliances, and the composition of the government formed after the elections and approved by the Parliament (confidence vote) in terms of swearing-in date, composition and status (minority or majority). Failures in forming a government are determined from unsuccessful confidence votes or from dissolutions of parliament without any government having been formed. All data were cross-checked on different datasets whenever possible.

As Table 1 shows, the number of elections considered for each country varies between 6 and 13, and we decided to start the analysis from election years around 1980 or later, given the difficulty to find information about the elections and/or parties involved before that year. However, some events considered to be particularly influential have moved the starting point for some countries, see Table 2. Our data collection stops at Fall 2020.

4.2 Construction of parliamentary networks

A parliamentary network is an undirected, complete graph in which each MP is a node. MPs from the same party are linked through positive edges of weight equal to +1, while MPs from different parties are linked through signed edges that can be chosen according to the following scenarios (see Fig. 1a). **I:** unweighted all-against-all (all weights equal to −1), see Fig. 1b; **II:** weighted according to the ideological spectrum (computed according to the rile index [33, 44]) and pre-electoral coalitions (MPs from parties involved in a coalition have positive edge weights), see Fig. 8a (top); **III:** weighted according to both a semi-randomized (and optimized) left-right grid and pre-electoral coalitions, see Fig. 8a (bottom). In this last case we considered 10000 different sets of values for the party positions

in the left-right political spectrum (which we kept fixed for all the elections) and computed the corresponding values of correlation between frustration and duration of government negotiation talks for each set. For each country the “optimal” choice of political positions corresponds to the one giving highest correlation. See Section A.1 for the details.

We assume that parties are homogeneous entities: each MP relates (in terms of cooperation or antagonism) to other parties’ MPs according to his/her party line. The resulting signed adjacency matrices are consequently always block matrices, see Fig. 1b and Section A.1 for more details.

4.3 Measuring the frustration of signed parliamentary networks

The notion of structural balance [22, 26] captures the idea that it is possible to split a graph into two subgraphs such that all edges on each subgraph are positive, while all edges through the cut set that splits the graph are negative. In our parliamentary network it could represent a two-party parliament or, in a scenario including electoral coalitions, a parliament split into two coalitions. In general, the signed parliamentary networks we obtain (in all scenarios) are not structurally balanced. In this case, we use the notion of “frustration” to measure the distance of a signed network from a structurally balanced state. Let \mathcal{L} be the normalized Laplacian corresponding to the signed graph, defined as $\mathcal{L} = I - (\text{diag}\{|A|\mathbf{1}\})^{-1}A$, where A is the adjacency matrix of the network, $|\cdot|$ is the element-wise absolute value, $\mathbf{1}$ is the vector of 1s and $\text{diag}\{x\}$ is a diagonal matrix having the vector x on the diagonal. By associating a binary variable $s_i = \pm 1$ (spin up and down) to each party, one can define an energy-like quantity

$$e(S) = \frac{1}{2} \sum_{i,j \neq i} [|\mathcal{L}| + S\mathcal{L}S]_{ij}, \quad (1)$$

where $S = \text{diag}\{S_1, \dots, S_{n_p}\}$ with $S_i = s_i I_{c_i}$ and $s_i = \pm 1$ ($i = 1, \dots, n_p$), n_p is the number of parties in the parliament and c_i , $i = 1, \dots, n_p$, is the number of seats for the i -th party ($\sum_{i=1}^{n_p} c_i = n$ where n is the number of seats in parliament). Notice that $e(S) = e(-S)$. The frustration of the parliamentary network corresponds to the ground state energy, i.e., the minimal of this functional over all possible combinations of parties

$$\zeta = \min_{\substack{S=\text{diag}\{S_1, \dots, S_{n_p}\}, \\ S_i=s_i I_{c_i}, s_i=\pm 1}} e(S). \quad (2)$$

Denote $\pm S_{\text{best}}$ the diagonal matrices achieving the optimum in (2), with S_{best} (“success”) having more +1 entries than $-S_{\text{best}}$ (“failure”).

4.4 Frustration vs fractionalization index

The fractionalization index F is one of classical measures of parliament fragmentation [12]. It can be expressed in terms of the Laasko-Taagepera effective number

of parties N_2 as $F = 1 - \frac{1}{N_2}$, where N_2 is defined as

$$N_2 = \frac{1}{\sum_{i=1}^{n_p} \left(\frac{c_i}{n}\right)^2} = \frac{n^2}{\sum_{i=1}^{n_p} c_i^2}, \quad (3)$$

with c_i the size of the i -th party, n_p the total number of parties and n the total number of MPs [11].

In our scenario **I** (all-against-all, no pre-electoral coalitions), the fractionalization index F corresponds to the fraction of negative entries in the adjacency matrix of the network. In fact, since all the party-party weights are equal to -1 and A is a $n \times n$ block matrix where each block has size $c_i \times c_j$ ($i, j = 1, \dots, n_p$), it follows that the total number of negative elements in A is equal to $\sum_{i,j=1}^{n_p} c_i c_j = n^2 - \sum_{i=1}^{n_p} c_i^2 = n^2 \left(1 - \frac{\sum_{i=1}^{n_p} c_i^2}{n^2}\right) = n^2 F$ and that the fraction of negative entries is equal to F .

As shown in Fig. 2d and Fig. 16, the frustration of the unweighted signed networks of scenario **I** has a very high correlation (0.99 in average) with F . To explain this correlation, we can observe that the frustration of a signed unweighted parliamentary network corresponds (up to a multiplicative constant) to a difference of three terms, the fractionalization index F , a constant term and a third term interpretable as the “distance” of the minimum winning coalition in correspondence of S_{best} from the 50% of the seats:

$$\zeta = \frac{n^2}{n-1} \cdot \left(F - \frac{1}{2} + \underbrace{\frac{1}{2} \left(\frac{E_{\text{best}}}{n/2} \right)^2}_{\text{“distance” of } \mathcal{P}_{\text{best}} \text{ from 50\%}} \right), \quad (4)$$

where $\mathcal{P}_{\text{best}}$ describes the majority coalition in correspondence of S_{best} (see also Section 4.5) and $E_{\text{best}} \in [0, \frac{n}{2}]$ is the number of seats in excess of $\mathcal{P}_{\text{best}}$ with respect to 50% of the total number of seats: $\sum_{i:p_i \in \mathcal{P}_{\text{best}}} c_i = \frac{n}{2} + E_{\text{best}}$, where p_i represents the i th party (see also Section B.4 for more details). This “distance” attains its minimum value when S_{best} corresponds to 50% +1 of the seats in a parliament, and grows when the seat excess corresponding to S_{best} grows.

4.5 Minimum energy government coalition

For each country and parliamentary election we compute the energy (1) for all 2^{n_p} (n_p is the number of parties in the parliament) possible party configurations S (see Fig. 11 for the energy landscape in scenario **I**) and the frustration (2) of the signed parliamentary network as the minimum of such energies. Our predicted government coalition, denoted $\mathcal{P}_{\text{best}}$, is given by the corresponding group of parties achieving a majority in S_{best} : $\sum_{i:p_i \in \mathcal{P}_{\text{best}}} c_i > \frac{n}{2}$ where c_i is the number of seats gained by the party p_i and n is the total number of MPs in parliament. Similarly, let S_{gov} be the party coalition that effectively obtained a confidence vote for that election and \mathcal{P}_{gov} the list of corresponding parties, $S_{\text{gov}} = \text{diag}\{s_1 I_{c_1}, \dots, s_{n_p} I_{c_{n_p}}\}$ with $s_i = +1$ if $p_i \in \mathcal{P}_{\text{gov}}$ or $s_i = -1$ otherwise. Denoting the cardinality of a set as

$\text{card}(\cdot)$, the fraction of government coalition captured by the ground state party coalition $\mathcal{P}_{\text{best}}$ can be defined as

$$\rho_{\text{gov}} = \frac{\text{card}(\mathcal{P}_{\text{best}} \cap \mathcal{P}_{\text{gov}})}{\text{card}(\mathcal{P}_{\text{gov}})}, \quad (5)$$

see Fig. 2b (and Fig. 10c, for scenario **I** only). The closer the value of ρ_{gov} is to 1, the better our estimate represents the actual cabinet composition. In scenario **I** the ground state may be “degenerate”, i.e., multiple signature matrices S_{best} could give the same value of frustration, and/or the corresponding party group $\mathcal{P}_{\text{best}}$ may hold exactly half of the seats in the parliament ($\sum_{i:p_i \in \mathcal{P}_{\text{best}}} c_i = \frac{n}{2}$). In the first case, we consider the matrix S_{best} giving maximum value of ρ_{gov} . In the second, we take as S_{best} the matrix that yields the least value in (2) while satisfying the condition $\sum_{i:p_i \in \mathcal{P}_{\text{best}}} c_i > \frac{n}{2}$ (see Section B.2 for the details). The case in which S_{best} is unique and its $\mathcal{P}_{\text{best}}$ corresponds to exactly $\frac{n}{2}$ is also degenerate as the ground state has no majority.

It is also possible to compare the frustration ζ with the actual government “energy”:

$$\eta_{\text{gov}} = 1 - \frac{e(S_{\text{gov}}) - \zeta}{\max_S e(S) - \zeta}. \quad (6)$$

η_{gov} expresses how close the energy of the actual government is compared to the frustration, i.e., the theoretical minimum of such quantity over all possible party combinations, see Fig. 2c (and Fig. 10d, for scenario **I** only).

When a minority government \mathcal{P}_{gov} needs additional support to win a confidence vote, we denote $\mathcal{P}_{\text{supp}}$ the set of parties which support the cabinet in the parliament without being explicitly part of it and $S_{\text{gov+supp}} = \text{diag}\{s_1 I_{c_1}, \dots, s_{n_p} I_{c_{n_p}}\}$ the corresponding party configuration, with $s_i = +1$ if $p_i \in \mathcal{P}_{\text{gov}} \cup \mathcal{P}_{\text{supp}}$ or $s_i = -1$ otherwise. The intuition is that $e(S_{\text{gov+supp}})$, i.e., the value of energy functional corresponding to the (majority) coalition supporting the government in the parliament (computed according to (1)), should be energetically closer to our ground state than $e(S_{\text{gov}})$, i.e., the energy functional corresponding to the (minority) government. See for example the 1996 election in Italy in Fig. 5. This analysis is carried out only for Italy, however we expect similar results for the other countries where minority governments are common, such as Denmark, Norway or Sweden (see Fig. 14b).

4.6 Dynamical model of government formation

We can describe the government formation process by the nonlinear interconnected model of collective decision-making introduced in [30–32],

$$\dot{x} = -\Delta x + \pi A \psi(x), \quad x \in \mathbb{R}^n. \quad (7)$$

In (7), n is the number of elected MPs, the vector $x = [x_1 \dots x_n]^T$ represents their opinions, A is the adjacency matrix of the signed parliamentary network, $\Delta = \text{diag}\{\delta_1, \dots, \delta_n\}$ is a diagonal matrix such that $\delta_i = \sum_{j \neq i} |a_{ij}|$ for all i , $\psi(x) =$

$[\psi_1(x_1) \cdots \psi_n(x_n)]^T$ is a vector of sigmoidal and saturated nonlinearities (expressing how the MPs transmit their opinion to the other MPs in the network) and π is a positive scalar parameter acting as bifurcation parameter. More precisely, the bifurcation occurs at

$$\pi_1 = \frac{1}{1 - \lambda_1(\mathcal{L})}, \quad (8)$$

see Fig. 1c, where $\lambda_1(\mathcal{L})$ is the smallest eigenvalue of the normalized signed Laplacian \mathcal{L} associated to (7), see Section A.3 for the details.

4.7 Rationale behind the correlation between frustration and duration of the government negotiation phase

If we denote τ the duration of the government negotiation phase, from election day to the day the government is sworn in (and, for the cases of failure, to the day in which the formateur loses a confidence vote or to the day in which the parliament is dissolved without a government having been formed), then τ and the frustration ζ are related by a form of direct proportionality (here indicated by the symbol “~”): $\tau \sim \zeta$. This can be explained in terms of the following chain of relationships linking τ to ζ via the bifurcation parameter π_1 and the algebraic conflict $\lambda_1(\mathcal{L})$:

$$\tau \sim \pi_1, \quad \pi_1 = \frac{1}{1 - \lambda_1(\mathcal{L})}, \quad \lambda_1(\mathcal{L}) \sim \zeta$$

While the second on these relationships is analytical, the first and third are only numerical. The correlation plots of Fig. 3 (Fig. 12 and Fig. 13 for scenario **II** and **III**, respectively) are an expression of this $\tau \sim \zeta$ relationship.

4.8 Leave-one-out analysis for scenario **III**

In scenario **III**, differently from **I** and **II**, the political positions of the parties are fixed in time and are chosen to maximize the Pearson correlation between frustration of the parliamentary networks and duration of the government negotiations. To validate this method we have performed a leave-one-out analysis, whose idea is to use only $N - 1$ elections (where N is the number of elections for each country) to determine the political positions of each party and then check how well the model “fits” the excluded election, by comparing the Pearson correlation coefficients (frustration of the parliamentary networks v.s. duration of the government negotiations) obtained with and without the data point corresponding to this election.

The procedure is described in detail as follows. For each country, the data on the parliamentary elections are divided into two sets: a validation set, made of a single election, and a training set, made of the remaining ($N - 1$) elections. The training set is used to find the “optimal” choice of political positions by maximizing the correlation frustration v.s. duration of the government negotiation talks. 1000 sets of values for the left-right political positions of the parties (see (10) in the Appendix) are randomly selected on the preassigned left-right grid, and for each choice of political positions the Pearson correlation (frustration v.s.

days) is calculated: the “optimal” choice for the weights corresponds to the one giving the best correlation index (among the 1000). This choice for the weights is used to build the parliamentary network corresponding to the excluded election (the validation set), whose frustration is then computed. Finally, the correlation (frustration v.s. days), when all the N elections are considered, is calculated. This process is repeated N times, each time by selecting a different election as validation set, obtaining (for each country) $N \times 1000$ correlation indexes calculated using the training sets and N “optimal” choices for the weights and corresponding values of correlation.

The values of correlation (frustration v.s. days) obtained when optimizing for $N - 1$ elections (i.e., leave-one-out analysis) are then compared with those obtained when optimizing for N elections (i.e., scenario **III**). Figure 4 illustrates these results. A particular case of this analysis is the “leave-the-last-one-out” validation, where the validation set consists of the last election. Figure 17 shows the regression plots between duration of government negotiations and frustration where the weights of the parliamentary networks are chosen so as to maximize the correlation (days v.s. frustration) in the first $N - 1$ elections.

4.9 Yearly trends

To obtain the yearly trends shown in Fig. 7 we compute and plot, for each country, the normalized values $\frac{x}{\|x\|_2}$ over the election years, where x is the N -dimensional vector of interest (frustration, government negotiation days, number of parties in the parliament and maximum number of MPs per party), N is the number of elections, and $\|\cdot\|_2$ the Euclidean norm.

Country (structure of parliament)	Election Dates	N. of Seats (n)	Relevant Events	
Albania (U)	1992, 1996, 1997, 2001, 2005, 2009, 2013, 2017	140 155 (1997 only)	1991	Dissolution of Social Republic, start of the 4th Republic.
Andorra (U)	1993, 1997, 2001, 2005, 2009, 2011, 2015, 2019	28	1993	Adoption of a new Constitution.
Austria (B)	1979, 1983, 1986, 1990, 1994, 1995, 1999, 2002, 2006, 2008, 2013, 2017, 2019	183	—	
Belgium (B)	1995, 1999, 2003, 2007, 2010, 2014, 2019	150	1993	Belgian Constitution, Belgium becomes a federal state.
Bosnia and Herzegovina (B)	1996, 1998, 2000, 2002, 2006, 2010, 2014, 2018	42	1992	Independence from SFR Yu- goslavia.
Bulgaria (U)	1991, 1994, 1997, 2001, 2005, 2009, 2013, 2014, 2017	240	1995	New Constitution.
Croatia (U)	1992, 1995, 2000, 2003, 2007, 2011, 2015, 2016, 2020	140 +M+D ^(a) (from 2000) 127 (1995) 138 (1992)	1991	Independence from SFR Yu- goslavia.
Czech Republic (B)	1992, 1996, 1998, 2002, 2006, 2010, 2013, 2017	200	1993	Independence from Czechoslo- vakia.
Denmark (U)	1981, 1984, 1987, 1988, 1990, 1994, 1998, 2001, 2005, 2007, 2011, 2015, 2019	179	—	
Estonia (U)	1992, 1995, 1999, 2003, 2007, 2011, 2015, 2019	101	1991	Independence from the Soviet Union.
Finland (U)	1991, 1995, 1999, 2003, 2007, 2011, 2015, 2019	200	1991	Constitution amendments: the powers of the President are di- minished.
			2000	New constitution; from semi- presidential to parliamentary republic [54].
Germany (B)	1990, 1994, 1998, 2002, 2005, 2009, 2013, 2017	598+O+L ^(b)	1990	German reunification.
Greece (U)	1990, 1993, 1996, 2000, 2004, 2007, 2009, May 2012, Jun 2012, Jan 2015, Sept 2015, 2019	300	—	
Hungary (U)	1990, 1994, 1998, 2002, 2006, 2010 ^(c)	386	1989	Third Republic.
Iceland (U)	1995, 1999, 2003, 2007, 2009, 2013, 2016, 2017	63	1991	From Bicameral to Unicameral Parliament.
Ireland (B)	1992, 1997, 2002, 2007, 2011, 2016, 2020	159 (2020) 157 (2016) 166 (until 2011) 630	—	
Italy (B)	1992, 1994, 1996, 2001, 2006, 2008, 2013, 2018		1993	New electoral system.
Latvia (U)	1993, 1995, 1998, 2002, 2006, 2010, 2011, 2014, 2018	100	1991	Independence from the Soviet Union.
Luxembourg (U)	1984, 1989, 1994, 1999, 2004, 2009, 2013, 2018	60 (from 1989) 64 (1984)	—	
North Macedonia (U)	1990, 1994, 1998, 2002, 2006, 2008, 2011, 2014, 2016, 2020	120 + A ^(d)	1991	Independence from SFR Yu- goslavia (officially recognized in 1993).
Moldova (U)	1994, 1998, 2001, 2005, July 2009, 2010, 2014, 2019 ^(e)	101 (from 1998) 104 (1994)	1991	Independence from the Soviet Union.
Netherlands (B)	1981, 1982, 1986, 1989, 1994, 1998, 2002, 2003, 2006, 2010, 2012, 2017	150	—	

Continued on next page

Table 2 – continued from previous page

Norway (U)	1981, 1985, 1989, 1993, 1997, 2001, 2005, 2009, 2013, 2017	169 (from 2005) 165 (1989-2001) 157 (1985) 155 (1981)	–
Serbia (U)	2007, 2008, 2012, 2014, 2016, 2020	250	2006 Independence declared from the Union of Serbia and Montenegro.
Slovakia (U)	1992, 1994, 1998, 2002, 2006, 2010, 2012, 2016, 2020	150	1993 Independence from Czechoslovakia.
Slovenia (B)	1992, 1996, 2000, 2004, 2008, 2011, 2014, 2018	90	1991 Independence from SFR Yugoslavia.
Spain (B)	1989, 1993, 1996, 2000, 2004, 2008, 2011, 2015, 2016, Apr. 2019, Nov. 2019	350	–
Sweden (U)	1982, 1985, 1988, 1991, 1994, 1998, 2002, 2006, 2010, 2014, 2018	349	–
United Kingdom (B)	1983, 1987, 1992, 1997, 2001, 2005, 2010, 2015, 2017, 2019	649 +1 speaker (from 2010) 645 +1 (2005) 658 +1 (2001, 1997) 651 (1992) 650 (1987, 1983)	–

(a) M = minority (8), D = diaspora seats (3 to 6).

(b) O = overhangs (since 2005), L = leveling seats (since 2013).

(c) The elections dated 2014 and 2018 have not been considered since the new Constitution of Hungary (2012) changed considerably the number of MPs (hence the frustration is not comparable).

(d) A = seats for Macedonians living abroad (in case of sufficient voter turnout).

(e) The April 2009 election is not considered since no president was elected and new elections had to be held in July 2009.

Table 2: List of countries considered in this work. For each country, we list: the structure of the Parliament (B = Bicameral or U = Unicameral), the election dates considered in this study, the number of seats in the Parliament (in case of changes, the year is specified), and the significant events which determined the starting point in our analysis.

Appendix

A Methods

A.1 Parliamentary network construction

For each country and parliamentary election listed in Table 2 we consider an undirected graph $\mathcal{G}_{\text{country}} = (\mathcal{V}, \mathcal{E}, A)$, where \mathcal{V} ($\text{card}(\mathcal{V}) = n$) is the vertex set, \mathcal{E} is the edge set, and $A = [a_{ij}] \in \mathbb{R}^{n \times n}$ is the adjacency matrix, with a_{ij} representing the weight of the edge $(j, i) \in \mathcal{E}$. Each of the n vertices in \mathcal{V} represents an elected Member of the Parliament (MP), and each edge in \mathcal{E} the relationship between two MPs, which can be cooperative or antagonistic. A positive edge $a_{ij} > 0$ means cooperation between the i -th and j -th MPs, a negative edge $a_{ij} < 0$ antagonism. All MPs of a party are always assigned the same weights, i.e., we treat a party as a homogeneous cluster of nodes. As shown in Fig. 1b, the network $\mathcal{G}_{\text{country}} = (\mathcal{V}, \mathcal{E}, A)$ is therefore composed of n_p clusters, represented by the political parties,

where n_p is the total number of parties which gained seats in the parliament after the election. This implies that the adjacency matrix A , as well as all the other matrices of interest, can be seen as a $n_p \times n_p$ block matrix:

$$A = \begin{bmatrix} A_{11} & \dots & A_{1n_p} \\ \ddots & \ddots & \\ A_{n_p 1} & \dots & A_{n_p n_p} \end{bmatrix}, \quad \text{with } A_{ij} = \begin{cases} (E_{c_i} - I_{c_i}) w_{ii}, & j = i \\ E_{c_i c_j} w_{ij}, & j \neq i, \end{cases} \quad (9)$$

where c_i is the number of seats gained by the i -th party, $E_{c_i c_j} = \mathbf{1}_{c_i} \mathbf{1}_{c_j}^T$ (simplified to E_{c_i} if $j = i$) is the matrix of all 1s ($\mathbf{1}_{c_i}$ is the vector of size c_i having all elements equal to 1), and $W = [w_{ij}] \in \mathbb{R}^{n_p \times n_p}$ is the matrix of party-party weights. Its signed entries w_{ij} describe the interaction between MPs of the party p_i and party p_j , in terms of political affinity.

In order to choose the matrix W , we consider different party grouping criteria and different weight assignment methods. The party grouping criteria are:

1. **All-against-all.** All parties compete against all parties:

$$w_{ij} \begin{cases} = 1 & \text{if } i = j \text{ (} p_i \text{ and } p_j \text{ are the same party; hereafter: } p_i = p_j \text{)} \\ < 0 & \text{if } p_i \text{ and } p_j \text{ are different parties.} \end{cases}$$

Electoral coalitions are not taken into account in this case. Germany is an exception to this rule, in that the Christian Democratic Union of Germany (CDU) and Christian Social Union in Bavaria (CSU) are always considered as a single party.

2. **Pre-electoral coalitions.** Parties in pre-electoral coalitions are cooperating:

$$w_{ij} \begin{cases} = 1 & \text{if } p_i = p_j \\ > 0 & \text{if parties } p_i \text{ and } p_j \text{ belong to the same coalition} \\ < 0 & \text{if parties } p_i \text{ and } p_j \text{ do not belong to the same coalition.} \end{cases}$$

Pre-electoral coalitions (explicit or implicit) are obtained from references and datasets such as [4, 55–59], WIKIPEDIA and the new Parline [46].

In the case of countries where double rounds of elections are common (for example in Hungary until 2010), we consider also the electoral coalitions made before the second round (so the ones made before the first round plus the ones made between the first and second round).

We always assume that $w_{ii} = 1$ and that $w_{ij} = w_{ji}$, i.e., that W (and hence A) is symmetric. In turn, this means that $\mathcal{G}_{\text{country}}$ is undirected.

As for the weights themselves, we consider both the cases of unweighted (but signed) W and weighted (and signed) W . In the **unweighted case**, $w_{ij} \in \{-1, +1\}$, with +1 only on the diagonal (all-against-all case). The resulting graph is complete, undirected and signed.

In the **weighted case**, the general philosophy is that off-diagonal weights between parties which are not in pre-electoral coalition should be negative, small (in absolute value) for ideologically close parties and approaching -1 for ideologically

antipodal parties. Instead, off-diagonal weights between parties in pre-electoral coalition should be positive and close to 1 for ideologically close parties. In order to define the matrix of weights $W = [w_{ij}]$, the first step consists in assigning to each political party a position in the left-right ideological spectrum. For this, we follow two different criteria, as shown in Fig. 8a: (1) use data from the Manifesto Project Database, in particular the so-called *rile* index [33]; and (2) place the parties on a predetermined left-right grid and assign to them specific positions randomly. It is convenient to start the description by the second approach.

2. Predetermined left-right grid with randomly assigned positions. For each country and each political election, we consider each political party gaining seats in the Parliament and classify its political position as one of the following: Far-Left (FL), Left to Far-Left (LFL), Left (L), Centre-Left to Left (CLL), Centre-Left (CL), Centre to Centre-Left (CCL), Centre (C), Centre to Centre-Right (CCR), Centre-Right (CR), Centre-Right to Right (CRR), Right (R), Right to Far-Right (RFR), Far-Right (FR), Big Tent (BT). Each of these labels X (except BT) occupies an ordered position q_X in the interval $[q_{\text{FL}}, q_{\text{FR}}] = [-0.5, 0.5]$, i.e., we have the following grid of ordered coarse-grained political positions between FL and FR:

$$[q_{\text{FL}} \ q_{\text{LFL}} \ q_L \ q_{\text{CLL}} \ q_{\text{CL}} \ q_{\text{CCL}} \ q_C \ q_{\text{CCR}} \ q_{\text{CR}} \ q_{\text{CRR}} \ q_R \ q_{\text{RFR}} \ q_{\text{FR}}], \quad (10)$$

see also Fig. 8b.

We assume that the left-right scale is symmetric around the central position $q_C = 0$. The specific values of q_{CCR} , q_{CR} , q_{CRR} , q_R , q_{RFR} are chosen randomly, sorting 5 values drawn from a uniform probability distribution in the interval $[0, 0.5]$. The values of q_{LFL} , q_L , q_{CLL} , q_{CL} and q_{CCL} are then obtained by symmetry, see Fig. 8b. Big tent parties do not fit into such a left-right grid, because they tend to attract voters from the entire ideological spectrum. Examples of big tent parties are the Italian party *Movimento 5 Stelle* (5 Star Movement), in the 2018 elections. Consequently, the distance d_{BT} between a big tent party and any other party must not be “too small” nor “too big” compared with all the possible differences $|q_{p_i} - q_{p_j}|$ for all non big tent parties p_i and p_j : d_{BT} is then chosen as the median of the differences, taken in absolute value, between all possible combinations of the pairwise distances of the positions $q_{\text{FL}}, \dots, q_{\text{FR}}$,

$$d_{\text{BT}} = \text{median}\{|q_{\text{FL}} - q_{\text{LFL}}|, |q_{\text{FL}} - q_L|, \dots, |q_{\text{FL}} - q_{\text{FR}}|, \dots, |q_{\text{RFR}} - q_{\text{FR}}|\}.$$

With these conventions we can proceed to assigning numerical values to the weights.

Since we take into account also pre-electoral coalitions, we need to assign weights that are positive to parties in the same coalition. The weights w_{ij} , $i, j = 1, \dots, n_p$, are then chosen with the following rule:

- if $p_i = p_j$ (same party): $w_{ij} = 1$;

- if $p_i \neq p_j$ (different party):

$$w_{ij} = \text{coal}(p_i, p_j) - \begin{cases} \frac{1}{2} \cdot \min_{\substack{x,y \in \{C, CCR, \dots, FR\} \\ s.t. q_x \neq q_y}} |q_x - q_y| & q_{p_i} = q_{p_j} \\ & (\text{parties placed} \\ & \quad \text{on the same position}) \\ |q_{p_i} - q_{p_j}|, & q_{p_i} \neq q_{p_j}; p_i, p_j \text{ not big tent} \\ & (\text{parties on different} \\ & \quad \text{positions, none is big tent}) \\ d_{BT}, & q_{p_i} \neq q_{p_j}; p_i \text{ or } p_j \text{ big tent} \\ & (\text{parties on different positions,} \\ & \quad \text{one of them is big tent}) \end{cases} \quad (11)$$

The function $\text{coal}(p_i, p_j)$ is equal to 1 if p_i and p_j are in electoral coalition, 0 otherwise.

Notice that when two parties p_i and p_j are not in electoral coalition but are located in the same position (i.e., $q_{p_i} = q_{p_j}$), then we still assume that the w_{ij} weight is negative, but small (equal to half the least nonzero difference between any two parties). The weights are kept constant throughout the party history.

To study how different choices of the values q_x can influence our analysis, we decided to consider 10000 different vectors of random values for each country, retaining the best value (i.e., the value that leads to the highest correlation between frustration and negotiation days, see below for more details).

1. **Rile index from Manifesto Project Database.** For each country and parliamentary election, the Manifesto Project Database [44] collects information on the electoral manifestos of the parties. The index denoted *rile* [33] summarizes their ideology according to various criteria (e.g., its position on economy, military, international relations, education, welfare, etc.). The political position of party p_i is given by the rile value of the party, properly rescaled to fit our $[-0.5, 0.5]$ normalization: $q_{p_i} = \text{rile}(p_i)$. In case of missing rile for a party p_i , its position q_{p_i} is determined using the method discussed above, only considering a vector (10) which is equispaced in $[-0.5, 0.5]$ (i.e., the random assignment to the q_x value is skipped). The weight matrix W is then given by the formula (11).

In the paper we consider three different combinations of edge weight assignments and party grouping methods listed in Fig. 1a.

A.2 Structural balance and frustration for signed networks

Consider a signed, undirected, simple and connected network $\mathcal{G} = (\mathcal{V}, \mathcal{E}, A)$. The *normalized signed Laplacian* of \mathcal{G} is defined as $\mathcal{L} = I - \Delta^{-1}A$ where the diagonal matrix $\Delta = \text{diag}\{\delta_1, \dots, \delta_n\}$ has elements $\delta_i = \sum_{j=1}^n |a_{ij}| > 0$, $i = 1, \dots, n$.

A signed network is *structurally balanced* if all its cycles are positive, i.e., if each cycle contains an even number of negative edges: in a social network context, every length-3 signed cycle of a structurally balanced network describes one of

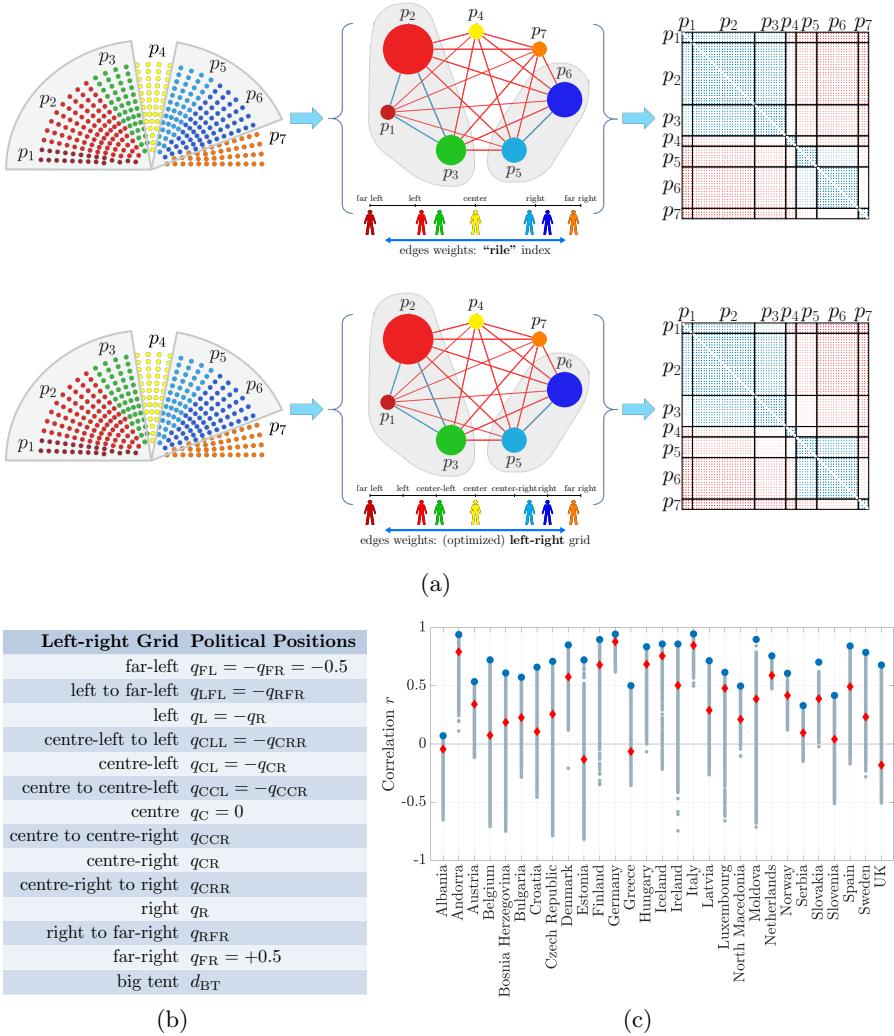


Figure 8: (a): Constructing a parliamentary network and the corresponding adjacency matrix for scenario **II** (top) and **III** (bottom). Scenario **I** is shown in Fig. 1b. (b): "Political positions" in the left-right political spectrum. (c): Correlation r for scenario **III**. 10000 sets of values for the left-right political positions of the parties were randomly selected on the preassigned left-right grid, as explained in Section A.1. Here the corresponding correlation values for each set are shown (gray circles), together with the overall country mean value (red diamond) and maximum value (blue circle). For each country, the optimal choice for the weights corresponds to the one giving best correlation index (blue circle).

the following concepts: "the friend of my friend is my friend", "the enemy of my friend is my enemy", "the friend of my enemy is my enemy", "the enemy of my

enemy is my friend”, see [26] for more details. The notion of structural balance captures the idea that it is possible to split a graph into two subgraphs such that all edges on each subgraph are positive, while all edges through the cut set that splits the graph are negative. In our parliamentary network it could represent a two-party parliament or, in the electoral coalition scenario, a parliament split into two coalitions. Equivalent conditions to structural balance are (i) $\lambda_1(\mathcal{L}) = 0$, and (ii) there exists a signature matrix $S = \text{diag}\{s_1, \dots, s_n\}$, with $s_i = \pm 1$, such that $S\mathcal{L}S$ has all nonpositive off-diagonal entries [60]. It follows that a network \mathcal{G} is structurally unbalanced if and only if $\lambda_1(\mathcal{L}) > 0$.

When a network is not structurally balanced, it is of interest to understand how “far” it is from a structurally balanced state. One idea is to use $\lambda_1(\mathcal{L})$ (the so-called “algebraic conflict” [37, 61]), which is strictly positive for structurally unbalanced networks, to measure such a distance; however, in the literature another standard measure, called *frustration index*, is more frequently adopted [24, 26, 62]. It is defined as the minimum (weighted, if \mathcal{G} is a weighted graph) sum of the positive edges over all signature similarity transformations of \mathcal{L} , $S\mathcal{L}S$, with S signature matrix [32]:

$$\epsilon(\mathcal{G}) = \min_{\substack{S=\text{diag}\{s_1, \dots, s_n\} \\ s_i=\pm 1}} \frac{\sum_{i,j \neq i} [|\mathcal{L}| + S\mathcal{L}S]_{ij}}{2}. \quad (12)$$

The computation of $\epsilon(\mathcal{G})$ constitutes a NP-hard problem: however, the intuition is that $\lambda_1(\mathcal{L})$ approximates well (up to a scaling factor) the value of the frustration index, and in particular that $\lambda_1(\mathcal{L})$ grows linearly with $\epsilon(\mathcal{G})$ [32]. Hence, they can both be used to measure the structural imbalance of a signed network.

The frustration index, as defined in (12), is also the minimum of a weighted energy functional over all possible signature matrices S . This terminology is inherited from Statistical Physics, where a (unweighted) signed graph is interpretable as an Ising spin glass, and the various spin configurations (“spin up” and “spin down” at the nodes) determine the energy of the spin glass. The least energy that can be achieved by any configuration (called the ground state) corresponds to the frustration index (12). The definition of energy functional introduced in [26] for this purpose can be adapted to networks that are weighted. The weighted energy functional can be defined as follows,

$$e(S) = \frac{\sum_{i,j} (|h_{ij}| - h_{ij}s_i s_j)}{2}, \quad (13)$$

where S is a signature matrix, $S = \text{diag}\{s_1, \dots, s_n\}$ with $s_i = \pm 1 \forall i$, and $H = [h_{ij}] = \Delta^{-1}A$ is the normalized adjacency matrix. Notice that $e(S) = e(-S)$. Notice further that by definition of normalized signed Laplacian, it follows that

$$e(S) = \frac{\sum_{i,j} [|H| - SHS]_{ij}}{2} = \frac{\sum_{i,j \neq i} [|\mathcal{L}| + S\mathcal{L}S]_{ij}}{2}.$$

Hence, the weighted frustration index is the minimum of the weighted energy functional over all possible signature matrices S (i.e., again, the ground state):

$$\epsilon(\mathcal{G}) = \min_{\substack{S=\text{diag}\{s_1, \dots, s_n\} \\ s_i=\pm 1}} e(S).$$

As described above, each network $\mathcal{G}_{\text{country}}$ can be partitioned into n_p clusters (the political parties), with n_p being the number of parties involved and gaining seats in the elections, making the adjacency matrix A a $n_p \times n_p$ blocks matrix. Under the assumption that all MPs of a party follow loyally and unanimously the designated party line, the definition of frustration index given in (12) can be specialized to a party-wise (or cluster-wise) frustration index, consisting of the minimum of all the energy functionals given by (13) over all block diagonal signature matrices S :

$$\zeta = \min_{\substack{S=\text{diag}\{S_1, \dots, S_{n_p}\} \\ S_i=s_i \cdot I_{c_i}, s_i=\pm 1}} e(S) \quad (14)$$

where c_i , $i = 1, \dots, n_p$, is the number of seats gained by the party p_i at the election. The difference with (12) is that now the signature matrix S is a block diagonal matrix (with n_p blocks): only the relationships between different parties, and not the ones between single MPs, are taken into account. Computing (14) is much faster than computing (12). In fact, all energy levels of our parliamentary networks can be exhaustively explored in a systematic way. The resulting energy landscape can be represented as histograms, having in the leftmost point the ground state energy. See Fig. 11 for the 29 countries of Table 2 in one of the scenarios discussed in the paper (scenario I).

A.3 Dynamical model of decision-making in presence of frustration

Let $\mathcal{G} = (\mathcal{V}, \mathcal{E}, A)$ be a signed network whose node set \mathcal{V} represents a community of agents. To represent a process of decision-making for these agents, we consider the following nonlinear interconnected dynamical model, previously used in [30, 31, 63],

$$\dot{x} = -\Delta x + \pi A \psi(x), \quad x \in \mathbb{R}^n, \quad (15)$$

where the state vector $x \in \mathbb{R}^n$ represents the decisions of the agents and $A \in \mathbb{R}^{n \times n}$ is the adjacency matrix of the network \mathcal{G} . Each element of the diagonal matrix $\Delta = \{\delta_1, \dots, \delta_n\}$ is given by $\delta_i = \sum_{j=1}^n |a_{ij}|$, $i = 1, \dots, n$, while π is a scalar positive parameter representing the “social effort” or strength of the commitment among the agents. $\psi(x) = [\psi_1(x_1) \dots \psi_n(x_n)]^T$ are sigmoidal nonlinear functions. We assume that each function $\psi_i(x_i) : \mathbb{R} \rightarrow \mathbb{R}$ is monotone, i.e., $\frac{\partial \psi_i}{\partial x_i}(x_i) > 0 \forall x_i \in \mathbb{R}$, with unit slope at the origin, $\frac{\partial \psi_i}{\partial x_i}(0) = 1$, and saturated behavior: $\lim_{x_i \rightarrow \pm\infty} \psi_i(x_i) = \pm 1$. From previous works, such as [31, 32], we know that the existence of equilibrium points is determined by the social effort parameter π : when $\pi < \pi_1$ the origin is the unique (and globally asymptotically stable) equilibrium point of the system (15). When $\pi = \pi_1$ the system (15) undergoes a pitchfork bifurcation (i.e., the number of equilibria “jumps” from one to three) and for $\pi > \pi_1$ two alternative equilibrium points x^* and $-x^*$ appear, which are locally asymptotically stable, while the origin becomes unstable. Finally, when $\pi = \pi_2$ the system (15) undergoes a second pitchfork bifurcation and for $\pi > \pi_2$ it admits multiple equilibria. In the context

Threshold values	Structurally balanced \mathcal{G}	Structurally unbalanced \mathcal{G}
$\pi_1 = \frac{1}{1-\lambda_1(\mathcal{L})}$	= 1, fixed	grows with the frustration $\epsilon(\mathcal{G})$
$\pi_2 = \frac{1}{1-\lambda_2(\mathcal{L})}$	depends on the algebraic connectivity of \mathcal{G} , $\lambda_2(\mathcal{L})$	independent from the frustration $\epsilon(\mathcal{G})$

Table 3: Description of the bifurcation behavior of the system (15).

of social networks, where each equilibrium point corresponds to a possible decision made by the agents, the behavior of the system can be interpreted as follows: when $\pi < \pi_1$ the social effort of the agents is not enough to reach a nontrivial decision. When π grows past the first threshold value π_1 , the higher strength of commitment among the agents leads to two possible alternative decisions. Finally, when $\pi > \pi_2$, the “overcommitment” of the agents leads to a situation in which multiple decisions are possible. We are interested in the case when π belongs to the interval $[0, \pi_2]$, since the agents have to choose among no nontrivial decision (in $[0, \pi_1]$) and among two alternative nontrivial decisions (in $[\pi_1, \pi_2]$).

The threshold values π_1 and π_2 are functions of the two smallest eigenvalues of the normalized signed Laplacian of the network, $\mathcal{L} = I - \Delta^{-1}A$: $\pi_1 = \frac{1}{1-\lambda_1(\mathcal{L})}$ and $\pi_2 = \frac{1}{1-\lambda_2(\mathcal{L})}$. Moreover, at $\pi = \pi_1$ the nontrivial equilibria $\pm x^*$ appear on $\text{span}\{v_1(\mathcal{L})\}$, where $v_1(\mathcal{L})$ is the eigenvector relative to $\lambda_1(\mathcal{L})$, see [32]. The values of $\lambda_1(\mathcal{L})$ and $\lambda_2(\mathcal{L})$ depend not only on the network structure but also on its *signature*: as explained in Section A.2, $\lambda_1(\mathcal{L}) = 0$, and consequently π_1 is fixed and constant to $\pi_1 = 1$, if and only if the network \mathcal{G} is structurally balanced. When this is not the case, $\lambda_1(\mathcal{L})$ grows with the frustration of the network, implying that also π_1 increases. This means that networks which are structurally unbalanced need a higher social effort (π_1) from the agents in order to converge to a nontrivial decision. Regarding π_2 , its value depends on $\lambda_2(\mathcal{L})$, which is the algebraic connectivity of the network in the structurally balanced case. If the network is structurally unbalanced, unlike $\lambda_1(\mathcal{L})$, $\lambda_2(\mathcal{L})$ remains nearly independent from the frustration of the network, i.e., even if $\epsilon(\mathcal{G})$ grows, $\lambda_2(\mathcal{L})$ remains almost constant. This means that the interval $[\pi_1, \pi_2]$ in which nontrivial decisions appear for the social effort parameter π shrinks as the frustration of the network $\epsilon(\mathcal{G})$ increases. These results, shown in [32], can be summarized as in Table 3.

As introduced in Section A.1, each network $\mathcal{G}_{\text{country}}$ considered in this work is composed of n_p clusters, where n_p is the total number of parliamentary political parties. Nonetheless, we still expect a behavior similar to the one described in Table 3 for the first threshold value π_1 also for the party-wise frustration of the network, whose definition was given in (14).

This can be shown with a numerical example, where we consider a network $\mathcal{G} = (\mathcal{V}, \mathcal{E}, A)$ with $n = \text{card}(\mathcal{V}) = 500$ individuals and n_p clusters, representing an all-against-all scenario (scenario I), whose adjacency matrix A is hence a $n_p \times n_p$ block matrix described by (9), where $w_{ii} = 1$ for all $i = 1, \dots, n_p$ and $w_{ij} = -1$ for all $i, j = 1, \dots, n_p$ and $j \neq i$. We decided to vary the number of parties and, for each

n_p , to consider 1000 (unique) randomly selected vectors $[c_1, \dots, c_{n_p}]$ s.t. $c_i \in [1, n]$ and $\sum_{i=1}^{n_p} c_i = n$, where c_i is the size of each party, $i = 1, \dots, n_p$. Figure 9 shows the results for $n_p = 3, 4, \dots, 20$. Figure 9a illustrates how factors such as the number of political parties and the size of each party i (c_i) influence the frustration: an increase in the number of parties (n_p , left panel) or a decrease in the maximum number of MPs per party ($\max_i c_i$, right panel) both lead to an increase in the frustration in average, although the variance is extremely large. Instead, Fig. 9b shows that as the frustration of the network increases so does the threshold value π_1 . As a consequence, a higher social effort will be required from the agents to achieve a nontrivial decision.

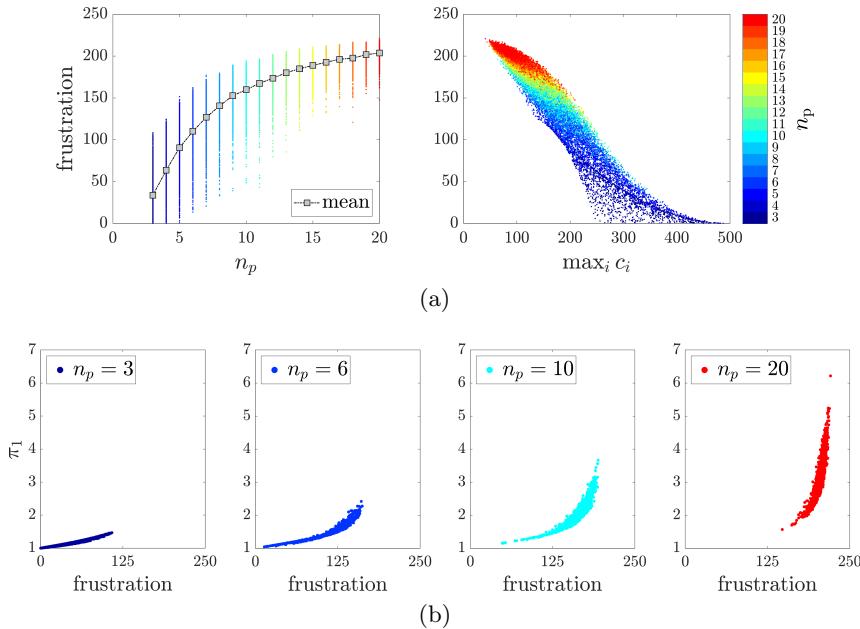


Figure 9: Numerical example of computation of the frustration as the number of parties and of seats per party varies. (a): Party-wise frustration ζ of the network \mathcal{G} vs number of parties (n_p) and vs maximum number of MPs per party ($\max_i c_i$), as the number of parties is changed and the number of seats per party is varied randomly. The artificial networks we consider here are “all-against-all” networks with $n_p \in \{3, 4, \dots, 20\}$ parties and size $n = 500$. (b): Behavior of π_1 as a function of the frustration ζ , as the number of seats per party changes randomly and $n_p \in \{3, 6, 10, 20\}$.

B Application: from parliamentary networks to government formation

For a given country and general election, we aim to use our parliamentary networks to predict:

1. The duration of the negotiation phase that leads to the formation of a post-electoral cabinet.
2. The composition of the governmental coalition that sustains such a successful post-electoral cabinet.

In both cases we are interested only in the cabinet that is being formed immediately after a general election. Inter-election government formation processes are often following different rules, see [5, 6].

B.1 Frustration vs government negotiation days

We are interested in the possible correlation between the frustration of each network $\mathcal{G}_{\text{country}}$, measured by ζ (formula (14)), and the numbers of days between the election date and the date the government is sworn in. The *rationale* behind this is that when no clear majority has emerged from the electoral ballot, there is uncertainty in the composition of the candidate cabinet, and the political parties require more time in order to overcome their differences and tensions if they want to establish coalitions which can ensure a majority in the parliament. Mathematically, the link between the two properties is given by the dynamical model described in Section A.3. Lack of a clear electoral winner corresponds to a parliamentary network with high frustration index ζ .

In our model (15), a high frustration raises the value of the bifurcation point π_1 , meaning that achieving a nontrivial decision (i.e., giving a confidence vote to a government) requires a high “social effort”, here interpreted as duration of the negotiation process among the parties. A graphical representation can be seen in Fig. 1c.

To check if this hypothesis is valid for our data, we compute the Pearson correlation (r) between ζ and the number of days between the general election and the date the government is sworn in. The higher the value for r (which ranges between -1 and 1), the “stronger” the evidence that frustration indeed influences the dynamics of the government formation process. The resulting values are shown in Fig. 2a.

B.2 Minimum energy government coalition

For each parliamentary network $\mathcal{G}_{\text{country}} = (\mathcal{V}, \mathcal{E}, A)$, the energy landscape is obtained by computing the energies in all the 2^{n_p} possible block signature matrices $S = \text{diag}\{S_1, \dots, S_{n_p}\}$, with $S_i = s_i \cdot I_{c_i}$ and $s_i = \pm 1 \forall i$. The party-wise frustration

index is the minimum of such energies, as computed in (14). Denote

$$\begin{aligned} S_{\text{best}} &= \arg \min_S e(S) \\ \text{s.t. } (i) \quad &S = \text{diag}\{s_1 \cdot I_{c_1}, \dots, s_{n_p} \cdot I_{c_{n_p}}\}, s_i = \pm 1, \\ (ii) \quad &\sum_{i: s_i=+1} c_i \geq \sum_{i: s_i=-1} c_i. \end{aligned} \quad (16)$$

the block diagonal signature matrix (with more diagonal elements equal to +1 than -1) which gives the minimum of the energy functional, i.e., $\zeta = e(S_{\text{best}})$. Consequently, $-S_{\text{best}}$ is a block diagonal signature matrix with more diagonal elements equal to -1 than +1; observe that $e(-S_{\text{best}}) = e(S_{\text{best}}) = \zeta$. In the paper, S_{best} is associated with “success” of a confidence vote to form a government cabinet, and $-S_{\text{best}}$ to “failure”. Such ground state can in general be degenerate (that is, several pairs $\pm S_{\text{best}}$ may give the same frustration index ζ).

Let $S_{\text{best}} = \text{diag}\{s_{1,\text{best}} \cdot I_{c_1}, \dots, s_{n_p,\text{best}} \cdot I_{c_{n_p}}\}$. The two party sets $\{p_i : s_{i,\text{best}} = +1\}$ and $\{p_i : s_{i,\text{best}} = -1\}$, where p_i is the i -th party, form a minimum energy partition of the set $\mathcal{P} = \{p_i, i = 1, \dots, n_p\}$ corresponding to (the positive and negative diagonal elements of) the block diagonal signature matrix S_{best} .

For the case of “success” (the vast majority of our data), the intuition is that the set $\mathcal{P}_{\text{best}} := \{p_i \in \mathcal{P} : s_{i,\text{best}} = +1\}$, which corresponds to a majority coalition of parties, should contain the possible new government coalition, plus the parties that support it in parliament without officially participating to it. In other words, $\mathcal{P}_{\text{best}}$ should be a superset of the set $\mathcal{P}_{\text{gov}} = \{p_i \in \mathcal{P} : p_i \in \text{gov}\}$, i.e., the post-electoral cabinet that actually took place.

Let $S_{\text{gov}} = \text{diag}\{s_{1,\text{gov}} I_{c_1}, \dots, s_{n_p,\text{gov}} I_{c_{n_p}}\}$ be a block diagonal signature matrix whose elements $s_{i,\text{gov}}$, $i = 1, \dots, n_p$, are defined as follows,

$$s_{i,\text{gov}} = \begin{cases} +1, & p_i \in \mathcal{P}_{\text{gov}} \\ -1, & p_i \notin \mathcal{P}_{\text{gov}}. \end{cases}$$

To evaluate how our predictions (S_{best}) reflect the actually formed government, we introduce the following indexes:

$$\rho_{\text{gov}} = \frac{\text{card}(\mathcal{P}_{\text{best}} \cap \mathcal{P}_{\text{gov}})}{\text{card}(\mathcal{P}_{\text{gov}})} \quad (17)$$

$$\eta_{\text{gov}} = 1 - \frac{e(S_{\text{gov}}) - \zeta}{\max_S e(S) - \zeta}. \quad (18)$$

The closer the value of ρ_{gov} is to 1, the better our estimate represents the actual cabinet composition. η_{gov} represents instead how close “energetically” our guess (i.e., ζ) is to the true government energy $e(S_{\text{gov}})$. In (18), $e(S)$ and ζ are given by (13) and (14) respectively, $e(S_{\text{gov}})$ is the energy in correspondence of S_{gov} , and S in $e(S)$ is used as in (14) to indicate a signature matrix with n_p blocks, i.e., $S = \text{diag}\{s_1 I_{c_1}, \dots, s_{n_p} I_{c_{n_p}}\}$, $s_i = \pm 1 \forall i$.

As already mentioned, several $\pm S_{\text{best}}$ may give the same frustration index ζ (i.e., the ground state is degenerate). Moreover, among those, there might

exist signature matrices S_{best} determining party sets $\{p_i \in \mathcal{P} : s_{i,\text{best}} = +1\}$ and $\{p_i \in \mathcal{P} : s_{i,\text{best}} = -1\}$ with equal total number of seats, $\sum_{i:s_{i,\text{best}}=+1} c_i = \sum_{i:s_{i,\text{best}}=-1} c_i = \frac{n}{2}$ (equality holds in (16)(ii)). We consider the latter as inconclusive cases, meaning that our analysis is not able to identify the possible government. More specifically, these issues are handled as follows.

1. In presence of multiple S_{best} and multiple valid values for ρ_{gov} , we consider the signature matrix S_{best} giving maximum possible value for ρ_{gov} .
2. In presence of multiple S_{best} all representing inconclusive cases for the analysis, we define $S'_{\text{best}} = \text{diag}\{s'_{1,\text{best}} \cdot I_{c_1}, \dots, s'_{n_p,\text{best}} \cdot I_{c_{n_p}}\}$ as the signature matrix giving minimum possible energy while satisfying a “non-inconclusive” condition, $\sum_{i:s'_{i,\text{best}}=+1} c_i > \sum_{i:s'_{i,\text{best}}=-1} c_i$, that is

$$\begin{aligned} S'_{\text{best}} = \arg \min_S e(S) \\ \text{s.t. (i)} \quad S = \text{diag}\{s_1 I_{c_1}, \dots, s_{n_p} I_{c_{n_p}}\}, s_i = \pm 1; \\ \text{(ii)} \quad \sum_{i:s_i=+1} c_i > \sum_{i:s_i=-1} c_i. \end{aligned} \tag{19}$$

Then, $\mathcal{P}_{\text{best}} := \{p_i \in \mathcal{P} : s'_{i,\text{best}} = +1\}$. If S'_{best} is degenerate, we follow point 1. Fig. 15 illustrates how often degenerate or inconclusive cases occur in scenario I.

B.3 Frustration vs smallest eigenvalue of the normalized signed Laplacian

Our analysis is based on two observations. The first, already mentioned in Section A.3, is a well-studied fact from the literature [32, 37, 61], namely that the smallest eigenvalue of the normalized signed Laplacian ($\lambda_1(\mathcal{L})$) grows linearly with the frustration of the network (ζ). The second is that there is a high overlap between the signature of the eigenvector relative to $\lambda_1(\mathcal{L})$ ($v_1(\mathcal{L})$) and the signature matrix corresponding to the ground state, which we denote $S_{\text{best}} = \text{diag}\{s_1 I_{c_1}, \dots, s_{n_p} I_{c_{n_p}}\}, s_i = \pm 1$ (16). Notice that since the adjacency matrix of a parliamentary network is a block matrix, see (9), then $v_1(\mathcal{L})$ is also a block vector, $v_1(\mathcal{L}) = [v_{1,1} \mathbf{1}_{c_1}^T \ v_{1,2} \mathbf{1}_{c_2}^T \ \cdots \ v_{1,n_p} \mathbf{1}_{c_{n_p}}^T]^T$.

To check if the first hypothesis holds for our data, we calculate the Pearson correlation between the smallest eigenvalue of the normalized signed Laplacian, $\lambda_1(\mathcal{L})$, and the frustration of the parliamentary networks, ζ . To check if the second assumption is satisfied, we calculate the overlap between $\pm S_{\text{best}}$ and the signature of $v_1(\mathcal{L})$, defined as

$$\begin{aligned} \% \text{ overlap} &= \max \left\{ 1 - \frac{\|\text{sign}(v_1(\mathcal{L})) - S_{\text{best}}\|_1}{n}, 1 - \frac{\|\text{sign}(v_1(\mathcal{L})) + S_{\text{best}}\|_1}{n} \right\} \cdot 100 \\ &= \max \left\{ 1 - \frac{\sum_{i=1}^{n_p} c_i |\text{sign}(v_{1,i}) - s_i|}{n}, 1 - \frac{\sum_{i=1}^{n_p} c_i |\text{sign}(v_{1,i}) + s_i|}{n} \right\} \cdot 100, \end{aligned} \tag{20}$$

where $n = \sum_{i=1}^{n_p} c_i$ is the total number of MPs and $\|\cdot\|_1$ indicates the 1-norm.

The average values of correlation between $\lambda_1(\mathcal{L})$ and the frustration of the parliamentary networks ζ , and the average overlap (20) between S_{best} and the signature of $v_1(\mathcal{L})$ (the eigenvector related to $\lambda_1(\mathcal{L})$) are reported in Table 4 for each scenario **(I, II, III)**.

Scenario	Correlation between $\lambda_1(\mathcal{L})$ and frustration	Overlap (average) between S_{best} and sign($v_1(\mathcal{L})$)
I	0.956	88.31 %
II	0.939	98.12 %
III	0.904	97.69 %

Table 4: Frustration vs smallest eigenvalue of the normalized signed Laplacian for scenario **I, II, III**. Left column: average correlation between the frustration of the parliamentary networks ζ and the smallest eigenvalue of the normalized signed Laplacian $\lambda_1(\mathcal{L})$. Right column: average overlap between S_{best} and the signature of $v_1(\mathcal{L})$, the eigenvector associated to $\lambda_1(\mathcal{L})$, as defined in equation (20).

B.4 Frustration vs fractionalization index

As shown in previous studies such as [5, 8] (and as intuitively clear), parliamentary fragmentation is one of the main factors influencing the duration of the government negotiation process. Such fragmentation is often measured in terms of the Laakso-Taagepera effective number of parties [11], or in terms of the strictly related fractionalization index [12]. In this section we investigate (analytically) how the frustration of the signed parliamentary networks (of scenario **I**) is related to the fractionalization index (which carries the same information of the effective number of parties).

The effective number of parties (denoted N_2) is defined as the inverse of the sum of squares of the shares of seats of each party, see (3), while the fractionalization index (denoted F) is defined as $F = 1 - \frac{1}{N_2}$.

To compare the frustration of the signed networks to these indexes, we first need to rewrite the expression of the weighted energy functional $e(S)$, see (13), by taking into consideration that the parliamentary networks, as introduced in Section A.1 and represented by $\mathcal{G}_{\text{country}}$, can be partitioned in n_p clusters (corresponding to the political parties), each of size c_i ($i = 1, \dots, n_p$). Indeed, this means that the adjacency matrix A (see (9)) and the diagonal matrix $\Delta = \text{diag}\{|A|\mathbf{1}\}$ (introduced in Section A.2) are block matrices. In particular, $\Delta = \text{diag}\{\delta_1 I_{c_1}, \dots, \delta_{n_p} I_{c_{n_p}}\}$ where each block $\delta_i I_{c_i}$, $i \in \mathcal{I} := \{1, \dots, n_p\}$, satisfies

$$(\delta_i I_{c_i})\mathbf{1}_{c_i} = \sum_{j \in \mathcal{I}} |A_{ij}| \mathbf{1}_{c_j} = \sum_{\substack{j \in \mathcal{I} \\ j \neq i}} |w_{ij}| E_{c_i c_j} \mathbf{1}_{c_j} + (E_{c_i} - I_{c_i})\mathbf{1}_{c_i} = \left(\sum_{j \in \mathcal{I}} |w_{ij}| c_j - 1 \right) \mathbf{1}_{c_i}.$$

Then the weighted energy functional $e(S)$, where $S = \text{diag}\{s_1 I_{c_1}, \dots, s_{n_p} I_{c_{n_p}}\}$ with $s_i = \pm 1 \forall i \in \mathcal{I}$, can be rewritten as follows:

$$e(S) = \frac{1}{2} \mathbf{1}_n^T \Delta^{-1} (|A| - SAS) \mathbf{1}_n = \frac{1}{2} \sum_{i,j \in \mathcal{I}} \frac{c_i c_j}{\delta_i} (|w_{ij}| - s_i w_{ij} s_j), \quad (21)$$

where w_{ij} is the weight between the i -th and j -th party.

In scenario **I** all party-party weights are negative and equal to -1 ($w_{ij} = -1$ for all $j \neq i$), which implies that $\delta_i = n - 1$ for all $i \in \mathcal{I}$ and that (21) becomes

$$e(S) = \frac{1}{2(n-1)} \sum_{i,j \in \mathcal{I}, j \neq i} c_i c_j (1 + s_i s_j) = \frac{1}{n-1} \cdot \sum_{\substack{i,j \in \mathcal{I}, j \neq i \\ s.t. s_i s_j > 0}} c_i c_j.$$

Let $\mathcal{C}_+ = \{i \in \mathcal{I} : s_i = +1\}$ and $\mathcal{C}_- = \{i \in \mathcal{I} : s_i = -1\}$ be the two node subsets defined by a signature matrix $S = \text{diag}\{s_1 I_{c_1}, \dots, s_{n_p} I_{c_{n_p}}\}$, and let $n_{\mathcal{C}_+} = \sum_{i \in \mathcal{C}_+} c_i$ and $n_{\mathcal{C}_-} = \sum_{i \in \mathcal{C}_-} c_i$. Notice that $\mathcal{C}_+ \cap \mathcal{C}_- = \emptyset$ and that $n_{\mathcal{C}_+} + n_{\mathcal{C}_-} = n$. Then, the calculation of the frustration of a signed parliamentary network in scenario **I** reduces to

$$\begin{aligned} \zeta &= \frac{1}{n-1} \cdot \min_{\substack{\text{diag}\{s_1, \dots, s_{n_p}\} \\ s_i = \pm 1 \\ s.t. s_i s_j > 0}} \sum_{i,j \in \mathcal{I}, j \neq i} c_i c_j = \frac{1}{n-1} \cdot \min_{\mathcal{C}_+ \subseteq \mathcal{I}} \left\{ \sum_{\substack{i,j \in \mathcal{C}_+ \\ j \neq i}} c_i c_j + \sum_{\substack{i,j \in \mathcal{I} \setminus \mathcal{C}_+ \\ j \neq i}} c_i c_j \right\} \\ &= \frac{1}{n-1} \cdot \left(\min_{\mathcal{C}_+ \subseteq \mathcal{I}} \left\{ \sum_{i,j \in \mathcal{C}_+} c_i c_j + \sum_{i,j \in \mathcal{I} \setminus \mathcal{C}_+} c_i c_j \right\} - \sum_{i \in \mathcal{I}} c_i^2 \right) \\ &= \frac{1}{n-1} \cdot \left(\min_{\mathcal{C}_+ \subseteq \mathcal{I}} \left(n_{\mathcal{C}_+}^2 + (n - n_{\mathcal{C}_+})^2 \right) - \sum_{i \in \mathcal{I}} c_i^2 \right) \\ &= \frac{1}{n-1} \cdot \left(2 \min_{\mathcal{C}_+ \subseteq \mathcal{I}} \left(n_{\mathcal{C}_+}^2 - n \cdot n_{\mathcal{C}_+} \right) + n^2 - \sum_{i \in \mathcal{I}} c_i^2 \right) \\ &= \frac{n^2}{n-1} \cdot \left(-2 \max_{\mathcal{C}_+ \subseteq \mathcal{I}} \left(\frac{n_{\mathcal{C}_+}}{n} - \frac{n_{\mathcal{C}_+}^2}{n^2} \right) + \underbrace{1 - \frac{\sum_{i \in \mathcal{I}} c_i^2}{n^2}}_{=1 - \frac{1}{N_2} = F} \right). \end{aligned} \tag{22}$$

Equation (22) shows that the frustration of the signed networks of scenario **I** is proportional to the difference between the fractionalization index F and a term which is related to the size of the minimum winning coalition. To obtain some insight on the frustration of scenario **I** (and in particular on the first term of equation (22)), let $\mathcal{P}_{\text{best}}$ be the solution of the minimization - or, maximization, depending on the sign - problem in (22) (see also Section B.2) and $n_{\mathcal{P}_{\text{best}}} = \frac{n}{2} + E_{\text{best}}$ where $E_{\text{best}} \in [0, \frac{n}{2}]$ is the *number of seats in excess* (with respect to 50% of the total number of seats). Then the frustration (up to a constant multiplicative term) can be written as (4).

In conclusion, equation (4) shows that the frustration of the signed parliamentary networks of scenario **I** is proportional to the frustration index F and to the “distance” of the the majority $\mathcal{P}_{\text{best}}$ (in correspondence of minimum winning coalition S_{best}) to 50% of the total number of seats. Figure 16 shows the linear regression plots between the fractionalization index and the frustration of the parliamentary networks for each country and election. The average value of correlation is 0.99 (see Fig. 2d). Moreover, Fig. 16 shows that countries with a lower value of correlation (such as Albania, North Macedonia, Moldova, UK) are

characterized by minimum winning coalitions S_{best} with higher excess of seats E_{best} .

B.5 Description of the results

A summary of the results obtained for all the scenarios mentioned in Fig. 1a, in terms of correlation between duration of government negotiations and frustration, indexes evaluating how good our estimates of the cabinet composition are, and correlation between fractionalization index and frustration is given in Fig. 2.

For scenario **I**, the number of government negotiation days and the frustration for each country are shown in Fig. 10a and Fig. 10b, respectively. The percentages of correct government predictions, in terms of ρ_{gov} and η_{gov} , are given in Fig. 10c and Fig. 10d, respectively. Fig. 11 shows the energy functionals for each country and election year; the red line represents the energy corresponding to the government coalition, $e(S_{\text{gov}})$.

Regression plots (see Section B.6 for more details on the linear regression analysis) between the government negotiation days and the frustration of the parliamentary networks with the corresponding values of correlation are shown in Fig. 3, Fig. 12 and Fig. 13, respectively for scenario **I**, **II** and **III**.

The correlation r_{country} computed for each country and scenario, between the duration of coalition negotiations and the party-wise frustration of the network, is given in Fig. 2a: overall, the values of correlation for all scenarios are positive and above 0.4 with few exceptions which often admit a clear explanation and will be discussed in Section B.7. In scenario **III** we consider 10000 different values for the political positions (fixed in time: the same for all elections) in the left-right scale, as explained in Section A.1: for each country, the “optimal” choice of positions is identified as the one giving the best value of correlation, depicted in Fig. 8c with a blue circle. This value is then used in Fig. 2 (scenario **III**) and Fig. 13. As can be seen in Fig. 8c, the positions we assign to the left-right grid significantly affect the correlation, hence the optimal r_{country} may be far from the average of the correlations obtained for each country. This could also explain why the results obtained in scenario **II** are worse, in terms of correlation, than the ones obtained in scenario **I**, where the networks we consider are unweighted.

B.6 Influential points of the regression are important

Consider the “frustration vs. days” regression plots for a country (depicted in Fig. 3, Fig. 12 and Fig. 13 for scenarios **I**, **II** and **III**), and let x_{date} and y_{date} be, respectively, the party-wise frustration of the network $\mathcal{G}_{\text{country}}$ and the duration of coalition negotiations after the election specified by “date”. We say that a point $(x_{\text{date}}, y_{\text{date}})$ is an *outlier* for that country if the value of y_{date} does not follow the linear regression line, meaning that this value is unusual given x_{date} . Instead, it is a *high leverage* point if it has “extreme” x_{date} value, i.e., this value is unusual given y_{date} . Finally, we say that the point is *influential* if it influences the slope of the regression line. See [64, 65] for a detailed explanation of outliers, high leverage and influential points. Tests such as residuals, leverage and delete-1 statistics are used

to identify the possible outliers, high leverage and influential points. In practice, given the linear regression model, outliers are observations whose studentized (also denoted externally studentized or studentized deleted) residuals are greater than 3 in absolute value and whose standardized (or internally studentized) residuals are greater than 2 in absolute value. Observations whose leverage statistics have values greater than $\frac{2p}{N}$, where $p = 2$ is the number of regression coefficients and N is the total number of observations (in this case the number of elections considered for each country), are identified as high leverage. Finally, observations whose Cook's distance is greater than three times the average Cook's distance, and whose Dffits (Difference in fits) values are greater than $2\sqrt{\frac{p}{N}}$, are identified as influential points.

We are particularly interested in observations marked as outliers and as both high leverage and influential, because they often represent unexpected events or difficult situations, in terms of either frustration or government negotiation duration, or both. In the vast majority of cases indeed they correspond to both high frustration and long negotiation times.

B.7 A brief discussion on national rules and traditions influencing the duration of the government negotiation talks

It is beyond the scope of this paper to make a thorough analysis of the additional factors influencing the duration of coalition negotiations, and we refer the reader to [9, 50, 66, 67]. However, to shed some light on the systematic differences in the government negotiation times across the various countries (see Fig. 10a), it is worth mentioning that the countries characterized by “negative” parliamentarism tend to have short cabinet formation processes. We say that a certain country has “negative” parliamentarism if the government (in order to rule) does not need to win a vote of confidence in the Parliament (as in the UK), or if the majority of the parties does not vote against it in the Parliament (as in Sweden), while it has “positive” parliamentarism otherwise [67]. The presence of minority governments and, most importantly, average short government formation processes is related to countries having “negative” parliamentarism: for instance Denmark, Finland, Iceland, Norway, Sweden and United Kingdom, see [5, 66, 67] and Fig. 14b.

In other countries the duration of the negotiation period is defined (and limited) by the constitution. It is worth observing that Estonia and Greece, two of the three countries showing a negative correlation between frustration and negotiation days in our scenario **I** and **II**, are among them. In Estonia, according to Article 89 of the Constitution, a limited time is given to the candidate Prime Minister to form a new Government: the President has 14 days to appoint a candidate Prime Minister, who in turn has to win a confidence vote in not more than 24 days. In case of failure the President can nominate (within 7 days) another candidate and the procedure repeats. Again, in case of failure, the Parliament nominates a candidate who has 14 days to win a vote of confidence. For Greece, as Article 37 of the Constitution states, the leader of the party with relative majority receiving from the President of the Republic the task to form a coalition (that has to enjoy

the confidence of the Parliament) has 3 days to succeed. In case of failure, the task is given to the leader of the second, and then third, largest party in the Parliament. If all the rounds of government negotiations fail new elections are called.

B.8 A brief discussion on elections resulting in a hung parliament

It is interesting to describe in some detail some cases of hung parliament mentioned in Section 3, whose corresponding data points in Fig. 3 (for scenario **I**, Fig. 12 and 13 for scenarios **II** and **III**) are both influential and high leverage points.

The 2006 Czech election saw a perfect partition of the parliament into parties of the left, Czech Social Democratic Party (ČSSD) and Communist Party of Bohemia and Moravia (KSČM) (that together won 100 of the 200 seats), and centre-right parties, Civic Democratic Party (ODS), Green Party (SZ) and Christian and Democratic Union - Czechoslovak People's Party (KSU-ČSL) (that collectively won the other 100 seats) [68]. In our framework, a parliament split into two identical halves represents a degenerate situation which escapes classification: the frustration can be very low but no majority exists. Indeed in our method the partition mentioned above corresponds to the ground state, i.e., our S_{best} consists of exactly 50% of + and 50% of -. The duration of the negotiation process cannot be predicted in degenerate cases like this because the two “real-life” equal size factions are also matching ideological polarization.

The 2016 and 2020 elections in Ireland, unlike the previous elections, resulted in no clear majority (or possible government coalition) and produced instead the most fragmented Dáils in history (which explains the higher levels of frustration): the number of parties in the parliament increased to 12 (from an average of 8 in the previous elections) and, while in the previous elections the percent distance between the first two parties (i.e., difference between their shares of seats in the parliament) had been at least of 19%, in the following elections it decreased, with the two biggest parties differing by as little as 3.2% in 2016 and 2.3% in 2020. However, differently from the 2016 election, in the 2020 election three parties (Fianna Fáil, Sinn Féin, Fine Gael) won roughly the same number of seats, which is reflected in a fractionalization index (or effective number of parties, see (3)) for the 2020 election higher than the one for the 2016 election: this meant that, for the first time, at least three parties were needed to form a majority coalition [69]. After the 2016 election, most of the parties did not want to participate in the government coalition and after 70 days of negotiations a “confidence and supply” deal was signed between Fine Gael and Fianna Fáil, with Fianna Fáil abstaining from voting against Fine Gael [70], while after the 2020 election, both Fine Gael and Fianna Fáil did not want to collaborate with Sinn Féin and decided instead to form a grand coalition with the Green Party.

In Spain, after the 2015 general election no agreement was reached between the parties before the deadline imposed by the Article 99 of the Spanish constitution (corresponding to two months from the first vote for investiture). A snap election was then called in June 2016 and after 131 days the Rajoy II government, composed by members of the People's Party (PP) and independents, was sworn in. Single-

party (minority) governments have always been common in Spain, however these elections saw a change from a nearly two-party system, i.e., majority of votes won by two parties, PP and Spanish Socialist Workers' Party (PSOE), to an effectively multi-party system [71] which increased the complexity of the cabinet formation process. Similarly, the government negotiations failed after the April 2019 legislative election, and the King had to call new elections in November 2019, which gives us a second “failure point” for Spain.

In 2017 in Germany 171 days passed from the election date before the Merkel IV cabinet was sworn in. Neither of the possible coalitions, Christian Democratic Union/Christian Social Union in Bavaria (CDU/CSU) and Free Democratic Party (FDP) or Social Democratic Party (SPD) and The Greens (GRÜNE), obtained a majority of seats in the parliament, while a new party, the far-right Alternative for Germany (AfD), managed to enter the parliament for the first time [72]. Since none of the parties wanted to collaborate with AfD, a compromise between parties with different ideological views had to be found: after a failed attempt to form a “Jamaica Coalition” (CDU/CSU, GRÜNE and FDP) the new government comprised a grand coalition between CDU/CSU and SPD.

B.9 A brief discussion on elections where the government negotiations failed

In our analysis of government formation, most of the legislative elections (255 out of 260 elections) are followed by successful cabinet negotiation talks, after which a cabinet is approved by the parliament and sworn in. In three countries, Czech Republic, Greece and Spain, we found however instances of failure of the government negotiation talks, in the form of a failed vote of investiture or expiration of the deadline set by the Constitution.

In the Czech Republic the candidate government needs to pass an investiture vote (of simple majority) within 30 days after its appointment (Article 68 of the Czech Constitution). In 2006, Topolánek (leader of ODS) decided to form a minority government even without managing to secure support from the other political parties [68]. Topolánek's first cabinet was sworn in on September 4 but lost a confidence vote on October 3, and cabinet negotiations had to restart. The deadlock was broken only after 220 days, thanks to the abstention of some MPs from ČSSD, and the Topolánek II cabinet (centre-right coalition of ODS, SZ and KSU-ČSL) was sworn in [68]. A similar situation (failure and then success at the same election) occurred again in the 2017 Czech election. In 2017, none of the political parties wanted to cooperate with Babiš due to the criminal fraud charges he was facing [73] and the Babiš first cabinet (ANO 2011 only), formed on December 13, failed to pass an investiture vote on January 16. In the cabinet negotiations that followed, ANO 2011 managed to reach a coalition agreement with ČSSD and to obtain external support from KSCM, and after 249 days the Babiš second cabinet was sworn in.

In Greece after the May 2012 legislative election none of the three leaders tasked to form a government managed to negotiate a cabinet coalition and, following the Article 37 of the Greek Constitution (see also Section B.7), 10 days after the

election day a caretaker cabinet was appointed and new elections were announced for June.

In Spain, see Article 99 of the Spanish Constitution, if the candidate government does not succeed in obtaining the confidence of the parliament (i.e., if it loses both a first vote of absolute majority and, 48 hours after, a second vote of simple majority) then a new candidate cabinet can try to obtain the confidence of the parliament. If within two months after the first vote for investiture no cabinet has managed to win a vote of confidence, the King can call new elections. After the 2008, 2016 and November 2019 elections the candidate cabinets managed to win the confidence of the parliament (with a vote of simple majority) while, after the 2015 and April 2019 elections, the government negotiations failed and the King had to call new elections.

For each of these three countries, Fig. 14a reports the election dates, the government sworn in dates and the government negotiations failure dates, determined for the Czech Republic as the date the cabinet (which later failed to pass the investiture vote) was formed, for Greece as the date the caretaker cabinet was sworn in, and for Spain as the date corresponding to two months after the first vote for investiture.

Figure 3 plots the duration of the government negotiation talks (calculated as number of days between the election date and the sworn in date and denoted *days*, see Fig. 14a) vs the frustration of the parliamentary networks, for scenario **I** (scenario **II** and **III** are reported respectively in Fig. 12 and Fig. 13). For Czech Republic, Greece and Spain, the blue data points (and corresponding regression line and value of correlation) correspond to elections followed by successful votes of investiture (the May 2012 Greek election and the 2015 and April 2019 Spanish elections are hence excluded), while the yellow data points (and corresponding regression line and value of correlation between parentheses) include all the elections: in this case the variable *days* indicates the number of days between the election date and the government negotiations failure date (see Fig. 14a).

B.10 Analysis of the type of governments formed after the elections

We distinguish between four types of governments (two being minority governments, two majority) that may be formed after a legislative election: first, we consider minority governments where the cabinet coalition has a strict minority in the parliament; second, minority governments whose coalition of parties holds exactly half of the seats in the parliament; third, majority governments which are minimal winning coalitions; and last, majority governments which are surplus coalitions. A coalition of parties is minimal winning if removing a party implies loss of majority in the parliament [2], and it is denoted surplus otherwise.

Fig 14b shows the type of government the countries we have considered in our analysis have formed: of 257 analyzed governments, 61 (23.74%) are minority, 5 (1.95%) are cabinet coalitions holding exactly 50% of the seats in the parliament, 131 (50.97%) are minimal winning, and 60 (23.35%) are surplus coalitions.

In Scenario **I**, computing the frustration is equivalent to find the minimum

winning coalition (represented by the group of parties achieving majority in S_{best}), that is, a coalition of parties that is minimal winning and also minimizing the functional $e(S)$, meaning that between all possible minimal winning coalitions it is the one with the lowest amount of total MPs. The expectation that governments should be minimal winning is standard in the literature [2, 5], even if it has been recently observed that, in reality, it is not uncommon for minority or surplus governments to form [74, 75].

B.11 Analysis of the Italian bicameral parliamentary system

Under the bicameral system of Italy the candidate cabinet needs to win the confidence of both the Chamber of Deputies and the Senate of the Republic [76]. In what follows we extend the results obtained for the Chamber of Deputies by considering also the network described by the parties winning seats in the Senate of the Republic.

For each election let \mathcal{G}_C and \mathcal{G}_S indicate the signed networks of the lower and upper chamber, respectively, and let the party set be defined as $\mathcal{P} = \mathcal{P}_C \cup \mathcal{P}_S$, where $\mathcal{P}_C = \{p_i : p_i \in \text{Chamber of Deputies}\}$ and $\mathcal{P}_S = \{p_i : p_i \in \text{Senate of the Republic}\}$. For each subset $\mathcal{P}(S)$ of \mathcal{P} described by the party configuration $S = \text{diag}\{s_1, \dots, s_{\text{card}(\mathcal{P})}\}$, where $s_i = 1$ if the party p_i belongs to the subset $\mathcal{P}(S)$ or $s_i = -1$ otherwise, we define the energy of the configuration S as the couple $[e_C(S_C), e_S(S_S)]$, where S_C (resp., S_S) and $e_C(S_C)$ (resp., $e_S(S_S)$) indicate the corresponding party configuration and energy functional (13) of the network \mathcal{G}_C (resp., \mathcal{G}_S), that is

$$S_X = \text{diag}\{s_1 I_{c_1}, \dots, s_{\text{card}(\mathcal{P}_X)} I_{c_{\text{card}(\mathcal{P}_X)}}\}$$

$$s_i = \begin{cases} +1, & p_i \in \mathcal{P}_X \cap \mathcal{P}(S) \\ -1, & p_i \in \mathcal{P}_X \setminus (\mathcal{P}_X \cap \mathcal{P}(S)) \end{cases}, \quad X \in \{C, S\},$$

where c_i is the number of seats won by party p_i (correspondingly in the Chamber or the Senate), and

$$e_X(S_X) = \frac{1}{2} \sum_{i,j \neq i} [|L(\mathcal{G}_X)| + S_X L(\mathcal{G}_X) S_X]_{ij}, \quad X \in \{C, S\},$$

where $L(\mathcal{G}_X)$ indicates the normalized signed Laplacian of the network \mathcal{G}_X . We say that a party configuration S is a majority configuration if its parties hold the majority of seats in both chambers of the parliament, that is, if $\sum_{i: p_i \in \mathcal{P}_C \cap \mathcal{P}(S)} c_i > \frac{630}{2}$ and $\sum_{i: p_i \in \mathcal{P}_S \cap \mathcal{P}(S)} c_i > \frac{315}{2}$, where 630 and 315 are the total number of seats in the Chamber and Senate, respectively.

For each election, the energy of all configurations S is depicted in Fig. 6, which shows the values $e_S(S_S)$ vs $e_C(S_C)$.

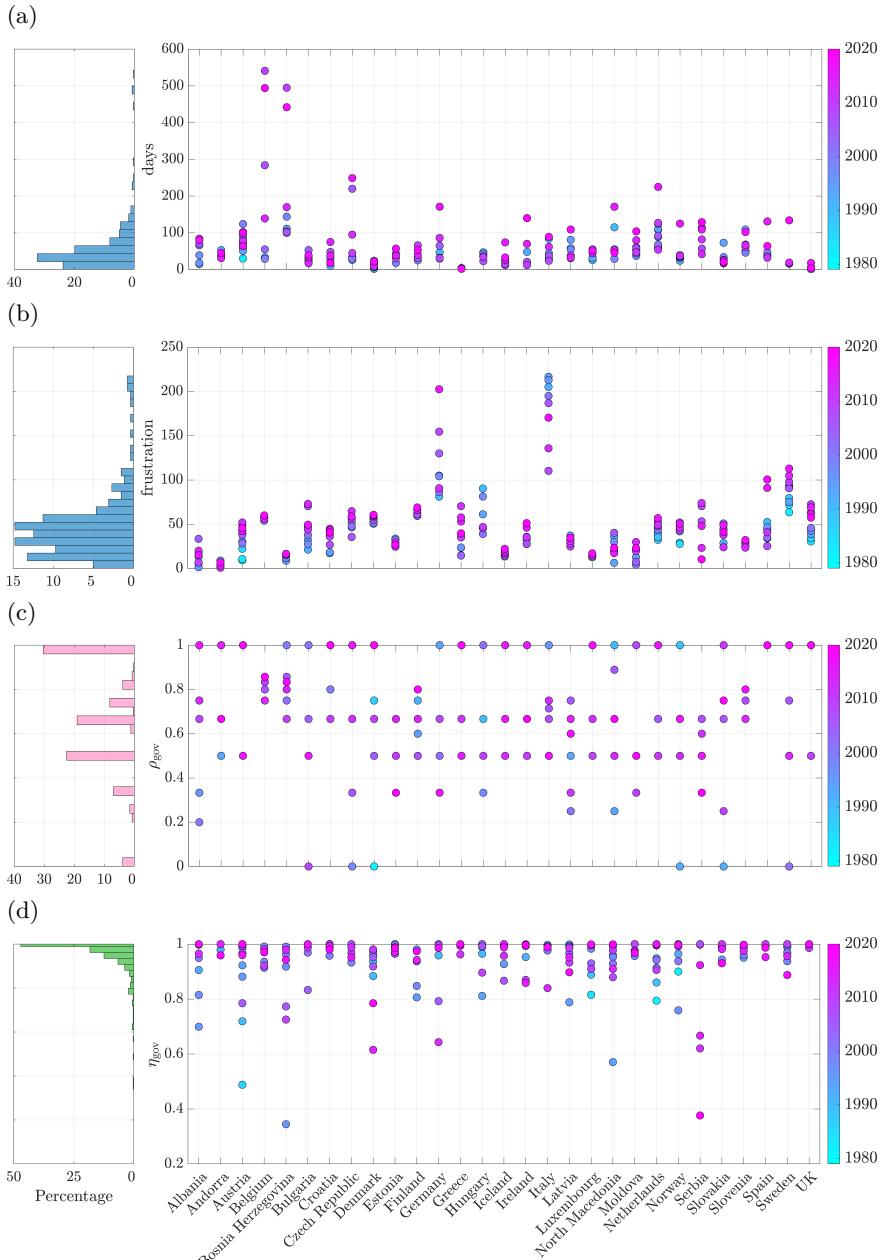


Figure 10: (a): Number of days of government coalition negotiations for each country and election. (b): Party-wise frustration for each country and election in scenario I. (c): ρ_{gov} , summary for each country and election. (d): η_{gov} , summary for each country and election. Left panels in (c), and (d): corresponding histograms. A colormap is used to differentiate data from different years.

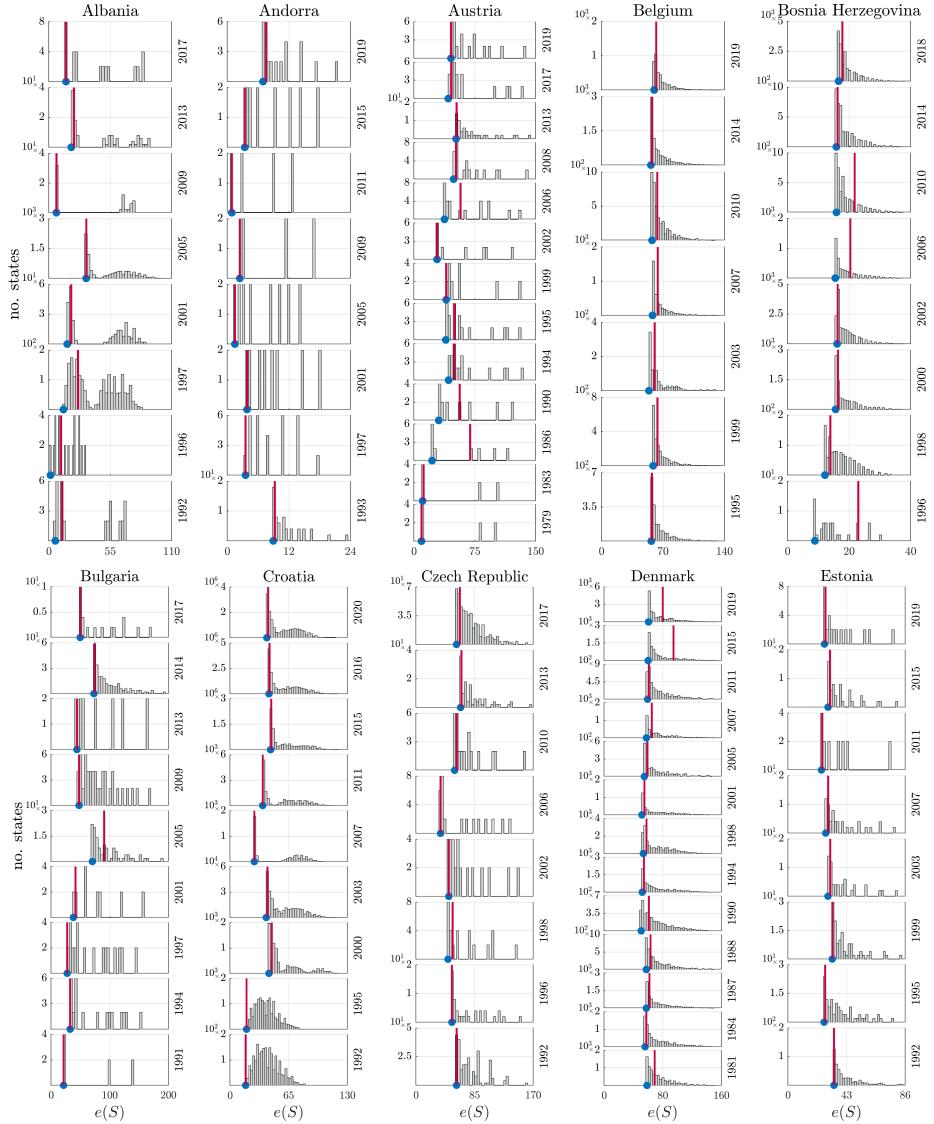
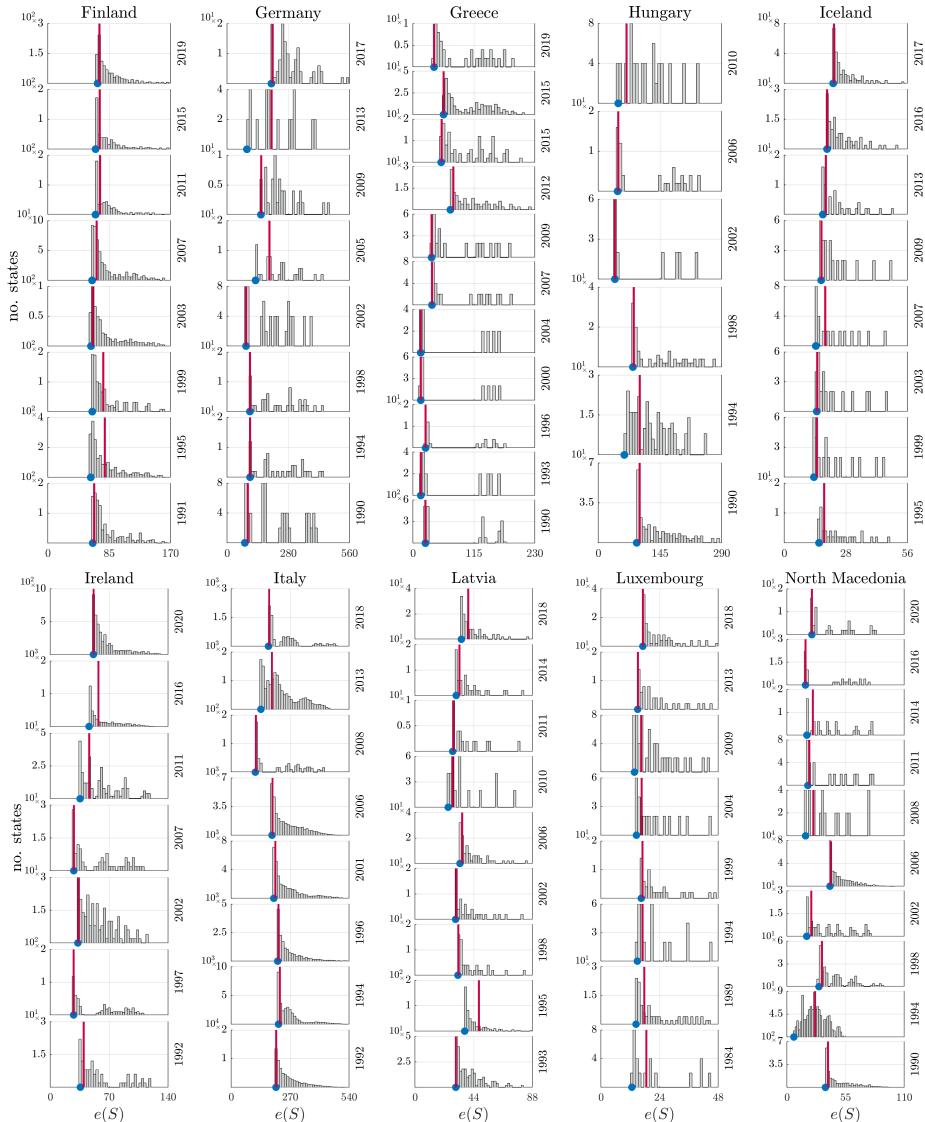


Figure 11: Continued on the next page.

**Figure 11:** Continued on the next page.

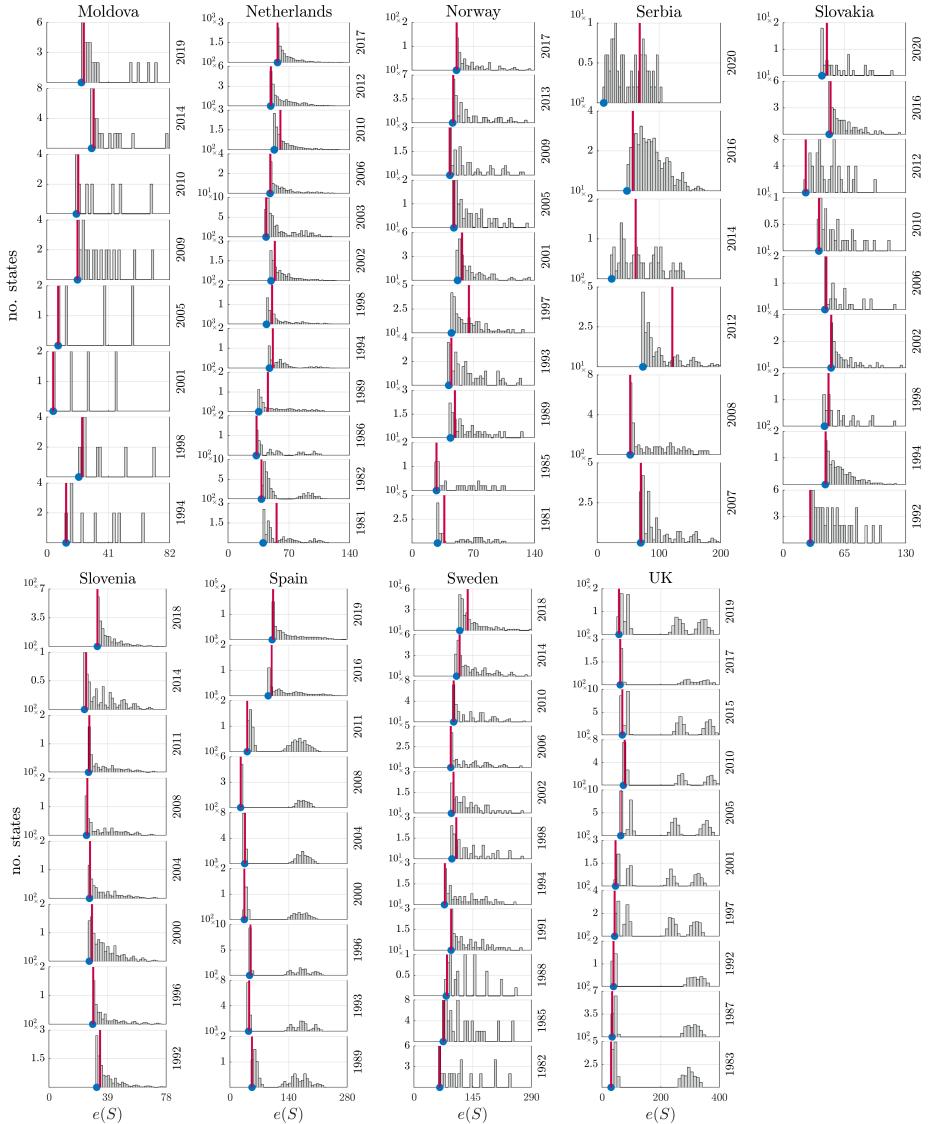


Figure 11: Energy landscape for scenario I and for each country, i.e., values of the energy $e(S)$ (see (13)) as $S = \text{diag}\{s_1 I_{c_1}, \dots, s_{n_p} I_{c_{n_p}}\}$, $s_i = \pm 1$, is varied: the value of the energy functional corresponding to the government, $e(S_{gov})$, is highlighted with the red line, while the minimum of the energy functionals, $e(S_{best}) = \zeta$, is indicated with the blue dot.

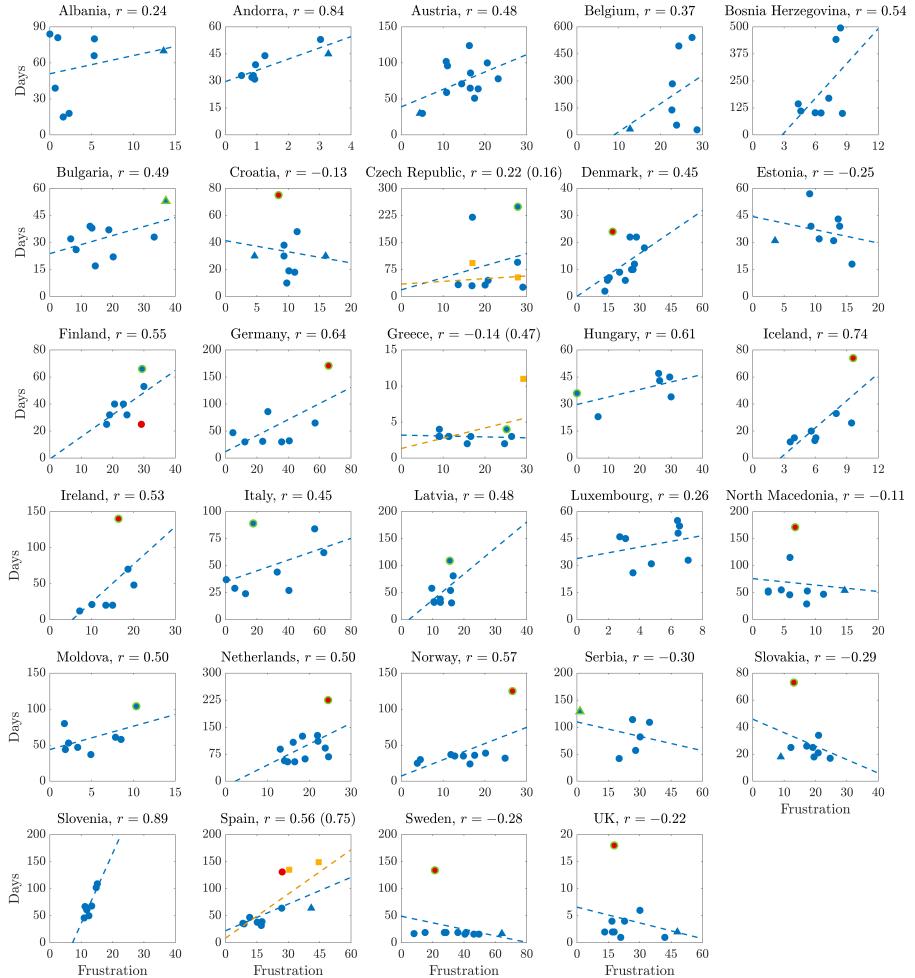


Figure 12: Scenario II. Frustration of the parliamentary networks v.s. duration of the government negotiation talks (days) and corresponding linear regression line, for all countries of Table 2. The value of correlation (r) for each country is reported in the plot heading. Legend: blue circles represent points that are neither outliers, nor high leverage nor influential. A red symbol indicates an outlier, a triangle a high leverage point and a symbol with green outline an influential point. Residual analysis, leverage statistic and delete-1 statistics are used to identify outliers, high leverage and influential points, respectively. Yellow square data points indicate elections corresponding to failure of government negotiations resulting in votes of no-confidence (Czech Republic in 2006 and 2017) and new elections (Spain in December 2015 and April 2019, Greece in May 2012), see Fig. 14a. Blue regression lines consider only the successful government formations. Including also the failure points we obtain the yellow regression lines.

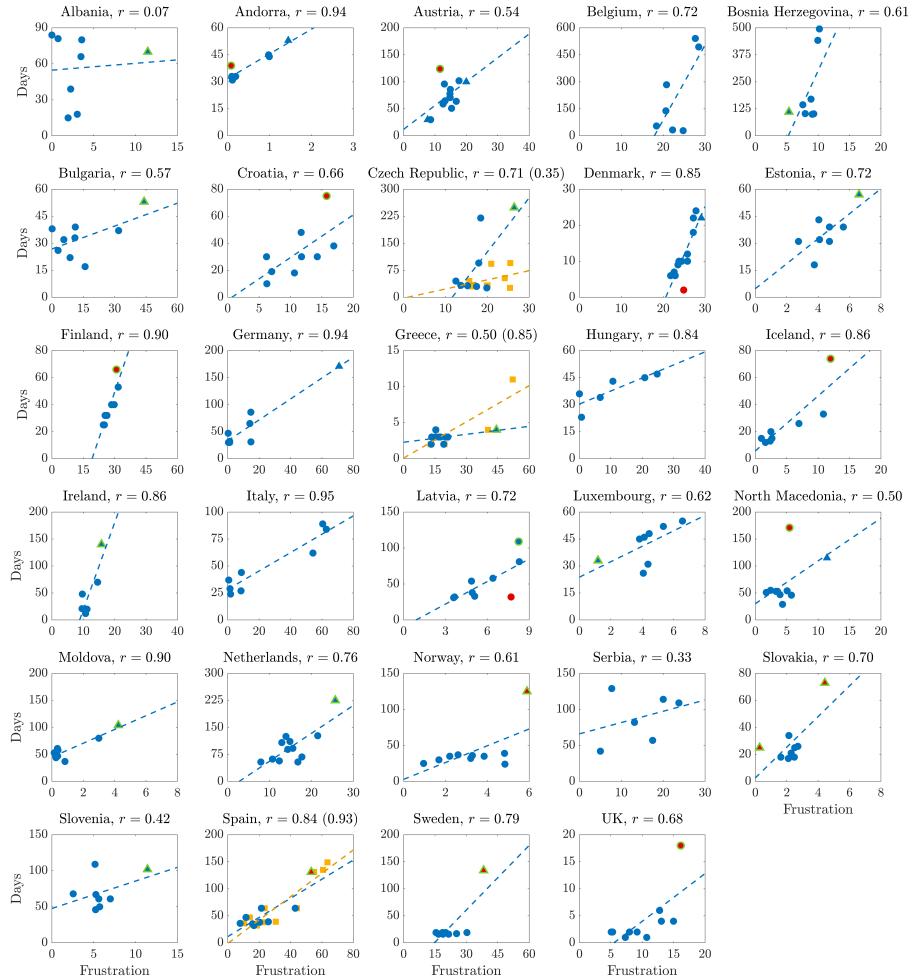
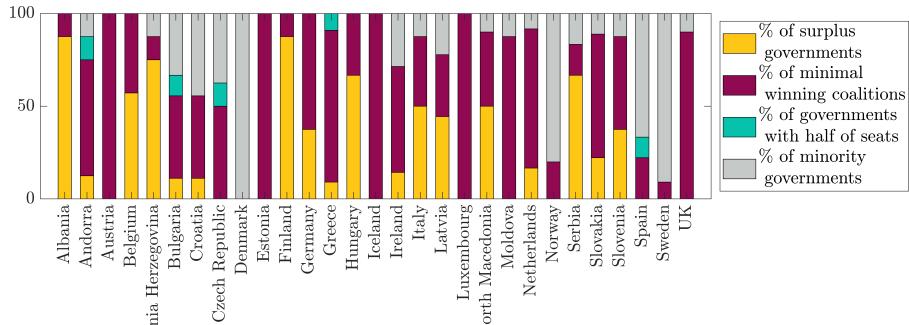


Figure 13: Scenario III. Frustration of the parliamentary networks v.s. duration of the government negotiation talks (days) and corresponding linear regression line, for all countries of Table 2. The value of correlation (r) for each country is reported in the plot heading. Legend: blue circles represent points that are neither outliers, nor high leverage nor influential. A red symbol indicates an outlier, a triangle a high leverage point and a symbol with green outline an influential point. Residual analysis, leverage statistic and delete-1 statistics are used to identify outliers, high leverage and influential points, respectively. Yellow square data points include also elections corresponding to failure of government negotiations resulting in votes of no-confidence (Czech Republic in 2006 and 2017) and new elections (Spain in December 2015 and April 2019, Greece in May 2012), see Fig. 14a. In these 3 countries all edge weights have been recomputed (hence frustration values are different with respect to the blue data points). Blue regression lines consider only the successful government formations. Including also the failure points we obtain the yellow regression lines.

Country	election	Date failure	sworn-in	N. of Days
Czech Republic	2006-06-03	2006-09-04	2007-01-09	93 (failure), 220 (success)
	2017-10-21	2017-12-13	2018-06-27	53 (failure), 249 (success)
Greece	2012-05-06	2012-05-17	–	11 (failure)
	2012-06-17	–	2012-06-21	4 (success)
Spain	2015-12-20	2016-05-03	–	135 (failure)
	2016-06-26	–	2016-11-04	131 (success)
	2019-04-28	2019-09-24	–	149 (failure)
	2019-11-10	–	2020-01-13	64 (success)

(a)



(b)

Figure 14: (a): Legislative elections related to failure of government negotiation talks in Czech Republic, Greece and Spain. (b): Type of government coalition formed after the election (minority, minority but yielding exactly half of the seats in the parliament, minimal winning, surplus), for each country of Table 2. Observe that to be classified as “majority” a cabinet coalition needs to hold (strictly) more than half of the seats in the parliament.

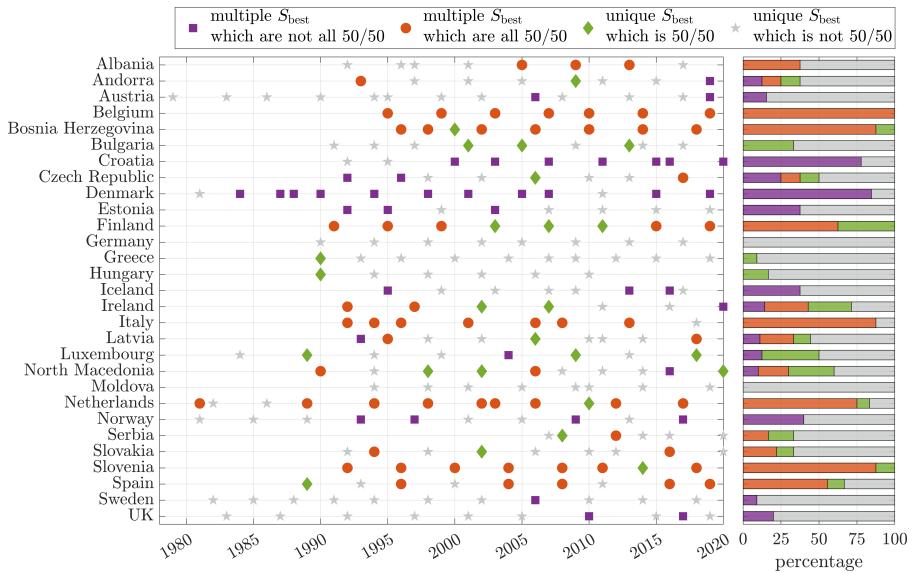


Figure 15: Scenario I. Elections for which the ground state (S_{best}) is “degenerate” (purple, orange, green) and not “degenerate” (light gray). Left panel: cases in which the ground state (S_{best}) is “degenerate” for each country and election; Right panel: percentage of elections for which the ground state is “degenerate”, for each country. Legend: gray stars are used to represent cases where the ground state is not “degenerate”. We consider three different “degenerate” cases: (1) multiple signature matrices S_{best} give the same value of frustration but there exists a S_{best} whose corresponding group of parties holds a majority of seats in the parliament (purple squares); (2) multiple signature matrices S_{best} give the same value of frustration and all corresponding party groups hold exactly half of the seats (50/50) in the parliament (orange circles); (3) unique signature matrix S_{best} whose corresponding party group holds exactly half of the seats (50/50) in the parliament (green diamond).

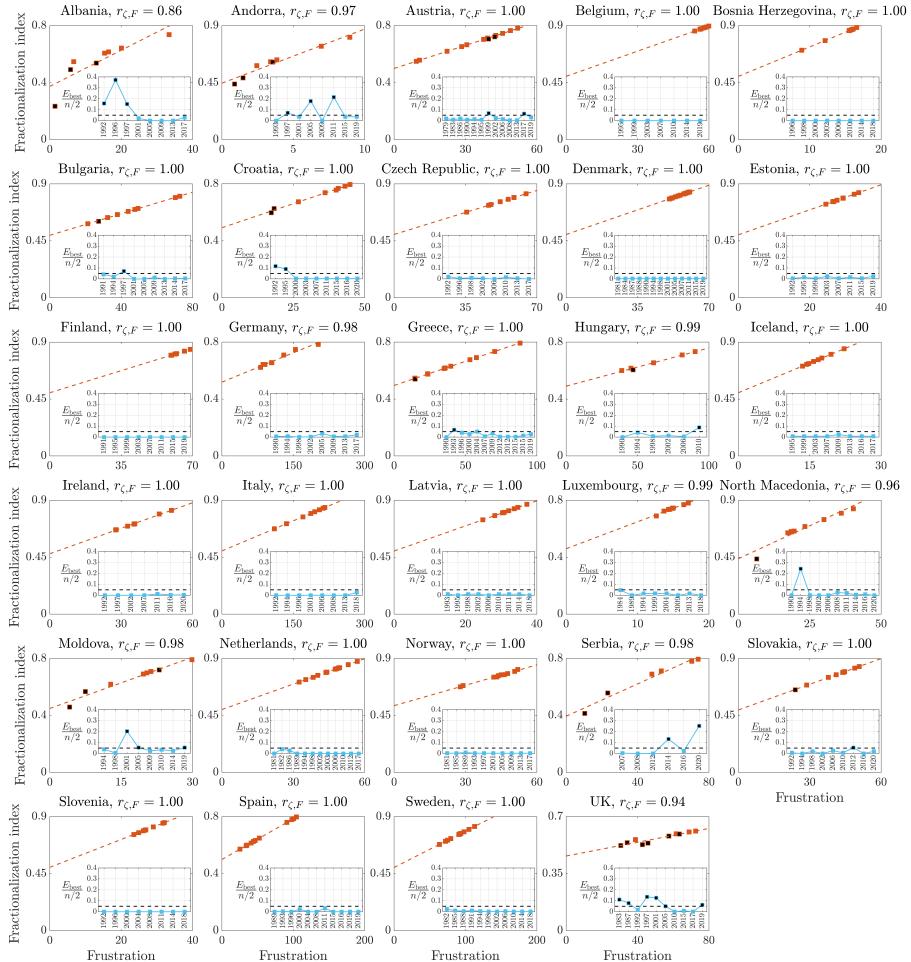


Figure 16: Scenario I. Fractionalization index vs frustration of the parliamentary networks and corresponding linear regression line, for all countries of Table 2. The value of correlation ($r_{\zeta,F}$) for each country is reported in the plot heading. Inset: distance of the minimum winning coalition S_{best} to 50% (of the total number of seats), calculated as $\frac{E_{best}}{n/2}$ where E_{best} is the number of seats in excess of S_{best} (see Section B.4). In all panels black squares indicate minimum winning coalitions S_{best} whose distance from 50% is greater than 5%.

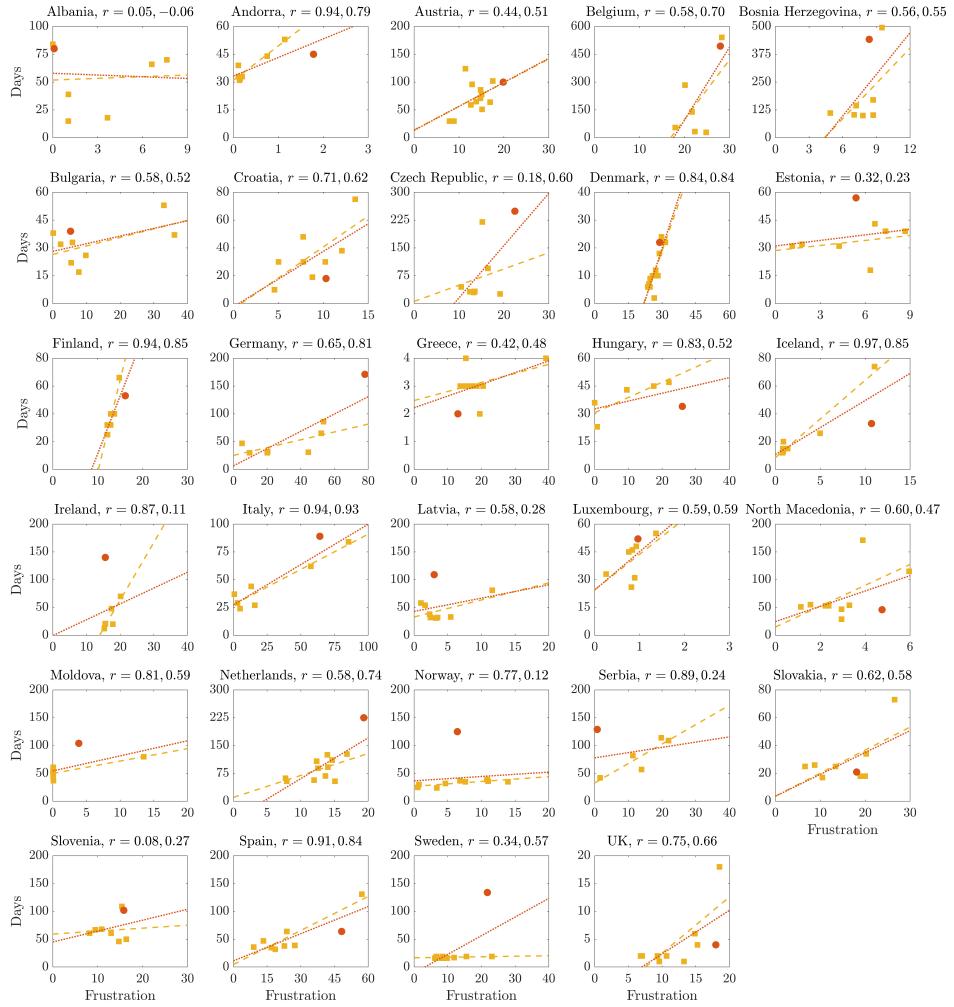


Figure 17: Leave-the-last-one-out analysis and linear regression plots between the duration of the government negotiation talks (days) and the frustration of the parliamentary networks in scenario III, for all countries. The last election (red circle) is used for the validation set while the remaining elections (yellow squares) are used to calculate the optimal choice of political positions in the left-right grid (see Fig. 4). A yellow dashed line represents the regression calculated on the first $N-1$ elections, while a red dotted line the regression calculated on all elections. The corresponding values of correlation r are reported in the plot heading.

Bibliography

- [1] A. Fontan and C. Altafini, "A signed network perspective on the government formation process in parliamentary democracies," *Scientific Reports*, vol. 11, no. 5134, dec 2021.
- [2] W. H. Riker, *The Theory of Political Coalitions*. New Haven: Yale University Press, 1962.
- [3] M. Laver, "MODELS OF GOVERNMENT FORMATION," *Annual Review of Political Science*, vol. 1, no. 1, pp. 1–25, jun 1998.
- [4] M. Debus, "Pre-electoral commitments and government formation," *Public Choice*, vol. 138, no. 1-2, pp. 45–64, jan 2009.
- [5] S. N. Golder, "Bargaining Delays in the Government Formation Process," *Comparative Political Studies*, vol. 43, no. 1, pp. 3–32, jan 2010.
- [6] A. Ecker and T. M. Meyer, "The duration of government formation processes in Europe," *Research and Politics*, vol. 2, no. 4, pp. 1–9, 2015.
- [7] D. Diermeier and P. van Roozendaal, "The Duration of Cabinet Formation Processes in Western Multi-Party Democracies," *British Journal of Political Science*, vol. 28, no. 4, pp. 609–626, 1998.
- [8] D. Diermeier, H. Eraslan, and A. Merlo, "A Structural Model of Government Formation," *Econometrica*, vol. 71, no. 1, pp. 27–70, jan 2003.
- [9] L. W. Martin and G. Vanberg, "Wasting Time? The Impact of Ideology and Size on Delay in Coalition Formation," *British Journal of Political Science*, vol. 33, no. 2, pp. 323–332, apr 2003.
- [10] A. Rusinowska and H. De Swart, "Negotiating a stable government: An application of bargaining theory to a coalition formation model," *Group Decision and Negotiation*, vol. 17, no. 5, pp. 445–464, 2008.
- [11] M. Laakso and R. Taagepera, "The ‘Effective’ number of parties: a measure with application to West Europe," *Comparative Political Studies*, vol. 12, no. 1, pp. 3–27, 1979.
- [12] D. W. Rae and M. Taylor, *The Analysis of Political Cleavages*. Yale University Press, 1970.
- [13] D. Austen-Smith and J. Banks, "Elections, Coalitions, and Legislative Outcomes," *American Political Science Review*, vol. 82, no. 2, pp. 405–422, jun 1988.
- [14] D. P. Baron and J. A. Ferejohn, "Bargaining in Legislatures," *American Political Science Review*, vol. 83, no. 4, pp. 1181–1206, dec 1989.
- [15] A. Merlo, "Bargaining over Governments in a Stochastic Environment," *Journal of Political Economy*, vol. 105, no. 1, pp. 101–131, feb 1997.

- [16] D. Diermeier and A. Merlo, “Government turnover in parliamentary democracies,” *Journal of Economic Theory*, vol. 94, no. 1, pp. 46–79, 2000.
- [17] A. Rusinowska, H. D. Swart, and J.-W. van der Rijt, “A new model of coalition formation,” *Social Choice and Welfare*, vol. 24, no. 1, pp. 129–154, feb 2005.
- [18] D. Ray, *A game-theoretic perspective on coalition formation*. Oxford University Press, 2007.
- [19] D. Diermeier and A. Merlo, “An empirical investigation of coalitional bargaining procedures,” *Journal of Public Economics*, vol. 88, no. 3-4, pp. 783–797, 2004.
- [20] S. Wasserman and K. Faust, *Social Network Analysis: Methods and applications*, ser. Structural Analysis in the Social Sciences. Cambridge University Press, 1994.
- [21] D. Easley and J. Kleinberg, *Networks, Crowds, and Markets: Reasoning about a Highly Connected World*. Cambridge: Cambridge University Press, 2010.
- [22] D. Cartwright and F. Harary, “Structural balance: a generalization of Heider’s theory.” *Psychological Review*, vol. 63, no. 5, pp. 277–293, 1956.
- [23] T. Zaslavsky, “Signed graphs,” *Discrete Applied Mathematics*, vol. 4, no. 1, pp. 47–74, 1982.
- [24] S. Aref and M. C. Wilson, “Balance and frustration in signed networks,” *Journal of Complex Networks*, vol. 7, no. 2, pp. 163–189, apr 2019.
- [25] J. Kunegis, “Applications of Structural Balance in Signed Social Networks,” *arXiv:1402.6865v1*, pp. 1–37, 2014.
- [26] G. Facchetti, G. Iacono, and C. Altafini, “Computing global structural balance in large-scale signed social networks,” *Proceedings of the National Academy of Sciences*, vol. 108, no. 52, pp. 20953–20958, 2011.
- [27] M. Mezard, G. Parisi, and M. Virasoro, *Spin Glass Theory and Beyond. An Introduction to the Replica Method and Its Applications*, ser. World Scientific Lecture Notes in Physics: Volume 9. World Scientific, nov 1986.
- [28] K. Binder and A. P. Young, “Spin glasses: Experimental facts, theoretical concepts, and open questions,” *Reviews of Modern Physics*, vol. 58, no. 4, pp. 801–976, oct 1986.
- [29] N. E. Leonard, “Multi-agent system dynamics: Bifurcation and behavior of animal groups,” *Annual Reviews in Control*, vol. 38, no. 2, pp. 171–183, 2014.
- [30] R. Gray, A. Franci, V. Srivastava, and N. E. Leonard, “Multiagent Decision-Making Dynamics Inspired by Honeybees,” *IEEE Transactions on Control of Network Systems*, vol. 5, no. 2, pp. 793–806, jun 2018.

- [31] A. Fontan and C. Altafini, “Multiequilibria Analysis for a Class of Collective Decision-Making Networked Systems,” *IEEE Transactions on Control of Network Systems*, vol. 5, no. 4, pp. 1931–1940, dec 2018.
- [32] A. Fontan and C. Altafini, “Achieving a decision in antagonistic multi agent networks: frustration determines commitment strength,” in *57th IEEE Conference on Decision and Control*. Miami Beach, FL, USA: IEEE, dec 2018, pp. 109–114.
- [33] M. J. Laver and I. Budge, Eds., *Party Policy and Government Coalitions*. London: Palgrave Macmillan UK, 1992.
- [34] I. Budge, “The Standard Right-Left Scale,” 2013. https://manifesto-project.wzb.eu/down/papers/budge{__}right-left-scale.pdf
- [35] N. Merz, S. Regel, and J. Lewandowski, “The manifesto corpus: A new resource for research on political parties and quantitative text analysis,” *Research and Politics*, vol. 3, no. 2, pp. 1–8, 2016.
- [36] I. D. Couzin, J. Krause, N. R. Franks, and S. A. Levin, “Effective leadership and decision-making in animal groups on the move,” *Nature*, vol. 433, no. 7025, pp. 513–516, 2005.
- [37] S. Aref and M. C. Wilson, “Measuring partial balance in signed networks,” *Journal of Complex Networks*, vol. 6, no. 4, pp. 566–595, aug 2018.
- [38] P. A. Sabatier, Ed., *Theories of the Policy Process*, 1st ed. Boulder, CO: Routledge, jun 2007.
- [39] P. Leifeld, *Discourse Network Analysis: policy debates as dynamic networks*, J. N. Victor, A. H. Montgomery, and M. Lubell, Eds. Oxford University Press, aug 2017, vol. 1.
- [40] P. Leifeld and L. Brandenberger, “Endogenous Coalition Formation in Policy Debates,” *arXiv:1904.05327v1*, 2019.
- [41] Z. Maoz and Z. Somer-Topcu, “Political polarization and cabinet stability in multiparty systems: A social networks analysis of European parliaments, 1945–98,” *British Journal of Political Science*, vol. 40, no. 4, pp. 805–833, 2010.
- [42] S. Aref and Z. Neal, “Detecting coalitions by optimally partitioning signed networks of political collaboration,” *Scientific Reports*, vol. 10, no. 1, pp. 1–10, 2020.
- [43] S. L. Feld and B. Grofman, “The Laakso-Taagepera index in a mean and variance framework,” *Journal of Theoretical Politics*, vol. 19, no. 1, pp. 101–106, 2007.

- [44] A. Volkens, T. Burst, W. Krause, P. Lehmann, T. Matthieß, N. Merz, S. Regel, B. Weßels, and L. Zehnter, “The Manifesto Data Collection. Manifesto Project (MRG/CMP/MARPOR). Version 2020b,” Berlin, 2020. <https://doi.org/10.25522/manifesto.mpds.2020b>
- [45] D. Holger and P. Manow, “Parliaments and governments database (ParlGov): Information on parties, elections and cabinets in modern democracies. Development version.” 2018. <http://www.parlgov.org>
- [46] I. Inter-Parliamentary Union, “New Parline: the IPU’s Open Data Platform,” 2019. <https://data.ipu.org>
- [47] R. Bakker, C. de Vries, E. Edwards, L. Hooghe, S. Jolly, G. Marks, J. Polk, J. Rovny, M. Steenbergen, and M. A. Vachudova, “Measuring party positions in Europe,” *Party Politics*, vol. 21, no. 1, pp. 143–152, 2015.
- [48] J. Polk, J. Rovny, R. Bakker, E. Edwards, L. Hooghe, S. Jolly, J. Koedam, F. Kostelka, G. Marks, G. Schumacher, M. Steenbergen, M. Vachudova, and M. Zilovic, “Explaining the salience of anti-elitism and reducing political corruption for political parties in Europe with the 2014 Chapel Hill Expert Survey data,” *Research and Politics*, vol. 4, no. 1, pp. 1–9, 2017.
- [49] F. Casal Bértoa, “Database on WHO GOVERNS in Europe and beyond, PSGo.” 2019. www.whogoverns.eu
- [50] B. E. Rasch, S. Martin, and J. A. Cheibub, Eds., *Parliaments and Government Formation*. Oxford University Press, nov 2015.
- [51] T. Lansford, Ed., *Political Handbook of the World 2016-2017*. SAGE Publications, 2017.
- [52] “Political Data Yearbook interactive.” <http://www.politicaldatayearbook.com/>
- [53] A. Bågenholm and A. Clark, Eds., *Political Data Yearbook 2017*. Wiley Online Library, 2018, vol. 57, no. 1.
- [54] J. Nousiainen, “From semi-presidentialism to parliamentary government: Political and constitutional developments in Finland,” *Scandinavian Political Studies*, vol. 24, no. 2, pp. 95–109, 2001.
- [55] K. Strøm and J. Y. Leipart, “Policy, Institutions, and Coalition Avoidance: Norwegian Governments, 1945–1990.” *American Political Science Review*, vol. 87, no. 4, pp. 870–887, dec 1993.
- [56] E. H. Allern, “Survival of a Majority Coalition: The Norwegian Parliamentary Election of 14 September 2009,” *West European Politics*, vol. 33, no. 4, pp. 904–912, jul 2010.
- [57] E. H. Allern and R. Karlsen, “A Turn to the Right: The Norwegian Parliamentary Election of September 2013,” *West European Politics*, vol. 37, no. 3, pp. 653–663, 2014.

- [58] E. H. Allern and N. Aylott, “Overcoming the fear of commitment: Pre-electoral coalitions in Norway and Sweden,” *Acta Politica*, vol. 44, no. 3, pp. 259–285, 2009.
- [59] N. Aylott and N. Bolin, “Polarising Pluralism: The Swedish Parliamentary Election of September 2014,” *West European Politics*, vol. 38, no. 3, pp. 730–740, may 2015.
- [60] C. Altafini, “Consensus Problems on Networks With Antagonistic Interactions,” *IEEE Transactions on Automatic Control*, vol. 58, no. 4, pp. 935–946, apr 2013.
- [61] J. Kunegis, S. Schmidt, A. Lommatzsch, J. Lerner, E. W. De Luca, and S. Albayrak, “Spectral Analysis of Signed Graphs for Clustering, Prediction and Visualization,” in *2010 SIAM International Conference on Data Mining*. Philadelphia, PA: Society for Industrial and Applied Mathematics, apr 2010, pp. 559–570.
- [62] S. Aref, A. J. Mason, and M. C. Wilson, “An exact method for computing the frustration index in signed networks using binary programming,” *arXiv:1611.09030v2*, nov 2016.
- [63] P. U. Abara, F. Ticozzi, and C. Altafini, “Spectral Conditions for Stability and Stabilization of Positive Equilibria for a Class of Nonlinear Cooperative Systems,” *IEEE Transactions on Automatic Control*, vol. 63, no. 2, pp. 402–417, feb 2018.
- [64] G. James, D. Witten, T. Hastie, and R. Tibshirani, *An Introduction to Statistical Learning*, ser. Springer Texts in Statistics. Springer, 2013.
- [65] M. H. Kutner, C. J. Nachtsheim, J. Neter, and W. Li, *Applied Linear Statistical Models Fifth Edition*, 5th ed. McGraw-Hill Irwin, 2005.
- [66] A. Ecker and T. M. Meyer, “Coalition Bargaining Duration in Multiparty Democracies,” *British Journal of Political Science*, no. 1998, pp. 1–20, dec 2017.
- [67] T. Bergman, “Constitutional rules and party goals in coalition formation: an analysis of winning minority governments in Sweden,” PhD Dissertation, Umeå universitet, 1995.
- [68] L. Linek, “Czech Republic,” *European Journal of Political Research*, vol. 46, no. 7-8, pp. 929–937, 2007.
- [69] C. Little, “Change gradually, then all at once: the general election of February 2020 in the Republic of Ireland,” *West European Politics*, vol. 44, no. 3, pp. 714–723, 2021.
- [70] E. O’Malley, “70 Days: Government Formation in 2016,” in *How Ireland Voted 2016: The Election that Nobody Won*, M. Gallagher and M. Marsh, Eds. Cham: Springer International Publishing, 2016, pp. 255–276.

- [71] T. D. Lancaster, “The Spanish general elections of 2015 and 2016: a new stage in democratic politics?” *West European Politics*, vol. 40, no. 4, pp. 919–937, jul 2017.
- [72] T. Bräuninger, M. Debus, J. Müller, and C. Stecker, “Party Competition and Government Formation in Germany: Business as Usual or New Patterns?” *German Politics*, vol. 28, no. 1, pp. 80–100, 2019.
- [73] S. Hanley and M. A. Vachudova, “Understanding the illiberal turn: democratic backsliding in the Czech Republic,” *East European Politics*, vol. 34, no. 3, pp. 276–296, 2018.
- [74] A. Bassi, “Policy preferences in coalition formation and the stability of minority and surplus governments,” *Journal of Politics*, vol. 79, no. 1, pp. 250–268, 2017.
- [75] M. Golder, S. N. Golder, and D. A. Siegel, “Modeling the institutional foundation of parliamentary government formation,” *Journal of Politics*, vol. 74, no. 2, pp. 427–445, 2012.
- [76] F. Russo, “Government Formation in Italy,” in *Parliaments and Government Formation*. Oxford University Press, nov 2015, pp. 136–152.

Paper E

On the properties of Laplacian pseudoinverses

Authors: Angela Fontan and Claudio Altafini

© [2021] IEEE. *Edited version of the paper:*

A. Fontan and C. Altafini, “On the properties of Laplacian pseudoinverses,” in *60th IEEE Conference on Decision and Control*. Austin, TX, USA: IEEE, 2021.

On the properties of Laplacian pseudoinverses

Angela Fontan^{*} and Claudio Altafini^{*}

^{*}Dept. of Electrical Engineering
Linköping University
SE-581 83 Linköping, Sweden
angela.fontan@liu.se, claudio.altafini@liu.se

Abstract

The pseudoinverse of a graph Laplacian is used in many applications and fields, such as for instance in the computation of the effective resistance in electrical networks, in the calculation of the hitting/commuting times for a Markov chain and in continuous-time distributed averaging problems. In this paper we show that the Laplacian pseudoinverse is in general not a Laplacian matrix but rather a signed Laplacian with the property of being an eventually exponentially positive matrix, i.e., of obeying a strong Perron-Frobenius property. We show further that the set of signed Laplacians with this structure (i.e., eventual exponential positivity) is closed with respect to matrix pseudoinversion. This is true even for signed digraphs, and provided that we restrict to Laplacians that are weight balanced also stability is guaranteed.

1 Introduction

For a network or a networked system, the Laplacian matrix is a fundamental object that captures information about e.g., connectivity and spectrum [2, 3], as well as properties of the dynamics that live on the graph [4–7]. Associated to the Laplacian is also a Laplacian pseudoinverse, typically a Moore-Penrose pseudoinverse, which has also been used extensively to describe graph-related quantities. For instance it is used to build an effective resistance matrix for the graph, a distance measure that exploits the analogy between graphs and electrical networks [8–10], and to compute hitting/commuting times in Markov chains [11–14]. It is also used to estimate the \mathcal{H}_2 norm in networked dynamical systems [15–17].

If we consider a graph with nonnegative edge weights, it is well-known that the Laplacian L is an M-matrix (i.e., a matrix with nonpositive off-diagonal entries, such that $-L$ is marginally stable, see below for proper definitions). It is also easy to show that the Laplacian pseudoinverse does not belong to the same class, not even when the graph is undirected. Consider for instance the following Laplacian

matrix

$$L = \begin{bmatrix} 0.8 & -0.7 & -0.1 \\ -0.7 & 0.9 & -0.2 \\ -0.1 & -0.2 & 0.3 \end{bmatrix}.$$

Its pseudoinverse is

$$L^\dagger = \begin{bmatrix} 0.773 & 0.048 & -0.821 \\ 0.048 & 0.628 & -0.676 \\ -0.821 & -0.676 & 1.498 \end{bmatrix}$$

which has an anomalous sign in the (1,2) entry, even though it has the same stability properties of L .

The aim of this paper is to investigate the algebraic properties of Laplacian pseudoinverses. Even though L^\dagger is not an M-matrix, it has nevertheless most of the properties of M-matrices, most notably it obeys to a strong Perron-Frobenius property: the pair formed by the eigenvalue 0 and eigenvector $\mathbf{1} = [1 \dots 1]^T$ is the “dominant” eigenpair for $-L^\dagger$ (just like it is for $-L$) in spite of the presence of positive off-diagonal entries in L^\dagger . Such matrices are called *eventually exponentially positive* in the linear algebra literature [5, 18–20]. We show in the paper that this argument can be extended to signed Laplacians, i.e., Laplacians associated to signed digraphs¹: the pseudoinverse of an eventually exponentially positive signed Laplacian is an eventually exponentially positive signed Laplacian. In other words, the class of eventually exponentially positive signed Laplacians is closed with respect to pseudoinversion.

Under the assumption of edge weight balance, such class is also closed with respect to stability, i.e., $-L$ and $-L^\dagger$ eventually exponentially positive are also marginally stable (and of corank 1).

When we restrict further the class from weight balanced L to normal L , then it coincides also with the class of Laplacians and Laplacian pseudoinverses whose symmetric part is positive semidefinite of corank 1. Such restriction is particularly useful in the context of effective resistance, which, being a distance, has to be symmetric. For normal signed Laplacians we obtain a natural way to extend the notion of effective resistance to digraphs, alternative to the definitions already appeared in the literature, see e.g. [22].

2 Preliminaries

2.1 Linear algebraic preliminaries

Given a matrix $A = [a_{ij}] \in \mathbb{R}^{n \times n}$, $A \geq 0$ means element-wise nonnegative, i.e., $a_{ij} \geq 0$ for all $i, j = 1, \dots, n$, while $A > 0$ means element-wise positive, i.e., $a_{ij} > 0$ for all $i, j = 1, \dots, n$. The spectrum of A is denoted $\Lambda(A) = \{\lambda_1(A), \dots, \lambda_n(A)\}$, where $\lambda_i(A)$, $i = 1, \dots, n$, are the eigenvalues of A . In this paper we use the ordering $\operatorname{Re}[\lambda_1(A)] \leq \operatorname{Re}[\lambda_2(A)] \leq \dots \leq \operatorname{Re}[\lambda_n(A)]$, where $\operatorname{Re}[\lambda_i(A)]$ indicates the real part

¹As signed Laplacian here we use the so-called “repelling Laplacian” in the terminology of [21], see Section 2 for a precise definition.

of $\lambda_i(A)$. The spectral radius of A is the smallest real nonnegative number such that $\rho(A) \geq |\lambda_i(A)|$ for all $i = 1, \dots, n$ and $\lambda_i(A) \in \Lambda(A)$. A matrix A is called *Hurwitz stable* if $\text{Re}[\lambda_n(A)] < 0$, and *marginally stable* if $\text{Re}[\lambda_n(A)] = 0$ is a simple root of the minimal polynomial of A .

A matrix A is called positive semidefinite (psd) if $x^T Ax = \frac{x^T A + A^T}{2} x \geq 0 \forall x \in \mathbb{R}^n$ and it is called positive definite (pd) if $x^T Ax = \frac{x^T A + A^T}{2} x > 0 \forall x \in \mathbb{R}^n \setminus \{0\}$.

A matrix A is called irreducible if there does not exist a permutation matrix P s.t. $P^T AP$ is block triangular.

A matrix B is called a Z-matrix if it can be written as $B = sI - A$, where $A \geq 0$ and $s > 0$, and it is called a M-matrix if, in addition, $s \geq \rho(A)$, which implies that all the eigenvalues of B have nonnegative real part. If $s > \rho(A)$ then B is nonsingular and $-B$ is Hurwitz stable. If $s = \rho(A)$ then B is singular, and if A is irreducible then $-B$ is marginally stable.

If A is a singular matrix, the Moore-Penrose pseudoinverse of A , denoted A^\dagger , is the unique $n \times n$ matrix that satisfies $AA^\dagger A = A$, $A^\dagger AA^\dagger = A^\dagger$, $(A^\dagger A)^T = A^\dagger A$, and $(AA^\dagger)^T = AA^\dagger$. A singular matrix A is said to have index 1 if the range of A , $\text{range}(A)$, and the kernel of A , $\ker(A)$, are complementary subspaces, i.e., $\text{range}(A) \cap \ker(A) = \emptyset$. For index 1 singular matrices, the Drazin inverse and the group inverse coincide. A singular M-matrix has always index 1 [23]. A matrix is *normal* if it commutes with its transpose: $AA^T = A^T A$. A matrix A is said an *EP matrix* (Equal Projector, also called a *range symmetric matrix* [23]) if $\ker(A) = \ker(A^T)$ (and hence $\text{range}(A) = \text{range}(A^T)$). EP matrices generalize normal matrices, and like normal matrices have many equivalent characterizations, see [23]. For instance an EP matrix A is such that A commutes with its Moore-Penrose pseudoinverse A^\dagger . If A is an EP-matrix, then $\exists U$ orthogonal such that

$$A = U \begin{bmatrix} 0 & 0 \\ 0 & B \end{bmatrix} U^T$$

with B nonsingular of dimension $r = \text{rank}(A)$. Singular EP matrices have index 1, and for them the Moore-Penrose pseudoinverse, the Drazin inverse and the group inverse coincide.

2.2 Signed graphs

Let $\mathcal{G}(A) = (\mathcal{V}, \mathcal{E}, A)$ be the (weighted) digraph with vertex set \mathcal{V} ($\text{card}(\mathcal{V}) = n$), $\mathcal{E} = \mathcal{V} \times \mathcal{V}$ and adjacency matrix $A = [a_{ij}] \in \mathbb{R}^{n \times n}$: $a_{ij} \in \mathbb{R} \setminus \{0\}$ iff $(j, i) \in \mathcal{E}$. Since each edge of the digraph is labeled by a sign (i.e., $\text{sign}(a_{ij}) = \pm 1$), $\mathcal{G}(A)$ is called a signed digraph. In the particular case where $A \geq 0$, the digraph $\mathcal{G}(A)$ is called nonnegative. For digraphs $\mathcal{G}(A)$ which are strongly connected and without self-loops, the matrix A is irreducible with null diagonal.

The weighted in-degree and out-degree of node i are denoted $\delta_i^{\text{in}} = \sum_{j=1}^n a_{ij}$ and $\delta_i^{\text{out}} = \sum_{j=1}^n a_{ji}$, respectively. The (signed) *Laplacian* of a graph $\mathcal{G}(A)$ is the (in general non-symmetric) matrix $L = \Delta - A$ where $\Delta = \text{diag}\{\delta_1^{\text{in}}, \dots, \delta_n^{\text{in}}\}$. This definition of signed Laplacian corresponds to the so-called “repelling Laplacian” in the terminology of [21]. By construction, this Laplacian is a singular matrix

with $\ker(L) = \text{span}\{\mathbf{1}\}$ where $\mathbf{1} \in \mathbb{R}^n$ is the vector of 1s. However, $-L$ need not be marginally stable and its symmetric part $L_s = \frac{L+L^T}{2}$ need not be positive semidefinite, as we show in the examples in Section 3. Moreover, L irreducible (or, $\mathcal{G}(A)$ strongly connected) need not imply $\text{corank}(L) = 1$. For instance, consider a complete, undirected, signed graph $\mathcal{G}(A)$ whose Laplacian is

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

It is $\Lambda(L) = \{0, 0, 4, 4\}$ and $\mathbf{1}, [0, 1, -1, 0]^T \in \ker(L)$, i.e., L is marginally stable of corank 2. Similarly, $\text{corank}(L)\mathbf{1}$ need not imply L irreducible.

A digraph $\mathcal{G}(A)$ is *weight balanced* if in-degree and out-degree coincide for each node, i.e., $\delta_i^{\text{in}} = \sum_{j=1}^n a_{ij} = \sum_{j=1}^n a_{ji} = \delta_i^{\text{out}}$ for all $i = 1, \dots, n$. As we show in Lemma 2, $\text{corank}(L) = 1$ and L weight balanced imply L irreducible.

2.3 Eventual exponential positivity

Definition 1. A matrix $A \in \mathbb{R}^{n \times n}$ has the (strong) Perron-Frobenius property if $\rho(A)$ is a simple positive eigenvalue of A s.t. $\rho(A) > |\lambda(A)|$ for every $\lambda(A) \in \Lambda(A)$, $\lambda(A) \neq \rho(A)$, and χ , the right eigenvector relative to $\rho(A)$, is positive. _____

The set of matrices which possess the Perron-Frobenius property will be denoted \mathcal{PF} , and it is known (see e.g. [24, Thm 8.4.4]) that irreducible nonnegative matrices are part of this set. However, it has been shown (see [18]) that matrices having negative elements can also possess this property, provided that they are eventually positive.

Definition 2. A matrix $A \in \mathbb{R}^{n \times n}$ is called *eventually positive* (denoted $A \stackrel{\vee}{>} 0$) if $\exists k_0 \in \mathbb{N}$ s.t. $A^k > 0$ for all $k \geq k_0$. _____

Theorem 1 (2.2 in [18]). Let $A \in \mathbb{R}^{n \times n}$. Then the following statements are equivalent:

- (i) Both $A, A^T \in \mathcal{PF}$;
- (ii) $A \stackrel{\vee}{>} 0$;
- (iii) $A^T \stackrel{\vee}{>} 0$.

Definition 3. A matrix $A \in \mathbb{R}^{n \times n}$ is called *eventually exponentially positive* if $\exists t_0 \in \mathbb{N}$ s.t. $e^{At} > 0$ for all $t \geq t_0$. _____

Lemma 1 (Thm 3.3 in [19]). A matrix $A \in \mathbb{R}^{n \times n}$ is eventually exponentially positive if and only if $\exists d \geq 0$ s.t. $A + dI \stackrel{\vee}{>} 0$.

2.4 Kron reduction for undirected networks

Consider an undirected, connected, weighted graph $\mathcal{G}(A) = (\mathcal{V}, \mathcal{E}, A)$ with adjacency matrix $A = [a_{ij}] \in \mathbb{R}^{n \times n}$. Let $\alpha \subset \{1, \dots, n\}$ (with $\text{card}(\alpha) \geq 2$) and

$\beta = \{1, \dots, n\} \setminus \alpha$ be a partition of the node set $\mathcal{V} = \{1, \dots, n\}$. After an adequate permutation of its rows and columns, the Laplacian L of the graph $\mathcal{G}(A)$ can be rewritten as $L = \begin{bmatrix} L[\alpha] & L[\alpha, \beta] \\ L[\beta, \alpha] & L[\beta] \end{bmatrix}$, where we denote $L[\alpha, \beta]$ the submatrix of L determined by the index sets α and β , and $L[\alpha] := L[\alpha, \alpha]$ the principal submatrix of L determined by the index set α .

If $L[\beta]$ is nonsingular, the Schur complement of $L[\beta]$ in L is given by $L/L[\beta] := L[\alpha] - L[\alpha, \beta]L[\beta]^{-1}L[\beta, \alpha]$. In the context of electrical networks, where α and β are referred to as boundary (or terminal) and interior nodes, this procedure is denoted Kron reduction (see e.g., [25, 26]) and it yields a matrix $L_r := L/L[\beta]$, denoted Kron-reduced matrix, which is still a Laplacian of a weighted, undirected graph \mathcal{G}_r (see [25] for details and properties of L_r).

If $\mathcal{G}(A)$ is signed, when α is chosen as the set of nodes incident to edges with negative weight it is shown in [27] that $L[\beta]$ is positive definite and that L is psd of corank 1 if and only if L_r is psd of corank 1.

3 Pseudoinverse of eventually exponentially positive Laplacians

In this section we study the connection between the marginal stability and eventual positivity of the Laplacian L and of its pseudoinverse L^\dagger . If the network is undirected and signed, or if the network is directed, signed and weight balanced, we show that $-L$ is eventually exponentially positive if and only if L^\dagger is eventually exponentially positive.

3.1 Directed signed network case

Assume that the graph $\mathcal{G}(A) = (\mathcal{V}, \mathcal{E}, A)$ is directed and without loops, which means that the adjacency matrix A is with null diagonal.

When the graph is weight balanced, the Laplacian is a EP-matrix since $\ker(L) = \ker(L^T) = \text{span}\{\mathbf{1}\}$. In this case, it is shown in [28] that $-L$ is eventually exponentially positive if and only if $-L$ is marginally stable (of corank 1). In addition, if the Laplacian is a normal matrix, then eventual exponential positivity of L is equivalent to that of its symmetric part.

Theorem 2. Consider a signed digraph $\mathcal{G}(A)$ such that the corresponding Laplacian L is weight balanced. Then, the following conditions are equivalent:

- (i) $-L$ is eventually exponentially positive;
- (ii) $-L$ is marginally stable of corank 1.

Furthermore, if L is normal then (i) and (ii) are equivalent to

$$(iii) \quad L_s = \frac{L+L^T}{2} \text{ is psd of corank 1.}$$

Proof. (i) \iff (ii) See [28, Corollary 2].

(ii) \iff (iii): If L is normal, then there exists an orthonormal matrix U such that $L = UDU^T$, where, if μ_1, \dots, μ_k are the real eigenvalues of L and $\nu_1 \pm i\omega_1, \dots, \nu_\ell \pm$

$i\omega_\ell$ are its complex conjugate eigenvalues:

$$D = \mu_1 \oplus \cdots \oplus \mu_k \oplus \begin{bmatrix} \nu_1 & \omega_1 \\ -\omega_1 & \nu_1 \end{bmatrix} \oplus \cdots \oplus \begin{bmatrix} \nu_\ell & \omega_\ell \\ -\omega_\ell & \nu_\ell \end{bmatrix}$$

where \oplus indicates direct sum. It follows that $L_s = \frac{1}{2}(L + L^T) = \frac{1}{2}U(D + D^T)U^T$ and therefore that $\text{Re}[\lambda_i(L)] = \lambda_i(L_s)$. \square

Observe that Theorem 2 does not explicitly assume that $\mathcal{G}(A)$ is strongly connected. However, as we will show later in Lemma 2, any of the conditions (i) or (ii) implies that L is irreducible (i.e., that $\mathcal{G}(A)$ is strongly connected).

Remark 1. Corollary 2 of [28] claims that the equivalence (ii) \iff (iii) is valid in the more general case of weight balance L . Unfortunately that result is not true as the following Example 1 shows. A complication arises for instance from the fact that for L weight balanced but not normal L_s may acquire negative diagonal elements even if $-L$ is marginally stable. L_s with negative diagonal elements obviously cannot be psd. However, even when L_s has positive diagonal it is not guaranteed to be psd, see Example 2.

Example 1

In correspondence of

$$L = \begin{bmatrix} 0.15 & 0 & 0 & -0.15 \\ -0.23 & 0.15 & 0.15 & -0.07 \\ 0.01 & -0.12 & -0.03 & 0.14 \\ 0.07 & -0.03 & -0.12 & 0.08 \end{bmatrix}$$

it is $\Lambda(L) = \{0, 0.0901 \pm 0.199i, 0.169\}$, i.e., $-L$ is marginally stable of corank 1. Moreover, $L\mathbf{1} = L^T\mathbf{1} = 0$ and, for $d > 0.2647$, $B = dI - L \succ^{\vee} 0$. However, $\Lambda(L_s) = \{-0.0402, 0, 0.1248, 0.2655\}$, i.e., L_s is not psd.

Example 2

For

$$L = \begin{bmatrix} 0.23 & 0 & -0.28 & 0.05 \\ -0.01 & 0.03 & 0.02 & -0.04 \\ 0.05 & -0.03 & 0.04 & -0.06 \\ -0.27 & 0 & 0.22 & 0.05 \end{bmatrix}$$

it is $\Lambda(L) = \{0, 0.1443 \pm 0.1859i, 0.0514\}$, i.e., $-L$ is marginally stable of corank 1. Moreover, $L\mathbf{1} = L^T\mathbf{1} = 0$ and, for $d > 0.1919$, $B = dI - L \succ^{\vee} 0$. However, $\Lambda(L_s) = \{-0.0446, 0, 0.0404, 0.3441\}$, i.e., L_s is not psd.

As already mentioned, for signed Laplacians, irreducibility does not imply corank 1. When we have weight balance, however, the opposite is true.

Lemma 2. Let $\mathcal{G}(A)$ be a signed digraph with Laplacian L . If $-L$ is eventually exponentially positive or if L is weight balanced and of corank 1, then L is irreducible.

Proof. In both statements assume, by contradiction, that L is reducible, i.e., there exists a permutation matrix P s.t. $P^T L P = \begin{bmatrix} L_{11} & L_{12} \\ 0 & L_{22} \end{bmatrix}$.

Assume that $-L$ is eventually exponentially positive, i.e., $\exists d \geq 0$ s.t. $B = dI - L \stackrel{\vee}{>} 0$ (see Lemma 1). Then B is also reducible, since $P^T B P = \begin{bmatrix} dI - L_{11} & -L_{12} \\ 0 & dI - L_{22} \end{bmatrix}$.

It follows that $(P^T B P)^k = \begin{bmatrix} (dI - L_{11})^k & * \\ 0 & (dI - L_{22})^k \end{bmatrix}$ for all $k \geq 1$, i.e., $P^T B P$ is not eventually positive and, consequently, B is not eventually positive.

Assume that L is weight balanced of corank 1. Then $L\mathbf{1} = L^T\mathbf{1} = 0$ implies that $0 \in \Lambda(L_{11}^T) = \Lambda(L_{11})$ and that $0 \in \Lambda(L_{22})$. Consequently, L is not of corank 1. \square

Remark 2. For a signed digraph $\mathcal{G}(A)$ it holds that if L_s is psd of corank 1 then L is EP (see [29]) and hence weight balanced, and $-L$ is marginally stable of corank 1. Therefore, in Theorem 2, the assumption that the Laplacian L is a normal matrix is sufficient to prove that L_s is psd of corank 1 but not necessary. For example, for

$$L = \begin{bmatrix} 1 & 1 & -1 & -1 \\ -1 & 1 & 0 & 0 \\ -1 & -1 & 2 & 0 \\ 1 & -1 & -1 & 1 \end{bmatrix},$$

which is not normal, it is $\Lambda(L) = \{0, 1.5 \pm 1.323i, 2\}$, i.e., $-L$ is marginally stable of corank 1, and $\Lambda(L_s) = \{0, 0.7192, 1.5, 2.7808\}$, i.e., L_s is psd of corank 1. \square

We now show that the same statements of Theorem 2 hold also for the pseudoinverse L^\dagger of L . Moreover, we show that $-L$ is eventually exponentially positive (and marginally stable) if and only if $-L^\dagger$ is. These results are summarized in the following theorem.

Theorem 3. Let $\mathcal{G}(A)$ be a directed signed network such that the corresponding Laplacian L is weight balanced. Let L^\dagger be the weight balanced pseudoinverse of L . Then, the following conditions are equivalent:

- (i) $-L$ is eventually exponentially positive;
- (ii) $-L^\dagger$ is marginally stable of corank 1;
- (iii) $-L^\dagger$ is eventually exponentially positive.

Furthermore, if L (equivalently, L^\dagger) is normal then (i) \div (iii) are equivalent to

- (iv) $L_s^\dagger = \frac{L^\dagger + (L^\dagger)^T}{2}$ is psd of corank 1.

The proof of Theorem 3 relies on some considerations and propositions that we state first.

Since a weight balanced L (of corank 1) is a EP-matrix, its left and right orthogonal projectors onto $\text{range}(L)$ are identical and given by $\Pi = I - J$, with $J = \frac{\mathbf{1}\mathbf{1}^T}{n}$. Furthermore it is $\lim_{t \rightarrow \infty} e^{-Lt} = J$. The following properties for J can be found in [30] ([23, 31]) or computed straightforwardly.

Lemma 3. The matrix J has the following properties:

1. $J^k = J \forall k \in \mathbb{N}$ which implies that $(I - J)^k = (I - J) \forall k \in \mathbb{N}$;

2. $JL = LJ = 0$ which implies that $e^{-(L+J)} = e^{-L}e^{-J}$ and $Je^{-L} = e^{-L}J = J$;
3. $e^{-Jt} = I - J + Je^{-t}$ which implies that $Je^{-Jt} = e^{-Jt}J = Je^{-t}$.

We have the following properties for the Laplacian pseudoinverse.

Lemma 4. *If L is weight balanced and of corank 1, then L^\dagger is weight balanced and of corank 1. For it*

$$LL^\dagger = L^\dagger L = \Pi \quad (1)$$

$$L^\dagger \mathbf{1} = (L^\dagger)^T \mathbf{1} = 0 \quad (2)$$

$$L^\dagger \Pi = \Pi L^\dagger = L^\dagger \quad (3)$$

$$L^\dagger = (L + \gamma J)^{-1} - \frac{1}{\gamma} J \quad \forall \gamma \neq 0. \quad (4)$$

Furthermore, if L is normal then L^\dagger is normal.

Proof. Assume that L is weight balanced. Eq. (1)-(4) are all well-known for L symmetric, and follow easily also for EP matrices. They are proven here only for sake of completeness. Eq. (2) is a consequence of L commuting with L^\dagger . As for eq. (2), from $(L^\dagger L)^T = L^\dagger L$ it follows that $L^T(L^\dagger)^T \mathbf{1} = 0$. Since L is irreducible, $L^T v \neq 0$ for $v \neq c\mathbf{1}$ ($c \in \mathbb{R}$), hence it must be $(L^\dagger)^T \mathbf{1} = 0$ or $\mathbf{1}^T L^\dagger = 0$, i.e., L^\dagger has $\mathbf{1}$ as left eigenvector relative to 0. The proof for the right eigenvector is identical. Concerning eq. (3), from $L^\dagger \mathbf{1} = 0$ it is $L^\dagger \Pi = L^\dagger(I - \frac{\mathbf{1}\mathbf{1}^T}{n}) = L^\dagger$, and similarly for $\Pi L^\dagger = L^\dagger$. For eq. (4), since $L + \gamma J$ is nonsingular, as in [25], it is enough to show the following:

$$(L + \gamma J)(L^\dagger + \frac{1}{\gamma} J) = LL^\dagger + \gamma JL^\dagger + \frac{1}{\gamma} LJ + J^2 = \Pi + J = I - J + J = I,$$

where we have used the properties of Lemma 3.

Then, $\ker(L) = \ker(L^T) = \ker(L^\dagger) = \ker((L^\dagger)^T) = \text{span}\{\mathbf{1}\}$ and (4) imply that L^\dagger is weight balanced of corank 1. Notice that irreducibility of L and L^\dagger follows from Lemma 2.

Finally, we need to show that if L is normal then L^\dagger is normal. L normal, J symmetric and $LJ = L^TJ = JL = JL^T = 0$ imply $L + \gamma J$ normal, which means that $(L + \gamma J)^{-1}$ is also normal. Since J is symmetric (hence normal) and satisfies the properties of Lemma 3, to show that L^\dagger is normal it is sufficient to observe that $(L + \gamma J)^{-1}J = \frac{1}{\gamma}J = J(L + \gamma J)^{-1}$. \square

We are now ready to prove Theorem 3.

Proof. (i) \implies (ii) To show marginal stability of $-L^\dagger$, denote $\lambda_i(L)$ the eigenvalues of L , of eigenvectors $\mathbf{1}, v_2, \dots, v_n$. Using Theorem 2, since $-L$ is eventually exponentially positive then $-L$ is also marginally stable of corank 1, meaning that $0 = \lambda_1(L) < \text{Re}[\lambda_2(L)] \leq \dots \leq \text{Re}[\lambda_n(L)]$. Consider eq. (4) of Lemma 4. Choosing $\gamma \neq 0$, since J is the orthogonal projection onto $\ker(L) = \ker(L^T) = \text{span}\{\mathbf{1}\}$, the effect of adding γJ to L is only to shift the 0 eigenvalue to γ , while $\lambda_2(L), \dots, \lambda_n(L)$ are unchanged (see [24, Thm 2.4.10.1]). For the nonsingular $L + \gamma J$ the inverse $(L + \gamma J)^{-1}$ has eigenvalues $\frac{1}{\gamma}, \frac{1}{\lambda_2(L)}, \dots, \frac{1}{\lambda_n(L)}$ of eigenvectors $\mathbf{1}, v_2, \dots, v_n$. From

orthogonality, $(L + \gamma J)^{-1} - \frac{1}{\gamma}J$ only shifts the $\frac{1}{\gamma}$ eigenvalue back to the origin without touching the other eigenvalues.

(i) \Rightarrow (iii) Assume that $-L$ is eventually exponentially positive, that is, $-L$ is marginally stable of corank 1 (see Theorem 2). Then $-L^\dagger$ is also marginally stable of corank 1, see Lemma 4 and proof (i) \Rightarrow (ii). To prove that $-L^\dagger$ is eventually exponentially positive, we can use Theorem 2. The proof is here reported for completeness. In particular, from Lemma 4, we know that $-L^\dagger$ is marginally stable with $0 = \lambda_1(L^\dagger) < \operatorname{Re}[\lambda_2(L^\dagger)] \leq \dots \leq \operatorname{Re}[\lambda_n(L^\dagger)]$ and with $\mathbf{1}$ as left/right eigenvector for 0. If we choose $d > \max_{i=2,\dots,n} \frac{|\lambda_i(L^\dagger)|^2}{2\operatorname{Re}[\lambda_i(L^\dagger)]}$, then $B = dI - L^\dagger$ has $\rho(B) = d$ as a simple eigenvalue of eigenvector $\mathbf{1}$ and so does B^T . Hence $B, B^T \in \mathcal{PF}$, or, from Theorem 1, $B \stackrel{\vee}{>} 0$, i.e., B is eventually positive. Hence from Lemma 1 L^\dagger is eventually exponentially positive.

(iii) \Rightarrow (i) Since L^\dagger is weight balanced of corank 1 with $\operatorname{span}\{\mathbf{1}\} = \ker(L^\dagger) = \ker((L^\dagger)^T)$, it is itself a signed Laplacian. The argument can be proven in a similar way as the opposite direction, observing that $L = (L^\dagger)^\dagger$.

(iv) Assume now that L is normal or, equivalently, that L^\dagger is normal (see Lemma 4). Since L normal implies L weight balanced, the statements (i), (ii), and (iii) are still equivalent. To show the equivalence with (iv) it is sufficient to apply Theorem 2 on L^\dagger since L^\dagger is itself a normal signed Laplacian of corank 1. \square

The following corollary, characterizing the class of eventually exponentially positive Laplacian matrices, follows directly from Theorem 3.

Corollary 1. *The class of eventually exponentially positive, weight balanced Laplacian matrices is closed under the pseudoinverse operation.*

The class of eventually exponentially positive, normal Laplacian matrices is closed under the pseudoinverse and the symmetrization operation (the latter intended as the operation of taking the symmetric part).

Remark 3. Notice that the operations of pseudoinverse and of symmetrization do not commute, i.e., $L_s^\dagger = \frac{L^\dagger + (L^\dagger)^T}{2} \neq (L_s)^\dagger = (\frac{L+L^T}{2})^\dagger$, even in the case of a normal Laplacian L . Indeed, let $L = UDU^T$ with U orthonormal and D as in the proof of Theorem 2. Without lack of generality, assume that the first column of U is $\mathbf{1}\frac{1}{\sqrt{n}}$, which means that $D = 0 \oplus \bar{D}$ where $\bar{D} = \mu_2 \oplus \dots \oplus \mu_k \oplus \begin{bmatrix} \nu_1 & \omega_1 \\ -\omega_1 & \nu_1 \end{bmatrix} \oplus \dots \oplus \begin{bmatrix} \nu_\ell & \omega_\ell \\ -\omega_\ell & \nu_\ell \end{bmatrix}$ is nonsingular (\oplus denotes the direct sum). Then

$$L_s = U \begin{bmatrix} 0 & 0 \\ 0 & \frac{\bar{D} + \bar{D}^T}{2} \end{bmatrix} U^T, \quad L^\dagger = U \begin{bmatrix} 0 & 0 \\ 0 & \bar{D}^{-1} \end{bmatrix} U^T, \quad L_s^\dagger = U \begin{bmatrix} 0 & 0 \\ 0 & \frac{\bar{D}^{-1} + \bar{D} - T}{2} \end{bmatrix} U^T,$$

and

$$(L_s)^\dagger = U \begin{bmatrix} 0 & 0 \\ 0 & (\frac{\bar{D} + \bar{D}^T}{2})^{-1} \end{bmatrix} U^T \neq L_s^\dagger.$$

Example 3

For the Laplacian L of Example 1, we obtain

$$L^\dagger = \begin{bmatrix} 2.25 & -1.86 & -0.19 & -0.19 \\ -1.42 & 1.58 & -5.64 & 5.47 \\ 1.92 & 0.47 & 4.36 & -6.75 \\ -2.75 & -0.19 & 1.47 & 1.47 \end{bmatrix}.$$

It is $\Lambda(L^\dagger) = \{0, 1.8888 \pm 4.1709i, 5.8891\}$ (notice that $\lambda_i(L^\dagger) = \frac{1}{\lambda_i(L)}$, $i = 2, 3, 4$), i.e., $-L^\dagger$ is marginally stable of corank 1. Moreover, $L^\dagger \mathbf{1} = (L^\dagger)^T \mathbf{1} = 0$ and, for $d > 5.5495$, $B = dI - L^\dagger \stackrel{\vee}{>} 0$. However, $\Lambda(L_s^\dagger) = \{-1.1164, 0, 2.0926, 8.6904\}$, i.e., L_s^\dagger is not psd.

Example 4

The signed Laplacian

$$L = \begin{bmatrix} 0.282 & -0.072 & 0.191 & -0.401 \\ -0.072 & 0.252 & 0.008 & -0.189 \\ -0.401 & -0.189 & 0.297 & 0.293 \\ 0.191 & 0.008 & -0.496 & 0.297 \end{bmatrix}$$

is normal, irreducible and it is $\Lambda(L) = \{0, 0.3983 \pm 0.5920i, 0.3311\}$, i.e., $-L$ is marginally stable of corank 1. In accordance with Theorem 3, $\Lambda(L^\dagger) = \{0, 0.7823 \pm 1.1628i, 3.0204\}$, i.e., $-L^\dagger$ is marginally stable of corank 1, $\Lambda(L_s) = \{0, 0.3983, 0.3983, 0.3311\}$, i.e., L_s is psd of corank 1, and $\Lambda(L_s^\dagger) = \{0, 0.7823, 0.7823, 3.0204\}$, i.e., L_s^\dagger is psd of corank 1.

When the weight balanced digraph is nonnegative, the normality assumption in Theorem 3(iv) can be dropped.

Theorem 4. Let $\mathcal{G}(A)$ be a strongly connected nonnegative ($A \geq 0$) digraph such that the corresponding Laplacian L is weight balanced. Let L^\dagger be the irreducible weight balanced pseudoinverse of L . Then $L_s^\dagger = \frac{L^\dagger + (L^\dagger)^T}{2}$ is psd of corank 1.

Proof. In [28, Corollary 1] it is shown that when $A \geq 0$, $L_s = \frac{L+L^T}{2}$ is psd of corank 1 if and only if $\mathcal{G}(A)$ is weight balanced. Using (4) (see Lemma 4) we can write

$$\begin{aligned} L_s^\dagger &= \frac{(L + \gamma J)^{-1} + (L^T + \gamma J)^{-1}}{2} - \frac{J}{\gamma} \\ &= (L + \gamma J)^{-1} \frac{L^T + \gamma J + L + \gamma J}{2} (L^T + \gamma J)^{-1} - \frac{J}{\gamma} \\ &= (L + \gamma J)^{-1} (L_s + \gamma J) (L^T + \gamma J)^{-1} - (L + \gamma J)^{-1} \frac{(L + \gamma J) J (L^T + \gamma J)}{\gamma} (L^T + \gamma J)^{-1} \\ &\stackrel{*}{=} (L + \gamma J)^{-1} (L_s + \gamma J - \gamma J) (L^T + \gamma J)^{-1} \\ &= (L + \gamma J)^{-1} L_s (L + \gamma J)^{-T}, \end{aligned}$$

where in the step marked * we have used the properties of J listed in Lemma 3. Hence, if L_s is psd (of corank 1) so is L_s^\dagger . \square

3.2 Undirected signed network case

Assume that the graph $\mathcal{G}(A) = (\mathcal{V}, \mathcal{E}, A)$ is undirected, connected and without self-loops, which means that the adjacency matrix A is irreducible, symmetric and with null diagonal. Consequently, the Laplacian $L = \Delta - A$ is a symmetric matrix, meaning that $\ker(L^T) = \ker(L) = \text{span}\{\mathbf{1}\}$ and that $\text{range}(L) = \text{span}\{\mathbf{1}\}^\perp$. In this case, Theorem 3 proves the equivalence between eventual exponential positivity of $-L$ and positive semidefiniteness of L (of corank 1).

For undirected networks the conditions discussed in Theorem 3 can be applied to the Kron reduction of $\mathcal{G}(A)$, as described in Section 2.4, showing that if $-L$ is eventually exponentially positive then $-L_r$ (where L_r is the Kron reduced matrix) is also eventually exponentially positive.

Theorem 5. *Let $\mathcal{G}(A)$ be an undirected, connected, signed network with Laplacian L . Let α (with $\text{card}(\alpha) \in [2, n-1]$) and $\beta = \{1, \dots, n\} \setminus \alpha$ be a partition of the node set \mathcal{V} . Let \mathcal{G}_r be the signed undirected graph obtained by applying the Kron reduction on \mathcal{G} , and let $L_r = L/L[\beta]$ be its (symmetric) Laplacian.*

Consider the following conditions:

- (i) $-L$ is eventually exponentially positive;
- (ii) $-L_r$ is psd of corank 1;
- (iii) $-L_r$ is eventually exponentially positive.

If L satisfies (i) then L_r satisfies (ii) and (iii).

Furthermore, if α is the set of nodes incident to negatively weighted edges and $\beta = \{1, \dots, n\} \setminus \alpha$, then the conditions (i), (ii), (iii) are equivalent.

Proof. Let α (with $\text{card}(\alpha) \in [2, n-1]$) and $\beta = \{1, \dots, n\} \setminus \alpha$ be a partition of the node set \mathcal{V} meaning that, after an adequate permutation, L can be rewritten as $L = \begin{bmatrix} L[\alpha] & L[\alpha, \beta] \\ L[\beta, \alpha] & L[\beta] \end{bmatrix}$. Let $L_r = L/L[\beta] = L[\alpha] - L[\alpha, \beta]L[\beta]^{-1}L[\beta, \alpha] \in \mathbb{R}^{\text{card}(\alpha) \times \text{card}(\alpha)}$ be the Kron reduced matrix. Observe that L_r is symmetric and that $\mathbf{1}_{\text{card}(\alpha)} \in \ker(L_r)$ (see also [25, Lemma II.1]), meaning that L_r is itself a signed Laplacian.

(i) \Rightarrow (ii) \Leftrightarrow (iii). Assume that $-L$ is eventually exponentially positive or, equivalently, that L is psd of corank 1 (see Theorem 2). Then $L[\beta]$ is also psd as it is a principal submatrix of L . In what follows we prove (by contradiction) that $L[\beta]$ is actually pd, since L is irreducible and has the row and column inclusion property.

Let $\text{card}(\beta) = 1$ and assume, by contradiction, that $L[\beta] = \Delta[\beta] = 0$. However, L psd means that L has the row and column inclusion property, i.e., if the diagonal element $\Delta[\beta]$ is zero then $A[\alpha, \beta] = 0$ and $A[\beta, \alpha] = 0$, which contradicts the hypothesis that L (and A) is irreducible. Hence, $L[\beta] > 0$ (pd). Now we repeat the same argument for $1 < \text{card}(\beta) \leq n-2$: suppose by contradiction that $L[\beta]$

is psd, i.e., there exists a vector $v \in \mathbb{R}^{\text{card}(\beta)}$ s.t. $L[\beta]v = 0$. Then $\bar{v} = \begin{bmatrix} 0 \\ v \end{bmatrix}$ is s.t. $L\bar{v} = 0$ (since $\bar{v}^T L \bar{v} = 0$), which contradicts the hypothesis that L has corank 1 since $\mathbf{1} \in \ker(L)$ and $\bar{v} \notin \text{span}\{\mathbf{1}\}$ (notice that if $v = \mathbf{1}_{\text{card}(\beta)}$, then either $A[\beta, \alpha]$ is the zero matrix - in contradiction with the hypothesis that L is irreducible -, or $\begin{bmatrix} \mathbf{1}_{\text{card}(\alpha)} \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ \mathbf{1}_{\text{card}(\beta)} \end{bmatrix} \in \ker(L)$ - in contradiction with the hypothesis that L has corank 1). Therefore, $L[\beta]$ is pd.

Rewrite L as

$$L = \begin{bmatrix} I & L[\alpha, \beta]L[\beta]^{-1} \\ 0 & I \end{bmatrix} \begin{bmatrix} L_r & 0 \\ 0 & L[\beta] \end{bmatrix} \begin{bmatrix} I & 0 \\ L[\beta]^{-1}L[\beta, \alpha] & I \end{bmatrix}$$

where $L[\alpha, \beta]L[\beta]^{-1} = (L[\beta]^{-1}L[\beta, \alpha])^T$. Then, applying Sylvester's law of inertia, L psd of corank 1 and $L[\beta]$ pd imply L_r psd of corank 1 or, equivalently (from Theorem 2), $-L_r$ eventually exponentially positive.

(i) \iff (ii) \iff (iii). Let α be the set of nodes incident to negatively weighted edges. In what follows, the steps marked by the symbol \star follow from Theorem 2 while the step marked by the symbol \triangle from [32, Theorem 1].

$$\begin{aligned} & -L \text{ eventually exponentially positive} \\ \stackrel{\star}{\iff} & L \text{ psd of corank 1} \\ \stackrel{\triangle}{\iff} & L_r \text{ psd of corank 1} \\ \stackrel{\star}{\iff} & -L_r \text{ eventually exponentially positive.} \end{aligned} \quad \square$$

Similarly to Corollary 1, from Theorem 3 and Theorem 5 we obtain the following characterization of the class of eventually exponentially positive Laplacian matrices of undirected graphs.

Corollary 2. *The class of eventually exponentially positive, irreducible, symmetric Laplacian matrices is closed under the pseudoinverse operation and the operation of Kron reduction.*

4 Electrical networks and effective resistance

A resistive electrical network can be represented as a graph $\mathcal{G}(A) = (\mathcal{V}, \mathcal{E}, A)$ where each weight a_{ij} represents the inverse of the resistance between the two nodes (i.e., the conductance of the transmission): $a_{ij} = \frac{1}{r_{ij}}$, see [8, 10] and [26] for an overview. The notion of effective resistance between a pair of nodes (see e.g. [26]) is related to the pseudoinverse of the Laplacian associated to the electrical network. When the network is connected, undirected and nonnegative, its Laplacian (and its pseudoinverse) is known to be psd of corank 1, which means that the effective resistance between two nodes is well-defined (see e.g. [10] for its properties). Extensions to signed graphs and negative resistances have been investigated in [27, 32–35], where positive semidefiniteness of the Laplacian is expressed in terms of effective resistance.

In what follows we make use of L_s^\dagger to extend the notion of effective resistance to directed (strongly connected) signed networks whose Laplacian L is a normal matrix and $-L$ is eventually exponentially positive.

Definition 4. The effective resistance between two nodes $i, j \in \{1, \dots, n\}$ of a signed digraph whose corresponding Laplacian L is normal and $-L$ is eventually exponentially positive, is given by

$$R_{ij} = (e_i - e_j)^T L_s^\dagger (e_i - e_j), \quad (5)$$

where $L_s^\dagger = \frac{L^\dagger + (L^\dagger)^T}{2}$ and L^\dagger is the pseudoinverse of L . The effective resistance matrix $R = [R_{ij}]$ is defined as

$$R = D_{L_s^\dagger} \mathbf{1} \mathbf{1}^T + \mathbf{1} \mathbf{1}^T D_{L_s^\dagger} - 2L_s^\dagger \quad (6)$$

where $D_{L_s^\dagger} = \text{diag}\{[L_s^\dagger]_{11}, \dots, [L_s^\dagger]_{nn}\}$ is a diagonal matrix whose elements are the diagonal elements of L_s^\dagger . The total effective resistance is defined as

$$R_{\text{tot}} = \frac{1}{2} \mathbf{1}^T R \mathbf{1}. \quad (7)$$

In the literature on undirected networks, the total effective resistance (7) is also called “weighted effective graph resistance” [36] or “Kirchhoff index” [9], and represents the overall transport capability of the graph [14].

Remark 4. Observe that if the graph is undirected, eq. (5) reduces to the standard notion of effective resistance since $L_s^\dagger = L^\dagger$.

The effective resistance (5), as its counterpart for undirected graphs (see [8, 10, 26]), is still nonnegative and symmetric, its square root is a metric, and the effective resistance matrix (6) is a Euclidean distance matrix, i.e., it has nonnegative elements, zero diagonal elements and it is negative semidefinite on $\mathbf{1}^\perp$ [10]. The last part of the proof of the next lemma follows [10, Section 2.8] and is here reported for completeness.

Lemma 5. *The square root of the effective resistance (5) between two nodes $i, j \in \{1, \dots, n\}$ of a signed digraph with normal Laplacian L is a metric: it is nonnegative, symmetric and it satisfies the triangle inequality. The effective resistance matrix (6) is a Euclidean distance matrix.*

Proof. Theorem 3 shows that for a signed digraph with normal Laplacian L s.t. $-L$ is eventually exponentially positive, the matrix L_s^\dagger is itself a signed Laplacian and it is psd of corank 1 with $\ker(L_s^\dagger) = \text{span}\{\mathbf{1}\}$. Since R_{ij} is a quadratic form generated by L_s^\dagger , then

$$\begin{aligned} R_{ij} &= (e_i - e_j)^T L_s^\dagger (e_i - e_j) = \|(L_s^\dagger)^{\frac{1}{2}}(e_i - e_j)\|_2^2 \\ &= \|(L_s^\dagger)^{\frac{1}{2}}(e_j - e_i)\|_2^2 = (e_j - e_i)^T L_s^\dagger (e_j - e_i) = R_{ji} \end{aligned}$$

for all $i, j = 1, \dots, n$, and

$$R_{ij} = (e_i - e_j)^T L_s^\dagger (e_i - e_j) = \|(L_s^\dagger)^{\frac{1}{2}}(e_i - e_j)\|_2^2 \geq 0$$

for all $i, j = 1, \dots, n$, with $R_{ij} = 0$ if and only if $i = j$ (since $e_i - e_j \in \text{span}\{\mathbf{1}^\perp\}$ when $i \neq j$). Moreover,

$$\begin{aligned} \sqrt{R_{ik}} + \sqrt{R_{kj}} &= \|(L_s^\dagger)^{\frac{1}{2}}(e_i - e_k)\|_2 + \|(L_s^\dagger)^{\frac{1}{2}}(e_k - e_j)\|_2 \\ &\geq \|(L_s^\dagger)^{\frac{1}{2}}(e_i - e_k) + (L_s^\dagger)^{\frac{1}{2}}(e_k - e_j)\|_2 \\ &= \|(L_s^\dagger)^{\frac{1}{2}}(e_i - e_j)\|_2 = \sqrt{R_{ij}} \end{aligned}$$

for all $i, j, k = 1, \dots, n$, i.e., the triangle inequality holds.

Finally, to prove that R is an Euclidean distance matrix we need to show that $x^T Rx \leq 0$ for all $x \perp \mathbf{1}$:

$$x^T Rx = x^T (D_{L_s^\dagger} \mathbf{1} \mathbf{1}^T + \mathbf{1} \mathbf{1}^T D_{L_s^\dagger} - 2L_s^\dagger) x = -2x^T L_s^\dagger x \leq 0,$$

since L_s^\dagger is psd with $\ker(L_s^\dagger) = \text{span}\{\mathbf{1}\}$. \square

Note that if we consider only nonnegative digraphs then the normality assumption of the Laplacian can be replaced by the less restrictive weight balanced assumption when defining the effective resistance in (5). Indeed, Theorem 4 shows that if the digraph is nonnegative and strongly connected then L_s^\dagger is psd of corank 1.

Proposition 1. Consider a nonnegative strongly connected weight balanced digraph $\mathcal{G}(A)$ (with $A \geq 0$). Then $R_{ij} \geq 0$ for all $i, j = 1, \dots, n$ and $R_{\text{tot}} \geq 0$.

Another generalization of the notion of effective resistance for directed, strongly connected, nonnegative networks is introduced in [15, 22]. The authors use the fact that the Laplacian L is marginally stable and its projection on $\mathbf{1}^\perp$, denoted $\bar{L} = QLQ^T$ (where the rows of $Q \in \mathbb{R}^{n-1 \times n}$ form an orthonormal basis for $\mathbf{1}^\perp$), is Hurwitz stable, to define the effective resistance between nodes i and j as $\tilde{R}_{ij} = (e_i - e_j)^T X (e_i - e_j)$, where $X = 2Q^T S Q$ and S is the pd solution of the Lyapunov equation $\bar{L}S + S\bar{L}^T = I_{n-1}$. The Kirchhoff index is then defined as $K_f = \sum_{i < j} \tilde{R}_{ij}$.

If we consider digraphs $\mathcal{G}(A)$ whose Laplacian is a normal matrix, K_f reduces to $K_f = n \sum_{i=2}^n \frac{1}{\text{Re}[\lambda_i(L)]}$ and we can show that it provides an upper bound for R_{tot} defined in (7).

$$\begin{aligned} R_{\text{tot}} &= n \cdot \text{trace}(L_s^\dagger) = n \cdot \sum_{i=2}^n \lambda_i(L_s^\dagger) \\ &= n \cdot \sum_{i=2}^n \text{Re}[\lambda_i(L^\dagger)] = n \cdot \sum_{i=2}^n \text{Re}\left[\frac{1}{\lambda_i(L)}\right] \\ &= n \cdot \sum_{i=2}^n \frac{\text{Re}[\lambda_i(L)]}{|\lambda_i(L)|^2} \\ &= n \cdot \sum_{i=2}^n \frac{1}{\text{Re}[\lambda_i(L)] \left(1 + \left(\frac{\text{Im}[\lambda_i(L)]}{\text{Re}[\lambda_i(L)]}\right)^2\right)} \\ &\leq n \cdot \sum_{i=2}^n \frac{1}{\text{Re}[\lambda_i(L)]} = K_f, \end{aligned}$$

with equality only if $\mathcal{G}(A)$ is undirected (notice that L normal and non-symmetric means $n \geq 3$).

Example 5

Let $\mathcal{G}(A)$ be a nonnegative, unweighted, directed, cycle graph, whose Laplacian L is a normal matrix with eigenvalues $1 + e^{i\theta_k}$, with $\theta_k = \pi(1 - \frac{2k}{n})$, $k = 1, \dots, n-1$. Then, $K_f = \frac{n(n^2-1)}{6}$ (see e.g. [15]),

$$R_{\text{tot}} = n \cdot \sum_{k=2}^n \text{Re}\left[\frac{1}{\lambda_k(L)}\right] = n \cdot \sum_{k=2}^n \frac{1 + \cos\theta_k}{(1 + \cos\theta_k)^2 + \sin^2\theta_k} = n \cdot \sum_{k=2}^n \frac{1}{2} = \frac{n(n-1)}{2},$$

and we obtain $R_{\text{tot}} \leq K_f$ for all $n \geq 2$.

5 Conclusions and future work

For signed Laplacians which are weight balanced, marginal stability (of corank 1) is equivalent to eventual exponential stability. This work shows that the class of eventually exponentially positive, weight balanced Laplacians is closed under the pseudoinverse operation and, therefore, it provides a natural embedding for the usual nonnegative Laplacian. As a byproduct we get conditions for checking the marginal stability of the pseudoinverse of signed Laplacians. Moreover, closure under the symmetrization operation can be proven when this class is restricted to Laplacians that are also normal matrices. The normality assumption is a sufficient condition and it remains to be investigated if it can be relaxed.

In addition, we would like to gain a better understanding of the set of eventually exponentially positive, weight balanced Laplacians and its properties. For instance, it is easy to observe that it is not a convex cone, not even if we consider normal matrices (but the intuition is that this set is actually a convex cone, without the origin, if we restrict to undirected graphs). However, similarly to [20], it is possible to show that it is path-wise connected. These considerations, among other directions, will be investigated in a future paper.

Bibliography

- [1] A. Fontan and C. Altafini, “On the properties of Laplacian pseudoinverses,” in *60th IEEE Conference on Decision and Control*. Austin, TX, USA: IEEE, 2021.
- [2] F. R. K. Chung, *Spectral Graph Theory*, ser. CBMS Number 92. American Mathematical Society, 1997.
- [3] R. Agaev and P. Chebotarev, “On the spectra of nonsymmetric Laplacian matrices,” *Linear Algebra and Its Applications*, vol. 399, no. 1-3, pp. 157–168, 2005.
- [4] R. Olfati-Saber and R. Murray, “Consensus Problems in Networks of Agents With Switching Topology and Time-Delays,” *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1520–1533, sep 2004.
- [5] C. Altafini and G. Lini, “Predictable dynamics of opinion forming for networks with antagonistic interactions,” *IEEE Transactions on Automatic Control*, vol. 60, no. 2, pp. 342–357, feb 2015.
- [6] J. C. Bronski and L. DeVille, “Spectral Theory for Dynamics on Graphs Containing Attractive and Repulsive Interactions,” *SIAM Journal on Applied Mathematics*, vol. 74, no. 1, pp. 83–105, jan 2014.
- [7] L. Pan, H. Shao, and M. Mesbahi, “Laplacian dynamics on signed networks,” in *55th IEEE Conference on Decision and Control*. Las Vegas, USA: IEEE, dec 2016, pp. 891–896.
- [8] D. J. Klein and M. Randić, “Resistance distance,” *Journal of Mathematical Chemistry*, vol. 12, no. 1, pp. 81–95, 1993.
- [9] W. Xiao and I. Gutman, “Resistance distance and Laplacian spectrum,” *Theoretical Chemistry Accounts*, vol. 110, no. 4, pp. 284–289, 2003.
- [10] A. Ghosh, S. Boyd, and A. Saberi, “Minimizing Effective Resistance of a Graph,” *SIAM Review*, vol. 50, no. 1, pp. 37–66, jan 2008.
- [11] A. K. Chandra, P. Raghavan, W. L. Ruzzo, R. Smolensky, and P. Tiwari, “The electrical resistance of a graph captures its commute and cover times,” *Computational Complexity*, vol. 6, no. 4, pp. 312–340, 1996.
- [12] J. L. Palacios, “Resistance distance in graphs and random walks,” *International Journal of Quantum Chemistry*, vol. 81, no. 1, pp. 29–33, 2001.
- [13] D. Boley, G. Ranjan, and Z. L. Zhang, “Commute times for a directed graph using an asymmetric Laplacian,” *Linear Algebra and Its Applications*, vol. 435, no. 2, pp. 224–242, 2011.
- [14] P. Van Mieghem, K. Devriendt, and H. Cetinay, “Pseudoinverse of the Laplacian and best spreader node in a network,” *Physical Review E*, vol. 96, no. 3, pp. 1–22, 2017.

- [15] G. F. Young, L. Scardovi, and N. E. Leonard, “Robustness of noisy consensus dynamics with directed communication,” in *2010 American Control Conference (ACC)*. Baltimore, MD, USA: IEEE, jun 2010, pp. 6312–6317.
- [16] G. F. Young, L. Scardovi, and N. E. Leonard, “Rearranging trees for robust consensus,” in *50th IEEE Conference on Decision and Control and European control Conference*. Orlando, FL, USA: IEEE, 2011, pp. 1000–1005.
- [17] G. Lindmark and C. Altafini, “Investigating the effect of edge modifications on networked control systems,” arXiv:2007.13713, 2020.
- [18] D. Noutsos, “On Perron–Frobenius property of matrices having some negative entries,” *Linear Algebra and its Applications*, vol. 412, no. 2-3, pp. 132–153, jan 2006.
- [19] D. Noutsos and M. J. Tsatsomeros, “Reachability and Holdability of Nonnegative States,” *SIAM Journal on Matrix Analysis and Applications*, vol. 30, no. 2, pp. 700–712, jan 2008.
- [20] C. R. Johnson and P. Tarazaga, “On matrices with Perron-Frobenius properties and some negative entries,” *Positivity*, vol. 8, no. 4, pp. 327–338, 2004.
- [21] G. Shi, C. Altafini, and J. S. Baras, “Dynamics over Signed Networks,” *SIAM Review*, vol. 61, no. 2, pp. 229–257, jan 2019.
- [22] G. F. Young, L. Scardovi, and N. E. Leonard, “A New Notion of Effective Resistance for Directed Graphs—Part I: Definition and Properties,” *IEEE Transactions on Automatic Control*, vol. 61, no. 7, pp. 1727–1736, jul 2016.
- [23] C. D. Meyer, *Matrix Analysis and Applied Linear Algebra*. Society for Industrial & Applied Mathematics, 2000.
- [24] R. A. Horn and C. R. Johnson, *Matrix analysis*, 2nd ed. Cambridge University Press, 2013.
- [25] F. Dörfler and F. Bullo, “Kron reduction of graphs with applications to electrical networks,” *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 60, no. 1, pp. 150–163, 2013.
- [26] F. Dorfler, J. W. Simpson-Porco, and F. Bullo, “Electrical Networks and Algebraic Graph Theory: Models, Properties, and Applications,” *Proceedings of the IEEE*, vol. 106, no. 5, pp. 977–1005, 2018.
- [27] Y. Chen, S. Z. Khong, and T. T. Georgiou, “On the definiteness of graph Laplacians with negative weights: Geometrical and passivity-based approaches,” in *2016 American Control Conference (ACC)*. Boston, MA, USA: IEEE, jul 2016, pp. 2488–2493.
- [28] C. Altafini, “Investigating stability of Laplacians on signed digraphs via eventual positivity,” in *58th IEEE Conference on Decision and Control*. Nice, France: IEEE, dec 2019, pp. 5044–5049.

- [29] T. O. Lewis and T. G. Newman, “Pseudoinverses of Positive Semidefinite Matrices,” *SIAM Journal on Applied Mathematics*, vol. 16, no. 4, pp. 701–703, 1968.
- [30] F. Bullo, *Lectures on Nonlinear Network Systems* (ed. 1.4). Kindle Direct Publishing, 2020. <http://motion.me.ucsb.edu/book-lns>
- [31] A. Ben-Israel and T. N. E. Greville, *Generalized Inverses*, 2nd ed., ser. CMS Books in Mathematics. New York: Springer-Verlag, 2003.
- [32] W. Chen, D. Wang, J. Liu, Y. Chen, S. Z. Khong, T. Basar, K. H. Johansson, and L. Qiu, “On Spectral Properties of Signed Laplacians with Connections to Eventual Positivity,” *IEEE Transactions on Automatic Control*, vol. 9286, no. c, pp. 1–1, 2020.
- [33] D. Zelazo and M. Burger, “On the definiteness of the weighted Laplacian and its connection to effective resistance,” in *53rd IEEE Conference on Decision and Control*. IEEE, dec 2014, pp. 2895–2900.
- [34] D. Zelazo and M. Burger, “On the robustness of uncertain consensus networks,” *IEEE Transactions on Control of Network Systems*, vol. 4, no. 2, pp. 170–178, 2017.
- [35] W. Chen, J. Liu, Y. Chen, S. Z. Khong, D. Wang, T. Basar, L. Qiu, and K. H. Johansson, “Characterizing the positive semidefiniteness of signed Laplacians via Effective Resistances,” in *55th IEEE Conference on Decision and Control*. Las Vegas, USA: IEEE, dec 2016, pp. 985–990.
- [36] W. Ellens, F. M. Spieksma, P. Van Mieghem, A. Jamakovic, and R. E. Kooij, “Effective graph resistance,” *Linear Algebra and Its Applications*, vol. 435, no. 10, pp. 2491–2506, 2011.

PhD Dissertations
Division of Automatic Control
Linköping University

- M. Millnert:** Identification and control of systems subject to abrupt changes. Thesis No. 82, 1982. ISBN 91-7372-542-0.
- A. J. M. van Overbeek:** On-line structure selection for the identification of multivariable systems. Thesis No. 86, 1982. ISBN 91-7372-586-2.
- B. Bengtsson:** On some control problems for queues. Thesis No. 87, 1982. ISBN 91-7372-593-5.
- S. Ljung:** Fast algorithms for integral equations and least squares identification problems. Thesis No. 93, 1983. ISBN 91-7372-641-9.
- H. Jonson:** A Newton method for solving non-linear optimal control problems with general constraints. Thesis No. 104, 1983. ISBN 91-7372-718-0.
- E. Trulsson:** Adaptive control based on explicit criterion minimization. Thesis No. 106, 1983. ISBN 91-7372-728-8.
- K. Nordström:** Uncertainty, robustness and sensitivity reduction in the design of single input control systems. Thesis No. 162, 1987. ISBN 91-7870-170-8.
- B. Wahlberg:** On the identification and approximation of linear systems. Thesis No. 163, 1987. ISBN 91-7870-175-9.
- S. Gunnarsson:** Frequency domain aspects of modeling and control in adaptive systems. Thesis No. 194, 1988. ISBN 91-7870-380-8.
- A. Isaksson:** On system identification in one and two dimensions with signal processing applications. Thesis No. 196, 1988. ISBN 91-7870-383-2.
- M. Viberg:** Subspace fitting concepts in sensor array processing. Thesis No. 217, 1989. ISBN 91-7870-529-0.
- K. Forsman:** Constructive commutative algebra in nonlinear control theory. Thesis No. 261, 1991. ISBN 91-7870-827-3.
- F. Gustafsson:** Estimation of discrete parameters in linear systems. Thesis No. 271, 1992. ISBN 91-7870-876-1.
- P. Nagy:** Tools for knowledge-based signal processing with applications to system identification. Thesis No. 280, 1992. ISBN 91-7870-962-8.
- T. Svensson:** Mathematical tools and software for analysis and design of nonlinear control systems. Thesis No. 285, 1992. ISBN 91-7870-989-X.
- S. Andersson:** On dimension reduction in sensor array signal processing. Thesis No. 290, 1992. ISBN 91-7871-015-4.
- H. Hjalmarsson:** Aspects on incomplete modeling in system identification. Thesis No. 298, 1993. ISBN 91-7871-070-7.
- I. Klein:** Automatic synthesis of sequential control schemes. Thesis No. 305, 1993. ISBN 91-7871-090-1.
- J.-E. Strömbärg:** A mode switching modelling philosophy. Thesis No. 353, 1994. ISBN 91-7871-430-3.
- K. Wang Chen:** Transformation and symbolic calculations in filtering and control. Thesis No. 361, 1994. ISBN 91-7871-467-2.
- T. McKelvey:** Identification of state-space models from time and frequency data. Thesis No. 380, 1995. ISBN 91-7871-531-8.
- J. Sjöberg:** Non-linear system identification with neural networks. Thesis No. 381, 1995. ISBN 91-7871-534-2.
- R. Germundsson:** Symbolic systems – theory, computation and applications. Thesis No. 389, 1995. ISBN 91-7871-578-4.

- P. Pucar:** Modeling and segmentation using multiple models. Thesis No. 405, 1995. ISBN 91-7871-627-6.
- H. Fortell:** Algebraic approaches to normal forms and zero dynamics. Thesis No. 407, 1995. ISBN 91-7871-629-2.
- A. Helmersson:** Methods for robust gain scheduling. Thesis No. 406, 1995. ISBN 91-7871-628-4.
- P. Lindskog:** Methods, algorithms and tools for system identification based on prior knowledge. Thesis No. 436, 1996. ISBN 91-7871-424-8.
- J. Gunnarsson:** Symbolic methods and tools for discrete event dynamic systems. Thesis No. 477, 1997. ISBN 91-7871-917-8.
- M. Jirstrand:** Constructive methods for inequality constraints in control. Thesis No. 527, 1998. ISBN 91-7219-187-2.
- U. Forssell:** Closed-loop identification: Methods, theory, and applications. Thesis No. 566, 1999. ISBN 91-7219-432-4.
- A. Stenman:** Model on demand: Algorithms, analysis and applications. Thesis No. 571, 1999. ISBN 91-7219-450-2.
- N. Bergman:** Recursive Bayesian estimation: Navigation and tracking applications. Thesis No. 579, 1999. ISBN 91-7219-473-1.
- K. Edström:** Switched bond graphs: Simulation and analysis. Thesis No. 586, 1999. ISBN 91-7219-493-6.
- M. Larsson:** Behavioral and structural model based approaches to discrete diagnosis. Thesis No. 608, 1999. ISBN 91-7219-615-5.
- F. Gunnarsson:** Power control in cellular radio systems: Analysis, design and estimation. Thesis No. 623, 2000. ISBN 91-7219-689-0.
- V. Einarsson:** Model checking methods for mode switching systems. Thesis No. 652, 2000. ISBN 91-7219-836-2.
- M. Norrlöf:** Iterative learning control: Analysis, design, and experiments. Thesis No. 653, 2000. ISBN 91-7219-837-0.
- F. Tjärnström:** Variance expressions and model reduction in system identification. Thesis No. 730, 2002. ISBN 91-7373-253-2.
- J. Löfberg:** Minimax approaches to robust model predictive control. Thesis No. 812, 2003. ISBN 91-7373-622-8.
- J. Roll:** Local and piecewise affine approaches to system identification. Thesis No. 802, 2003. ISBN 91-7373-608-2.
- J. Elbornsson:** Analysis, estimation and compensation of mismatch effects in A/D converters. Thesis No. 811, 2003. ISBN 91-7373-621-X.
- O. Härkegård:** Backstepping and control allocation with applications to flight control. Thesis No. 820, 2003. ISBN 91-7373-647-3.
- R. Wallin:** Optimization algorithms for system analysis and identification. Thesis No. 919, 2004. ISBN 91-85297-19-4.
- D. Lindgren:** Projection methods for classification and identification. Thesis No. 915, 2005. ISBN 91-85297-06-2.
- R. Karlsson:** Particle Filtering for Positioning and Tracking Applications. Thesis No. 924, 2005. ISBN 91-85297-34-8.
- J. Jansson:** Collision Avoidance Theory with Applications to Automotive Collision Mitigation. Thesis No. 950, 2005. ISBN 91-85299-45-6.
- E. Geijer Lundin:** Uplink Load in CDMA Cellular Radio Systems. Thesis No. 977, 2005. ISBN 91-85457-49-3.
- M. Enqvist:** Linear Models of Nonlinear Systems. Thesis No. 985, 2005. ISBN 91-85457-64-7.

- T. B. Schön:** Estimation of Nonlinear Dynamic Systems — Theory and Applications. Thesis No. 998, 2006. ISBN 91-85497-03-7.
- I. Lind:** Regressor and Structure Selection — Uses of ANOVA in System Identification. Thesis No. 1012, 2006. ISBN 91-85523-98-4.
- J. Gillberg:** Frequency Domain Identification of Continuous-Time Systems Reconstruction and Robustness. Thesis No. 1031, 2006. ISBN 91-85523-34-8.
- M. Gerdin:** Identification and Estimation for Models Described by Differential-Algebraic Equations. Thesis No. 1046, 2006. ISBN 91-85643-87-4.
- C. Grönwall:** Ground Object Recognition using Laser Radar Data – Geometric Fitting, Performance Analysis, and Applications. Thesis No. 1055, 2006. ISBN 91-85643-53-X.
- A. Eidehall:** Tracking and threat assessment for automotive collision avoidance. Thesis No. 1066, 2007. ISBN 91-85643-10-6.
- F. Eng:** Non-Uniform Sampling in Statistical Signal Processing. Thesis No. 1082, 2007. ISBN 978-91-85715-49-7.
- E. Wernholt:** Multivariable Frequency-Domain Identification of Industrial Robots. Thesis No. 1138, 2007. ISBN 978-91-85895-72-4.
- D. Axehill:** Integer Quadratic Programming for Control and Communication. Thesis No. 1158, 2008. ISBN 978-91-85523-03-0.
- G. Hendeby:** Performance and Implementation Aspects of Nonlinear Filtering. Thesis No. 1161, 2008. ISBN 978-91-7393-979-9.
- J. Sjöberg:** Optimal Control and Model Reduction of Nonlinear DAE Models. Thesis No. 1166, 2008. ISBN 978-91-7393-964-5.
- D. Törnqvist:** Estimation and Detection with Applications to Navigation. Thesis No. 1216, 2008. ISBN 978-91-7393-785-6.
- P-J. Nordlund:** Efficient Estimation and Detection Methods for Airborne Applications. Thesis No. 1231, 2008. ISBN 978-91-7393-720-7.
- H. Tidefelt:** Differential-algebraic equations and matrix-valued singular perturbation. Thesis No. 1292, 2009. ISBN 978-91-7393-479-4.
- H. Ohlsson:** Regularization for Sparseness and Smoothness — Applications in System Identification and Signal Processing. Thesis No. 1351, 2010. ISBN 978-91-7393-287-5.
- S. Moberg:** Modeling and Control of Flexible Manipulators. Thesis No. 1349, 2010. ISBN 978-91-7393-289-9.
- J. Wallén:** Estimation-based iterative learning control. Thesis No. 1358, 2011. ISBN 978-91-7393-255-4.
- J. D. Hol:** Sensor Fusion and Calibration of Inertial Sensors, Vision, Ultra-Wideband and GPS. Thesis No. 1368, 2011. ISBN 978-91-7393-197-7.
- D. Ankelhed:** On the Design of Low Order H-infinity Controllers. Thesis No. 1371, 2011. ISBN 978-91-7393-157-1.
- C. Lundquist:** Sensor Fusion for Automotive Applications. Thesis No. 1409, 2011. ISBN 978-91-7393-023-9.
- P. Skoglar:** Tracking and Planning for Surveillance Applications. Thesis No. 1432, 2012. ISBN 978-91-7519-941-2.
- K. Granström:** Extended target tracking using PHD filters. Thesis No. 1476, 2012. ISBN 978-91-7519-796-8.
- C. Lyzell:** Structural Reformulations in System Identification. Thesis No. 1475, 2012. ISBN 978-91-7519-800-2.
- J. Callmer:** Autonomous Localization in Unknown Environments. Thesis No. 1520, 2013. ISBN 978-91-7519-620-6.
- D. Petersson:** A Nonlinear Optimization Approach to H2-Optimal Modeling and Control. Thesis No. 1528, 2013. ISBN 978-91-7519-567-4.

- Z. Sjanic:** Navigation and Mapping for Aerial Vehicles Based on Inertial and Imaging Sensors. Thesis No. 1533, 2013. ISBN 978-91-7519-553-7.
- F. Lindsten:** Particle Filters and Markov Chains for Learning of Dynamical Systems. Thesis No. 1530, 2013. ISBN 978-91-7519-559-9.
- P. Axelsson:** Sensor Fusion and Control Applied to Industrial Manipulators. Thesis No. 1585, 2014. ISBN 978-91-7519-368-7.
- A. Carvalho Bittencourt:** Modeling and Diagnosis of Friction and Wear in Industrial Robots. Thesis No. 1617, 2014. ISBN 978-91-7519-251-2.
- M. Skoglund:** Inertial Navigation and Mapping for Autonomous Vehicles. Thesis No. 1623, 2014. ISBN 978-91-7519-233-8.
- S. Khoshfetrat Pakazad:** Divide and Conquer: Distributed Optimization and Robustness Analysis. Thesis No. 1676, 2015. ISBN 978-91-7519-050-1.
- T. Ardestiri:** Analytical Approximations for Bayesian Inference. Thesis No. 1710, 2015. ISBN 978-91-7685-930-8.
- N. Wahlström:** Modeling of Magnetic Fields and Extended Objects for Localization Applications. Thesis No. 1723, 2015. ISBN 978-91-7685-903-2.
- J. Dahlin:** Accelerating Monte Carlo methods for Bayesian inference in dynamical models. Thesis No. 1754, 2016. ISBN 978-91-7685-797-7.
- M. Kok:** Probabilistic modeling for sensor fusion with inertial measurements. Thesis No. 1814, 2016. ISBN 978-91-7685-621-5.
- J. Linder:** Indirect System Identification for Unknown Input Problems: With Applications to Ships. Thesis No. 1829, 2017. ISBN 978-91-7685-588-1.
- M. Roth:** Advanced Kalman Filtering Approaches to Bayesian State Estimation. Thesis No. 1832, 2017. ISBN 978-91-7685-578-2.
- I. Nielsen:** Structure-Exploiting Numerical Algorithms for Optimal Control. Thesis No. 1848, 2017. ISBN 978-91-7685-528-7.
- D. Simon:** Fighter Aircraft Maneuver Limiting Using MPC: Theory and Application. Thesis No. 1881, 2017. ISBN 978-91-7685-450-1.
- C. Veibäck:** Tracking the Wanders of Nature. Thesis No. 1958, 2018. ISBN 978-91-7685-200-2.
- C. Andersson Naesseth:** Machine learning using approximate inference: Variational and sequential Monte Carlo methods. Thesis No. 1969, 2018. ISBN 978-91-7685-161-6.
- Y. Jung:** Inverse system identification with applications in predistortion. Thesis No. 1966, 2018. ISBN 978-91-7685-171-5.
- Y. Zhao:** Gaussian Processes for Positioning Using Radio Signal Strength Measurements. Thesis No. 1968, 2019. ISBN 978-91-7685-162-3.
- R. Larsson:** Flight Test System Identification. Thesis No. 1990, 2019. ISBN 978-91-7685-070-1.
- P. Kasebzadeh:** Learning Human Gait. Thesis No. 2012, 2019. ISBN 978-91-7519-014-3.
- K. Radnosrati:** Time of flight estimation for radio network positioning. Thesis No. 2054, 2020. ISBN 978-91-7929-884-5.
- O. Ljungqvist:** Motion planning and feedback control techniques with applications to long tractor-trailer vehicles. Thesis No. 2070, 2020. ISBN 978-91-7929-858-6.
- G. Lindmark:** Controllability of Complex Networks at Minimum Cost. Thesis No. 2074, 2020. ISBN 978-91-7929-847-0.
- K. Bergman:** Exploiting Direct Optimal Control for Motion Planning in Unstructured Environments. Thesis No. 2133, 2021. ISBN 978-91-7929-677-3.
- P. Boström-Rost:** Sensor Management for Target Tracking Applications. Thesis No. 2137, 2021. ISBN 978-91-7929-672-8.