Polar Opinion Dynamics in Social Networks

Victor Amelkin†, Francesco Bullo‡, and Ambuj K. Singh†

***Abstract*—For decades, scientists have studied opinion formation in social networks, where information travels via word of mouth. The particularly interesting case is when polar opinions—Democrats vs. Republicans or iOS vs. Android—compete in the network. The central problem is to design and analyze a model that captures how polar opinions evolve in the real world.**

In this work, we propose a *general non-linear model of polar opinion dynamics*, rooted in several theories of sociology and social psychology. The model's key distinguishing trait is that, unlike in the existing linear models, such as DeGroot and Friedkin-Johnsen models, *the individuals' susceptibility to persuasion is a function of their current opinions*. For example, a person holding a neutral opinion may be rather malleable, while "extremists" may be strongly committed to their current beliefs. We also study three specializations of our general model, whose susceptibility functions correspond to different socio-psychological theories.

We provide a comprehensive theoretical analysis of our non-linear models' behavior using several tools from non-smooth analysis of dynamical systems. To study convergence, we use non-smooth max-min Lyapunov functions together with the generalized Invariance Principle. For our general model, we derive a general sufficient condition for the convergence to consensus. For the specialized models, we provide a full theoretical analysis of their convergence—whether to consensus or disagreement. Our results are rather general and easily apply to the analysis of other non-linear models defined over directed networks, with Lyapunov functions constructed out of convex components.

I. Introduction

The central goal of this work is modeling the evolution of opinions of a group of people—the agents—connected in a directed social network. We assume that the objective means for opinion evaluation are limited, and the agents evaluate their opinions by comparison with the opinions of others [25]. Thus, the process of opinion formation in a group is a *network process*, where each agent's opinion changes due to the agent's interaction with his or her neighbors in the network.

In particular, we focus on *polar opinions*, which describe either degrees of proclivity toward one of two competing alternatives (e.g., Democrats vs. Republicans or iOS vs. Android) or an attitude—from extreme unfavorable to neutral to extreme favorable—toward a single issue (e.g., using nuclear power as an energy source). We will use the terms *opinion* and *attitude* interchangeably, and refer to them both as *an agent's state*. Our emphasis on polar opinions will manifest itself in that the agents in our non-linear models will change their opinion-adoption behavior as their opinions shift toward one or another pole of the opinion spectrum.

This work was supported by the U. S. Army Research Laboratory and the U. S. Army Research Office under grant number W911NF-15-1-0577.

†Victor Amelkin (victor@cs.ucsb.edu) and Ambuj K. Singh (ambuj@cs.ucsb.edu) are with the Department of Computer Science, University of California, Santa Barbara.

‡Francesco Bullo (bullo@engineering.ucsb.edu) is with the Department of Mechanical Engineering, University of California, Santa Barbara.

In what follows, we will review the main components of an opinion dynamics model and connect them with the existing theories of sociology and social psychology, preparing a foundation upon which our models will be constructed.

Opinion formation via weighted averaging: The most basic network model of opinion dynamics is the weighted averaging model of DeGroot [21]:

$$x(t+1) = Wx(t),$$

where t is time, $x(t)$ is a vector of agent states, and W is the row-stochastic adjacency matrix of the social network, with W_{ij} indicating the relative extent to which agent j influences the opinion of agent i, or, alternatively, the relative share of $x_j(t)$ in $x_i(t+1)$. According to this model, each agent forms his or her opinion as a weighted average of all the opinions available in the agent's out-neighborhood in the network.

The appeal of DeGroot model stems from its consistency with such theories of social psychology as social comparison theory [25], cognitive dissonance theory [26], and balance theory [13], [37], whose unifying idea is that the agents act to achieve balance with other group members or, alternatively, to relieve psychological discomfort from their disagreement with others. However, one limitation of DeGroot model is that the agents' "behavior" does not change depending on the agent, its current state—opinion or attitude—and the issue at hand.

Models with susceptibility to persuasion: At the very least, the strength of an agent's attachment to his or her opinion depends on the extent to which the issue is important to that agent and is representative of his or her values. Such a dimension of the strength of attitude—a function of the agent and the issue—has arisen in multiple studies under the names of embeddedness [61], ego preoccupation and ego involvement [2], [46], among others. Friedkin-Johnsen model [29] addresses the limitation of DeGroot model by allowing the agents to have different *susceptibilities to persuasion*:

$$x(t+1) = AWx(t) + (I-A)x(0),$$

where t, $x(t)$ and W are defined as before, I is the identity matrix, and A is a constant diagonal matrix whose diagonal element A_{ii} describes the extent to which agent i's opinion is affected by the opinions of other agents as opposed to his or her own initial opinion. The diagonal elements of matrix $(I-A)$ are usually referred to as the agents' degrees of *stubbornness*. Friedkin-Johnsen model improves upon DeGroot model not only in terms of the model's interpretation, but also in terms of the model's behavior—while the typical asymptotic behavior of DeGroot model in a "well-connected" social network is the convergence to consensus, in case of Friedkin-Johnsen model, agents usually disagree. The latter behavior usually occurs in the real world [45].

State-dependent susceptibility in linear models: The "state-oblivious" definition of the agents' constant susceptibilities to persuasion of Friedkin-Johnsen model does not capture another component of the agents' strength of conviction, known in social psychology as commitment [61], [2], [49] or certainty [46], which is determined by each agent's attitude toward the issue. The dependency of susceptibility to persuasion on the agents' beliefs has been studied in the context of Friedkin-Johnsen model [28], where the asymptotic behavior of the model was empirically evaluated under several definitions of susceptibility $A(x(0))$ as a function of the initial opinions $x(0)$ of the agents.

Among the existing theories of social psychology, there is no agreement upon a single correct definition of susceptibility A based on the agents' beliefs. One factor that has arisen in multiple studies as closely related to the strength of conviction is the attitude extremity or polarity [47], [7], [61], [59], [54], [22]. The conclusion that can be drawn from these works is that extreme opinions are more resistant to change, possibly, due to the preferential evaluation of attitude-congruent information by the agents holding extreme opinions. Alternatively, social comparison theory [25] suggests that, when the majority of the agents hold a certain, say, neutral, opinion, establishing a social norm, then the agents with opinions close to that norm have relatively weaker tendencies to change their positions, while the extreme opinions are unstable.

Existence of multiple alternative theories regarding the factors determining the strength of the agents' commitment to their opinions is not surprising, particularly, because these factors have been shown to be domain-specific [61]. Hence, it is rational to either study the opinion formation process under the most general definition of the strength of the agents' attitudes or to use multiple definitions of the attitude strength based on the existing socio-psychological theories. In this work, we will study both the most general definition of agent susceptibility as well as several specialized definitions consistent with different socio-psychological theories.

State-dependent susceptibility in non-linear models: The definition of susceptibility A as a function of the *initial state* $x(0)$ is beneficial in that it does not take away the model's linearity and, hence, allows application of the existing linear-algebraic techniques to the formal analysis of the model's asymptotic behavior. However, definition of A as a function of the *current state* $x(t)$, while would make the model non-linear, has at least two advantages. For one thing, the definition $A = A(x(t))$ may be more appropriate when the evolution of opinions is studied at a large time scale, as in the case when a group of people is working on a year-long project, and hardly anyone remembers what their opinions were a year ago. For another thing, and more importantly, in several existing studies [8], [63], [75], the agents' attitudes are posited to be "constructed on demand", and, in particular, according to the potentiated recruitment model [8], the strength of attitude is an emergent property of the process of attitude construction occurring when the attitude is recruited. Thus, in our models, we adopt the definitions of agent susceptibility dependent on the agents' current states.

This work's summary and contributions: In our work, we propose novel non-linear models of polar opinion dynamics and formally analyze their behavior. More specifically, we make the following contributions.

(i) Novel Models: We propose a general non-linear model of polar opinion dynamics, where the agents' susceptibilities to persuasion are general functions of the agents' current beliefs. Additionally, we propose three specialized instances of the general model, having different definitions of agent susceptibility $A(x(t))$ corresponding to different theories of social psychology. The proposed models are novel in that they capture more traits of the opinion formation process than the existing models, and manifest a behavior unobserved in their linear counterparts—we can generally observe either the agents' convergence to consensus or their persistent disagreement, depending on the agents' initial beliefs $x(0)$.

(ii) Analysis of the General Model: For our general model of polar opinion dynamics we prove a contraction property of its trajectories, and provide a sufficient condition for the convergence to consensus. That sufficient condition is rather general and, roughly, states that convergence to consensus takes place if the agents non-responsive to persuasion have identical states. The latter entails that, quite naturally, a disagreement among the agents may arise only if there are multiple agents having different beliefs and unwilling to change them.

(iii) Analysis of the Specialized Models: We provide a comprehensive theoretical analysis of the asymptotic behavior of our specialized models, characterizing all their states of equilibrium—corresponding to either the states of consensus or disagreement—through the analysis of certain partitions of the network, and prove each model's convergence. For the cases when a model converges to a state of disagreement, we provide an explicit expression for that limiting state, which depends on the network's structure as well as the beliefs and locations of the agents non-responsive to persuasion, yet, does not depend on the initial beliefs of the susceptible agents.

(iv) Novel Analysis of Convergence: The standard tools for the analysis of convergence of non-linear continuous-time models, such as Lyapunov's Second Method and LaSalle Invariance Principle, require existence of a smooth Lyapunov function, which may not and, sometimes, provenly does not exist [55] for a model defined over a directed network. In this work, we use max-min non-smooth Lyapunov functions along with several existing non-smooth analysis techniques to prove convergence of our non-linear models. While such Lyapunov functions have appeared in existing literature, to the best of our knowledge, we are the first to provide a full formal analysis of convergence of continuous-time non-linear systems defined over directed networks using such functions together with the generalized Invariance Principle.

Paper's organization: Having motivated our work in Section I, we proceed with a review of existing opinion dynamics models in Section II. Then, in Section III, we define our models of polar opinion dynamics, and, subsequently, analyze them in Section IV. The analysis of our models relies on some notions from non-smooth analysis reviewed in the Appendix. We conclude with a discussion of our results in Section V.

II. General literature review

The numerous existing opinion dynamics models can be roughly divided into two groups: *analytic* and *algorithmic*.

Analytic models are represented as systems of difference or differential equations

$$x(t+1) = W(x(t),t)x(t) \quad \text{or} \quad dx/dt = \dot{x} = W(x(t),t)x$$

and describe a process of agent interaction usually targeting a certain form of agreement among the agents. These models mainly differ in the extra properties of the agent interaction process besides the agents' effort toward reaching agreement. All our models proposed in this paper belong to this group.

Analytic models have long been studied by sociologists, starting with the works of French [27] and Harary [35]. Nowadays, the basic formulation of the weighted averaging process is usually referred to as DeGroot model [21]. A variation of DeGroot model with some agents' states kept constant has been studied by Pirani and Sundaram [60]. An improvement upon DeGroot model was proposed by Friedkin and Johnsen [29], who enabled the agents to have individual levels of susceptibility to persuasion by other agents. Variations of Friedkin-Johnsen model have been studied in a context of a group's discussing a sequence of issues [41], [72]. The question of the dependence of A upon the agents' beliefs has been empirically studied by Friedkin in [28]. A variation of Friedkin-Johnsen model with time-dependent $A(t)$ and its connection to the underlying notion of dissonance minimization was discussed in work [34] by Groeber et al. A Friedkin-Johnsen-type model with stubborn agents has been studied as a local interaction game by Ghaderi et al. [30].

Another type of analytic models—close in spirit to our models—are the models with state-dependent agent interaction, $W(x)$, and, in particular, the bounded confidence models [14], [52], whose key idea is that only the agents with close enough states can interact. The two popular representatives of such models are Hegselmann-Krause (HK) [36] model and the model of Deffuant et al. [20]. Some convergence results for HK model have been proven by Blondel et al. [9]. Sufficient convergence conditions for a more general model with state-dependent agent interaction, that includes HK model as a special case, have been studied by Lorenz [51].

A special subtype of bounded confidence models are those that allow for stubbornness, leadership, antagonism, or zealotry, and whose behavior has been investigated through simulation. In particular, in [19], Deffuant et al. study the behavior of Deffuant's model with smooth confidence bounds in the presence of stubborn agents. Kurmyshev et al. [48] extend Deffuant's model with two types of agents characterized by "friendly" and "antagonistic" interaction, respectively. Jalili [40] has studied the effect of the choice of the subset of stubborn leaders as well as a particular network structure upon the the bounded confidence model's convergence rate. Sobkowicz [64] has considered the "Deffuant model with emotions", where different opinions have varying resilience to change. In particular, the author considered the cases when the extreme opinions are more resilient to change than the neutral opinions, as well as the case of an asymmetric dependency of the opinion resilience on the opinion value. These opinion resilience mechanisms are similar to some of those we use in our specialized models in Section III. Chen et al. [15] investigated how stubborn leaders can attract followers in the context of a bounded confidence model that incorporates the leaders' reputation, stubbornness, appeal, as well as the extremity of their opinions. Finally, Tucci et al. [67] have studied the bounded confidence model with stubborn leaders and investigated the effect of the number of leaders on the opinion dynamics profile.

Similar analytic models are studied in the control systems and robotics communities, in the context of multi-agent coordination problems [11], [57]. The models with time-varying topology $W(t)$ of the network have been studied in [53], [56], [55], [58]; the models for signed networks, allowing for agents' friendly and antagonistic interaction, have been studied in [5], [39]; the models with randomized agent interaction have been studied in [24], [65].

The final class of analytic models are the models considered in the context of the naming game. Specifically, Waagen et al. [69] design a naming game model with discrete opinions and zealots—who do not change their opinion—and study the effect of the number of zealots on the opinion dynamics of the entire population. Verma et al. [68] has also studies the effect of the presence of zealots in the naming game.

Algorithmic models for opinion dynamics are usually defined as combinatorial algorithms—probabilistic or deterministic—describing how the agents update their states. These models usually operate with discrete agent states and in discrete time. A notable difference of algorithmic models from their analytic counterparts is that algorithmic models are usually data-driven, that is, such models are usually to be fit to data, whereas the analytic models are "prescriptive". One model in this group is the Independent Cascade Model [31], where the agents get "activated" with an opinion by their neighbors in a probabilistic fashion. The basic version of this model uses binary opinions—indicating presence or absence of an opinion—and is usually used in the context of the influence maximization problem [42]. A version of the Independent Cascade Model for the case of multiple competing opinions has been proposed in [12]; a version with asynchronous communications has been studied in [62]. A related, yet more general model, allowing for competing opinions, is the switching-selection model of [32].

Two other types of algorithmic models are the Voter model [17], [23], [44], [73] and the Linear Threshold model [33], [70], where in the former model, each agent is activated in a probabilistic fashion based on the number of active agents in the neighborhood, and in the latter model, agents become active as soon as the number of active neighbors surpasses a constant threshold. Versions of the Linear Threshold model for the case of competing opinions have been studied in [10]. The extensions of the discrete-opinion Voter model with stubborn agents have been considered in works [3], [74], [73], where the authors studied the models' long-run behavior as well as the problem of influence maximization.

Finally, there are Bayesian algorithmic models [4], whose agent state update rules are based on the Bayes rule.

III. Model of polar opinion dynamics

General model: We define the *general model of polar opinion dynamics* as follows:

$$\boxed{\dot{x} = -A(x)Lx.} \tag{1}$$

where $x(t) \in [-1,1]^n$ represents the agents' states, $A(x(t)) \in \operatorname{diag}([0,1]^n)$ is a diagonal matrix whose diagonal elements are the agents' state-dependent and possibly different susceptibility functions locally Lipschitz in $[-1,1]^n$, $L = I - W$ is the network's Laplacian matrix, and $W \in [0,1]^{n\times n}$ is the row-stochastic adjacency matrix of the directed network, with its edge weight W_{ij} measuring the amount of relative influence of agent j upon agent i.

The mathematical interpretation of the above defined model is as follows. In (1), the negative Laplacian $-L$, when applied to $x(t)$, measures how much, on (weighted) average, the agent's state is smaller than the states of the agents in its out-neighborhood $N_{out}(i) = \{j \mid j \neq i \wedge W_{ij} > 0\}$

$$(-Lx)_i = \sum_{j\in N_{out}(i)} w_{ij}(x_j - x_i).$$

When $(-Lx)_i > 0$ and agent i is *open* to persuasion, that is, $A_{ii}(x) > 0$, then $\dot{x}_i > 0$, and x_i grows, "following" its out-neighbors. Conversely, if $(-Lx)_i < 0$, the state of an open agent i decreases. If either an agent's state is in balance with the states of its out-neighbors, or the agent is *closed* to persuasion, that is, $A_{ii}(x) = 0$, then this agent's state does not change.

Model (1) can also be thought of as a non-linear generalization of the heat diffusion model $\dot{x} = \alpha\Delta x$, where the negative Laplacian $-L$ of (1) corresponds to the finite-difference approximation of the continuous-space Laplace operator Δ, and the rate $A(x)$ at which the state of the model evolves may be thought of as the temperature-dependent thermal diffusivity—a naturally occurring phenomenon [71].

The general model for polar opinion dynamics $\dot{x} = -A(x)Lx$ consists of two conceptual components: the averaging component $-Lx$ drives the agents towards agreement, while the susceptibility component $A(x)$ impedes this convergence process. The averaging component is based on such theories of social psychology as social comparison theory [25], cognitive dissonance theory [26], and balance theory [13], [37], whose unifying idea is that the agents act to achieve balance with other group members. The general idea of agents' susceptibility or stubbornness to persuasion comes from the socio-psychological studies of the strength of attachment to one's opinion [61], [2], [46]. The dependency of the agents' susceptibility $A(x)$ on their current beliefs agrees with the socio-psychological studies [8], [63], [75], that posit that the agents' attitudes are "constructed on demand".

Specialized models: In addition to the general model (1), we will consider three specialized models, each with a different definition of state-dependent susceptibility $A(x(t))$, having different socio-psychological interpretations.

(i) Model with stubborn extremists: The first specialized model draws from the socio-psychological studies [47], [7], [61], [59], [54], [22] of the attitude extremity as being a major factor defining the strength of conviction, and whose definition of agent susceptibility $A(x) = (I - \operatorname{diag}(x)^2)$ assumes that extreme opinions are more resistant to change than neutral opinions.

$$\boxed{\dot{x} = -(I - \operatorname{diag}(x)^2)Lx.} \tag{2}$$

This model is appropriate when the extreme opinions *compete* in that an agent's strong preference of one extreme implies this agent's likely rejection of the opposite extreme. For example, this may be the case when agents' states describe the degrees of support for one of the two major political parties in the US—inveterate Republicans or Democrats are unlikely to change their political affiliation, while neutral voters can be successfully attracted toward one or another pole of the opinion spectrum.

(ii) Model with stubborn positives: The second specialized model is a variation of the model with stubborn extremists with the asymmetric susceptibility function $A(x) = (I - \operatorname{diag}(x))/2$, where the agents only at one end of the opinion spectrum are stubborn.

$$\boxed{\dot{x} = -\tfrac{1}{2}(I - \operatorname{diag}(x))Lx.} \tag{3}$$

This definition—inspired by the "Stubborn Left" and "Stubborn Right" susceptibility functions considered by Friedkin in [28]—fits those cases when the agents at one, say, negative extreme of the opinion spectrum have no reason to reject the alternative opinion, while the agents having the opposite, positive, opinion have an incentive to maintain their position. For example, the opinion may describe the degree of liking one of two smartphone brands, where opinion -1 corresponds to the neutrally marketed brand, while opinion $+1$ corresponds to the brand that is aggressively marketed not just as the best, but also as the only viable option.

(iii) Model with stubborn neutrals: Finally, in our third specialized model, drawing from the social comparison theory [25] and the studies of social norms [28], we defined agent susceptibility as A(x) = $\operatorname{diag}(x)^2$, assuming that the neutral opinions are resilient to change, while the extreme opinions are unstable, thereby, making this model the opposite of the model with stubborn extremists.

$$\boxed{\dot{x} = -\operatorname{diag}(x)^2 Lx.} \tag{4}$$

This model assumes that the neutral opinion 0 correspond to a social norm, and the agents may not feel comfortable deviating from it and going against what is acceptable in their society.

Specialized models' justification: Stating our specialized models, we have provided three particular definitions of the agents' susceptibility $A(x)$. While these definitions agree with several well-established socio-psychological theories, the latter theories do not provide any specifics about the particular mathematical form of $A(x)$, besides giving a general idea of its behavior. In our definitions, we use low-degree polynomials, making sure $A(x)$ fits the socio-psychological theories and, at the same time, is simple enough to allow a clear analysis. Similarly, quadratic polynomials were used by Taylor [66] who extended Abelson's linear models [1] with "variable resistance". Alternatively, Friedkin in his recent study [28], when defining the constant susceptibility $A(x(0))$ of the agents based on their initial beliefs, used functions of similar "shape", yet, expressed them using exponential functions.

Nevertheless, despite this lack of a single correct mathematical form for each version of $A(x)$, the convergence results we obtain in the next Section IV are derived independently of a particular mathematical form of $A(x)$, only relying on the zeros of $A(x)$ as well as our ability to analytically compute them. Thus, while we will further provide our analysis for the specialized models using the particular susceptibility functions defined above, this analysis is easy to adapt to other susceptibility functions behaving similarly, yet, possibly, having different mathematical form.

IV. Analysis of the models

In this section, we will analyze well-posedness, equilibrium points, and the asymptotic behavior of our models. The convergence proofs will rely on several notions from non-smooth analysis reviewed in the Appendix.

A. Well-posedness

In order for our general model $\dot{x} = -A(x)Lx$ to be well-posed, its solutions must exist, be unique, and must never escape the state space $[-1,1]^n$. These properties of the solutions are stated in the following theorem and its corollary.

Theorem 1 (Well-posedness of the general model)**.** *If $x(0) \in [-1,1]^n$, and the evolution of $x(t)$ is governed by the general model of polar opinion dynamics $\dot{x} = -A(x)Lx$, then $x(t) \in [-1,1]^n$ for any $t \geq 0$.*

Corollary (Existence, uniqueness, smoothness of solutions)**.** *The general model of polar opinion dynamics $\dot{x} = -A(x)Lx$, $x(0) \in [-1,1]^n$, has a unique continuously-differentiable solution $x(t)$ defined for all $t > 0$.*

Since both Theorem 1 and its Corollary are standard results in control theory, we will omit their proofs and only mention that the validity of Theorem 1 immediately follows from the general contraction Lemma 3, while the validity of the Corollary follows from Theorem 3.3 of [43] used together with local Lipschitz-continuity of $A(x)$ and Theorem 1.

B. Equilibrium points

Prior to studying the equilibrium points of our models, we will prove a basic lemma.

Lemma 1 (Properties of some network partitions)**.** *Let $W \in \mathbb{R}^{n\times n}$ be a row-stochastic adjacency matrix of a strongly connected network $G(W)$. If $G(W)$'s nodes are partitioned into two non-empty sets $\{1,\dots,n\} = I_1 \cup I_2$, $I_1 \cap I_2 = \varnothing$, and P is any permutation matrix such that nodes I_1 precede nodes I_2 in*

$$PWP^\intercal = \begin{array}{c} \\ I_1 \\ I_2 \end{array} \begin{array}{c} \begin{array}{cc} I_1 & I_2 \end{array} \\ \begin{bmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{bmatrix} \end{array},$$

then both $(I - W_{11})$ and $(I - W_{22})$ are invertible, and both $(I - W_{11})^{-1}W_{12}\mathbb{1} = \mathbb{1}$ and $(I - W_{22})^{-1}W_{21}\mathbb{1} = \mathbb{1}$.

Proof. Since $G(W)$ is strongly connected, W is irreducible. Consequently, since both I_1 and I_2 and non-empty, W_{11} is substochastic, so there exists $\ell \in I_1$, such that

$$\sum_{j\in I_1} (W_{11})_{\ell j} < 1. \qquad (*)$$

Notice that $\sum_{j\in I_1} (W_{11})^k_{ij}$ can be interpreted as the likelihood of a k-hop random walk on $G(W)$ to start at node i and end at any node of I_1, so $(*)$ implies that there is a positive likelihood for a 1-hop random walk starting at ℓ to escape I_1. If we define $d(i,j)$ to be the length—in hops—of the shortest path from node i to node j in $G(W)$, and $d_{max}(\ell) = \max_{i\in I_1} d(i,\ell)$, then

$$\forall k > d_{max}(\ell)\ \forall i \in I_1 : \sum_{j\in I_1} (W_{11})^k_{ij} < 1,$$

since for each $i \in I_1$, there is at least one k-hop walk passing through ℓ, and, as a result, there is a positive likelihood of any such walk's escaping I_1. Hence, for all $k > d_{max}(\ell)$, W_{11}^k is convergent, and its spectral radius $\rho(W_{11}^k) < 1$, which immediately entails $\rho(W_{11}) < 1$. Hence, for the spectrum of $(I - W_{11})$, we have $\sigma(I - W_{11}) \subset (0,2)$. Thus, $(I - W_{11})$ is non-singular and, as such, invertible. Consequently, matrix $(I - W_{11})^{-1}W_{12}$ is well-defined, and its row-sums are

$$\begin{aligned}(I - W_{11})^{-1}W_{12}\mathbb{1} &= (\text{since } W \text{ is row-stochastic})\\ &= (I - W_{11})^{-1}(\mathbb{1} - W_{11}\mathbb{1}) =\\ &= (I - W_{11})^{-1}(I - W_{11})\mathbb{1} = \mathbb{1}.\end{aligned}$$

Applying the same reasoning to blocks W_{22} and W_{21} in place of blocks W_{11} and W_{12}, we obtain the existence of $(I - W_{22})^{-1}$, and equality $(I - W_{22})^{-1}W_{21}\mathbb{1} = \mathbb{1}$. □

Theorem 2 (Equilibrium points)**.** *Suppose the network's adjacency matrix W is row-stochastic, and network $G(W)$ is strongly connected. Then, the following holds.*

1) The equilibrium points of the stubborn positives model $\dot{x} = -\frac{1}{2}(I - \operatorname{diag}(x))Lx$ and the stubborn neutrals model $\dot{x} = -\operatorname{diag}(x)^2Lx$ are

$$x^* = \alpha\mathbb{1}_n,\ \alpha \in [-1,1].$$

2) Consider an arbitrary agent set partition $\{1,\dots,n\} = I_1 \cup I_2$, $I_1 \cap I_2 = \varnothing$, $2 \leq |I_1| \leq n$, and an arbitrary permutation matrix P such that the agents are ordered as

$$PWP^\intercal = \begin{array}{c} \\ I_1 \\ I_2 \end{array} \begin{array}{c} \begin{array}{cc} I_1 & I_2 \end{array} \\ \begin{bmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{bmatrix} \end{array}, \quad Px = \begin{array}{c} I_1 \\ I_2 \end{array} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

Then, the equilibrium points of the stubborn extremists model $\dot{x} = -(I - \operatorname{diag}(x)^2)Lx$ are

$$x^* = \alpha\mathbb{1}_n,\ \alpha \in [-1,1], \quad \textit{and} \quad x^* = P^\intercal[x_1^{*\intercal}, x_2^{*\intercal}]^\intercal,$$

where

$$\begin{aligned} x_1^* &\in \{-1,1\}^{|I_1|} \setminus \{-\mathbb{1}_{|I_1|}, \mathbb{1}_{|I_1|}\},\\ x_2^* &= \begin{cases} (I - W_{22})^{-1}W_{21}x_1^*, & \textit{if } I_2 \neq \varnothing,\\ [\ ]_{0\times 1}, & \textit{if } I_2 = \varnothing. \end{cases}\end{aligned}$$

Proof. 1) First, let us deal with the model with stubborn positives $\dot{x} = f(x) = -\frac{1}{2}(I - \operatorname{diag}(x))Lx$. It is easy to see that

$x^* = \alpha\mathbb{1},\ \alpha \in [-1,1]$ are equilibrium points of the system. Now, let us look for equilibrium points corresponding to the states of disagreement. Consider such a candidate point $x \in [-1,1]^n$, $x \neq \alpha\mathbb{1}$, $n > 1$. Since $x \neq \alpha\mathbb{1}$, there exists agent $i \in \{1,\dots,n\}$, such that $x_i = \min(x) < 1$, and exists agent $j \in \{1,\dots,n\}$ such that $x_j > x_i$ and $W_{ij} > 0$. Because $x_i < 1$, $A_{ii}(x) = \frac{1}{2}(1-x_i) > 0$; due to the existence of such agent j, $(Lx)_i \neq 0$. As a result, $(f(x))_i \neq 0$ and, thus, x is not a point of equilibrium. Hence, $x^* = \alpha\mathbb{1}$ are the only points of equilibrium of the model with stubborn positives.

In the remainder of the proof, we will consider different partitions $\{1,\dots,n\} = I_1 \cup I_2$, $I_1 \cap I_2 = \varnothing$ of the agent set, and P being any permutation matrix such that agents I_1 precede agents I_2 in PWP^T and Px. For readability, for each partition $I_1 \cup I_2$, we will omit P in the expressions for W and x, and will apply the right agent ordering later.

Let us proceed to the model with stubborn neutrals $\dot{x} = -\operatorname{diag}(x)^2 Lx$. A candidate equilibrium point x is defined w.r.t. an agent set partition $\{1,\dots,n\} = I_1 \cup I_2$, $I_1 \cap I_2 = \varnothing$, $0 \leq |I_1| \leq n$ as follows: $x_i = 0$ if $i \in I_1$; and $|x_i| > 0$ if $i \in I_2$. If $I_1 = \varnothing$, then $f(x) = \mathbb{0} \Leftrightarrow Lx = \mathbb{0} \Leftrightarrow x^* = \alpha\mathbb{1}$, where $\alpha \in [-1,1]$. If $I_2 = \varnothing$, then, clearly, $x^* = \mathbb{0}$. Finally, if both I_1 and I_2 are non-empty, then, w.l.o.g., assuming that agents I_1 precede agents I_2, $f(x) = \mathbb{0} \Leftrightarrow (I - W_{22})x_2 = \mathbb{0}$. However, from Lemma 1, we know that $(I - W_{22})$ is invertible. Hence, the obtained equation has only the trivial solution $x_2 = \mathbb{0}$, and the corresponding equilibrium point is $x^* = \mathbb{0}$. We have proven that the equilibrium points of the model with stubborn neutrals are $x^* = \alpha\mathbb{1},\ \alpha \in [-1,1]$.

2) For the model with stubborn extremists $\dot{x} = -(I - \operatorname{diag}(x)^2)Lx$, we define candidate equilibrium points w.r.t. partition $\{1,\dots,n\} = I_1 \cup I_2$, $I_1 \cap I_2 = \varnothing$, $0 \leq |I_1| \leq n$ as follows: $|x_i| = 1$ if $i \in I_1$; and $|x_i| < 1$ if $i \in I_2$. If $I_1 = \varnothing$, then $f(x) = \mathbb{0} \Leftrightarrow Lx = \mathbb{0}$, and, thus, $x^* = \alpha\mathbb{1}$, $\alpha \in [-1,1]$. If $I_2 = \varnothing$, then $x^* \in \{-1,1\}^n$. If both I_1 and I_2 are non-empty, then, again, assuming that agents I_1 precede agents I_2 in W and x, equation $f(x) = \mathbb{0}$ is rewritten as

$$
\begin{aligned}
&(I - \operatorname{diag}(x_2)^2)\begin{bmatrix} -W_{21} & (I - W_{22}) \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \mathbb{0} \\
&\Leftrightarrow (\text{since } |x_2| < \mathbb{1}) \Leftrightarrow (I - W_{22})x_2 = W_{21}x_1.
\end{aligned}
$$

From Lemma 1, we know that $(I - W_{22})$ is invertible. Thus, $x_2 = (I - W_{22})^{-1}W_{21}x_1$. Among the $x = [x_1^\mathsf{T}, x_2^\mathsf{T}]^\mathsf{T}$ satisfying the obtained equation, we would like to separate those corresponding to consensus and those corresponding to disagreement. If $x_1 = \mathbb{1}$, then, again from Lemma 1, we derive $x_2 = \mathbb{1}$, and $x^* = \mathbb{1}$. Similarly, if $x_1 = -\mathbb{1}$, then $x^* = -\mathbb{1}$. Notice, that all the equilibrium points discovered so far correspond to consensus and are independent of matrix P.

Finally, if $x_1 \in \{-1,1\}^{|I_1|} \setminus \{\mathbb{1}, -\mathbb{1}\}$ (which implies $|I_1| \geq 2$), and, hence, the agents necessarily disagree, then $x_2 = (I - W_{22})^{-1}W_{21}x_1$, and the corresponding equilibrium points, under the partition-defined agent order P, are $x^* = P^\mathsf{T}\begin{bmatrix} x_1^\mathsf{T} & x_2^\mathsf{T} \end{bmatrix}^\mathsf{T}$. □

C. Convergence analysis

Having studied the equilibrium points of our specialized models, we will now study these models' convergence. We will, first, establish sufficient conditions for convergence to consensus of the general model of polar opinion dynamics and, then, use this result to prove convergence of the three specialized models.

In the proofs of convergence, we will need establishing forward invariance of certain subsets of the state space with respect to a model at hand. To that end, we will need the following Lemma, being an immediate consequence of the solution uniqueness stated in the Corollary of Theorem 1.

Lemma 2 (Agent subset invariance)**.** *If $x(t) \in [-1,1]^n$ evolves according to one of the specialized models*

$$\dot{x} = -\tfrac{1}{2}(I - \operatorname{diag}(x))Lx, \tag{3}$$

$$\dot{x} = -\operatorname{diag}(x)^2 Lx, \tag{4}$$

$$\dot{x} = -(I - \operatorname{diag}(x)^2)Lx, \tag{2}$$

and the agents are partitioned into $I_{closed}(t) = \{i \mid A_{ii}(x(t)) = 0\}$ and $I_{open}(t) = \{i \mid A_{ii}(x(t)) > 0\}$, then, for all $t \geq 0$, $I_{closed}(t) = I_{closed}(0) = I_{closed}$ and $I_{open}(t) = I_{open}(0) = I_{open}$.

The following lemma will be instrumental in proving forward invariance of subsets of the state space as well as in the construction of Lyapunov functions in the convergence proofs.

Lemma 3 (General contraction lemma)**.** *Suppose that W is a row-stochastic adjacency matrix of the network, and agent states $x(t) \in [-1,1]^n$ evolve according to the general model of polar opinion dynamics*

$$\dot{x} = f(x) = -A(x)Lx. \tag{1}$$

Then, $V_{max}(x) = \max(x)$ is non-increasing and $V_{min}(x) = \min(x)$ is non-decreasing along the trajectories of (1).

Proof. Let us consider $V_{max}(x) = \max(x)$ and define $I_{max}(x) = \{i \mid x_i = \max(x)\}$. According to Lemma 2.2 of Lin et al. [50], the upper Dini derivative of V_{max} along the trajectories of (1) is defined as

$$
\begin{aligned}
D_f^+ V_{max}(x) &= \max_{i \in I_{max}(x(t))} \dot{x}_i(t) \\
&= \max_{i \in I_{max}(x(t))} \underbrace{A_{ii}(x(t))}_{\in[0,1]} \sum_{j \in N^{out}(i)} w_{ij} \underbrace{(x_j - x_i)}_{\leq 0} \leq 0.
\end{aligned}
$$

Hence, V_{max} is non-increasing along the trajectories of (1). The proof for V_{min} is similar and, hence, is omitted. □

We have laid out all the necessary preliminaries, and are ready to prove convergence of our models. In the following theorem, we will establish a sufficient condition for the convergence to consensus of the general model of polar opinion dynamics.

Theorem 3 (General convergence to consensus)**.** *Suppose that W is a row-stochastic adjacency matrix of a strongly connected network $G(W)$, and agent states $x(t) \in [-1,1]^n$ evolve according to the general model of polar opinion dynamics*

$$\dot{x} = f(x) = -A(x)Lx, \tag{1}$$

with the agents' having potentially different susceptibility functions $A_{ii}(x)$.

Let $S \subseteq [-1,1]^n$ be a non-empty compact set, forward invariant w.r.t. system (1), and

$$N = S \cap \{\alpha\mathbb{1} \mid \alpha \in [-1,1]\}$$

be its non-empty subset of consensus states.

Further, assume that in S, the agents' susceptibility functions $A_{ii}(x)$ agree upon their zeros in that

$$\forall x \in S\ \forall i, j \in \{1,\dots,n\} : A_{ii}(x) = A_{jj}(x) = 0 \to x_i = x_j.$$

Then, all trajectories $x(t)$ of (1) starting in S converge to N as $t \to \infty$.

Proof. We will prove convergence using the Invariance Principle given as Theorem 9 in the Appendix. To apply it, we will, first, need to find a suitable Lyapunov function for system (1). Consider the following Lyapunov function candidate

$$V_{max-min}(x) = \max(x) - \min(x).$$

While it immediately follows from the general contraction Lemma 3 that $V_{max-min}(x)$ is non-increasing along the trajectories of system (1), in order to prove convergence to a set, we need a more detailed analysis that would distinguish the cases when $V_{max-min}(x)$ decreases and when it does not change along the system's trajectories.

From Theorem 8, it follows that

$$\partial V_{max-min}(x) = \begin{cases} \{P^\top[\alpha^\top, -\beta^\top, \mathbb{0}^\top]^\top\}, & \text{if } x \in S \setminus N, \\ \{P^\top(\alpha - \beta)\}, & \text{if } x \in N, \end{cases}$$

where convex combination coefficients α_i correspond to the agents in $I_{max}(x) = \{i \mid x_i = \max(x)\}$, convex combination coefficients β_i correspond to the agents in $I_{min}(x) = \{i \mid x_i = \min(x)\}$, $\mathbb{0}$ correspond to the rest of the agents $I_{mid}(x) = \{1,\dots,n\} \setminus I_{max}(x) \setminus I_{min}(x)$, and permutation matrix $P^\top$ restores the original agent order.

Now, for each $\xi \in \partial V_{max-min}(x)$, we are interested in the values of inner products $\langle \xi, f(x) \rangle$, which, according to Definition 2, comprise the set-valued Lie derivative

$$\widetilde{\mathscr{L}}_f V_{max-min}(x) = \{a \in \mathbb{R} \mid \forall \xi \in \partial V_{max-min}(x) : \langle \xi, f(x) \rangle = a\}$$

of $V_{max-min}(x)$ at x along the trajectories of system (1). Our immediate goal is to understand when $\max \widetilde{\mathscr{L}}_f V_{max-min}(x)$ is negative and when it is zero, depending on the chosen $x \in S$.

If $x \in N$, that is, $x = \alpha\mathbb{1}$ for some $\alpha \in [-1,1]$, then $f(x) = -A(x)Lx = -A(x)\alpha(L\mathbb{1}) = \mathbb{0}$ and, thus, $\forall \xi \in \partial V_{max-min}(x)$: $\langle \xi, f(x) \rangle = 0$, so $\widetilde{\mathscr{L}}_f V_{max-min}(x) = \{0\}$.

If $x \in S \setminus N$, then let us investigate the possible values of $\langle \xi, f(x) \rangle$, w.l.o.g., dropping P in the expression for ξ, for readability, and using the same agent order in $f(x)$ as in ξ:

$$\begin{aligned} &\langle \xi, f(x) \rangle = -\xi^\top A(x) L x \\ &= -\begin{bmatrix} \alpha \\ -\beta \\ \mathbb{0} \end{bmatrix}^\top \begin{bmatrix} A_{max}(x) & \mathbb{0} & \mathbb{0} \\ \mathbb{0} & A_{min}(x) & \mathbb{0} \\ \mathbb{0} & \mathbb{0} & A_{mid}(x) \end{bmatrix} \times \\ &\times \begin{bmatrix} (I - W_{11}) & -W_{12} & -W_{13} \\ -W_{21} & (I - W_{22}) & -W_{23} \\ -W_{31} & -W_{32} & (I - W_{33}) \end{bmatrix} \begin{bmatrix} x_{max} \\ x_{min} \\ x_{mid} \end{bmatrix} \\ &= -\big(\alpha^\top A_{max}(x)(x_{max} - [W_{11} W_{12} W_{13}]x) + \\ &\quad \beta^\top A_{min}(x)([W_{21} W_{22} W_{23}]x - x_{min})\big), \end{aligned}$$

where x_{max}, x_{min}, and x_{mid} are the states of the agents from $I_{max}(x)$, $I_{min}(x)$, and $I_{mid}(x)$, respectively; $A_{max}(x)$, $A_{min}(x)$, and $A_{mid}(x)$ are the diagonal matrices of susceptibilities of the agents from these three agent subsets; and adjacency matrix W is partitioned ordering the agents as $I_{max}(x)$, $I_{min}(x)$, $I_{mid}(x)$.

Since $x \in S \setminus N$, then $x \neq \alpha\mathbb{1}$ and $\max(x) > \min(x)$. Hence, from the assumption of the theorem about the agreement of $A_{ii}(x)$ upon zeros, it follows that at least one of the inequalities $\operatorname{diag}(A_{max}(x)) > \mathbb{0}$ and $\operatorname{diag}(A_{min}(x)) > \mathbb{0}$ holds. Let us assume, for now, that $\operatorname{diag}(A_{max}(x)) > \mathbb{0}$ and focus on the first term $\alpha^\top A_{max}(x)(x_{max} - [W_{11} W_{12} W_{13}]x)$ of the obtained expression for $\langle \xi, f(x) \rangle$.

Due to row-stochasticity of W and strong connectivity of $G(W)$, for $(x_{max} - [W_{11} W_{12} W_{13}]x)_i = 0$ to hold, all out-neighbors of agent $i \in I_{max}(x)$ in $G(W)$ must also be from $I_{max}(x)$. If there are no such agents i, then $x_{max} - [W_{11} W_{12} W_{13}]x > \mathbb{0}$. Additionally, since α is comprised of a convex combination's coefficients, and $\operatorname{diag}(A_{max}(x)) > \mathbb{0}$, then at least one element of $\alpha^\top A_{max}(x)$ is positive. Hence, $\alpha^\top A_{max}(x)(x_{max} - [W_{11} W_{12} W_{13}]x) > 0$, and $\max \widetilde{\mathscr{L}}_f V_{max-min}(x) < 0$.

If there is an agent $i \in I_{max}(x)$ with its entire out-neighborhood consisting of the members of $I_{max}(x)$, then $(x_{max} - [W_{11} W_{12} W_{13}]x)_i = 0$. However, $x_{max} - [W_{11} W_{12} W_{13}]x \neq \mathbb{0}$, as the opposite would be possible either if agents $I_{max}(x)$ were disconnected from the rest of the network (which is impossible due to the network's strong connectivity assumption), or x was a consensus state (which is impossible, as such states are absent from $S \setminus N$). Thus, there is $j \in I_{max}(x)$ such that $(x_{max} - [W_{11} W_{12} W_{13}]x)_j = \delta > 0$. Now, however, if we put $\alpha_1 = e_i$ and $\alpha_2 = (e_i + e_j)/2$, with e_k being the k'th element of the standard basis, we will have

$$\begin{aligned} &\alpha_1^\top A_{max}(x)(x_{max} - [W_{11} W_{12} W_{13}]x) = 0, \\ &\alpha_2^\top A_{max}(x)(x_{max} - [W_{11} W_{12} W_{13}]x) = \delta/2 (A_{max}(x))_{jj} > 0. \end{aligned}$$

Consequently, for a given $x \in S \setminus N$, term $\alpha^\top A_{max}(x)(x_{max} - [W_{11} W_{12} W_{13}]x)$ takes at least two different values, depending on the choice of α. It can be analogously shown that, if $\operatorname{diag}(A_{min}(x)) > \mathbb{0}$, then $\beta^\top A_{min}(x)([W_{21} W_{22} W_{23}]x - x_{min})$ also takes at least two different values, for different β. Hence, since α and β can be chosen independently, and at least one of the terms $\alpha^\top A_{max}(x)$ and $\beta^\top A_{min}(x)$ is not $\mathbb{0}$, we conclude that, if $x \in S \setminus N$, and there are some agents in $I_{max}(x)$ whose entire out-neighborhood is also in $I_{max}(x)$, then

$$\exists \xi_1 \neq \xi_2 : \langle \xi_1, f(x) \rangle \neq \langle \xi_2, f(x) \rangle,$$

which entails $\widetilde{\mathscr{L}}_f V_{max-min}(x) = \varnothing$ and, by convention, $\max \widetilde{\mathscr{L}}_f V_{max-min}(x) = \max \varnothing = -\infty < 0$.

To summarize, we have so far shown that, if $x \in S \setminus N$, then $\max \widetilde{\mathscr{L}}_f V_{max-min}(x) < 0$, and, if $x \in N$, then $\widetilde{\mathscr{L}}_f V_{max-min}(x) = \{0\}$. Additionally, it immediately follows from Theorem 7 that $V_{max-min}(x)$ is Lipschitz and regular on S. Thus, $V_{max-min}(x)$ is a Lyapunov function for system (1).

Finally, we notice that, by assumption, S is compact and forward invariant w.r.t. system (1). Additionally, N, in which $0 \in \widetilde{\mathscr{L}}_f V_{max-min}(x)$, is forward invariant w.r.t. system (1)—as it entirely consists of equilibrium points—and, clearly, is the largest closed subset of itself. These two facts, taken together with the existence of Lyapunov function $V_{max-min}(x)$, allow us to conclude that, by Invariance Principle, all trajectories $x(t)$ of system (1) starting in S converge to N as $t \to \infty$. □

Having proven a sufficient condition for the convergence to consensus of the general model, we will proceed with a comprehensive analysis of convergence for the three specialized models, starting with the model with stubborn positives.

Theorem 4 (Convergence—Stubborn Positives)**.** *Suppose that W is a row-stochastic adjacency matrix of a strongly connected network, and $x(t)$ evolves according to the model with stubborn positives*

$$\dot{x} = f(x) = -A(x)Lx, \quad A(x) = \tfrac{1}{2}(I - \mathrm{diag}(x)). \tag{3}$$

Then,

- *if $x(0) < \mathbb{1}$, then $\lim_{t\to\infty} x(t) = \alpha\mathbb{1}$, $\alpha \in [0,1)$;*
- *if exists i such that $x_i(0) = 1$, then $\lim_{t\to\infty} x(t) = \mathbb{1}$.*

In other words, in the absence of the agents initially having extreme states, $x(t)$ converges to some consensus state $\alpha\mathbb{1}$, and if there is at least one agent initially holding the extreme state of 1, then all agents approach that state as $t \to \infty$.

Proof. Since the convergence behavior of model (3) varies across the state space, let us, first, partition the latter and, then, prove convergence in each part individually. Consider the following state space partition (see Fig. 1):

$$[-1,1]^n = \lim_{\epsilon\to+0} S_0(\epsilon) \cup S_1,$$
$$S_0(\epsilon) = [-1, 1-\epsilon]^n,$$
$$S_1 = \{P^\intercal x \mid x \in \cup_{k=1}^{n} \{1\}^k \times [-1,1)^{n-k}\},$$

where $N_0(\epsilon) = \{\alpha\mathbb{1} \mid \alpha \in [-1, 1-\epsilon]\}$ and $N_1 = \{\mathbb{1}\}$ are the sets of consensus states in $S_0(\epsilon)$ and S_1, respectively, convergence to which is expected.

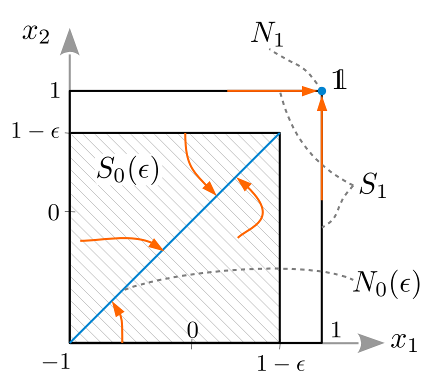

Fig. 1. Convergence behavior of the model with stubborn positives in two dimensions, as well as the partition of the state space. Several trajectories representative of the model's behavior are displayed as solid arrows.

(i) Convergence from $S_0(\epsilon)$ to $N_0(\epsilon)$: We will prove convergence using the general convergence Theorem 3, with $S = S_0(\epsilon)$ and $N = N_0(\epsilon)$. To apply the theorem, we need to prove the agreement upon zeros of the susceptibility functions $A_{ii}(x)$ and forward invariance of $S_0(\epsilon)$. (Theorem 3 also requires both S and N to be non-empty, and S to be compact. However, whenever we use Theorem 3, non-emptiness trivially follows from the definition of these sets, and compactness of S immediately follows from Heine-Borel theorem, as we always choose $S \subseteq \mathbb{R}^n$, $n < \infty$ to be both bounded and closed. Thus, we will further omit the discussion of these two statements about S and N from our proofs.)

Firstly, as $A(x) = \frac{1}{2}(I - \mathrm{diag}(x))$ and, thus, $A_{ii}(x) = \frac{1}{2}(1 - x_i)$, it is clear that, if $A_{ii}(x) = A_{jj}(x) = 0$, then $x_i = x_j = 1$, which proves the zero-agreement property

$$\forall i,j \in 1,\dots,n : A_{ii}(x) = A_{jj}(x) = 0 \to x_i = x_j.$$

In order to prove forward invariance of $S_0(\epsilon)$ w.r.t. system (3), we notice that, according to contraction Lemma 3, $V_{max}(x) = \max(x)$ is non-increasing along the trajectories of system (3), and, at the same time, from the well-posedness Theorem 1, we know that $\min(x) \geq -1$ for all $x \in [-1,1]^n$. Consequently, all the trajectories of the system starting inside cube $S_0(\epsilon)$ remain in it as $t \to \infty$.

Now, by invoking the general convergence Theorem 3, we conclude that all trajectories of system (3) starting in $S_0(\epsilon)$ converge to $N_0(\epsilon)$ as $t \to \infty$.

(ii) Convergence from S_1 to N_1: The agreement upon zeros property of $A_{ii}(x)$ has already been proven above. As to forward invariance of S_1 w.r.t. system (3), it follows immediately from Lemma 2 about the invariance of the closed agent subset. Thus, by the general convergence Theorem 3, all trajectories of system (3) starting in S_1 converge to N_1 as $t \to \infty$. □

Theorem 5 (Convergence—Stubborn Neutrals)**.** *Suppose that W is a row-stochastic adjacency matrix of a strongly connected network, and $x(t)$ evolves according to the model with stubborn neutrals*

$$\dot{x} = f(x) = -A(x)Lx, \quad A(x) = \mathrm{diag}(x)^2. \tag{4}$$

Then,

- *if $x(0) > \mathbb{0}$, then $\lim_{t\to\infty} x(t) = \alpha\mathbb{1}$, $\alpha \in (0,1]$;*
- *if $x(0) < \mathbb{0}$, then $\lim_{t\to\infty} x(t) = \alpha\mathbb{1}$, $\alpha \in [-1,0)$;*
- *otherwise, $\lim_{t\to\infty} x(t) = \mathbb{0}$.*

In other words, if the initial states of all agents are positive, then $x(t)$ converges to an element-wise positive consensus state; if the initial states are all negative, then the convergence is to a negative consensus; finally, if either there are some closed agents, or some open agents' states have opposite signs, then $x(t)$ converges to $\mathbb{0}$ as $t \to \infty$.

Proof. Let us, first, partition the state space and, then, prove convergence for each part individually (see Fig. 2).

$$[-1,1]^n = S_0 \cup \lim_{\epsilon\to+0} S_-(\epsilon) \cup \lim_{\epsilon\to+0} S_+(\epsilon),$$
$$S_-(\epsilon) = [-1,-\epsilon]^n, \quad N_-(\epsilon) = \{\alpha\mathbb{1} \mid \alpha \in [-1,-\epsilon]\},$$
$$S_+(\epsilon) = [\epsilon,1]^n, \qquad N_+(\epsilon) = \{\alpha\mathbb{1} \mid \alpha \in [\epsilon,1]\},$$
$$S_0 = \{x \in [-1,1]^n \mid \prod_{i=1}^{n} x_i = 0 \vee \exists i,j : \mathrm{sgn}(x_i x_j) = -1\},$$
$$N_0 = \{\mathbb{0}\},$$

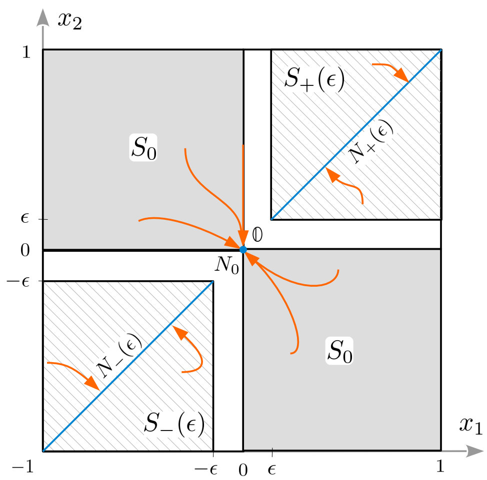

Fig. 2. Convergence behavior of the model with stubborn neutrals in two dimensions, as well as the partition of the state space. Several trajectories representative of the model's behavior are shown as solid arrows.

where we expect convergence from $S_-(\epsilon)$, $S_+(\epsilon)$, and S_0 to $N_-(\epsilon)$, $N_+(\epsilon)$, and N_0, respectively.

(i) Convergence from $S_-(\epsilon)$ to $N_-(\epsilon)$: We will prove convergence using the general convergence Theorem 3, which requires that we prove the agreement upon zeros of functions $A_{ii}(x)$ and forward invariance of $S_-(\epsilon)$.

Since $A(x) = \mathrm{diag}(x)^2$ and $A_{ii}(x) = x_i^2$, it is clear that $A_{ii}(x) = 0 \Leftrightarrow x_i = 0$, thus, proving the zero-agreement property

$$\forall i, j \in 1, \ldots, n : A_{ii}(x) = A_{jj}(x) = 0 \to x_i = x_j.$$

Forward invariance of $S_-(\epsilon)$ immediately follows from the facts that $V_{max}(x)$ is non-increasing along the trajectories of the system due to the general contraction Lemma 3, and $\max(x) \geq -1$ following from the well-posedness Theorem 1.

We, now, can invoke the general convergence Theorem 3 and conclude that all trajectories of (4) starting in $S_-(\epsilon)$ converge to $N_-(\epsilon)$ as $t \to \infty$.

(ii) Convergence from $S_+(\epsilon)$ to $N_+(\epsilon)$: The proof is identical to the proof for the case of $S_-(\epsilon)$ and $N_-(\epsilon)$ and, as such, is omitted.

(iii) Convergence from S_0 to N_0: The agreement of $A_{ii}(x)$ upon zeros has already been proven in part (i). As to forward invariance of S_0, there are two qualitatively different ways a trajectory of the system can leave one of the mixed-sign orthants S_0 consists of: either a trajectory leaves cube $[-1,1]^n$ or it escapes into either the positive or the negative orthant. The former is impossible due to the well-posedness Theorem 1, which states that the trajectories cannot leave the state space $[-1,1]^n$. The latter is also impossible, because, in order for a continuous trajectory $x(t)$ to leave from a mixed-sign to the negative or the positive orthant, the closed agent subset $I_{closed}(x)$ has to change when a trajectory passes a mixed-sign orthant's boundary, which would contradict the agent subset invariance Lemma 2. Thus, by the general convergence Theorem 3, all trajectories of (4) starting in S_0 converge to N_0 as $t \to \infty$. □

Theorem 6 (Convergence—Stubborn Extremists)**.** *Suppose that W is a row-stochastic matrix of a strongly-connected network, and state $x(t)$ is governed by the model with stubborn extremists*

$$\dot{x} = -A(x)Lx, \quad A(x) = (I - \mathrm{diag}(x)^2). \tag{2}$$

Further, assume that the agent set is partitioned as

$$I_{open} = \{i \mid A_{ii}(x(0)) > 0\},$$
$$I_{closed} = \{i \mid A_{ii}(x(0)) = 0\}.$$

Then, the following holds:

- If $I_{closed} = \varnothing$, then $\lim_{t\to\infty} x(t) = \alpha\mathbb{1}$, for some $\alpha \in [-1,1]$, that is, if there are no closed agents, then the system converges to a consensus.

- If $I_{closed} \neq \varnothing$, yet, $\forall i, j \in I_{closed} : x_i = x_j = \alpha \in \{-1,1\}$, then $\lim_{t\to\infty} x(t) = \alpha\mathbb{1}$. In other words, if there are some closed agents, all of whom agree on the state α, then the system converges to that consensus value.

- If $I_{closed} \neq \varnothing$, $\exists i, j \in I_{closed} : x_i \neq x_j$, and a permutation matrix P structures the adjacency matrix W of the network so that the closed agents I_{closed} precede the open agents I_{open} in in $PWP^\intercal$, then $\lim_{t\to\infty} x(t) = x^* = P^\intercal[x_1^{*\intercal}, x_2^{*\intercal}]^\intercal$, where x_1^* are the initial states of the closed agents, and $x_2^* = (I - W_{22})^{-1} W_{21} x_1^*$ if $|I_{open}| > 0$ and $x_2^* = [\;]_{0\times 1}$ otherwise. In other words, if there are multiple closed agents disagreeing on the state, then the system converges to the defined above state x^* of disagreement.

Proof. As the behavior of the system varies across the state space (see Fig. 3), let us, first, partition the latter and, then, prove convergence for each part individually.

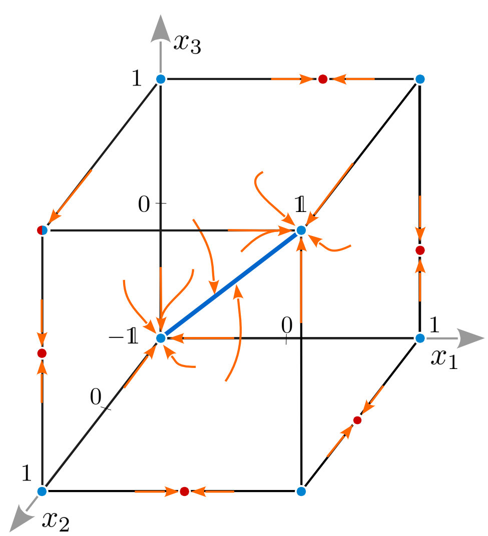

Fig. 3. Convergence behavior of the model with stubborn extremists in three dimensions. A few representative trajectories are displayed as solid arrows; the solid diagonal correspond to the consensus equilibrium states of the system; the circles correspond to the disagreement equilibrium states; the corners are the states of disagreement where all agents are closed, while in the states in the interior of the cube's edges, some agents remain open (depending on the location of closed agents in the network, some edges may have no such internal points of equilibrium, like in the case of edge [-1,-1,1]-[-1,1,1]).

$$\begin{aligned}
[-1,1]^n = &\lim_{\epsilon\to+0} S_0(\epsilon) \cup \lim_{\epsilon\to+0} S_-(\epsilon) \cup \lim_{\epsilon\to+0} S_+(\epsilon) \\
&\bigcup_{|I_{max}\cup I_{min}|\geq 2} S_*(I_{max}, I_{min}), \\
&S_0(\epsilon) = [-1+\epsilon, 1-\epsilon]^n, \\
&S_-(\epsilon) = \{Px \mid x \in \cup_{k=1}^n \{-1\}^k \times [-1, 1-\epsilon]^{n-k}\}, \\
&S_+(\epsilon) = \{Px \mid x \in \cup_{k=1}^n \{\ \ 1\}^k \times [-1+\epsilon, 1]^{n-k}\}, \\
&S_*(I_{max},\, I_{min}) = \\
&\qquad \{x \mid i \in I_{max} \leftrightarrow x_i = 1 \wedge i \in I_{min} \leftrightarrow x_i = -1\},
\end{aligned}$$

Above, $S_0(\epsilon)$ corresponds to the interior of the state space, comprised of the states without extremist agents; $S_-(\epsilon)$ and $S_+(\epsilon)$ correspond to the parts of the state space's surface where all extremist agents are either in state -1 or in state 1, respectively; and $S_*(I_{max}, I_{min})$ $(|I_{max} \cup I_{min}| \geq 2)$ are the "edges" in which there are necessarily multiple extremist agents having different opinions. Now, we will study convergence of system (2) inside each of the above defined parts of the state space.

(i) Convergence in $S_0(\epsilon)$ (see Fig. 4): From the general

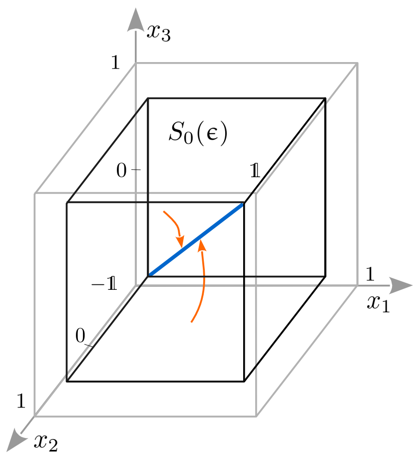

Fig. 4. Convergence in $S_0(\epsilon)$.

contraction Lemma 3, it follows that $S_0(\epsilon)$ is forward invariant w.r.t. (2). Additionally, inside $S_0(\epsilon)$, the agents cannot hold extreme opinions, so $A_{ii}(x) > 0$. Thus, it immediately follows from the general convergence Theorem 3 that all trajectories of system (2) starting in $S_0(\epsilon)$ converge to the latter's subset of consensus states, that is, $N_0(\epsilon) = \{\alpha\mathbb{1} \mid \alpha \in [-1+\epsilon, 1-\epsilon]\}$, as $t \to \infty$.

(ii) Convergence in $S_-(\epsilon)$ (see Fig. 5): forward invariance of $S_-(\epsilon)$ follows from the general contraction Lemma 3. Additionally, since the states of $S_-(\epsilon)$ can only have extremist agents in state -1, then $A_{ii}(x) = 0 \Leftrightarrow x_i = -1$, and, hence the zero-agreement property

$$\forall i, j \in 1, \ldots, n : A_{ii}(x) = A_{jj}(x) = 0 \to x_i = x_j$$

holds in $S_-(\epsilon)$. Thus, by the general convergence Theorem 3, all trajectories of system (2) starting in $S_-(\epsilon)$ converge to $N_-(\epsilon) = S_-(\epsilon) \cap \{\alpha\mathbb{1} \mid \alpha \in [-1, 1]\} = \{-\mathbb{1}\}$ as $t \to \infty$.

(iii) Convergence in $S_+(\epsilon)$: The case is identical to the case of $S_-(\epsilon)$, with the extreme state 1 replacing the extreme state

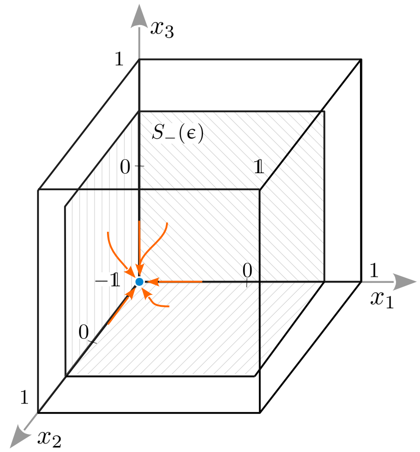

Fig. 5. Convergence in $S_-(\epsilon)$.

-1, and the set to which convergence is expected to occur being $N_+(\epsilon) = \{\mathbb{1}\}$.

(iv) Convergence in $S_*(I_{max}, I_{min})$ (see Fig. 6): In this case, the general convergence Theorem 3 is not applicable, as we expect convergence to a state $x^* = P^\intercal[x_1^{*\intercal}, x_2^{*\intercal}]^\intercal$ of disagreement, with x_1^* corresponding to the initial states of the closed agents of $I_{closed} = I_{max} \cup I_{min} = \{i \mid |x_i| = 1\}$, $x_2^* = (I - W_{22})^{-1} W_{21} x_1^*$, and the adjacency matrix W being structured according to the partition I_{closed}, I_{open}. Notice, that if either $|I_{min} \cup I_{max}| = n$, or

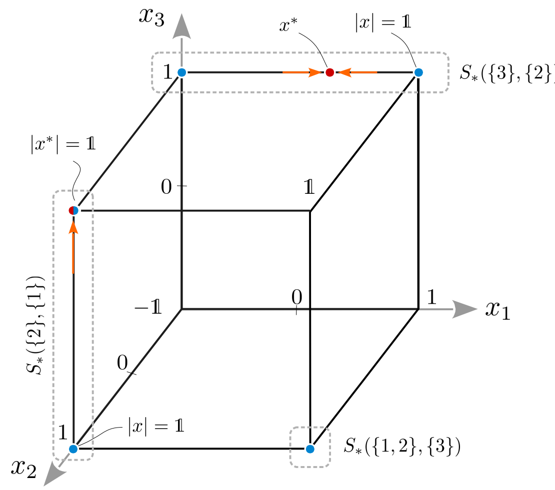

Fig. 6. Convergence in $S_*(I_{max}, I_{min})$.

if the open agents can be reached only from the extremists in one state, then $|x^*| = \mathbb{1}$, that is, the states of all the agents may asymptotically become extreme.

First, we will construct a Lyapunov function out of max-min functions, then, re-use it to prove invariance and, eventually, convergence to a state of disagreement using the Invariance Principle.

Consider function $V_{*max}(x) = \max(x - x^*)$, where x^* is the state of disagreement defined above. From Theorem 8, it follows that

$$\partial V_{*max}(x) = \{P^\intercal[\alpha^\intercal, \mathbb{0}^\intercal]^\intercal\},$$

where α, consisting of the coefficients of a convex combination, correspond to the agents of $I_{*max}(x) = \{i \mid (x-x^*)_i = \max(x-x^*)\}$, $\mathbb{0}$ correspond to the rest of the agents, and $P^\intercal$ is a permutation matrix restoring the original agent order. Now, for any $x \in S_*(I_{max}, I_{min})$ and $\xi \in \partial V_{*max}(x)$, up to reordering of the agents, we will have

$$\langle \xi, f(x) \rangle = -\alpha^\intercal A_{*max}(x)(x_{*max} - [W_{11}, W_{12}]x), \tag{5}$$

where $A_{*max}(x)$ are the susceptibilities of the agents of $I_{*max}(x)$, and the adjacency matrix is structured according to the agent set partition $I_{*max}(x)$, $\{1,\dots,n\} \setminus I_{*max}(x)$. Our immediate goal is to determine the sign of the obtained expression (5). To that end, consider factor $(x_{*max} - [W_{11}, W_{12}]x)$, letting x^*_{*max} be the part of x^* corresponding to the agents of $I_{*max}(x)$

$$\begin{aligned} &x_{*max} - [W_{11}, W_{12}]x \\ &= x_{*max} \underbrace{-x^*_{*max} + x^*_{*max}}_{\mathbb{0}} - [W_{11}, W_{12}](x \underbrace{-x^* + x^*}_{\mathbb{0}}) \\ &= (x_{*max} - x^*_{*max}) - [W_{11}, W_{12}](x - x^*) + \underbrace{(x^*_{*max} - [W_{11}, W_{12}]x^*)}_{\mathbb{0}} \\ &= (x-x^*)_{*max} - [W_{11}, W_{12}](x-x^*). \end{aligned}$$

It is clear from the definition of $I_{*max}(x)$ that $(x-x^*)_{*max} - [W_{11}, W_{12}](x-x^*) \geq 0$ and, hence $\langle \xi, f(x) \rangle \leq 0$. Furthermore, we can apply the argument from the proof of the general convergence Theorem 3, to establish that when there is an agent i such that $((x-x^*)_{*max} - [W_{11}, W_{12}](x-x^*))_i = 0$, we can vary α to make $\langle \xi, f(x) \rangle$ take different values for the same x. Thus, we can conclude that $\widetilde{\mathscr{L}_f} V_{*max}(x) = \{0\}$ when $|x| = \mathbb{1}$ (as $\xi = \mathbb{0}$) or when $x = x^*$, and $\max \widetilde{\mathscr{L}_f} V_{*max}(x) < 0$ for the other $x \in S_*(I_{max}, I_{min})$.

We can repeat the same reasoning to establish that, for

$$V_{-*min} = -\min(x - x^*),$$

it holds that $\widetilde{\mathscr{L}_f} V_{-*min}(x) = \{0\}$ when $|x| = \mathbb{1}$ or $x = x^*$, and $\max \widetilde{\mathscr{L}_f} V_{-*min}(x) < 0$ for the rest of $x \in S_*(I_{max}, I_{min})$.

Our reasoning about $V_{*max}(x)$ and $V_{-*min}(x)$ allow us to conclude that function

$$\begin{aligned} V_{*max-min}(x) &= V_{*max}(x) + V_{-*min}(x) \\ &= \max(x-x^*) - \min(x-x^*) \end{aligned}$$

is a Lyapunov function for system (2), as required by the Invariance Principle. Additionally, $S_*(I_{max}, I_{min})$ is forward invariant, which immediately follows from the agent subset invariance Lemma 2. Thus, by Invariance Principle, all trajectories of system (2) starting in $S_*(I_{max}, I_{min})$ converge to set

$$N_*(I_{max}, I_{min}) = \{x \mid |x| = 1\} \cup \{x^*\},$$

in which $0 \in \widetilde{\mathscr{L}_f} V_{*max-min}(x)$. What remains to show is what element of $N_*(I_{max}, I_{min})$ the system converges to.

Clearly, if all the agents are initially closed, that is, $|x(0)| = \mathbb{1}$, then $\lim_{t \to \infty} x(t) = x(0)$, which follows from the agent subset invariance Lemma 2. Now, assume that $|x(0)| \neq \mathbb{1}$. In such a case, a trajectory cannot approach any element of $\{x \mid |x| = \mathbb{1}\}$ (except, possibly, x^* in the case when the open agents are only reachable by the closed agents having the same state, and, as a result, $|x^*| = \mathbb{1}$), as, generally, approaching one of these states would violate at least one of the above proven inequalities

$$\begin{aligned} \max \widetilde{\mathscr{L}_f} V_{*max}(x) &= \max \widetilde{\mathscr{L}_f} \max(x-x^*) \leq 0, \\ \max \widetilde{\mathscr{L}_f} V_{-*min}(x) &= \max \widetilde{\mathscr{L}_f} (-\min(x-x^*)) \leq 0. \end{aligned}$$

Hence, if $|x(0)| \neq \mathbb{1}$, then the trajectories of (2) converge to x^* as $t \to \infty$. □

V. Discussion

In this section, we summarize and interpret the obtained results, as well as assess where they fit in and how contribute to the existing body of research.

New models: In this work, we have defined the general model of polar opinion dynamics

$$\dot{x} = -A(x)Lx, \tag{1}$$

that, depending on how we define $A(x)$, has interpretation in terms of one of the socio-psychological theories. Model (1) can be viewed as a non-linear analog of DeGroot [21] and Friedkin-Johnsen [29] models, with the dependence of the agents' susceptibilities $A(x)$ to persuasion upon their current opinions being the key distinguishing trait of our model. Mathematically, model (1) is also related to the class of bounded confidence models [64], [14], [52], [36], [20], in which the opinion-adoption behavior of the agents also depends on the agent's current beliefs, yet, this dependence is based upon the socio-psychological principles different from the ones we consider. A notable exception is the work of Sobkowicz [64], in which, the author uses the opinion resilience mechanisms similar to the ones we use in our models with stubborn positives and stubborn extremists.

Behavior of the general model: For the general model (1), in Theorem 3, we have provided a sufficient condition for the convergence to consensus. Roughly speaking, a trajectory starting inside a forward invariant set approaches a state of consensus if all closed agents have similar states, that is, $A_{ii}(x) = A_{jj}(x) = 0 \to x_i = x_j$. From the sociological perspective, it means that, as long as all the ultimately stubborn agents in the network agree upon their states, the agents will eventually agree, as there is no force that would drive the system to disagreement. The observed behavior is different from Friedkin-Johnsen model in that, in our model, the presence of non-fully open agents, having $A_{ii}(x) < 1$, does not immediately lead to an asymptotic disagreement. The obtained sufficient condition for convergence to consensus, compared to its analogs derived in [51], [9], is better interpretable from the sociological point of view, and is not more restrictive.

Behavior of the specialized models: In addition to the general model, we have considered three specialized models

$$\begin{aligned} \dot{x} &= -\tfrac{1}{2}(I - \operatorname{diag}(x))Lx, && \textit{(stubborn positives)} \\ \dot{x} &= -\operatorname{diag}(x)^2 Lx, && \textit{(stubborn neutrals)} \\ \dot{x} &= -(I - \operatorname{diag}(x)^2)Lx. && \textit{(stubborn extremists)} \end{aligned}$$

The behavior of these models, studied in Theorems 4, 5, and 6, can be summarized as follows.

Convergence to consensus (closed agents absent): If there are no closed agents in the network, that is, if $A(x(0)) > \mathbb{0}$, and, as a result, every agent is at least to some degree susceptible to persuasion, then the system converges to a state of consensus, as shown for the example of the model with stubborn positives in Fig. 7. In the absence of closed agents, the particular consensus value is known only for the model with stubborn neutrals when the agents with both positive and negative states are present—in this case, the system converges to $\mathbb{0}$.

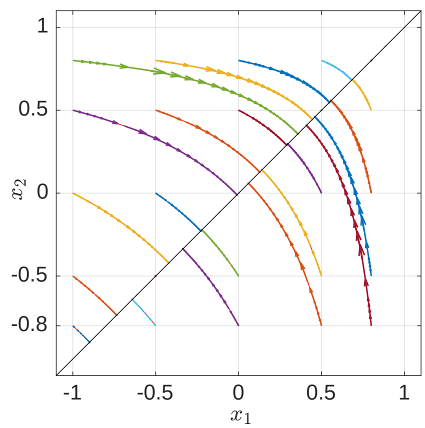

Fig. 7. Phase portrait: convergence to consensus of the model with stubborn positives in the absence of closed agents.

This behavior is not surprising, since, when term $A(x)$ of the vector field $-A(x)Lx$ does not prevent any agent from changing its state, the negative Laplacian expectedly drives the state toward a consensus. Thus, if there are no ultimately stubborn agents, then the group will asymptotically reach an agreement.

Convergence to consensus (closed agents present): In the presence of closed agents all of whom agree on their state $\alpha \in [-1,1]$, the system converges to consensus $\alpha\mathbb{1}$. In other words, if the ultimately stubborn agents are present and share the same opinion, they will persuade the rest of the group to adopt that opinion. A representative example of such behavior is given in Fig. 8, for the model with stubborn neutrals. A natural conclusion is that the only force that can counteract the persuasion efforts of the ultimately stubborn agents agreeing on an opinion is the ultimately stubborn agents having a different opinion.

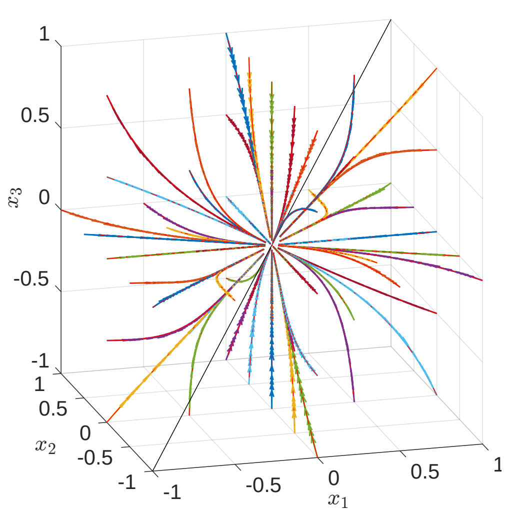

Fig. 8. Phase portrait: convergence to consensus of the model with stubborn neutrals when closed agents are present.

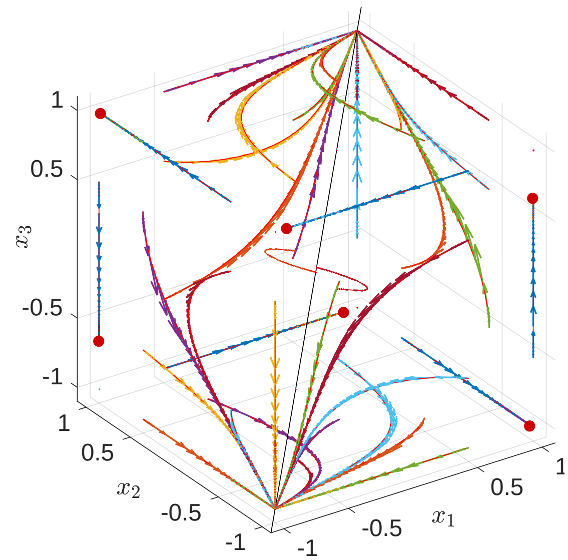

Fig. 9. Phase portrait: asymptotic behavior of the model with stubborn extremists. Strictly inside the state space, trajectories converge to a consensus. On the surface of the state space, when all closed agents have the same state α, trajectories converge to that consensus state $\alpha\mathbb{1}$. In the presence of multiple closed agents holding different opinions, the system converges to a state of disagreement (such states are displayed as solid circles).

Convergence to disagreement: Finally, in the presence of multiple closed agents holding different opinions, which is possible only for the model with stubborn extremists, the system converges to a state of disagreement. Fig. 9 shows a full range of qualitatively different asymptotic behaviors of the model with stubborn extremists. If, in addition to the closed agents holding different opinions, there are some open agents in the network, then the closed agents will persuade the open agents to adopt a combination of their opinions. In this case, the particular limiting state to which the system will converge will depend on the structure of the network and the locations and states of the closed agents, yet, not on the initial opinions of the open agents. (A similar behavior has also been observed in the context of the Voter model with stubborn agents [74, Sec. 4]). It is particularly interesting that the opinions of the open agents in the latter limiting state have a rather simple expression $(I-W_{22})^{-1}W_{21}x_1^*$, given in Theorem 6, where W_{22} is the block of the adjacency matrix corresponding to the cluster of the open agents, W_{21} is the block responsible for the influence of the closed agents upon the open agents, and x_1^* are the closed agents' (initial) states.

Model analysis: The bulk of our theoretical analysis of the models' behavior is comprised of the proofs of convergence. The standard tools for the analysis of convergence of nonlinear models, such as LaSalle Invariance Principle, require existence of a smooth Lyapunov function, with quadratic functions being a popular choice. The latter, however, may be hard and, sometimes, provenly impossible [55] to find for a model defined over a general directed network. In this work, we show, using several existing tools from non-smooth analysis, how to apply non-smooth max-min functions to prove convergence of our models. Such Lyapunov functions have been considered in the literature [55], [38], however, this work is the first to provide a full formal analysis of such functions used along with the generalized Invariance Principle. Due to the generality of the non-smooth analysis tools we have used, our analysis can be easily adapted to other non-linear models defined over directed networks, with Lyapunov functions constructed out of convex components.

VI. Acknowledgement

We express our gratitude to Noah Friedkin for pointing out the usefulness of having multiple definitions of the agent susceptibility, providing critique of an early version of this paper, as well as introducing us to the field of social psychology. We are also grateful to Andrew R. Teel for a useful discussion.

References

[1] R. P. Abelson. Mathematical models of the distribution of attitudes under controversy. *Contributions to Mathematical Psychology*, 14:1–160, 1964.

[2] R. P. Abelson. Conviction. *American Psychologist*, 43(4):267, 1988.

[3] D. Acemoglu, G. Como, F. Fagnani, and A. Ozdaglar. Opinion fluctuations and disagreement in social networks. *Mathematics of Operations Research*, 38(1):1–27, 2013.

[4] D. Acemoglu and A. Ozdaglar. Opinion dynamics and learning in social networks. *Dynamic Games and Applications*, 1(1):3–49, 2011.

[5] C. Altafini. Consensus problems on networks with antagonistic interactions. *IEEE Transactions on Automatic Control*, 58(4):935–946, 2013.

[6] A. Bacciotti and F. Ceragioli. Stability and stabilization of discontinuous systems and nonsmooth Lyapunov functions. *ESAIM: Control, Optimisation & Calculus of Variations*, 4:361–376, 1999.

[7] J. N. Bassili. Meta-judgmental versus operative indexes of psychological attributes: The case of measures of attitude strength. *Journal of Personality and Social Psychology*, 71(4):637, 1996.

[8] J. N. Bassili and R. D. Brown. Implicit and explicit attitudes: Research, challenges, and theory. In D. Albarracn, B. T. Johnson, and M. P. Zanna, editors, *Handbook of Attitudes and Attitude Change*, chapter 13, pages 543–574. Lawrence Erlbaum Associates, 2005.

[9] 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*, 54(11):2586–2597, 2009.

[10] A. Borodin, Y. Filmus, and J. Oren. Threshold models for competitive influence in social networks. In *Workshop on Internet and Network Economics*, pages 539–550. Springer, 2010.

[11] Y. Cao, W. Yu, W. Ren, and G. Chen. An overview of recent progress in the study of distributed multi-agent coordination. *IEEE Transactions on Industrial Informatics*, 9(1):427–438, 2013.

[12] T. Carnes, C. Nagarajan, S. M. Wild, and A. Van Zuylen. Maximizing influence in a competitive social network: A follower's perspective. In *International Conference on Electronic Commerce*, pages 351–360. ACM, 2007.

[13] D. Cartwright and F. Harary. Structural balance: A generalization of Heider's theory. *Psychological Review*, 63(5):277, 1956.

[14] C. Castellano, S. Fortunato, and V. Loreto. Statistical physics of social dynamics. *Reviews of Modern Physics*, 81(2):591, May 2009.

[15] S. Chen, D. H. Glass, and M. McCartney. Characteristics of successful opinion leaders in a bounded confidence model. *Physica A: Statistical Mechanics and its Applications*, 449:426–436, 2016.

[16] F. H. Clarke. *Optimization and Nonsmooth Analysis*, volume 5 of *Classics in Applied Mathematics*. SIAM, 1990.

[17] P. Clifford and A. Sudbury. A model for spatial conflict. *Biometrika*, 60(3):581–588, 1973.

[18] J. Cortés. Discontinuous dynamical systems – A tutorial on solutions, nonsmooth analysis, and stability. *IEEE Control Systems Magazine*, 28(3):36–73, June 2008.

[19] G. Deffuant, F. Amblard, and G. Weisbuch. Modelling group opinion shift to extreme: The smooth bounded confidence model. *arXiv preprint cond-mat/0410199*, 2004.

[20] G. Deffuant, D. Neau, F. Amblard, and G. Weisbuch. Mixing beliefs among interacting agents. *Advances in Complex Systems*, 3(1/4):87–98, 2000.

[21] M. H. DeGroot. Reaching a consensus. *Journal of the American Statistical Association*, 69(345):118–121, 1974.

[22] A. H. Eagly and S. Chaiken. *The Psychology of Attitudes*. Harcourt Brace Jovanovich College Publishers, 1993.

[23] E. Even-Dar and A. Shapira. A note on maximizing the spread of influence in social networks. In *Workshop on Internet and Network Economics*, pages 281–286. Springer, 2007.

[24] F. Fagnani and S. Zampieri. Randomized consensus algorithms over large scale networks. *IEEE Journal on Selected Areas in Communications*, 26(4):634–649, 2008.

[25] L. Festinger. A theory of social comparison processes. *Human Relations*, 7(2):117–140, 1954.

[26] L. Festinger. *A Theory of Cognitive Dissonance*, volume 2. Stanford University Press, 1962.

[27] J. R. French Jr. A formal theory of social power. *Psychological Review*, 63(3):181–194, 1956.

[28] N. E. Friedkin. The problem of social control and coordination of complex systems in sociology: A look at the community cleavage problem. *IEEE Control Systems Magazine*, 35(3):40–51, June 2015.

[29] N. E. Friedkin and E. C. Johnsen. Social influence networks and opinion change. *Advances in group processes*, 16(1):1–29, 1999.

[30] J. Ghaderi and R. Srikant. Opinion dynamics in social networks with stubborn agents: Equilibrium and convergence rate. *Automatica*, 50(12):3209–3215, 2014.

[31] J. Goldenberg, B. Libai, and E. Muller. Talk of the network: A complex systems look at the underlying process of word-of-mouth. *Marketing Letters*, 12(3):211–223, 2001.

[32] S. Goyal and M. Kearns. Competitive contagion in networks. In *44th STOC Symposium on Theory of Computing Conference*, pages 759–774. ACM, 2012.

[33] M. Granovetter. Threshold models of collective behavior. *American Journal of Sociology*, 83(6):1420–1443, 1978.

[34] P. Groeber, J. Lorenz, and F. Schweitzer. Dissonance minimization as a microfoundation of social influence in models of opinion formation. *The Journal of Mathematical Sociology*, 38(3):147–174, 2014.

[35] F. Harary. A criterion for unanimity in French's theory of social power. In D. Cartwright, editor, *Studies in Social Power*, pages 168–182. University of Michigan, 1959.

[36] R. Hegselmann and U. Krause. Opinion dynamics and bounded confidence models, analysis, and simulation. *Journal of Artificial Societies and Social Simulation*, 5(3), 2002.

[37] F. Heider. Attitudes and cognitive organization. *The Journal of Psychology*, 21(1):107–112, 1946.

[38] J. M. Hendrickx. *Graphs and networks for the analysis of autonomous agent systems*. PhD thesis, Université Catholique de Louvain, Belgium, February 2008.

[39] J. M. Hendrickx. A lifting approach to models of opinion dynamics with antagonisms. In *IEEE Conf. on Decision and Control*, pages 2118–2123, 2014.

[40] M. Jalili. Effects of leaders and social power on opinion formation in complex networks. *Simulation*, 89(5):578–588, 2013.

[41] P. Jia, A. MirTabatabaei, N. E. Friedkin, and F. Bullo. Opinion dynamics and the evolution of social power in influence networks. *SIAM Review*, 57(3):367–397, 2015.

[42] D. Kempe, J. Kleinberg, and É. Tardos. Maximizing the spread of influence through a social network. In *SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 137–146. ACM, 2003.

[43] H. K. Khalil. *Nonlinear Systems*. Prentice Hall, 3 edition, 2002.

[44] M. Kimura, K. Saito, K. Ohara, and H. Motoda. Learning to predict opinion share in social networks. In *AAAI Conference on Artificial Intelligence*, 2010.

[45] D. Krackhardt. A plunge into networks. *Science*, 326(5949):47–48, 2009.

[46] J. A. Krosnick, D. S. Boninger, Y. C. Chuang, M. K. Berent, and C. G. Carnot. Attitude strength: One construct or many related constructs? *Journal of Personality and Social Psychology*, 65(6):1132–1151, 1993.

[47] J. A. Krosnick and R. E. Petty. Attitude strength: An overview. In *Attitude Strength: Antecedents and Consequences*, chapter 1, pages 1–24. Lawrence Erlbaum Associates, Mahwah, NJ, 1995.

[48] E. Kurmyshev, H. A. Juárez, and R. A. González-Silva. Dynamics of bounded confidence opinion in heterogeneous social networks: Concord against partial antagonism. *Physica A: Statistical Mechanics and its Applications*, 390(16):2945–2955, 2011.

[49] J. L. Lastovicka and D. M. Gardner. Components of involvement. In J. C. Maloney and B. Silverman, editors, *Attitude Research Plays for High Stakes*, pages 53–73. American Marketing Association, 1979.

[50] Z. Lin, B. Francis, and M. Maggiore. State agreement for continuous-time coupled nonlinear systems. *SIAM Journal on Control and Optimization*, 46(1):288–307, 2007.

[51] J. Lorenz. A stabilization theorem for dynamics of continuous opinions. *Physica A: Statistical Mechanics and its Applications*, 355(1):217–223, 2005.

[52] J. Lorenz. Continuous opinion dynamics under bounded confidence: A survey. *International Journal of Modern Physics C*, 18(12):1819–1838, 2007.

[53] S. Martin and J. Hendrickx. Continuous-time consensus under non-instantaneous reciprocity. *IEEE Transactions on Automatic Control*, PP(99), 2015.

[54] A. G. Miller, J. W. McHoskey, C. M. Bane, and T. G. Dowd. The attitude polarization phenomenon: Role of response measure, attitude extremity, and behavioral consequences of reported attitude change. *Journal of Personality and Social Psychology*, 64(4):561, 1993.

[55] L. Moreau. Stability of continuous-time distributed consensus algorithms. In *IEEE Conf. on Decision and Control*, volume 4, pages 3998–4003, Nassau, Bahamas, 2004.

[56] L. Moreau. Stability of multiagent systems with time-dependent communication links. *IEEE Transactions on Automatic Control*, 50(2):169–182, 2005.

[57] R. Olfati-Saber, A. Fax, and R. M. Murray. Consensus and cooperation in networked multi-agent systems. *Proceedings of the IEEE*, 95(1):215–233, 2007.

[58] R. Olfati-Saber and R. M. Murray. Consensus problems in networks of agents with switching topology and time-delays. *IEEE Transactions on Automatic Control*, 49(9):1520–1533, 2004.

[59] C. E. Osgood and P. H. Tannenbaum. The principle of congruity in the prediction of attitude change. *Psychological Review*, 62(1):42–55, January 1955.

[60] M. Pirani and S. Sundaram. Spectral properties of the grounded Laplacian matrix with applications to consensus in the presence of stubborn agents. In *American Control Conference*, pages 2160–2165. IEEE, 2014.

[61] E. M. Pomerantz, S. Chaiken, and R. S. Tordesillas. Attitude strength and resistance processes. *Journal of Personality and Social Psychology*, 69(3):408–419, 1995.

[62] K. Saito, M. Kimura, K. Ohara, and H. Motoda. Selecting information diffusion models over social networks for behavioral analysis. In *ECML PKDD European Conference on Machine Learning and Knowledge Discovery in Databases*, pages 180–195. Springer, September 2010.

[63] N. Schwarz and G. Bohner. The construction of attitudes. In A. Tesser and N. Schwarz, editors, *Blackwell Handbook of Social Psychology: Intraindividual Processes*, chapter 20, pages 436–457. Blackwell Publishing, 2001.

[64] P. Sobkowicz. Extremism without extremists: Deffuant model with emotions. *Frontiers in Physics*, 3:17, 2015.

[65] A. Tahbaz-Salehi and A. Jadbabaie. A necessary and sufficient condition for consensus over random networks. *IEEE Transactions on Automatic Control*, 53(3):791–795, 2008.

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

[67] K. Tucci, J. González-Avella, and M. Cosenza. Rise of an alternative majority against opinion leaders. *Physica A: Statistical Mechanics and its Applications*, 446:75–81, 2016.

[68] G. Verma, A. Swami, and K. Chan. The impact of competing zealots on opinion dynamics. *Physica A: Statistical Mechanics and its Applications*, 395:310–331, 2014.

[69] A. Waagen, G. Verma, K. Chan, A. Swami, and R. D'Souza. Effect of zealotry in high-dimensional opinion dynamics models. *Physical Review E*, 91(2):022811, 2015.

[70] D. J. Watts and P. S. Dodds. Influentials, networks, and public opinion formation. *Journal of Consumer Research*, 34(4):441–458, 2007.

[71] A. G. Whittington, A. M. Hofmeister, and P. I. Nabelek. Temperature-dependent thermal diffusivity of the Earth's crust and implications for magmatism. *Nature*, 458(7236):319–321, 2009.

[72] Z. Xu, J. Liu, and T. Başar. On a modified DeGroot-Friedkin model of opinion dynamics. In *American Control Conference*, pages 1047–1052. IEEE, July 2015.

[73] E. Yildiz, D. Acemoglu, A. Ozdaglar, A. Saberi, and A. Scaglione. Discrete opinion dynamics with stubborn agents. *Submitted to Operations Research, OPRE-2011-01-026*, 2011.

[74] E. Yildiz, A. Ozdaglar, D. Acemoglu, A. Saberi, and A. Scaglione. Binary opinion dynamics with stubborn agents. *ACM Transactions on Economics and Computation*, 1(4):19, 2013.

[75] J. Zaller. *The Nature and Origins of Mass Opinion*. Cambridge University Press, 1992.

APPENDIX
REVIEW OF NON-SMOOTH ANALYSIS

This section contains a review of several tools from non-smooth analysis that prove useful when dealing with non-smooth Lyapunov functions in the proofs of convergence. In what follows, we rely on the standard definitions of *locally Lipschitz* [16, p.9] and *regular* [16, pp.39–40] *functions*.

Definition 1 ((Clarke) Generalized gradient [16]). *Let $V : \mathbb{R}^n \to \mathbb{R}$ be locally Lipschitz, and Ω_V be the set of points where V fails to be differentiable. Then, the* generalized gradient of V *is defined as follows*

$$\partial V(x) = \mathrm{co}\left\{ \lim_{i\to\infty} \nabla V(x_i) \mid x_i \to x, x_i \notin Z \cup \Omega_V \right\},$$

where $\mathrm{co}\{\cdot\}$ is the convex hull, and Z is any set of Lebesgue measure zero. Thus, the generalized gradient of V at x is the convex hull of all gradient values around and approaching x where V is differentiable.

Theorem 7 (Properties of generalized gradient [18]). *Let $V_1, V_2 : \mathbb{R}^n \to \mathbb{R}$ be locally Lipschitz and regular at $x \in \mathbb{R}^n$, $a, b \in [0, \infty) \subset \mathbb{R}$. Then,*

(i) [Scaling rule] *$\partial(a \cdot V_1)(x) = a \cdot \partial V_1(x)$ and $(a \cdot V_1)$ is locally Lipschitz and regular at x.*

(ii) [Sum rule] *$\partial(a \cdot V_1 + b \cdot V_2)(x) = a\partial V_1(x) + b\partial V_2(x)$, and $a \cdot V_1 + b \cdot V_2$ is locally Lipschitz and regular at x. The sum of sets in the expression above is understood in the sense of $A + B = \{a + b \mid a \in A, b \in B\}$.*

(iii) [Max-min functions] *Let $V_i : \mathbb{R}^n \to \mathbb{R}$, $k \in \{1, \ldots, m\} < \infty$ be locally Lipschitz at $x \in \mathbb{R}^n$, and*

$$V_{max}(y) \triangleq \max\{V_k(y) \mid k \in \{1, \ldots, m\}\},$$
$$V_{min}(y) \triangleq \min\{V_k(y) \mid k \in \{1, \ldots, m\}\}.$$

Also, let

$$I_{max}(x) = \{i \mid V_i(x) = V_{max}(x)\},$$
$$I_{min}(x) = \{i \mid V_i(x) = V_{min}(x)\}.$$

Then,

- *V_{max} and V_{min} are locally Lipschitz at x.*
- *If V_i is regular at x for each $i \in I_{max}(x)$, then*

  $$\partial V_{max}(x) = \mathrm{co} \cup \{\partial V_i(x) \mid i \in I_{max}(x)\}$$

  and V_{max} is regular at x.
- *If $-V_i$ is regular at x for each $i \in I_{min}(x)$, then*

  $$\partial V_{min}(x) = \mathrm{co} \cup \{\partial V_i(x) \mid i \in I_{min}(x)\}$$

  and $-V_{min}$ is regular at x.

The following theorem is as a corollary of Theorem 7.

Theorem 8 (Generalized gradients of max-min functions). *Consider functions $V_{max}(x) = \max(x)$, $V_{-min}(x) = -\min(x)$,*

$V_{max-min}(x) = V_{max}(x) + V_{-min}(x)$, where $x \in S \subseteq [-1,1]^n$. Also, let $N = S \cap \{\alpha\mathbb{1} \mid \alpha \in [-1,1]\}$, and define $I_{max}(x)$ and $I_{min}(x)$ as in Theorem 7, and $I_{mid}(x) = \{1,\ldots,n\} \setminus I_{max}(x) \setminus I_{min}(x)$. Then,

$$\partial V_{max}(x) = \{P^\mathsf{T}[\alpha^\mathsf{T}, \mathbb{0}^\mathsf{T}]^\mathsf{T}\},$$

$$\partial V_{-min}(x) = \{P^\mathsf{T}[-\beta^\mathsf{T}, \mathbb{0}^\mathsf{T}]^\mathsf{T}\},$$

$$\partial V_{max-min}(x) = \begin{cases} \{P^\mathsf{T}[\alpha^\mathsf{T}, -\beta^\mathsf{T}, \mathbb{0}^\mathsf{T}]^\mathsf{T}\}, & \text{if } x \in S \setminus N, \\ \{P^\mathsf{T}(\alpha - \beta)\}, & \text{if } x \in N, \end{cases}$$

where α and β are vectors whose elements comprise coefficients of convex combinations ($\alpha_i, \beta_j \geq 0$, $\sum_i \alpha_i = \sum_j \beta_j = 1$), α_i correspond to the agents from $I_{max}(x)$, β_j correspond to the agents from $I_{min}(x)$, $\mathbb{0}_k$ correspond to the agents from $I_{mid}(x)$, and permutation matrices P^T restore the original order of $\{x_i\}$.

Proof. Notice that both $V_{max}(x) = \max(x)$ and $V_{min}(x) = \min(x)$ can been viewed as, respectively, the maximum and the minimum of a finite number of functions $V_i(x) = x_i$, $i \in \{1,\ldots,n\}$. Since each $V_i(x)$ is continuously-differentiable and, thus, locally Lipschitz and regular on S, Theorem 7 allows us to apply the rule (iii) for computing the generalized gradient for max-min functions, followed by the application of the (i) scaling and (ii) sum rules. The statement of the theorem, then, follows immediately. □

Definition 2 (Set-valued Lie derivative [6], [18])**.** *For a locally Lipschitz $V : \mathbb{R}^n \to \mathbb{R}$ and system $\dot{x} = f(x)$, the* set-valued Lie derivative *$\widetilde{\mathscr{L}}_f V(x)$ of V along the trajectories of the system is defined as*

$$\widetilde{\mathscr{L}}_f V(x) = \{a \in \mathbb{R} \mid \forall \xi \in \partial V(x) : \langle \xi, f(x) \rangle = a\}.$$

The following theorem is an analog of the generalized Invariance Principle [6], [18], specialized for the case of a continuous vector field, while the original was stated for differential inclusions.

Theorem 9 (Invariance Principle [6], [18])**.** *If*

(i) $V : \mathbb{R}^n \to \mathbb{R}$ is locally Lipschitz and regular,

(ii) $S \subset \mathbb{R}^n$ is compact and invariant w.r.t. $\dot{x} = f(x)$, and

(iii) $\max \widetilde{\mathscr{L}}_f V(x) \leq 0$ for each $x \in S$, then all solutions $x(t) : [0,\infty) \to \mathbb{R}^n$ starting in S converge to the largest invariant subset M of

$$S \cap \overline{\{x \in \mathbb{R}^n \mid 0 \in \widetilde{\mathscr{L}}_f V(x)\}},$$

where $\overline{\{\cdots\}}$ is set closure. If M is finite, then the limit of each solution $x(0) \in S$ exists and is an element of M.

Victor Amelkin received BSc and MSc degrees in Applied Mathematics and Computer Science from Tula State University, Tula, Russia, in 2006 and 2008, respectively. He is currently a PhD Candidate in the Department of Computer Science at the University of California, Santa Barbara. Earlier, he had spent several years as a software engineer in industry. Victor's research interests include machine learning, linear algebra, combinatorial algorithms, dynamical systems, and their application to modeling of dynamic processes occurring in networks.

Francesco Bullo (IEEE S'95-M'99-SM'03-F'10) is a Professor with the Mechanical Engineering Department at the University of California, Santa Barbara. He received a Laurea in Electrical Engineering from the University of Padova in 1994, and a PhD in Control and Dynamical Systems from the California Institute of Technology in 1999. From 1998 to 2004, he was affiliated with the University of Illinois, Urbana-Champaign. He is the coauthor of the books "Geometric Control of Mechanical Systems" (Springer, 2004) and "Distributed Control of Robotic Networks" (Princeton, 2009). He received the 2008 IEEE CSM Outstanding Paper Award, the 2011 Hugo Schuck Best Paper Award, the 2013 SIAG/CST Best Paper Prize, and the 2014 Automatica Best Paper Award. His main research interest is dynamics and control of multi-agent network systems, with applications to robotic coordination, power systems, distributed computing, and social networks.

Ambuj K. Singh is a Professor of Computer Science at the University of California, Santa Barbara, with part-time appointments in the Biomolecular Science and Engineering Program and the Technology Management Program. He received a B.Tech. degree from the Indian Institute of Technology, Kharagpur, and a PhD degree from the University of Texas at Austin. His research interests are broadly in in the areas of databases, data mining, and bioinformatics. He has published 200 technical papers over his career. He has led a number of multidisciplinary projects including UCSB's Information Network Academic Research Center funded by the US Army. He is currently directing UCSB's Interdisciplinary Graduate Education Research and Training (IGERT) program on Network Science funded by the National Science Foundation (NSF), and the Multidisciplinary University Research Initiative (MURI) on Network Science of Teams funded by the US Army. He has graduated approximately 40 PhD, MS, and postdoctoral students over his career, including 20 PhD students.