

# Patterns of Nonlinear Opinion Formation on Networks

Anastasia Bizyaeva, Ayanna Matthews, Alessio Franci, and Naomi Ehrlich Leonard

**Abstract**—When communicating agents form opinions about a set of possible options, agreement and disagreement are both possible outcomes. Depending on the context, either can be desirable or undesirable. We show that for nonlinear opinion dynamics on networks, and a variety of network structures, the spectral properties of the underlying adjacency matrix fully characterize the occurrence of either agreement or disagreement. We further show how the corresponding eigenvector centrality, as well as any symmetry in the network, informs the resulting patterns of opinion formation and agent sensitivity to input that triggers opinion cascades.

**Index Terms**—Multi-agent systems, decision making, opinion dynamics, consensus, graph theory

## I. INTRODUCTION

Multi-agent systems that perform distributed control tasks in uncertain or dynamic contexts benefit when agents use network communications to form and change opinions about context-dependent options. For example, network communications can help autonomous multi-robot teams make better navigational choices among alternative routes or better allocation choices among alternative tasks. Mathematical models of opinion dynamics over networks are often used to bring a group to a desired opinion configuration. In a task-allocation context, agreement is not necessarily the only desirable opinion configuration, as sometimes agents are better off exploring different routes or doing different tasks.

A general model of opinion dynamics for distributed agents on a network was recently introduced in [1], [2]. In this multi-agent multi-option model, real-valued opinions evolve in continuous time according to a nonlinear update rule that saturates network exchanges. A key feature of the model is the emergence of consensus and dissensus as equilibrium opinion configurations, even when agents are homogeneous, receive no input, and communicate over an all-to-all network. The emergence of consensus and dissensus depends on a small number of parameters that distinguish the interactions between agents as cooperative or competitive.

In [1] the behavior of the general model is examined with particular attention to all-to-all and vertex-transitive cycle network topologies. The work identifies parameter regimes

that correspond to consensus and dissensus (a heterogeneous opinion configuration with neutral average opinion). In the present paper we take first steps to analyze opinion formation on other classes of networks. We consider opinion formation for two options and homogeneous agents that communicate over networks that include  $K$ -regular, bipartite, and strongly connected directed graphs. We examine how network structure influences the group outcome of the opinion formation process, and we prove that, generically, agreement and disagreement arise on these networks. We show that the parameter regimes associated with consensus and dissensus for complete graphs in [1] correspond precisely to agreement and disagreement regimes for more general networks.

The engineering literature on distributed opinion dynamics typically associates formation of opinions in continuous time to the spectral properties of the Laplacian matrix of the network graph. In linear consensus protocols, the governing equations take on the structure of the graph Laplacian, and consensus is achieved as opinions converge to its kernel [3]. Analogous distributed Laplacian schemes are considered with antagonistic (signed) interconnections in, e.g., [4]–[6]. When graphs are structurally balanced, i.e., when the signed graph Laplacian has a zero eigenvalue [4], such schemes give rise to clustered disagreement on the network. A nonlinear model of distributed consensus formation with saturated network interactions studied in [7]–[11] also relies on Laplacian-like structure of the governing equations, with each agent weighting its opinion state based on its in-degree on the network graph. Linearization of this model about the unopinionated state yields linear Laplacian consensus dynamics.

We study opinion dynamics of homogeneous agents with governing equations that do not necessarily have a Laplacian structure. We show how spectral properties of the *adjacency matrix* of the underlying graph, rather than those of its Laplacian, play a key role in characterizing the opinion formation process. In [8] the spectral properties of the adjacency matrix are tied to a necessary and sufficient condition for existence of multi-stable consensus equilibria. However, unlike our model, the model in [8] relies on a Laplacian structure, and so the dynamics do not inherit the full symmetry properties of the network graph. For our model, we show that the largest and smallest eigenvalues of the adjacency matrix determine if nonzero opinions form, and the associated eigenspaces select the sign and relative magnitude of resultant opinions on the network. With the addition of dynamics on an attention parameter, the eigenvectors also determine which agents in

Supported by NSF grant CMMI-1635056, ONR grant N00014-19-1-2556, ARO grant W911NF-18-1-0325, DGAPA-UNAM PAPIIT grant IN102420, and Conacyt grant A1-S-10610, and NSF Grad. Res. Fell. DGE-2039656.

A. Bizyaeva, A. Matthews and N.E. Leonard are with the Department of Mechanical and Aerospace Engineering, Princeton University, Princeton, NJ, 08544 USA; e-mail: bizyaeva@princeton.edu, naomi@princeton.edu.

A. Franci is with the National Autonomous University of Mexico, 04510 Mexico City, Mexico. e-mail: afranci@ciencias.unam.mx

the group are maximally sensitive to inputs. We illustrate how this can be used to trigger opinion cascades.

In Section II we present the opinion dynamics model of [1] on a network in the case of homogeneous agents and two options. In Section III we prove that agreement and disagreement equilibria, of which consensus and dissensus are special cases, arise on networks as bifurcations from an unopinionated equilibrium. We prove relationships between the magnitude of an agent's equilibrium opinion and its centrality on the network in Sections IV, and to the symmetry properties of the network graph in Section V. In Section VI we introduce attention feedback dynamics and demonstrate how the group's sensitivity to input relates to the adjacency eigenvectors. We conclude in Section VII.

## II. OPINION DYNAMICS MODEL

We study a nonlinear model of  $N_a$  homogeneous agents forming opinions about two options, a specialization of the general heterogeneous multi-agent, multi-option, model introduced in [1], [2]. The *opinion of each agent  $i$*  is captured by a real-valued variable  $x_i \in \mathbb{R}$ . When  $x_i = 0$  agent  $i$  has a *neutral opinion*, and when  $x_i > 0$  ( $< 0$ ) agent  $i$  *favors* option A (option B). A greater magnitude  $|x_i|$  corresponds to a stronger commitment of agent  $i$  to one of the options. We define a threshold  $\vartheta > 0$  to formalize what is meant by "close"; the value of  $\vartheta$  can be chosen for purposes of interpretation. We call an agent *opinionated* when  $|x_i| > \vartheta > 0$ , and *unopinionated* otherwise.

The *directed* graph  $\bar{G} = (V, \bar{E})$  encodes which agents can communicate with which other agents. The set of vertices  $V = \{1, \dots, N_a\}$  represents the set of  $N_a$  agents, and edges  $\bar{E} \subseteq V \times V$  represent interactions between agents. If edge  $\bar{e}_{ik} \in \bar{E}$ , then agent  $k$  is a *neighbor* of agent  $i$ . When inter-agent communication is bidirectional,  $G$  is *undirected*: if  $\bar{e}_{ik} \in \bar{E}$ , then  $\bar{e}_{ki} \in \bar{E}$ . The graph adjacency matrix  $\bar{A} \in \mathbb{R}^{N_a \times N_a}$  encodes interaction weights with element  $\bar{a}_{ik} \neq 0$  if and only if  $\bar{e}_{ik} \in \bar{E}$ . The sign of inter-agent weights determines whether agents are *cooperative* ( $\bar{a}_{ik} > 0$ ) or *competitive* ( $\bar{a}_{ik} < 0$ ). In this paper we specialize  $\bar{a}_{ik} \in \{0, 1\}$ .

Let  $\mathbf{x} = (x_1, \dots, x_{N_a}) \in \mathbb{R}^{N_a}$  be the *opinion state of the group*. When  $\mathbf{x} = \mathbf{0}$ , the group is in the *neutral state*. When all agents are unopinionated, the group is in an *unopinionated state*. A pair of opinionated agents  $i, k$  *agree* if both share the same qualitative opinion state, i.e.,  $\text{sign}(x_i) = \text{sign}(x_k)$ . When all agents agree, the group is an *agreement state*. The *consensus state* is a special type of agreement state in which opinions are close in value, i.e.  $|x_i - x_k| < \vartheta$  for all  $i, k \in V$ . A pair of opinionated agents *disagrees* if each has a different qualitative opinion state. If at least one pair of agents disagrees, then the group is in a *disagreement state*. The *dissensus state* is a special type of disagreement state in which individual agents may be opinionated but the group is unopinionated on average, i.e.,  $|\sum_{i=1}^{N_a} x_i| < \vartheta N_a$ .

Each agent's opinion is updated in continuous time as a function of three key terms: a linear damping term, a nonlin-

ear network interaction term that includes self-reinforcement, and an additive input term:

$$\dot{x}_i = -dx_i + u_i S \left( \alpha x_i + \gamma \sum_{\substack{k=1 \\ k \neq i}}^{N_a} \bar{a}_{ik} x_k \right) + b_i := h_i(\mathbf{x}). \quad (1)$$

$S : \mathbb{R} \rightarrow \mathbb{R}$  is an odd saturating function satisfying  $S(0) = 0$ ,  $S'(0) = 1$ ,  $\text{sign}(S''(z)) = -\text{sign}(z)$ . The saturation function is applied to the sum of two terms. The first is a self-reinforcement term with weight  $\alpha \geq 0$ . The second is the sum of opinions of agent  $i$ 's neighbors with weight  $\gamma \in \mathbb{R}$ . That the network interactions are saturated using  $S$  in (1) means the opinion dynamics of agent  $i$  are proportionally sensitive to changes in these interactions when the corresponding opinion magnitudes are small, but that the influence of these interactions levels off when the corresponding opinion magnitudes become large. In all simulations we let  $S = \tanh$ ; however the results hold qualitatively for more general sigmoidal functions. For the general model, motivation, and relation to the broader literature see [1], [2].

The parameter  $d > 0$  is the *damping coefficient*. In the absence of network interactions, agent  $i$ 's opinion  $x_i$  will converge exponentially to  $b_i/d$  at a rate determined by  $d$ . The control parameter  $u_i > 0$  is the *attention* of agent  $i$  to network interactions. Attention  $u_i$  governs the influence of the saturated network interactions, relative to damping, on the opinion dynamics of agent  $i$ . When  $u_i$  is above a critical threshold that grows with  $d$ , the magnitude of  $x_i$  grows nonlinearly to a value much larger than  $b_i/d$ . Attention  $u_i$  can be fixed, time-varying, or defined to evolve according to state-dependent (closed-loop) dynamics. The *input*  $b_i \in \mathbb{R}$  is an external signal or internal bias that stimulates an agent in favor of option A (option B) when  $b_i > 0$  ( $< 0$ ). For most of this paper we focus on the zero-input case:  $b_i = 0$  for all  $i$ . We consider nonzero input in Section VI.

Let  $W(\lambda_i)$  be the generalized eigenspace of  $\bar{A}$  relative to its eigenvalue  $\lambda_i$ . Let  $\lambda_{max}$  and  $\lambda_{min}$  be the eigenvalues of  $\bar{A}$  with largest and smallest real parts. In the following lemma we state several useful properties of  $\lambda_{min}$ ,  $\lambda_{max}$ , and their associated eigenspaces.

**Lemma 1. A.** *When  $\bar{G}$  is strongly connected (directed),  $\lambda_{max} > 0$  is real with multiplicity 1, and for any nonzero vector  $\mathbf{v} = (v_1, \dots, v_{N_a}) \in W(\lambda_{max})$ ,  $v_i \neq 0$  and  $\text{sign}(v_i) = \text{sign}(v_k)$  for all  $i, k \in V$ ;*

**B.** *When  $G$  is connected (undirected),  $\lambda_{min} < 0$  is real and for any nonzero vector  $\mathbf{w} = (w_1, \dots, w_{N_a}) \in W(\lambda_{min})$ ,  $\text{sign}(w_i) = -\text{sign}(w_k)$  for at least one pair of  $i, k \in V$ .*

*Proof.* Observe that  $\text{tr } \bar{A} = 0$  so  $\text{Re}(\lambda_{min}) < 0$ . When  $\bar{G}$  is undirected,  $\bar{A}$  is symmetric and  $\lambda_{min}$  is real. Further,  $\bar{A}$  is a nonnegative matrix. Parts A and B follow from the Perron-Frobenius Theorem [12, Theorem 11].  $\square$

## III. AGREEMENT AND DISAGREEMENT

In this section we study opinion dynamics (1) with static  $u_i := u \geq 0$  for all  $i \in V$  and show how cooperative agents

( $\gamma > 0$ ) give rise to agreement, whereas competitive agents ( $\gamma < 0$ ) give rise to disagreement. In the following theorem, we expand upon the result in [1, Theorem IV.1] for two-option networks, and describe the steady-state solutions that arise from (1) in different parameter regimes.

**Theorem 1.** *The following hold true for (1) with  $u_i := u \geq 0$  and  $b_i = 0$  for all  $i = 1, \dots, N_a$ :*

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

$$u_a = \frac{d}{\alpha + \gamma \lambda_{max}}. \quad (2)$$

*At  $u = u_a$ , branches of agreement equilibria,  $x_i \neq 0$ ,  $\text{sign}(x_i) = \text{sign}(x_k)$  for all  $i, k \in V$ , emerge in a steady-state bifurcation off of  $\mathbf{x} = \mathbf{0}$  along  $W(\lambda_{max})$ ;*

**B. Competition leads to disagreement:** *Let  $\bar{G}$  be a connected undirected graph. If  $\gamma < 0$  the neutral state  $\mathbf{x} = \mathbf{0}$  is a locally exponentially stable equilibrium for  $0 < u < u_d$  and unstable for  $u > u_d$ , with*

$$u_d = \frac{d}{\alpha + \gamma \lambda_{min}}. \quad (3)$$

*At  $u = u_d$ , branches of disagreement equilibria,  $\text{sign}(x_i) = -\text{sign}(x_k)$  for at least one pair  $i, k \in V$ ,  $i \neq k$ , emerge in a steady-state bifurcation off of  $\mathbf{x} = \mathbf{0}$  along  $W(\lambda_{min})$ .*

*Proof.* The Jacobian of (1) evaluated at  $\mathbf{x} = \mathbf{0}$  is  $J(\mathbf{0}) = (u\alpha - d)\mathcal{I} + u\gamma\bar{A}$ , where  $\mathcal{I}$  is the identity matrix. The eigenvalues of  $J(\mathbf{0})$  are  $\mu_i = u(\alpha + \gamma\lambda_i) - d$  where  $\lambda_i$  is an eigenvalue of  $\bar{A}$ . When  $0 \leq u < \min_i \frac{d}{\alpha + \gamma \text{Re}(\lambda_i)}$ ,  $\text{Re}(\mu_i)$  is negative for all  $i \in V$  and  $\mathbf{x} = \mathbf{0}$  is locally exponentially stable. For values of  $u$  above this bound the origin is unstable. When  $\gamma > 0$  this bound  $u_a$  corresponds to  $\lambda_i = \lambda_{max}$  and is given by (2). When  $\gamma < 0$  the bound  $u_d$  corresponds to  $\lambda_i = \lambda_{min}$  and is given by (3). Thus, when  $\gamma > 0$  ( $\gamma < 0$ ) a steady-state bifurcation happens at  $u = u_a$  ( $u = u_d$ ) along  $W(\lambda_{max})$  ( $W(\lambda_{min})$ ). The rest follows from Lemma 1.  $\square$

**Remark 1.** *Due to space constraints, we defer to a future publication detailed analysis of stability of the agreement and disagreement equilibria emerging at agreement and disagreement opinion-forming bifurcations. For one-dimensional kernels, we expect these equilibria to be generically stable, as easily verifiable using center-manifold reduction arguments.*

A main takeaway of Theorem 1 is that the spectral properties of  $\bar{A}$  inform the opinion formation outcomes on the network. Characterizing the eigenvalues  $\lambda_{min}$ ,  $\lambda_{max}$  along with their associated eigenspaces  $W(\lambda_{min})$ ,  $W(\lambda_{max})$  is equivalent to characterizing the primary branches of opinionated steady-state solutions of (1) emerging at bifurcations from the neutral state. We use this approach to systematically classify the opinion patterns that arise for various networks with spectral properties that are well known or easily established. For larger and less structured networks these quantities can

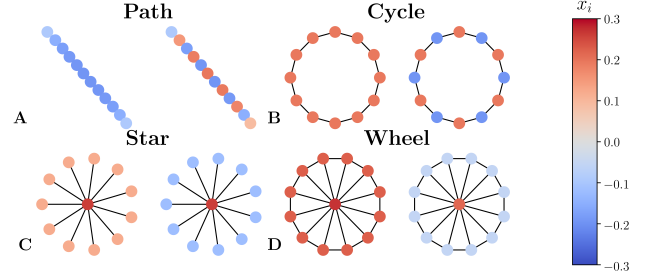


Fig. 1: Steady-state patterns of agreement with  $\gamma = 1.3$  (left) and disagreement with  $\gamma = -1.3$  (right) from simulation of opinion dynamics (1) for four undirected graph types. Node color represents opinion  $x_i$  at  $t = 500$ . All nodes have  $b_i = 0$  and randomized initial opinions drawn from  $U(-1, 1)$ . Parameters:  $d = 1$ ,  $\alpha = 1.2$ ,  $u = 0.31$  for path and cycle,  $u = 0.26$  for star and wheel.

be easily computed numerically. Figures 1 and 2 illustrate agreement and disagreement equilibria of (1) on a variety of graphs, with  $u$  slightly above the bifurcation point.

Network consensus and dissensus are special cases of agreement and disagreement described in Theorem 1. Let the *consensus space* be  $W_c := \text{span}\{\mathbf{1}\}$  and the *dissensus space* be  $W_d := \{\mathbf{x} \in \mathbb{R}^{N_a} : \sum_{i=1}^{N_a} x_i = 0\}$ . Observe that  $W_c$  and  $W_d$  are orthogonal complements in  $\mathbb{R}^{N_a}$ . Outside a  $\vartheta$ -neighborhood of the origin, the  $\vartheta$ -neighborhood of  $W_c$  is made of consensus solutions, whereas the  $\vartheta$ -neighborhood of  $W_d$  is made of dissensus solutions. We show next that for graphs in which all agents have the same number of neighbors, the agreement and disagreement equilibria of Theorem 1 correspond to consensus and dissensus.

**Definition 1.**  $\hat{G} = (\hat{V}, \hat{E})$  is a  $K$ -regular graph if every vertex  $i \in \hat{V}$  has exactly  $K$  neighbors.

**Lemma 2.** *If  $\bar{G}$  is undirected, connected, and  $K$ -regular, all vectors  $\mathbf{x} = (x_1, \dots, x_{N_a}) \in W(\lambda_{min})$  satisfy  $\sum_{i=1}^{N_a} x_i = 0$ .*

*Proof.* Observe that for a connected,  $K$ -regular graph,  $\lambda_{max} = K$  and  $W(\lambda_{max}) = W_c$ . Because  $\bar{A}$  is symmetric, all its generalized eigenspaces are orthogonal and, thus,  $W(\lambda_{min}) \subseteq W_d$ . The statement of the lemma follows.  $\square$

**Theorem 2 (Consensus and Dissensus).** *If  $\bar{G}$  is undirected, connected, and  $K$ -regular, the agreement bifurcations at  $u = u_a$  with  $\gamma > 0$  give rise to consensus solutions  $|x_i - x_k| < \vartheta$  for all  $i, k \in V$ , and the disagreement bifurcations at  $u = u_d$  with  $\gamma < 0$  give rise to dissensus solutions  $|\sum_{i=1}^{N_a} x_i| < \vartheta N_a$ .*

*Proof.* From Theorem 1 and Lemma 2, an opinion-forming bifurcation emerges along the consensus space for  $\gamma > 0$  and dissensus space for  $\gamma < 0$ . This means that the resulting equilibrium solutions are arbitrarily close to the consensus (dissensus) space as  $u \searrow u_a$  ( $u \searrow u_d$ ), where the notation  $\searrow$  signifies approaching in value from above.  $\square$

Figure 1B and Figure 2A,B are examples of consensus and dissensus on 2-regular graphs (a cycle with an even and odd

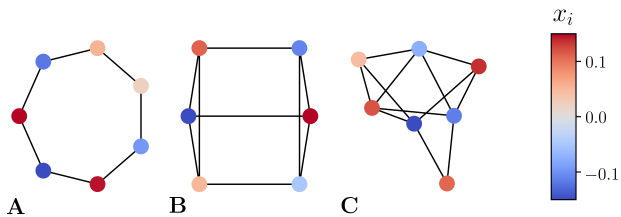


Fig. 2: Disagreement patterns on odd cycle (A), 3-regular (B) and randomly generated (C) graphs. Parameters are  $d = 1$ ,  $\alpha = 0.5$ ,  $\gamma = -0.5$ ,  $u = u_d + 0.01$ . All else is as in Fig. 1.

number of nodes) and a 3-regular graph.

#### IV. AGENT CENTRALITY

In this section we examine how equilibrium opinions of agents depend on their location in the graph. We show that at an opinion-forming bifurcation, an agent's opinion strength is often determined by its relative location in the network as quantified by a suitable centrality measure.

A *centrality measure* ranks how central each node is in a network, i.e., measures its influence over some emergent network property. We recall the definition of a well known network centrality measure originally proposed in [13]:

**Definition 2** (Eigenvector Centrality). *The entries of the normalized positive left eigenvector corresponding to the eigenvalue  $\lambda_{max}$  of  $\hat{A}$  for a (directed or undirected) graph  $\hat{G}$  provide a centrality measure for the nodes of the graph.*

It is shown in [14] that eigenvector centrality, deriving from the adjacency matrix and not the Laplacian, is particularly useful for graphs on which agents with high in-degree have many neighbors of low in-degree, and is related to several other common graph centrality measures. Let  $\mathbf{v}^c = (v_1^c, \dots, v_{N_a}^c)$  be the centrality eigenvector for  $\bar{A}$  from Definition 2. For undirected graphs, this centrality vector determines the opinion strength of each agent at an agreement equilibrium predicted by Theorem 1A, i.e., the larger an agent's centrality, the stronger its opinion at agreement.

**Theorem 3.** *Consider opinion dynamics (1) with undirected, connected  $\bar{G}$ ,  $u_i := u \geq 0$ ,  $b_i = 0$ , and  $\alpha > 0$ . Agreement equilibria  $\mathbf{x} = (x_1, \dots, x_{N_a})$  described in Theorem 1A satisfy  $|x_i| < |x_k|$  if  $v_i^c < v_k^c$  and  $|x_i| = |x_k|$  if  $v_i^c = v_k^c$  for all  $i, k = 1, \dots, N_a$ .*

*Proof.* Let  $\mathbf{v} = (v_1, \dots, v_{N_a})$  be the normalized right  $\lambda_{max}$ -eigenvector of  $\bar{A}$ . By symmetry of  $\bar{A}$ ,  $v_i = v_i^c$  for all  $i = 1, \dots, N_a$  and the theorem follows from Theorem 1.A.  $\square$

We next state an analogous result for disagreement equilibria on a common class of graphs called *bipartite graphs*.

**Definition 3.** *Undirected  $\hat{G} = (\hat{V}, \hat{E})$  is a bipartite graph if  $\hat{V}$  can be subdivided into disjoint subsets  $\hat{V}_1, \hat{V}_2$  such that every edge  $\hat{e}_{ik} \in \hat{E}$  connects a vertex in  $\hat{V}_1$  to one in  $\hat{V}_2$ .*

In the following, we show that for disagreement on bipartite graphs, an agent's partition membership determines the sign

of its equilibrium opinion and an agent's centrality in the network determines the strength of its equilibrium opinion.

**Lemma 3** ( $\lambda_{min}$ -Eigenvector of Bipartite Graph). *Suppose  $\bar{G}$  is a bipartite graph and let  $\hat{V}_1, \hat{V}_2$  be the two vertex subsets of  $V$  from Definition 3. Let  $\mathbf{w} = (w_1, \dots, w_{N_a})$  be the eigenvector corresponding to  $\lambda_{min}$  of  $\bar{A}$ . Then  $w_i, w_k \neq 0$  and  $\text{sign}(w_i) = -\text{sign}(w_k)$  for all  $i \in \hat{V}_1, k \in \hat{V}_2$ . Moreover,  $|w_i| = v_i^c$  for all  $i = 1, \dots, N_a$ .*

*Proof.* For all bipartite graphs,  $\dim W(\lambda_{min}) = 1$ . By symmetry of  $\bar{A}$ , the  $\lambda_{max}$ -eigenvector is equal to  $\mathbf{v}^c$ . The lemma then follows from [15, Theorem 1.2], which states that the terms of the  $\lambda_{min}$ -eigenvector are equal in magnitude to the terms of the  $\lambda_{max}$ -eigenvector, with the sign structure reflecting the bipartition.  $\square$

**Theorem 4** (Disagreement Opinion Strength Reflects Agent Centrality on Bipartite Graphs). *Consider (1) with undirected bipartite  $\bar{G}$ ,  $u_i := u \geq 0$ ,  $b_i = 0$ ,  $\alpha > 0$ , and  $\gamma < 0$ . Let  $\hat{V}_1, \hat{V}_2$  be the two subsets of  $V$  from Definition 3. Disagreement equilibria  $\mathbf{x} = (x_1, \dots, x_{N_a})$  described in Theorem 1.B satisfy  $|x_i| < |x_k|$  if  $v_i^c < v_k^c$  and  $|x_i| = |x_k|$  if  $v_i^c = v_k^c$  for all  $i, k = 1, \dots, N_a$ . Moreover,  $\text{sign}(x_i) = -\text{sign}(x_k)$  for all  $i \in \hat{V}_1, k \in \hat{V}_2$ .*

*Proof.* This follows by Theorem 1B and Lemma 3.  $\square$

All graphs shown in Figure 1 have a simple  $\lambda_{min}$ . Further, the cycle, path, and star are bipartite graphs and the sign distribution of nodes across options reflects their bipartition. Observe that the magnitude of opinions reflects relative centrality on all the shown graphs in both the agreement and disagreement parameter regimes, including the wheel which is not bipartite. This suggests that the  $\lambda_{min}$ -eigenvector can sometimes be related to a notion of agent centrality even when it does not precisely equal  $\mathbf{v}^c$ . In contrast, graphs in Figure 2 are not bipartite, and graphs A, B have a two-dimensional  $W(\lambda_{min})$ . The distribution of disagreement opinions on these graphs is more heterogeneous in magnitude, despite the first two graphs being regular and all agents being equally central. Additional equilibrium distributions of opinions are possible in the disagreement regime.

#### V. GRAPH SYMMETRY

In this section we relate the patterns of emergent opinions on solution branches described in Theorem 1 to the symmetry of the underlying graph. We first state several key definitions.

Recall that a *graph automorphism* of  $\bar{G}$  is a permutation of vertices that preserves adjacency (i.e., maps edges to edges and non-edges to non-edges). The *automorphism group*  $\text{Aut}(\bar{G}) := \Gamma$  is the set of all graph automorphisms of  $\bar{G}$ . Define the group of permutations of a set of  $n$  symbols  $\{1, \dots, n\}$  as  $S_n$ . Then for graph  $\bar{G}$  with  $N_a$  agent vertices,  $\Gamma \subseteq S_{N_a}$ . The graph automorphism group of  $\bar{G}$  can also be interpreted as the group of permutation matrices which commute with its adjacency matrix  $\bar{A}$ . Each element  $\rho \in \Gamma$  can be represented by a permutation matrix  $P_\rho$  which satisfies

$P_\rho \bar{A} = \bar{A} P_\rho$  [16, Proposition 3.8.1]. We commonly refer to  $\rho$  as a *symmetry* of  $\bar{G}$  and  $\Gamma$  as its *symmetry group*.

A connected notion is the *equivariance* of a dynamical system with respect to a symmetry group. Consider the opinion dynamics (1) as a dynamical system  $\dot{\mathbf{x}} = \mathbf{h}(\mathbf{x})$ , where the map  $\mathbf{h} : \mathbb{R}^{N_a} \rightarrow \mathbb{R}^{N_a}$  is  $\mathbf{h}(\mathbf{x}) = (h_1(\mathbf{x}), \dots, h_{N_a}(\mathbf{x}))$ . Let  $\Sigma$  be a compact Lie group with elements  $\sigma$  that act on  $\mathbb{R}^{N_a}$ . Then  $\mathbf{h}$  is  $\sigma$ -*equivariant* for some  $\sigma \in \Sigma$  if  $\sigma \mathbf{h}(\mathbf{x}) = \mathbf{h}(\sigma \mathbf{x})$ , and  $\mathbf{h}$  is  $\Sigma$ -*equivariant* if this holds true for all  $\sigma \in \Sigma$  [17, Definition 1.7]. In the following theorem we show that the symmetry group of the graph  $\bar{G}$  is also a symmetry group of the opinion dynamics (1) with zero input.

**Theorem 5** ( $\Gamma$ -equivariance). *Consider opinion dynamics (1) with  $b_i = 0$  for  $i = 1, \dots, N_a$ . Let  $\Gamma$  be the automorphism group of  $\bar{G}$ . The opinion dynamics are  $\Gamma$ -equivariant.*

*Proof.* Let  $\rho$  be an element of  $\Gamma$ , and define the function  $\hat{\mathbf{S}} : \mathbb{R}^{N_a} \rightarrow \mathbb{R}^{N_a}$  as  $\hat{\mathbf{S}}(\mathbf{y}) = (S(y_1), \dots, S(y_{N_a}))$ . Every permutation  $\rho$  can be decomposed into a product of transpositions  $\rho_{ik}$  that interchange elements in positions  $i$  and  $k$  of a set. For an arbitrary transposition, it is easy to see that  $P_{\rho_{ik}} \hat{\mathbf{S}}(\mathbf{y}) = \hat{\mathbf{S}}(P_{\rho_{ik}} \mathbf{y})$ , and by iteratively applying transpositions we see that for all permutations  $\rho$ ,  $P_\rho \hat{\mathbf{S}}(\mathbf{y}) = \hat{\mathbf{S}}(P_\rho \mathbf{y})$ . Using this and the definition of graph automorphism we get

$$\begin{aligned} P_\rho \mathbf{h}(\mathbf{x}) &= -dP_\rho \mathbf{x} + u \hat{\mathbf{S}}(P_\rho(\alpha \mathcal{I} + \gamma \bar{A}) \mathbf{x}) \\ &= -dP_\rho \mathbf{x} + u \hat{\mathbf{S}}((\alpha \mathcal{I} + \gamma \bar{A}) P_\rho \mathbf{x}) = \mathbf{h}(P_\rho \mathbf{x}). \end{aligned}$$

This holds for all  $\rho \in \Gamma$ , and the theorem follows.  $\square$

In the remainder of this section we illustrate how graph symmetry dictates patterns of opinions on the network.

**Definition 4.** *Let  $\Gamma$  be the automorphism group of  $\bar{G}$ , and  $i \in V$  a vertex. An orbit of  $i$  is  $O_i = \{k \in V \mid k = \rho i \text{ for some } \rho \in \Gamma\}$ . The orbits are equivalence classes that partition  $V$  through the equivalence relation*

$$i \sim k \text{ if } k = \rho i \text{ for some } \rho \in \Gamma. \quad (4)$$

**Theorem 6.** *Consider opinion dynamics (1) with  $u_i := u \geq 0$  and  $b_i = 0$  for all  $i \in V$ . Let  $\Gamma$  be the automorphism group of the undirected graph  $\bar{G}$ , and for any two vertices  $i, k \in V$  define the equivalence relation  $i \sim k$  as in (4).*

**A.** *Suppose  $\gamma > 0$ . For the agreement equilibria  $\mathbf{x} = (x_1, \dots, x_{N_a})$  from Theorem 1.A, if  $i \sim k$ , then  $x_i = x_k$ .*

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

*Proof.* In A and B, the solutions  $\mathbf{x}$  appear along the 1-dimensional subspace  $\ker J(\mathbf{0})$ . For any  $\rho \in \Gamma$  and  $\mathbf{v} \in \ker J(\mathbf{0})$ ,  $P_\rho \mathbf{v} \in \ker J(\mathbf{0})$  [17, Remark 1.25], and  $P_\rho \mathbf{v} = \pm \mathbf{v}$  [16, Lemma 3.8.2]. The only way for this to be true is if for any  $i, k \in V$  for which  $i \sim k$ ,  $|v_i| = |v_k|$ .  $\square$

All graphs shown in Figure 1 have a nontrivial symmetry and a simple  $\lambda_{min}$ . The vertices contained in the same orbit

under the action of its symmetry have equal magnitude opinion in both agreement and disagreement parameter regimes (e.g. all the vertices of a cycle, the outer vertices of the star and wheel). The randomly generated graph in Figure 2C also has a simple  $\lambda_{min}$  but only a trivial symmetry. Each node is its own orbit and the equilibrium opinion magnitude is different at each node. In contrast, both graphs in Figure 2A,B have a nontrivial symmetry but the dimension of  $W(\lambda_{min})$  is 2 and the conditions of Theorem 5 are not met. The disagreement opinion magnitudes on these graphs do not reflect the orbit structure of the graph's symmetry group.

## VI. AGENT SENSITIVITY AND OPINION CASCADES

In [1, Section VI] a state feedback law is introduced for the attention parameter  $u_i$  in (1), which enables the group to become opinionated in response to input  $b_i$ . Each agent's attention parameter  $u_i$  tracks a saturated norm of the system state observed by agent  $i$ , with dynamics defined by

$$\tau_s \dot{u}_i = -u_i + S_u \left( x_i^2 + \sum_{k=1}^{N_a} (\bar{a}_{ik} x_k)^2 \right). \quad (5)$$

Here,  $\tau_s > 0$  is the time scale of the integration and  $S_u(y) = u_f(F(g(y - y_m)) - F(-gy_m))$ , with  $F(x) = \frac{1}{1+e^{-x}}$ . Design parameters  $g, u_f, y_m > 0$  tune the system response. The positive feedback between opinion strength and attention defined by dynamics (5) provides the resulting opinion-attention dynamics with a threshold for the input strength needed to let an originally weakly opinionated ( $|x_i| \ll \vartheta$ , for all  $i = 1, \dots, N_a$ ), weakly attentive ( $|u_i| \ll \vartheta$ , for all  $i = 1, \dots, N_a$ ) network transition to a strongly opinionated ( $|x_i| \gg \vartheta$ , for all  $i = 1, \dots, N_a$ ), strongly attentive ( $|u_i| \gg \vartheta$ , for all  $i = 1, \dots, N_a$ ) state [18].

Such an *opinion cascade* is illustrated in Figures 3 and 4 top left. An opinion cascade is of agreement (disagreement) type if the final opinion configuration is of agreement (disagreement) type. A rigorous analysis of opinion cascade over networks is the subject of ongoing work based on unfolding theory [19] of agreement and disagreement bifurcations. Here, we illustrate with numerical examples. We show in particular how the spectral properties of the adjacency matrix predict which agents are more likely to trigger a cascade when excited by an input.

We conjecture that the norm of the projection of the vector of inputs  $\mathbf{b}$  onto the eigenspaces  $W(\lambda_{max})$  or  $W(\lambda_{min})$ , depending on the cooperative or competitive regime, is directly related to whether or not the opinion cascade gets triggered. To illustrate this, in Figures 3 and 4 we show two examples of a disagreement cascade for graphs with a simple  $\lambda_{min}$ . In Figure 3 the graph is a balanced tree, which is bipartite. Thus, the magnitude of each entry of its  $\lambda_{min}$ -eigenvector corresponds to the corresponding agent's relative centrality, as proved in Lemma 3. When an input is given to the most central agent an opinion cascade gets triggered, whereas when the same input is given to an agent located on an outer leaf, the graph remains neutral.

Analogously, in Figure 4 a cascade is triggered on a randomly generated network when an input is given to the agent with the largest magnitude entry in the  $\lambda_{min}$ -eigenvector, and is not triggered when the same input is given to the agent with the smallest entry. The post-cascade pattern of opinions on this network is very close to the pattern for the solution of the zero-input system shown in Figure 2C. In this way, the spectral properties of  $\bar{A}$  can be used not only to predict the patterns of opinion formation for opinion dynamics (1) both with a static  $u_i = u$  and coupled with (5), but also to inform which agent should receive a control signal or carry a sensor in order to have maximal likelihood of an informed opinion cascade spreading on the network.

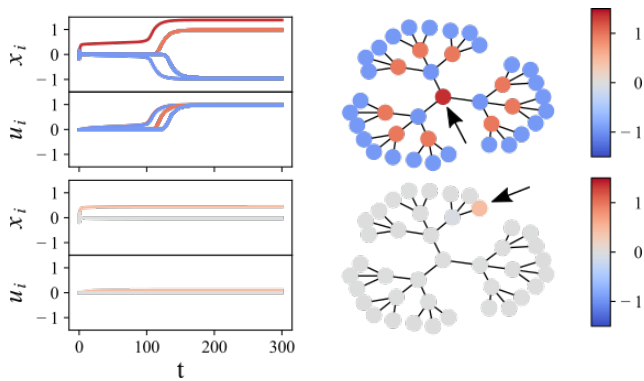


Fig. 3: Triggering a cascade on the balanced tree. Input  $b_i = 0.4$  for agent marked with arrow;  $b_i = 0$  for all other agents. Simulations of (1) start from small random initial opinions drawn from  $\mathcal{N}(0, 0.1)$ . Left: time trajectories of the dynamics. Right: final opinion, indicated by color, of simulation at  $t = 300$ . Parameters:  $d = 1$ ,  $\alpha = 1$ ,  $\gamma = -1$ ,  $u_i(0) = 0$  for all  $i \in V$ ,  $g = 10$ ,  $y_m = 0.4$ ,  $u_f = 1$ ,  $\tau_s = 10$ . Only in the top plots does the model undergo a disagreement opinion cascade, with  $\vartheta = 0.5$ .

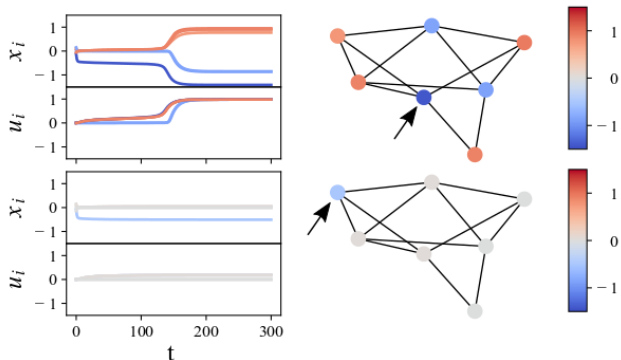


Fig. 4: Triggering a cascade on a randomly generated graph. Input  $b_i = -0.45$  for agent marked with arrow;  $b_i = 0$  for all other agents. Simulations of (1) start from small random initial opinions drawn from  $\mathcal{N}(0, 0.1)$ . Left: time trajectories of the dynamics. Right: final opinion, indicated by color, of simulation at  $t = 300$ . Parameters:  $d = 1$ ,  $\alpha = 0.5$ ,  $\gamma = -0.5$ ,  $u_i(0) = 0$  for all  $i \in V$ ,  $g = 10$ ,  $y_m = 0.4$ ,  $u_f = 1$ ,  $\tau_s = 10$ . Only in the top plots does the model undergo a disagreement opinion cascade, with  $\vartheta = 0.5$ .

## VII. DISCUSSION

We have shown how patterns of nonlinear opinion formation on two-option networks can be studied systematically using spectral properties of the graph adjacency matrix. We proved that agent centrality and symmetry of the underlying graph select the pattern of opinions in agreement and disagreement for a large set of networks. We illustrated how the eigenvectors of the adjacency matrix inform sensitivity of individual nodes to input that triggers opinion cascades; proving the influence of distributed input on opinion cascades is the subject of ongoing work. Other future work includes the natural generalization to multi-option networks with opinion dynamics defined in [1].

## REFERENCES

- [1] A. Bizyaeva, A. Franci, and N. E. Leonard, "A general model of opinion dynamics with tunable sensitivity," *arXiv:2009.04332 [math.OC]*, Sep. 2020.
- [2] A. Franci, M. Golubitsky, A. Bizyaeva, and N. E. Leonard, "A model-independent theory of consensus and dissensus decision making," *arXiv:1909.05765 [math.OC]*, Sep. 2020.
- [3] R. Olfati-Saber and R. M. Murray, "Consensus problems in networks of agents with switching topology and time-delays," *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1520–1533, 2004.
- [4] C. Altafani, "Consensus problems on networks with antagonistic interactions," *IEEE Trans. Aut. Ctrl.*, vol. 58, no. 4, pp. 935–946, 2013.
- [5] J. Liu, X. Chen, T. Başar, and M. A. Belabbas, "Exponential convergence of the discrete- and continuous-time altafani models," *IEEE Trans. Automatic Control*, vol. 62, no. 12, pp. 6168–6182, 2017.
- [6] G. Shi, C. Altafani, and J. S. Baras, "Dynamics over signed networks," *SIAM Review*, vol. 61, no. 2, pp. 229–257, 2019.
- [7] A. Franci, V. Srivastava, and N. E. Leonard, "A realization theory for bio-inspired collective decision making," *arXiv:1503.08526 [math.OC]*, Mar. 2015.
- [8] A. Fontan and C. Altafani, "Multiequilibria analysis for a class of collective decision-making networked systems," *IEEE Transactions on Control of Network Systems*, vol. 5, no. 4, pp. 1931–1940, 2017.
- [9] R. Gray, A. Franci, V. Srivastava, and N. E. Leonard, "Multiagent decision-making dynamics inspired by honeybees," *IEEE Transactions on Control of Network Systems*, vol. 5, no. 2, pp. 793–806, 2018.
- [10] P. U. Abara, F. Ticozzi, and C. Altafani, "Spectral conditions for stability and stabilization of positive equilibria for a class of nonlinear cooperative systems," *IEEE Transactions on Automatic Control*, vol. 63, no. 2, pp. 402–417, 2018.
- [11] A. Fontan and C. Altafani, "Achieving a decision in antagonistic multi agent networks: frustration determines commitment strength\*," in *IEEE Conference on Decision and Control*, 2018, pp. 109–114.
- [12] L. Farina and S. Rinaldi, *Positive linear systems: theory and applications*. John Wiley & Sons, 2011, vol. 50.
- [13] P. Bonacich, "Factoring and weighting approaches to status scores and clique identification," *J. Math. Sociol.*, vol. 2, no. 1, pp. 113–120, 1972.
- [14] —, "Some unique properties of eigenvector centrality," *Social Networks*, vol. 29, no. 4, pp. 555–564, 2007.
- [15] D. Stevanovic, *Spectral radius of graphs*. Academic Press, 2014.
- [16] D. Cvetkovic, S. Simic, and P. Rowlinson, *An Introduction to the Theory of Graph Spectra*. Cambridge University Press, 2009.
- [17] M. Golubitsky and I. Stewart, *The Symmetry Perspective*, 1st ed., ser. Progress in Mathematics. Birkhäuser Basel, 2002, vol. 200.
- [18] A. Franci, A. Bizyaeva, S. Park, and N. Leonard, "Analysis and control of agreement and disagreement opinion cascades," *arXiv*, 2021.
- [19] M. Golubitsky and D. G. Schaeffer, *Singularities and Groups in Bifurcation Theory*, ser. Applied Mathematical Sciences. New York, NY: Springer-Verlag, 1985, vol. 51.