

# A biased assimilation model on signed graphs

Lingfei Wang, Hong Yiguang, Shi Guodong and Claudio Altafini

## Conference Paper

Cite this paper as:

Wang, L., Yiguang, H., Guodong, S., Altafini, C. A biased assimilation model on signed graphs, In, 2020 59th IEEE Conference on Decision and Control (CDC): IEEE; 2020, ISBN: 978-1-7281-7447-1

DOI: <https://doi.org/10.1109/CDC42340.2020.9304000>

The self-archived postprint version of this conference paper is available at Linköping University Institutional Repository (DiVA):

<http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-179271>

Original publication available at:

<https://doi.org/10.1109/CDC42340.2020.9304000>

Copyright: IEEE

<http://www.ieee.org/>

©2020 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.



# A biased assimilation model on signed graphs

Lingfei Wang, Yiguang Hong, Guodong Shi, Claudio Altafini

**Abstract**—This work introduces antagonistic interactions into the so-called biased assimilation model of opinion dynamics, a nonlinear model expressing the bias of the agents towards their own opinions. In this model, opinions exchanged in a signed network are multiplied by a state dependent term having the bias as exponent. For small values of the bias parameters, while for structurally balanced networks polarization always occurs, we show that for structurally unbalanced networks also a state of indecision (corresponding to the centroid of the opinion hypercube) is a local attractor. When instead the biases are all large, the opinions usually polarize. In particular, for large enough biases if all agents take the same initial stance (i.e., are all on the same side of the indecision state), then the opinions polarize all to the same extreme value for both structurally balanced network and structurally unbalanced network.

## I. INTRODUCTION

Social opinion dynamics explores the formation of opinions in communities of individuals interacting over a social network, i.e., a network in which nodes represent individuals and state variables associated to the nodes their opinions [1], [2], [3], [4]. In one basic social opinion dynamics model, called the DeGroot model, each node uses a convex combination of its neighbors' states to update its own state, and consensus among all these nodes is asymptotically reached under some simple assumption [5]. Many variants of the DeGroot model have been proposed recently, including some that consider the special influence exerted by the nodes' own opinions during the opinion evolution [6], [7], [8], [10]. For example, in the Friedkin-Johnson model, each node incorporates its initial condition in the terms of the convex combination when updating the state [6]; in the Hegselmann-Krause model, each node keeps a confidence bound, and only interacts with the nodes that hold opinions closer than that bound from its own opinion [7]; there are also some computational models describing echo chambers, i.e., how an individual's opinion reinforces itself due to repeated interactions with like-minded individuals [8], [9]. These models express, more or less, an important character of opinion dynamics – bias.

\*Work supported in part by grants from the National Natural Science Foundation of China (Grant 61733018 to Y.H.), by the Australian Research Council (Grant DP190103615 to G.S.) and by the Swedish Research Council (Grant 2015-04390 to C.A.).

L. Wang and Y. Hong are with the Key Laboratory of Systems and Control, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, and also with the University of Chinese Academy of Sciences, Beijing 100190, China. E-mail: wlf@amss.ac.cn, yghong@iss.ac.cn

G. Shi is with the Australian Center for Field Robotics, School of Aerospace, Mechanical and Mechatronic Engineering, The University of Sydney, NSW 2008, Sydney. E-mail: guodong.shi@sydney.edu.au

C. Altafini is with the Division of Automatic Control, Department of Electrical Engineering, Linköping University, SE-58183 Linköping, Sweden, E-mail: claudio.altafini@liu.se

Biased assimilation is a common cognitive phenomenon that exists widely in real life [11], [12], [13], [14]. It characterizes the fact that people are more likely to accept supportive opinions but scrutinize opposite ones more cautiously. Biased assimilation has received some attention as a topic of opinion dynamics in recent years [10], [22], [27]. In [10], the authors proposed a model which extends the DeGroot update rule by multiplying the weighted sum of the neighbors' opinions with a own-state dependent factor. The state dependent term is characterized by an exponent called the bias: the greater the bias, the more importance an individual gives to his/her opinion when updating it. This model, denoted biased assimilation model in the following, is investigated in [10], [22], [27], [28]. According to these papers, the biased assimilation model generally produces opinion polarization, i.e., the state vector asymptotically converges to one vertex of the unit hypercube (the opinion domain). An intermediate equilibrium point (the centroid of the unit hypercube) which always exist in these models and which can be considered as a “state of indecision” is in fact unstable. In [28] it is shown that this intermediate equilibrium can become locally stable for negative bias, but negative bias is unrealistic because it means that each individual is mistrusting his/her own opinion.

Another limitation of the biased assimilation model of [10] is that it only considers cooperative interactions among the individuals. However, interactions among individuals cannot be always considered of cooperative nature, especially when it comes to exchanging opinions. Non-cooperative (or antagonistic) interactions are usually characterized as negative edges on a network, which is then called a signed social network [15], [16]. Several recent studies have investigated social opinion dynamics over signed networks, showing how the presence of antagonistic interactions can alter both a consensus and a collective decision process, see [21], [18], [19], [17], [20].

In this paper, we extend the model of [10] to negative interactions, and study its behavior over signed networks. Due to the high degree of non-linearity, a complete analysis of the reformulated model is extremely hard. Nevertheless, we are still able to characterize the most salient aspects of the dynamics, and for general network topology. We prove that if all the nodes are endowed with a small bias, the opinions polarize to a vertex of the unit hypercube for structurally balanced networks, while all the opinions converge to the indecision state (centroid of the unit hypercube) for structurally unbalanced networks. We also prove that if all nodes have large biases and hold the same stance initially (i.e., the initial conditions are on the same side of the indecision

state), then all opinions polarize to the same side, regardless of structural balance or unbalance. For all the asymptotic results we give quantitative bounds that the individual bias parameters of the agents must obey.

The rest of the paper is organized as follows: some preliminary concepts and the model formulation are introduced in Section II; the main results of asymptotic behavior for different ranges of the bias parameter are discussed in Section III; Section IV contains some of the proofs of our main results (other proofs had to be omitted for lack of space and will appear in an extended journal version of this paper).

*Notations.* The real number, integer and complex number sets are denoted by  $\mathbb{R}$  and  $\mathbb{Z}$ ,  $\mathbb{C}$ , respectively. In general, real numbers and imaginary numbers are both denoted by lowercase letters  $x, y, a, b, \dots$  and lowercase Greek letters  $\alpha, \beta, \dots$ . All vectors are real column vectors denoted by bold lowercase letters  $\mathbf{x}, \mathbf{y}, \mathbf{z}, \dots$ . The  $i$ -th entry of a vector  $\mathbf{x}$  is denoted by  $[\mathbf{x}]_i$ . Matrices are denoted by upper case letters such as  $A, B, P, Q, \dots$ . All matrices are real. The identity matrix is denoted by  $I$ , with dimension depending on the context. Given a matrix  $A$ ,  $A^\top$  denotes its transpose and  $A^k$  denotes the  $k$ th power of  $A$  when it is a square matrix. The  $(i, j)$ -th entry of a matrix  $A$  is denoted by  $A_{ij}$  or  $[A]_{ij}$ ; if  $A$  is symmetric, its largest (smallest) eigenvalue is denoted by  $\lambda_{\max}(A)$  ( $\lambda_{\min}(A)$ ). In the  $n$ -dimensional Euclidean space, we use  $\mathcal{B}(d)$  to denote the ball centered at the origin with radius  $d$ . The  $n$ -order vector, matrix and tensor with all entries being 0 or 1 are all written as  $\mathbf{0}$  or  $\mathbf{1}$ . The modulus of a complex number is denoted by  $|\cdot|$ . The Euclidean norm of a vector and 2-norm of a matrix are both denoted by  $\|\cdot\|$ . For any  $x \in \mathbb{R}$ , we write  $\text{sgn}(x)$  to denote the sign of  $x$  and let  $\text{sgn}(0) = 0$ .

## II. PROBLEM FORMULATION

### A. Signed Social Networks

Consider a network with  $n$  nodes indexed in the set  $V = \{1, \dots, n\}$ . The structure of the network is represented as a directed graph  $G = (V, E)$ , where an ordered pair  $(j, i) \in E$  denotes a link from node  $j$  to node  $i$  over the set  $V$ . Each link  $(j, i)$  in  $E$  is associated with a positive or negative sign, defining  $G$  as a signed graph. The positive and negative links are collected in the sets  $E^+$  and  $E^-$ , respectively. Then  $G^+ = (V, E^+)$  and  $G^- = (V, E^-)$  are, respectively, the positive and negative subgraphs of  $G$ . In this paper, we suppose  $(i, i) \notin E^-$  for all  $i \in V$ .

**Assumption 1** (i)  $G$  is strongly connected;  
(ii)  $G^+, G^-$  both contain at least one edge.

For a node  $i \in V$ , its positive neighbors are the nodes in the set  $N_i^+ := \{j : (j, i) \in E^+\} \setminus \{i\}$ . Similarly, the negative neighbor set of the node  $i$  is denoted by  $N_i^- = \{j : (j, i) \in E^-\}$ . The set  $N_i = N_i^+ \cup N_i^-$  then contains all nodes connecting to  $i$  over the graph  $G$ .

To each  $(j, i) \in E$  we associate a weight  $w_{ij} > 0$ . All these weights generate a  $n$ -order nonnegative weight matrix  $W$ , with  $W_{ij} = w_{ij}$  if  $(j, i) \in E$  and  $W_{ij} = 0$  otherwise.

Define  $d_i^+ := \sum_{j \in N_i^+} W_{ij}$  and  $d_i^- := \sum_{j \in N_i^-} W_{ij}$  as the total link weight from the positive and negative neighbors of node  $i$ , respectively. Let  $d_i := d_i^+ + d_i^-$  be the total link weight from all of  $i$ 's neighbors.

**Definition 1 (Structural Balance)** (i) A signed graph  $G$  is called **structurally balanced** if there is a partition of the node set into  $V = V_1 \cup V_2$  with  $V_1$  and  $V_2$  being nonempty and mutually disjoint, where any link between the two node subsets  $V_1$  and  $V_2$  is negative, and any link within each  $V_i$  is positive;  
(ii) A network which is not structurally balanced is said to be **structurally unbalanced**.

### B. The Opinion Dynamics Model

Time is slotted at  $t = 0, 1, \dots$ . Each node  $i$  holds a opinion  $x_i(t) \in [0, 1]$  at time  $t$  and interacts with its neighbors at each time to update its opinion. Define  $s_i^+(t) := \sum_{j \in N_i^+} W_{ij} x_j(t)$ , resp.  $s_i^-(t) := \sum_{j \in N_i^-} W_{ij} (1 - x_j(t))$ , as the weighted sum of the opinions of the positive, resp. negative, neighbors of  $i$ . Let  $s_i(t) := s_i^+(t) + s_i^-(t)$ . The opinion dynamics of  $x_i(t)$  is described by

$$x_i(t+1) = \frac{W_{ii}x_i(t) + x_i(t)^{b_i}s_i(t)}{W_{ii} + x_i(t)^{b_i}s_i(t) + (1 - x_i(t))^{b_i}(d_i - s_i(t))}, \quad (1)$$

where  $b_i \geq 0$  represents the (individual) bias of agent  $i$ .  $x_i(t)^{b_i}$  can be viewed as an additional factor by which node  $i$  weights  $x_j(t)$  or  $1 - x_j(t)$ , where  $j$  is one of  $i$ 's neighbors. When  $b_i = 0$  for all  $i \in V$ , under the transformation  $z_i(t) = x_i(t) - \frac{1}{2}$ ,  $i \in V$ , (1) will be identical to the discrete-time model of [21], in which all opinions range from  $-\frac{1}{2}$  to  $\frac{1}{2}$ .

The model (1) is a simple extension of the biased assimilation model proposed in [10], with the addition of negative interactions. It can be also considered a generalization of the model of [21]. In the original model of [21], all agents update their own opinions to the weighted average of their neighbors' opinions over signed networks. At every updating step of (1), each node  $i$  weights the weighted average of its neighbors' opinions by the state-dependent factor  $x_i(t)^{b_i}$ , and the opinion centroid is  $\frac{1}{2}$ . Moreover, a larger  $b_i$  corresponds to a more biased agent  $i$ .

The novel feature of (1) is that for any network  $G$  the indecision state (i.e., the opinion centroid) can become an attractor for suitable values of the biases. Such a feature can occur for the model [10] only for very special topologies (two-island networks, see [10]). It can occur also in the model [27], [28], but at the cost of letting the agent have a negative bias, which is a not very realistic assumption, as it implies that an agent is mistrusting his/her own opinion.

### C. Problems of interests

As mentioned above, our model (1) is a combination of two existing models. On one hand, biased assimilation has been shown in the literature to cause opinion polarization. The underlying networks are mostly assumed to be some special graphs such as the two-island network, the star network,

and so on. On the other hand, the model of [21] can generate bipartite consensus or convergence to the origin, respectively for structurally balanced network and structurally unbalanced network [21]. A natural problem that comes for our model is then the following: When the bias and the negative interactions coexist, will polarization occur for general network topology? If so, what kind of polarization? How will bias and negative interactions influence the evolution of opinions? For the sake of accuracy, we give our definition of polarization first.

**Definition 2 (Unilateral/Bipartite Polarization)** (i)

For opinion dynamics with  $x_i(t) \in [0, 1]$  for all  $i \in V$ , we say **unilateral polarization** occurs if  $\lim_{t \rightarrow \infty} x_i(t) = 0, i \in V$  or  $\lim_{t \rightarrow \infty} x_i(t) = 1, i \in V$ .

(ii) For opinion dynamics with  $x_i(t) \in [0, 1]$  for all  $i \in V$ , we say **bipartite polarization** occurs if there exists nonempty, mutually disjointed  $V_1, V_2$  such that  $V = V_1 \cup V_2$  and the following equations hold

$$\lim_{t \rightarrow \infty} x_i(t) = 0, i \in V_1; \lim_{t \rightarrow \infty} x_i(t) = 1, i \in V_2.$$

Note that Definition 2 is different from the definition of polarization in [10]. See also [28], where a similar definition of polarization is discussed in detail.

### III. MAIN RESULTS

In this section we present our main results. Before showing the results, we introduce some definitions and one more assumption.

Define  $W_G$  as the normalized matrix of  $W$ , i.e.,  $[W_G]_{ij} := \frac{W_{ij}}{W_{ii} + d_i}$  for all  $i, j \in V$ . Let  $\sigma_{ij}$  be the sign of link  $(j, i) \in E$ , and  $\sigma_{ij} = 0$  if  $(j, i) \notin E$ . Construct the signed weight matrix  $S_G$  corresponding to  $G$  as  $[S_G]_{ij} := \sigma_{ij}[W_G]_{ij}$  for all  $i, j \in V$ . We also define the transformed opinion vector  $\mathbf{z}(t) = (z_1(t), \dots, z_n(t))^T$  as  $z_i(t) := x_i(t) - \frac{1}{2}$  and

$$m(t) := \max_{j \in V} \{|z_j(t)|\}$$

for all  $i \in V, t \geq 0$ . These notations will be used throughout the paper.

A directed path is a concatenation of directed links of  $E$ :  $P = \{(v_{i_1}, v_{i_2}), \dots, (v_{i_{p-1}}, v_{i_p})\} \subset E$ . The length of  $P$  is  $p-1$ . A directed cycle of  $G$  is a directed path with the same beginning and ending node, i.e.,  $v_{i_1} = v_{i_p}$ . A graph is said to be aperiodic if there is no integer  $k > 1$  that divides the length of every cycle of the graph.

**Assumption 2**  $G$  is aperiodic.

Assumption 2 is easy to satisfy. For example, once there exists a node  $i \in V$  with a self arc, i.e.,  $W_{ii} > 0$ ,  $G$  is then aperiodic if Assumption 1 is also satisfied.

Under Assumption 1 and 2, the normalized weight matrix  $W_G$  becomes a primitive matrix, defined as a square nonnegative matrix some power of which is positive [26], where nonnegative (positive) means that all entries of the matrix is nonnegative (positive). It is known from [25]

that  $W_G$  has a simple nonzero eigenvalue with maximum modulus, and the corresponding left (right) eigenvector is positive. Specifically, the only eigenvalue with maximum modulus of  $W_G$  is 1. Denote the left eigenvector of  $W_G$  corresponding to eigenvalue 1 by  $\mathbf{v} = (v_1, \dots, v_n)^T$  such that  $\sum_{j=1}^n v_j = 1$ .

The following proposition is a corollary of Theorem 1 in [22].

**Proposition 1** Let Assumption 1 hold. Suppose  $G$  is structurally balanced under the partition  $V = V_1 \cup V_2$ . If  $x_i(0) < \frac{1}{2}$  for all  $i \in V_1$  and  $x_i(0) > \frac{1}{2}$  for all  $i \in V_2$ , then for (1), it holds for all  $b_i > 0$  that

$$\lim_{t \rightarrow \infty} x_i(t) = 0, \forall i \in V_1; \lim_{t \rightarrow \infty} x_j(t) = 1, \forall j \in V_2,$$

i.e., bipartite polarization occurs. Moreover, the convergence rate is exponential.

With Assumption 1, 2 holding and with  $G$  being structurally balanced under the partition  $V = V_1 \cup V_2$ , we denote  $f_G(\mathbf{v}) := \sum_{j \in V_1} v_j z_j(0) - \sum_{j \in V_2} v_j z_j(0)$ . By Theorem 2 in [24] and Theorem 1 in [23], it holds

$$\lim_{t \rightarrow \infty} [S_G^t \mathbf{z}(0)]_i = \operatorname{sgn}\left(\frac{3}{2} - k\right) f_G(\mathbf{v}), \quad i \in V_k$$

for  $k = 1, 2$ . We then let  $\hat{t}$  be the smallest positive integer satisfying

$$\operatorname{sgn}\left(\frac{3}{2} - k\right) f_G(\mathbf{v}) \cdot [S_G^{\hat{t}} \mathbf{z}(0)]_i > 0, \quad i \in V_k, k = 1, 2$$

if  $f_G(\mathbf{v}) \neq 0$ . Given  $\hat{t}$ , let  $\hat{\delta}, \hat{\delta}_1, \hat{\delta}_2$  be some positive numbers satisfying

$$\hat{\delta} = \frac{3}{2} \hat{\delta}_1 + \hat{\delta}_2 < \min\left\{\frac{1}{2} - m(0), \min_{i \in V} \{|[S_G^{\hat{t}} \mathbf{z}(0)]_i|/2\}\right\}.$$

Moreover, denote

$$\hat{\mu} := \max \left\{ \ln \left( \frac{\frac{1}{2} + m(0) + \hat{\delta}}{\frac{1}{2} - m(0) - \hat{\delta}} \right), 2m(0) + 2\hat{\delta} \right\}. \quad (2)$$

**Theorem 1** Let Assumption 1, 2 hold. Suppose  $G$  is structurally balanced under the partition  $V = V_1 \cup V_2$ . Suppose  $f_G(\mathbf{v}) \neq 0$ . Let  $x_i(0) \in (0, 1)$  for all  $i \in V$ . For (1), there exists a threshold

$$\hat{b} = \min \left\{ \frac{\hat{\delta}_1}{\hat{\mu} \hat{t}}, \frac{\ln(1 - \frac{\hat{\delta}_2}{2\hat{t}})}{\ln(\frac{1}{2} - m(0) - \hat{\delta})} \right\} < 1$$

such that if  $0 < b_i < \hat{b}$  for all  $i \in V$ , bipartite polarization takes place in the sense that

$$\lim_{t \rightarrow \infty} x_i(t) = \frac{1 + \operatorname{sgn}(f_G(\mathbf{v}))}{2}, \quad i \in V_1;$$

$$\lim_{t \rightarrow \infty} x_i(t) = \frac{1 - \operatorname{sgn}(f_G(\mathbf{v}))}{2}, \quad i \in V_2.$$

According to the proof of Theorem 1 given in the next section, we know that for any initial state inside the opinion domain, the opinion evolution can be divided into two stages. In the first stage, the negative interactions drive the opinions

grouped by  $V_1, V_2$  towards two opposite sides of  $\frac{1}{2}$ . In the second stage, under the effect of the bias, the two groups of opinions converge to 0, 1 respectively. This can also be illustrated by the following example.

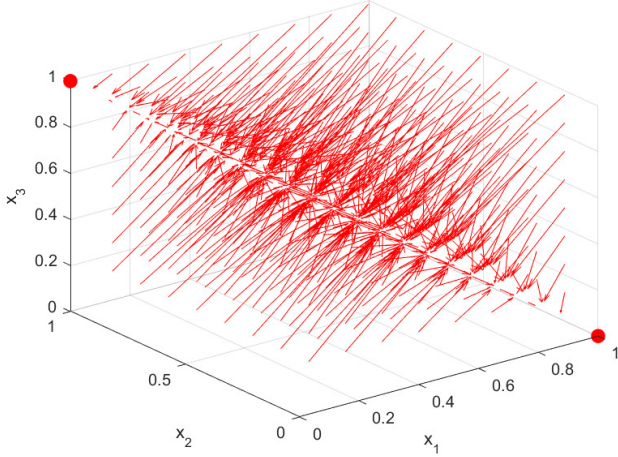


Fig. 1. Phase portraits for Example 1, with  $b_i = 0.1$  for all  $i \in V$ .

**Example 1** Let  $V = \{1, 2, 3\}$ . Consider a structurally balanced complete graph with self-arcs, whose corresponding signed weight matrix is

$$S_G = \begin{pmatrix} \frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ -\frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}.$$

Let  $b_i = 0.1$  for all  $i \in V$ . The phase portraits of the system (1) is shown in Fig 1. We see that for almost all the initial states, the trajectory of (1) gets close to the line  $x_2 = x_3 = 1 - x_1$ , and then converges to  $(0, 1, 1)$  or  $(1, 0, 0)$ . This confirms Theorem 1 and the statements above.  $\square$

We now consider structurally unbalanced networks with small bias. Denote  $\rho_{\max}(S_G)$  ( $\rho_{\min}(S_G)$ ) the maximum (minimum) square root of the eigenvalues of  $S_G^\top S_G$ . When  $G$  is structurally unbalanced, by the Gershgorin circle theorem and the proof of Theorem 1 in [23], all eigenvalues of  $S_G$  have magnitudes less than 1. Therefore, the infinite sum  $\sum_{r=0}^{\infty} (S_G^r)^\top S_G^r$  is well defined. Let  $p_0$  be some number larger than  $\|\sum_{r=0}^{\infty} (S_G^r)^\top S_G^r\|$ . Denote  $\tilde{d} := \sqrt{\frac{3}{20p_0}}$ . By Theorem 2 in [21], we have  $\lim_{t \rightarrow \infty} S_G^t \mathbf{z}(t) = \mathbf{0}$ . There exists  $\tilde{t}$  such that  $S_G^{\tilde{t}} \mathbf{z}(0) \in \mathcal{B}(\frac{\tilde{d}}{2})$ . Suppose  $m(0) < \frac{1}{2}$ . Given  $\tilde{t}$ , let  $\tilde{\delta}, \tilde{\delta}_1, \tilde{\delta}_2$  be some positive numbers satisfying

$$\tilde{\delta} = \frac{3}{2}\tilde{\delta}_1 + \tilde{\delta}_2 < \min\left\{\frac{1}{2} - m(0), \frac{\tilde{d}}{2}\right\}.$$

Moreover, denote

$$\tilde{\mu} := \max\left\{\ln\left(\frac{\frac{1}{2} + m(0) + \tilde{\delta}}{\frac{1}{2} - m(0) - \tilde{\delta}}\right), 2m(0) + 2\tilde{\delta}\right\}. \quad (3)$$

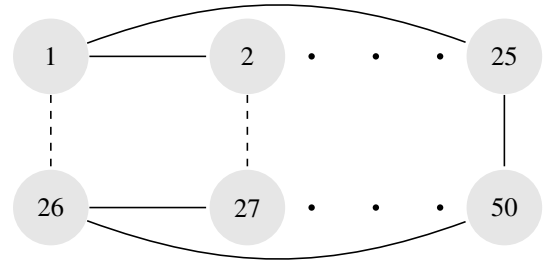


Fig. 2. The structurally unbalanced graph used in Example 2. Solid line represents positive edge, while dashed line represents negative edge.

**Theorem 2** Let Assumption 1 hold. Suppose  $G$  is structurally unbalanced. If  $x_i(0) \in (0, 1)$  for all  $i \in V$ , then for (1), there exists a shreshold

$$\tilde{b} = \min\left\{\frac{1}{25\sqrt{np_0^2\rho_{\max}(S_G)^2 + n}}, \frac{\tilde{\delta}_1}{\tilde{\mu}\tilde{t}}, \frac{\ln(1 - \frac{\tilde{\delta}_2}{2\tilde{t}})}{\ln(\frac{1}{2} - m(0) - \tilde{\delta})}\right\},$$

such that with  $0 < b_i < \tilde{b}$  for all  $i \in V$ , there holds  $\lim_{t \rightarrow \infty} x_i(t) = \frac{1}{2}$  for all  $i \in V$ .

The proof is omitted for lack of space. Also in this case, the process of convergence can be divided into two stages. In the first stage, the influence of structural unbalance overwhelms that of the bias and drives the state trajectories towards a neighborhood of the indecision state, where the influence of the bias is weakened even more by the opinions themselves. All the opinions then converge to the indecision state exponentially fast.

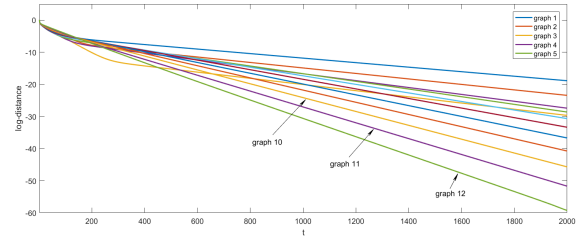


Fig. 3. Simulation for Example 2, with  $b_i = 0.001$  for all  $i \in V$ .

**Example 2** Let  $V = \{1, \dots, 50\}$ . Consider a group of undirected signed graphs  $\{G_k = (V, E_k) | k = 1, \dots, 12\}$  with  $E_k = \{(i, i \bmod 25 + 1) | i = 1, \dots, 25\} \cup \{(i, i \bmod 25 + 26) | i = 25, \dots, 50\} \cup \{(i, i + 25) | i = 1, \dots, 25\}$  and  $E_k^- = \{(i, i + 25) | i = 1, \dots, k\}$ , where the symbol  $\bmod$  is the operator of taking remainder. The network topology of each  $G_k$  is shown as Fig. 2. For every edge of these graphs, the associated weight is  $\frac{1}{3}$ . Let  $b_i = 0.001$  for all  $i \in V$ . Let the initial state be  $x_i(0) = 0.2 + 0.016(i - 1)$  for  $i = 1, \dots, 25$  and  $x_i(0) = 0.4 + 0.016(i - 26)$  for  $i = 26, \dots, 50$ . Define the log-distance between  $\mathbf{x}(t)$  and the opinion centroid as  $d(t) = \ln(\sum_{i=1}^{50} (x_i(t) - \frac{1}{2})^2)$ . For each  $k = 1, \dots, 12$ , we start from the same initial state and plot  $d(t)$  as a function of time  $t$  in Fig. 3. All the

trajectories are shown to converge to  $\frac{1}{2}$ , as predicted by Theorem 2. We also see that the convergence process can be divided in two stages as described above. The convergence rate of the second stage is increasing with the graph label. In fact, by construction, the graphs  $G_k$  are characterized by  $\lambda_{\max}(G_k) > \lambda_{\max}(G_j)$  for  $k < j$ ,  $j, k = 1, \dots, 12$ . According to [19],  $1 - \lambda_{\max}(G_k)$  represents the ‘‘disorder’’ induced by the negative edges (i.e., the ‘‘distance to structural balance’’) in the graph  $G_k$ . Therefore, Fig. 3 indicates that a larger disorder of the underlying network generates a larger convergence rate towards the indecision state in the second stage.

We already know that bilateral polarization occurs for structurally balanced networks and convergence to the opinion centroid occurs for structurally unbalanced networks when the biases are all small. Next we show that when all biases are large and all the initial opinions are on the same side of  $\frac{1}{2}$ , unilateral polarization always occurs. Denote

$$m^*(t) := \min\{\max_{i \in V}\{x_i(t)\}, 1 - \min_{i \in V}\{x_i(t)\}\}. \quad (4)$$

**Theorem 3** *Let Assumption 1 hold. Assume that each agent has at least one positive neighbor. For (1), if all  $x_i(0) - \frac{1}{2}$ ,  $i \in V$  have the same sign, there exists a threshold*

$$b^* = \max_{i \in V} \left\{ \frac{\ln(2d_i) - \ln(d_i^+ m^*(0))}{\ln(1 - m^*(0)) - \ln(m^*(0))} \right\} > 1,$$

such that if  $b_i > b^*$  for all  $i \in V$ , unilateral polarization takes place in sense that

$$\lim_{t \rightarrow \infty} x_i(t) = \frac{1 + \text{sgn}(x_i(0) - \frac{1}{2})}{2}, \quad i \in V. \quad (5)$$

The convergence rate is exponential.

#### IV. ANALYSIS

In this section, we give the proof of Theorem 1. The proof of Theorems 2 and 3 is omitted due to the limitation of space. Without loss of generality, we let  $W_{ii} + d_i = 1$  for all  $i \in V$ .

Denote  $u_i^+(t) := \sum_{j \in N_i^+} W_{ij} z_j(t)$ ,  $u_i^-(t) := \sum_{j \in N_i^-} W_{ij} z_j(t)$  and  $u_i(t) := u_i^+(t) - u_i^-(t)$ . Denote  $l_i(t) := W_{ii} + (\frac{1}{2} + z_i(t))^{b_i} (\frac{d_i}{2} + u_i(t)) + (\frac{1}{2} - z_i(t))^{b_i} (\frac{d_i}{2} - u_i(t))$ . The update of  $z_i(t)$  reads as

$$\begin{aligned} z_i(t+1) &= \frac{(\frac{1}{2} + z_i(t))^{b_i} (\frac{d_i}{2} + u_i(t))}{l_i(t)} - \frac{1}{2} \\ &= \frac{W_{ii} z_i(t)}{l_i(t)} + \frac{(\frac{1}{2} + z_i(t))^{b_i} (\frac{d_i}{2} + u_i(t))}{2l_i(t)} \\ &\quad - \frac{(\frac{1}{2} - z_i(t))^{b_i} (\frac{d_i}{2} - u_i(t))}{2l_i(t)} - \frac{(W_{ii} z_i(t) + u_i(t)) l_i(t)}{l_i(t)} \\ &\quad + W_{ii} z_i(t) + u_i(t) := W_{ii} z_i(t) + u_i(t) + g_i(t) \end{aligned} \quad (6)$$

Rearranging the form of  $g_i(t)$  with  $W_{ii} + d_i = 1$ , we get

$$\begin{aligned} g_i(t) &= \frac{(\frac{d_i}{2} + u_i(t))(\frac{d_i}{2} - u_i(t))}{l_i(t)} \left[ \left(\frac{1}{2} + z_i(t)\right)^{b_i} - \left(\frac{1}{2} - z_i(t)\right)^{b_i} \right] \\ &\quad + \frac{W_{ii} d_i - W_{ii} u_i(t) z_i(t)}{l_i(t)} \left[ \left(\frac{1}{2} + z_i(t)\right)^{b_i} - \left(\frac{1}{2} - z_i(t)\right)^{b_i} \right] \\ &\quad + \frac{W_{ii} u_i(t) - W_{ii} d_i z_i(t)}{2l_i(t)} \left[ \left(\frac{1}{2} + z_i(t)\right)^{b_i} + \left(\frac{1}{2} - z_i(t)\right)^{b_i} - 2 \right] \end{aligned} \quad (7)$$

Denote  $\mathbf{g}(t) := (g_1(t), \dots, g_n(t))^\top$ . Rewrite the iteration of  $z_i(t)$  for  $i \in V$  in compact form as

$$\mathbf{z}(t+1) = S_G \mathbf{z}(t) + \mathbf{g}(t). \quad (8)$$

**Proof of Theorem 1.** We only consider the case that  $f_G(\mathbf{v}) < 0$ , and the proof for  $f_G(\mathbf{v}) > 0$  is similar.

Suppose  $b_i < \hat{b} < 1$  for all  $i \in V$ . Consider  $t$  steps of iterations

$$\mathbf{z}(t) = S_G^t \mathbf{z}(0) + \sum_{r=0}^{t-1} S_G^{t-1-r} \mathbf{g}(r). \quad (9)$$

We estimate the second part on the right-hand side of (9). By construction, there holds  $\sum_{j \in N_i} |A_{ij}| = 1$ . Let  $\hat{\delta}(t) := \max_{i \in V} \{|g_i(t)|\}$  be the  $\infty$ -norm of  $\mathbf{g}(t)$ . We then have  $|\sum_{r=0}^{t-1} S_G^{t-1-r} \mathbf{g}(r)|_i \leq \hat{\delta}(r)$  for all  $i \in V$  and  $r = 0, 1, \dots$ .

Hence,

$$\left| \sum_{r=0}^{t-1} S_G^{t-1-r} \mathbf{g}(r) \right|_i \leq \sum_{r=0}^{t-1} \hat{\delta}(r), \quad i \in V. \quad (10)$$

It is easy to verify that  $z_i(t) \in (-\frac{1}{2}, \frac{1}{2})$  and  $u_i(t) \in (-\frac{1}{2}, \frac{1}{2})$  for all  $i \in V$  and  $t \geq 0$ , which indicates that

$$0 < \left(\frac{1}{2} + z_i(t)\right) \left(\frac{d_i}{2} + u_i(t)\right) + \left(\frac{1}{2} - z_i(t)\right) \left(\frac{d_i}{2} - u_i(t)\right) < l_i(t)$$

holds for all  $b_i \in (0, 1)$ . Thus,

$$\begin{aligned} \frac{(\frac{d_i}{2} + u_i(t))(\frac{d_i}{2} - u_i(t))}{l_i(t)} &< \\ &\frac{(\frac{d_i}{2} + u_i(t))(\frac{d_i}{2} - u_i(t))}{\left(\frac{1}{2} + z_i(t)\right) \left(\frac{d_i}{2} + u_i(t)\right) + \left(\frac{1}{2} - z_i(t)\right) \left(\frac{d_i}{2} - u_i(t)\right)}. \end{aligned} \quad (11)$$

The right hand side is less than 1 no matter if  $\frac{u_i(t)}{d_i} \leq -z_i(t)$  or  $\frac{u_i(t)}{d_i} > -z_i(t)$ . Recalling the definition of  $g_i(t)$  and  $z_i(t), u_i(t) \in (-\frac{1}{2}, \frac{1}{2}), d_i \in (0, 1]$ , we have

$$\begin{aligned} |g_i(t)| &< \frac{3}{2} \left| \left(\frac{1}{2} + z_i(t)\right)^{b_i} - \left(\frac{1}{2} - z_i(t)\right)^{b_i} \right| \\ &\quad + \frac{1}{2} \left| \left(\frac{1}{2} + z_i(t)\right)^{b_i} + \left(\frac{1}{2} - z_i(t)\right)^{b_i} - 2 \right| \\ &\leq \frac{3}{2} \left[ \left(\frac{1}{2} + m(t)\right)^{\hat{b}} - \left(\frac{1}{2} - m(t)\right)^{\hat{b}} \right] \\ &\quad + \frac{1}{2} \left[ 2 - \left(\frac{1}{2} + m(t)\right)^{\hat{b}} - \left(\frac{1}{2} - m(t)\right)^{\hat{b}} \right] := f(m(t), \hat{b}). \end{aligned} \quad (12)$$

By the arbitrariness of  $i$ ,

$$\hat{\delta}(t) < f(m(t), \hat{b}), \quad t \geq 0. \quad (13)$$

With the definition of  $\hat{b}$ , we claim that the following equations are satisfied:

$$\left(\frac{1}{2} + m(0) + \hat{\delta}\right)^{\hat{b}} - \left(\frac{1}{2} - m(0) - \hat{\delta}\right)^{\hat{b}} < \frac{\hat{\delta}_1}{\hat{t}}, \quad (14a)$$

$$2 - \left(\frac{1}{2} + m(t) + \hat{\delta}\right)^{\hat{b}} - \left(\frac{1}{2} - m(t) - \hat{\delta}\right)^{\hat{b}} < \frac{\hat{\delta}_2}{\hat{t}}. \quad (14b)$$

In fact, by analyzing the monotonicity, we obtain

$$\left(\frac{1}{2} + m(0) + \hat{\delta}\right)^b - \left(\frac{1}{2} - m(0) - \hat{\delta}\right)^b \leq \hat{\mu}b, \quad b \in (0, 1),$$

where  $\hat{\mu}$  is defined by (2). Therefore,  $\hat{b} < \frac{\hat{\delta}_1}{\hat{\mu}\hat{t}}$  is sufficient to make (14a) hold. On the other hand, to make (14b) hold, we only need  $2\left(\frac{1}{2} - m(t) - \hat{\delta}\right)^{\hat{b}} > 2 - \frac{\hat{\delta}_2}{\hat{t}}$ , which is equivalent to

$$\hat{b} < \ln\left(1 - \frac{\hat{\delta}_2}{2\hat{t}}\right) / \ln\left(\frac{1}{2} - m(0) - \hat{\delta}\right).$$

By the definition of  $\hat{b}$ , we know that (14) holds. Combining (9), (10), (13) and by induction, we obtain  $\hat{\delta}(t) \leq \frac{\hat{\delta}}{\hat{t}}$  and  $m(t) \leq m(0) + \frac{\hat{t}}{\hat{t}}\hat{\delta}$  for all  $0 \leq t \leq \hat{t}$ . Consequently, there holds

$$\left\| \sum_{r=0}^{\hat{t}-1} S_G^{\hat{t}-1-r} \mathbf{g}(r) \right\|_i \leq \sum_{r=0}^{\hat{t}-1} \hat{\delta}(r) \leq \hat{\delta}$$

for all  $i \in V$ . Combining (9) with the definition of  $\hat{\delta}, \hat{t}$ , we have  $z_i(\hat{t}) \leq \frac{[S_G^{\hat{t}} \mathbf{z}(0)]_i}{2} < 0$  for  $i \in V_1$  and  $z_i(\hat{t}) \geq \frac{[S_G^{\hat{t}} \mathbf{z}(0)]_i}{2} > 0$  for  $i \in V_2$ . This also means that  $x_i(\hat{t}) < \frac{1}{2}$  for all  $i \in V_1$ , and  $x_i(\hat{t}) > \frac{1}{2}$  for all  $i \in V_2$ . Then Proposition 1 can be applied, generating the desired conclusion.  $\square$

## V. CONCLUSION

In this paper, we have extended the biased assimilation model to antagonistic interactions and analyzed its asymptotic behavior for different ranges of the bias parameter over the resulting signed networks. We have shown that structurally balance plays a key role of the asymptotic behavior for small bias, i.e., bipartite polarization occurs for structurally balanced networks while convergence to the indecision state occurs for structurally unbalanced networks. We have also shown that for large bias, unilateral polarization always occurs if all the initial opinions are on the same side of the indecision state.

## REFERENCES

- [1] D. Acemoglu, and O. Asuman, "Opinion dynamics and learning in social networks", *Dynamic Games and Applications*, vol 1, no. 1, pp. 3-49, 2011.
- [2] A. V. Proskurnikov, and R. Tempo, "A tutorial on modeling and analysis of dynamic social networks. Part I", *Annual Reviews in Control*, vol 43, pp. 65-79, 2017.
- [3] J. Scott, *Social Network Analysis*, SAGE, 2017.
- [4] B. D. O. Anderson, and M. Ye. "Recent advances in the modelling and analysis of opinion dynamics on influence networks", *International Journal of Automation and Computing*, vol 16, no. 2, pp. 129-149, 2019.
- [5] M. H. DeGroot, "Reaching a consensus", *Journal of the American Statistical Association*, vol. 69, no. 345, pp. 118-121, 1974.
- [6] N. E. Friedkin, and C. J. Eugene, "Social influence and opinions", *Journal of Mathematical Sociology*, vol. 15, no. 3-4, pp. 193-206, 1990.

- [7] R. Hegselmann, and U. Krause, "Opinion dynamics and bounded confidence models, analysis, and simulation", *Journal of artificial societies and social simulation*, vol. 5, no. 3, 2002.
- [8] F. Baumann, P. Lorenz-Spreen, I. M. Sokolov, and M. Starnini, "Modeling echo chambers and polarization dynamics in social networks", *Physical Review Letters*, vol. 124, no. 4, 2020.
- [9] D. Balsamo, V. Gelardi, C. Han, D. Rama, A. Samantray, C. Zucca and M. Starnini, "Inside the Echo Chamber: Disentangling network dynamics from polarization", arXiv preprint arXiv:1906.09076, 2019.
- [10] P. Danderar, A. Goel, and D. T. Lee, "Biased assimilation, homophily, and the dynamics of polarization", *Proceedings of the National Academy of Sciences* vol. 110, no. 15, pp. 5791-5796, 2013.
- [11] C. G. Lord, L. Ross, and M. R. Lepper, "Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence", *Journal of personality and social psychology*, vol. 37, no. 11, pp. 2098, 1979.
- [12] G. D. Munro, P. H. Ditto, L. K. Lockhart, A. Fagerlin, M. Gready, and E. Peterson, "Biased assimilation of sociopolitical arguments: Evaluating the 1996 US presidential debate", *Basic and Applied Social Psychology*, vol. 24, no. 1, pp. 15-26, 2002.
- [13] G. A. Boysen, and D. L. Vogel, "Biased assimilation and attitude polarization in response to learning about biological explanations of homosexuality", *Sex Roles*, vol. 57, no. 9-10, pp. 755-762, 2007.
- [14] E. Suhay, and C. Erisen, "The role of anger in the biased assimilation of political information", *Political Psychology*, vol. 39, no. 4, pp. 793-810, 2018.
- [15] S. Wasserman, and K. Faust, *Social network analysis: Methods and applications*, vol. 8, Cambridge university press, 1994.
- [16] D. Easley, and J. Kleinberg, *Networks, crowds, and markets*, vol. 8, Cambridge: Cambridge university press, 2010.
- [17] Z. Meng, G. Shi, KH. Johansson, M. Cao, and Y. Hong, "Behaviors of networks with antagonistic interactions and switching topologies", *Automatica*, vol. 73, pp. 110-116, 2016.
- [18] G. Shi, M. Johansson, and K. H. Johansson, "How agreement and disagreement evolve over random dynamic networks", *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 6, pp. 1061-1071, 2013.
- [19] A. Fontan, and C. Altafini, "Achieving a decision in antagonistic multi agent networks: frustration determines commitment strength", *In 2018 IEEE Conference on Decision and Control (CDC)*, pp. 109-114. IEEE, 2018.
- [20] C. Altafini, and G. Lini, "Predictable dynamics of opinion forming for networks with antagonistic interactions", *IEEE Transactions on Automatic Control*, vol. 60, no. 2, pp. 342-357, 2015.
- [21] C. Altafini, "Consensus problems on networks with antagonistic interactions", *IEEE Transactions on Automatic Control*, vol. 58, no. 4, pp. 935-946, 2013.
- [22] Z. Chen, J. Qin, B. Li, H. Qi, B. Peter, and G. Shi, "Dynamics of opinions with social biases", *Automatica*, vol. 106 pp. 374-383, 2019
- [23] G. Shi, C. Altafini, and J. S. Baras, "Dynamics over signed networks", *SIAM Review*, vol. 61, no. 2, pp. 229-257, 2019.
- [24] R. Olfati-Saber, J. A. Fax, and R. M. Murray, "Consensus and cooperation in networked multi-agent systems", *Proceedings of the IEEE*, vol. 95, no. 1, pp. 215-233, 2007.
- [25] R. A. Horn, and C. R. Johnson, *Matrix analysis*, Cambridge university press, 2012.
- [26] H. Charles, and L. Q. Zamboni, "Directed graphs and substitutions", *Theory of Computing Systems*, vol. 34, no. 6, pp. 545-564, 2001.
- [27] W. Xia, J. Liu, M. Cao, and X. M. Sun, "Stability Analysis of a Nonlinear Opinion Dynamics Model for Biased Assimilation", *In 2018 IEEE Conference on Decision and Control (CDC)*, pp. 1287-1292, IEEE, 2018.
- [28] W. Xia, M. Ye, J. Liu, M. Cao, and X. Sun, "Analysis of a Nonlinear Opinion Dynamics Model with Biased Assimilation", arXiv preprint arXiv:1912.01778, 2019.
- [29] H. K. Khalil, *Nonlinear systems*, Upper Saddle River, 2002.
- [30] C. Chen, *Linear system theory and design*, Oxford University Press, Inc., 1998.