Signed Social Networks With Biased Assimilation
Lingfei Wang, Yiguang Hong, Guodong Shi and Claudio Altafini

The self-archived postprint version of this journal article is available at Linköping University Institutional Repository (DiVA):
http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-189321

N.B.: When citing this work, cite the original publication.
hits://doi.org/10.1109/TAC.2021.3121221

Original publication available at:
https://doi.org/10.1109/TAC.2021.3121221

http://www.ieee.org/index.html

©2022 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.
Abstract—A biased assimilation model of opinion dynamics is a nonlinear model in which opinions exchanged in a social network are multiplied by a state-dependent term having the bias as exponent, and expressing the bias of the agents towards their own opinions. The aim of this work is to extend the bias assimilation model to signed social networks. We show that while for structurally balanced networks polarization to an extreme value of the opinion domain (the unit hypercube) always occurs regardless of the value of the bias, for structurally unbalanced networks a stable state of indecision (corresponding to the centroid of the opinion domain) also appears, at least for small values of the bias. When the bias grows and passes a critical threshold which depends on the amount of “disorder” encoded in the signed graph, then a bifurcation occurs and opinions become again polarized.

Index Terms—Opinion dynamics, biased assimilation, signed social networks.

I. INTRODUCTION

Understanding how social interactions influence the formations of opinions in individuals has been a topic of interests for many years, and especially in the last decade it has attracted a lot of attention due to the explosion of social media and the possibilities they offer [1]–[4]. Social opinions dynamics is often modeled on networks, in which a node corresponds to an individual and is endowed with a state variable representing its opinion on a subject. The evolution of the opinions is mediated by the interactions with the neighbouring nodes. One such model for opinion dynamics is the well-known DeGroot model: each node updates its own state to the average of its neighbors’ opinions, producing an asymptotic agreement for all nodes under a connectivity assumption [5]. In the real world, however, instead of agreement, the differences of opinions, attitudes, and behavior among individuals often persist over time [6]. In order to capture this diversity of opinions, several ingredients have been added to the DeGroot model—stubbornness of the initial opinions [7], confidence bound [8], untruthfully expressed opinions [4], bias [9], non-cooperative interactions [10], and so on [11], [12].

In particular, non-cooperative interactions are usually represented by negative edges in the interaction network, which is then called a signed social network [13]–[15]. Signed networks attract interest because of many real examples of non-cooperative relationships, e.g., depending on the context, friendly/unfriendly links, trust/mistrust, collaborative/competitive interactions [16], [17]. Several studies have investigated social opinion dynamics on a signed social network, see e.g. [10], [18]–[21]. Particularly, the author of [10] proposed a linear consensus-like model which, when the graph is structurally balanced, converges to a so-called bipartite consensus, i.e., the opinions split into two consensus values symmetric with respect to the origin, see also [18] for a similar model over random networks. For signed graphs, structural unbalance corresponds to inserting “disorder” into the graph [13]. This disorder can give rise to different behaviors for different dynamical models: in [10], [18] it leads to opinions that converge to 0 (“state of indecision”) while in e.g. [47] it leads to opinion fluctuations and in [46] to multiple clusters. In [19] the amount of such disorder is shown to modulate the bifurcation point in certain nonlinear models.

In the real world, people are often biased when inspecting others’ opinions. This phenomenon in social science is termed biased assimilation, and describes how people tend to accept evidence which is supportive of their own opinion while critically examining a contrarian one [22]. Biased assimilation has been observed in many social experiments [22]–[25]. Several papers have also tried to describe the biased assimilation phenomenon mathematically [9], [26]–[28]. Especially, Pranav et al. [9] proposed a non-linear model in which every node updates its own opinion via a multiplication of the average of its neighbors’ opinions with a state-dependent factor possessing the bias parameter as exponent. The higher the exponent is, the more biased the individual will be. For the model in [9], several papers have investigated the set of equilibria and tested their stability for some special network topologies, see [26]–[28]. In spite of these efforts, due to the non-linear character of the model in [9], the overall phase space analysis of its dynamical behavior is still incomplete. Furthermore, a limitation of the model in [9] is that it only supports a polarized behavior, i.e., the asymptotic opinions generically converge to the extreme values of the state space (the unit hypercube $[0, 1]^n$). The intermediate equilibrium...
point (centroid of the unit hypercube) always present in the model and interpretable as a “state of indecision” (in which the agents do not take any side) is in fact unstable. In [27] it is shown that for negative values of the bias parameter local stability of this intermediate equilibrium point is possible. However, negative values of the bias (“antibias”) seem difficult to justify in practice, as they correspond to individuals not supporting their own opinions. One of the scopes of this paper is to show that a locally stable state of indecision naturally appears when disorder is introduced in the social network in the form of structurally unbalanced antagonistic interactions.

In this paper we reformulate the model in [9] with negative interactions and consider the model over signed networks. As with [9], the resulting signed biased assimilation model is highly nonlinear, meaning that an exhaustive analysis is out with [9], the resulting model is.

- We investigate the asymptotic behavior of our model for different ranges of the bias parameter. We prove that for small bias the opinions polarize in groups over structurally balanced networks, but converge to the indecision state over structurally unbalanced networks. We also show that when the bias is large, polarization occurs regardless of structural balance/unbalance, and that the value to which each node converges only depends on its initial state. Explicit bounds or critical points for all these cases are given. Note that our results apply for general strongly connected network topologies, instead of special network topologies as [9], [26], [27] do for the model in [9]. Note further that also most of the arguments used in the proofs of the various results (e.g. for the computation of the bounds) are novel.

- We show that for structurally unbalanced networks, when the bias parameter grows, bifurcation may take place at the opinion centroid. In particular, we show that the disorder encoded in the signed network (quantified by the largest eigenvalue of the signed adjacency matrix of the network [19]) influences the values of the bias parameter at which the bifurcation occurs. Our theoretical analysis and simulation study show that for small amount of disorder (e.g. one or a few edge sign flips with respect to the balanced case) a very small value of bias is required for polarization to happen. When the disorder increases so does the least value of bias needed for polarization.

- The equilibrium properties for different special combinations of positive and negative subgraphs (see below for details on these notions) are investigated. The values of the bias parameter we explore in this part are large, and in most of the cases the indecision state is the only interior equilibria. This is a situation which is similar to the cases of only cooperative interactions [26], [28].

The rest of this paper is organized as follows: the model formulation and its sociological interpretation are introduced in Section II; the asymptotic behavior for different ranges of the bias parameter is discussed in Section III; equilibria and stability analysis for several special network topologies are investigated in Section IV.

A preliminary version of this paper appears in the proceedings of CDC 2020 [29]. This conference paper deals with a slightly different version of the model (different individual bias coefficients) and mainly explores the asymptotic behavior we discuss in Section III. The bifurcation analysis and the equilibrium analysis of Section IV are presented here for the first time.

Notations. The sets of real, non-negative real, complex and integer numbers are denoted by \( \mathbb{R}, \mathbb{R}_+, \mathbb{C} \) and \( \mathbb{Z} \), respectively. In general, real numbers and imaginary numbers are both denoted by lowercase letters \( x, y, a, b, \ldots \) and lowercase Greek letters \( \alpha, \beta, \ldots \) All vectors are real column vectors denoted by bold lowercase letters \( \mathbf{x}, \mathbf{y}, \mathbf{z}, \ldots \) Given a complex number \( x \), let \( \text{Re}(x) \) and \( \text{Im}(x) \) be its real and imaginary parts. The \( i \)-th entry of a vector \( \mathbf{x} \) is denoted by \( x_i \) or \( |x_i| \). Matrices are denoted by upper case letters such as \( A, B, P, Q \ldots \) 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 \( A \).

II. MODEL FORMULATION AND SOCIOLOGICAL INTERPRETATION

A. Signed Social Networks

Consider a network with \( n \) nodes indexed in the set \( V = \{1, \ldots, n\} \). 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, denoting a directed edge from \( i \) to \( j \) with strength \( A_{ij} \) (or \(-A_{ij} \)).

A directed path is a concatenation of directed links of \( E \):

\[
P_{1:i,p} = \{(v_{i_1}, v_{i_2}), \ldots, (v_{i_{p-1}}, v_{i_p}) \} \subset E.
\]

The length of \( P_{1:i,p} \) 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 1

1) \( G \) is strongly connected;
2) \( G^+ \) contains at least one edge and \((i, i) \notin E^- \) for all \( i \in V \).

Assumption 2 \( G \) is aperiodic.

For a node \( i \in V \), its positive neighbors are the nodes in the set \( N^+_i := \{j : (j, i) \in E^+ \} \). 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 \( W \in \mathbb{R}^{n \times n}_{\geq 0} \), 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 \( V_2 \) mutually disjoint (and one of them possibly empty), where any edge between the two node subsets \( V_1 \) and \( V_2 \) is negative, and any edge within each \( V_i, i = 1, 2 \) is positive; (ii) A network which is not structurally balanced is said to be structurally unbalanced.

**B. The Signed Biased Assimilation Model**

Time is slotted at \( t = 0, 1, \ldots \) Each node \( i \) holds an opinion \( x_i(t) \in [0, 1] \) at time \( t \) and interacts with its neighbors at each time to update its opinion. All agents update their opinions simultaneously at each time step. Define \( s_i^+(t) := \sum_{j \in N_i^+} W_{ij} x_j(t), s_i^-(t) := \sum_{j \in N_i^-} W_{ij} (1 - x_j(t)) \) as the weighted sum of \( i \)'s positive neighbors' opinions and the weighted sum of the opposite of its negative neighbors' opinions. Let \( s_i(t) := s_i^+(t) + s_i^-(t) \). The opinion dynamics of \( x_i(t) \) is described by the signed biased assimilation model:

\[
x_i(t + 1) = \frac{x_i(t)^b s_i(t)}{x_i(t)^b s_i(t) + (1 - x_i(t))^b (d_i - s_i(t))},
\]

where \( b > 0 \) represents the bias of \( i \). All nodes are associated with the same bias, \( x_i(t)^b \) 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 \( b = 0 \), under the transformation \( z_i(t) = x_i(t) - \frac{1}{2} \), \( i \in V \), the signed biased assimilation model (1) is identical to the discrete-time model of [10], [35], in which all opinions range from \(-\frac{1}{2}\) to \(\frac{1}{2}\).

The model (1) is obtained by adding negative interactions into the biased assimilation model proposed in [9], resulting in a model that can be seen as a combination of the models in [9] and [10]. Though simple, this combination provides an extra feature with respect to [9], namely to make the intermediate indecision state (i.e., the opinion centroid \( \left( \frac{1}{2}, \ldots, \frac{1}{2} \right) \)) a local asymptotically stable equilibrium for the model (1). Such indecision state is normally unstable in biased assimilation models, see [9], [27], [28]. Another way to make the indecision state a local attractor, proposed in [27], is to let all agents have negative biases. Given the interpretation of the bias parameter in the model, this approach appears less realistic, as it means that no agent is supporting its own opinion.

Technically, the dynamical form of the signed biased assimilation model (1) has two differences with respect to that in [9]. Firstly, in the update rule, at time \( t \) each node \( i \) in [9] considers \( x_j(t) \) for each \( j \) of its neighbor set, whereas in the model (1) agent \( i \) considers \( x_j(t) \) for \( j \in N_i^+ \) and \( 1 - x_j(t) \) for \( j \in N_i^- \). This update rule encodes a basic social principle of trusting your friends and mistrusting your enemies. Secondly, unlike the formulation in [9], agents of the signed biased assimilation model (1) can be their own positive neighbors, i.e., \( i \in N_i^+ \) if \( W_{ii} > 0 \). Accordingly, at every update, agent \( i \) weights both its neighbors’ and its own opinions by a state-dependent factor, while in [9] the factor weighting each node’s own opinion is constant. This modification does not affect the qualitative behavior of the model, see [29]. As in the analysis of [27], here we also take the values of the bias parameters identical for all nodes. An extension to individual bias parameters is proposed in [29].

**C. Problems of interests**

As mentioned above, the signed biased assimilation model (1) is a combination of two existing models. On the one hand, biased assimilation has been shown in the literature to cause opinion polarization, not only as a mathematical result [9], [26], but also as the real world phenomena [22], [39]. For what concerns the mathematical analysis, the results of [26]–[28] show that the opinion centroid is always unstable whereas the boundary equilibrium points can be stable. Notice that when carrying out the analysis 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 [10] can generate bipartite consensus or convergence to the indecision state, respectively for structurally balanced networks and structurally unbalanced networks [10]. 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?

Since the model (1) is a discrete-time system with high degree of non-linearity, it is important to study its dynamics, including the locations of equilibria, their stability, and so on. In this work, these topics are discussed for signed networks. We give now the definition of polarization used in this paper.

**Definition 2 (Unilateral/Bipartite Polarization)** For opinion dynamics with \( x_i(t) \in [0, 1] \) for all \( i \in V \), we say:

(i) unilateral polarization occurs if \( \lim_{t \to \infty} x_i(t) = 0 \), for all \( i \in V \) or \( \lim_{t \to \infty} x_i(t) = 1 \), for all \( i \in V \).

(ii) bipartite polarization occurs if there exists nonempty, mutually disjointed \( V_1, V_2 \) such that \( V = V_1 \cup V_2 \) and

\[
\lim_{t \to \infty} x_i(t) = 0, \quad i \in V_1; \quad \lim_{t \to \infty} x_i(t) = 1, \quad i \in V_2.
\]

Note that Definition 2 is different from the definition of polarization in [9], and in our opinion, more natural. See also [28], where a similar definition of polarization is discussed in detail.

**D. Outline and sociological interpretation of the main results**

The main results we obtain in the paper can be divided according to two different regimes of the bias parameter.
1) **Small bias.** When the bias is small, the model behaves qualitatively similarly to the model [10], in the sense that
   a) if the graph is structurally balanced and negative interactions exist then bipartite polarization follows, regardless of the initial values of the opinions;
   b) if the graph is not structurally balanced then the opinions converge to the indecision state.

2) **Large bias.** When the bias is large, it dominates over the pattern of signs of the graph, and the opinions become polarised only according to the initial conditions of the agents, in a fashion similar to [9].

These two regimes are well-distinct and the corresponding behaviors admit a reasonable sociological interpretation.

1) Small biases let the pattern of friendly/unfriendly interactions among the agents override possible discordant opinions, at least in the structurally balanced case. The structurally unbalanced case, which in the model [10] leads to collapse to the origin, here admits a similar interpretation in terms of the indecision state. Unlike [9], this state becomes an attractor, meaning that the “tension” and social stress induced by imbalance make it more difficult for the community of agents to reach the “ordinary” opinion steady states of the model, i.e., its corner points \{0, 1\}. In spite of its symmetry, the indecision state should be interpreted, similarly to [10], as impossibility of the agents to become polarized (opinions are “neutralized” when they converge to the indecision state). What can be seen in simulation, and only partially in analysis, is that the more of this social stress is encoded in the network, the higher is the value of the bias parameter for which the indecision state remains an attractor. This is coherent with the results of e.g. [19] and with our intuition that “disorder” makes it difficult to form clear-cut opinions (as the \{0, 1\} points of our opinion space can be thought of).

2) Large biases instead lend more importance to the opinions themselves than to the friendly/unfriendly influences that other agents can exert. Consequently, there is no room for indecision (the opinion centroid becomes unstable), and the polarization pattern follows essentially the pattern of the initial opinions of the agents.

In between the two distinct regimes the behavior of the system is much more difficult to understand and analyze, and only partial results for specific parameter values can be obtained.

### III. ASYMPTOTIC BEHAVIOR DEPENDS ON THE BIAS PARAMETER

In this section, we explore the asymptotic behavior of the signed biased assimilation model (1), for different ranges of the bias parameter and initial opinions. Before showing our results, we give some definitions that will be used throughout the paper.

Define \(W_G \in \mathbb{R}_{\geq 0}^{n \times n}\) as the normalized matrix of \(W\), i.e., \(W_G[i,j] := W[i,j]/\sum_i W[i,j]\) if \((j,i) \in E\), and \(W_G[i,j] = 0\) otherwise. Let \(\sigma_{ij}\) be the sign of link \((j,i) \in E\). Construct the signed weight matrix \(S_G \in \mathbb{R}^{n \times n}\) corresponding to \(G\) as

\[
[S_G]_{ij} = \begin{cases} 
\sigma_{ij}[W_G]_{ij}, & \text{if } (j,i) \in E, \\
0, & \text{otherwise.}
\end{cases}
\]

We also define the transformed opinion vector \(z(t) = (z_1(t), \ldots, z_n(t))^\top\) as

\[
z_i(t) := x_i(t) - \frac{1}{2}, \quad i \in V, t \geq 0,
\]

and

\[
m_z(t) := \max_{j \in V}\{|z_j(t)|\}.
\]

Under Assumptions 1 and 2, the normalized weight matrix \(W_G\) becomes a primitive matrix, defined as a square non-negative matrix some power of which is positive [34], where nonnegative (positive) means that all entries of the matrix are nonnegative (positive). It is known from [33] that \(W_G\) has a simple nonzero eigenvalue with maximum modulus, and the corresponding left (right) eigenvector is positive. Specifically, due to the fact that \(W_G\) is row stochastic, the only eigenvalue with maximum modulus of \(W_G\) is 1. Denote the left eigenvector of \(W_G\) corresponding to eigenvalue 1 by \(v = (v_1, \ldots, v_n)^\top\) such that \(\sum_{j=1}^n v_j = 1\).

#### A. Structurally balanced network with small bias

We first give a basic result that will be used in the following.

**Proposition 1** Let Assumption 1 hold. Suppose that \(G\) 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 the signed biased assimilation model (1), it holds for all \(b > 0\) that

\[
\lim_{t \to \infty} x_i(t) = 0, \forall i \in V_1; \quad \lim_{t \to \infty} x_j(t) = 1, \forall j \in V_2,
\]

i.e., bipartite polarization occurs for nonempty \(V_1, V_2\), while unilateral polarization occurs if \(V_1\) or \(V_2\) is empty. Moreover, the convergence rate is exponential.

The proof is inspired by that of Theorem 1 in [26], and it is given in Appendix I.

We need some more notations for the next theorem, with Assumptions 1, 2 holding and with \(G\) being structurally balanced under the partition \(V = V_1 \cup V_2\). Denote

\[
f_G(v) := \sum_{j \in V_1} v_j z_j(0) - \sum_{j \in V_2} v_j z_j(0).
\]

By Theorem 2 in [31] and Theorem 1 in [30],

\[
\lim_{t \to \infty} [S_G^b z(0)]_i = f_G(v), \quad i \in V_1;
\]

\[
\lim_{t \to \infty} [S_G^b z(0)]_i = -f_G(v), \quad i \in V_2.
\]

We then let \(\ell\) be the smallest positive integer satisfying

\[
\text{sgn}(f_G(v))[S_G^\ell z(0)]_i > 0, \quad i \in V_1,
\]

\[
\text{sgn}(f_G(v))[S_G^\ell z(0)]_i < 0, \quad i \in V_2,
\]
if \( f_G(v) \neq 0 \). Given \( \hat{\nu} \), let \( \hat{\delta} \) be some positive number satisfying
\[
\hat{\delta} < \min \{ \frac{1}{2} - m(z(0)), \min_{i \in V} \{ ||S_G^i z(0)||/2 \} \}.
\]
Moreover, denote
\[
\hat{\mu} := \max \left\{ \log \left( \frac{1 + m(z(0)) + \hat{\delta}}{2 - m(z(0)) - \hat{\delta}} \right), 2m(z(0)) + 2\hat{\delta} \right\}.
\]
(3)

Let \( \hat{b} = \hat{\delta}/(\hat{\mu} \hat{\ell}) \). We know \( \hat{b} < 1 \), since \( \hat{\ell} \geq 1 \) and \( \hat{\mu} > 2\hat{\delta} \).

**Theorem 1** Let Assumptions 1 and 2 hold. Suppose that \( G \) is structurally balanced under the partition \( V = V_1 \cup V_2 \). Suppose \( f_G(v) \neq 0 \). Let \( x_i(0) \in (0, 1) \) for all \( i \in V \). For the signed biased assimilation model (1), polarization takes place for any \( b < \hat{b} \), in the sense that
\[
\lim_{t \to \infty} x_i(t) = \frac{1 + sgn(f_G(v))}{2}, \quad i \in V_1;
\]
\[
\lim_{t \to \infty} x_i(t) = \frac{1 - sgn(f_G(v))}{2}, \quad i \in V_2.
\]

Note that the polarization is unilateral only if either \( V_1 \) or \( V_2 \) is empty, otherwise it is bipartite. The proof is given in Appendix I.

According to the proof, at first, the opinion evolution is dominated by the influence matrix \( S_G \), which drives the opinions grouped by \( V_1, V_2 \) towards two opposite sides of \( \frac{1}{2} \). At a later stage, under the effect of the bias, the two groups of opinions converge to 0,1 exponentially fast. Note from the proof that the upper bound \( \hat{b} \) given in Theorem 1 is conservative, and it can be loosened to any value satisfying equation (17).

**B. Structurally unbalanced network with small bias**

We now consider structurally unbalanced networks with small bias, under Assumption 1. In this subsection, we only consider graphs in which each node has a self-arc.

By Theorem 1 of [45], we have \(\lim_{t \to \infty} S_G^s y(z(0)) = 0\) for any \( z(0) \), thus all eigenvalues of \( S_G \) are strictly inside in the unit circle. By the statements after Theorem 5.55 of [32], the infinite sum \( \sum_{r=0}^{\infty} (S_G')^r S_G \) is well defined. Let \( p_0 \) be some number larger than \( \| \sum_{r=0}^{\infty} (S_G')^r S_G \| \). Denote \( \tilde{\delta} := \sqrt{\frac{3}{2p_0}} \) and let \( \tilde{\ell} \) be such that
\[
S_G^{\tilde{\delta}} z(0) \in B(\frac{\tilde{\delta}}{2}).
\]

Suppose \( m(z(0)) < \frac{1}{2} \) (i.e., \( x_i(0) \neq \{0, 1\} \)). Define
\[
\tilde{\mu} := \max \left\{ \log \left( \frac{1 + m(z(0)) + \tilde{\delta}}{2 - m(z(0)) - \tilde{\delta}} \right), 2m(z(0)) + 2\tilde{\delta} \right\},
\]
where \( \tilde{\delta} \) is any positive number satisfying
\[
\tilde{\delta} < \min \{ \frac{1}{2} - m(z(0)), \frac{\tilde{\delta}}{2} \}.
\]
Let
\[
\hat{b} = \min \left\{ \frac{1}{10\sqrt{p_0} ||S_G|| + 1}, \frac{\delta}{\mu \ell} \right\} < 1.
\]
With these notations, we present the following result.

**Theorem 2** Let Assumption 1 hold. Suppose that \( G \) is structurally unbalanced and that each node has a self-arc, i.e., \( W_{ii} > 0 \) for all \( i \). If \( x_i(0) \in (0, 1) \) for all \( i \in V \), then for the signed biased assimilation model (1) with \( b < \hat{b} \), \( \lim_{t \to \infty} x_i(t) = \frac{1}{2} \) holds for all \( i \in V \).

The proof of Theorem 2 is given in Appendix II. Observing the opinion evolution process shown in this proof, for \( b \) small enough the influence of structural unbalance dominates 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. The conclusion of Theorem 2 is that polarization will not happen under the given conditions provided that the bias is small enough.

**C. Discussion about the opinion centroid for structurally unbalanced networks**

In this subsection, for small bias over structurally unbalanced networks, we discuss the stability of the opinion centroid and the bifurcation that may take place at it. As in the last subsection, we suppose each node has a self-arc.

Before proceeding to the discussion, we give some definitions. For any \( x = (x_1, \ldots, x_n) \in \mathbb{R}^n \)
\[
s_i = \sum_{j \in N^+} W_{ij} x_j + \sum_{j \in N^-} W_{ij} (1 - x_j),
\]
define a map \( h(x, b) := (h_1(x, b), \ldots, h_n(x, b)) \) with \( b \) being a fixed bias parameter and
\[
h_i(x, b) = \frac{x_i^b s_i}{x_i^b s_i + (1 - x_i)^b (d_i - s_i)}.
\]
If we rewrite the signed biased assimilation model (1) into compact form, it corresponds to the map \( x \mapsto h(x, b) \). Linearizing \( h \) at the indecision state \( (\frac{1}{2}, \ldots, \frac{1}{2}) \), we obtain its Jacobian matrix, denoted by \( J_{h_b} \),
\[
J_{h_b} = bI + S_G.
\]
Let \( \lambda_i, i \in V \), be the eigenvalues of \( S_G \), with multiplicity counted. Define
\[
b_0 = \min_{i \in V} \left\{ \sqrt{1 - \text{Im}(\lambda_i)^2} - \text{Re}(\lambda_i) \right\}.
\]
We have the following result.

**Proposition 2** The equilibrium point \( (\frac{1}{2}, \ldots, \frac{1}{2}) \) of the signed biased assimilation model (1) is locally stable if \( b < b_0 \), and unstable if \( b > b_0 \).

The proof of Proposition 2 is in Appendix III. Note that \( b_0, \hat{b} \) are both small values. In general, as a bound \( \hat{b} \) is conservative and smaller than \( b_0 \).

**Remark 1** Since the sum of \( S_G \)'s eigenvalues is equal to \( \text{tr}(S_G) \) and \( \text{tr}(S_G) > 0 \), there is at least one eigenvalue of \( S_G \)
that has a positive real part, which yields $b_0 < 1$. Therefore, by Proposition 2, $\left(\frac{1}{2}, \ldots, \frac{1}{2}\right)^T$ is unstable for $b \geq 1$.

Remark 2 If $S_G$ is symmetric, it holds that $b_0 = 1 - \lambda_{\text{max}}(S_G)$, where $\lambda_{\text{max}}(S_G)$ is called the algebraic conflict of $S_G$, and can be used to represent the “disorder” of a signed graph (i.e., its “distance to structural balance”) [19]. Therefore, the intuition behind Proposition 2 is that a network with higher disorder gives a wider range of bias values for which the state of indecision is stable, see also Example 3.

In Remark 2, we show that if $S_G$ is symmetric, $1 - \lambda_{\text{max}}(S_G)$ is a critical point where stability of the opinion centroid changes. Now we go further, and show that a bifurcation happens at $1 - \lambda_{\text{max}}(S_G)$. One assumption is needed as follows.

Assumption 3 $S_G$ is symmetric and the maximum eigenvalue of $S_G$ is simple.

Define $f(x, b) := h(x, b) - x$. Following [41], a bifurcation point (with respect to $b$) for the signed biased assimilation model (1) is a point $(x_0, b_0)$ at which the number of solutions of

$$f(x, b) = 0$$

changes when $b$ crosses $b_0$ (see [41] for more details).

The derivatives of $f$ with respect to $x$ and $b$ are denoted respectively $f_x$ and $f_b$. We also define two operators $\text{rank}()$ and $\text{range}()$ over any square matrix, to represent its rank and the vector space spanned by its columns.

Theorem 3 Let Assumptions 1 and 3 hold for the signed biased assimilation model (1). Suppose that $G$ structurally unbalanced with every node having a self-arc. When $b$ crosses $1 - \lambda_{\text{max}}(S_G)$, bifurcation occurs and new equilibrium points appear in some neighborhood of $\left(\frac{1}{2}, \ldots, \frac{1}{2}\right)$.

The proof of Theorem 3 is in Appendix III. Notice that the bifurcation in Theorem 3 is a pitchfork bifurcation (this follows from $a_1 = 0$ in the proof of Theorem 3, see Section 5.5 of [41]). Pitchfork bifurcations are seen also in other multiagent contexts such as social opinion formation [42] and animal group motion [43].

D. Large bias

We already know that when $b$ is small, polarization will occur for structurally balanced networks and not occur for structurally unbalanced networks. In this section, we continue to explore the asymptotic behavior of the signed biased assimilation model (1) in the case that the bias is large. Without loss of generality we assume that $d_i = 1$ for all $i \in V$.

For all $i \in V$, denote

$$y_j^i(t) = x_j(t), j \in N_i^+; \quad y_j^i(t) = 1 - x_j(t), j \in N_i^-; \quad N_i^0 = \{j | \text{sgn}(y_j^i(0)) - \frac{1}{2} \} = \text{sgn}(x_i(0) - \frac{1}{2})$$

where $N_i^0$ is the set of positive neighbors with the same standpoint of $i$ and negative neighbors with opposite standpoints to $i$. Here “same standpoint” means at the same side of $\frac{1}{2}$. Let

$$p_i(0) = \log \left( \frac{1 - \sum_{j \in N_i^0} W_{ij}(1 - y_j^i(0))}{\sum_{j \in N_i^0} W_{ij}(y_j^i(0))} \right) / \log \left( \frac{1 - x_i(0)}{x_i(0)} \right),$$

$$q_i(0) = \log \left( \frac{1 - \sum_{j \in N_i^0} W_{ij}y_j^i(0)}{\sum_{j \in N_i^0} W_{ij}y_j^i(0)} \right) / \log \left( \frac{x_i(0)}{1 - x_i(0)} \right),$$

and

$$b^* = \max \left\{ \max_{x_i(0) < \frac{1}{2}} \{ p_i(0) \}, \max_{x_i(0) > \frac{1}{2}} \{ q_i(0) \} \right\} + 1. \tag{5}$$

Theorem 4 Suppose $x_i(0) \neq \frac{1}{2}$ and $N_i^0 \neq \emptyset$ for all $i \in V$. For the signed biased assimilation model (1) with $b > b^*$, it holds

$$\lim_{t \to \infty} x_i(t) = \frac{1 + \text{sgn}(x_i(0) - \frac{1}{2})}{2}, \quad i \in V. \tag{6}$$

The convergence rate is exponential.

The proof of Theorem 4 is given in Appendix IV. Polarization is claimed by Theorem 4. Whether the polarization is unilateral or bipartite depends on the initial opinions. Intuitively, if we divide the opinions into two kinds (i.e., positive neighbors hold the same standpoints or negative neighbors hold opposite ones), then they will polarize their opinions towards the extreme values determined by the initial standpoints. However, if no supporting evidence is observed initially for node $i$ (i.e., $N_i^0 = \emptyset$), the opinion evolution is much more complicated and not covered by our theorem. Note that we do not use any property of structural balance/unbalance in the proof.

The lower bound $b^*$ in Theorem 4 leading to polarization depends on both the network topology and the initial opinions. Conversely, given $W$ and $b > 1$, Theorem 4 also gives a condition on the initial opinions such that (6) holds. Define

$$w^* = \min_{i \in V} \sum_{j \in N_i^0} W_{ij};$$

$$m^* = \max \left\{ \max_{x_i(0) < \frac{1}{2}} \{ x_i(0) \}, 1 - \min_{x_i(0) > \frac{1}{2}} \{ x_i(0) \} \right\}. \tag{6}$$

From the definition, we have $m^* = \| x(0) - x^* \|_\infty$, where $x_i^* = \frac{1 + \text{sgn}(x_i(0) - \frac{1}{2})}{2}$ for all $i \in V$.

Corollary 1 Suppose $N_i^0 \neq \emptyset$ for all $i \in V$. Given $b > 1$, (6) holds with an exponential convergence rate if

$$m^* \leq \frac{w^*}{(2 - w^*)^{\frac{1}{\gamma + 1}}} \tag{6},$$

The proof of Corollary 1 is given in Appendix IV. Intuitively, $m^*$ represents an estimation of the domain of attraction of the polarized equilibrium $x^*$. From this point of
view, Corollary 1 gives an upper bound of $m^*$ under which opinion polarization happens. Stability of some of the boundary equilibrium points can be inferred from Corollary 1, as shown in the following corollary. Here boundary equilibrium point refers to $x^*$ with all entries as 0 or 1.

**Corollary 2** Given $b > 1$, a boundary equilibrium point $x^*$ is locally exponentially stable if for all $i \in V$, there exists one $j \in N_i$ such that

$$\text{sgn}(x^*_i - \frac{1}{2}) \text{sgn}(x^*_j - \frac{1}{2}) = \sigma_{ij}.$$  

The proof of Corollary 2 is given in Appendix IV.

Let us apply Corollary 2 to a special case, in which $G$ is a complete graph with no negative interactions. We know that $[x_1^*, \ldots, x_n^*]^T$ with $2 \leq \sum_{i=1}^n x_i^* \leq n - 2$ and $x_i^* \in \{0, 1\}$ for all $i \in V$ are locally exponentially stable if $b > 1$. This is close to Theorem 5 in [28], which claims for the model in [9] that $[x_1^*, \ldots, x_n^*]^T$ with $2 \leq \sum_{i=1}^n x_i^* \leq n - 2$ and $x_i^* \in \{0, 1\}$ for all $i \in V$ are locally exponentially stable if all the bias parameters are larger than 1. Note that [28] assumes that all weights associated to $G$ are equal to 1, while here we do not make any restriction on edge weights.

Finally, notice that the bounds on $b$ obtained in Theorems 1, 2 and 4 are conservative.

**E. Numerical examples**

In this subsection we do some numerical simulations to verify the theorems above. At first, we consider a structurally balanced network with a small bias.

**Example 1** Let $V = \{1, 2, 3, 4\}$. Consider the structurally balanced digraph shown by Figure 1. The corresponding signed weight matrix is

$$S_G = \begin{pmatrix}
\frac{1}{3} & 0 & \frac{1}{3} & \frac{1}{3} \\
0 & \frac{1}{3} & \frac{1}{3} & 0 \\
\frac{1}{3} & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\
\frac{1}{3} & 0 & \frac{1}{3} & \frac{1}{3}
\end{pmatrix}.$$  

Let the initial opinion vector be $x(0) = [0.1, 0.4, 0.4, 0.7]'$, or equivalently, $z(0) = [-0.4, -0.1, -0.1, 0.2]'$ and $n_e(0) = 0.4$. Following the notations in Theorem 1, we take $\ell = 2$, $\delta = 0.06$ and $b = 0.04$. The conditions given by equation (17) are then verified to be satisfied. The opinion evolutions are displayed in Fig. 2, showing that the opinions of node 1, 2 converge to 0, and the opinions of node 3, 4 converge to 1, as expected from Theorem 1.

Next we turn to structurally unbalanced networks.

![Fig. 1](image)

**Fig. 1.** Left: the structurally balanced graph used in Example 1; Right: the structurally unbalanced graph used in Example 2. Solid lines represent cooperative interactions, and dashed lines represents antagonistic interactions.

![Fig. 2](image)

**Fig. 2.** Simulation for Example 1. The red lines represent opinions of node 1, 2, and the green lines represents opinions of node 3, 4.

![Fig. 3](image)

**Fig. 3.** Simulations for Example 2. Panels (a)-(c) display the opinion evolution for different $b$. Panel (d) shows the bifurcation diagram in the $x_1 - x_3$ plane as a function of $b$.

**Example 2** Let $V = \{1, 2, 3, 4\}$. Consider the structurally unbalanced graph shown by Fig. 1, with the corresponding signed weighed matrix

$$S_G = \begin{pmatrix}
\frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\
0 & \frac{1}{2} & \frac{1}{2} & 0 \\
\frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\
\frac{1}{2} & 0 & \frac{1}{2} & \frac{1}{2}
\end{pmatrix}.$$  

We perform three groups of simulations:

1) Let $b = 0.1$. The initial opinion vector is $x(0) = [0.1, 0.4, 0.6, 0.7]'$. The opinion evolutions are plotted in Fig. 3(a). We see that all the opinions converge to $\frac{1}{2}$, as indicated by Theorem 2.

2) Let $b = 2.5$. The initial opinion vectors are set as $x(0) = [0.49, 0.49, 0.49, 0.49]$ and $x(0) = [0.49, 0.51, 0.49, 0.51]$. The opinion evolutions are plotted in Fig. 3(b) and 3(c). All the opinions in Fig. 3(b) converge to 0, while the opinions of node 1, 3 converge to 0 and the opinions of node 2, 4 converge to 1 in Fig. 3(c).
Since all agents of G have at least one positive neighbor, Theorem 5 can be applied.

3) Let \( b \) vary from 0.31 to 0.38 with stepsize 0.001. For each \( b \), take two initial opinion vectors as \([0.3, 0.3, 0.3, 0.7, 0.7]^{\top}\) and \([0.7, 0.7, 0.3, 0.3]^{\top}\). Then we update the opinions. For all the updates, we observe that opinions has stabilized by step 2000. The blue curves in Figure 3(d) are the opinions of node 1 and node 3 at the 2000th step, as \( b \) varies, for the two different initial opinion vectors. The green line corresponds to \( x_1 = x_3 = 0.5 \). We see that the bifurcation happen when \( b \) crosses 0.333. The opinion centroid is stable for \( b < 0.333 \) and unstable for \( b > 0.333 \). Notice that \( 1 - \lambda_{\text{max}}(S_G) \) is exactly 0.333, as predicted by Theorem 3.

For a linear system, the convergence rate depends on the spectral radius of the state-transition matrix. Now we step a little further and in next example we explore the relation between the spectral radius of \( S_G \) and the convergence rate in Theorem 2. The example also shows that a larger disorder in the signed graph leads to a larger domain of bias values for which we have convergence to the opinion centroid.

**Example 3** Consider a sequence of undirected signed complete graphs \( G_k = (V, E_k) \), with \( V = \{1, \ldots, 50\} \), \( E_k = \{(i,j) | i,j = 1, \ldots, 50\} \) for \( k = 1, \ldots, 25 \). The edge set is split into \( E_k = E_k^{+} \cup E_k^{-} \), with \( E_k^{+} = \{(50,j) | j = 1, \ldots, 50\} \cup \{(i,50) | i = 1, \ldots, 50\} \) and \( E_k^{-} = E_k \setminus E_k^{+} \). Every entry of \( W_{G_k} \) is equal to \( \frac{1}{50} \).

1) Let \( b = 0.001 \) and \( x_i(0) = 0.6 + \frac{i - 1}{49} \times 0.3 \) for \( i = 1, \ldots, 50 \). Define the log-distance at time \( t \) from \( x(t) \) to \( (\frac{1}{2}, \ldots, \frac{1}{2}) \) as \( d(t) := \log(\sum_{i=1}^{50}(x_i(t) - \frac{1}{2})^2) \). For graph \( G_k \), once \( d(t) < -30 \), we stop the iteration and record the stopping time. We also calculate the largest eigenvalues of \( S_{G_k} \). The variations of the stopping time and the largest eigenvalue for different \( k \) are shown in Fig. 4(a). We see that all opinion trajectories enter the ball centered at \( (\frac{1}{2}, \ldots, \frac{1}{2}) \) and of radius \( e^{-30} \). This follows from Theorem 2. Moreover, in this special case, the stop time has a positive correlation with the largest eigenvalue, indicating that a bigger largest eigenvalue of \( S_{G_k} \) generates a smaller rate of convergence to the state of indecision.

2) Let all the \( x_i(0) \)’s be 0.3. For every \( G_k \), we try different values of \( b \) in order to estimate the largest \( \tilde{b} \) in Theorem 2. These estimates for different \( k \) are plotted against the largest eigenvalue of \( S_{G_k} \) in Fig. 4(b). The estimated largest \( \tilde{b} \) is shown to have negative correlations with the spectral radius, which matches Remark 2.

**IV. INVESTIGATING EQUILIBRIUM POINTS AND THEIR PROPERTIES**

When studying a complex nonlinear system like the signed biased assimilation model (1), it is important to understand its loci of equilibrium points. For the model (1), it is not hard to verify that all boundary points \( (a_1, \ldots, a_n) \) with \( a_i = 0 \) or 1, and the interior point \( (\frac{1}{2}, \ldots, \frac{1}{2}) \) are always equilibrium points for any underlying network topology. The difficult part is to understand whether these are the only equilibrium points or if there are others, and how they change with the network topology and with the bias parameter. In this section we are going to explore this problem for some special cases.

**Assumption 4** 1) \( G \) is undirected, i.e., \( \sigma_{ij} = \sigma_{ji} \) for all \( i, j \in V \); 2) \( W_{ij} = 1 \) for all \( (j, i) \in E \) and \( (i, i) \notin E \) for all \( i \in V \).

We remind that a graph \( G \) is complete if \( (i, j) \in E \) for all \( i, j \in V \). It is a star graph if \( E = \{(i,n) : i = 1, \ldots, n-1\} \), and a cycle graph if \( E = \{(1,2), \ldots, (n-1,n), (n,1)\} \). The interior equilibria set (i.e., set of equilibrium points with all entries located in \((0,1)\)) of the signed biased assimilation model (1) is denoted by \( W_{\text{int}} \).

**A. Theoretical results**

When \( G \) is a complete graph and the subgraph \( G^- \) is a star graph, \( (\frac{1}{2}, \ldots, \frac{1}{2}) \) will be the only interior equilibrium point.

**Theorem 5** Suppose that \( G \) is a complete graph and \( G^- \) is a star graph with \( n \geq 3 \). Let Assumption 4 hold. Then,

\[
W_{\text{int}} = \{(\frac{1}{2}, \ldots, \frac{1}{2})^\top\}
\]

if \( b \leq 1 \) or \( b \geq 2 \). The equilibrium point \( (\frac{1}{2}, \ldots, \frac{1}{2})^\top \) is unstable.

The proof of Theorem 5 is given in Appendix V.

Note that \( G \) is structurally balanced. Theorem 5 is an extension of the results of [26]. In fact, under the transformation (7), all results about locations of equilibrium points for cooperative networks can be used for structurally balanced networks.

More interesting for us are the results for structurally unbalanced networks. In what follows, we consider highly antagonistic environments. Two special cases are going to be explored. One is when there is a “coordinator” keeping positive interactions with all the remain nodes, the other is when there exists sparse positive interactions, represented by a positive cycle subgraph.

**Theorem 6** Consider the signed biased assimilation model (1). Suppose that \( G \) is a complete graph with \( n \geq 3 \). Suppose \( G^+ \) is a star graph with node \( n \) being the center. Let Assumption 4 hold. Then,

1) if \( b \in [1,2) \cup (2, \infty) \), \( W_{\text{int}} = \{(\frac{1}{2}, \ldots, \frac{1}{2})^\top\} \), and \( (\frac{1}{2}, \ldots, \frac{1}{2})^\top \) is unstable; 2) if \( b = 2 \), \( W_{\text{int}} = \{(a, \ldots, a, 1-a)^\top : \forall a \in (0,1)\} \).

**Theorem 7** Consider the signed biased assimilation model (1). Suppose that \( G \) is a complete graph and \( G^+ \) is a cycle graph with \( n > 5 \). Let Assumption 4 hold. Then,

\[
W_{\text{int}} = \{(\frac{1}{2}, \ldots, \frac{1}{2})^\top\}
\]

holds if \( b = 1 \) or \( b \geq 2 \).

The proofs of Theorems 6, 7 are given respectively in Appendix VI and VII.
are subject to positive network \([26], [28]\). We remind that all these cases equilibrium points of the model in \([9]\) over the explored some existing results also show the instability of the interior points. However, it is far from clear whether this is true for other positive network topologies. positive networks reduce the number of interior equilibrium means that the negative interactions added to the two special cycle network) contains the one of the model \((1)\) over the pure positive star network (resp. compared to \(\text{Theorem 6}\) (resp. \(\text{Theorem 7}\)), for any bias parameter in the range given (resp. the pure positive cycle) network. Compared to \(\text{Theorem 6}\) (resp. \(\text{Theorem 7}\)), for any bias parameter in the range given \((\text{Theorem 6})\) over the pure positive star network (resp. pure positive cycle network) contains the one of the model \((1)\) over the pure positive cycle network. Compared to \(\text{Theorem 6}\) (resp. \(\text{Theorem 7}\)), the interior equilibrium set \(\{a, a, 1 - a\} a \in (0, 1)\}. \]

In \([26]\), the authors explored the locations of interior equilibrium points for the model in \([9]\) over the pure positive star (resp. the pure positive cycle) network. Compared to \(\text{Theorem 6}\) (resp. \(\text{Theorem 7}\)), for any bias parameter in the range given by \(\text{Theorem 6}\) (resp. \(\text{Theorem 7}\)), the interior equilibrium set of the model in \([9]\) over the pure positive star network (resp. cycle network) contains the one of the model \((1)\) over the signed network given in \(\text{Theorem 6}\) (resp. \(\text{Theorem 7}\)), and there exist some cases for which the containment is strict. This means that the negative interactions added to the two special positive networks reduce the number of interior equilibrium points. However, it is far from clear whether this is true for other positive network topologies.

On the other hand, for all the values of bias we consider in the above theorems, every interior point is unstable. Some existing results also show the instability of the interior equilibrium points of the model in \([9]\) over the explored positive network \([26], [28]\). We remind that all these cases are subject to \( b \geq 1\). The intuition behind is that when the bias is large, the inclination of each agent to “keep close to” or “stay away from” its neighbors is much weaker than the inclination to stick to its own stance, and the agents’ opinions will necessarily polarize if not initialized at an equilibrium point.

### B. Numerical examples

For most of the cases considered in this section, the opinion centroid is proved to be unstable. Now we use numerical examples to illustrate this. At first we choose a simple case described by \(\text{Theorem 6}\) to show the opinion evolution when the system is initialized close to the opinion centroid.

**Example 4** Let \( V = \{1, 2, 3\} \). Consider the structurally unbalanced graph with signed weight matrix

\[
S_G = \begin{pmatrix}
0 & -\frac{1}{2} & \frac{1}{2} \\
-\frac{1}{2} & 0 & \frac{1}{2} \\
\frac{1}{2} & \frac{1}{2} & 0
\end{pmatrix},
\]

which is a special case of the graphs in \(\text{Theorem 6}\). We perform \(9^3\) simulations for the same bias parameter \(b = 1\) and different initial states. For each of the simulations, \(x_i(0)\) is taken from \(\{0.45 + 0.0125 \times k|k = 0, \ldots, 8\}\) for \(i \in V\). We collect all the opinion evolution trajectories as a phase portrait in Fig. 5. The phase portrait is shown to be symmetric about the blue plane which crosses the green line. We see that almost every trajectory converges to some boundary equilibrium point, which indicates the instability of \((\frac{1}{2}, \ldots, \frac{1}{2})\) for this special case.

Next we take more samples, to validate the instability of the opinion centroid.

**Example 5** Let \( G = (V, E) \) with \( V = \{1, \ldots, 500\} \). We perform 120 groups of simulations, tagged by \((p, q)\), for randomly chosen edge sets with different probabilities. Here \(p \in \{0.01, 0.1, 0.19, \ldots, 1\}\) represents the probability that there is a edge from \(i\) to \(j\) for any \(i, j \in V\) and \(q \in \{0.01, 0.11, \ldots, 0.91\}\) represents the probability that any off-diagonal entry of \(S_G\) is negative. The absolute values of all non-zero entries of \(S_G\) are equal to 1. The bias \(b\) is
constantly set to be 1. For every group of simulations, the initial states are chosen from \([0.49,0.51]\) for 10 times and we plot the variations of \(d(t)\) each time, where \(d(t)\) is the log-distance defined as \(d(t) := \log(\sum_{i=1}^{500} (x_i(t) - \frac{1}{2})^2)\). All the plots are collected in Fig. 6. We see that for every simulation, \(d(t)\) increases to \(\log(125)\), which is indeed the log-distance from \(\left\{\frac{1}{2}, \ldots, \frac{1}{2}\right\}\) to any boundary equilibrium point in \(\{a_1, \ldots, a_8\}|a_i = 0\) or 1 for \(i \in V\). The result means that polarization always occur no matter how close the initial state is to the opinion centroid, indicating that \(\left\{\frac{1}{2}, \ldots, \frac{1}{2}\right\}\) is not stable for all these cases.

\[\text{V. CONCLUSION}\]

When the bias assimilation model is reformulated on signed networks, a novel feature appears. The state of indecision can become a local attractor in opinion space. The sociological interpretation of this result is that lack of structural balance plays against polarization, because it becomes unclear which nodes are “friends” and which are “enemies” on a global scale. This can happen only when the network is structurally unbalanced and the bias parameter is small. When the bias parameter grows the model bifurcates, and for sufficiently large bias values polarization occurs regardless of how unbalanced the network is, i.e., self-opinions become more relevant than any pattern of friendly/unfriendly ties. Interestingly, the exact point where the bifurcation occurs seems to be a function of the amount of disorder present in the signed graph, which is coherent with what was observed in [19]. Although a thorough theoretical analysis of this behavior is very complicated, we plan to investigate it in more detail in the future.

\[\text{APPENDIX I}\]

\[\text{PROOF OF PROPOSITION 1 AND THEOREM 1}\]

\[\text{Proof of Proposition 1}.\quad \text{Under the transformation}\]

\[y_i(t) = x_i(t), \quad i \in V_1; \quad y_i(t) = 1 - x_i(t), \quad i \in V_2,\]

\[\text{the evolution of } y_i(t), i \in V, \text{ is also in the form of the}\]

\[\text{model (1), with only positive edges in the associated network.}\]

\[\text{Therefore, we only need to prove that for a positive graph } G,\]

\[\text{if } x_i(0) < \frac{1}{2} \text{ for all } i \in V, \text{ } x_i(t) \text{ converges to } 0 \text{ exponentially for all } i \in V.\]

Rearranging the model (1), we have

\[1 - x_i(t + 1) = \left(1 - x_i(t)\right)^b + \frac{u_i(t)}{a_i}.\]

\[\text{Define } m_x(t) = \max_{i \in V} x_i(t). \text{ Since } x_i(0) \leq \frac{1}{2}, i \in V, \text{ by (8),}\]

\[\frac{1 - x_i(1)}{x_i(1)} = \left(\frac{1 - m_x(0)}{m_x(0)}\right)^{1+b} > 1 - m_x(0) > \frac{1 - m_x(0)}{m_x(0)}.\]

As a consequence, \(x_i(1) < m_x(0) < \frac{1}{2}.\) By the arbitrariness of \(i,\) we obtain \(m_x(1) < m_x(0).\) This process can be iterated, which yields

\[1 - m_x(t) > m_x(0)^{(1+b)^t} \].

Therefore, the desired conclusion is proved.

\[\text{Proof of Theorem 1}.\quad \text{We only consider the case that } f_G(v) < 0, \text{ and the proof for } f_G(v) > 0 \text{ is similar.}\]

\[\text{Suppose } b < 1. \text{ Denote}\]

\[u_i^+(t) := \sum_{j \in N_i^+} W_{ij} z_j(t), \quad u_i^-(t) := \sum_{j \in N_i^-} W_{ij} z_j(t)\]

\[\text{and } u_i(t) := u_i^+(t) + u_i^-(t) \text{. Denote}\]

\[l_i(t) := \left(1 + z_i(t)\right)^b \left(\frac{1}{2} + \frac{u_i(t)}{d_i}\right) + \left(1 - z_i(t)\right)^b \left(\frac{1}{2} - \frac{u_i(t)}{d_i}\right).\]

\[\text{The update of } z_i(t) \text{ reads as}\]

\[z_i(t + 1) = \left(\frac{1}{2} + z_i(t)\right)^b \frac{d_i l_i(t)}{d_i l_i(t)} - \frac{1}{2} \]

\[= \left(\frac{1}{2} + \frac{u_i(t)}{d_i}\right) \frac{1}{d_i} \frac{1}{d_i} \left(\frac{1}{2} + z_i(t)\right)^b - \left(1 - z_i(t)\right)^b\]

\[\frac{u_i(t)}{d_i} := \frac{u_i(t)}{d_i} + g_i(t) \]

\[\text{Denote } g_i(t) := (g_1(t), \ldots, g_n(t))^T. \text{ Rewrite the iteration of } z_i(t) \text{ for } i \in V \text{ in compact form as}\]

\[z(t + 1) = S_C z(t) + g(t).\]

Consider \(t \) steps of iteration

\[z(t) = S_C^t z(0) + \sum_{r=0}^{t-1} S_C^{t-1-r} g(r).\]

\[\text{We estimate the second part on the right-hand side of (11). By construction, it holds } \sum_{j \in V} ||S_G^{t-1} g|| = 1. \text{ Let } m_g(t) := \max_{i \in V} \{||g_i(t)||\} \text{ be the } \infty \text{-norm of } g(t). \text{ We then have}\]

\[||S_C^{t-1} g||_1 \leq m_g(r), \quad i \in V, \quad r = 0, 1, \ldots\]

\[\text{Hence,}\]

\[\sum_{r=0}^{t-1} m_g(r), \quad i \in V.\]

\[\text{By (1) we know that if } 0 < x_i(t) < 1 \text{ for all } i \in V, 0 < x_i(t+1) < 1 \text{ hold for all } i \in V. \text{ Therefore, we have } z_i(t) \in (-\frac{1}{2}, \frac{1}{2}) \text{ and } u_i(t) < \frac{1}{2} \text{ for all } i \in V \text{ and } t \geq 0, \text{ which indicates that}\]

\[0 < \left(\frac{1}{2} + z_i(t)\right)^b \left(\frac{1}{2} + \frac{u_i(t)}{d_i}\right) + \left(1 - z_i(t)\right)^b \left(\frac{1}{2} - \frac{u_i(t)}{d_i}\right) < l_i(t) \]

\[\text{holds for all } b \in (0, 1). \text{ Thus,}\]

\[\left(\frac{1}{2} + \frac{u_i(t)}{d_i}\right) \frac{1}{d_i} \frac{1}{d_i} \left(\frac{1}{2} + z_i(t)\right)^b + \left(1 - z_i(t)\right)^b \left(\frac{1}{2} - \frac{u_i(t)}{d_i}\right) < l_i(t) \]

\[\text{The right hand side is less than 1, since the numerator is no larger than the second part of the denominator if } u_i(t) \leq -z_i(t) \text{ and less than the first part of the denominator if } u_i(t) > -z_i(t). \text{ Recalling the definition of } g_i(t), \text{ we have}\]

\[|g_i(t)| < \left(\frac{1}{2} + z_i(t)^b - \frac{1}{2} - z_i(t)^b\right)\]

\[< \left(\frac{1}{2} + m_i(t)^b - \frac{1}{2} - m_i(t)^b\right).\]
By the arbitrariness of $i$,
\[ m_y(t) < \left( \frac{1}{2} + m_x(t) \right)^b - \left( \frac{1}{2} - m_x(t) \right)^b, \quad t \geq 0. \quad (15) \]
Moreover, from (10) we have
\[ m_z(t + 1) \leq m_z(t) + m_y(t). \quad (16) \]

Recalling the definition of $\mu$ in (3), by Lemma 1 (given below), we have
\[ (\frac{1}{2} + m_z(0) + \delta)^b - (\frac{1}{2} - m_z(0) - \delta)^b < \mu b \]
for all $b \in (0, 1)$. By the definition of $\hat{b}$,
\[ (\frac{1}{2} + m_z(0) + \hat{\delta})^b - (\frac{1}{2} - m_z(0) - \hat{\delta})^b < \hat{\delta}/\hat{\ell} \quad (17) \]
for all $0 < b < \hat{b}$. Combining (15) and (17), we have
\[ m_y(1) < \left( \frac{1}{2} + m_z(0) \right)^b - \left( \frac{1}{2} - m_z(0) \right)^b < \hat{\delta}/\hat{\ell}. \]
By (16), we obtain
\[ m_z(1) < m_z(0) + \hat{\delta}/\hat{\ell} < m_z(0) + \hat{\delta}. \]
Again by (15) and (17), it holds $m_y(2) < \hat{\delta}/\hat{\ell}$. This process can be iterated. We then have $m_z(t) \leq m_z(0) + (t/\hat{\ell})\hat{\delta}$ and $m_y(t) \leq \hat{\delta}/\hat{\ell}$ for all $t \leq \hat{\ell}$. Therefore,
\[ \|S_{G}^{t-1}r g(r)\|_1 \leq \sum_{r=0}^{t-1} m_y(r) \leq \hat{\delta} < \min_{x \in (0, \hat{\ell})} \{ \|S_{G}^{t}z(0)\|_1/2 \}. \quad (18) \]
Thus, if $b < \hat{b}$, we obtain
\[ z_i(\hat{\ell}) \leq \left[ \frac{S_{G}^{\hat{\ell}}z(0)i}{2} \right]_1 < 0, \quad i \in V_1; \quad z_i(\hat{\ell}) \geq \left[ \frac{S_{G}^{\hat{\ell}}z(0)i}{2} \right]_1 > 0, \quad i \in V_2. \]
This also means that $x_i(\hat{\ell}) < \frac{1}{2}$ for all $i \in V_1$, and $x_i(\hat{\ell}) > \frac{1}{2}$ for all $i \in V_2$. Then Proposition 1 can be applied, generating the desired conclusion.

Lemma 1 It holds
\[ \left( \frac{1}{2} + a \right)^x - \left( \frac{1}{2} - a \right)^x < \log(\frac{\frac{1}{2} + a}{\frac{1}{2} - a})x \]
for all $a \in [0, \frac{1}{2})$ and $x \in (0, 1)$.

Proof: Let
\[ f(x) = \left( \frac{1}{2} + a \right)^x - \left( \frac{1}{2} - a \right)^x - \log(\frac{\frac{1}{2} + a}{\frac{1}{2} - a})x \]
for $x \in (0, 1)$. At first, we have $f(0) = 0$ and $f(1) = 2a - \log(\frac{\frac{1}{2} + a}{\frac{1}{2} - a})$. Let $q(a) = 2a - \log(\frac{\frac{1}{2} + a}{\frac{1}{2} - a})$ for $a \in [0, \frac{1}{2})$. It holds that $q(0) = 0$ and $q(a') < 2 - \frac{3a^2}{3a^2 - 1} < 0$. Therefore, $q(a) \leq 0$ for $a \in [0, \frac{1}{2})$, which gives $f(1) \leq 0$.

Secondly, we have
\[ f(x)' = \left( \frac{1}{2} + a \right)^x \log(\frac{1}{2} + a) - \left( \frac{1}{2} - a \right)^x \log(\frac{1}{2} - a) - \log(\frac{\frac{1}{2} + a}{\frac{1}{2} - a}) \]
and
\[ f(x)'' = \left( \frac{1}{2} + a \right)^x [\log(\frac{1}{2} + a)]^2 - \left( \frac{1}{2} - a \right)^x [\log(\frac{1}{2} - a)]^2. \]
f(x)'' < 0 is equal to $c$ for some $c$ that depends on $a$. Observe $f(0)' = 0$. We know that $f(x)' < 0$ for $x \in (0, 1)$ or $f(x)' < 0$ for $(0, c')$ and $f(x') > 0$ for $x \in (c', 1)$ ($c'$ is some number between 0 and 1). Therefore, either: (i) $f(x)$ is strictly decreasing; (ii) $f(x)$ is strictly decreasing over $(0, c')$ and increasing over $(c', 1)$. No matter which case, as $f(0) = 0$ and $f(1) \leq 0$. $f(x) < 0$ hold for $x \in (0, 1)$. We then completed the proof.

APPENDIX II
PROOF OF THEOREM 2
To prove Theorem 2, we introduce first a lemma, which is a simple extension of Example 9.1 in [40].
Consider a system
\[ z_{k+1} = Az_k + g(z_k), \quad z_k \in \mathbb{R}^n, \quad (19) \]
where $g$ is a map from $\mathbb{R}^n$ to $\mathbb{R}^n$ and $A$ is an $n$-order matrix with all the eigenvalues having magnitudes less than 1. By Theorem 5.D5 in [32], there exists a positive definite matrix $P$ such that $P - A^TPA = I$.

Lemma 2 For the system (19), assume that there exist $d, \bar{d}, \epsilon > 0$ such that

- If $z \in B(d)$, $\|g(z)\| \leq \epsilon \|z\|$ holds, with
\[ \epsilon < \frac{1}{2\sqrt{\|P\|\|A\|^2 + 1}. \quad (20) \]
- $\bar{d} \leq d$. Once the trajectory enters in $B(\bar{d})$, it will stay in $B(d)$.

If $z_0 \in B(\bar{d})$, we have $\lim_{k \to \infty} z_k = 0$.

Proof: Denote $g_k = g(z_k)$. Construct a candidate Lyapunov function as $v(z) := z^TPz$. If $z_k \in B(d)$, then
\[ v(z_{k+1}) - v(z_k) = -\|z_k\|^2 + 2\langle P g_k, A z_k \rangle + \|g_k\|^2 \]
\[ \leq -\epsilon \|P\|\|A\|^2 - \epsilon^2 \|z_k\|^2. \quad (21) \]
By (20), we obtain $1 - \epsilon \|P\|\|A\|^2 - \epsilon^2 > 0$, which gives $v(z_{k+1}) \leq v(z_k)$. Further, we have
\[ v(z_{k+1}) \leq (1 - \epsilon \|P\|\|A\|^2 - \epsilon^2) v(z_k). \]
This derivation can be iterated, since the trajectory will not go out of $B(d)$. Hence we have
\[ v(z_t) \leq (1 - \epsilon \|P\|\|A\|^2 - \epsilon^2)^{t-k} v(z_k), \quad t \geq k. \]
Let $t \to \infty$, then $v(z_t) \to 0$. Therefore, $\lim_{t \to \infty} z_t = 0$. ■

Proof of Theorem 2. When $b = 0$, the conclusion will hold by Theorem 2 in [10].
Suppose $0 < b < 1$. We follow the notation of $g(t)$ in the proof of Theorem 1. At first, we verify all conditions in Lemma 2:
- From (10) and (14),
\[ z(t + 1) = S_G z(t) + g(t), \]
where

$$|g_i(t)| < \left| \frac{1}{2} + z_i(t) \right|^b - \left( \frac{1}{2} - z_i(t) \right)^b. $$

We can prove that if $z \in (-d, d)$ with $d = \sqrt{\frac{1}{2} \ell}$,

$$\left| \frac{1}{2} + z \right|^b - \left( \frac{1}{2} - z \right)^b < 5b|z|, \quad b \in (0, 1). \quad (22)$$

In fact, let $f(z) = \left( \frac{1}{2} + z \right)^b - \left( \frac{1}{2} - z \right)^b - 5bz$. Taking the first derivative of $f$ and by simple calculations, we have $f'(z) < 0$ for $z \in [0, \sqrt{\frac{1}{2} \ell}]$. In view of $f(0) = 0$, (22) is proved. Therefore, if $z(t) \in B(d)$, it holds $\|g(t)\| < 5b|z(t)|$.

As in the proof of Lemma 2, let $P$ be the only positive definite matrix satisfying $P - S_G^TPS_G = I$ and define the candidate Lyapunov function $v(\cdot)$ as $v(z) := z^T Pz$. By Theorem 5.6 in [32], the explicit form of $P$ is $P = \sum_{r=0}^{\infty} (S_G')^r (S_G')^r$. Recalling $d = \sqrt{\frac{1}{2} \ell}$ and $p_0 > \|P\|$, let

$$\tilde{d} = \sqrt{\frac{1}{2p_0} < \frac{\lambda_{\min}(P)}{\lambda_{\max}(P)}. d.}$$

The right inequality holds because of $\lambda_{\min}(P) > 1$ and $\lambda_{\max}(P) = \|P\|$ (the equality holds since $P$ is symmetric).

We then prove that once the trajectory of system (10) enters in $B(\tilde{d})$, it will stay in $B(d)$. Suppose $z(k) \in B(\tilde{d})$ for some $k \geq 0$. We have $\|g(k)\| < 5b|z(k)|$, and

$$v(z(k)) \leq \lambda_{\max}(P) \|z(k)\|^2 < \lambda_{\min}(P) d^2.$$ 

According to the proof of Lemma 2, if $b$ satisfies

$$5b < \frac{1}{2 \sqrt{p_0} \|S_G\| + 1}, \quad (23)$$

it holds $v(z(k + 1)) \leq v(z(k))$. We then have $\lambda_{\min}(P) \|z(k + 1)\|^2 \leq v(z(k + 1)) \leq v(z(k)) < \lambda_{\min}(P) d^2$, yielding $\|z(k + 1)\| < d$. This again gives that

$$v(z(k + 2)) \leq v(z(k + 1)) \leq v(z(k)) < \lambda_{\min}(P) d^2,$$

yielding $\|z(k + 2)\| < d$. The process can be iterated. Therefore, $\|z(t)\| < d$ for $t \geq k$.

Then Lemma 2 can be applied, i.e., if $z(k) \in B(\tilde{d})$ for some $k \geq 0$. Notice that we only require (23) to hold here.

Next we prove that when $b$ is small enough, the trajectory of system (10) will enter in $B(\tilde{d})$ at some time $k$. Take the same definitions of $m_g(t)$ as in the proof of Theorem 1. Note that in the derivation of (14)-(17), we only use the definition of $\hat{\mu}$ and $\tilde{\delta} < \frac{1}{2} - m_z(0), \tilde{b} < \frac{\hat{\mu}}{\tilde{\delta}}$. Therefore, if $b < \frac{\hat{\mu}}{\tilde{\delta}}$, (14)-(17) also hold here, with $\hat{\mu}, \ell, \delta, \tilde{b}$ substituted by $\tilde{\mu}, \ell, \hat{\delta}, \tilde{b}$. Further, noting the similarity between $\tilde{\delta}$ and $\hat{\delta}$, we get an inequality similar to (18), i.e.,

$$|\sum_{r=0}^{\tilde{l}-1} S_G^{-1-r} g(r)| \leq \sum_{r=0}^{\tilde{l}-1} m_g(r) \leq \tilde{\delta} < \frac{\ell}{\tilde{d}}. \quad (24)$$

Recalling the definition of $\tilde{\ell}$, we have $S_G^\ell \zeta(0) \in B(\tilde{d}/2).$ Therefore, if $b < \frac{\hat{\mu}}{\tilde{\delta}}$, by (11), we have $z(\tilde{\ell}) \in B(\tilde{d})$. Therefore, choosing $b$ such that $b < \frac{\hat{\mu}}{\tilde{\delta}}$ and (23) hold, we will have $\lim_{t \to \infty} \|z(t)\| \to 0$. By the definition of $\tilde{b}$, we know that $\lim_{t \to \infty} \|z(t)\| \to 0$ if $b < \tilde{b}$. Then the proof is completed, since $\lim_{t \to \infty} \|z(t)\| \to 0$ is equivalent to $\lim_{t \to \infty} x_i(t) = \frac{1}{2}, i \in V$. \hfill \blacksquare

**APPENDIX III**

**Proof of Proposition 2 and Theorem 3**

**Proof of Proposition 2.** By the definition of $\lambda_i$, the eigenvalues of $J_{hs}$ are $b + \lambda_i$. According to Section 5.7 in [44], $(\frac{1}{2}, \ldots, \frac{1}{2})^T$ is stable if all eigenvalues of $J_{hs}$ are in the unit circle, i.e., $|b + \lambda_i| < 1$ for all $i \in V$. This is equivalent to $b < b_0$. On the other hand, $(\frac{1}{2}, \ldots, \frac{1}{2})^T$ is unstable if one of the eigenvalues of $J_{hs}$ is out of the unit circle, which is equivalent to $b > b_0$. We then completed the proof. \hfill \blacksquare

**Proof of Theorem 3.** Following the notations in [41], we use $f_0^u$ and $f_0^v$ to denote the derivatives of $f$ with respect to $x$ and $x$ at $(\frac{1}{2}, \ldots, \frac{1}{2})$ respectively. Define $f_{xx}^0$, the second-order derivative of $f$ with respect to $x$ at $(\frac{1}{2}, \ldots, \frac{1}{2})$, and $f_{xx}^0, f_{xx}^0$ are the similarly defined second-order derivatives. We remind that $f_0^u, f_0^v$ are vectors, while $f_{xx}^0, f_{xx}^0$ are matrices and $f_{xx}^0$ is a tensor.

By Assumption 3, the maximum eigenvalue 0 of $f_0^u$ is simple, since $f_0^u = -\lambda_{\max}(S_G) I + S_G$. Consider the left eigenvector and the right eigenvector with respect to the eigenvalue 0, denoted by $u_i, u_i$, rescaled such that $u_i^T u_i = 1$. According to Theorem 5.7 in [41], the sufficient conditions for $((\frac{1}{2}, \ldots, \frac{1}{2}), 1 - \lambda_{\max}(S_G))$ to be a bifurcation point are:

1. $f_0^u \in$ range$(f_{xx}^0)$,
2. $a_2 - a_1 a_4 > 0$, with
   $$a_1 := u_i^T f_{xx}^0 u_i, \quad a_2 := u_i^T (f_{xx}^0 v + f_{xx}^0 u_i),$$
   $$a_3 := u_i^T (f_{xx}^0 v v + 2 f_{xx}^0 v + f_{xx}^0),$$
   where $v$ is the vector satisfying $f_{xx}^0 v + f_{xx}^0 = 0, a_4 v = 0$.

In our case, we obtain

$$f_0^u = 0, \quad f_{xx}^0 = 0, \quad f_{xx}^0 = I,$$

yielding $a_1 = 0, a_2 = 1$. Therefore, the conditions (1), (2) are both satisfied for any $v$, which completes the proof. \hfill \blacksquare

**APPENDIX IV**

**Proof of Theorem 4 and Corollary 1**

**Proof of Theorem 4.** The conclusion obviously holds for $x_i(0) = 0$ and $x_i(0) = 1$. Thus, we only need to consider node $i$ with $x_i(0) \in (0, 1)$. By (1), we have

$$\frac{1 - x_i(t + 1)}{x_i(t + 1)} = \left( \frac{1 - x_i(t)}{x_i(t)} \right)^b \frac{1 - s_i(t)}{s_i(t)}. \quad (25)$$

For the case $x_i(0) > \frac{1}{2}, \frac{1 - x_i(0)}{x_i(0)} < 1$ holds. For each $j \in N_i^o$, we have $y_j < \frac{1}{2}$. By the definition of $s_i(0)$, we obtain

$$s_i(0) \geq \sum_{j \in N_i^o} W_{ij} y_j(0),$$

which leads to

$$\frac{1 - s_i(0)}{s_i(0)} \leq \frac{1 - \sum_{j \in N_i^o} W_{ij} y_j(0)}{\sum_{j \in N_i^o} W_{ij} y_j(0)}.$$
By the definition of $b^*$, we have

$$b^* - 1 \geq \log \left( \frac{1 - \sum_{j \in N_i} W_{ij} y_j^i(0)}{\sum_{j \in N_i} W_{ij} y_j^i(0)} \right) / \log \left( \frac{x_i(0)}{1 - x_i(0)} \right),$$

i.e.,

$$\left( \frac{1 - x_i(0)}{x_i(0)} \right) b^* - 1 - \frac{1 - \sum_{j \in N_i} W_{ij} y_j^i(0)}{\sum_{j \in N_i} W_{ij} y_j^i(0)} \leq 1. \quad (26)$$

If $b > b^*$, by (25),

$$\frac{1 - x_i(1)}{x_i(1)} < \left( \frac{1 - x_i(0)}{x_i(0)} \right) b^* - 1 - \frac{1 - \sum_{j \in N_i} W_{ij} y_j^i(0)}{\sum_{j \in N_i} W_{ij} y_j^i(0)} \leq \frac{1 - x_i(0)}{x_i(0)}, \quad (27)$$

which means $x_i(1) > x_i(0) > \frac{1}{2}$.

For the case $x_i(0) < \frac{1}{2}$, i.e., $\frac{1 - x_i(0)}{x_i(0)} > 1$ holds. We have $y_j^i < \frac{1}{2}$ for $j \in N_i$, and $s_i(0) \leq 1 - \sum_{j \in N_i} W_{ij} (1 - y_j^i(0))$, which leads to

$$1 - s_i(0) \geq \frac{\sum_{j \in N_i} W_{ij} (1 - y_j^i(0))}{1 - \sum_{j \in N_i} W_{ij} (1 - y_j^i(0))}. \quad (28)$$

By the definition of $b^*$, we have

$$b^* - 1 \geq \log \left( \frac{1 - \sum_{j \in N_i^0} W_{ij} (1 - y_j^i(0))}{\sum_{j \in N_i^0} W_{ij} (1 - y_j^i(0))} \right) / \log \left( \frac{1 - x_i(0)}{x_i(0)} \right),$$

If $b > b^*$, similarly to (26)-(27), we have $\frac{1 - x_i(1)}{x_i(1)} > \frac{1 - s_i(0)}{s_i(0)}$, which means $x_i(1) < x_i(0) < \frac{1}{2}$.

Take an arbitrary $b$ such that $b > b^*$. By the discussion above, we know: (1) $x_i(1) < x_i(0)$ if $x_i(0) < \frac{1}{2}$, (2) $x_i(1) > x_i(0)$ if $x_i(0) > \frac{1}{2}$. Accordingly, for $j \in N_i^0$, we have: (1) $y_j^i(1) < y_j^i(0) < \frac{1}{2}$ if $x_i(0) < \frac{1}{2}$; (2) $y_j^i(1) > y_j^i(0) > \frac{1}{2}$ if $x_i(0) > \frac{1}{2}$.

Consider $i$ with $x_i(0) > \frac{1}{2}$. It holds $\sum_{j \in N_i^0} W_{ij} y_j^i(1) > \sum_{j \in N_i^0} W_{ij} y_j^i(0)$. By (25) and (27), we obtain

$$\frac{1 - x_i(0)}{x_i(1)} < \left( \frac{1 - x_i(0)}{x_i(1)} \right) b^* - 1 - \frac{1 - \sum_{j \in N_i^0} W_{ij} y_j^i(0)}{\sum_{j \in N_i^0} W_{ij} y_j^i(0)} \leq \frac{1 - x_i(0)}{x_i(1)}, \quad (29)$$

which means $x_i(2) > x_i(1) > \frac{1}{2}$. Similarly to (28), for $i$ with $x_i(0) < \frac{1}{2}$, we have

$$\frac{1 - x_i(2)}{x_i(1)} \cdots > \frac{1 - x_i(1)}{x_i(1)}, \quad (29)$$

which means $x_i(2) < x_i(1) < \frac{1}{2}$. This process can be iterated. We then obtain $x_i(t + 1) < x_i(t)$ for all $t > 0$ if $x_i(0) < \frac{1}{2}$, and $x_i(t + 1) > x_i(t)$ for all $t > 0$ if $x_i(0) > \frac{1}{2}$. Furthermore, by (28), if $x_i(0) > \frac{1}{2}$, it holds

$$\frac{1 - x_i(t)}{x_i(t)} < \cdots < \left( \frac{1 - x_i(0)}{x_i(0)} \right)^{(b^* - 1)^t},$$

which yields $\lim_{t \to \infty} x_i(t) = 1$. Similarly, by (29), if $x_i(0) < \frac{1}{2}$, we have

$$\frac{1 - x_i(t)}{x_i(t)} > \left( \frac{1 - x_i(0)}{x_i(0)} \right)^{(b^* - 1)^t},$$

which yields $\lim_{t \to \infty} x_i(t) = 0$. We then completed the proof.

**Proof of Corollary 1.** Observing

$$\sum_{j \in N_i^0} W_{ij} y_j^i(0) \leq w^*/2 \quad \text{if} \quad x_i(0) \leq \frac{1}{2},$$

and

$$\min_{x_i(0) < \frac{1}{2}} \frac{1 - x_i(0)}{x_i(0)}, \min_{x_i(0) > \frac{1}{2}} \frac{x_i(0)}{1 - x_i(0)} = \frac{1 - m^*}{m^*},$$

we have

$$b^* < \log \left( \frac{2 - w^*}{w^*} \right) \log \left( \frac{1 - m^*}{m^*} \right).$$

By Theorem 4, (6) holds if

$$b > \log \left( \frac{2 - w^*}{w^*} \right) \log \left( \frac{1 - m^*}{m^*} \right) + 1. \quad (30)$$

Rearranging (30), we get the desired conclusion.

**Proof of Corollary 2.** Given $\epsilon > 0$, take $\sigma$ as

$$\sigma = \min \left\{ \epsilon, \frac{(w^*)^{1 - \epsilon}}{2 - w^* + (w^*)^{1 - \epsilon}} \right\}.$$

If $\|x(0) - x^*\| < \sigma$, we have $N_i^0 \neq \emptyset, i \in V$ and

$$m^* = \|x(0) - x^*\|_{\infty} \leq \|x(0) - x^*\| < \sigma.$$

Then by Corollary 1, we get the desired conclusion.

**Appendix V**

**Proof of Theorem 5**

Let $(x_1, \ldots, x_n)^T$ be an interior equilibrium point and $s_i = \sum_{j \in N_i} x_j + \sum_{j \in N_i} (1 - x_j)$. Then,

$$x_i = x_i b s_i / [x_i b s_i + (1 - x_i)^b (d_i - s_i)],$$

holds for all $i \in V$. Rearranging this equation, we obtain

$$s_i / d_i = (1 - x_i)^{b - 1} / [x_i b^{b - 1} + (1 - x_i)^{b - 1}], \quad i \in V. \quad (31)$$

Note that $G$ is structurally balanced under the partition $V_1 = \{1, \ldots, n - 1\}, V_2 = \{n\}$. By the transformation (7), the dynamics of $y_i(t)$ is (1) over a graph with no negative interactions. Observe that $(\frac{1}{2}, \ldots, \frac{1}{2})$ keeps the same under the transformation. Now with a little abuse of symbols, for the transformed system, we still use $(x_1, \ldots, x_n)^T$ to represent an interior equilibrium point, and then $s_i = \sum_{j \neq i} x_j$.

By (31),

$$\frac{\sum_{j \neq i} x_j}{n - 1} = \frac{(1 - x_i)^{b - 1}}{x_i^{b - 1} + (1 - x_i)^{b - 1}}, \quad i \in V. \quad (32)$$

Suppose $b \leq 1$ or $b = 2$. Note that (31) has the same form as Eq. C.1 in [26]. The analysis and conclusion of Theorem 1 in [26] then hold, which gives $x_i = \frac{1}{2}$ for all $i \in V$. 


• Suppose $b > 2$. Denote $x_{\min}, x_{\max}$ as the minimum and maximum of \{x_1, \ldots, x_n\}. We prove $x_{\min} = x_{\max} = \frac{1}{2}$ by contradiction. Assume $x_{\max} > \frac{1}{2}$. For $x \in (0, 1)$, we have

$$
\frac{(1 - x)^{b - 1}}{x^{b - 1} + (1 - x)^{b - 1}} = x \frac{(1 - x)(1 - x)^{b - 2} - x^{b - 2}}{x^{b - 1} + (1 - x)^{b - 1}}. 
$$

(33)

Therefore,

$$
(1 - x)^{b - 1}/[x^{b - 1} + (1 - x)^{b - 1}] < 1 - x \text{ if } x < 1/2. 
$$

(34)

Taking $x_i = x_{\max}$ into (32) and applying (34), we know that

$$
x_{\min} \leq \sum_{j=1}^{n-1} x_j/(n - 1) < 1 - x_{\max} < 1/2. 
$$

(35)

Taking $x_i = x_{\min}$ into (32) and applying (34) again, we get

$$
\frac{1}{2} < 1 - x_{\min} < \sum_{j \neq i} x_j/n - 1 \leq x_{\max}. 
$$

(36)

Combining (35) with (36), we get a contradiction. Therefore, $x_{\max} \leq \frac{1}{2}$. Symmetrically, we can prove $x_{\min} \geq \frac{1}{2}$. Therefore, $x_{\min} = x_{\max} = \frac{1}{2}$. That is, $x_1 = \cdots = x_{n-1} = \frac{1}{2}$.

By Proposition 1, $(\frac{1}{2}, \ldots, \frac{1}{2})$ is unstable, which completes the proof.

**APPENDIX VI**

**PROOF OF THEOREM 6**

Let $(x_1, \ldots, x_n)$ be an equilibrium point. Note that (31) is independent of the underlying network structure. We then have

$$
\frac{\sum_{j=1}^{n-1} x_j}{n - 1} = \frac{(1 - x_n)^{b - 1}}{x^{b - 1} + (1 - x_n)^{b - 1}}, 
$$

(37a)

$$
\sum_{j \neq i, n} (1 - x_j) + x_n/n - 1 = \frac{(1 - x_i)^{b - 1}}{x^{b - 1} + (1 - x_i)^{b - 1}}, i = 1, \ldots, n - 1. 
$$

(37b)

Taking (37a) into (37b) and rearranging the equation, we have

$$
x_i^{b - 1}/(x^{b - 1} + (1 - x_i)^{b - 1}) + x_i/n - 1 = \frac{(1 - x_n)^{b - 1}}{x^{b - 1} + (1 - x_n)^{b - 1}} + 1/n - 1. 
$$

(38)

Note that (38) holds for all $i = 1, \ldots, n - 1$. Define

$$
f_b(x) := \frac{x^{b - 1}}{x^{b - 1} + (1 - x)^{b - 1}} + \frac{x}{n - 1}. 
$$

(39)

We have $f_b(x_1) = \cdots = f_b(x_{n-1})$. Calculating the derivative of $f_b$, it is easy to get $\frac{d}{dx}f_b(x) > 0$ for $x \in (0, 1)$ and $b \geq 1$. Therefore, $x_1 = \cdots = x_{n-1}$. By (37a), we have

$$
x_i = (1 - x_n)^{b - 1}/[x^{b - 1} + (1 - x_n)^{b - 1}]. 
$$

(40)

Combining (38) with (40),

$$
\frac{(1 - x_n)^{b - 1}/x^{b - 1} + (1 - x_n)^{b - 1}}{x^{b - 1} + (1 - x_n)^{b - 1}} = \frac{n - 2}{n - 1} \frac{(1 - x_n)^{b - 1}}{x^{b - 1} + (1 - x_n)^{b - 1}} + \frac{1 - x_n}{n - 1}. 
$$

(41)

(i) $1 \leq b < 2$. We prove $x_n = \frac{1}{2}$ by contradiction. Assume $x_n < \frac{1}{2}$ first. To get a contradiction, we just need to show

$$
\frac{(1 - x_n)^{b - 1}/x^{b - 1} + (1 - x_n)^{b - 1}}{x^{b - 1} + (1 - x_n)^{b - 1}} \leq \frac{(1 - x_n)^{b - 1}}{x^{b - 1} + (1 - x_n)^{b - 1}} < 1 - x_n, 
$$

since (41) requires that the value of $\frac{(1 - x_n)^{b - 1}/x^{b - 1} + (1 - x_n)^{b - 1}}{x^{b - 1} + (1 - x_n)^{b - 1}}$ to be between $\frac{1}{x^{b - 1} + (1 - x_n)^{b - 1}}$ and $1 - x_n$. In fact, by (33), the right inequality of (42) holds. As for the left inequality, it is equivalent to

$$
[x_n/(1 - x_n)]^{(b - 1)/b} \geq [x_n/(1 - x_n)]^{b - 1}. 
$$

This inequality certainly holds, since $b - 1 \in (0, 1)$ and $x_n < \frac{1}{2}$. Therefore, (42) holds and gives the contradiction. As a consequence, the previous assumption does not hold, and we have $x_n \geq \frac{1}{2}$. Assume $x_n > \frac{1}{2}$. By similar analysis, we are able to deduce (42) with the signs of inequality flipped, which gives the contradiction. Thus, $x_n \leq \frac{1}{2}, x_n = \frac{1}{2}$ is then obtained, which also yields $x_1 = \cdots = x_{n-1} = \frac{1}{2}$ by (40).

(ii) $b > 2$. We have

$$
[x_n/(1 - x_n)]^{(b - 1)/b} > [x_n/(1 - x_n)]^{b - 1} \text{ if } x_n > 1/2. 
$$

(43)

Combining (43) and (34), we obtain

$$
\frac{(1 - x_n)^{b - 1}/x^{b - 1} + (1 - x_n)^{b - 1}}{x^{b - 1} + (1 - x_n)^{b - 1}} < \frac{(1 - x_n)^{b - 1}}{x^{b - 1} + (1 - x_n)^{b - 1}} \geq 1 - x_n 
$$

for $x_n > 1/2$. As a consequence, by the proof of contradiction as in (i), we can get $x_i = \frac{1}{2}, i \in V$.

(iii) $b = 2$. By (40), we have $x_1 = \cdots = x_{n-1} = 1 - x_n$. Therefore, the interior equilibria set is $E_{int} = \{(a, \ldots, a, 1 - a)|a \in (0, 1)\}$.

Next, we will prove that $(\frac{1}{2}, \ldots, \frac{1}{2})$ is unstable for the signed biased assimilation model (1) with $b \leq 1$. At first, we have

$$
S_G = \frac{1}{n-1} \begin{pmatrix} 
0 & -1 & \cdots & -1 & 1 \\
-1 & 0 & \cdots & -1 & 1 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
-1 & -1 & \cdots & 0 & 1 \\
1 & 1 & \cdots & 1 & 0 
\end{pmatrix}.
$$

It is known from the standard linear algebra that $S_G$ only has two eigenvalues, i.e., $-1$ and $\frac{1}{n-1}$. Therefore, $\lambda_{\max}(J_{b_i}) = b + \frac{1}{n-1} > 1$ for $b \geq 1$, where $J_{b_i}$ is the Jacobian matrix at $(\frac{1}{2}, \ldots, \frac{1}{2})^T$ which is defined in (4). By Proposition 2, $(\frac{1}{2}, \ldots, \frac{1}{2})^T$ is unstable for $b \geq 1$. We then completed the proof.

**APPENDIX VII**

**PROOF OF THEOREM 7**

For convenience, let node $kn + i$ be identified with node $i \in V$ for all $k = 0, 1, 2, \ldots$. Before proving Theorem 7, we introduce the following lemma.
Lemma 3 Let $x_i$ be the value kept by node $i$, $i \in V$. Let $p, q, r$ be nonzero real numbers. Suppose that $|\frac{-p^{2}\sqrt{q}^2\sqrt{p}-4qr}{2q}| \neq 1$ or 0 and $n \geq 3$. If the values satisfy
\[ px_i + q x_{i+1} + r x_{i-1} = \frac{(p+q+r)}{2} \tag{44} \]
for all $i \in V$, $x_i = \frac{1}{2}$ holds for all $i \in V$.

Proof: Using the method of undetermined coefficients, we are going to find $k_1, k_2, k_3 \in \mathbb{C}$ such that the following equation holds for all $i \in V$:
\[ q(x_{i+1} + k_1 x_i + k_2) = k_3(x_i + k_1 x_{i-1} + k_2) \tag{45} \]
Compared to (44), we get two groups of solutions: (i) $k_1 = \frac{p+\sqrt{p^2-4qr}}{2q}, k_2 = \frac{-p-2q+\sqrt{p^2-4qr}}{4q}, k_3 = \frac{-p+2q+\sqrt{p^2-4qr}}{4q};$ (ii) $k_1 = \frac{p-\sqrt{p^2-4qr}}{2q}, k_2 = \frac{-p+2q+\sqrt{p^2-4qr}}{4q}, k_3 = \frac{-p-2q+\sqrt{p^2-4qr}}{4q}.$

For both solutions, it holds $\frac{k_a q}{q} \neq 1$ or 0 according to the given condition. From (45), by iteration, we have
\[ x_i + k_1 x_{i-1} + k_2 = x_{i+n} + k_1 x_{i+n-1} + k_2 = (k_3/q)^n (x_i + k_1 x_{i-1} + k_2), i \in V. \]
Therefore, $x_i + k_1 x_{i-1} + k_2 = 0$ for all $i \in V$, which gives
\[ \begin{cases} x_i + \frac{p+\sqrt{p^2-4qr}}{2q} x_{i-1} + \frac{-p-2q+\sqrt{p^2-4qr}}{4q} = 0 \\ x_i + \frac{p-\sqrt{p^2-4qr}}{2q} x_{i-1} + \frac{-p+2q+\sqrt{p^2-4qr}}{4q} = 0 \end{cases} \tag{46} \]
The only solution to (46) is $x_i = x_{i-1} = \frac{1}{2}$. We then conclude that $x_1 = \cdots = x_n = \frac{1}{2}$.

Proof of Theorem 7. Let $(x_1, \ldots, x_n)$ be an interior equilibrium point. By (31), we have
\[ x_i - x_{i+1} + \sum_{j \neq i, i+1} (1-x_j) \frac{n-1}{n} = \frac{(1-x_i)^{b-1}}{x_i^{b-1} + (1-x_i)^{b-1}} \tag{47} \]
for all $i \in V$. We will discuss the cases with $b=1, b=2$ and $b > 2$ separately, and then prove instability of $(\frac{1}{2}, \ldots, \frac{1}{2})$.

(i) $b = 1$. Considering (47) with $b = 1$, it holds
\[ [x_{i-1} + x_{i+1} + \sum_{j \neq i, i+1} (1-x_j)]/(n-1) = 1/2, \quad i \in V. \tag{48} \]
Summing from 1 to $n$, we obtain
\[ [n(n-3)-(n-5) \sum_{j=1}^{n} x_j]/(n-1) = n/2, \]
which yields $\sum_{j=1}^{n} x_j = \frac{n}{2}$. Here we use the condition $n > 5$. Hence by (48), we have
\[ x_i + 2x_{i-1} + 2x_{i+1} = 5/2, \quad i \in V, \]
in which the coefficients satisfy the conditions of Lemma 3. Therefore, $x_i = \frac{1}{2}$ for all $i \in V$.

(ii) $b = 2$. The right hand of (47) becomes $1 - x_i$. Summing (47) from 1 to $n$, it gives that $\sum_{j=1}^{n} x_j = \frac{n}{2}$, which inserted on (47) gives
\[ nx_i + 2x_{i+1} + 2x_{i-1} = n/2 + 2, \quad i \in V. \]

By Lemma 3, we have $x_i = \frac{1}{2}$ for all $i \in V$.

(3) $b > 2$. Let $x_{\max} = \max_{i \in V} x_i, x_{\min} = \min_{i \in V} x_i$. Similar to the proof of Theorem 5, we use contradiction. Suppose $x_{\max} \geq \frac{1}{2}$ first. Combining (34) with (47), we have
\[ x_{\min} \leq (1-x_{\max})^{b-1} / [x_{\max}^{b-1} + (1-x_{\max})^{b-1}] < 1 - x_{\max} < 1/2, \]
and then
\[ 1/2 < 1 - x_{\min} < (1-x_{\min})^{b-1} / [x_{\min}^{b-1} + (1-x_{\min})^{b-1}] \leq x_{\max}. \]
Combining the two inequalities above, we get a contradiction. Therefore, $x_{\max} \leq \frac{1}{2}$. By similar argument, we can prove $x_{\min} \geq \frac{1}{2}$. As a consequence, we have $x_{\max} = x_{\min} = \frac{1}{2}$, i.e., $x_1 = \cdots = x_n = \frac{1}{2}$. This completes our proof.

References


