Analysis of a Nonlinear Opinion Dynamics Model with Biased Assimilation ⋆

Weiguo Xia [a], Mengbin Ye [b,d], Ji Liu [c], Ming Cao [d], Xi-Ming Sun [a]

[a] *Key Laboratory of Intelligent Control and Optimization for Industrial Equipment of Ministry of Education and School of Control Science and Engineering, Dalian University of Technology, China (e-mails: {wgxiaseu,sunxm}@dlut.edu.cn).*

[b] *Optus-Curtin Centre of Excellence in Artificial Intelligence, Curtin University, Australia (e-mail: mengbin.ye@curtin.edu.au).*

[c] *Department of Electrical and Computer Engineering, Stony Brook University, USA (e-mail: ji.liu@stonybrook.edu).*

[d] *Faculty of Science and Engineering, ENTEG, University of Groningen, the Netherlands (e-mail: m.cao@rug.nl).*

Abstract

This paper analyzes a nonlinear opinion dynamics model which generalizes the DeGroot model by introducing a bias parameter for each individual. The original DeGroot model is recovered when the bias parameter is equal to zero. The magnitude of this parameter reflects an individual's degree of bias when assimilating new opinions, and depending on the magnitude, an individual is said to have weak, intermediate, and strong bias. The opinions of the individuals lie between 0 and 1. It is shown that for strongly connected networks, the equilibria with all elements equal identically to the extreme value 0 or 1 is locally exponentially stable, while the equilibrium with all elements equal to the neutral consensus value of $\frac{1}{2}$ is unstable. Regions of attraction for the extreme consensus equilibria are given. For the equilibrium consisting of both extreme values 0 and 1, which corresponds to opinion polarization according to the model, it is shown that the equilibrium is unstable for all strongly connected networks if individuals all have weak bias, becomes locally exponentially stable for complete and two-island networks if individuals all have strong bias, and its stability heavily depends on the network topology when individuals have intermediate bias. Analysis on star graphs and simulations show that additional equilibria may exist where individuals form clusters.

1 Introduction

There has been a persistent interest in theoretical sociology over the past decades in the modeling and study of opinion formation processes (Ajzen [2001]). Various models have been proposed to study how the opinions of an interconnected social group evolve and how limiting phenomena arise, including consensus, polarization, and clustering.

The French–DeGroot model (DeGroot [1974]) is probably the most well-known (referred to as DeGroot henceforth for simplicity); each individual repeatedly updates his/her opinion to be a weighted average of the opinions of his/her neighbors (perhaps including him/herself), reflecting the subconscious human cognitive capability of taking convex combinations when processing new information (Anderson [1981]). The opinions of the individuals will eventually reach a consensus as long as the interaction network satisfies some appropriate connectivity requirements. Over the years, a number of discrete- and continuous-time variants of the DeGroot model have been proposed and studied extensively. Notable among them include the Friedkin–Johnsen model (Friedkin and Johnsen [1990]), the Hegselmann–Krause model (Hegselmann and Krause [2002]), and the Altafini model (Altafini [2013]), For recent advances in the modelling of opinion dynamics on influence networks, see Anderson and Ye [2019], Proskurnikov and Tempo [2017].

The phenomenon of extremization, which refers to the tendency for a group of individuals to eventually reach

⋆ This paper was not presented at any IFAC meeting. A preliminary version of this paper was presented at the 57th IEEE Conference on Decision and Control (Xia et al. [2018]). This work was supported in part by the National Key R&D Program of China under Grant 2018YFB1700102, the National Natural Science Foundation of China under Grants 61973051, 61911530696, the Fundamental Research Funds for the Central Universities under Grant DUT19ZD103, the Youth Star of Dalian Science and Technology (2018RQ51), the Australian Research Council under grants DP160104500 and DP190100887, Optus Business, the European Research Council (ERC-CoG-771687), and the Netherlands Organization for Scientific Research (NWO-vidi-14134).

opinions that are more extreme than their initial inclination (and perhaps polarizing into two opposite camps), has become of increasing relevance in the modern age, and is the focus of research from several scientific communities (Dimock et al. [2014]). Several recent models have been introduced which can predict polarization or extremization (Duggins [2017], G. Shi et al. [2017], Mäs et al. [2014]), but typically attributes this phenomenon to antagonistic interactions that increase in strength as the difference in opinions between individuals grow, and only G. Shi et al. [2017] has provided analysis for its proposed model.

Biased assimilation is the phenomenon in social psychology in which individuals tend to process new information with a bias towards their current position, accepting confirming evidence while evaluating disconfirming evidence critically (Lord et al. [1979]). This can result in an individual developing a more extreme opinion when exposed to information from a confirming and disconfirming source (Munro et al. [2002], Taber and Lodge [2006]). A new generalization of the DeGroot model was recently proposed in Dandekar et al. [2013], where a bias parameter helps to capture the cognitive processes described in the preceding two sentences. For homophilous networks, it has been shown that under some specific conditions, polarization arises if the individuals are sufficiently biased and consensus is reached under some other specific initial opinions for a small bias parameter close to zero (Dandekar et al. [2013]). However, the situation that the system converges to consensus is rarely observed for other initial states.

In this paper, we focus on strongly connected networks and further examine the model proposed in Dandekar et al. [2013] to elucidate the role of biased assimilation in shaping opinion formation processes. The level of biased assimilation is captured by the scalar b_i for individual i, and assumed to be heterogeneous among the individuals. First, we provide a detailed, quantitative argument to illustrate how biased assimilation is captured in the model when an individual is presented with two opposing opinions. As a consequence, we are able to clearly illustrate how the magnitude of b_i determines whether individual i has a weak, intermediate, or strong intensity of biased assimilation. We then concentrate on equilibria with meaningful social interpretations, such as *extreme consensus* equilibria, the *neutral consensus* equilibrium, and *extreme polarization* equilibria. The role that b_i, viz. the intensity of the biased assimilation, plays in determining the (local) stability or instability of these equilibria is explicitly identified. For *extreme consensus* equilibria, broad regions of attraction are obtained, and we identify further equilibria for star networks. This contrasts the work of Dandekar et al. [2013], which identifies regions of convergence for the polarization equilibria but does not consider stability or other equilibria, and the work of Chen et al. [2019], where stability results are only established for special classes of network topologies. Detailed discussions are provided for the findings on each of the above types of equilibria, and social interpretations and implications are examined.

The rest of the paper is organized as follows. Section 2 introduces the biased opinion dynamics model. Section 3 analyzes the equilibria and their stability for the model, with the proofs given in Section 4. Section 5 provides several simulations to illustrate the rich set of possible dynamic behaviors, including some not covered in the analysis. Conclusions are drawn in Section 6.

Notation: For a positive integer N, let $\mathbf{1}_N$ and $\mathbf{0}_N$ denote the N-dimensional all-one vector and all-zero vector, respectively. Let $I_{N\times N}$ and $O_{N\times N}$ denote the $N \times N$ identity matrix and zero matrix, respectively. We will use the terms "individual" and "agent" interchangeably.

2 The Model For Opinion Dynamics With Bias Assimilation

Consider a group of N agents labeled by 1 to N. Each agent can receive information only from its neighbors. The neighbor relationships among the N agents are characterized by an N-node directed graph represented by $\mathbb{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} = \{1, \dots, N\}$ is the node set and $\mathcal{E}$ is the edge set. The graph is associated with a weight matrix $W = (w_{ij})_{N\times N}$ where the self-weight $w_{ii} \geq 0$, and for $j \neq i$, one has $(j, i) \in \mathcal{E}$ if and only if[1] $w_{ij} > 0$. Let $\mathcal{N}_i = \{j \in \mathcal{V} : (j, i) \in \mathcal{E}\}$ be the set of neighbors of agent i, representing the other agents j that have influence on i. Note that no self-loop is allowed in the graph $\mathbb{G}$ and therefore $i \notin \mathcal{N}_i$ and $(i, i) \notin \mathcal{E}$ for $i = 1, \dots, N$, but the self-weight w_{ii} can be positive.

Each agent i has a real-valued opinion $x_i(k)$, on a given issue being discussed, which may change over time k. At every discrete time instant $k = 0, 1, \dots$, each agent i updates its opinion by setting

$$x_i(k+1) = \frac{w_{ii}x_i(k) + (x_i(k))^{b_i} s_i(k)}{w_{ii} + (x_i(k))^{b_i} s_i(k) + (1 - x_i(k))^{b_i}(d_i - s_i(k))}, \tag{1}$$

where $d_i = \sum_{j\in\mathcal{N}_i} w_{ij}$, $s_i(k) = \sum_{j\in\mathcal{N}_i} w_{ij}x_j(k)$, and $w_{ij}, i, j \in \mathcal{V}$, are the elements in the weight matrix W representing the influence weights. The bias of agent i is captured by the parameter b_i, and is assumed to be nonnegative except for a special scenario considered in the sequel. Observe that on the right hand side of (1), the numerator is nonnegative and the denominator is greater than or equal to the numerator for any $x_i(k) \in [0, 1]$. Thus, $x_i(0) \in [0, 1]$ for all $i \in \mathcal{V}$ guarantees that $x_i(k) \in$

[1] Note that some works define edges in the opposite direction, so that $(i, j) \in \mathcal{E} \Leftrightarrow w_{ij} > 0$. This is primarily a notational matter, and the analytical results are unchanged. We adopt the notation of Proskurnikov and Tempo [2017].

$[0,1]$ for all $k \geq 0$ and $i \in \mathcal{V}$. We assume from here on that $x_i(0) \in [0,1]$ for all $i \in \mathcal{V}$, and 0 and 1 represent the extreme opinions of opposing points of view on the given topic, respectively. By way of example, suppose the issue being discussed was the statement "recreational marijuana should be legalized", then $x_i = 0$ and $x_i = 1$ correspond to individual i totally opposing and totally supporting the legalization of marijuana. Consequently, $x_i(k)$ can be regarded as agent i's degree of support at time k for the extreme opinion represented by 1, and so correspondingly $1 - x_i(k)$ can be regarded as the degree of support for the extreme opinion represented by 0. The reader is referred to [Anderson and Ye, 2019, Section 2.2] for further details.

We now give an intuitive explanation on how the model (1) captures bias assimilation, and provide quantitative arguments in the next subsection. Readers may also refer to Dandekar et al. [2013], which first proposed the model. One can consider $s_i(k) = \sum_{j \in \mathcal{N}_i} w_{ij} x_j(k)$ and $d_i - s_i(k) = \sum_{j \in \mathcal{N}_i} w_{ij}(1 - x_j(k))$ to be the weighted average support for the position represented by 1 and 0, respectively. When $b_i > 0$ and supposing for example that $x_i(k) > 0$, (1) indicates that individual i applies a larger weight of $x_i(k)^{b_i}$ to $s_i(k)$, and a smaller weight of $(1 - x_i(k))^{b_i}$ to $d_i - s_i(k)$. This represents the biased assimilation phenomenon (Lord et al. [1979]), which explains that individuals may process new information with a bias, being more readily inclined to accept evidence confirming their existing views while evaluating disconfirming evidence critically, perhaps even rejecting it. When $b_i = 0$ for all $i \in \mathcal{V}$, (1) simplifies to the classical DeGroot model (DeGroot [1974]).

2.1 Exploring the Bias Parameter's Effect

In this subsection, we look closely at the effect of the bias parameter $b_i > 0$ on the dynamics in (1) and show each individual assimilates new information with a bias towards information supporting his or her current opinion, that increases with b_i. To do so, we construct a specific example for a single individual i in the presence of equal information from both ends of the opinion spectrum. The example imposes some additional assumptions, which are not restrictive; the same conclusions can be drawn with other similar assumptions.

Suppose that $w_{ij} = 1$ for all $i, j \in \mathcal{V}$, and that the neighbors of i have opinions that yield $s_i = \sum_{j \in \mathcal{N}_i} x_j = d_i/2 \triangleq s$, i.e. there is an equal influence from i's neighbors on both ends of the opinion spectrum [2]. The update equation (1) of individual i can be rewritten as

$$x_i(k+1) = p(b_i, x_i(k))$$

where

$$p(b_i, x_i) \triangleq \frac{x_i + x_i^{b_i} s}{1 + x_i^{b_i} s + (1 - x_i)^{b_i} s}.$$

Evidently, the DeGroot update equation of individual i is $x_i(k+1) = p(0, x_i(k)) = (x_i(k) + s)/(1 + 2s)$.

The following inequalities illustrate the role of the bias parameter b_i, with the precise calculations given in the arXiv extended version (Xia et al. [2020]). It can be shown that

(1) If $b_i > 0$, then $p(b_i, x_i) > p(0, x_i)$ if $x_i \in (0.5, 1)$, and $p(b_i, x_i) < p(0, x_i)$ if $x_i \in (0, 0.5)$.
(2) If $b_i = 1$, then $p(1, x_i) = x_i$.
(3) If $b_i > 1$, then $p(b_i, x_i) > x_i$ for $x_i \in (0.5, 1)$ and $p(b_i, x_i) < x_i$ for $x_i \in (0, 0.5)$.
(4) If $b_i < 1$, then $p(b_i, x_i) < x_i$ for $x_i \in (0.5, 1)$ and $p(b_i, x_i) > x_i$ for $x_i \in (0, 0.5)$.

Item 1) indicates that individual i's next opinion $x_i(k+1)$ under the bias model update rule (1) is closer to the polarized value of 0 (if $x_i(k) \in (0, 0.5)$) or 1 (if $x_i(k) \in (0.5, 1)$) when compared to $x_i(k+1) = p(0, x_i(k))$ of an individual i described by the DeGroot model. It is by this mechanism that (1) captures an individual who, for $b_i > 0$, assimilates a balanced mixture of influence with a bias, more readily accepting neighboring information that supports his or her current opinion, while placing a lower weight on neighboring information that opposes his or her current opinion.

Item 2) illustrates a biased individual whose non-neutral opinion remains unchanged in the presence of equal information from both ends of the opinion spectrum; there is a perfect balance between biased assimilation and social influence from neighbors' opinions. Item 3) indicates that when $b_i > 1$, biased assimilation overpowers the social influence, and the individual tends to an extreme opinion, *even though the overall social influence due to the neighbors' opinions is unchanged.* This represents how "biased assimilation causes individuals to arrive at more extreme opinions after being exposed to identical, inconclusive evidence" (Lord et al. [1979]). Item 4) shows an individual whose the level of biased assimilation is not sufficient to overcome the social influence from the neighbors' opinions. Thus, x_i tends to 0.5, where the social influence from both ends of the spectrum is equal. Based on this discussion, we say individual i has weak bias if $b_i < 1$, or strong bias if $b_i > 1$, or intermediate bias if $b_i = 1$.

Remark 1 *For some models, each individual has a parameter describing her susceptibility to external influence (the parameter is constant in Friedkin and Johnsen [1990] and opinion-dependent in Amelkin et al. [2017]). However, both models share the same property; when an individual i is exposed to two equal pieces of opinions from either end of the spectrum, the opinion furthest from*

[2] E.g., individual i has two neighbors, one having an opinion value of 1, and the other an opinion value of 0.

opinion x_i is more attractive. This contrasts our conclusion above; for an individual i with $b_i > 0$, the opinion closer to opinion x_i is more attractive.

3 Main Results

In this section, we present the theoretical results, interweaved with discussion and interpretation in the social context, with the proofs presented in the next section. We will study the equilibria (and also their stability) of the system (1) for both $b_i > 0$ and $b_i < 0$. It turns out that this is a challenging problem in general and some results we obtain only establish local stability.

Let

$$f_i(x) \triangleq \frac{w_{ii}x_i + (x_i)^{b_i}s_i}{w_{ii} + (x_i)^{b_i}s_i + (1-x_i)^{b_i}(d_i - s_i)}.$$

The dynamics of the opinions of the N individuals in the network is rewritten as

$$x(k+1) = F(x(k)), \tag{2}$$

where $x = [x_1, \dots, x_N]^\top$ and $F(x) = [f_1(x), \dots, f_N(x)]^\top$.

For system (1) with $b_i > 0$, note that if $x_i(k) = 1$ ($x_i(k) = 0$), then $x_i(k') = 1$ ($x_i(k') = 0$) for all $k' \geq k$. It can be verified that for $b_i > 0$, we have that $\mathbf{0}_N$, $\mathbf{1}_N$, and $\frac{1}{2}\mathbf{1}_N$ are equilibria of system (2). We refer to $x^* = \mathbf{0}_N$ and $x^* = \mathbf{1}_N$ as extreme consensus and $x^* = \frac{1}{2}\mathbf{1}_N$ as neutral consensus. Any vector with all entries either 0 or 1 is also an equilibrium; without loss of generality, we denote such an equilibrium as $[\mathbf{0}_{n_1}^T, \mathbf{1}_{n_2}^T]^T$ with $n_1+n_2 = N$ and represents polarization of the network. Let the extreme, neutral consensus, and polarization equilibria of the system (2) be respectively denoted by

$$x_a^* = \mathbf{0}_N,\ x_c^* = \mathbf{1}_N,\ x_d^* = \frac{1}{2}\mathbf{1}_N,\ x_e^* = [\mathbf{0}_{n_1}^T, \mathbf{1}_{n_2}^T]^T.$$

Besides the above equilibria, there may exist other equilibria of the system depending on the graph $\mathbb{G}$, the value of the bias parameter b_i, and the weights w_{ij}. We give examples in the sequel. If the bias parameter $b_i < 0$ but is close to 0, then $x_d^* = \frac{1}{2}\mathbf{1}_N$ is an equilibrium of system (2). Though a rigorous proof is missing, we conjecture the following based on numerous simulations.

Conjecture: For a given network topology and initial states, if the system (2) with $b_i = 0$ for all $i \in \mathcal{V}$ converges (DeGroot model), then the system (2) with $b_i > 0$ for all $i \in \mathcal{V}$ will also converge.

3.1 Extreme and Neutral Consensus Equilibria

We first discuss stability of equilibria corresponding to extreme consensus and neutral consensus.

Theorem 1 *Suppose that the neighbor graph $\mathbb{G}$ is strongly connected and $b_i > 0, \forall i \in \mathcal{V}$. Then, $x_a^* = \mathbf{0}_N$ and $x_c^* = \mathbf{1}_N$ are locally exponentially stable equilibria and $x_d^* = \frac{1}{2}\mathbf{1}_N$ is an unstable equilibrium of system (2).*

In the social context, the result of Theorem 1 indicates that individuals' biased assimilation makes it possible for a network *to reach a consensus that is more extreme* ($x(\infty) = \mathbf{1}_N$ and $x(\infty) = \mathbf{0}_N$) than any individual's initial opinion $x_i(0)$. For example, one could have $x_i(0) \in (1-\epsilon_1, 1-\epsilon_2)$ for all i, with sufficiently small $\epsilon_1 > \epsilon_2 > 0$, and we get $x(\infty) = \mathbf{1}_N$, which means that $\max_i x_i(0) < \max_i x_i(\infty)$. This points to the dangers of biased assimilation in a network of individuals who all begin with similar opinions. One could say the network of individuals is "self-extremizing". Theorem 1 also tells us that when individuals exhibit biased assimilation, it is unlikely for a network to reach the unstable state of neutral consensus (in which every individual adopts the neutral opinion). However, it might be possible that stable equilibria exist in which a subset of the individuals (but not all) adopt the neutral opinion. For many established models (Altafini [2013], DeGroot [1974], Friedkin and Johnsen [1990], Hegselmann and Krause [2002]), the example initial states above will yield $\max_i x_i(0) \geq \max_i x_i(\infty)$. Some models (Duggins [2017], G. Shi et al. [2017], Mäs et al. [2014]) can have $\max_i x_i(0) < \max_i x_i(\infty)$, but only exhibit extreme polarization (see Section 3.2 below) exists in Duggins [2017], G. Shi et al. [2017], Mäs et al. [2014], and not extreme consensus.

The paper Dandekar et al. [2013] showed that the biased assimilation model exhibits extreme polarization $[\mathbf{0}_{n_1}^T, \mathbf{1}_{n_2}^T]^T$ in a two-island network, which is also not present in the existing models. Since Dandekar et al. [2013] requires that $n_1 \neq 0$ and $n_2 \neq 0$, this means that Dandekar et al. [2013] does not study the stability of extreme consensus states, as in our paper (extreme consensus can be considered as a special case of the polarization equilibria, with $n_1 = 0$ or $n_2 = 0$). We now detail a result on when the neutral consensus equilibrium is stable.

Theorem 2 *Suppose that the neighbor graph $\mathbb{G}$ is strongly connected, $w_{ii} > 0$ and $b_i \in [-\epsilon, 0)$, $\forall i \in \mathcal{V}$. If $\epsilon > 0$ is sufficiently small, then $x_d^* = \frac{1}{2}\mathbf{1}_N$ is a locally exponentially stable equilibrium of the system (2).*

To get some idea on the region of attraction of equilibria corresponding to extreme consensus (i.e. those reported in Theorem 1), we present the following result with the system starting from some restricted initial states.

Theorem 3 *Consider the system* (2), *and let $b_i \geq 1$ for all $i \in \mathcal{V}$. Suppose that the neighbor graph $\mathbb{G}$ is strongly connected. Then,*

(1) If $x_i(0) \geq 0.5$ for all $i \in \mathcal{V}$ and there exists at least

one $j \in \mathcal{V}$ such that $x_j(0) > 0.5$, then $x_i(k)$ will asymptotically converge to 1 *for all $i \in \mathcal{V}$.*

(2) If $x_i(0) \leq 0.5$ for all $i \in \mathcal{V}$ and there exists at least one $j \in \mathcal{V}$ such that $x_j(0) < 0.5$, then $x_i(k)$ will asymptotically converge to 0 *for all $i \in \mathcal{V}$.*

Theorem 3 establishes that for $b_i \geq 1$, the region of attraction for extreme consensus is in fact quite large. In particular, a network of individuals with *intermediate or strong levels of biased assimilation* will "self-extremize" to a state of extreme consensus if all initial opinions are on the same side of the opinion spectrum ($x_i(0) \geq 0.5$ or $x_i(0) \leq 0.5$ for all i), even if initially the individuals have varying degrees of support for the position at 1 or 0. An echo chamber (Barberá et al. [2015]) is a scenario whereby an individual only has access to information that supports his or her current opinion, either as a deliberate result of the individual's actions, or an unintended consequence of enabling technology, e.g. recommender systems. Theorem 3 illustrates the dangerous consequence, viz. extreme consensus, of having individuals with intermediate/strong bias assimilation together in an echo chamber.

3.2 Polarization Equilibria

Stability results for the polarization equilibria $x_e^* = [\mathbf{0}_{n_1}^T, \mathbf{1}_{n_2}^T]^T$ on strongly connected graphs, complete graphs and two-island networks are now introduced.

Theorem 4 *If the neighbor graph $\mathbb{G}$ is strongly connected, and $b_i \in (0,1)$ for all $i \in \mathcal{V}$, then the equilibrium $x_e^* = [\mathbf{0}_{n_1}^T, \mathbf{1}_{n_2}^T]^T$ of the system* (2) *is unstable for every $n_1 = 1, \ldots, N-1$.*

Theorem 5 *For an undirected complete neighbor graph $\mathbb{G}$ with weights $w_{ij} = 1$, $i \neq j$, $i, j \in \mathcal{V}$, the equilibrium $x_e^* = [\mathbf{0}_{n_1}^T, \mathbf{1}_{n_2}^T]^T$ of the system* (2)*, with $n_1 = 2, \ldots, N-2$, is unstable when $b_i = 1$ for all $i \in \mathcal{V}$ and is locally exponentially stable when $b_i > 1$ for all $i \in \mathcal{V}$.*

Next we introduce the two-island network structure studied in Dandekar et al. [2013], modeling a *homophilous* network. Consider an undirected network in which the nodes in $\mathcal{V}$ are partitioned into two types, τ_1, τ_2. Let $\mathcal{V}_i$ denote the set of nodes of type τ_i and $|\mathcal{V}_i| = n_i$, $i = 1, 2$. Without loss of generality, assume that $\mathcal{V}_1 = \{1, \ldots, n_1\}$ and $\mathcal{V}_2 = \{n_1 + 1, \ldots, N\}$. Assume that each node in $\mathcal{V}_1$ has $n_1 p_s$ neighbors in $\mathcal{V}_1$ and $n_1 p_d$ neighbors in $\mathcal{V}_2$, and each node in $\mathcal{V}_2$ has $n_2 p_s$ neighbors in $\mathcal{V}_2$ and $n_2 p_d$ neighbors in $\mathcal{V}_1$, where $p_s, p_d \in (0,1)$ and $n_1 p_s, n_1 p_d, n_2 p_s, n_2 p_d$ are all integers.

Theorem 6 *Suppose that the neighbor graph $\mathbb{G}$ is a connected two-island network and $w_{ij} = 1$, $i \neq j$, $i, j \in \mathcal{V}$. Then, $x_e^* = [\mathbf{0}_{n_1}^T, \mathbf{1}_{n_2}^T]^T$ is a locally exponentially stable equilibrium of the system* (2) *when $b_i \geq 1$ for all $i \in \mathcal{V}$.*

Theorem 4 establishes a result of particular interest when considered in conjunction with Theorem 1. Specifically, although a network of individuals with weak bias, $b_i \in (0,1)$, can converge to an extreme consensus (which is undesirable), the same weak bias ensures that polarization (a different type of undesirable equilibrium) is an unstable phenomenon. Theorem 4 also tells us it is unlikely for a network to converge to a polarized state if individuals are only weakly biased; polarization equilibria are stable only when individuals have intermediate or strong levels of bias (Theorems 5 and 6). Efforts to reduce polarization could therefore focus on changing individual bias as opposed to e.g. changing network structure or the strategic introduction of agents.

Remark 2 *We observed from numerous simulations where we sampled $x_i(0)$ from a uniform distribution in* $[0,1]$*, that polarization occurs if b_i is much larger than 1 for a large set of strongly connected network topologies, such as regular, complete, random, and small-world graphs, while for other topologies like path, and star graphs, polarization does not occur.*

The system (2) can have other equilibria and can exhibit rich asymptotic behaviors as will be illustrated in Section 5. The following theorem establishes a case when other types of equilibria of the system (2) exist and their stability is discussed.

Theorem 7 *Let $b_i = 1, \forall\, i \in \mathcal{V}$. Consider an undirected star graph with the weights $w_{ij} = 1$, $i \neq j$, $i, j \in \mathcal{V}$. Without loss of generality, suppose that node 1 is the center node. The equilibria of system (2) include those vectors whose elements are either zero or one, and $x^* = [\frac{1}{2}, a_2, \ldots, a_N]^T$ with $a_i \in [0,1]$ and $\sum_{i=2}^N a_i = \frac{N-1}{2}$. If N is odd, the system has additional equilibria of the form $x^* = [c, \mathbf{0}_{\frac{N-1}{2}}^T, \mathbf{1}_{\frac{N-1}{2}}^T]^T$ with $c \in (0,1)$. Among these equilibria, $x_a^* = \mathbf{0}_N$, and $x_c^* = \mathbf{1}_N$ are locally exponentially stable and all the other equilibria are unstable.*

Consider equilibria of the form $x^* = [\frac{1}{2}, a_2, \ldots, a_N]^T$ with $\sum_{i=2}^N a_i = (N-1)/2$, or $x^* = [c, \mathbf{0}_{\frac{N-1}{2}}^T, \mathbf{1}_{\frac{N-1}{2}}^T]^T$ with $c \in (0,1)$. Theorem 7 establishes that it is possible under biased assimilation to split a star network so that the leaf nodes separate into 2 nonempty factions, one supporting the opinion represented by 1, and the other supporting the opinion represented by 0. In fact, the support can be of varying levels of intensity, with different faction sizes, since one only requires that $a_i \in [0,1]$ and $\sum_{i=2}^N a_i = \frac{N-1}{2}$. The centre node acts as a mediating individual to the two factions at an unstable equilibrium.

4 Analyses

In this section, we prove the theorems in the previous section. We linearize the system (2) to analyze the local stability of these equilibria. Let $g_i(x) \triangleq w_{ii} + x_i^{b_i} s_i + (1-x_i)^{b_i}(d_i - s_i)$ for $i = 1, \ldots, N$. By calculation, one obtains that the Jacobian of $F(x(k))$ in (2), $\frac{\partial F}{\partial x} = (\frac{\partial f_i}{\partial x_j})_{N\times N}$, has entries

$$\begin{aligned}\frac{\partial f_i}{\partial x_i} =& \frac{1}{g_i^2(x)}\Big[(w_{ii} + b_i x_i^{b_i-1} s_i) g_i(x) - (w_{ii}x_i + x_i^{b_i} s_i) \\ &\times \Big[b_i x_i^{b_i-1} s_i - b_i(1-x_i)^{b_i-1}(d_i - s_i)\Big]\Big] \quad (3)\end{aligned}$$

and

$$\begin{aligned}\frac{\partial f_i}{\partial x_l} =& \frac{1}{g_i^2(x)}\Big[x_i^{b_i} w_{il} g_i(x) \\ &- (w_{ii}x_i + x_i^{b_i} s_i)(x_i^{b_i} w_{il} - (1-x_i)^{b_i} w_{il})\Big] \quad (4)\end{aligned}$$

for $l \neq i$ and $i, l \in \mathcal{V}$.

A real matrix $M = (m_{ij})_{N\times N}$ is called a *nonnegative* matrix if $m_{ij} \geq 0,\ i, j = 1, \ldots, N$. The spectral radius of square matrix M is denoted as $\rho(M)$. Before proving the theorems, the following lemma is first introduced that will be used later.

Lemma 1 *(Horn and Johnson [1985]) Suppose* $M \in \mathbb{R}^{N\times N}$ *and* M *is a nonnegative matrix. Then* $\rho(M)$ *is an eigenvalue of* M *and* $\min_{1\leq i\leq N} \sum_{j=1}^N m_{ij} \leq \rho(M) \leq \max_{1\leq i\leq N} \sum_{j=1}^N m_{ij}$.

Proof of Theorem 1: Consider the equilibrium $x_a^* = \mathbf{0}_N$. For all $i \in \mathcal{V}$, one knows that $s_i = 0$ and $g_i(x_a^*) = w_{ii} + d_i$. One can derive using (3) and (4) that

$$\left.\frac{\partial f_i}{\partial x_i}\right|_{x_a^*} = \frac{w_{ii}}{g_i(x_a^*)}\ , \text{ and } \left.\frac{\partial f_i}{\partial x_l}\right|_{x_a^*} = 0, \text{ for } l \neq i.$$

Thus, the Jacobian at the equilibrium $x_a^* = \mathbf{0}_N$ becomes

$$P \triangleq \left.\frac{\partial F}{\partial x}\right|_{x_a^*} = \text{diag}\left\{\frac{w_{11}}{g_1(x_a^*)}, \frac{w_{22}}{g_2(x_a^*)}, \ldots, \frac{w_{NN}}{g_N(x_a^*)}\right\}.$$

Note that $g_i(x_a^*) = w_{ii} + \sum_{j\in\mathcal{N}_i} w_{ij}$. The eigenvalues of P are $w_{ii}/(w_{ii} + \sum_{j\in\mathcal{N}_i} w_{ij}),\ i \in \mathcal{V}$, which lie in the interval $[0, 1)$ as long as each agent has at least one neighbor. Since $\mathbb{G}$ is strongly connected, $\rho(P) < 1$ and thus the equilibrium $x_a^* = \mathbf{0}_N$ is locally exponentially stable.

For the equilibrium $x_c^* = \mathbf{1}_N$, observe that for $i \in \mathcal{V}$, one has $g_i(x_c^*) = w_{ii} + d_i$. This yields

$$\left.\frac{\partial f_i}{\partial x_i}\right|_{x_c^*} = \frac{w_{ii}}{g_i(x_c^*)}\ , \text{ and } \left.\frac{\partial f_i}{\partial x_l}\right|_{x_c^*} = 0, \text{ for } l \neq i$$

Thus the Jacobian matrix at the equilibrium $x_c^* = \mathbf{1}_N$ becomes

$$P \triangleq \left.\frac{\partial F}{\partial x}\right|_{x_c^*} = \text{diag}\left\{\frac{w_{11}}{g_1(x_c^*)}, \frac{w_{22}}{g_2(x_c^*)}, \ldots, \frac{w_{NN}}{g_N(x_c^*)}\right\}.$$

The eigenvalues of P are $w_{ii}/(w_{ii} + d_i),\ i \in \mathcal{V}$, which lie in the interval $[0, 1)$ as in the previous case. Thus the equilibrium $x_c^* = \mathbf{1}_N$ is locally exponentially stable.

For the equilibrium $x_d^* = \frac{1}{2}\mathbf{1}_N$, one obtains $g_i(x_d^*) = w_{ii} + d_i/2^{b_i}$, for all $i \in \mathcal{V}$. One further calculates that

$$\left.\frac{\partial f_i}{\partial x_i}\right|_{x_d^*} = \frac{w_{ii} + \frac{b_i d_i}{2^{b_i}}}{g_i(x_d^*)}, \quad \left.\frac{\partial f_i}{\partial x_l}\right|_{x_d^*} = \frac{w_{il}}{2^{b_i} g_i(x_d^*)}, \text{ for } l \neq i. \quad (5)$$

The above implies that the Jacobian matrix $P \triangleq \frac{\partial F}{\partial x}|_{x_d^*}$ at $x_d^* = \frac{1}{2}\mathbf{1}_N$ is a nonnegative matrix. Using Lemma 1 and (5), one can compute that the spectral radius obeys

$$\rho(P) \geq \min_{i=1,\ldots,N} \sum_{j=1}^N p_{ij} = 1 + \min_{i=1,\ldots,N} \frac{b_i \frac{1}{2^{b_i}} d_i}{w_{ii} + \frac{1}{2^{b_i}} d_i} > 1.$$

Thus $x_d^* = \frac{1}{2}\mathbf{1}_N$ is an unstable equilibrium. □

Proof of Theorem 2: Similar calculations to the proof of Theorem 1 shows that the Jacobian matrix evaluated at $x_d^* = \frac{1}{2}\mathbf{1}_N$, denoted $P \triangleq \frac{\partial F}{\partial x}|_{x_d^*}$, has the same entries as in (5). The off-diagonal elements of P are nonnegative. Since $w_{ii} > 0$ and $b_i \in [-\epsilon, 0)$, for all $i \in \mathcal{V}$, with $\epsilon > 0$ sufficiently small, one has

$$\frac{w_{ii} + b_i \frac{1}{2^{b_i}} d_i}{g_i(x_d^*)} = \frac{w_{ii} + b_i \frac{1}{2^{b_i}} d_i}{w_{ii} + \frac{1}{2^{b_i}} d_i} \geq 0$$

for all $i \in \mathcal{V}$, and hence P is a nonnegative matrix. By Lemma 1, the spectral radius of P satisfies that

$$\rho(P) \leq \max_{i=1,\ldots,N} \sum_{j=1}^N p_{ij} = 1 + \max_{i=1,\ldots,N} \frac{\frac{b_i}{2^{b_i}} \sum_{j\in\mathcal{N}_i} w_{ij}}{g_i(x_d^*)} < 1.$$

Therefore $x_d^* = \frac{1}{2}\mathbf{1}_N$ is a locally exponentially stable equilibrium of the system (2) for $b_i \in [-\epsilon, 0)$ when ϵ is sufficiently small. □

Proof of Theorem 3: We first prove item 1). Consider any $i \in \mathcal{V}$, and observe that

$$\begin{aligned}&x_i(k+1) - x_i(k) \\ &\quad = \frac{\zeta_i(x(k))}{w_{ii} + (x_i(k))^{b_i} s_i(k) + (1 - x_i(k))^{b_i}(d_i - s_i(k))}.\end{aligned}$$

where $\zeta_i(x) = x_i^{b_i}s_i - x_i^{b_i+1}s_i - x_i(1-x_i)^{b_i}(d_i - s_i)$.

Proving $x_i(k+1) - x_i(k) \geq 0$ is equivalent to proving that $\zeta_i(x(k)) \geq 0$ since the equation above has a positive denominator. Rearranging terms in $\zeta_i(x)$ and recalling that $d_i = \sum_{j\in\mathcal{N}_i} w_{ij}$ and $s_i = \sum_{j\in\mathcal{N}_i} w_{ij}x_j$, yields

$$\zeta_i(x) = \sum_{j\in\mathcal{N}_i} w_{ij}x_i^{b_i}\Big(x_j(1-x_i) - \frac{x_i}{x_i^{b_i}}(1-x_i)^{b_i}(1-x_j)\Big).$$

Since $x_i \in [0.5, 1], i \in \mathcal{V}$, $x_i^{b_i} > 0$, implying that $\zeta_i(x) \geq 0$ if $x_j(1-x_i) - x_i^{-b_i}x_i(1-x_i)^{b_i}(1-x_j) \geq 0$, or equivalently:

$$\frac{1-x_i}{x_i} \geq \left(\frac{1-x_i}{x_i}\right)^{b_i} \frac{1-x_j}{x_j} \tag{6}$$

holds for all $j \in \mathcal{N}_i$. Trivially, (6) holds if $x_i = 1$, so let us consider $x_i \in [0.5, 1)$. Notice that $x_i \in [0.5, 1) \Rightarrow (1-x_i)/x_i \leq 1$ with equality if and only if $x_i = 0.5$. Thus, (6) holds if $x_i \in [0.5, 1)$, with equality if and only if $x_j = 0.5$ and either (i) $b_i = 1$ or (ii) $x_i = 0.5$. With $x_j \in [0.5, 1], j \in \mathcal{N}_i$, we can then conclude that $\zeta_i(x) > 0$ if (i) $\exists j \in \mathcal{N}_i : x_j > 0.5$, or (ii) $b_i > 1$ and $x_i \in (0.5, 1)$. If $x_i(0) \geq 0.5$ for all $i \in \mathcal{V}$, then $x_i(k+1) \geq x_i(k)$ for all $i \in \mathcal{V}$ and all time k. Moreover, since there exists at least one $j \in \mathcal{V}$ such that $x_j(0) > 0.5$ and $\mathbb{G}$ is strongly connected, unless $x_i(0) = 1$ for all $i \in \mathcal{V}$, there exists a $p \in \mathcal{V}$ such that $p \neq j$ and $x_p(1) > x_p(0) \geq 0.5$. Repeating this argument, one concludes that there exists a finite τ such that $x_i(k) > 0.5$ for all $i \in \mathcal{V}$ and $k \geq \tau$.

Consider the Lyapunov function $V(x(k)) = 1 - \min_{i\in\mathcal{V}} x_i(k)$. From (1), if $x_i(k) = 1$, then $x_i(k+1) = 1$, which implies that if $x_i(0) = 1$, then $x_i(k) = 1$ for all time k. Thus, if $x_i(k) = 1$ for all $i \in \mathcal{V}$ at some time k, then $V(x(k)) = 0$ and $V(x(k+1)) = 0$. Suppose that there exists at least one agent p such that $x_p(k) < 1$ at a specific time k. Without loss of generality, assume $k \geq \tau$. From the preceding discussion, $x_p(k+1) > x_p(k)$, which implies that $\min_{i\in\mathcal{V}} x_i(k+1) > \min_{i\in\mathcal{V}} x_i(k)$, and thus $V(x(k+1)) < V(x(k))$. By Lyapunov's stability theorem for discrete-time autonomous systems [Haddad and Chellaboina, 2008, Theorem 13.2], $\lim_{k\to\infty} x_i(k) = 1$ asymptotically for all $i \in \mathcal{V}$.

Item 2) can be proved using arguments similar to those in the proof of item 1), with the Lyapunov function $V(x(k)) = \max_{i\in\mathcal{V}} x_i(k)$. □

Before proving Theorems 4-6, we calculate the elements of the Jacobian matrix of the system (2) at the equilibrium $x_e^* = [\mathbf{0}_{n_1}^T, \mathbf{1}_{n_2}^T]^T$.

Let x_{ei}^* be the ith entry of x_e^*, $\mathcal{N}_i^{(0)} = \{j : j \in \mathcal{N}_i, x_{ej}^* = 0\}$ and $\mathcal{N}_i^{(1)} = \{j : j \in \mathcal{N}_i, x_{ej}^* = 1\}$ for $i \in \mathcal{V}$, and

$$d_i^{(0)} = \sum_{j\in\mathcal{N}_i^{(0)}} w_{ij},\ d_i^{(1)} = \sum_{j\in\mathcal{N}_i^{(1)}} w_{ij}.$$

For agent i that satisfies $x_{ei}^* = 0$, it is easy to see that $g_i(x_e^*) = w_{ii} + d_i^{(0)}$. Calculations show that

$$\frac{\partial f_i}{\partial x_i}\Big|_{x_e^*} = \frac{(w_{ii} + {x_{ei}^*}^{b_i-1}b_i d_i^{(1)})g_i(x_e^*) - {x_{ei}^*}^{2b_i-1}b_i(d_i^{(1)})^2}{g_i^2(x_e^*)},$$

and $\frac{\partial f_i}{\partial x_l}|_{x_e^*} = \frac{1}{g_i^2(x_e^*)}{x_{ei}^*}^{b_i} = 0$ for $l \neq i$. For agent i that satisfies $x_{ei}^* = 1$, one has $g_i(x_e^*) = w_{ii} + d_i^{(1)}$. Eq. (3) then yields

$$\frac{\partial f_i}{\partial x_i}\Big|_{x_e^*} = \frac{1}{g_i(x_e^*)}\left[w_{ii} + (1-x_{ei}^*)^{b_i-1}b_i d_i^{(0)}\right], \tag{7}$$

and for $l \neq i$, (4) evaluates to be

$$\frac{\partial f_i}{\partial x_l}\Big|_{x_e^*} = \frac{1}{g_i^2(x_e^*)}[w_{il}g_i(x_e^*) - w_{il}(w_{ii} + d_i^{(1)})] = 0. \tag{8}$$

Proof of Theorem 4: Since the graph is strongly connected, for any equilibrium $x_e^* = [\mathbf{0}_{n_1}^T, \mathbf{1}_{n_2}^T]^T$ with $n_1 + n_2 = N$ and a given $n_1 \in \{1, \dots, N-1\}$, there exists an agent l such that $x_{el}^* = 1$ and $d_l^{(0)} > 0$. When $0 < b_i < 1$ for all $i \in \mathcal{V}$, it follows from (7) that

$$\frac{\partial f_l}{\partial x_l}\Big|_{x_e^*} = \frac{1}{g_l(x_e^*)}\left[w_{ll} + (1-x_{el}^*)^{b_i-1}b_i d_l^{(0)}\right] = +\infty.$$

In view of (8), the Jacobian matrix at the equilibrium $x_e^* = [\mathbf{0}_{n_1}^T, \mathbf{1}_{n_2}^T]^T$ is a diagonal matrix with at least one element equal to $+\infty$. Therefore, the equilibrium $x_e^* = [\mathbf{0}_{n_1}^T, \mathbf{1}_{n_2}^T]^T$ is unstable. □

Proof of Theorem 5: When $b_i = 1$ for all $i \in \mathcal{V}$, for agent i with $x_{ei}^* = 0$ and agent l with $x_{el}^* = 1$, one has

$$\frac{\partial f_i}{\partial x_i}\Big|_{x_e^*} = \frac{w_{ii} + d_i^{(1)}}{w_{ii} + d_i^{(0)}}, \text{ and } \frac{\partial f_l}{\partial x_l}\Big|_{x_e^*} = \frac{w_{ll} + d_l^{(0)}}{w_{ll} + d_l^{(1)}},$$

respectively. Then the Jacobian matrix $P = \frac{\partial F}{\partial x}|_{x_e^*}$ at the equilibrium $x_e^* = [\mathbf{0}_{n_1}^T, \mathbf{1}_{n_2}^T]^T$ is

$$P = \text{diag}\left\{\frac{w_{11}+d_1^{(1)}}{w_{11}+d_1^{(0)}}, \frac{w_{22}+d_2^{(1)}}{w_{22}+d_2^{(0)}}, \dots, \frac{w_{NN}+d_N^{(0)}}{w_{NN}+d_N^{(1)}}\right\}. \tag{9}$$

Suppose that $n_1 \geq n_2$. For the i-th agent with $x_{ei}^* = 1$, in view of (9), the i-th diagonal element of P is given by $p_{ii} = (w_{ii} + n_1)/(w_{ii} + n_2 - 1) > 1$. If $n_1 < n_2$, one can similarly show that there exists a diagonal element of P that is greater than 1. In both cases, P has an eigenvalue greater than 1. Therefore the equilibrium $x_e^* = [\mathbf{0}_{n_1}^T, \mathbf{1}_{n_2}^T]^T$ is unstable when $b_i = 1$ for all $i \in \mathcal{V}$.

When $b_i > 1$ for all $i \in \mathcal{V}$, for agent i with $x_{ei}^* = 0$, one has $\frac{\partial f_i}{\partial x_i}|_{x_e^*} = w_{ii}/(w_{ii}+d_i^{(0)})$. Since the graph is complete and $n_1 \geq 2$, $d_i^{(0)} > 0$ and therefore $0 \leq w_{ii}/(w_{ii}+d_i^{(0)}) < 1$. For agent l with $x_{el}^* = 1$, one has $\frac{\partial f_l}{\partial x_l}|_{x_e^*} = w_{ll}/(w_{ll} + d_l^{(1)})$. Similarly, one derives that $0 \leq w_{ll}/(w_{ll}+d_l^{(1)}) < 1$. Then the Jacobian matrix $P = \frac{\partial F}{\partial x}|_{x_e^*}$ evaluated at the equilibrium $x_e^* = [\mathbf{0}_{n_1}^T, \mathbf{1}_{n_2}^T]^T$ is

$$P = \mathrm{diag}\left\{\frac{w_{11}}{w_{11}+d_1^{(0)}}, \frac{w_{22}}{w_{22}+d_2^{(0)}}, \ldots, \frac{w_{NN}}{w_{NN}+d_N^{(1)}}\right\} \quad (10)$$

and has spectral radius $\rho(P) < 1$. It follows that $x_e^* = [\mathbf{0}_{n_1}^T, \mathbf{1}_{n_2}^T]^T$ is locally exponentially stable when $b_i > 1$ for all $i \in \mathcal{V}$. □

Proof of Theorem 6: From the definition of the two-island network model, the following inequalities hold

$$d_1^{(0)} > d_1^{(1)}, d_2^{(0)} > d_2^{(1)}, \ldots, d_N^{(1)} > d_N^{(0)}. \quad (11)$$

For $b_i = 1$ and $b_i > 1$, the Jacobian matries are given by (9) and (10), respectively. In both cases, one can see that the eigenvalues of P lie in the interval $[0, 1)$ and thus $x_e^* = [\mathbf{0}_{n_1}^T, \mathbf{1}_{n_2}^T]^T$ is locally exponentially stable. □

Due to spatial constraints, the proof of Theorem 7 is given in the arXiv extended version (Xia et al. [2020]) of this paper.

5 Numerical simulations

In this section, we perform several simulations to show the rich asymptotic behaviors of the system (2), including some equilibria not studied in Section 3. Each of the following simulations consider a two-island network model with each island consisting of 50 nodes. For each node, the number of neighbors of the same type and of the other type are $n_1 p_s = n_2 p_s = 4$ and is $n_1 p_d = n_2 p_d = 2$, respectively. Edges are bidirectional, i.e. $(j, i) \in \mathcal{E} \Leftrightarrow (i, j) \in \mathcal{E}$, but w_{ij} and w_{ji} are not necessarily equal, thus making the graph directed. In particular, if $(j, i) \in \mathcal{E}$, we draw w_{ij} from a uniform distribution from $[0.5, 1.5]$, and set $w_{ii} = 0$ for all $i \in \mathcal{V}$.

In the first case, we consider when b_i for all $i \in \mathcal{V}$ are chosen randomly from a uniform distribution in the interval $[1.01, 1.5]$, i.e. all individuals have strong bias. The initial states of the agents are chosen randomly from a uniform distribution in the interval $[0, 1]$, and the evolution of the states of the agents are illustrated in Fig. 1, from which one can see that the system reaches an extreme polarization equilibrium. If b_i for all $i \in \mathcal{V}$ are much larger than 1, we observe from extensive simulations that extreme polarization is also observed for a large class of strongly connected network topologies such as regular graphs, random graphs, and small-world graphs. This illustrates the important role of individuals with strong bias in creating a polarized network state.

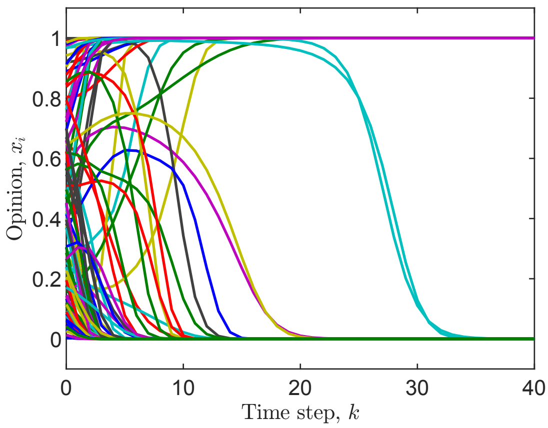

Fig. 1. The system state under a two-island network with $b \in [1.01, 1.5]$ and randomized initial states.

To illustrate that there are other equilibria which are very different to those analyzed in Section 3, we present the following simulations. We now draw b_i for all $i \in \mathcal{V}$ from a uniform distribution of interval $\in [0.5, 1.5]$, so that some individuals have weak bias and some have strong bias. If the initial states are randomly chosen from a uniform distribution in $[0, 1]$, then Fig. 2 illustrates that the states of most of the agents converge either to 0 or 1 and the final states of the remaining agents lie in between. Again, similar results to Fig. 2 can be observed in other network topology types, including path networks, regular networks and small-world networks.

Now consider the case where b_i for all $i \in \mathcal{V}$ belongs to a uniform distribution of interval $\in (0, 1)$. For initial states uniformly randomly chosen from the interval $[0, 1]$, two situations are typically observed for the state evolution of the system. In the first situation, the states of all agents converge either to 1 or 0, and in the other situation, the states of most of the agents converge to values close either to 0 or 1 and the final states of the remaining agents lie in $[0, 1]$. As all b_i values tend closer to 0, the situation in which the states of all the agents converge to an extreme consensus equilibrium occurs more frequently. When b_i is close to 1, the agents may converge to two clusters of opinions close to the extreme polarization equilibria for certain initial states. For example, we consider $b_i \in [0.8, 0.9]$ under the two-island network, and the initial states of agents 1 to 50 are randomly chosen from a uniform distribution of interval $[0.15, 0.2]$ and

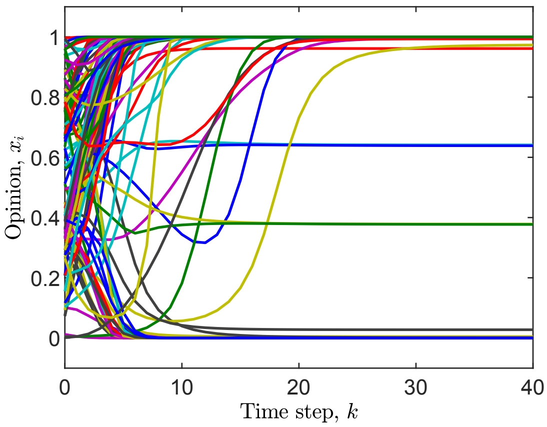

Fig. 2. The system state under a two-island network with $b_i \in [0.5, 1.5]$ and randomized initial states.

the remaining agents have initial states from a uniform distribution of interval $[0.75, 0.8]$. Fig. 3 shows that the network converges to a steady state in which the two islands have states close to the extreme values of 0 and 1.

Remark 3 *We have shown that there are equilibria other than those studied in Section 3. Although not shown, we also observed that heterogeneous b_i can generate equilibria that does not exist for a homogeneous bias parameter. Similarly, there may be equilibria for undirected networks which do not exist for directed graphs, and vice versa.*

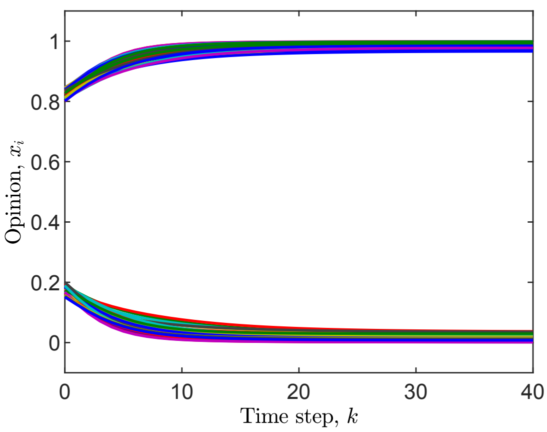

Fig. 3. The system state under a two-island network with $b_i \in [0.8, 0.9]$ and the initial state $x_i(0) \in [0.15, 0.2], i = 1, \dots, 50$ and $x_i(0) \in [0.8, 0.85], i = 51, \dots, 100$.

6 Conclusion

In this paper, we have studied the equilibria and their stability for a recently proposed nonlinear opinion dynamics model with biased assimilation in which each agent is associated with a bias parameter. We have shown that, with heterogeneous bias parameter values, the stability of certain equilibria depend on the degree of bias and the topology of the neighbor relationships among the agents. Both theoretical analyses and numerical simulations have shown that both the value of the bias parameter and the network topology play a key role in determining the limiting opinion distributed in the network. For future work, we aim to further study the region of attraction of the different equilibria and explore the general convergence condition for arbitrary strongly connected networks and arbitrary initial states, though a conjecture was given in Section 3.

References

I. Ajzen. Nature and operation of attitudes. *Annual Review of Psychology*, 52:27–58, 2001.

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

V. Amelkin, F. Bullo, and A. K. Singh. Polar opinion dynamics in social networks. *IEEE Transactions on Automatic Control*, 62(11):5650–5665, 2017.

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*, 16(2):129–149, 2019.

N. H. Anderson. *Foundations of Information Integration Theory*. Academic Press, 1981.

P. Barberá, J. T. Jost, J. Nagler, J. A. Tucker, and R. Bonneau. Tweeting from left to right: Is online political communication more than an echo chamber? *Psychological Science*, 26(10):1531–1542, 2015.

Z. Chen, J. Qin, B. Li, H. Qi, P. Buchhorn, and G. Shi. Dynamics of Opinions with Social Biases. *Automatica*, 106(8):374–383, 2019.

P. Dandekar, A. Goel, and D. T. Lee. Biased assimilation, homophily, and the dynamics of polarization. *Proceedings of the National Academy of Sciences*, 110: 5791–5796, 2013.

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

M. Dimock, J. Kiley, S. Keeter, and C. Doherty. Political Polarization in the American Public. Technical report, Pew Research Center, 2014.

P. Duggins. A psychologically-motivated model of opinion change with applications to american politics. *Journal of Artificial Societies and Social Simulation*, 20(1):1–13, 2017. ISSN 1460-7425. doi: 10.18564/jasss.3316.

N. E. Friedkin and E. C. Johnsen. Social influence and opinions. *Journal of Mathematical Sociology*, 15(3-4): 193–205, 1990.

G. Shi, A. Proutiere, M. Johansson, J. S. Baras, and K. H. Johansson. Emergent behaviors over signed random dynamical networks: Relative-state-flipping model. *IEEE Transactions on Control of Network Systems*, 4(2):369–379, 2017.

W. M. Haddad and V. Chellaboina. *Nonlinear Dynamical Systems and Control: A Lyapunov-Based Approach*. Princeton University Press, Princeton, 2008.

R. Hegselmann and U. Krause. Opinion dynamics and bounded confidence: Models, analysis and simulation. *Journal of Artificial Societies and Social Simulation*, 5:1–24, 2002.

R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, Cambridge, U.K., 1985.

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*, 37(11):2098–2109, 1979.

M. Mäs, A. Flache, and J. A. Kitts. Cultural integration and differentiation in groups and organizations. In *Perspectives on Culture and Agent-based Simulations*, pages 71–90. Springer, 2014.

G. D. Munro, P. H. Ditto, L. K. Lockhart, A. Fagerlin, M. Gready, and E. Peterson. Biased assimilation of sociopolitical arguments: Evaluating the 1996 U.S. presidential debate. *Basic and Applied Social Psychology*, 24(1):15–26, 2002.

A. V. Proskurnikov and R. Tempo. A tutorial on modeling and analysis of dynamic social networks. Part I. *Annual Reviews in Control*, 43:65–79, 2017.

C. S. Taber and M. Lodge. Motivated skepticism in the evaluation of political beliefs. *American Journal of Political Science*, 50(3):755–769, 2006.

W. Xia, J. Liu, M. Cao, and X.-M. Sun. Stability analysis of a nonlinear opinion dynamics model for biased assimilation. In *Proc. of the 57th IEEE Conference on Decision and Control*, pages 1287–1292, 2018.

W. Xia, M. Ye, J. Liu, M. Cao, and X.-M. Sun. Analysis of a nonlinear opinion dynamics model with biased assimilation. 2020. arXiv:1912.01778v2.