



HAL
open science

Consensus in Models for Opinion Dynamics with Generalized-Bias

Juan Paz, Camilo Rocha, Luis Tobòn, Frank Valencia

► **To cite this version:**

Juan Paz, Camilo Rocha, Luis Tobòn, Frank Valencia. Consensus in Models for Opinion Dynamics with Generalized-Bias. COMPLEX NETWORKS 2024 - 13th International Conference on Complex Networks & Their Applications, Dec 2024, Istanbul, Turkey. pp.253-269, 10.48550/arXiv.2409.10809 . hal-04918975

HAL Id: hal-04918975

<https://inria.hal.science/hal-04918975v1>

Submitted on 29 Jan 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Consensus in Models for Opinion Dynamics with Generalized-Bias

Juan Paz¹, Camilo Rocha¹, Luis Tobón¹, and Frank Valencia^{2,1}

¹ Pontificia Universidad Javeriana Cali, Colombia

² CNRS-LIX, École Polytechnique de Paris, France.

Abstract. Interest is growing in social learning models where users share opinions and adjust their beliefs in response to others. This paper introduces generalized-bias opinion models, an extension of the DeGroot model, that captures a broader range of cognitive biases. These models can capture, among others, dynamic (changing) influences as well as in-group favoritism and out-group hostility, a bias where agents may react differently to opinions from members of their own group compared to those from outside. The reactions are formalized as arbitrary functions that depend, not only on opinion difference, but also on the particular opinions of the individuals interacting. Under certain reasonable conditions, all agents –despite their biases– will converge to a consensus if the influence graph is strongly connected, as in the original DeGroot model. The proposed approach combines different biases, providing deeper insights into the mechanics of opinion dynamics and influence within social networks.

Keywords: Multi-Agent Systems, Social Networks, Cognitive Bias, Consensus, Intergroup Bias

1 Introduction

Social networks have played a significant role in shaping users’ opinions, often influencing democratic processes and contributing to social polarization. Broadly, the dynamics of opinion formation in social networks involve users expressing their views, encountering the opinions of others, and potentially updating their own beliefs based on these interactions. There is a growing interest in developing models of social learning [21] capturing these dynamics to gain insights on how opinions form and spread within social networks.

The DeGroot model [9] is perhaps the most prominent framework for social learning and opinion formation dynamics. In this model, a community is represented as a weighted directed graph, known as the *influence graph*, with edges indicating how much individuals (*agents*) influence one another. Each agent has an opinion represented as a value in the interval $[0, 1]$, indicating the strength of the agreement with an underlying proposition (e.g., “*AI poses a threat to humanity*”). Agents revise their opinions by “averaging” them with those of their contacts, weighted by their influence. Studies support the validity of opinion formation, in many cases, by averaging as demonstrated in controlled sociological experiments [6].

A key theoretical result about the model states that the agents will converge to *consensus* if the influence graph is *strongly connected* and the agents have non-zero self-influence (*puppet freedom* is assumed) [12]. The significance of this result lies in the fact that consensus is a central problem in social learning. Indeed, the inability to reach consensus is a sign of a polarized community. The DeGroot model is valued for its tractability, derived from its connection with matrix powers and Markov chains; it remains a significant focus of study for understanding of opinion evolution [12].

The DeGroot model, however, assumes two constraints that could be overly demanding in the context of social networks. It assumes *homogeneity* and *linearity* of opinion update dynamics. In social scenarios, two agents may update their opinions differently depending on their individual *cognitive biases* on disagreement (i.e., on how they interpret and react towards the level of disagreements with others). This results in more involved updates that may rely on non-linear, even non-monotonic, functions. For example, an individual under *confirmation (cognitive) bias* [4] may ignore the opinion of those whose level of disagreement is over a certain threshold. A recent work [2] introduces an extension of the DeGroot model that allows for some form of *heterogeneous* and *non-linear* opinion updates. The model takes *disagreement* between agents as one of its key parameters.

Nevertheless, representing cognitive biases solely as functions of disagreement [2] overlooks key behaviors where reactions to the same opinion can vary greatly depending on whether the opposing view comes from someone within their own political group. This can be attributed to *in-group favoritism* (or out-group hostility), which are aspects of inter-group bias and social identity theory [19]. In order to capture these nuances, models of cognitive biases need to incorporate factors beyond mere, isolated, disagreement.

This paper introduces *generalized-bias opinion models*, a framework that generalizes the aforementioned models: it can capture a broader range of cognitive biases formalized as *arbitrary* functions that depend, not only on opinion disagreement, but also on the opinions of the individuals involved. It is demonstrated that the proposed framework can effectively represent standard cognitive biases of significant relevance in social networks. These include, among others, confirmation and identity-based *inter-group biases* [19], where users are more inclined to favor opinions from people within their own ideological group. It also allows for modelling *dynamic* (changing) influences. The approach allows for the combination and integration of these biases, enabling a more comprehensive and formal understanding of opinion dynamics and influence in social networks.

Furthermore, the key result of the DeGroot model is extended by demonstrating that, under certain reasonable conditions, all agents—despite being subject to the aforementioned cognitive biases—will converge to a consensus if the influence graph is strongly connected. This highlights the robustness of the proposed framework in achieving consensus even when accounting for complex, arbitrary bias-driven dynamics under some reasonable conditions. Additionally, it is also shown that the proposed framework allows for more expressive models to those introduced in [1,2,9]. These results contribute to a richer representation of human cognition and interaction in social networks. Indeed, the framework is illustrated with a case-study modelling inter-group bias.

2 Preliminaries

This paper uses notation and background from linear algebra [15]. This section summarizes the main aspects of the notation and the background. Further details on special types of matrices and dynamical systems [5] can be found in Appendix A.

The term N is a positive integer and is used to identify the number of agents in a network. The expressions \mathbb{R}^N and $\mathbb{R}^{N \times N}$ denote the collections, respectively, of real-valued vectors of length N and real-valued squared matrices of side N . Boldface lowercase letters identify vectors in \mathbb{R}^N and boldface uppercase letters identify matrices in $\mathbb{R}^{N \times N}$. The expression \mathbf{I} denotes the identity matrix in $\mathbb{R}^{N \times N}$ and $\det(_)$ the determinant (partial) function on matrices. Recall that if the inverse \mathbf{A}^{-1} of \mathbf{A} exists, it satisfies $\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$. The L1 norm of a vector \mathbf{x} is denoted $\|\mathbf{x}\|_1$. When the values of a vector \mathbf{x} need to be made explicit, it can be denoted as $\mathbf{x} = [x_i]_{i \in I}$ or $[x_i]$ when the index set I can be inferred from the context. Similarly, when the values of a matrix \mathbf{A} need to be made explicit, it can be denoted as $\mathbf{A} = [a_{ij}]_{(i,j) \in I \times J}$ or $[a_{ij}]$ when the index sets I and J can be inferred from the context. It is said that \mathbf{A} is *non-negative*, denoted $\mathbf{A} \geq 0$, iff $a_{ij} \geq 0$ for $1 \leq i, j \leq N$. Likewise, \mathbf{A} is *positive*, denoted $\mathbf{A} > 0$, iff $a_{ij} > 0$ for $1 \leq i, j \leq N$. A matrix \mathbf{A} is primitive iff $\mathbf{A}^m > 0$ for some $m \in \mathbb{N}$ [15]. A (square) matrix \mathbf{P} is called a *permutation matrix* when it has exactly one entry of 1 in each row and each column, and 0s elsewhere. A matrix \mathbf{A} is *reducible* if there is a permutation matrix \mathbf{P} such that $\mathbf{P}^T \mathbf{A} \mathbf{P} = \begin{bmatrix} \mathbf{B} & \mathbf{C} \\ \mathbf{0} & \mathbf{D} \end{bmatrix}$ where \mathbf{B} and \mathbf{D} are square matrices (possibly of different sizes), \mathbf{C} is a matrix of appropriate size, and $\mathbf{0}$ is the zero matrix of appropriate size. If no such permutation matrix \mathbf{P} exists, then \mathbf{A} is called *irreducible*.

A non-zero vector \mathbf{x} is an *eigenvector* of a matrix \mathbf{A} with *eigenvalue* $\lambda \in \mathbb{C}$ iff $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$ (i.e., applying \mathbf{A} to the eigenvector \mathbf{x} only scales the eigenvector by the scalar value λ). The *spectrum* of \mathbf{A} is the set $\sigma(\mathbf{A})$ of all eigenvalues λ satisfying the equation $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$. The *geometric multiplicity* of $\lambda \in \sigma(\mathbf{A})$, denoted $m_g(\lambda)$, is the number of eigenvectors associated to λ in the (spectral) decomposition of \mathbf{A} . The *algebraic multiplicity* of $\lambda \in \sigma(\mathbf{A})$, denoted $m_a(\lambda)$, is the number of times λ appears as a root of the characteristic polynomial of \mathbf{A} . It is known that $m_g(\lambda) \leq m_a(\lambda)$ for any $\lambda \in \sigma(\mathbf{A})$. The *spectral radius* of \mathbf{A} is the real value $\rho(\mathbf{A})$ defined by $\rho(\mathbf{A}) = \max_{\lambda \in \sigma(\mathbf{A})} |\lambda|$ identifying the greatest magnitude among all eigenvalues of \mathbf{A} . If $\lambda \in \sigma(\mathbf{A})$ and it is the only eigenvalue that satisfies $|\lambda| = \rho(\mathbf{A})$, then λ is called the *dominant eigenvalue* of \mathbf{A} .

The *Perron-Frobenius Theorem* [15] is a fundamental result about non-negative matrices. For irreducible $\mathbf{A} \in \mathbb{R}^{N \times N} \geq 0$, there exists a real and positive eigenvalue $\lambda_1 = \rho(\mathbf{A})$, called the *Perron-Frobenius eigenvalue*, there exists a corresponding eigenvector \mathbf{v} to λ_1 , called the *Perron-Frobenius eigenvector*, such that \mathbf{v} has positive components (it can always be picked to have Euclidean norm equal to 1), and the eigenvalue λ_1 is *simple* in the sense that $m_a(\lambda_1) = 1$.

3 Generalized-bias Opinion Models

This section presents a generalization of the DeGroot model to capture biases on opinion differences. It can account, among others, for dynamic influence/weights and inter-group biases [14] where agents favor one's own group over other groups.

In social learning models, a community, society, or network is typically represented as a directed weighted graph. The edges between individuals (agents) specify the direction and strength of the influence that one carries over the other.

Definition 1. An (N -agent) influence graph is a directed weighted graph $G = (A, E, I)$ with vertices $A = \{1, \dots, N\}$, edges $E \subseteq A \times A$, and weight function $I = A \times A \rightarrow [0, 1]$ satisfying $I(i, j) = 0$ iff $(i, j) \notin E$.

The vertices A represent the N agents of a given community or network. Whenever $(i, j) \in E$, it means that agent i influences agent j . The value $I(i, j)$, for simplicity written $I_{i,j}$, denotes the strength of the influence: 0 means no influence and a higher value means stronger influence. The expression A_i is used to denote the set of (inbound-)neighbors of agent i , i.e., the set of agents j with $I_{j,i} > 0$. Recall that a graph is *strongly connected* if there is at least one path between any two vertices. In a strongly connected influence graph, all the agents influence –directly or indirectly– one another.

The evolution of agents’ opinions, about some underlying statement or proposition, is modeled similar to the DeGroot-like models in [9,1,2].

An *opinion state* (or *belief state*) of the N agents is represented as a vector in $[0, 1]^N$. For any opinion state $\mathbf{x} = [x_1, \dots, x_N]$ and $1 \leq i \leq N$, the expression $\mathbf{x}_i = x_i$ denotes the opinion of agent i . If $\mathbf{x}_i = 0$, then i completely disagrees with the proposition; if $\mathbf{x}_i = 1$, then i completely agrees with it. The higher the value of \mathbf{x}_i , the stronger the agreement.

Definition 2. An opinion model is a tuple $M = (G, \mathbf{x}^0, \mu_G)$, where G is an N -agent influence graph, \mathbf{x}^0 is the initial state of opinion, and $\mu_G : [0, 1]^N \rightarrow [0, 1]^N$ is a state-transition function, called update function. For every $t \in \mathbb{N}$, $\mathbf{x}^{t+1} = \mu_G(\mathbf{x}^t)$ is the state of opinion at time $t + 1$ in M .

An update function can be used to express any deterministic and discrete transition from one opinion state to the next, possibly taking into account the influence graph. Intuitively, these update functions specify the reaction of an agent to the opinion disagreements with each one of its influencers.

Consensus is a central property in social learning (e.g., the inability to reach consensus is a sign of a polarized society).

Definition 3. Let (G, \mathbf{x}^0, μ_G) be an opinion model with $G = (A, E, I)$. A subset of agents $B \subseteq A$ is said to converge to an opinion value v iff for every $i \in B$, $\lim_{t \rightarrow \infty} \mathbf{x}_i^t = v$. The subset B is said to converge to consensus iff B converges to some opinion value v .

3.1 DeGroot Update

The *normalized influence* $\overline{I}_{j,i} = \frac{I_{j,i}}{\sum_{k \in A_i} I_{k,i}}$ is used to model the opinion update of agent i as a function of its neighbors’ opinions [2]. Then, the standard DeGroot model [9] is obtained by the update function $\mu_G(\mathbf{x})_i = \sum_{j \in A_i} \overline{I}_{j,i} \mathbf{x}_j$ which can be equivalently formulated as in Equation 1:

$$\mu_G(\mathbf{x})_i = \mathbf{x}_i + \sum_{j \in A_i} \overline{I}_{j,i} (\mathbf{x}_j - \mathbf{x}_i). \quad (1)$$

If j influences i , then i 's opinion would tend to move closer to j 's. The *disagreement term* $(\mathbf{x}_j - \mathbf{x}_i) \in [-1, 1]$ in Equation 1 realizes this intuition. If $(\mathbf{x}_j - \mathbf{x}_i)$ is a negative term in the sum, the disagreement can be thought of as contributing with a magnitude of $|\mathbf{x}_j - \mathbf{x}_i|$ (multiplied by $\overline{I_{j,i}}$) to *decreasing* i 's belief in the underlying proposition. Similarly, if $(\mathbf{x}_j - \mathbf{x}_i)$ is positive, the disagreement contributes with the same magnitude but to *increasing* i 's belief.

It is known that in standard DeGroot, two conditions on the influence graph are enough for the agents converge to consensus. First, the all the agents are required to influence, directly or indirectly, each other. Second, each agent has to have some self-influence.

Theorem 1 ([9]). *Let (G, \mathbf{x}^0, μ_G) be an opinion model with $G = (A, E, I)$ and μ_G defined by Eq. 1. If G is (1) strongly connected and (2) for each $i \in A$, $\overline{I_{j,i}} < 1$ for some $j \in A_i$, then A converges to consensus.*

3.2 Updates with Bias Factors

A broad class of update functions, generalizing the original one in DeGroot, is introduced in this section. More precisely, updates for an agent i can *weight* disagreements with each one of its neighbors j using functions α_{ij} from opinion states to $[0, 1]$ referred to as (*generalized*) *bias factors*. The same opinion difference can then be weighted differently by bias factors, depending of opinions of agents i and j . Thus, intuitively α_{ij} may also be seen as *changing* the influence of j over i depending on their opinions.

As an example, assume that 0.5 represents a neutral opinion about an underlying proposition. If $\mathbf{x}_i = 0.4$ and $\mathbf{x}_j = 0.1$, the agents can be viewed as somewhat disagreeing and strongly disagreeing, respectively, with the underlying proposition. If $\mathbf{x}_i = 0.7$ and $\mathbf{x}_j = 0.4$, agent i seems to agree somewhat with the proposition, while j does not. Alas, in both cases $(\mathbf{x}_j - \mathbf{x}_i) = 0.3$. It could be desirable to have $\alpha_{i,j}(\mathbf{x})$ assigning a higher weight in the first situation because these opinions come from a group that disagrees with the proposition. (This cannot be achieved with the influence graph alone since the opinion of agents may change while their influence in an opinion model does not.) As a matter of fact, this is an instance of in-group bias.

Definition 4. *Let (G, \mathbf{x}^0, μ_G) be an opinion model with $G = (A, E, I)$. The function $\mu_G : [0, 1] \rightarrow [0, 1]$ is an generalized bias update iff, for each $i \in A$, it is defined by:*

$$\mu_G(\mathbf{x})_i = \mathbf{x}_i + \sum_{j \in A_i} \overline{I_{j,i}} \alpha_{i,j}(\mathbf{x})(\mathbf{x}_j - \mathbf{x}_i), \quad (2)$$

with each bias factor $\alpha_{i,j} : [0, 1]^N \rightarrow [0, 1]$ satisfying:

1. $\alpha_{i,j}$ is a continuous function,
2. if $\mathbf{x}_i \neq \mathbf{x}_j$, then $\alpha_{i,j}(\mathbf{x}) > 0$, and
3. for each $i \in A$, there exists $j \in A_i$, such that $\alpha_{i,j}(\mathbf{x}) < 1$.

The first condition on the bias factor $\alpha_{i,j}$ guarantees that small changes in the state of opinion lead to small changes in the factor result. Without the second condition, the

bias factor would cancel the influence of the contact j , which may break the strong connectivity of the graph. The third and last condition reproduces the second condition in the Theorem 1 in the original DeGroot model.

Note that the update function in Equation 2 can be expressed equivalently as:

$$\mu_G(\mathbf{x})_i = \left(1 - \sum_{j \in A_i} \overline{I_{j,i}} \alpha_{i,j}(\mathbf{x}) \right) \mathbf{x}_i + \sum_{j \in A_i} \overline{I_{j,i}} \alpha_{i,j}(\mathbf{x}) \mathbf{x}_j, \quad (3)$$

Furthermore, Equation 2 is a *convex combination* of $[0, 1]$ elements and thus $\mu_G(\mathbf{x})_i \in [0, 1]$ because $[0, 1]$ is a convex set. Consequently, the generalized-bias update function can be expressed in vectorial form as:

$$\mu_G(\mathbf{x}) = \mathbf{A}(\mathbf{x})\mathbf{x}, \quad (4)$$

where $\mathbf{A} : [0, 1]^N \rightarrow \mathbb{R}^{N \times N}$ is a square matrix with entries

$$a_{i,j}(\mathbf{x}) = \begin{cases} 1 - \sum_{k \in A_i} \overline{I_{k,i}} \alpha_{i,k}(\mathbf{x}) & , \text{ if } i = j \\ \overline{I_{j,i}} \alpha_{i,j}(\mathbf{x}) & , \text{ if } i \neq j. \end{cases} \quad (5)$$

This section is concluded by introducing the notion of generalized bias opinion models.

Definition 5. Let $M = (G, \mathbf{x}^0, \mu_G)$ be an opinion model with $G = (A, E, I)$. Then, M is an generalized-bias (opinion) model iff μ_G can be expressed as in Equation 4.

4 Consensus and Expressiveness

This section presents a consensus result for generalized-bias opinion models (Definition 5). It is the main theoretical result of the paper and extends the DeGroot's consensus result (see Theorem 1). This section also shows how generalized-bias models capture the opinion specification and dynamics of other models, including state-of-the-art [1,2,9]. It highlights dynamics that generalized-bias models can handle, but the other models cannot.

Lemma 1. Let (G, \mathbf{x}^0, μ_G) be an generalized-bias model, with $G = (A, E, I)$ and $\mu_G(\mathbf{x}) = \mathbf{A}(\mathbf{x})\mathbf{x}$. The matrix $\mathbf{A}(\mathbf{x})$ is (right-)stochastic for all $\mathbf{x} \in [0, 1]^N$.

Recall that a matrix $\mathbf{A} = [a_{ij}] \in \mathbb{R}^{N \times N}$ is (right-)stochastic if its entries satisfy $a_{ij} \geq 0$ and $\sum_{j=1}^N a_{ij} = 1$. This guarantees that *consensual* state of opinion vectors (i.e., vectors in which all agents have the same opinion) in generalized-bias models are *fixed points* of their update function. Intuitively, this means that once the agents are in consensus, they will remain in consensus.

Lemma 2. Let (G, \mathbf{x}^0, μ_G) be a generalized-bias model, with $G = (A, E, I)$ and $\mu_G(\mathbf{x}) = \mathbf{A}(\mathbf{x})\mathbf{x}$. If G is a strongly connected graph, then for every $t > 0$, $\mathbf{A}(\mathbf{x}^t)$ is primitive.

Generalized-bias models have *unique* fixed points. The primitivity of $\mathbf{A}(\mathbf{x}_t)$, for $t > 0$, indicates that they are consensual vectors. This is a key observation for proving that generalized-bias models converge to consensus, as it ensures that for each state opinion vector \mathbf{x}^t , its update \mathbf{x}^{t+1} is closer to consensus than \mathbf{x}^t .

Theorem 2 (Consensus for Generalized-Bias Models). *Let (G, \mathbf{x}^0, μ_G) be a generalized-bias model, with $G = (A, E, I)$ and $\mu_G(\mathbf{x}) = \mathbf{A}(\mathbf{x})\mathbf{x}$. If G is a strongly connected graph, then A converges to consensus.*

The connectivity constraint on G ensures that each pair of agents (i, j) is connected, directly or indirectly. Hence, the opinion of i is in a certain way influenced by the opinion of j and viceversa. The non-zero condition of the bias factors ensures that no matter the difference of opinions between two agents, if there is a connection between them, this influence, and hence their disagreement, will not be ignored but it would be dynamically weighted. Intuitively, the generalized-bias models guarantee that in an environment where all users can freely express their opinions and be heard, they will eventually come to an agreement even if there are certain cognitive biases that regulate the influences between them.

The proofs of the above lemmas and Theorem 2 can be found in Appendix B.

Expressiveness of Generalized-Bias Models. The work in [2] introduces *disagreement-bias (opinion) models*. They are extensions of DeGroot's of the form (G, \mathbf{x}^0, μ_G) with

$$\mu_G(\mathbf{x})_i = \mathbf{x}_i + \sum_{j \in A_i} \overline{I}_{j,i} \beta_{i,j}(\mathbf{x}_j - \mathbf{x}_i), \quad (6)$$

where each $\beta_{i,j} : [-1, 1] \rightarrow [-1, 1]$ is a function referred to as *disagreement bias*. Intuitively, $\beta_{i,j}(\mathbf{x}_j - \mathbf{x}_i)$ is a factor that depends only on the difference of opinion between j and i .

It is shown that the agents in a disagreement-bias model converge to consensus if its influence graph is strongly connected and each $\beta_{i,j}$ function is continuous and in the region $\mathbf{R} = \{(x, y) \in [-1, 1]^2 : xy > 0 \text{ and } |y| < |x|\} \cup \{(0, 0)\}$. The models with all their bias functions in \mathbf{R} can capture a wide range of biases, such as variations of confirmation bias as well as the dynamics of receptive and resistant agents (i.e., agents willing to change their opinion towards their contacts but with some skepticism) [4].

Theorem 3 ([2]). *Let (G, \mathbf{x}^0, μ_G) with $G = (A, E, I)$ be a disagreement-bias model whose disagreement bias functions are continuous and in \mathbf{R} . If G is strongly connected, A converges to consensus.*

Theorem 3 is actually a consequence of Theorem 2 from this paper. It can be shown that the update functions of disagreement-bias models, whose continuous bias functions are in \mathbf{R} , can be expressed as update functions of generalized-bias models. It can also be shown that there are updates functions from generalized-bias models that cannot be expressed as the update functions of a disagreement-bias model. More precisely,

Theorem 4 (Expressiveness). *(1) Let (G, \mathbf{x}^0, ν_G) be a disagreement-bias model whose disagreement bias functions are continuous and in \mathbf{R} . Then, there exists a generalized-bias model (G, \mathbf{x}^0, μ_G) , such that $\mu_G = \nu_G$. (2) There exists a generalized-bias model (G, \mathbf{x}^0, μ_G) such that every disagreement-bias model (G, \mathbf{x}^0, ν_G) satisfies $\mu_G \neq \nu_G$.*

Proof. Part (1). Given any disagreement-bias model (G, \mathbf{x}^0, ν_G) whose disagreement biases are continuous and in \mathbf{R} , define the generalized-bias model (G, \mathbf{x}^0, μ_G) where:

$$\alpha_{i,j}(\mathbf{x}) = \begin{cases} \frac{\beta_{i,j}(\mathbf{x}_j - \mathbf{x}_i)}{\mathbf{x}_j - \mathbf{x}_i} & , \text{ if } \mathbf{x}_j \neq \mathbf{x}_i, \\ 0 & , \text{ if } \mathbf{x}_j = \mathbf{x}_i. \end{cases} \quad (7)$$

Here, $\beta_{i,j}$ is the corresponding disagreement bias in (G, \mathbf{x}^0, ν_G) . It can be verified that $0 < \alpha_{i,j}(\mathbf{x}) < 1$. Also, $\alpha_{i,j}$ is continuous because $\beta_{i,j}$ is continuous. Thus, $\alpha_{i,j}$ satisfies the conditions for the bias factors in Definition 4; it can be verified that $\mu_G = \nu_G$.

For Part (2), consider the generalized-bias model (G, \mathbf{x}^0, μ_G) , with $G = (A, E, I)$, $A = \{1, 2\}$, $E = \{(1, 2), (2, 1)\}$, and $I_{12} = I_{21} = 1.0$, and:

$$\mu_G(\mathbf{x})_1 = \mathbf{x}_1 + \alpha_{12}(\mathbf{x}_1, \mathbf{x}_2)(\mathbf{x}_2 - \mathbf{x}_1), \quad \mu_G(\mathbf{x})_2 = \mathbf{x}_2 + \alpha_{21}(\mathbf{x}_1, \mathbf{x}_2)(\mathbf{x}_2 - \mathbf{x}_1). \quad (8)$$

Suppose that $\alpha_{12}(\mathbf{x}_1, \mathbf{x}_2) = \alpha_{21}(\mathbf{x}_1, \mathbf{x}_2) = \alpha(\mathbf{x}_1, \mathbf{x}_2)$, where

$$\alpha(\mathbf{x}) = \alpha(\mathbf{x}_1, \mathbf{x}_2) = 2(\mathbf{x}_1 - \mathbf{x}_1^2 + 0.25)(1 - 0.9|\mathbf{x}_1 - \mathbf{x}_2|). \quad (9)$$

Towards of contradiction, suppose that there exists a disagreement-bias model (G, \mathbf{x}^0, ν_G) such that $\nu_G = \mu_G$. Then, the update function ν_G must have the form

$$\nu_G(\mathbf{x})_1 = \mathbf{x}_1 + \beta_{12}(\mathbf{x}_2 - \mathbf{x}_1) \quad \nu_G(\mathbf{x})_2 = \mathbf{x}_2 + \beta_{21}(\mathbf{x}_1 - \mathbf{x}_2), \quad (10)$$

with $\beta_{12}(\mathbf{x}_1 - \mathbf{x}_2) = \alpha_{12}(\mathbf{x}_1, \mathbf{x}_2)(\mathbf{x}_2 - \mathbf{x}_1)$ and $\beta_{21}(\mathbf{x}_2 - \mathbf{x}_1) = \alpha_{21}(\mathbf{x}_1, \mathbf{x}_2)(\mathbf{x}_2 - \mathbf{x}_1)$. However, for the opinion states $\mathbf{y} = [0, 0.2]$ and $\mathbf{z} = [0.5, 0.7]$, it follows that $-0.2\alpha_{12}(0, 0.2) = \beta_{21}(-0.2) = -0.2\alpha_{12}(0.5, 0.7)$. This is a contradiction because $\alpha_{12}(0, 0.2) = 0.41$ and $\alpha_{12}(0.5, 0.7) = 0.82$. \square

5 Intergroup Bias: A Case Study

Intergroup bias is a well-documented phenomenon where individuals tend to evaluate members of their own group more favorably than members of other groups [14]. This section will illustrate opinion dynamics under this bias as a particular case of a generalized-bias model.

Consider a scenario involving ideologically distinct groups that traditionally clash, such as *progressives* (left-wingers), *moderates* (centrists), and *conservatives* (right-wingers). To illustrate how these groups might be divided by their opinions on a controversial proposition, the statement “free markets promote prosperity” is used. Division is represented by splitting the interval $[0, 1]$ into disjoint sub-intervals, corresponding to each group’s agreement with the proposition. Define $I_1 = [0, 0.45)$, $I_2 = [0.45, 0.55]$ and $I_3 = (0.55, 1]$ for the progressive, moderate and conservative opinion groups.

Assessment Function. To capture inter-group bias, the bias factor α_{ij} is defined as an *assessment* function $a(x, y)$ assigning a weight depending on the groups x and y belong to. Let $w_i(x)$ be the characteristic function of the set I_i (i.e., it is 1 if $x \in I_i$ and otherwise 0) and $u_{ij}(x, y) = w_i(x)w_j(y)$ (i.e., it is 1 if $x \in I_i$ and $y \in I_j$, and otherwise 0).

Characteristic functions are used so that the assessment function, $a(x, y)$, first detects the group of the first agent and then detects the group of the second agent. This will help in calculating the value that the first agent assigns to the opinion of the second. For detecting the group of the evaluating agent, the $w_i(x)$ function is used. To calculate the assessment that this agent gives to the second agent, the $u_{ij}(x, y)$ function is used. It then proceeds by defining the functions that calculate the assessments that the first agent has of the second agent depending on their groups: $a_1(x, y)$ when x is progressive, $a_2(x, y)$ when x is moderate, and $a_3(x, y)$ when x is conservative. These functions are combined by multiplying each one by the function that detects the group of x .

First, start with x in the progressives group (I_1). The group of y is identified and the corresponding value is defined according to function a_1 :

$$a_1(x, y) = u_{11}(x, y) + u_{12}(x, y)\varphi(y) + 0.5u_{13}(x, y), \quad (11)$$

with $\varphi(y) = 1 - 5(y - 0.45)$. If y is a progressive opinion as well, then $u_{11}(x, y) = 1$ and $u_{12}(x, y) = 0 = u_{13}(x, y)$; thus the weight is 1. If y is a moderate opinion, the weight is given by $\varphi(y)$; a decreasing linear function stating that when the distance grows, the weight decreases. Notice that $\varphi(0.45) = 1$ and $\varphi(0.55) = 0.5$. Finally, if y is a conservative opinion the weight is equal to 0.5. Clearly, $a_1(x, y)$ is continuous in $x \in I_1$ and $y \in [0, 1]$.

Second is the case when x is a moderate opinion (I_2), handled with a_2 :

$$a_2(x, y) = \varphi(x)u_{21}(x, y) + \kappa(x, y)u_{22}(x, y) + \psi(x)u_{23}(x, y), \quad (12)$$

with $\psi(x) = 0.5 + 5(x - 0.45)$ and $\kappa(x, y) = 100xy - 50x - 50y + 25.75$. It can be checked that $\kappa(0.45, y) = \varphi(y)$ and $\kappa(0.55, y) = \psi(y)$; thus, $a_2(x, y)$ is continuous function in $x \in I_2$ and $y \in [0, 1]$.

Third, the case when x is a conservative opinion (I_3). The assessment is given by function a_3 :

$$a_3(x, y) = 0.5u_{31}(x, y) + \psi(y)u_{32}(x, y) + u_{33}(x, y), \quad (13)$$

which is continuous in $x \in I_3$ and $y \in [0, 1]$.

With this setup, the continuous assessment function a (see Figure 1):

$$a(x, y) = w_1(x)a_1(x, y) + w_2(x)a_2(x, y) + w_3(x)a_3(x, y). \quad (14)$$

Note that function $a(x, y)$ is continuous, positive for all $(x, y) \in [0, 1]^2$, and if each agent in the graph is connected to a member of another group as in the example below, then the assessment function is indeed a bias factor (Definition 4).

The update function for inter-group bias is finally defined as:

$$\mu_G(\mathbf{x}) = \mathbf{x}_i + \sum_{j \in A_i} \bar{I}_{j,i} \alpha_{ij}(\mathbf{x}_i, \mathbf{x}_j)(\mathbf{x}_i - \mathbf{x}_j), \quad (15)$$

where $\alpha_{ij}(\mathbf{x}_i, \mathbf{x}_j) = a(\mathbf{x}_i, \mathbf{x}_j)$.

The following example illustrates evolution of opinion under this form of inter-group bias.

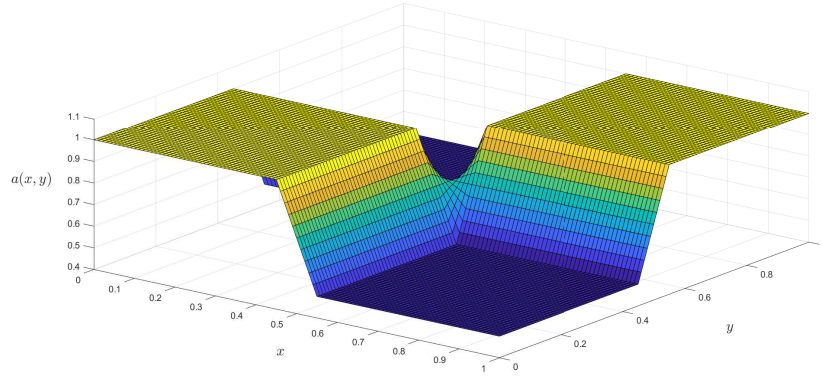
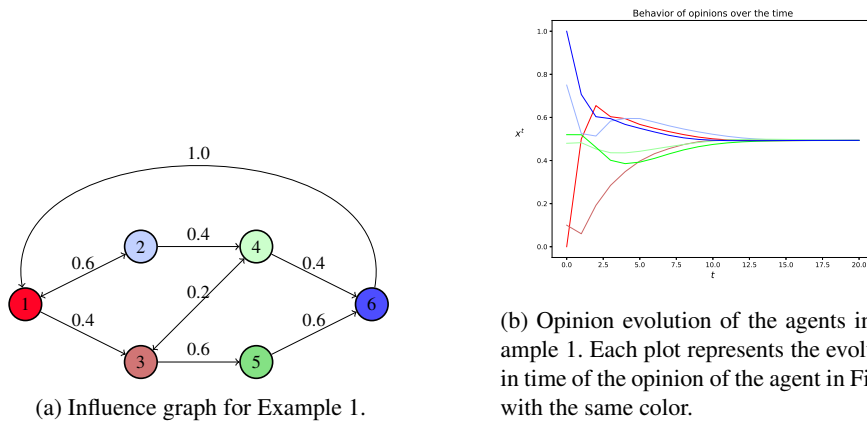


Fig. 1: Surface of function $a(x, y)$ defined in Eq. 14.

Example 1. Let (G, \mathbf{x}^0, μ_G) be a generalized-bias model for six agents where G is the graph in Figure 2a. The initial state of opinion is $\mathbf{x}^0 = [0.0, 0.75, 0.1, 0.48, 0.52, 1.0]$ and μ_G is the inter-group bias update in Equation 15. Initially, there are two agents that are progressives (agents in red), two moderates (agents in green), and two conservatives (agents in blue). Note that G is strongly connected. Because of the influence graph and the initial state opinions, there are interactions between agents of different groups, as well as in the same group. Figure 2b illustrates the evolution of opinion of the agents converging to consensus (around a moderate opinion) as predicted by the consensus result in Theorem 2.



(a) Influence graph for Example 1.

(b) Opinion evolution of the agents in Example 1. Each plot represents the evolution in time of the opinion of the agent in Fig. 2a with the same color.

Fig. 2: Influence graph and opinion evolution.

6 Concluding Remarks

This paper introduced the notion of generalized-bias opinion dynamic model, a generalization of the classic DeGroot model, in which agents may react differently to the same opinion difference depending on whether the person holding the opposing view is from the same group. These reactions were formalized as arbitrary functions—parametric to the models—that depend, not only on opinion difference, but also on the opinions of the individuals interacting. This is a well-known bias in social psychology that cannot be modeled by state-of-the-art DeGroot-based models, as shown in Theorem 4. Furthermore, a consensus result was provided (Theorem 2), extending the classic result of the DeGroot model [9] for a wide range of social biases. Finally, an application example involving a particular case study of inter-group bias was used to illustrate the expressiveness of the new model.

The relevance of biased reasoning in human interactions has been extensively studied in [22], [18], [20], and others. Previous research in social psychology has explored intergroup bias [11,10,17], concluding that fostering dialogue between members of different groups reduces the disparity in their opinions. These findings align with the consensus results and the case study presented in the current work. The present model may bring further insights to social scientists about cognitive bias in opinion formation.

There is a great deal of work on formal models for belief change in social networks; we focus on the work on biased belief updates, which is the focus of this paper. Some models have previously been proposed to generalize the DeGroot model and introduce bias. For instance, [8], [7], and [23] analyze the effects of incorporating a bias factor for each agent to represent biased assimilation—how much of the external opinions the agent will take into consideration. The main difference between these models and the generalized-bias models in this paper is that in the previous models, biases are not represented as arbitrary continuous functions satisfying two conditions. Instead, biases are either modeled as an exponential factor that reduces the impact of neighbors’ opinions or by dynamically changing the weights in the DeGroot model with a specific function. Thus, the new model provides greater flexibility for capturing a wider range of biases.

The generalized-bias opinion models introduced here assume synchronous communication among the agents in the network, meaning all agents update their opinions simultaneously. In the line of [16,3,13], agents can also communicate asynchronously or via a hybrid blend where both types of communication can coexist. It would be interesting to see if the results presented here can be extended to these two settings. Moreover, it would be worth pursuing the use of rewriting logic to specify generalized-bias opinion models and perform several forms of formal analysis, including reachability analysis, model checking, and statistical model checking of concrete instances and desired properties of these models. Finally, the model could be extended to include agents that can learn by exchanging beliefs and lies as done in [13].

References

1. M. S. Alvim, B. Amorim, S. Knight, S. Quintero, and F. Valencia. A formal model for polarization under confirmation bias in social networks. *LMCS*, 2023.

2. M. S. Alvim, A. G. da Silva, S. Knight, and F. Valencia. A multi-agent model for opinion evolution in social networks under cognitive biases. In *Formal Techniques for Distributed Objects, Components, and Systems*, volume 14678 of *LNCS*, pages 3–19. Springer, 2024.
3. J. Aranda, S. Betancourt, J. F. Díaz, and F. Valencia. Fairness and consensus in an asynchronous opinion model for social networks. In *35th International Conference on Concurrency Theory, CONCUR*. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2024.
4. E. Aronson, T. Wilson, and R. Akert. *Social Psychology*. Upper Saddle River, NJ : Prentice Hall, 7 edition, 2010.
5. M. Bernardo, C.J.Budd, A.R.Champneys, and P.Kowalczyk. *Piecewise-smooth Dynamical Systems Theory and Applications*. Springer Science, 2008.
6. A. G. Chandrasekhar, H. Larreguy, and J. P. Xandri. Testing models of social learning on networks: Evidence from a lab experiment in the field. Working Paper 21468, National Bureau of Economic Research, August 2015.
7. Z. Chen, J. Qin, B. Li, H. Qi, P. Buchhorn, and G. Shi. Dynamics of opinions with social biases. *Automatica*, 106:374–383, 2019.
8. P. Dandekar, A. Goel, and D. Lee. Biased assimilation, homophily and the dynamics of polarization. *Proceedings of the National Academy of Sciences of the United States of America*, 110, 03 2013.
9. M. H. DeGroot. Reaching a consensus. *Journal of the American Statistical association*, 69(345):118–121, 1974.
10. S. L. Gaertner and J. F. Dovidio. *Reducing intergroup bias: The common ingroup identity model*. Psychology Press, 2014.
11. S. L. Gaertner, J. F. Dovidio, M. C. Rust, J. A. Nier, B. S. Banker, C. M. Ward, G. R. Mottola, and M. Houlette. Reducing intergroup bias: elements of intergroup cooperation. *Journal of personality and social psychology*, 76(3):388, 1999.
12. B. Golub and E. Sadler. Learning in social networks. Available at SSRN 2919146, 2017.
13. S. Haar, S. Perchy, C. Rueda, and F. Valencia. An algebraic view of space/belief and extrusion/utterance for concurrency/epistemic logic. In *17th International Symposium on Principles and Practice of Declarative Programming (PPDP 2015)*, 2015.
14. M. Hewstone, M. Rubin, and H. Willis. Intergroup bias. *Annual review of psychology*, 53(1):575–604, 2002.
15. C. D. Meyer. *Matrix Analysis and Applied Linear Algebra*. SIAM, 2000.
16. C. Olarte, C. Ramírez, C. Rocha, and F. Valencia. Unified opinion dynamic modeling as concurrent set relations in rewriting logic. In *15th International Workshop, WRLA, Revised Selected Papers*. Springer, 2024.
17. T. F. Pettigrew and L. R. Tropp. How does intergroup contact reduce prejudice? meta-analytic tests of three mediators. *European journal of social psychology*, 38(6), 2008.
18. V. J. Ramos. Analyzing the role of cognitive biases in the decision-making process. *Advances in Psychology, Mental Health, and Behavioral Studies*, 2018.
19. H. Tajfel. An integrative theory of intergroup conflict. *The social psychology of intergroup relations/Brooks/Cole*, 1979.
20. B. M. Tappin and S. Gadsby. Biased belief in the bayesian brain: A deeper look at the evidence. *Consciousness and Cognition*, 68:107–114, 2019.
21. S. Wasserman and K. Faust. Social network analysis in the social and behavioral sciences. In *Social Network Analysis: Methods and Applications*. Cambridge University Press, 1994.
22. D. Williams. Hierarchical bayesian models of delusion. *Consciousness and Cognition*, 61:129–147, 2018.
23. W. Xia, M. Ye, J. Liu, M. Cao, and X.-M. Sun. Analysis of a nonlinear opinion dynamics model with biased assimilation. *Automatica*, 120:109113, 2020.

A Matrices and Dynamical Systems

A.1 Matrices

Square matrices in $\mathbb{R}^{N \times N}$ can be seen as representations of directed graphs with N vertices and vice-versa.

Definition 6. The graph $G_{\mathbf{A}}$ associated to a matrix $\mathbf{A} = [a_{ij}] \in \mathbb{R}^{N \times N}$ is defined by $G_{\mathbf{A}} = (V_{\mathbf{A}}, E_{\mathbf{A}})$ with N nodes $V_{\mathbf{A}} = \{1, \dots, N\}$ and edges $E_{\mathbf{A}} = \{(i, j) \mid a_{ij} = 1\}$.

Structural properties of (square) matrices may be studied from the structure and connectivity in the associated graphs. Recall that a graph is *strongly connected* if it has exactly one strongly connected component (i.e., there is a path between any pair of its nodes).

Theorem 5. A matrix \mathbf{A} is irreducible iff $G_{\mathbf{A}}$ is strongly connected.

Stochastic matrices are a special case of non-negative matrices. They are also known as a probability matrices, transition matrices, or Markov matrices, and are used to describe the transitions of a Markov chain.

Definition 7. A matrix $\mathbf{A} = [a_{ij}] \in \mathbb{R}^{N \times N}$ is called (right-)stochastic iff the following conditions hold on \mathbf{A} :

1. It is non-negative.
2. The sum of each one of its rows is 1, i.e., for each $1 \leq i \leq N$, it satisfies:

$$\sum_{j=1}^N a_{ij} = 1.$$

It can be checked that any stochastic matrix \mathbf{A} has an eigenvalue equal to 1, which corresponds to its spectral radius. This follows from the fact that the sum of each row of \mathbf{A} is 1, and hence there is an eigenvalue equal to 1 associated to a stationary distribution vector (i.e., a probability vector) \mathbf{v} satisfying $\mathbf{A}\mathbf{v} = \mathbf{v}$.

A.2 Dynamical Systems

A dynamical system is a formal concept used to describe a system that evolves over time according to a specific set of rules.

Definition 8. A (discrete) dynamical system is a pair $\mathcal{S} = (S, \phi)$ where S is a set and $\phi = \{\phi_t\}_{t \in T}$ is a family of T -indexed functions $\phi_t : S \rightarrow S$, with T a semigroup, satisfying for any $x \in S$ and $s, t \in T$:

1. $\phi_0(x) = x$.
2. $\phi_s(\phi_t(x)) = \phi_{s+t}(x)$.

For convenience, the expression $\phi_t(x)$ is abbreviated x_t , for any $x \in S$ and $t \in T$.

In a dynamical system $\mathcal{S} = (S, \phi)$, the set S represents all possible states of the system and the iterative application of ϕ defines the evolution of the system. The *orbit* of a state $x \in S$ is the set $\{x_t \mid t \in T\}$ of all states reachable from x . A set $\Lambda \subseteq S$ is called an *invariant set* of \mathcal{S} iff $\phi_t(x) \in \Lambda$ for any $x \in \Lambda$ and $t \in T$. An invariant set is a subset of the state space that, once entered by a trajectory of the system, it cannot be left. In other words, if the system's state ever enters an invariant set, it will remain in that set for all future time steps. Since an invariant set is a subset of the state space that is closed under transitions, it can include the orbits of many elements in the state space.

For the purpose of this work, the set S is assumed to be \mathbb{R}^N and the index set T is the set of natural numbers \mathbb{N} .

Example 2. Consider the dynamical system $\mathcal{S} = (\mathbb{R}^2, \phi)$, with $\phi = \{f^n\}_{n \in \mathbb{N}}$ and $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ defined by:

$$f(\mathbf{x}) = \mathbf{Ax} = \begin{bmatrix} 0.5 & 0.5 \\ 0.2 & 0.8 \end{bmatrix} \mathbf{x}.$$

Note that f is the application of the irreducible stochastic matrix \mathbf{A} on a given vector \mathbf{x} . By the Perron-Frobenius Theorem (see Section 2), the spectral radius $\rho(\mathbf{A}) = 1$ has multiplicity 1. Therefore, there is exactly one eigenvector \mathbf{v} (modulo scalar multiplication) satisfying $\mathbf{x} = f(\mathbf{x})$. Consequently, the identity over \mathbb{R}^2 is an invariant set for \mathcal{S} .

An attractor in a dynamical system is an invariant set toward which the system tends to evolve from a wide variety of initial conditions. Attractors can play a crucial role in understanding the long-term behavior of dynamical systems.

Definition 9. Let $\mathcal{S} = (S, \phi)$ be a dynamical system. An attractor is a set of states $U \subseteq S$ with the following properties:

1. U is an invariant set.
2. For any point $x \in U$, the distance from x_n to U tends to zero as $n \in \mathbb{N}$ tends to infinity.

Attractors can be fixed points, periodic orbits, limit cycles, or more complex structures like strange attractors. They are invariant and have a *basin of attraction* that draws nearby trajectories, making them key to understanding the system's long-term behavior.

An *ω -limit set* (often called an *ω -set*) is a set of accumulation points that a trajectory approaches as time goes to infinity. It also provides important information about the long-term behavior of the system. Formally, for $\mathcal{S} = (S, \phi)$ a dynamical system, the *ω -limit set* of a point $x \in S$, denoted $\omega(x)$, is defined as the set of all points $y \in S$, such that there exists an (strictly) increasing sequence $\{n_k\}_{k \in \mathbb{N}}$ satisfying $\lim_{k \rightarrow \infty} \phi_{n_k}(x) = y$.

Example 3. Consider the dynamical system $\mathcal{S} = ([0..1], \phi)$, with $\phi = \{f^n\}_{n \in \mathbb{N}}$ and $f : \mathbb{R} \rightarrow \mathbb{R}$ the *logistic map* defined by $f(x) = \mu x(1 - x)$.

Figure 3 depicts the orbit of x_0 in \mathcal{S} when setting $\mu = 3.4$ and $x_0 = 0.3$. It converges to the set $U = \{0.4519, 0.8421\}$ (the red dots). The even indexes converge to 0.4519 and the odd indexes converge to 0.8421. That is, $\omega(x_0) = U$.

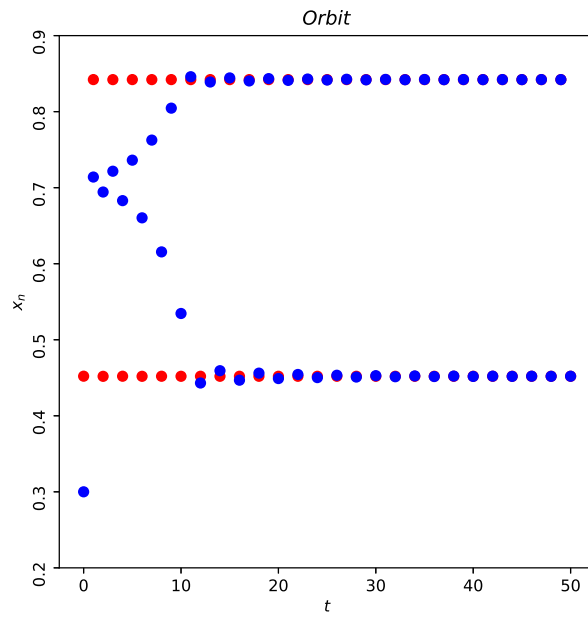


Fig. 3: Orbit of point x_0 under map L , with $\mu = 3.4$ in blue. Set $\{0.4519, 0.8421\}$ in red.

B Proofs

In this appendix, we present the proofs of our theoretical results for consensus.

B.1 Lemma 1

To prove this result, we have to prove that for all $\mathbf{x} \in [0, 1]^N$, the entries $a_{ij}(\mathbf{x})$ of matrix $\mathbf{A}(\mathbf{x})$, satisfy

- $a_{ij}(\mathbf{x}) \geq 0$, for all $i, j = 1, \dots, N$.
- $\sum_{j=1}^N a_{ij}(\mathbf{x}) = 1$, for all $i = 1, \dots, N$.

We begin the proof with the definition of these entries, given in equation (5)

$$a_{i,j}(\mathbf{x}) = \begin{cases} 1 - \sum_{k \in A_i} \bar{I}_{k,i} \alpha_{i,k}(\mathbf{x}) & , \text{ if } i = j \\ \bar{I}_{j,i} \alpha_{i,j}(\mathbf{x}) & , \text{ if } i \neq j. \end{cases} \quad (16)$$

Note that $a_{i,i} = 1 - (a_{i,1} + \dots + a_{i,i-1} + a_{i,i+1} + \dots + a_{i,N})$, then, it is easy to see that $a_{i,1} + \dots + a_{i,i-1} + a_{i,i} + a_{i,i+1} + \dots + a_{i,N} = 1$.

Now, note that $a_{i,j} \geq 0$, $i \neq j$, since $\bar{I}_{j,i} \geq 0$ and $\alpha_{i,j}(\mathbf{x}) \geq 0$, for all $i, j = 1, \dots, N$.

Recall that $\sum_{j \in A_i} \bar{I}_{j,i} = 1$ and $\alpha_{i,j}(\mathbf{x}) \leq 1$, furthermore, for each $i \in A$, there exist a neighbor $j \in A_i$ such that $\alpha_{i,j}(\mathbf{x}) < 1$, then

$$\sum_{k \in A_i} I_{k,i} \alpha_{i,k}(\mathbf{x}) < \sum_{k \in A_i} I_{k,i} = 1,$$

with this, we can check that $a_{i,i} = 1 - \sum_{k \in A_i} I_{k,i} \alpha_{i,k}(\mathbf{x}) > 0$.

Hence $\mathbf{A}(\mathbf{x})$ is stochastic, for all $\mathbf{x} \in [0, 1]^N$.

B.2 Lemma 2

To prove that matrix $\mathbf{A}(\mathbf{x})$ is primitive, we have to check several conditions

- $\mathbf{A}(\mathbf{x})$ is irreducible for all $\mathbf{x} \in [0, 1]^N$.
- There exists $m > 0$, such that $\mathbf{A}^m(\mathbf{x}) > 0$, for all $\mathbf{x} \in [0, 1]^N$.

By Theorem 5, prove that for all $\mathbf{x} \in [0, 1]^N$, $\mathbf{A}^m(\mathbf{x})$ is irreducible, is equivalent to prove that the graph $G_{\mathbf{A}(\mathbf{x})}$ associated at $\mathbf{A}(\mathbf{x})$, is strongly connected.

By hypothesis, the graph G , is strongly connected, then for each pair of vertex i, j , there exist a path of edges $(i, v_2) \dots (v_k, j)$ which connects these vertex.

It is easy to see that $I_{j,i} > 0$, implies $a_{i,j} > 0$, for $i \neq j$. Then, for each edge (j, i) in E , there exists an edge between vertex (i, j) , in $G_{\mathbf{A}(\mathbf{x})}$. From this, we can conclude that for each pair of vertex (i, j) in $G_{\mathbf{A}(\mathbf{x})}$, exists a path of edges which connects these vertex, hence $G_{\mathbf{A}(\mathbf{x})}$ is strongly connected.

Above argument show that $\mathbf{A}(\mathbf{x})$ is irreducible for all $\mathbf{x} \in [0, 1]^N$.

To prove the second condition, we shall use the Frobenius test of primitivity with $m = |A|$. If we denote $a_{i,j}^{(m)}$ like the entry i, j of matrix \mathbf{A}^m , then we must prove that $a_{i,j}^{(m)} > 0$, for all $i, j = 1, \dots, N$.

We shall prove this with the next two claims

Claim 1: If exist a path of length l between i and j , then $a_{i,j}^{(l)} > 0$.

Proof: We prove this claim by induction over l .

The base case is $l = 1$, what is trivial, since the existence of an edge (i, j) implies $a_{i,j} > 0$.

Inductively, assume that if exist a path of length l between 2 nodes i, k , then $a_{i,k}^{(l)} > 0$.

Consider a pair of vertex i, j and a path of edges $(i, v_1)(v_1, v_2) \dots (v_{l-1}, v_l)(v_l, j)$ of length $l + 1$. For inductively hypothesis $a_{i,v_l}^{(l)} > 0$, then

$$a_{i,j}^{(l+1)} = \sum_{k=1}^N a_{i,k}^{(l)} a_{k,j} \geq a_{i,v_l}^{(l)} a_{v_l,j} > 0.$$

It is proved.

Claim 2: If $a_{i,j}^{(n)} > 0$, then $a_{i,j}^{(n+1)} > 0$, for all $n \geq 1$.

Proof: Recall that for all $j = 1, \dots, N$, $a_{j,j} > 0$, then, if we have $a_{i,j}^{(n)}$, for some $n \geq 1$, we have

$$a_{i,j}^{(n+1)} = \sum_{k=1}^N a_{i,k}^{(n)} a_{k,j} \geq a_{i,j}^{(n)} a_{j,j} > 0.$$

It is proved.

Since for all $\mathbf{x} \in [0, 1]^N$, $G_{\mathbf{A}(\mathbf{x})}$ is strongly connected, then for each pair of vertex (i, j) exists a path of length l , where $l \leq |A|$. By Claim 1, $a_{i,j}^{(l)} > 0$, and by Claim 2, $a_{i,j}^{(|A|)} > 0$. Then $\mathbf{A}^{(|A|)}(\mathbf{x}) > 0$, and by Frobenius test of primitivity, $\mathbf{A}(\mathbf{x})$ is primitive.

B.3 Theorem 2

The proof of this result requires several concepts from dynamical systems, as well as the lemmas stated above. Recall that an update function converges to consensus if there exists a $v \in [0, 1]$ such that, for every agent $i \in A$, their opinion \mathbf{x}_i^t satisfies $\lim_t \mathbf{x}_i^t = v$.

We equivalently prove that the sequence $\{\eta_k\} \subset \mathbb{R}$, defined by $\eta_k = \|\mathbf{x}^k - \bar{\mathbf{x}}^k\|$, where $\mathbf{x}^{k+1} = \mu_G(\mathbf{x}^k)$, converges at $\eta_k \rightarrow 0$.

Proof is composed of the next way

- Prove that the sequence $\eta_k = \|\mathbf{x}^k - \bar{\mathbf{x}}^k\|$ converges at one value $L \geq 0$ when $k \rightarrow \infty$.
- Prove that $L = 0$.

Consider the line $l \subset \mathbb{R}^N$, defined by $l = \{r\mathbf{1} : r \in \mathbb{R}\}$, where $\mathbf{1}$ is a vector whose all entries are equal to 1.

Claim 1: The point $\mathbf{y} = \bar{\mathbf{x}} \in l$, is the point who minimize the distance $\|\mathbf{x} - \mathbf{y}\|$.

Proof: Proof of above claim is a geometric property, since $\bar{\mathbf{x}}$ is the projection of \mathbf{x} over the line l .

Claim 2: The sequence $\eta_k = \left\| \mathbf{x}^k - \overline{\mathbf{x}^k} \right\|$, where $\mathbf{x}^{k+1} = \mu_G(\mathbf{x}^k)$, converges at converges at one value $L \geq 0$.

Proof: To prove this claim, we have to prove that the sequence η_k is monotonous and bounded.

First, we use the Claim 1 and that vectors in l are fixed points of map F

$$\begin{aligned} \left\| \mathbf{x}^{k+1} - \overline{\mathbf{x}^{k+1}} \right\| &< \left\| \mathbf{x}^{k+1} - \overline{\mathbf{x}^k} \right\|, \\ \left\| \mathbf{x}^{k+1} - \overline{\mathbf{x}^{k+1}} \right\| &< \left\| \mathbf{A}(\mathbf{x}^k)\mathbf{x}^k - \mathbf{A}(\mathbf{x}^k)\overline{\mathbf{x}^k} \right\|, \\ \left\| \mathbf{x}^{k+1} - \overline{\mathbf{x}^{k+1}} \right\| &< \left\| \mathbf{A}(\mathbf{x}^k) \right\| \left\| \mathbf{x}^k - \overline{\mathbf{x}^k} \right\|, \end{aligned}$$

Vector $\mathbf{x}^k - \overline{\mathbf{x}^k}$ is in the space generated by the eigenvectors $\mathbf{v}_2, \dots, \mathbf{v}_s$, $s \geq N$. This space is invariant for the operator \mathbf{A} , then the norm of this operator is $|\lambda_2| < 1$ (because \mathbf{A} is primitive), then

$$\begin{aligned} \left\| \mathbf{x}^{k+1} - \overline{\mathbf{x}^{k+1}} \right\| &< |\lambda_2| \left\| \mathbf{x}^k - \overline{\mathbf{x}^k} \right\|, \\ \left\| \mathbf{x}^{k+1} - \overline{\mathbf{x}^{k+1}} \right\| &< \left\| \mathbf{x}^k - \overline{\mathbf{x}^k} \right\|, \end{aligned}$$

With this, we have shown that the sequence is decreasing monotone.

By properties of the norm operator $\|\cdot\|$, this is non negative, for this reason and the monotonic property, the sequence is bounded. For each $k \geq 1$, we have $0 \leq \eta_k \leq \eta_0$. By the real analysis, we know that this sequence converges at some value L , where $0 \leq L \leq \eta_0$.

Claim 3: The limit L of the sequence η_k is $L = 0$.

The first item to prove this claim, is that the update function μ_G defined in (4) is continuous. We know this because this map is composed by continuous functions.

For properties of continuous dynamical systems presented in Appendix A.2, we have that for each initial condition \mathbf{x}^0 , the set $\omega_F(\mathbf{x}^0)$ is an invariant set; which means that for all $\mathbf{y} \in \omega_F(\mathbf{x}^0)$, $F(\mathbf{y}) \in \omega_F(\mathbf{x}^0)$.

Now, let us proceed by contradiction. Suppose that $L > 0$, then $\left\| \mathbf{x}^k - \overline{\mathbf{x}^k} \right\| \rightarrow L > 0$, when $k \rightarrow \infty$. This implies that $\omega_\mu(\mathbf{x}^0)$ is a subset of the sphere of radius L centered in a point $\mathbf{p} \in \mathbb{R}^N$ of the line l , denoted by S_L .

Let $\mathbf{y} \in \omega_\mu(\mathbf{x}^0)$, by our above reasoning, we have that $\left\| \mathbf{y} - \overline{\mathbf{y}} \right\| > \left\| \mu_G(\mathbf{y}) - \overline{\mu_G(\mathbf{y})} \right\|$, but this contradicts the invariance of $\omega_\mu(\mathbf{x}^0)$, because the distance of $F(\mathbf{y})$ is less than L .

In conclusion, L must be equal to zero.