Automatica 166 (2024) 111711

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Balancing homophily and prejudices in opinion dynamics: An extended Friedkin–Johnsen model ☆

Giorgia Disarò, Maria Elena Valcher *

Dipartimento di Ingegneria dell'Informazione Università di Padova, via Gradenigo 6B, 35131 Padova, Italy

ARTICLE INFO

Article history:
Received 6 April 2023
Received in revised form 15 December 2023
Accepted 3 April 2024
Available online 17 May 2024

Keywords:
Opinion dynamics
Social networks
Multi-agent systems
Signed graphs
Homophily
Prejudices

ABSTRACT

In this paper we propose an extended version of the Friedkin–Johnsen (FJ) model where we analyze the effects that a homophily-based influence matrix has on the opinion formation process. In particular, we assume that the influence matrix varies over time, and its entries (that represent the appraisals that individuals have of the other individuals) update based on the comparison between their opinions and can take both positive and negative values. This leads to a system of two difference equations, where the first one corresponds to the standard FJ model, except that the influence matrix is no longer constant, and the second one models the dynamics of this matrix. We show that a necessary and sufficient condition for the convergence of this modified version of the classical FJ model is that the influence matrix becomes constant in a finite number of steps. Moreover, we provide the explicit expression for the agents' asymptotic opinions in some special cases: the purely cooperative setting, the case of a structurally balanced network and the case of a single discussion topic. Finally, the case when the topics are correlated and the influence matrix depends on the average opinion vectors is investigated.

1. Introduction

In the last decades there has been a growing interest in modeling and describing the opinion formation processes. This is motivated by the fact that individuals' opinions drive their actions, and hence understanding how opinions evolve represents a crucial step in order to formulate predictions and control strategies within social networks. Opinion dynamics is a research topic that attracts many researchers from various fields, ranging from sociology and psychology to mathematics and control engineering. Indeed, the study of opinion dynamics requires both knowledge of the underlying sociological mechanisms and mathematical skills to formalize them rigorously. The challenge is to find a model that represents a reasonable compromise between the following conflicting needs: explain the effective complexity of social phenomena and be tractable from a mathematical point of view. The various models that have been proposed have provided many insights into the opinion formation processes, capturing important effects of mutual interactions (Proskurnikov & Tempo, 2017, 2018). A large number of opinion formation models is based on a linear weighted-averaging process, as originally introduced by DeGroot in DeGroot (1974). However, the weighted-averaging mechanism implicitly assumes that the attractiveness between two opinions is proportional to their mutual distance and thus, under mild assumptions, generates consensus, which in many practical situations is not a realistic outcome. To overcome this limitation, more recently the interest has shifted towards models that justify other observed behaviors such as disagreement, polarization and conflict (Altafini, 2012; Bizyaeva et al., 2023; Franci et al., 2023; Friedkin & Johnsen, 1999; Hegselmann & Krause, 2002; Hiller, 2017; Leonard et al., 2021; Mei et al., 2022; Santos et al., 2021). More specifically, in Hegselmann and Krause (2002) a bounded confidence mechanism is introduced, assuming that opinions attractiveness first increases proportionally with opinions distance but then vanishes, once the mutual distance crosses a certain threshold. Instead, in Altafini (2012) and Hiller (2017) the opinion formation occurs in a signed network to take into account that the relationships among the agents can also be antagonistic and thus repulsive, leading to polarization or conflict. Other solutions that have been proposed to overcome this unrealistic feature of the weighted-averaging mechanism include Bizyaeva et al. (2023), where a novel nonlinear extension of the classical averaging-based process that leverages the saturation function has been presented. In Mei et al. (2022), the authors solve the problem by substituting the

☆ A subset of the paper results is included in a paper presented at the 62nd IEEE Conference on Decision and Control, December 13–15, 2023, Marina Bay Sands, Singapore. This paper was recommended for publication in revised form by Associate Editor Julien M. Hendrickx under the direction of Editor Christos G. Cassandras.

* Corresponding author.
E-mail addresses: giorgia.disaro@phd.unipd.it (G. Disarò), meme@dei.unipd.it (M.E. Valcher).

https://doi.org/10.1016/j.automatica.2024.111711

weighted-averaging mechanism with a weighted-median mechanism, in which opinions distance and attractiveness are not intrinsically coupled.

In this regard, another simple and widely used model is the Friedkin–Johnsen (FJ) model (Friedkin & Johnsen, 1999), which introduces in the standard DeGroot model for consensus (DeGroot, 1974) a term that accounts for the individuals' stubbornness, namely their attitude to remain attached to their initial opinions. This leads to a more involved dynamics where at each time step the opinions of individuals (in the following also referred to as "agents") update on the basis of two main driving forces: the opinions of their neighbors and their own initial opinions (prejudices). In this setting the reaching of a consensus is a very rare occurrence, if not unfeasible: the final opinions are closer to each other, but not identical.

The original FJ model focuses on a single discussion topic and so individual opinions are scalars. Moreover, the influence matrix of the network is constant and row-stochastic. Later on, several extensions of the model have been proposed where multiple topics have been considered (Friedkin, 2015; Parsegov et al., 2015, 2017) or where the influence matrix is row-stochastic but time-varying (Proskurnikov et al., 2017). Recently, a version of the model that accounts also for the fact that the relationships among the agents do not need to be cooperative, and thus positive (see Altafini (2013), Pan et al. (2021) and Xia et al. (2016)), has been presented (He et al., 2022). In this modified version, the influence matrix is no longer a nonnegative matrix, but remains independent of the individuals' opinions, as in all the models cited above. This assumption does not seem to be realistic since in real life very often two individuals like or dislike each other on the basis of their opinions similarity. An extensive experimental literature (McPherson et al., 2001; Rivera et al., 2010) in social psychology establishes that individuals who share similar values, attitudes and beliefs tend to associate and interact more intensively, following what in Lazarsfeld and Merton (1954) is called *value homophily*. The effect of homophily in opinion dynamics has already been analyzed in existing models, such as Liu et al. (2020) and Mei et al. (2019). In Mei et al. (2019), the homophily mechanism employed is a person-to-person homophily that finds expression in the updating of the interpersonal evaluations of any two individuals based on how much they agree with each other on the evaluations of all agents in the group. Instead, in Liu et al. (2020), the authors consider a person-to-entity homophily (the aforementioned value homophily), i.e., two individuals like or dislike each other depending on whether agreement or disagreement on the topics of discussion prevails. More specifically, the authors proposed an interesting model of the interplay between homophily-based appraisal dynamics and influence-based opinion dynamics. The model analyzes for the first time how the agents' mutual appraisals impact the evolution of the opinions of the same agents on a certain number of issues/topics. More recently, a simplified version of the model, that accounts only for the signs of the mutual appraisals rather than their values, has been proposed in De Pasquale and Valcher (2022). It has been shown that this simpler model is still able to predict the asymptotic behavior of the individuals' opinions in small networks, as the ones we will consider in this paper.

In this contribution, we propose a modified version of the classical FJ model, which maintains all the fundamental characteristics of an FJ-type dynamics, namely influence of the initial opinions, determinism, continuance, decomposability and simultaneity (see Friedkin and Johnsen (1999)). The only relaxation concerns the fixed social structure: indeed, in our model the set of actors, their stubbornness coefficients and their initial opinions are fixed, but the influence network is time-varying (Proskurnikov et al., 2017) and signed (He et al., 2022). However, differently from Proskurnikov et al. (2017), we further assume that at each time step, the influence matrix is not arbitrary, but instead it is constructed based on the agents' mutual appraisals and specifically on their signs. The agents' mutual appraisals are in turn based on the comparison of the individuals' opinions, according to a value homophily mechanism, as in Liu et al. (2020).

In the general case, we have been able to prove that the opinion matrix of a group of n agents on m topics asymptotically converges to a constant solution, that strongly depends on the agents' initial opinions as well as on the agents' stubbornness, namely their attitude to remain attached to their original opinions, if and only if the influence matrix becomes constant in finite time. Furthermore, when the initial opinions of the agents lead to an initial nonnegative influence matrix, the relationships between pairs of individuals remain cooperative at all the subsequent times, creating a collaborative set-up that involves all the agents or disjoint groups of agents that are internally cooperative, depending on whether the initial graph is connected or not, respectively. In this situation we are able to express the asymptotic opinions of the agents in closed form. Then, we consider the special case of a single discussion topic. Finally, we introduce a slightly different dynamics for the influence matrix update that can be used to simplify the model when the topics are interrelated. In both these cases we provide an explicit expression of the agents' asymptotic opinions.

The paper is organized as follows: Section 2 introduces the model explaining the meaning of all the quantities involved. Section 3 provides the main results about the dynamics of the model. Section 4 shows that corresponding to certain initial opinion matrices the opinion dynamics of the whole group of individuals splits into disjoint subgroups that do not interact with each other. Section 5 analyzes the particular case of a cooperative social network. Section 6 addresses the single-topic case. In Section 7 we consider the case of interrelated topics. In Section 8 two numerical examples are proposed. Finally, Section 9 concludes the paper.

To ensure a smoother flow of arguments in the paper, all proofs have been moved to the Appendix.

The paper builds upon the conference paper (Disarò & Valcher, 2023), with which it shares Section 2, Section 3 and Section 6. Sections 4, 5 and 7 are original. Example 1 can be found also in Disarò and Valcher (2023), while the other example is new.

Notation. Given two integers k and n, with $k \leq n$, the symbol $[k, n]$ denotes the set $\{k, k+1, \dots, n\}$. We let $\mathbb{1}_k$ ($\mathbb{0}_k$, resp.) denote the k-dimensional vector with all unitary (zero, resp.) entries. The symbol $\mathbb{0}_{k\times l}$ represents the matrix of dimension $k \times l$ whose entries are all zero. We denote by $\mathbb{e}_i$ the ith canonical vector of dimension n, where n is always clear from the context. In the sequel, the (i, j)th entry of a matrix A is denoted by $[A]_{ij}$, while the ith entry of a vector v by v_i. The identity matrix of size n is denoted by I_n. When the size is clear from the context, sometimes we omit the subscript n. The function sgn: $\mathbb{R}^{n\times m} \to \{-1, 0, 1\}^{n\times m}$ is the function that maps a real matrix A into a matrix taking values in $\{-1, 0, 1\}$, in accordance with the sign of its entries, namely $[\mathrm{sgn}(A)]_{ij} = \mathrm{sgn}([A]_{ij})$ for every i, j. The expression $X = \mathrm{blockdiag}\{X_1, \dots, X_k\}$ denotes the block diagonal matrix whose diagonal blocks are $X_1, \dots, X_k$. If $X_1, \dots, X_k$ are scalars, we use the notation $X = \mathrm{diag}\{X_1, \dots, X_k\}$. A *signature matrix* is a diagonal matrix whose diagonal entries belong to $\{-1, 1\}$. The *max norm of a matrix* $A \in \mathbb{R}^{n\times n}$ is defined as $\|A\|_{max} := \max_{i,j\in[1,n]} |[A]_{ij}|$. The *max norm of a vector* $v \in \mathbb{R}^n$ is $\|v\|_{max} := \max_{i\in[1,n]} |v_i|$.

The *spectrum* of a matrix A, denoted as $\sigma(A)$, is the set of all its eigenvalues, and the *spectral radius*, $\rho(A)$, is defined as $\rho(A) := \max\{|\lambda| : \lambda \in \sigma(A)\}$.

Given a matrix A, the symbol $A \geq 0$ means that all the entries of the matrix are nonnegative, while the symbol $A > 0$ means that all the entries of the matrix are positive. A matrix $A \in \mathbb{R}^{n\times n}$ is said to be *irreducible* if no permutation matrix $P \in \mathbb{R}^{n\times n}$ can be found such that (s.t.) $P^\top AP$ is block triangular. If such a permutation matrix exists, the matrix is said to be *reducible*. A nonnegative matrix $A \in \mathbb{R}^{n\times n}$ is said to be *primitive* if $\exists\, k > 0$ such that $A^k > 0$. A primitive matrix is always irreducible, but the converse is not true.

In this paper by an *undirected and signed graph* we mean a triple $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{A})$, where $\mathcal{V} = [1, n]$ is the set of nodes (or vertices), $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ is the set of edges (or arcs) and $\mathcal{A} \in \{-1, 0, 1\}^{n\times n}$ is the *adjacency matrix* of the graph $\mathcal{G}$. An arc $(j, i) \in \mathcal{E}$ if and only if $[\mathcal{A}]_{ij} \neq 0$. When so, $[\mathcal{A}]_{ij}$ represents the (positive or negative) weight of the arc. Moreover, due to the fact that the graph is undirected, the matrix $\mathcal{A}$ is symmetric and so $(i, j) \in \mathcal{E}$ if and only if $(j, i) \in \mathcal{E}$. Since the adjacency matrix $\mathcal{A}$ uniquely identifies the graph, in the following we will use the notation $\mathcal{G}(\mathcal{A})$. A graph $\mathcal{G}$ is said to be *structurally balanced* (Altafini, 2013; Xia et al., 2016) if the set of its nodes can be partitioned into two disjoint subsets such that the weights of the edges between nodes belonging to the same subset are nonnegative, and the weights of the edges between nodes belonging to different subsets are nonpositive. If $\mathcal{G} = \mathcal{G}(\mathcal{A})$ for some matrix $\mathcal{A} \in \mathbb{R}^{n\times n}$, then $\mathcal{G}$ is structurally balanced if and only if (Altafini, 2013) there exists a signature matrix $D = \mathrm{diag}\{d_1, \dots, d_n\}, d_i \in \{-1, 1\},\ i \in [1, n]$, s.t. $D\mathcal{A}D \geq 0$.

2. The model

We consider a network with n agents expressing their opinions about m distinct topics. We denote by $Y(t) \in \mathbb{R}^{n\times m}$ the *opinion matrix at time t*, whose (i, j)th entry represents the opinion that agent i has about topic j at time $t \in \mathbb{Z}_+$ and by $W(t) \in \mathbb{R}^{n\times n}$ the *influence matrix at time t*, whose (i, j)th entry represents the weight that agent i gives to the opinion of agent j at time t. Specifically, we assume that:

- $[W(t)]_{ij} > 0 \Leftrightarrow i$ positively weights the opinion of j;
- $[W(t)]_{ij} < 0 \Leftrightarrow i$ negatively weights the opinion of j;
- $[W(t)]_{ij} = 0 \Leftrightarrow i$ neglects the opinion of j.

At every time t, the influence that agent j has on agent i coincides with agent i's appraisal of agent j. On the other hand, the appraisal that i has of j at time t depends on the comparison of the opinions that agents i and j have about all the topics at time $t-1$, according to a value homophily mechanism (Dandekar et al., 2013; Lazarsfeld & Merton, 1954; McPherson et al., 2001). As in De Pasquale and Valcher (2022), we consider only the signs of the mutual appraisals, rather than their values. This is motivated by the fact that from a practical viewpoint it is complicated to quantify the appraisal each individual has of the others, but, on the contrary, it is easy to recognize if the relationship between two agents is friendly or hostile. In addition, this choice turns out to be more robust to modeling errors and more realistic, because, even if we are able to establish that agent j influences positively or negatively agent i's opinion, this influence does not necessarily scale with the absolute value of their mutual appraisal.

Based on these premises, in this paper we propose the following model, representing the intertwining between an FJ-type opinion dynamics and a homophily-based appraisal mechanism:

$$Y(t+1) = (I_n - \Theta)W(t+1)Y(t) + \Theta Y(0), \tag{1}$$

$$W(t+1) = \frac{1}{n}\mathrm{sgn}\left(Y(t)Y(t)^\top\right), \tag{2}$$

where $\Theta := \mathrm{diag}\{\theta_1, \dots, \theta_n\} \in \mathbb{R}^{n\times n}$ is a diagonal matrix. For every $i \in [1, n]$, θ_i represents the stubbornness of agent i in preserving the original opinion.

Eq. (1) is a standard FJ-model, whose influence matrix, however, is time-varying and evolves through a feedback mechanism based on the opinion matrix itself. In fact, Eq. (2) describes the fact that the weight that agent i gives to agent j is the result of a comparison between the (real-valued) opinion vectors of agents i and j. Since we are considering small networks, it is reasonable to assume that all agents know each other or, in other words, that the underlying social network is all-to-all connected. Indeed, the update of the influence matrix in Eq. (2) is the combination of two mechanisms: one that gives the opinions of all other agents in the network the same relative weight equal to $\frac{1}{n}$, and another one that attributes a positive, negative or neutral appraisal to each agent on the basis of the opinions similarity. Specifically, the appraisal is positive (negative) if the "total amount of agreement" between the opinions of the two agents about the various topics is greater than (smaller than) the "total amount of disagreement". By combining these two mechanisms, we obtain the actual (signed) weight that each individual gives to the opinions of every other individual. Note that a null entry in the matrix $W(t)$, say $[W(t)]_{ij} = [W(t)]_{ji} = 0$, corresponds to the situation in which two agents i and j know each other, but do not find any correlation between their opinions and therefore they both decide to neglect the opinion of the other. This is an information that should be considered and motivates the choice of dividing each row of $\mathrm{sgn}\left(Y(t)Y(t)^\top\right)$ by n, instead of by the number of its non-zero entries. However, it is worth noticing that condition $[W(t)]_{ij} = 0$ is a very rare occurrence since it corresponds to the case where the opinion vectors of agent i and j at time $t-1$ (i.e., the ith and jth rows of $Y(t-1)$) are orthogonal. In other words the total amount of agreement and disagreement between the opinions of the two agents about the various topics coincide.

In the paper we will steadily assume:

Assumption 1. For every $i \in [1, n]$ the stubbornness of agent i satisfies $0 < \theta_i < 1$.

It is easy to see that if the ith row of $Y(0)$ is zero, then the ith row of $Y(t)$ is zero for every $t \geq 0$. So, in the following we will rule out this case, which is of no interest.

Assumption 2. The matrix $Y(0) \in \mathbb{R}^{n\times m}$ is devoid of zero rows.

Finally, it is worth noticing that the influence matrix $W(t+1)$ is a symmetric matrix for every $t \geq 0$.

3. General results

In order to investigate the asymptotic behavior of the opinion matrix, we first provide an alternative way to express the opinion matrix at time t, by introducing the transition matrix $M(t)$, relating $Y(t)$ to $Y(0)$. In the following we will steadily resort to the following notation:

$$S_0 := Y(0)Y(0)^\top. \tag{3}$$

Proposition 1. *For every $Y(0) \in \mathbb{R}^{n\times m}$, at every time $t \geq 0$ we have*

$$Y(t+1) = M(t+1)Y(0), \tag{4}$$

where

$$M(t+1) = (I_n - \Theta)W(t+1)M(t) + \Theta, \tag{5}$$

$$M(0) = I_n, \tag{6}$$

$$W(t+1) = \frac{1}{n}\mathrm{sgn}(M(t)S_0M(t)^\top). \tag{7}$$

Based on Proposition 1, we now derive the main result regarding the asymptotic behavior of the sequence $\{M(t)\}_{t\in\mathbb{Z}_+}$.

Theorem 2. *For every $Y(0) \in \mathbb{R}^{n\times m}$, the solution of the system in (5)–(6)–(7) is bounded, and specifically*[1] $\|M(t)\|_{max} \leq 1$ *for all* $t \in \mathbb{Z}_+$. *Moreover, the following conditions are equivalent:*

(i) There exists $M_\infty := \lim_{t\to+\infty} M(t)$;
(ii) There exists $T \in \mathbb{Z}_+, T \geq 1$, *such that* $W(t) = W(T) =: W_\infty$, *for every* $t \geq T$.

If either of the above equivalent conditions holds, then

$$M_\infty = (I_n - \Theta)W_\infty M_\infty + \Theta. \tag{8}$$

The main consequence of Theorem 2 is that, for every $Y(0) \in \mathbb{R}^{n\times m}$ for which $\exists M_\infty = \lim_{t\to+\infty} M(t)$, we can also ensure that $\exists Y_\infty := \lim_{t\to+\infty} Y(t)$ and $Y_\infty = M_\infty Y(0)$. Moreover, starting from some $T \in \mathbb{Z}_+, T \geq 1$, we have $W_\infty = W(T) \in \{-1/n, 0, 1/n\}^{n\times n}$, and hence

$$W_\infty = \frac{1}{n}\mathrm{sgn}\left(M(T-1)S_0 M(T-1)^\top\right). \tag{9}$$

We now explore some interesting properties of M_∞.

Proposition 3. *Given* $Y(0) \in \mathbb{R}^{n\times m}$, *if there exists* $M_\infty = \lim_{t\to+\infty} M(t)$, *then* M_∞ *is nonsingular and* $\forall i \in [1, n]$

(i) $\|M_\infty \mathbb{e}_i\|_{max} = \max_{j\in[1,n]} |[M_\infty]_{ji}| = |[M_\infty]_{ii}|$;
(ii) $[M_\infty]_{ii} > 0$.

The previous result means that, in case of convergence to a constant solution, the initial opinion of each agent impacts more on his/her own final opinion than on the final opinions of the other agents. In other words, the agent that weights more agent i's initial opinion is agent i himself/herself. Moreover, (and not unexpectedly!) such impact is always positive.

4. Model analysis can be carried on for an irreducible S_0

As $S_0 = Y(0)Y(0)^\top$ is symmetric, when it is reducible its normal form Gantmacher (1960) is a block diagonal matrix, i.e., there exists a permutation matrix $P \in \mathbb{R}^{n\times n}$ such that

$$P^\top S_0 P = \mathrm{blockdiag}\{S_{0,1}, \dots, S_{0,k}\} \tag{10}$$

for some $k \in \mathbb{Z}$, $k \geq 2$, and irreducible diagonal blocks $S_{0,i} \in \mathbb{R}^{n_i\times n_i}$, where the positive integers n_i, $i \in [1, k]$, satisfy $n_1 + \cdots + n_k = n$.

We want to show that in this case all matrices $Y(t+1)Y(t+1)^\top$ are block diagonal and the dynamics of each block is decoupled from the dynamics of the other blocks.

Proposition 4. *If the matrix S_0 is reducible, with block diagonal structure as in (10) (for $P = I_n$), then $\forall t \geq 0$ the matrices $W(t+1)$ and $M(t+1)$ have the same block diagonal structure as S_0, i.e.,*

$$W(t+1) = \mathrm{blockdiag}\{W_1(t+1), \dots, W_k(t+1)\}, \tag{11}$$
$$M(t+1) = \mathrm{blockdiag}\{M_1(t+1), \dots, M_k(t+1)\}, \tag{12}$$

where $W_i(t+1)$ and $M_i(t+1)$ belong to $\mathbb{R}^{n_i\times n_i}$.

As a consequence of the previous proposition, when the initial condition $Y(0)$ is such that (10) holds for some permutation matrix P, we can accordingly partition the matrix $Y(t)$ in k blocks as

$$P^\top Y(t) = \begin{bmatrix} Y_1(t)^\top & \dots & Y_k(t)^\top \end{bmatrix}^\top,$$

where $Y_i(t) \in \mathbb{R}^{n_i\times m}$, and express the diagonal matrix Θ as $P^\top \Theta P = \mathrm{blockdiag}\{\Theta_1, \dots, \Theta_k\}$ with $\Theta_i \in \mathbb{R}^{n_i\times n_i}$. Since the dynamics of each block is independent of the dynamics of the remaining ones, we can study each of them separately (but when evaluating $W_i(t+1)$ we still have to account for the size of the original network, n). This also implies that we can always confine ourselves to the case in which S_0 is irreducible.

[1] Note that $W(t+1), t \in \mathbb{Z}_+$, is always bounded, since it takes values in $\{-1/n, 0, 1/n\}$.

5. Collaborative set-up

We now consider the case where the initial condition $Y(0) \in \mathbb{R}^{n\times m}$ is such that $S_0 = Y(0)Y(0)^\top \geq 0$. This amounts to assuming that, at the initial time, for each pair of agents in the group the "agreement exceeds the disagreement" and hence each agent gives a positive weight to the opinions of the others. When so, we can provide a closed form solution to the opinion dynamics model. Indeed, we can always prove that the influence matrix W becomes constant in a finite number of steps, and hence there exists $M_\infty = \lim_{t\to+\infty} M(t)$

Proposition 5. *If $S_0 \geq 0$, then, for every $t \geq 1$*

$$\mathrm{sgn}(M(t)) = \mathrm{sgn}(S_0^{\alpha(t)}) = \mathrm{sgn}(M(t)^\top), \tag{13}$$
$$\mathrm{sgn}(W(t)) = \mathrm{sgn}(S_0^{\beta(t)}), \tag{14}$$

where

$$\begin{cases} \alpha(t+1) = 3\alpha(t) + 1 \\ \alpha(1) = 1 \end{cases} \tag{15}$$

$$\begin{cases} \beta(t+1) = 2\alpha(t) + 1 \\ \beta(1) = 1. \end{cases} \tag{16}$$

If $Y(t)Y(t)^\top \geq 0$ for $t = 0$, this situation persists also at the subsequent times. Consequently, the sequence of influence matrices $W(t)$ consists in turn of nonnegative matrices (the sign function allows to distinguish the zero entries of the influence matrices), and, depending on the graph of S_0, its limit value represents either a situation of complete cooperation or the case where the overall group splits into disjoint cooperative groups.

Theorem 6. *Suppose that $S_0 \geq 0$. Then, $Y(t)Y(t)^\top \geq 0$ for every $t \geq 0$, there exists $T \geq 1$ such that $W(t) = W(T) = W_\infty, \forall t \geq T$, and two cases may occur:*

(i) if S_0 is an irreducible matrix, then $W_\infty = \frac{1}{n}\mathbb{1}_n\mathbb{1}_n^\top$;
(ii) if S_0 is reducible, and hence (by the symmetry of S_0) there exists a permutation matrix $P \in \mathbb{R}^{n\times n}$ s.t.

$$P^\top S_0 P = \mathrm{blockdiag}\{S_{0,1}, \dots, S_{0,k}\}$$

where $S_{0,i} \in \mathbb{R}^{n_i\times n_i}$ is a nonnegative irreducible block, then

$$P^\top W_\infty P = \frac{1}{n}\mathrm{blockdiag}\{\mathbb{1}_{n_1}\mathbb{1}_{n_1}^\top, \dots, \mathbb{1}_{n_k}\mathbb{1}_{n_k}^\top\}.$$

Remark 7. If S_0 is an irreducible nonnegative matrix then $W_\infty = \frac{1}{n}\mathbb{1}_n\mathbb{1}_n^\top$ and hence, by making use of (8) and (9), and of elementary algebraic manipulations, we obtain

$$\begin{aligned} M_\infty &= \left[I_n - (I_n - \Theta)\frac{1}{n}\mathbb{1}_n\mathbb{1}_n^\top\right]^{-1}\Theta \\ &= \sum_{t=0}^{+\infty}\left[(I_n - \Theta)\frac{1}{n}\mathbb{1}_n\mathbb{1}_n^\top\right]^t\Theta \\ &= \left[I_n + \frac{1}{\sum_{i=1}^n \theta_i}(I_n - \Theta)\mathbb{1}_n\mathbb{1}_n^\top\right]\Theta \end{aligned}$$

and clearly, $Y_\infty = M_\infty Y(0)$.

If $S_0 \geq 0$ is reducible, the expression of M_∞ can be derived in a similar way.

The case where S_0 is not a nonnegative matrix but the associated graph is structurally balanced, can be easily reduced to the cooperative case by making use of the following technical lemma.

Lemma 8. *Given a signature matrix*

$$D = \mathrm{diag}\{d_1, \dots, d_n\}, \quad \text{with} \quad d_i \in \{-1, 1\}, \ i \in [1, n],$$

set $\tilde{Y}(t) := DY(t)$ *and* $\tilde{W}(t+1) := DW(t+1)D$. *Then*

$$\tilde{Y}(t+1) = (I_n - \Theta)\tilde{W}(t+1)\tilde{Y}(t) + \Theta\tilde{Y}(0)$$
$$\tilde{W}(t+1) = \frac{1}{n}\mathrm{sgn}\left(\tilde{Y}(t)\tilde{Y}(t)^\top\right).$$

Proposition 9 extends the result of Theorem 6 to the case where S_0 is irreducible and has a structurally balanced graph (Altafini, 2013; Xia et al., 2016). The extension to the case where S_0 is reducible and has a structurally balanced graph is analogous and omitted for the sake of brevity.

Proposition 9. *Suppose that* S_0 *is irreducible and its associated graph,* $\mathcal{G}(S_0)$, *is structurally balanced, so that there exists a signature matrix* $D = \mathrm{diag}\{d_1, \dots, d_n\}, d_i \in \{-1, 1\}, \ i \in [1, n]$, *s.t.* $DS_0D \geq 0$. *Then,*

$$W_\infty = \lim_{t\to+\infty} W(t+1) = \frac{1}{n}D\mathbb{1}_n(D\mathbb{1}_n)^\top,$$
$$M_\infty = \left[I_n + \frac{1}{\sum_{i=1}^n \theta_i}(I_n - \Theta)D\mathbb{1}_n(D\mathbb{1}_n)^\top\right]\Theta.$$

6. Single-topic case

We now address the case where $m = 1$, namely there is only one discussion topic. When so, the opinion matrix is a column vector, that we now denote by $y(t) \in \mathbb{R}^n$, containing the opinions of the agents on the topic. It is easy to see that if we define $v(t) := \mathrm{sgn}(y(t))$, then $\mathrm{sgn}\left(y(t)y(t)^\top\right) = v(t)v(t)^\top$, and model (1)–(2) becomes:

$$y(t+1) = (I - \Theta)W(t+1)y(t) + \Theta y(0) \tag{17}$$
$$W(t+1) = \frac{1}{n}\mathrm{sgn}\left(y(t)y(t)^\top\right) = \frac{1}{n}v(t)v(t)^\top, \tag{18}$$

leading to the difference equation:

$$y(t+1) = \frac{1}{n}(I_n - \Theta)v(t)v(t)^\top y(t) + \Theta y(0). \tag{19}$$

We also note that in this context Assumption 2 amounts to imposing that $y(0)$ is devoid of zero entries. In fact, condition $y_i(0) = 0$ would lead the ith agent to remain isolated and stick to the zero opinion.

Under the previous hypotheses, we can derive the following results.

Lemma 10. *For* $m = 1$,

$$v(t) = \mathrm{sgn}(y(t)) = \mathrm{sgn}(y(0)) = v(0), \qquad \forall\, t \in \mathbb{Z}_+.$$

Consequently, $W(t+1) = W(1) = \frac{1}{n}v(0)v(0)^\top$, $\forall t \geq 1$, *namely the influence matrix remains constant.*

As a consequence of the previous lemma, for $m = 1$ the model in (17)–(18) becomes time-invariant and the dynamics of $y(t)$ can be expressed as:

$$y(t+1) = \frac{1}{n}(I_n - \Theta)v(0)v(0)^\top y(t) + \Theta y(0). \tag{20}$$

Lemma 10 implies that the whole opinion dynamics evolves at each time step with an influence matrix that corresponds to a situation of structural balance, by this meaning that $\mathcal{G}(W(t+1))$ is structurally balanced for every $t \geq 0$. We can now derive the following result.

Theorem 11. *For* $m = 1$, *the matrix sequence* $\{M(t)\}_{t\in\mathbb{Z}_+}$ *always converges to a constant limit matrix* M_∞ *and*

$$M_\infty = \left[I_n + \frac{1}{\sum_{i=1}^n \theta_i}(I_n - \Theta)v(0)v(0)^\top\right]\Theta, \tag{21}$$
$$W_\infty = \frac{1}{n}v(0)v(0)^\top.$$

To conclude, we can provide an explicit expression for the agents' asymptotic opinions $y_\infty := \lim_{t\to+\infty} y(t)$, namely

$$y_\infty = \left[I_n + \frac{1}{\sum_{i=1}^n \theta_i}(I_n - \Theta)v(0)v(0)^\top\right]\Theta y(0).$$

Note, finally, that

$$W_\infty = \frac{1}{n}\mathrm{sgn}(y_\infty y_\infty^\top) = \frac{1}{n}\mathrm{sgn}(M_\infty y(0)y(0)^\top M_\infty^\top).$$

7. Interrelated topics: average-based appraisal matrix

To conclude, we propose a slightly modified version of the model that preserves Eq. (1), but replaces Eq. (2) with the following expression

$$W(t+1) = \frac{1}{n}\mathrm{sgn}\left(y_m(t)y_m(t)^\top\right), \tag{22}$$

where $y_m(t) := \frac{Y(t)\mathbb{1}_m}{m}$. This means that the appraisal matrix $W(t)$ updates based not on the opinions that agents express on each single topic, but on the average values of such opinions, which is a reasonable assumption provided that the topics on which agents express their opinions are correlated.

For what concerns the dynamics of $W(t)$ in (22), it is easy to see that the result in Lemma 10 applies also to this case and, in particular, we have

$$W(t) = W(1) = \frac{1}{n}\mathrm{sgn}\left(y_m(0)y_m(0)^\top\right) = \frac{1}{n}v(0)v(0)^\top$$

where $v(0) := \mathrm{sgn}(y_m(0))$. Thus, Eq. (1) can be rewritten as:

$$Y(t+1) = (I_n - \Theta)\frac{1}{n}v(0)v(0)^\top Y(t) + \Theta Y(0). \tag{23}$$

We can easily notice that in this case, in the influence matrix update, the role that used to be played by the matrix $S_0 = Y(0)Y(0)^\top$ is now played by the matrix $y_m(0)y_m(0)^\top$, which has clearly a structurally balanced graph. Therefore, we can adapt the analysis in the second part of Section 5 also to this model and deduce what follows.

Corollary 12. *Consider the opinion dynamics model* (1)–(22). *If* $y_m(0) \in \mathbb{R}^n$ *is devoid of zero entries, then,*

$$W_\infty = \lim_{t\to+\infty} W(t+1) = \frac{1}{n}v(0)v(0)^\top,$$
$$M_\infty = \left[I_n + \frac{1}{\sum_{i=1}^n \theta_i}(I_n - \Theta)v(0)v(0)^\top\right]\Theta,$$

where $v(0) = \mathrm{sgn}(y_m(0))$.

8. Examples

Example 1. Consider a group of $n = 6$ agents discussing $m = 6$ topics. Assume that $\theta_1 = \theta_6 = \frac{2}{3}, \theta_2 = \theta_5 = \frac{1}{2}, \theta_3 = \theta_4 = \frac{1}{3}$ and that $Y(0)$ is:

$$Y(0) = \begin{bmatrix} -0.1317 & 1.7035 & -0.2350 & 0.0802 & 0.7824 & -0.6380 \\ 0.2968 & -0.6272 & 0.9015 & -0.4425 & -0.1206 & -0.7040 \\ -0.6075 & -0.3453 & 0.3935 & -0.9496 & 0.5671 & -0.3654 \\ 0.5217 & -0.2691 & -0.2884 & -0.1193 & -0.3721 & -1.1914 \\ 0.0244 & -0.2168 & -0.2278 & 1.1211 & -0.3104 & -0.7398 \\ -0.3392 & 0.7993 & 0.1429 & -0.9816 & -1.4906 & 0.2002 \end{bmatrix}.$$

The evolutions of the opinions on the 6 topics as well as the evolution of the influence matrix and the final influence graph of the network, $\mathcal{G}(W_\infty)$, are illustrated in Fig. 1. In the graph solid (blue) lines represent positive edges, while dashed (red) lines represent negative edges.

Example 2. Consider a group of $n = 12$ agents discussing $m = 3$ topics. Assume that $\theta_1 = 0.7461$, $\theta_2 = 0.6625$, $\theta_3 = 0.5233$, $\theta_4 = 0.2599$, $\theta_5 = 0.9620$, $\theta_6 = 0.5402$, $\theta_7 = 0.0303$, $\theta_8 = 0.6963$, $\theta_9 = 0.5197$, $\theta_{10} = 0.0590$, $\theta_{11} = 0.8900$, $\theta_{12} = 0.3302$, and that $Y(0)$ is:

$$Y(0) = \begin{bmatrix} -30.3795 & 14.3698 & -10.4634 \\ -19.6148 & 36.0099 & 49.2175 \\ -1.6705 & -9.8117 & -9.7648 \\ -16.2188 & 13.19308 & 15.8857 \\ 29.8486 & 48.5237 & 40.1348 \\ 48.7488 & 5.9477 & 49.5382 \\ -34.0952 & 43.3592 & 15.3163 \\ -26.3120 & 22.0343 & -39.1564 \\ 20.2237 & -1.5961 & -46.3886 \\ -12.4528 & 13.9031 & 11.8091 \\ 47.3705 & 38.7637 & 6.7144 \\ 47.2306 & -30.1263 & 46.1965 \end{bmatrix}.$$

The evolutions of the opinions on the 3 topics as well as the evolution of the influence matrix and the final graph $\mathcal{G}(W_\infty)$ are illustrated in Fig. 2.

9. Conclusions

In this paper we proposed an extended version of the Friedkin–Johnsen model whose influence matrix takes values in a finite set and its entries update according to a homophily principle, namely their values depend on the correlations between the agents' opinion vectors. We proved that the opinion matrix asymptotically converges to a constant solution if and only if the influence matrix becomes constant in finite time. Moreover, we highlighted some interesting properties of the limit values of the transition matrix. Some special cases where the asymptotic behavior can be derived in closed form were proposed. Finally, the case where the topics are correlated and the influence matrix depends on the average opinion vectors was investigated.

Acknowledgment

The Authors are indebted with Mauro Bisiacco for part of the proof of Theorem 2.

Appendix. Proofs

Proof of Proposition 1. We prove the result by induction on t. We first show that the result is true for $t = 0$. We observe that

$$W(1) = \frac{1}{n}\text{sgn}(Y(0)Y(0)^\top) = \frac{1}{n}\text{sgn}(M(0)S_0M(0)^\top),$$

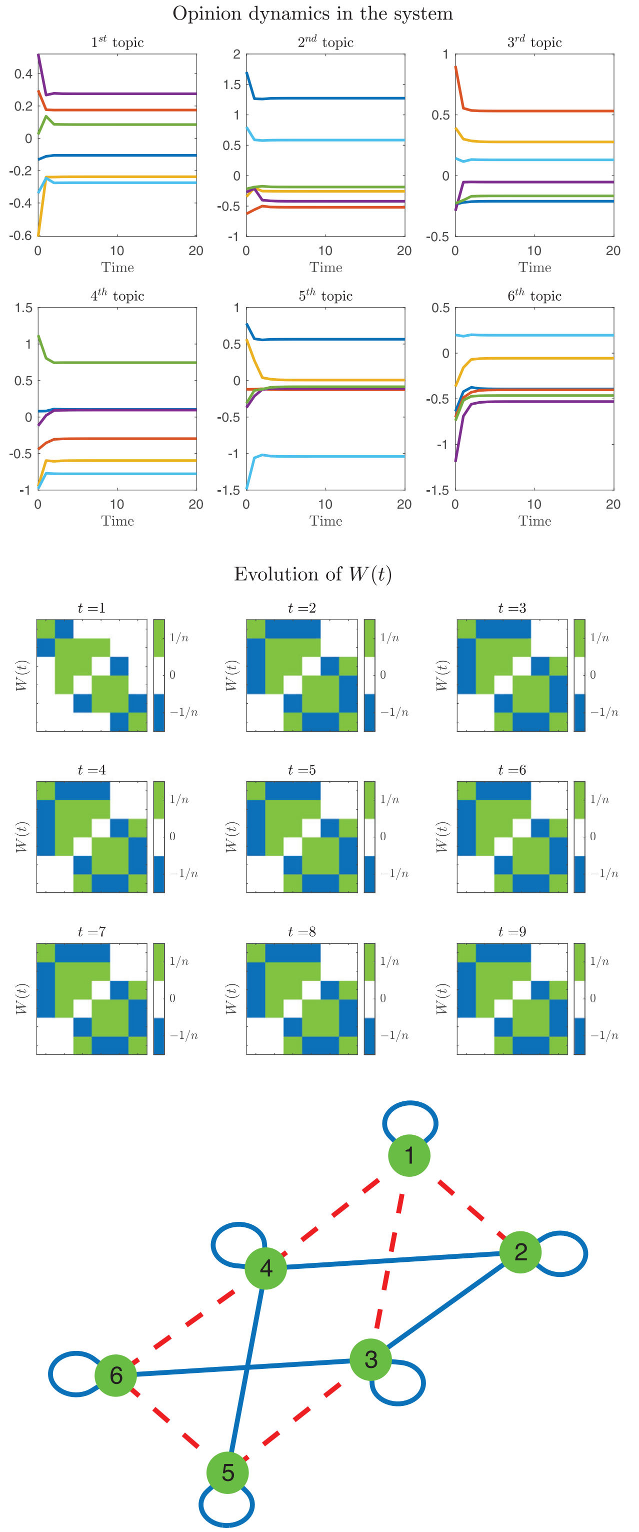

Fig. 1. Example 1: Evolutions of the opinions on the 6 topics (top); Evolution of the influence matrix (middle); Graph associated with W_∞ (bottom).

and hence

$$\begin{aligned} Y(1) &= [(I_n - \Theta)W(1) + \Theta]\,Y(0) \\ &= [(I_n - \Theta)W(1)M(0) + \Theta]\,Y(0) = M(1)Y(0), \end{aligned}$$

where

$$M(1) = (I_n - \Theta)W(1)M(0) + \Theta.$$

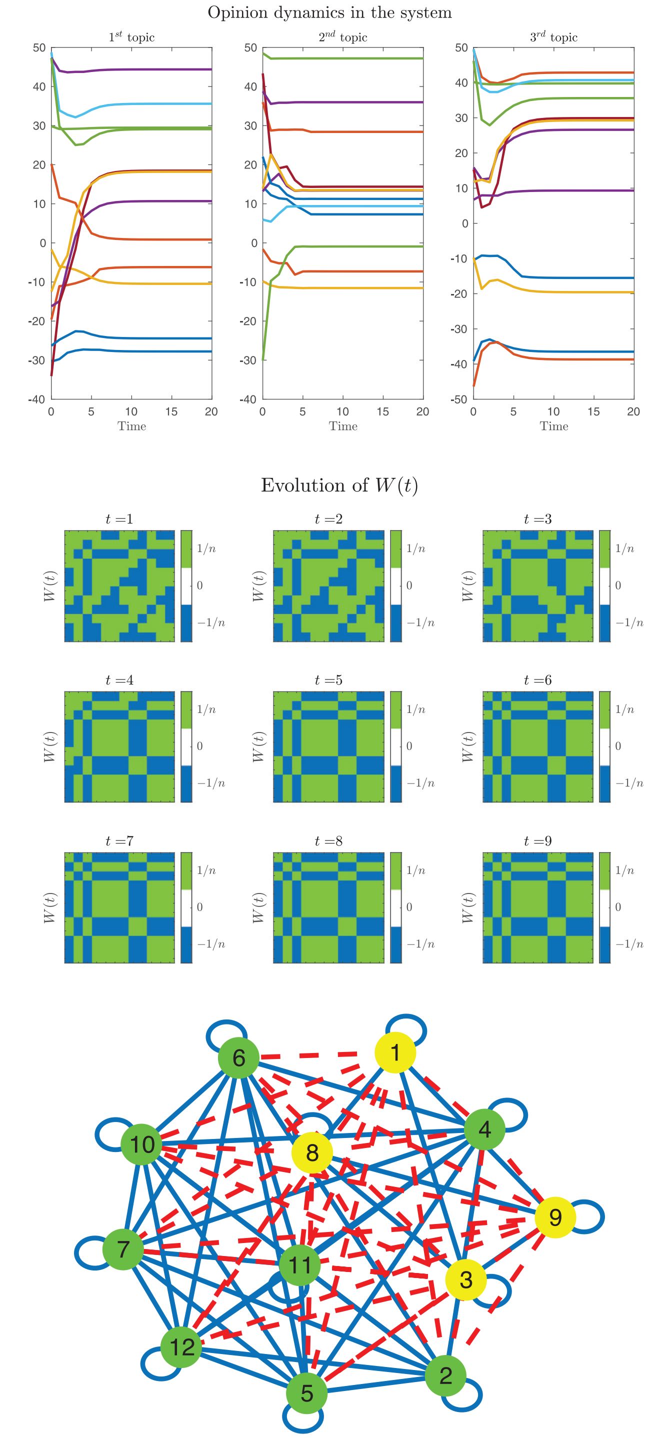

Fig. 2. Example 2: Evolutions of the opinions on the 3 topics (top); Evolution of the influence matrix (middle); Graph associated with W_∞ (bottom).

Now we assume that Eqs. (4), (5) and (7) are true for $t < \bar{t}$ and prove that they hold true also for $t = \bar{t}$.

From $W(\bar{t}+1) = \frac{1}{n}\mathrm{sgn}(Y(\bar{t})Y(\bar{t})^\top)$, by the inductive assumption (on the expression of Y), we obtain $W(\bar{t}+1) = \frac{1}{n}\mathrm{sgn}(M(\bar{t})S_0M(\bar{t})^\top)$. On the other hand,

$$\begin{aligned}Y(\bar{t}+1) &= (I_n-\Theta)W(\bar{t}+1)Y(\bar{t})+\Theta Y(0)\\ &= \left[(I_n-\Theta)W(\bar{t}+1)M(\bar{t})+\Theta\right]Y(0)\\ &= M(\bar{t}+1)Y(0),\end{aligned}$$

where $M(\bar{t}+1) = (I_n-\Theta)W(\bar{t}+1)M(\bar{t})+\Theta$. □

Proof of Theorem 2. We prove the result by induction on $t \in \mathbb{Z}_+$. We first observe that $\|M(0)\|_{max} = \|I_n\|_{max} = 1$. We now assume that $\|M(t)\|_{max} \le 1$ and prove that $\|M(t+1)\|_{max} \le 1$. From (5) we get

$$\begin{aligned}\|M(t+1)\|_{max} &= \|(I_n-\Theta)W(t+1)M(t)+\Theta\|_{max}\\ &\le \|(I_n-\Theta)\frac{1}{n}\mathbb{1}_n\mathbb{1}_n^\top\mathbb{1}_n\mathbb{1}_n^\top+\Theta\|_{max}\\ &= \|(I_n-\Theta)\mathbb{1}_n\mathbb{1}_n^\top+\Theta\|_{max} = 1.\end{aligned}$$

(i) ⇒ *(ii)* From (5) we easily deduce that

$$\begin{aligned}M(t+1)-M(t) &= (I_n-\Theta)W(t+1)M(t)\\ &\quad -(I_n-\Theta)W(t)M(t-1),\\ &= (I_n-\Theta)[W(t+1)-W(t)]M(t)\\ &\quad +(I_n-\Theta)W(t)[M(t)-M(t-1)].\end{aligned}$$

This implies

$$\begin{aligned}&\|\,M(t+1)-M(t)\,\|_{max}\\ &\quad\ge \|(I_n-\Theta)[W(t+1)-W(t)]M(t)\|_{max}\\ &\quad -\|(I_n-\Theta)W(t)[M(t)-M(t-1)]\|_{max}.\end{aligned}\tag{A.1}$$

We notice that

$$\begin{aligned}&\|\,(I_n-\Theta)\ \ W(t)[M(t)-M(t-1)]\,\|_{max}\\ &\quad\le n\|(I_n-\Theta)W(t)\|_{max}\|M(t)-M(t-1)\|_{max}\\ &\quad< \|M(t)-M(t-1)\|_{max},\end{aligned}\tag{A.2}$$

where we used the fact that $\|AB\|_{max} \le n\|A\|_{max}\|B\|_{max}$ and that $\|(I_n-\Theta)W(t)\|_{max} \le \max_i \frac{1-\theta_i}{n} < \frac{1}{n}$. So, (A.1) and (A.2) together lead to

$$\begin{aligned}&\|\,M(t+1)-M(t)\,\|_{max}+\|M(t)-M(t-1)\|_{max}\\ &\quad> \|(I_n-\Theta)[W(t+1)-W(t)]M(t)\|_{max}.\end{aligned}\tag{A.3}$$

From Eq. (5) we deduce

$$\begin{aligned}M(t) &= (I_n-\Theta)W(t)M(t-1)+\Theta\\ &= (I_n-\Theta)W(t)M(t)\\ &\quad+(I_n-\Theta)W(t)[M(t-1)-M(t)]+\Theta,\end{aligned}$$

that leads to

$$\begin{aligned}M(t) &= [I_n-(I_n-\Theta)W(t)]^{-1}[(I_n-\Theta)W(t)\\ &\quad\cdot\ (M(t-1)-M(t))+\Theta],\end{aligned}\tag{A.4}$$

where we used the fact that $(I_n-\Theta)W(t)$ is Schur[2] and hence $I_n-(I_n-\Theta)W(t)$ is nonsingular.

Assume, now, that there exists $\lim_{t\to+\infty} M(t)$. This means that for every $\varepsilon > 0$ there exists $T \in \mathbb{Z}_+$ such that for every $t \ge T$ we have $\|M(t)-M(t-1)\|_{max} < \varepsilon$. If ε is sufficiently small, this also ensures that, for every $t \ge T$, the matrix $M(t)-M(t+1)$ is infinitesimal, and that the matrix $(I_n-\Theta)W(t)(M(t-1)-M(t))+\Theta \approx \Theta$ is nonsingular. As a consequence, (A.4) leads to

$$\begin{aligned}M(t)^{-1} &= [(I_n-\Theta)W(t)(M(t-1)-M(t))+\Theta]^{-1}\\ &\quad\cdot\ [I_n-(I_n-\Theta)W(t)].\end{aligned}$$

To summarize, for every $t \ge T$ the matrix $M(t)^{-1}$ exists and is a bounded matrix (being the product of two bounded matrices). This guarantees that there exists $b > 0$, such that

$$\|M(t)^{-1}\|_{max} < b, \quad \forall t \ge T.\tag{A.5}$$

[2] The result can be easily proved by resorting to Gershgorin Circles Theorem (Horn & Johnson, 1985).

Now, we use the fact that

$$\begin{aligned}&\|(I_n-\Theta)[W(t+1)-W(t)]\|_{max}\\&=\|(I_n-\Theta)[W(t+1)-W(t)]M(t)M(t)^{-1}\|_{max}\\&\le n\|(I_n-\Theta)[W(t+1)-W(t)]M(t)\|_{max}\|M(t)^{-1}\|_{max}\end{aligned}$$

and hence

$$\|(I_n-\Theta)[W(t+1)-W(t)]M(t)\|_{max}\ge \frac{\|(I_n-\Theta)[W(t+1)-W(t)]\|_{max}}{n\|M(t)^{-1}\|_{max}}. \tag{A.6}$$

By replacing (A.6) in (A.3), and keeping into account (A.5), we obtain

$$\|M(t+1)-M(t)\|_{max}+\|M(t)-M(t-1)\|_{max} > \frac{1}{nb}\|(I_n-\Theta)[W(t+1)-W(t)]\|_{max}. \tag{A.7}$$

This guarantees that for every $t\ge T$

$$2\varepsilon > \frac{1}{nb}\|(I_n-\Theta)[W(t+1)-W(t)]\|_{max}.$$

Since the matrix W takes values in a finite set, and hence the nonzero entries of $W(t+1)-W(t)$ cannot be arbitrarily small, this ensures that $W(t+1)=W(t)$, $\forall t\ge T$.

(ii) ⇒ *(i)* If condition *(ii)* holds, then for every $t\ge T$ we have

$$\begin{aligned}M(t)&=(I_n-\Theta)W(t)M(t-1)+\Theta\\&=(I_n-\Theta)W_\infty M(t-1)+\Theta,\end{aligned}$$

and hence

$$M(t+1)-M(t)=(I_n-\Theta)W_\infty[M(t)-M(t-1)].$$

This implies that for every $t\ge T$

$$\begin{aligned}&\|M(t+1)-M(t)\|_{max}\\&=\|(I_n-\Theta)W_\infty[M(t)-M(t-1)]\|_{max}\\&\le n\|(I_n-\Theta)W_\infty\|_{max}\|M(t)-M(t-1)\|_{max}\\&\le \alpha\|M(t)-M(t-1)\|_{max}\\&\le \alpha^{t-T+1}\|M(T)-M(T-1)\|_{max},\end{aligned}$$

where we used the fact that $\|AB\|_{max}\le n\|A\|_{max}\|B\|_{max}$ and that $\|(I_n-\Theta)W_\infty\|_{max}\le \max_i\frac{1-\theta_i}{n}=\frac{\alpha}{n}$, $\alpha:=\max_{i\in[1,n]}(1-\theta_i)<1$. This ensures that

$$\lim_{t\to+\infty}\|M(t+1)-M(t)\|_{max}=0$$

and hence there exists $M_\infty=\lim_{t\to+\infty}M(t)$.

Condition (8) follows immediately. □

Proof of Proposition 3. We first prove that M_∞ is nonsingular. Suppose, by contradiction, that $v\in\mathbb{R}^n$, $v\ne \mathbb{0}_n$, belongs to the kernel of M_∞, i.e., $M_\infty v=\mathbb{0}_n$. Then, by making use of (8), we obtain

$$\mathbb{0}_n=M_\infty v=(I-\Theta)W_\infty M_\infty v+\Theta v \ \Rightarrow\ \Theta v=\mathbb{0}_n,$$

which is not possible as each $\theta_i\in(0,1)$, by Assumption 1.

(i) Let i be any index in $[1,n]$. Then

$$M_\infty \mathbb{e}_i=(I-\Theta)W_\infty M_\infty\mathbb{e}_i+\Theta\mathbb{e}_i.$$

If we permute the entries of $M_\infty\mathbb{e}_i$, using an $n\times n$ permutation matrix P, in such a way that $\tilde v:=P^\top M_\infty\mathbb{e}_i=\begin{bmatrix}\tilde v_1 & \dots & \tilde v_n\end{bmatrix}^\top$, with $|\tilde v_1|\ge|\tilde v_2|\ge\dots\ge|\tilde v_n|$, we obtain

$$\begin{aligned}\tilde v&=P^\top M_\infty\mathbb{e}_i=P^\top(I-\Theta)PP^\top W_\infty PP^\top M_\infty\mathbb{e}_i+\\&+P^\top\Theta PP^\top\mathbb{e}_i=(I-\tilde\Theta)\tilde W_\infty\tilde v+\tilde\Theta\mathbb{e}_j,\ \exists j\in[1,n],\end{aligned}$$

where $\tilde\Theta=P^\top\Theta P=\mathrm{diag}\{\tilde\theta_1,\dots,\tilde\theta_n\}$ and $\tilde W_\infty=P^\top W_\infty P$. The first component of $\tilde v$, i.e., $\tilde v_1$, satisfies

$$\tilde v_1=(1-\tilde\theta_1)\mathbb{e}_1^\top\tilde W_\infty\begin{bmatrix}\tilde v_1 & \dots & \tilde v_n\end{bmatrix}^\top+\tilde\theta_1\mathbb{e}_1^\top\mathbb{e}_j,$$

which implies that

$$|\tilde v_1|\le(1-\tilde\theta_1)\sum_{i=1}^n\frac{|\tilde v_i|}{n}+\tilde\theta_1\mathbb{e}_1^\top\mathbb{e}_j. \tag{A.8}$$

Therefore, if $j\ne1$, the right-hand side of (A.8) would be

$$(1-\tilde\theta_1)\sum_{i=1}^n\frac{|\tilde v_i|}{n}<\sum_{i=1}^n\frac{|\tilde v_i|}{n}\le|\tilde v_1|,$$

a contradiction. Thus, it must be $j=1$ and $\tilde\theta_1=\theta_i$. So, we have $\tilde v_1=[M_\infty\mathbb{e}_i]_i=[M_\infty]_{ii}$. This means that $\max_{j\in[1,n]}|[M_\infty]_{ji}|=|[M_\infty]_{ii}|$. Clearly, this is true for every index $i\in[1,n]$, namely for every column of M_∞.

(ii) We want to prove that $[M_\infty]_{ii}>0, \forall i\in[1,n]$, which is equivalent to showing that $\tilde v_1>0$, by referring to the notation adopted in part *(i)*. Suppose, by contradiction, that $\tilde v_1\le0$. Then, we get

$$\begin{aligned}\tilde v_1&=(1-\tilde\theta_1)\mathbb{e}_1^\top\tilde W_\infty\begin{bmatrix}\tilde v_1 & \dots & \tilde v_n\end{bmatrix}^\top+\tilde\theta_1\\&=(1-\tilde\theta_1)[\tilde W_\infty]_{11}\tilde v_1+(1-\tilde\theta_1)\sum_{j\ne1}[\tilde W_\infty]_{1j}\tilde v_j+\tilde\theta_1.\end{aligned}$$

Consequently,

$$\left(1-(1-\tilde\theta_1)[\tilde W_\infty]_{11}\right)\tilde v_1-\tilde\theta_1=(1-\tilde\theta_1)\sum_{j\ne1}[\tilde W_\infty]_{1j}\tilde v_j.$$

Note that since

$$\begin{aligned}[\tilde W_\infty]_{11}&=\left[\frac1n\mathrm{sgn}(P^\top Y(T)Y(T)^\top P)\right]_{11}\\&=\frac1n\mathrm{sgn}(\mathbb{e}_1^\top P^\top Y(T)Y(T)^\top P\mathbb{e}_1)\\&=\frac1n\mathrm{sgn}(\|Y(T)^\top P\mathbb{e}_1\|_2^2)\end{aligned}$$

for some $T\in\mathbb{Z}_+$, $[\tilde W_\infty]_{11}$ belongs to $\{0,\frac1n\}$, and $1-(1-\tilde\theta_1)[\tilde W_\infty]_{11}>0$. If $\tilde v_1\le0$, we get

$$\begin{aligned}&\left(1-(1-\tilde\theta_1)[\tilde W_\infty]_{11}\right)|\tilde v_1|+\tilde\theta_1\\&=\left|\left(1-(1-\tilde\theta_1)[\tilde W_\infty]_{11}\right)\tilde v_1-\tilde\theta_1\right|\\&=(1-\tilde\theta_1)\left|\textstyle\sum_{j\ne1}[\tilde W_\infty]_{1j}\tilde v_j\right|\le(1-\tilde\theta_1)\frac{n-1}{n}|\tilde v_1|,\end{aligned}$$

which implies that

$$\left[1-(1-\tilde\theta_1)\left([\tilde W_\infty]_{11}+\frac{n-1}{n}\right)\right]|\tilde v_1|+\tilde\theta_1\le0. \tag{A.9}$$

Since $[\tilde W_\infty]_{11}$ belongs to $\{0,\frac1n\}$, then $1-(1-\tilde\theta_1)\left([\tilde W_\infty]_{11}+\frac{n-1}{n}\right)>0$. So, all quantities on the left-hand side of (A.9) are nonnegative and, in particular, $\tilde\theta_1$ is positive. This contradicts inequality (A.9). Therefore, $\tilde v_1$ must be positive, which is equivalent to saying that $[M_\infty]_{ii}>0, \forall i\in[1,n]$. □

Proof of Proposition 4. We prove the result by induction on t. We first show that the result is true for $t=0$. For $t=0$, by the assumption on S_0, we have

$$W(1)=\frac1n\mathrm{sgn}\left(\begin{bmatrix}S_{0,1}&&\\&\ddots&\\&&S_{0,k}\end{bmatrix}\right)=\begin{bmatrix}W_1(1)&&\\&\ddots&\\&&W_k(1)\end{bmatrix}.$$

Moreover, as $M(0) = I_n$ and Θ is a diagonal matrix,

$$M(1) = (I_n - \Theta)\text{blockdiag}\{W_1(1), \ldots, W_k(1)\} + \Theta = \text{blockdiag}\{M_1(1), \ldots, M_k(1)\}.$$

We now suppose that the result is true for $t < \bar{t}$, and prove it for $t = \bar{t}$. By Proposition 1 and the inductive hypothesis, we have that $M(\bar{t})$ is block diagonal with k diagonal blocks of sizes $n_1, \ldots, n_k$, and hence so is $M(\bar{t})S_0M(\bar{t})^\top$. This ensures that $W(\bar{t}+1) = \frac{1}{n}\text{sgn}(M(\bar{t})S_0M(\bar{t})^\top)$ has the same block diagonal structure. Finally, as $M(\bar{t}+1) = (I_n - \Theta)W(\bar{t}+1)M(\bar{t}) + \Theta$, then $M(\bar{t}+1)$ is block diagonal with k diagonal blocks of sizes $n_1, \ldots, n_k$. This concludes the proof. □

Proof of Proposition 5. We prove the result by induction on t. For $t = 1$, we have:

$$W(1) = \frac{1}{n}\text{sgn}(S_0) \Rightarrow \text{sgn}(W(1)) = \text{sgn}(S_0^{\beta(1)}).$$

Moreover, from (5) and (6) we get $\text{sgn}(M(1)) = \text{sgn}(W(1)) = \text{sgn}(S_0) = \text{sgn}(S_0^{\alpha(1)}) = \text{sgn}(M(1)^\top)$. Now we assume that the result is true for $t \le \bar{t}$ and we prove that it is true also for $t = \bar{t}+1$. By making use of (5) and (7) for $t = \bar{t}$, we deduce that[3]

$$\begin{aligned}\text{sgn}(W(\bar{t}+1)) &= \text{sgn}(M(\bar{t})S_0M(\bar{t})^\top)\\ &= \text{sgn}(\text{sgn}(M(\bar{t}))\text{sgn}(S_0)\text{sgn}(M(\bar{t})^\top))\\ &= \text{sgn}(\text{sgn}(S_0^{\alpha(\bar{t})})\text{sgn}(S_0)\text{sgn}(S_0^{\alpha(\bar{t})}))\\ &= \text{sgn}(S_0^{\alpha(\bar{t})}S_0S_0^{\alpha(\bar{t})}) = \text{sgn}(S_0^{2\alpha(\bar{t})+1}).\end{aligned}$$

This proves that $\text{sgn}(W(\bar{t}+1)) = \text{sgn}(S_0^{\beta(\bar{t}+1)})$ where $\beta(\bar{t}+1) = 2\alpha(\bar{t})+1$. On the other hand, $\text{sgn}(M(\bar{t}+1)) = \text{sgn}(W(\bar{t}+1)M(\bar{t})) = \text{sgn}(S_0^{2\alpha(\bar{t})+1}S_0^{\alpha(\bar{t})}) = \text{sgn}(S_0^{3\alpha(\bar{t})+1}) = \text{sgn}(S_0^{\alpha(\bar{t}+1)}) = \text{sgn}(M(\bar{t}+1)^\top)$, where $\alpha(\bar{t}+1) = 3\alpha(\bar{t})+1$. □

Proof of Theorem 6. Condition $\text{sgn}(M(t)) = \text{sgn}(S_0^{\alpha(t)})$ ensures that $M(t) \ge 0$ for every $t \ge 0$ and hence also $Y(t)Y(t)^\top = M(t)S_0M(t)^\top \ge 0$ for every $t \ge 0$.
(i) If S_0 is an irreducible matrix, then having positive diagonal entries (by Assumption 2), it is also primitive. Therefore, there exists $T > 0$ such that $S_0^{\beta(t)}$ is strictly positive for every $t \ge T$. This implies that $W(t) = \frac{1}{n}\mathbb{1}_n\mathbb{1}_n^\top = W_\infty, \forall t \ge T$.
(ii) If S_0 is reducible, we can assume, for the sake of simplicity, that $P = I_n$, i.e., $S_0 = \text{blockdiag}\{S_{0,1}, \ldots, S_{0,k}\}$, where each $S_{0,i}$ is a symmetric, nonnegative and primitive block of size n_i. By Proposition 4 and subsequent comments, we can partition the dynamics of the matrix $Y(t)$ in k blocks. The dynamics of each block $Y_i(t)$ is independent of the dynamics of the other blocks, and its initial condition $Y_i(0)$ is such that $Y_i(0)Y_i(0)^\top$ is a nonnegative irreducible matrix. Therefore we can resort to case *(i)* to conclude the proof. □

Proof of Lemma 8. From (1)–(2) we obtain

$$\begin{aligned}\tilde{Y}(t+1) &= DY(t+1)\\ &= D(I_n - \Theta)W(t+1)D\tilde{Y}(t) + D\Theta Y(0)\\ &= (I_n - \Theta)[DW(t+1)D]\tilde{Y}(t) + \Theta\tilde{Y}(0).\end{aligned}$$

On the other hand,

$$\tilde{W}(t+1) = \frac{1}{n}D\,\text{sgn}\left(Y(t)Y(t)^\top\right)D = \frac{1}{n}\text{sgn}(\tilde{Y}(t)\tilde{Y}(t)^\top)$$

where the last equality holds since

$$\begin{aligned}\left[\text{sgn}\left(\tilde{Y}(t)\tilde{Y}(t)^\top\right)\right]_{ij} &= \text{sgn}\left(\mathbb{e}_i^\top DY(t)Y(t)^\top D\mathbb{e}_j\right)\\ &= \text{sgn}\left(d_i\mathbb{e}_i^\top Y(t)Y(t)^\top\mathbb{e}_jd_j\right) = d_i\left[\text{sgn}\left(Y(t)Y(t)^\top\right)\right]_{ij}d_j\\ &= \left[D\text{sgn}\left(Y(t)Y(t)^\top\right)D\right]_{ij}.\end{aligned}$$ □

Proof of Proposition 9. By making use of Lemma 8 and Theorem 6 part (i), we can claim that since $\tilde{Y}(0)\tilde{Y}(0)^\top = DS_0D$ is an irreducible nonnegative matrix with positive diagonal entries, and hence primitive, then

$$\tilde{W}_\infty := \lim_{t\to+\infty}\tilde{W}(t+1) = \frac{1}{n}\mathbb{1}_n\mathbb{1}_n^\top.$$

Therefore, $W_\infty = \lim_{t\to+\infty} D\tilde{W}(t+1)D = \frac{1}{n}D\mathbb{1}_n\mathbb{1}_n^\top D$. The rest immediately follows. □

Proof of Lemma 10. By induction on t. For $t = 1$, we have $v(1) = \text{sgn}(y(1)) = \text{sgn}\left[(I - \Theta)\frac{1}{n}v(0)v(0)^\top y(0) + \Theta y(0)\right] = \text{sgn}(y(0)) = v(0)$, where we used the fact that $v(0)^\top y(0) = \text{sgn}(y(0))^\top y(0) = \sum_{i=1}^n |y_i(0)| > 0$ (Assumption 2 rules out the case $y(0) = \mathbb{0}_n$). Suppose, now, that the result holds for $t < \bar{t}$. For $t = \bar{t}$:

$$\begin{aligned}v(\bar{t}+1) &= \text{sgn}(y(\bar{t}+1))\\ &= \text{sgn}\Big[(I_n - \Theta)\frac{1}{n}\overbrace{v(\bar{t})}^{=v(0)}\underbrace{v(\bar{t})^\top y(\bar{t})}_{\sum_{i=1}^n |y_i(\bar{t})|>0} + \Theta y(0)\Big] = v(0),\end{aligned}$$

where we used the inductive assumption $v(\bar{t}) = \text{sgn}(y(\bar{t})) = \text{sgn}(y(0)) = v(0)$ that ensures, in particular, $y(\bar{t}) \ne \mathbb{0}_n$. The second part immediately follows. □

Proof of Theorem 11. Lemma 10 ensures that $W(t) = \frac{1}{n}v(0)v(0)^\top$ for every $t \ge 1$. So, by Theorem 2, we can claim that $\exists M_\infty = \lim_{t\to+\infty} M(t)$ and that $W_\infty = \frac{1}{n}v(0)v(0)^\top$.

Moreover, from (8) we get $[I_n - (I_n - \Theta)W_\infty]M_\infty = \Theta$. By Gershgorin Circles Theorem (Horn & Johnson, 1985) and Assumption 1, we can claim that $(I_n - \Theta)W_\infty$ is Schur stable and hence $I_n - (I_n - \Theta)W_\infty$ is invertible. Consequently, $M_\infty = [I_n - (I_n - \Theta)W_\infty]^{-1}\Theta$. Finally,

$$\begin{aligned}[I_n - (I_n - \Theta)W_\infty]^{-1} &= I_n + \sum_{k=1}^{+\infty}[(I_n - \Theta)\tfrac{1}{n}v(0)v(0)^\top]^k\\ &= I_n + [(I_n - \Theta)\tfrac{1}{n}v(0)v(0)^\top]\sum_{k=1}^{+\infty}\left(\frac{\sum_{i=1}^n(1-\theta_i)}{n}\right)^{k-1}\\ &= I_n + \frac{1}{\sum_{i=1}^n\theta_i}(I_n - \Theta)v(0)v(0)^\top.\end{aligned}$$

Thus, M_∞ is expressed as in (22). □

References

Altafini, C. (2012). Dynamics of opinion forming in structurally balanced social networks. *PLoS One*, *7*(6), Article e38135.

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

Bizyaeva, A., Franci, A., & Leonard, N. (2023). Nonlinear opinion dynamics with tunable sensitivity. *IEEE Transactions on Automatic Control*, *68*(3), 1415–1430.

Dandekar, P., Goel, A., & Lee, D. T. (2013). Biased assimilation, homophily, and the dynamics of polarization. *Proceedings of the National Academy of Sciences*, *110*(15), 5791–5796.

De Pasquale, G., & Valcher, M. (2022). A bandwagon bias based model for opinion dynamics: Intertwining between homophily and influence mechanisms. *European Journal of Control*, *68*, file 100675.

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

Disarò, G., & Valcher, M. (2023). On an extension of the Friedkin-Johnsen model: The effects of a homophily-based influence matrix. In *Proceedings of the 62nd IEEE Conference on Decision and Control* (pp. 3384–3389). Singapore.

Franci, A., Golubitsky, M., Stewart, I., Bizyaeva, A., & Leonard, N. (2023). Breaking indecision in multiagent, multioption dynamics. *SIAM Journal on Applied Dynamical Systems*, *22*(3), 1780–1817.

[3] It is easy to see that if A, B and C are nonnegative matrices, then $\text{sgn}(ABC) = \text{sgn}(\text{sgn}(A)\text{sgn}(B)\text{sgn}(C))$.

Friedkin, N. (2015). The problem of social control and coordination of complex systems in sociology: A look at the community cleavage problem. *IEEE Control Systems Magazine*, *35*(3), 40–51.
Friedkin, N., & Johnsen, E. (1999). Social influence networks and opinion change. *Advances in Group Processes*, *16*, 1–29.
Gantmacher, F. (1960). *The theory of matrices*. Providence, Rhode Island: AMS Chelsea Pub..
He, G., Ci, Z., Wu, X., & Hu, M. (2022). Opinion dynamics with antagonistic relationship and multiple interdependent topics. *IEEE Access*, *10*, 31595–31606.
Hegselmann, R., & Krause, U. (2002). Opinion dynamics and bounded confidence: models, analysis and simulation. *Journal of the Artificial Societies and Social Simulation*, *5*, 1–24.
Hiller, T. (2017). Friends and enemies: a model of signed network formation. *Theoretical Economics*, *3*(12), 1057–1087.
Horn, R., & Johnson, C. (1985). *Matrix analysis*. Cambridge (GB): Cambridge Univ. Press.
Lazarsfeld, P., & Merton, R. (1954). Friendship as a social process: A substantive and methodological analysis. *Freedom and Control in Modern Society*, 18–66.
Leonard, N., Lipsitz, K., Bizyaeva, A., Franci, A., & Lelkes, Y. (2021). The nonlinear feedback dynamics of asymmetric political polarization. *Proceedings of the National Academy of Sciences*, *118*(50), Article e2102149118.
Liu, F., Cui, S., Mei, W., Dörfler, F., & Buss, M. (2020). Interplay between homophily-based appraisal dynamics and influence-based opinion dynamics: Modeling and analysis. *IEEE Control Systems Letters*, *5*(1), 181–186.
McPherson, M., Smith-Lovin, L., & Cook, J. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, *27*(1), file 415444.
Mei, W., Bullo, F., Chen, G., Hendrickx, J., & Dörfler, F. (2022). Micro-foundation of opinion dynamics: Rich consequences of the weighted-median mechanism. *Physical Review Research*, *4*, Article 023213.
Mei, W., Cisneros-Velarde, P., Chen, G., Friedkin, N., & Bullo, F. (2019). Dynamic social balance and convergent appraisals via homophily and influence mechanisms. *Automatica*, *110*, 61–67.
Pan, L., Shao, H., & Li, D. (2021). Peer selection in opinion dynamics on signed social networks with stubborn individuals. *Neurocomputing*, *477*, 104–113.
Parsegov, S., Proskurnikov, A., Tempo, R., & Friedkin, N. (2015). A new model of opinion dynamics for social actors with multiple interdependent attitudes and prejudices. In *Proceedings of the 54th IEEE Conference on Decision and Control* (pp. 3475–3480). Osaka, Japan.
Parsegov, S., Proskurnikov, A., Tempo, R., & Friedkin, N. (2017). Novel multidimensional models of opinion dynamics in social networks. *IEEE Transactions on Automatic Control*, *62*(5), 2270–2285.
Proskurnikov, A., & Tempo, R. (2017). A tutorial on modeling and analysis of dynamic social networks. Part I. *Annual Reviews in Control*, *43*, 65–79.
Proskurnikov, A., & Tempo, R. (2018). A tutorial on modeling and analysis of dynamic social networks. Part II. *Annual Reviews in Control*, *45*, 166–190.
Proskurnikov, A., Tempo, R., Cao, M., & Friedkin, N. E. (2017). Opinion evolution in time-varying social influence networks with prejudiced agents. *IFAC PapersOnLine*, *50*(1), 11896–11901.
Rivera, M., Soderstrom, S., & Uzzi, B. (2010). Dynamics of dyads in social networks: Assortative, relational, and proximity mechanisms. *Annual Review of Sociology*, *36*, file 91115.
Santos, F., Lelkes, Y., & Levin, S. (2021). Link recommendation algorithms and dynamics of polarization in online social networks. *Proceedings of the National Academy of Sciences*, *118*(50), Article e2102141118.
Xia, W., Cao, M., & Johansson, K. (2016). Structural balance and opinion separation in trust-mistrust social networks. *IEEE Transactions on Control of Network Systems*, *3*, 46–56.

Giorgia Disarò received the B.Sc. degree in Information Engineering and the M.Sc. degree in Control Systems Engineering, both from the University of Padova. She is currently a PhD student at the Department of Information Engineering in Padova. Her research interests focus on multiagent and networked systems, with applications in opinion dynamics and social networks. She also works on data-driven methods for state estimation in the presence of unknown inputs or disturbances, applied both to centralized and distributed scenarios.

Maria Elena Valcher received the Master degree (1991) and the PhD degree (1995) from the University of Padova. Since 2005 she is full professor at the University of Padova. She is author/co-author of 93 journal papers, 113 conference papers, 2 text-books and several book chapters. Her research interests include social networks, cooperative control and consensus, positive switched systems and Boolean control networks. She is the Founding Editor in Chief of the IEEE Control Systems Letters (2017-). She was IEEE CSS President (2015). She received the 2011 IEEE CSS Distinguished Member Award and she is an IEEE Fellow since 2012. She is an IFAC Fellow (2020–2023) and the 2024–2025 EUCA President.