UC Santa Barbara

Title

Opinion evolution in time-varying social influence networks with prejudiced agents

Authors

Proskurnikov, Anton V
Tempo, Roberto
Cao, Ming
et al.

Opinion evolution in time-varying social influence networks with prejudiced agents

Anton V. Proskurnikov [1,2,3] **Roberto Tempo** [4] **Ming Cao** [5]
Noah E. Friedkin [6]

[1] *Delft Center for Systems and Control (DCSC), Delft University of Technology, The Netherlands*
[2] *Institute for Problems of Mechanical Engineering (IPME RAS), St. Petersburg, Russia*
[3] *ITMO University, St. Petersburg, Russia*
[4] *CNR-IEIIT, Politecnico di Torino, Torino, Italy*
[5] *Engineering and Technology Institute (ENTEG), University of Groningen, The Netherlands*
[6] *Center for Control, Dynamical Systems and Computation, University of California Santa Barbara, Santa Barbara, USA*

Abstract: Investigation of social influence dynamics requires mathematical models that are "simple" enough to admit rigorous analysis, and yet sufficiently "rich" to capture salient features of social groups. Thus, the mechanism of iterative opinion pooling from (DeGroot, 1974), which can explain the generation of consensus, was elaborated in (Friedkin and Johnsen, 1999) to take into account individuals' ongoing attachments to their initial opinions, or *prejudices*. The "anchorage" of individuals to their prejudices may disable reaching consensus and cause disagreement in a social influence network. Further elaboration of this model may be achieved by relaxing its restrictive assumption of a time-invariant influence network. During opinion dynamics on an issue, arcs of interpersonal influence may be added or subtracted from the network, and the influence weights assigned by an individual to his/her neighbors may alter. In this paper, we establish new important properties of the (Friedkin and Johnsen, 1999) opinion formation model, and also examine its extension to time-varying social influence networks.

1. INTRODUCTION

During the past decades, there has been a substantial growth of interest in dynamics of social influence networks and opinion formation mechanisms in them. In contrast to the recent research emphasis on multi-agent consensus and coordination, models are being advanced that explain observed behaviors of social groups such as disagreement, polarization, and conflict (Friedkin, 2015; Proskurnikov and Tempo, 2017). An explanatory network science is advancing on the structural properties of social networks (Wasserman and Faust, 1994; Easley and Kleinberg, 2010) and some special dynamical processes over these networks, e.g. epidemic spread (Newman, 2003). At the same time, there is a growing recognition that systems and control theories may substantially broaden the scope of our understanding of the definitional problem of sociology—the coordination and control of social systems (Friedkin, 2015).

System-theoretic examination of social dynamics requires mathematical models that are capable of capturing the complex behavior of a social group yet simple enough to be rigorously examined. In this paper, we deal with one such model, proposed by Friedkin and Johnsen (Friedkin and Johnsen, 1999, 2011; Friedkin, 2015) and henceforth referred to as the FJ model. The FJ model extends the idea of iterative "opinion pooling" (DeGroot, 1974) by assuming that some agents are *prejudiced*. These agents have some level of "anchorage" on their initial opinions (*prejudices*) and factor them into any iteration of their opinions. Similar to continuous-time clustering protocols with "informed" leaders (Xia and Cao, 2011), the heterogeneity of the prejudices and its linkage to individuals' susceptibilities to interpersonal influence may lead to persistent disagreement of opinions and outcomes such as polarization and clustering. With the FJ model, the clustering of opinions does not require the existence of repulsive couplings, or "negative ties" among individuals (Fläche and Macy, 2011; Altafini, 2013; Proskurnikov et al., 2016a; Xia et al., 2016) whose ubiquity in interpersonal interactions is still waiting for supporting experimental evidence (Takács et al., 2016). Unlike models with discrete opinions (Castellano et al., 2009) and bounded confidence models (Hegselmann and Krause, 2002; Weisbuch et al., 2005; Blondel et al., 2009), the FJ model describes the opinion evolution by *linear* discrete-time equations, and is thus much simpler for mathematical analysis. At the same time, the FJ model has been confirmed by experiments with real social groups (Friedkin and Johnsen, 2011; Friedkin et al., 2016a). The FJ model is closely

related to the PageRank algorithm (Friedkin and Johnsen, 2014; Proskurnikov et al., 2016b) and has been given some elegant game-theoretic and electric interpretations (Bindel et al., 2011; Ghaderi and Srikant, 2014; Frasca et al., 2015). In the recent works (Parsegov et al., 2017; Proskurnikov and Tempo, 2017) necessary and sufficient conditions for the stability of the FJ model has been established; these conditions also provide convergence "on average" of its decentralized gossip-based counterpart (Frasca et al., 2013; Ravazzi et al., 2015; Frasca et al., 2015). A multidimensional extension of the FJ model has been used to describe the evolution of *belief systems* (Parsegov et al., 2017; Friedkin et al., 2016b), representing invidiuals' positions on several *mutually dependent* issues.

In this paper, we further develop the mathematical theory of the FJ model, obtaining explicit estimates for its convergence speed. We also examine an extension of the classical FJ model, describing a natural *time-varying* social influence process. Such an extension is important since during opinion dynamics on an issue, arcs of interpersonal influence may be added or subtracted from the network, and the influence weights assigned by an individual to his/her neighbors may alter. An example of such an evolution is the dynamics of individuals' *reflected appraisals* (Jia et al., 2015; Friedkin et al., 2016a; Chen et al., 2016).

2. PRELIMINARIES AND NOTATION

We denote matrices with capital letters $A = (a_{ij})$, using lower case letters for their scalar entries and vectors. The symbol $\mathbf{1}_n$ denotes the column vector of ones $(1, 1, \ldots, 1)^\top \in \mathbb{R}^n$, and I_n is the identity $n \times n$ matrix. For two vectors $x, y \in \mathbb{R}^n$ we write $x \le y$ if $x_i \le y_i \forall i$. The spectral radius of a square matrix A is denoted by $\rho(A)$, the matrix is *Schur* stable if $\rho(A) < 1$. A nonnegative matrix A is *substochastic* if $\sum_j a_{ij} \le 1$ for any i. Any such matrix has $\rho(A) \le 1$ due to the Gershgorin disk theorem (Horn and Johnson, 1985). A substochastic matrix A is *stochastic* if $\sum_j a_{ij} = 1 \forall i$; when A is sized $n \times n$, the stochasticity implies that $A\mathbf{1}_n = \mathbf{1}_n$ and $\rho(A) = 1$.

A (weighted directed) *graph* is a triple $\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$, where $\mathcal{V} = \{v_1, \ldots, v_n\}$ stands for the set of *nodes*, $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ is the set of *arcs*, and W is a (weighted) $n \times n$ *adjacency matrix*, i.e. $w_{ij} > 0$ when $(i, j) \in \mathcal{E}$ and otherwise $w_{ij} = 0$. Henceforth we assume that $\mathcal{V} = \{1, 2, \ldots, n\}$ and thus the graph $\mathcal{G} = \mathcal{G}(W)$ is uniquely defined by its adjacency matrix W. We denote an arc $(i, j) \in \mathcal{E}$ by $i \mapsto j$ and call the value w_{ij} its *weight*. A chain of arcs $i_0 \mapsto i_1 \mapsto \ldots \mapsto i_{r-1} \mapsto i_r$ is a *walk* of length r from node i_0 to node i_r.

3. THE FRIEDKIN-JOHNSEN MODEL

The FJ model describes a network of *social influence* (Friedkin and Johnsen, 2011), consisting of n individuals, or social *agents* indexed 1 through n. The agents opinions are represented by scalars $x_i \in \mathbb{R}$, constituting the vector of opinions $x = (x_1, \ldots, x_n)^\top$. The process of social influence is described by two matrices: a stochastic matrix of *interpersonal influences* $W \in \mathbb{R}^{n \times n}$ and a diagonal matrix $\Lambda = \text{diag}(\lambda_{11}, \ldots, \lambda_{nn})$ of individual *susceptibilities* $\lambda_{ii} \in [0; 1]$ to the interpersonal influence. At each step, the vector of opinions changes as follows

$$x(k + 1) = \Lambda W x(k) + (I_n - \Lambda)u, \quad k = 0, 1, \ldots. \quad (1)$$

The elements u_i of the constant vector u stand for the agents' *prejudices*; the original FJ model (Friedkin and Johnsen, 1999; Friedkin, 2015) assumed that $u_i = x_i(0)$.

In the special case where $\Lambda = I_n$ the model (1) reduces to DeGroot's iterative "opinion pooling" (DeGroot, 1974), providing a discrete-time *consensus algorithm* (Ren and Beard, 2008). At each step, an agent sets its new opinion to be the convex combination of its own and others' opinions

$$x_i(k + 1) = \sum_{j=1}^n w_{ij} x_j(k) \forall i \iff x(k + 1) = W x(k). \quad (2)$$

The influence weight w_{ij} shows the contribution of jth opinion on each stage to the ith opinion on the next stage.

The FJ model (1) also employs the mechanism of convex combination, allowing some agents to be prejudiced. If $\lambda_{ii} < 1$ then agent i is "attached" to its prejudice u_i and factors it into any opinion iteration, replacing (2) by

$$x_i(k + 1) = \lambda_{ii} \sum_{j=1}^n w_{ij} x_j(k) + (1 - \lambda_{ii})u_i \forall i. \quad (3)$$

When $\lambda_{ii} = 1$, the ith agent's opinion is formed by the DeGroot mechanism (2), otherwise its prejudice influences each stage of the opinion iteration. Agent i with $\lambda_{ii} = 0$ is "totally prejudiced" and its opinion is static $x_i(k) \equiv u_i$.

Under the assumption $u_i = x_i(0)$, adopted in the FJ model, any agent with $w_{ii} = 1$ (and thus $w_{ij} = 0 \forall j \ne i$) retains its opinion constant $x_i(k) = u_i$ independent of λ_{ii}, and one may suppose, without loss of generality, that

$$w_{ii} = 1 \iff \lambda_{ii} = 0. \quad (4)$$

In the original model from (Friedkin and Johnsen, 1999) an even stronger *coupling condition* $w_{ii} = 1 - \lambda_{ii} \forall i$ was adopted for parsimony in the model's empirical applications. In this paper, we do not assume this condition to hold, so Λ and W are independent except for the non-degeneracy condition (4). Notice that each FJ model corresponds to the substochastic matrix $A = \Lambda W$; for the models satisfying (4) this correspondence is one-to-one. A substochastic matrix A is decomposed as $A = \Lambda W$, where

$$\lambda_{ii} = \sum_j a_{ij} \quad \text{and} \quad w_{ij} = \begin{cases} a_{ij}/\lambda_{ii}, & \lambda_{ii} > 0, \\ 1, & i = j \text{ and } \lambda_{ii} = 0, \\ 0, & i \ne j \text{ and } \lambda_{ii} = 0. \end{cases}$$

The stability criteria for FJ models may thus be reformulated for substochastic matrices, and vice versa.

For us it will be convenient to discard the standard assumption $x(0) = u$ and consider u as some constant external "input", independent of the initial opinion [1] $x(0)$.

A central question concerned with the FJ dynamics (1) is its *convergence* of opinion vectors to a finite limit

$$x^\infty = \lim_{k \to \infty} x(k). \quad (5)$$

---

[1] Individuals prejudices may be explained (Friedkin and Johnsen, 1999) by the system "history", e.g. the effect of some exogenous factors, which influenced the community in the past. This motivates to introduce the explicit relation between the prejudice and initial condition of the social system $u = x(0)$. However, the prejudices can also be some non-trivial functions of the initial conditions $u = u(x(0))$ or be caused by external factors that are not related to the system's history, e.g. some information spread in social media.

A sufficient condition for convergence is the *Schur stability*: if $\rho(\Lambda W) < 1$ then the opinions converge to

$$x^\infty = Vu, \quad V = (I - \Lambda W)^{-1}(I - \Lambda). \quad (6)$$

It is known (Friedkin, 2015) that for any Schur stable FJ model the matrix V is stochastic and, obviously, x^∞ from (6) is the only equilibrium of the system (1). Generally, the Schur stability is not necessary for convergence, e.g. the DeGroot model (2) is never stable but converges when e.g. W is *primitive* (i.e. irreducible and aperiodic) (DeGroot, 1974; Gantmacher, 2000).

Henceforth we are primarily interested in Schur stable FJ models, where the steady opinion is unique and given by (6). The Schur stability is a "generic" condition if at least one prejudiced agent exists, and holds, for instance, for a strongly connected influence networks [2], as implied by the following lemma (Parsegov et al., 2017).

Lemma 1. The matrix ΛW is Schur stable if and only if each node in the graph $\mathcal{G}(W)$ either belongs to the set

$$S_\Lambda = \{i : \lambda_{ii} < 1\} \quad (7)$$

or connected to a node from S_Λ by a walk, i.e. any agent is either prejudiced or influenced by a prejudiced individual.

4. SCHUR STABLE FJ MODELS: OPINION CLUSTERING AND CONVERGENCE SPEED

In this section, we derive some advanced properties of Schur stable FJ models (1), satisfying the condition from Lemma 1. We answer the following two questions, related to such models' dynamics:

- do the final opinions x^∞ reach consensus or disagree?
- what is the convergence speed in (5)?

4.1 Consensus and disagreement in the FJ model

One can expect that for a general FJ model the consensus of the steady opinions $x_1^\infty, \ldots, x_n^\infty$ typically should disagree, whereas their consensus is an exceptional situation. This is confirmed by the following consensus criterion.

Theorem 2. Let the FJ model (1) be stable. Then the consensus of final opinions $x_1^\infty = \ldots = x_n^\infty$ is reached if and only if $u_i = u_0$ for some $u_0 \in \mathbb{R}$ and any $i \in S_\Lambda$. In this case $x_i^\infty = u_0 \, \forall i$. This holds e.g. when $S_\Lambda = \{i\}$ has only one element, i.e. only one agent is prejudiced.

The values u_i, where $i \notin S_\Lambda$ and thus $1 - \lambda_{ii} = 0$, obviously do not influence the value of x^∞ and may be arbitrary. Note that when consensus is not established, the number of "clusters" in the vectors u and x^∞ do not correlate. For instance, if $u_1 = 1$ and $u_i = 0 \, \forall i > 1$ then x^∞ is the first row of V and its elements are usually all different.

4.2 Convergence speed of the FJ model

In this subsection we give an explicit estimate of the spectral radius $\rho = \rho(\Lambda W)$, which also determines the convergence speed in (5): $|x(k) - x_*| = O(\rho^k)$ as $k \to \infty$. We start with introducing some definitions and notation.

---

[2] This property can be also reformulated as follows: an irreducible substochastic matrix is Schur stable (Meyer, 2000, Exercise 8.3.7).

Definition 3. An arc $i \mapsto j$ in $\mathcal{G}(W)$ with the weight $w_{ij} \geq \varepsilon$ is said to be an ε-arc. An ε-walk in the graph is a walk constituted by ε-arcs. Given a set $S \subseteq \{1, \ldots, n\}$ and node i, let $d_\varepsilon(i, S, W)$ stand for the length of the *shortest* ε-walk from i to S. By definition, $d_\varepsilon(i, S, W) = 0$ for any $i \in S$ and $d_\varepsilon(i, S, W) = \infty$ if no ε-walk from i to S exists.

For any diagonal matrix $\Lambda = \mathrm{diag}(\lambda_{11}, \ldots, \lambda_{nn})$ with $0 \leq \Lambda \leq I_n$ and $\delta > 0$, we introduce the set of indices

$$S_\Lambda^\delta = \{i : \lambda_{ii} \leq 1 - \delta\} \subseteq S_\Lambda. \quad (8)$$

Definition 4. The FJ model (or the pair (Λ, W)) belongs to the class $\mathbf{FJ}[\delta, \varepsilon, s]$ if $d_\varepsilon(i, S_\Lambda^\delta, W) \leq s$ for any node $i = 1, \ldots, n$. Here $\delta, \varepsilon > 0$ are real and $s \geq 0$ is an integer.

Any FJ model, belonging to $\mathbf{FJ}[\delta, \varepsilon, s]$ with $\delta, \varepsilon > 0$, is Schur stable due to Lemma 1. On the other hand, any Schur stable FJ model belongs to $\mathbf{FJ}[\delta_0, \varepsilon_0, n-1]$, where

$$\varepsilon_0 \triangleq \min\{w_{ij} : w_{ij} > 0\}, \quad \delta_0 \triangleq 1 - \max_{i \in S_\Lambda} \lambda_{ii}, \quad (9)$$

since $S_\Lambda^{\delta_0} = S_\Lambda$ and any walk in $\mathcal{G}(W)$ is an ε_0-walk.

The following theorem gives an explicit estimate for the spectral radius $\rho(\Lambda W)$ of a Schur stable FJ model (1).

Theorem 5. For any FJ model (1) from the class $\mathbf{FJ}[\delta, \varepsilon, s]$ one has $\rho(\Lambda W) \leq \rho_*(\delta, \varepsilon, s) \triangleq [1 - \delta \varepsilon^s]^{1/(1+s)}$.

For the case of undirected graph $\mathcal{G}(W)$ and special influence weights of arcs a similar estimate for the convergence speed has been obtained in (Ghaderi and Srikant, 2014). Unlike this paper, Theorem 5 deals with a general FJ model, where the matrix W can be arbitrary.

Corollary 6. For a stable FJ model (1), let ε_0, δ_0 be defined by (9). Then $\rho(\Lambda W) \leq \left[1 - \delta_0 \varepsilon_0^{n-1}\right]^{1/n}$.

Although the estimate from Theorem 5 is just an upper bound for $\rho(\Lambda W)$, this bound proves to be tight for special types of graphs. For instance, if $\Lambda = (1 - \delta)I_n$ then $S_\Lambda^\delta = \{1, \ldots, n\}$, $s = 0$ and hence $\rho(\Lambda W) = (1 - \delta)$ for any W. Another example is the cycle graph $1 \mapsto 2 \mapsto 3 \mapsto \ldots \mapsto n \mapsto 1$, where the weights of arcs are equal to $\varepsilon = 1$. If $\lambda_{11} = 1 - \delta$ and $\lambda_{22} = \ldots = \lambda_{nn} = 1$ then $s = n - 1$ and

$$\Lambda W = \begin{bmatrix} 0 \cdots 0 & 1-\delta \\ & 0 \\ I_{n-1} & \vdots \\ & 0 \end{bmatrix}, \; \rho(\Lambda W) = (1-\delta)^{1/n}.$$

5. TIME-VARYING FJ MODEL

A principal limitation of the standard FJ model (1) is the *time invariance* of social influence: the matrices Λ and W remain constant. In real social groups the structures of social influence may evolve over time as the interpersonal ties may emerge and disappear; even if their graph remains constant, the influence weights w_{ij} and susceptibilities λ_{ii} may change. One of the models, describing the evolution of the matrix W, is the dynamics of *reflected appraisals* (Jia et al., 2015; Chen et al., 2016; Friedkin et al., 2016a), where the self-confidence of a person depends on how he/she is evaluated by the others. In this section we consider a time-varying extension of the FJ model and study its properties.

The time-varying FJ model (TVFJ) is as follows

$$x(k+1) = \Lambda(k)W(k)x(k) + [I_n - \Lambda(k)]u. \qquad (10)$$

We assume that the matrices $\Lambda(k), W(k)$ on each stage of the opinion evolution are known and have the same structure, as for the classical model (1), i.e. $\Lambda(k)$ is diagonal, $0 \le \Lambda(k) \le I_n$ and $W(k)$ is stochastic. Given the initial condition $x(0) = x^0$ and the prejudice vector u, let $x(k|x^0, u)$ stand for the solution of (10). The averaging mechanism of (10) provides several useful properties.

Lemma 7. Any model (10) has the following properties:

(1) if $x^0 = u = u_* \mathbf{1}_n$, then $x(k|x^0, u) = u_* \mathbf{1}_n \, \forall k$;
(2) if $x^0, u \in [m, M]^n$, then $x(k|x^0, u) \in [m, M]^n \, \forall k$;
(3) more generally, if $x^0 \le x^1$ and $u \le u^1$, then $x(k|x^0, u) \le x(k|x^1, u^1) \, \forall k$;
(4) for any "perturbations" $\delta x^0, \delta u \in [m, M]^n$ one has $x(k|x^0 + \delta x^0, u + \delta u) - x(k|x^0, u) \in [m, M]^n$.

Here m, M, u_* stand for some real scalars.

Applied for $M = -m = \varepsilon$, statement 4) in Lemma 7 implies *robustness* of the trajectories against small perturbations in x^0 and u. Note that this property *does not depend* on the asymptotical (Schur) stability of the system (10). For a general neutrally stable system, such a robustness does not hold as illustrated by the simplest counterexample $x(k+1) = x(k) + u$.

Henceforth we are primarily interested in *asymptotically stable* TVFJ models, which means, as usual, that $x(k|x^0, 0) \xrightarrow[k \to \infty]{} 0$ for any initial condition x^0, i.e.

$$\Lambda(k)W(k)\Lambda(k-1)W(k-1)\dots\Lambda(0)W(0) \xrightarrow[k \to \infty]{} 0. \quad (11)$$

Unlike the stationary case, the asymptotical stability in general *does not* imply the convergence (5). For instance, let two stationary Schur stable FJ models with matrices (Λ_1, W_1) and (Λ_2, W_2) corresponding to different matrices V_1, V_2 (defined by (6)). Due to (6), when $(\Lambda(k), W(k))$ switches between (Λ_1, W_1) and (Λ_2, W_2) with sufficiently large dwell time, $x(k)$ oscillates between $V_1 u$ and $V_2 u$. Nevertheless, two "relaxed" versions of the convergence condition remain valid for asymptotically stable models (10).

Lemma 8. The following conditions are equivalent:

(1) *(stability)* the system (10) is asymptotically stable;
(2) *(containment)* for any $x(0), u \in \mathbb{R}^n$ one has
$$\min_j u_j \le \liminf_{k \to \infty} x_i(k) \le \limsup_{k \to \infty} x_i(k) \le \max_j u_j;$$
(3) *(consensus)* if $u = u_* \mathbf{1}_n$, then $x(k) \xrightarrow[k \to \infty]{} u \, \forall x(0)$.

Lemma 8 establishes an important relation between the TVFJ model and algorithms of multi-agent control, namely, protocols for *leader-following* consensus (Ren and Beard, 2008) and *containment control* (Ren and Cao, 2011). Adding a "virtual" agent $n + 1$ whose opinion is static $x_{n+1}(k) \equiv x_{n+1}(0)$ and the "augmented" opinion vector $\hat{x}(k) = (x_1(k), \dots, x_n(k), x_{n+1}(k))$, the system (10) with $u = x_{n+1}(0)\mathbf{1}_n$ can be rewritten as follows

$$\hat{x}(k+1) = \hat{A}(k)\hat{x}(k), \quad \hat{A}(k) = \begin{bmatrix} \Lambda(k)W(k) & (\mathbf{1}_n - \Lambda(k)\mathbf{1}_n) \\ 0_{1 \times n} & 1 \end{bmatrix}. \tag{12}$$

Lemma 8 states that stability of the model (10) is equivalent to establishing consensus in (12) $x_i(k) \xrightarrow[k \to \infty]{}$

$x_{n+1}(0) \, \forall i = 1, \dots, n$ for any initial condition $\hat{x}(0)$ ($i = 1, \dots, n$). This implies the following stability condition.

Lemma 9. Suppose that $\varepsilon > 0$ exists such that the matrix $\hat{A}(k) = (\hat{a}_{ij}(k))$ at any time $k \ge 0$ satisfies the conditions $\hat{a}_{ij}(k) \in \{0\} \cup [\varepsilon, 1]$ for any i, j and $\hat{a}_{ii}(k) \ge \varepsilon$ for any i. Then the model (10) is stable if a period $T \ge 1$ exists such that in the graph $\mathcal{G}[\hat{A}(k) + \dots + \hat{A}(k + T - 1)]$ each node is connected to node $n + 1$ by a walk. This holds e.g. if the condition from Lemma 1 is valid at any time.

Proof. Thanks to the standard consensus criterion for time-varying directed graphs (Blondel et al., 2005; Ren and Beard, 2008), the assumption of Lemma 9 entail consensus in the augmented network (12), which, in turn, is equivalent to stability of the model (10) due to Lemma 8.

The assumptions of Lemma 9, typically adopted to prove the convergence of multi-agent coordination algorithms (Ren and Beard, 2008; Ren and Cao, 2011), are however very restrictive for networks of social influence. Lemma 9, in particular, is not applicable to TVFJ models where some agents have zero levels of self-confidence $w_{ii} = 0$ or "totally prejudiced" $\lambda_{ii} = 0$. Unlike multi-agent *control* algorithms that are usually designed to have uniformly positive influence weights, such a positivity condition cannot be guaranteed for opinion dynamics. In particular, the process of reflected appraisal (Friedkin et al., 2016a) often leads to the situation where some self-confidence weights w_{ii} asymptotically vanish.

The following two counterexamples demonstrate that in presence of agents with $\lambda_{ii}(k)w_{ii}(k) = 0$ Lemma 9 is not valid, in particular, Schur stability of any matrix $\Lambda(k)W(k)$ does not imply the stability of the model (10). We start with two simple counterexamples: in one of them, the matrix Λ is fixed while W is switching, in the other one the matrix W is fixed and Λ switching.

Example 1. Consider $n = 3$ agents with $\Lambda(k) \equiv \mathrm{diag}(0, 1, 1)$ and let the matrix $W(k)$ switch as follows

$$W(2m) = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \ W(2m+1) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}, \quad m = 0, 1, \dots.$$

The dynamics (10) implies that $x_1(k) \equiv u_1$ and

$$(x_2(k+1), x_3(k+1)) = \begin{cases} (u_1, x_2(k)), & k = 2m, \\ (x_3(k), u_1), & k = 2m+1. \end{cases}$$

Therefore, we have $x_2(0) = x_2(2) = \dots = x_2(2m) \, \forall m$ and $x_2(k) \not\to 0$ as $k \to \infty$ when $u_1 = 0$ and $x_2(0) \ne 0$.

Example 2. Consider the TVFJ model with $n = 2$ and

$$W(k) \equiv \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \Lambda(2m) = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \Lambda(2m+1) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

The dynamics (10) can then be rewritten as follows

$$(x_1(k+1), x_2(k+1)) = \begin{cases} (u_1, x_2(k)), & k = 2m, \\ (x_1(k), u_2), & k = 2m+1, \end{cases}$$

entailing that $x_1(0) = x_1(2) = \dots = x_1(2m) \, \forall m$, and thus $x_1(k) \not\to 0$ as $k \to \infty$ when $u = 0$ and $x_1(0) \ne 0$.

In Examples 1 and 2 the switching model (10) appears to be not asymptotically stable in spite of the Schur stability of the two possible values $\Lambda(k)W(k)$: the joint spectral radius (Lin and Antsaklis, 2009) of these matrices equals to 1. This critical situation, where the results of classical switching systems theory (Lin and Antsaklis, 2009) are not

applicable, is typical for the TVFJ model. To guarantee its stability, special criteria are needed; one of such criteria, extending Theorem 5, is offered in this section.

We start with introducing a class $\mathbf{CFJ}[\delta, \varepsilon, s]$, where $\delta, \varepsilon > 0$ are real and $s \geq 0$ is an integer (acronym CFJ stands for "Chain of FJ models"). Unlike $\mathbf{FJ}[\delta, \varepsilon, s]$, constituted by pairs (Λ, W), the class $\mathbf{CFJ}[\delta, \varepsilon, s]$ consists of sequences $\{(\Lambda^{(k)}, W^{(k)})\}_{k=0}^s$. For such a sequence and $\delta, \varepsilon > 0$, we introduce the sets
$J_0 \triangleq S_{\Lambda^{(0)}}^\delta$, $J_k \triangleq S_{\Lambda^{(k)}}^\delta \cup \{i : w_{ij}^{(k)} \geq \varepsilon \text{ for some } j \in J_{k-1}\}$.
When $\Lambda^{(k)} = \Lambda$ and $W^{(k)} = W$ for any $k = 0, \ldots, s$, the set J_j contains all such indices i that $d_\varepsilon(i, S_\Lambda^\delta, W) \leq j$.

Definition 10. The class $\mathbf{CFJ}[\delta, \varepsilon, s]$ consists of all sequences $\{(\Lambda^{(k)}, W^{(k)})\}_{k=0}^s$ such that $J_s = \{1, 2, \ldots, n\}$.

The following result is proved similarly to Theorem 5.

Lemma 11. For any sequence $\{(\Lambda^{(k)}, W^{(k)})\}_{k=0}^s$ from the set $\mathbf{CFJ}[\delta, \varepsilon, s]$ the matrix $\mathcal{P} \triangleq \prod_{k=0}^s \Lambda^{(k)} W^{(k)}$ has row sums $\leq 1 - \delta \varepsilon^s$, that is, $\mathcal{P} \mathbf{1}_n \leq (1 - \delta \varepsilon^s) \mathbf{1}_n$.

Using Lemma 11, the following sufficient condition for asymptotic stability is immediate.

Theorem 12. Let real $\delta, \varepsilon > 0$ and an integer $s \geq 0$ exist such that the sequence $\{(\Lambda(k), W(k))\}_{k=0}^\infty$ contains infinitely many subsequences $\{(\Lambda(k), W(k))\}_{k=m}^{m+s}$ from $\mathbf{CFJ}[\delta, \varepsilon, s]$. Then the model (10) is asymptotically stable.

The condition of Theorem 12 can, evidently, be reformulated as follows: any infinite "tail" $\{(\Lambda(k), W(k))\}_{k=r}^\infty$ (where $r \geq 1$) contains a subsequence from $\mathbf{CFJ}[\delta, \varepsilon, s]$. This condition does not require stability of any matrix $\Lambda(k)W(k)$ and allows e.g. to have $\Lambda(k) = I_n$ for some k.

6. PROOFS

In this section, we prove our main results.

6.1 Proof of Theorem 2

We start with the sufficiency part. Suppose that $u_i = u_0 \, \forall i \in S_\Lambda$. One may assume that $u_i = u_0 \, \forall i$ since for $i \notin S_\Lambda$ the value of u_i has no effect on x^∞. Since V is row-stochastic (Friedkin, 2015), $x^\infty = V(u_0 \mathbf{1}_n) = u_0 \mathbf{1}_n$, which proves consensus. To prove necessity, assume that $x^\infty = u_0 \mathbf{1}_n$ for some $u_0 \in \mathbb{R}$. Using (6), $(I_n - \Lambda)u = u_0(I_n - \Lambda W)\mathbf{1}_n = u_0(I_n - \Lambda)\mathbf{1}_n$. In view of (7), $u_i = u_0 \, \forall i \in S_\Lambda$. ∎

6.2 Proofs of Theorems 5,12, Lemma 11 and Corollary 6

We start with a useful technical lemma. Given a sub-stochastic matrix A, the number $\hat{a}_i = 1 - \sum_j a_{ij} \geq 0$ is said to be the *deficiency* of the ith row. From the Gershgorin disk theorem (Horn and Johnson, 1985) it is obvious that $\rho(A) \leq \max\limits_{1 \leq i \leq n} [1 - \hat{a}_i]$.

Lemma 13. Let A and B be substochastic $n \times n$ matrices, $C = AB$ and $\hat{a}_i, \hat{b}_i, \hat{c}_i$ stand for the respective deficiencies. Then the following statements hold for any $i = 1, \ldots, n$

$$\hat{c}_i = \hat{a}_i + \sum_{j=1}^n a_{ij} \hat{b}_j. \tag{13}$$

Proof. Denote the ith row of B and C with respectively $b_{i\bullet}$ and $c_{i\bullet}$, we have $c_{i\bullet} \mathbf{1}_n = \sum_j a_{ij} b_{i\bullet} \mathbf{1}_n = \sum_j a_{ij}(1 - \hat{b}_j) = 1 - \hat{a}_i - \sum_j a_{ij} \hat{b}_j$, which entails (13).

Proof of Theorem 5. For brevity, we denote $A \triangleq \Lambda W$ and put $d(i) \triangleq d_\varepsilon(i, S_\Lambda^\delta, W)$. We are going to prove the following statement via induction on m: if $d(i) \leq m$ and $C = A^{1+m}$, then $\hat{c}_i \geq \delta \varepsilon^{d(i)}$. For $m = 0$ the claim is obvious: when $d(i) = 0$ one has $i \in S_\Lambda^\delta$ and hence $\hat{c}_i = \hat{a}_i \geq \delta$. Assuming that the statement has been proved for $m-1 \geq 0$, we have to prove it for m. Denoting $B = A^m$, one has $C = BA$. If $d(i) \leq m - 1$ then $\hat{c}_i \geq \hat{b}_i \geq \delta \varepsilon^{d(i)}$ thanks to (13). If $d(i) = m$, there exists j such that $d(j) \leq m - 1$ and $w_{ij} \geq \varepsilon$. Denoting $C' = WB$, (13) implies that $\hat{c}_i' \geq \varepsilon \hat{b}_j \geq \delta \varepsilon^m$. Since $C = \Lambda C'$, we have $\hat{c}_i \geq \hat{c}_i'$ which proves our statement for m. Substitution $m = s$ yields $\rho(A^{1+s}) = \rho(A)^{1+s} \leq 1 - \delta \varepsilon^s$ by definition of $\mathbf{FJ}(\delta, \varepsilon, s)$. ∎

Corollary 6 is immediate from Theorem 5 since a stable FJ model belongs to $\mathbf{FJ}[\delta_0, \varepsilon_0, n-1]$. ∎

Proof of Lemma 11. Similarly to proof of Theorem 5, one proves via induction on $m = 0, 1, \ldots, s$ that for any $i \in J_m$ the ith row of the matrix $\mathcal{P}_m \triangleq \prod_{k=0}^m \Lambda^{(k)} W^{(k)}$ has deficiency $\geq \delta(1 - \delta)^m \varepsilon^m$. For $m = 0$ the claim is trivial. Assuming that it holds for $m - 1$, let $A = \Lambda^{(m)} W^{(m)}$ and $B = \mathcal{P}_{m-1}$. If $i \in J_m$ then either $i \in S_{\Lambda^{(m)}}^\delta$ and hence $\hat{a}_i \geq \delta$ or such j exists for which $w_{ij}^{(m)} \geq \varepsilon$ and $\hat{b}_j \geq \delta \varepsilon^{m-1}$. Using (13), one now proves the claim for m in the same way as in Theorem 5. ∎

Proof of Theorem 12 is immediate from Lemma 11. Let $P(m) \triangleq \prod_{k=0}^m \Lambda(k) W(k)$. Notice that if $P(m-1)\mathbf{1}_n \leq \theta \mathbf{1}_n$ and $\{(\Lambda(k), W(k))\}_{k=m}^{m+s}$ belongs to $\mathbf{CFJ}[\delta, \varepsilon, s]$, then $P(m+s)\mathbf{1}_n \leq \theta(1 - \delta \varepsilon^s)\mathbf{1}_n$ due to Lemma 11. This implies, via induction on $r = 0, 1, \ldots$, that if the sequence $\{(\Lambda(k), W(k))\}_{k=0}^m$ contains r non-intersecting subsequences from $\mathbf{CFJ}[\delta, \varepsilon, s]$, then one has $P(m)\mathbf{1}_n \leq (1 - \delta \varepsilon^s)^r$. Therefore $P(m) \to 0$ as $m \to \infty$. ∎

6.3 Proofs of Lemmas 7 and 8

Proof of Lemma 7. Statements 1)-3) are proved using induction on k. For instance, if $x_i(0) \leq M$ and $u_i \leq M$ for any i, then $x_i(1) \leq M \, \forall i$ due to (10). Therefore, $x_i(2) \leq M$ and so on. Statement 4) follows from 2) due to the linearity of (10): $x(\cdot|x^0 + \delta x, u + \delta u) = x(\cdot|x^0, u) + x(\cdot|\delta x, \delta u)$. ∎

Proof of Lemma 8. Implications 2)$\Longrightarrow$3) and 3)$\Longrightarrow$1) are obvious (the first of them is proved by putting $u_i = u_*$ and the second one by taking $u_* = 0$). To prove the implication 1)$\Longrightarrow$2), note that the limits $\limsup x(k), \liminf x(k)$ do not depend on $x(0)$ due to stability. Assuming that $x(0) = 0$, the claim follows now from statement 2) in Lemma 7 by substituting $m = \min_i u_i$ and $M = \max_i u_i$. ∎

7. CONCLUSIONS

In this paper, important system-theoretic properties of the Friedkin-Johnsen (FJ) model of opinion dynamics (Friedkin and Johnsen, 1999) are considered such as stability

and convergence speed. We also examine the extension of the FJ model to the case of time-varying social influence and give sufficient conditions for its stability. The time-varying FJ model can be further extended to the case of multidimensional opinions, representing the agents' positions on several interrelated issues (*belief systems*); for static FJ model such an extension is discussed in (Parsegov et al., 2017; Friedkin et al., 2016b). In our future works we are going to validate the applicability of the FJ model to opinion dynamics in large-scale online social networks.

REFERENCES

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

Bindel, D., Kleinberg, J., and Oren, S. (2011). How bad is forming your own opinion? In *Proc. of IEEE Symp. on Foundations of Computer Science*, 57–66.

Blondel, V., Hendrickx, J., and Tsitsiklis, J. (2009). On Krause's multiagent consensus model with state-dependent connectivity. *IEEE Trans. Autom. Control*, 54(11), 2586–2597.

Blondel, V., Hendrickx, J., Olshevsky, A., and Tsitsiklis, J. (2005). Convergence in multiagent coordination, consensus, and flocking. In *Proc. IEEE Conf. Decision and Control*, 2996 – 3000.

Castellano, C., Fortunato, S., and Loreto, V. (2009). Statistical physics of social dynamics. *Rev. Modern Phys.*, 81, 591–646.

Chen, X., Liu, J., Belabbas, M.A., Xu, Z., and Basar, T. (2016). Distributed evaluation and convergence of self-appraisals in social networks (publ. online).

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

Easley, D. and Kleinberg, J. (2010). *Networks, Crowds and Markets. Reasoning about a Highly Connected World*. Cambridge Univ. Press, Cambridge.

Fläche, A. and Macy, M. (2011). Small worlds and cultural polarization. *J. Math. Sociology*, 35(1–3), 146–176.

Frasca, P., Ishii, H., Ravazzi, C., and Tempo, R. (2015). Distributed randomized algorithms for opinion formation, centrality computation and power systems estimation: A tutorial overview. *Europ. J. Control*, 24(7), 2–13.

Frasca, P., Ravazzi, C., Tempo, R., and Ishii, H. (2013). Gossips and prejudices: Ergodic randomized dynamics in social networks. In *Proc. of IFAC NecSys 2013 Workshop*, 212–219. Koblenz, Germany.

Friedkin, N. (2015). The problem of social control and coordination of complex systems in sociology: A look at the community cleavage problem. *IEEE Control Syst. Mag.*, 35(3), 40–51.

Friedkin, N., Jia, P., and Bullo, F. (2016a). A theory of the evolution of social power: natural trajectories of interpersonal influence systems along issue sequences. *Soc. Science*, 3, 444–472.

Friedkin, N. and Johnsen, E. (1999). Social influence networks and opinion change. In *Advances in Group Processes*, volume 16, 1–29.

Friedkin, N. and Johnsen, E. (2011). *Social Influence Network Theory*. Cambridge Univ. Press, New York.

Friedkin, N. and Johnsen, E. (2014). Two steps to obfuscation. *Social Networks*, 39, 12–13.

Friedkin, N., Proskurnikov, A., Tempo, R., and Parsegov, S. (2016b). Network science on belief system dynamics under logic constraints. *Science*, 354(6310), 321–326.

Gantmacher, F. (2000). *The Theory of Matrices*, volume 2. AMS Chelsea Publishing.

Ghaderi, J. and Srikant, R. (2014). Opinion dynamics in social networks with stubborn agents: Equilibrium and convergence rate. *Automatica*, 50(12), 3209–3215.

Hegselmann, R. and Krause, U. (2002). Opinion dynamics and bounded confidence models, analysis, and simulation. *J. Artifical Societies and Social Simulation*, 5(3), 2.

Horn, R. and Johnson, C. (1985). *Matrix Analysis*. Cambridge Univ. Press.

Jia, P., Mirtabatabaei, A., Friedkin, N., and Bullo, F. (2015). Opinion dynamics and the evolution of social power in influence networks. *SIAM Rev.*, 57(3), 367–397.

Lin, H. and Antsaklis, P. (2009). Stability and stabilizability of switched linear systems: A survey of recent results. *IEEE Trans. Autom. Control*, 54(2), 308–322.

Meyer, C. (2000). *Matrix Analysis and Applied Linear Algebra*. SIAM.

Newman, M. (2003). The structure and function of complex networks. *SIAM Rev.*, 45(2), 167–256.

Parsegov, S., Proskurnikov, A., Tempo, R., and Friedkin, N. (2017). Novel multidimensional models of opinion dynamics in social networks (publ. online). *IEEE Trans. Autom. Control*.

Proskurnikov, A., Matveev, A., and Cao, M. (2016a). Opinion dynamics in social networks with hostile camps: Consensus vs. polarization. *IEEE Trans. Autom. Control*, 61(6), 1524–1536.

Proskurnikov, A. and Tempo, R. (2017). A tutorial on modeling and analysis of dynamic social networks. Part I. *Annual Rev. Control*, 43, (accepted).

Proskurnikov, A., Tempo, R., and Cao, M. (2016b). PageRank and opinion dynamics: Missing links and extensions. In *Proc. of IEEE Conf. Norbert Wiener in the 21st Century*, 12–17. Melbourne.

Ravazzi, C., Frasca, P., Tempo, R., and Ishii, H. (2015). Ergodic randomized algorithms and dynamics over networks. *IEEE Trans. Control of Network Syst.*, 2(1), 78–87.

Ren, W. and Beard, R. (2008). *Distributed Consensus in Multi-Vehicle Cooperative Control: Theory and Applications*. Springer-Verlag, London.

Ren, W. and Cao, Y. (2011). *Distributed Coordination of Multi-agent Networks*. Springer.

Takács, K., Fläche, A., and Mäs, M. (2016). Discrepancy and disliking do not induce negative opinion shifts. *PLoS One*, 11(6), e0157948.

Wasserman, S. and Faust, K. (1994). *Social network analysis*. Cambridge Univ. Press, Cambridge.

Weisbuch, G., Deffuant, G., and Amblard, F. (2005). Persuasion dynamics. *Physica A*, 353, 555–575.

Xia, W. and Cao, M. (2011). Clustering in diffusively coupled networks. *Automatica*, 47(11), 2395–2405.

Xia, W., Cao, M., and Johansson, K. (2016). Structural balance and opinion separation in trust–mistrust social networks. *IEEE Trans. Control Netw. Syst.*, 3(1), 46–56.