Friedkin-Johnsen Model Is Distributed Gradient Descent

Akgün, Orhan Eren; Vékássy, Áron; Ballotta, Luca; Yemini, Michal; Gil, Stephanie

DOI
10.1109/LCSYS.2025.3581718

Publication date
2025

Document Version
Final published version

Published in
IEEE Control Systems Letters

Friedkin-Johnsen Model is Distributed Gradient Descent

Orhan Eren Akgün, Áron Vékássy, *Graduate Student Member, IEEE*, Luca Ballotta, Michal Yemini, *Member, IEEE*, and Stephanie Gil, *Member, IEEE*

***Abstract*—The Friedkin-Johnsen (FJ) model describes how agents adjust their opinions through repeated interactions while accounting for the influence of agents who are partially stubborn. In this letter, we demonstrate that the FJ model is *stepwise equivalent* to solving the average consensus problem via distributed gradient descent. This perspective provides a unifying framework that bridges opinion dynamics and optimization, enabling the application of well-established results from the optimization literature. To illustrate this, we examine the recently proposed FJ model with diminishing stubbornness and extend prior results that were concerned with fixed communication graphs to time-varying and jointly connected communication graphs. We derive convergence guarantees and analyze convergence rates under these relaxed assumptions. Finally, we present numerical experiments on random graphs to showcase the impact of diminishing stubbornness dynamics on convergence in both static and time-varying settings.**

***Index Terms*—Agents-based systems, modeling, optimization.**

I. Introduction

THE FRIEDKIN-JOHNSEN (FJ) model is widely studied in opinion dynamics for modeling how a network of interacting agents updates their opinions under persistent influences, referred to as "stubbornness" [1], [2]. Originally developed as a modification to the DeGroot consensus dynamics [3] to account for disagreement, the model has been generalized to a range of settings [2], [4], [5], [6], and has recently been used in the design of resilient control algorithms [7], [8]. In this letter, we study the FJ model from an optimization perspective, showing that it corresponds to solving a distributed optimization problem with quadratic costs via distributed gradient descent (DGD) [9]. This formal connection establishes a fundamental link between the FJ model and distributed optimization methods, which we use to derive new convergence results for the FJ dynamics.

The FJ model is studied in the literature through different lenses, each providing a unique perspective. The opinion dynamics literature typically studies the convergence of the model with given parameters [1], [2], [10], [11], and has explored multi-dimensional opinions [4], [12], random graphs [5], and noisy environments [6], [13]. In game theory, the FJ model is the best response of selfish agents myopically optimizing quadratic costs that favor partial cooperation [10], [14], [15].

Our work is closely related to recent studies of the FJ model for multi-agent control [7], [8], [16].The works [7], [8] treat the stubbornness parameter in the FJ model as a tunable design variable to enhance resilience of consensus, while the analysis in [16] focuses on convergence guarantees and rate in settings with diminishing stubbornness. Also, closely related is this letter [17], where the DeGroot-Friedkin model is shown to be equivalent to a centralized mirror descent algorithm, where each dimension of the state vector represents an agent's opinion. However, analyzing the FJ dynamics using standard consensus tools is nontrivial [16] and the centralized optimization perspective in [17] does not capture the decentralized nature of independent agents, such as the critical role of communication graphs in shaping the dynamics.

Our established equivalence complements the previous literature by connecting the FJ model to distributed optimization. Particularly, treating the stubbornness parameter as a stepsize aligns with the optimization and control focus on strategically designing parameters to enforce a desired converge behavior [9], [18], supporting the design of FJ-inspired algorithms, e.g., for resilient consensus [8]. Further, this link paves the path to extending current results on the FJ model to new settings, since DGD has been analyzed under a wide range of setups including noise, malicious agents, and non-quadratic costs.

In addition to establishing the equivalence between the DGD and the FJ model, we make the first steps towards fully leveraging this new perspective. Our novel contributions are: 1) to derive a necessary condition for consensus with time-varying FJ dynamics using a fixed-point analysis of the corresponding DGD iterates; 2) to establish convergence of the FJ model under time-varying graphs with uniformly diminishing

Received 17 March 2025; revised 12 May 2025; accepted 1 June 2025. Date of publication 20 June 2025; date of current version 11 July 2025. This work was supported by AFOSR under Award FA9550-22-1-0223. Recommended by Senior Editor I.-C. Morarescu. *(Corresponding author: Orhan Eren Akgün.)*

Orhan Eren Akgün, Áron Vékássy, and Stephanie Gil are with the School of Engineering and Applied Sciences, Harvard University, Cambridge, MA 02138 USA (e-mail: erenakgun@g.harvard.edu; avekassy@g.harvard.edu; sgil@seas.harvard.edu).

Luca Ballotta is with the Delft Center for Systems and Control, Delft University of Technology, 2628 CD Delft, The Netherlands (e-mail: l.ballotta@tudelft.nl).

Michal Yemini is with the Faculty of Engineering, Bar-Ilan University, Ramat Gan 5290002, Israel (e-mail: michal.yemini@biu.ac.il).

Digital Object Identifier 10.1109/LCSYS.2025.3581718

stubbornness; 3) to provide convergence rate results for a commonly used stepsize sequences in optimization; and 4) to discuss alternative model structures and noise, supported by numerical studies.

II. Time-Varying Friedkin-Johnsen Model

In this part, we present the time-varying Friedkin-Johnsen (FJ) model [2]. We consider a problem with n agents communicating over a possibly time-varying directed graph $\mathcal{G}_t = (\mathcal{V}, \mathcal{E}_t)$, where $\mathcal{V}$ with $|\mathcal{V}| = n$ denotes the set of agents and $\mathcal{E}_t$ denotes the communication edges at time t. We assume that $(i, i) \in \mathcal{E}_t$ for all $t \geq 0$. Existence of the edge $(j, i) \in \mathcal{E}_t$ indicates that agent j sends information to agent i at time t. Each agent $i \in \mathcal{V}$ stores a state variable $x_i(t) \in \mathbb{R}$ at time t with initial value $x_i(0)$. We consider the time-varying Friedkin-Johnsen (FJ) model with the following dynamics [2]:

$$x_i(t+1) = \lambda_i(t)x_i(0) + (1 - \lambda_i(t)) \sum_{j \in \mathcal{V}} w_{ij}(t)x_j(t), \quad (1)$$

where $w_{ij}(t)$ are the non-negative weights agent i assigns to agent j's opinions such that $w_{ij}(t) \geq 0$ and $w_{ij}(t) > 0$ if and only if $(j, i) \in \mathcal{E}_t$. We make two small modifications to the formulation in [2]. For the sake of simplicity, we set prejudices u_i in [2] to $x_i(0)$. Allowing for arbitrary u_i has interpretation in social networks, however, it does not affect the analysis. Furthermore, $\lambda_i(t) \in [0, 1]$ in our formulation represents *competition* or *stubbornness* of agent i at time t where a larger $\lambda_i(t)$ corresponds to a more stubborn agent as in [16]. The *susceptibility* in [2] corresponds to $1-\lambda_i(t)$ in our case. Letting $\lambda_i(t) \equiv \lambda_i > 0$ corresponds to the standard FJ model [19] and $\lambda(t) \equiv 0$ corresponds to the DeGroot model [3] and classical consensus algorithms in networked multi-agent control [20]. Moreover, imposing uniform stubbornness across all agents such that $\lambda_i(t) = \lambda(t)\ \forall i \in \mathcal{V}$ with fixed weights w_{ij} over static graphs corresponds to the competition model studied in [16].

In the next section, we show that the iterates of the FJ model given in (1) correspond to solving a specific distributed optimization problem using a variant of the distributed (sub)gradient descent (DGD) algorithm proposed in [9].

III. The Friedkin-Johnsen Model as Distributed Gradient Descent

First, we introduce a variant of the DGD algorithm proposed in [9], in its full generality. Later on, we show its connections to the time-varying FJ model.

A. Distributed Gradient Descent

In the DGD algorithm, we let each agent i store a $d-$dimensional state vector $x_i(t) \in \mathbb{R}^d$, initialized at $x_i(0) \in \mathbb{R}^d$. Agents have continuously differentiable private cost functions $f_i : \mathbb{R}^d \rightarrow \mathbb{R}$ with gradients $\nabla f_i(x)$. We make the following assumption on the cost functions:

Assumption 1: For all agents $i \in \mathcal{V}$, function f_i has L-Lipschitz continuous gradients, i.e., $\|\nabla f_i(x) - \nabla f_i(y)\| \leq L\|x - y\|$ for some $L > 0$ and for all $x, y \in \mathbb{R}^d$. Moreover, f_i are μ-strongly convex such that $\langle \nabla f_i(x) - \nabla f_i(y), x - y\rangle \geq \mu\|x - y\|^2$, for all $x, y \in \mathbb{R}^d$ for some $\mu > 0$.

Each agents' goal is to find an optimal solution to the following problem using only local gradient information and by exchanging state variables with its neighbors:

$$x^* \in \arg\min_{x \in \mathbb{R}^d} \sum_{i \in \mathcal{V}} f_i(x). \quad (2)$$

In the DGD algorithm, agents update their values following the iterates:

$$v_i(t+1) = \sum_{j \in \mathcal{V}} w_{ij}(t)x_j(t) \quad (3)$$

$$x_i(t+1) = v_i(t+1) - \lambda_i(t)\nabla f_i(v_i(t+1)). \quad (4)$$

We denote the weight matrix with elements $w_{ij}(t)$ by $W(t)$. The convergence propensities and fixed points of this algorithm depends on the choices of weights $W(t)$ as well as the stepsizes $\lambda_i(t)$. Here, we intentionally use $\lambda_i(t)$ to denote the stepsize since subsequently we will show its equivalence to the *stubbornness parameter* in the FJ model.

Note that the dynamics of the form (4) is commonly studied for the case of constrained distributed optimization (for example, see [18]). In the constrained setting, agents' goal is to solve the optimization problem (2) subject to the constraint $x \in \mathcal{X}$, where $\mathcal{X} \subseteq \mathbb{R}^d$ is a closed and convex set. To also take advantage of the results in the constrained optimization literature to analyze the FJ model, we introduce the distributed projected gradient (DPG) algorithm, as proposed in [18]:

$$x_i(t+1) = P_{\mathcal{X}}[v_i(t+1) - \lambda_i(t)\nabla f_i(v_i(t+1))], \quad (5)$$

where $P_{\mathcal{X}}[v]$ denotes the projection of vector x to the closed convex set $\mathcal{X}$ and $v_i(t+1)$ is the same as in Eq. (3). In the next part, we establish the equivalence between the DGD dynamics in (4) and the projected gradient dynamics in (5) with the FJ model, under appropriately chosen cost functions $f_i(x)$ and constraint set $\mathcal{X}$.

B. Equivalence of the FJ Model and DGD

In this section, we present our main result that establishes the formal equivalence between the DGD algorithm and the time-varying FJ model introduced in the previous sections.

Theorem 1: For each agent i, let $f_i(x) = \frac{1}{2}(x - x_i(0))^2$. Define the constraint set $\mathcal{X} \triangleq \text{conv}(x_1(0), \ldots, x_n(0))$, which is the convex hull of agents' initial values. Then, the following hold true:

1) The DGD algorithm in (3)–(4) is equivalent to the time-varying Friedkin-Johnsen model in (1) for all $\lambda_i(t) \in \mathbb{R}$.
2) If the weight matrix $W(t)$ is row-stochastic and $\lambda_i(t) \in [0, 1]$ for all $t \geq 0$, then the DPG algorithm in (5) is equivalent the time-varying Friedkin-Johnsen model in (1).

Moreover, both the unconstrained problem in (2) and its constrained counterpart with $x \in \mathcal{X}$ have a unique solution given by $x^* = \frac{1}{n}\sum_{i \in \mathcal{V}} x_i(0)$.

Proof: For each $f_i(x) = \frac{1}{2}(x - x_i(0))^2$, we have the gradient $\nabla f_i(x) = x - x_i(0)$.

Proof of 1) We rewrite the update rule in (4) for these choices of cost functions:

$$\begin{aligned} x_i(t+1) &= \upsilon_i(t+1) - \lambda_i(t)(\upsilon_i(t+1) - x_i(0)) \\ &= \sum_{j\in\mathcal{V}} w_{ij}(t)x_j(t) - \lambda_i(t)(\sum_{j\in\mathcal{V}} w_{ij}(t)x_j(t) - x_i(0)) \\ &= \lambda_i(t)x_i(0) + (1-\lambda_i(t))\sum_{j\in\mathcal{V}} w_{ij}(t)x_j(t), \end{aligned} \tag{6}$$

which is equivalent to the time-varying Friedkin-Johnsen dynamics in (1).

Proof of 2) Generally, the gradient step in distribution optimization schemes such as [9] can take us outside the set $\mathcal{X}$, and thus the projection operator is needed to ensure obeying to the constraint set $\mathcal{X}$. However, we next show that for (5) the projection operator is not activated. Let us define $z_i(t+1) \triangleq \lambda_i(t)x_i(0) + (1-\lambda_i(t))\sum_{j\in\mathcal{V}} w_{ij}(t)x_j(t)$, which is the expression inside the projection operator in (5) as shown in Eq.(6). Our goal is to prove that $z_i(t+1) \in \mathcal{X}$ for all $t \geq 0$, therefore, $P_{\mathcal{X}}[z_i(t+1)] = z_i(t+1)$. We have $x_i(0) \in \mathcal{X} = \text{conv}(x_1(0), \ldots, x_n(0))$ by definition. For any $z_i(t+1)$, if the weight matrix $W(t)$ is row-stochastic, then $\sum_{j\in\mathcal{V}} w_{ij}(t)x_j(t) \in \mathcal{X}$. Furthermore, if $\lambda_i(t) \in [0, 1]$, it follows that $z(t+1) \in \mathcal{X}$. Therefore, for all $t \geq 0$, we have $P_{\mathcal{X}}[z_i(t+1)] = z_i(t+1)$, and the DPG update in (5) is identical to the DGD update rule in (4), which we have already shown to be equivalent to the time-varying FJ dynamics in (1).

Finally, the local cost functions $f_i(x)$ are quadratic and strongly convex with $\mu = 1$. Therefore, the unconstrained optimization problem in (2) with this choice of $f_i(x)$ has a unique solution $x^* = \frac{1}{n}\sum_{i\in\mathcal{V}} x_i(0)$. Since $x^* \in \mathcal{X}$, it is also the solution to the constrained optimization problem. ■

Theorem 1 shows that the FJ model can be interpreted as a network of agents running the DGD (or DPG) algorithm with possibly uncoordinated and time-varying stepsizes $\lambda_i(t)$. The stubbornness parameters in the FJ model correspond exactly to the stepsizes of DPG. Given this equivalence, we use the terms "stubbornness" and "stepsizes" interchangeably throughout the rest of this letter. Since the equivalence holds under a precise choice of quadratic costs, we consider this case in the following.

Assumption 2: For each agent $i \in \mathcal{V}$, the local cost function is given by $f_i(x) = \frac{1}{2}(x - x_i(0))^2$. Moreover, the constraint set is defined as $\mathcal{X} \triangleq \text{conv}(x_1(0), \ldots, x_n(0))$.

The functions in Assumption 2 satisfies Assumption 1 with $\mu = L = 1$. Notably, the game-theoretic formulation of the FJ model considers different local cost functions of the form

$$f_i(x) = (x_i - u_i)^2 + \sum_{j\in\mathcal{V}} w_{ij}(t)(x_i - x_j)^2$$

where $u_i = x_i(0)$ in our case; see [10], [15]. The first term penalizes deviation from the agent's belief, while the second encourages cooperation. Minimizing this cost with respect to x_i at time t produces the FJ iterate (1). Coupling terms are absent in our formulation as they are enforced by the cooperative DGD and DPG update rules in (4) and (5), respectively.

C. Variants of DGD

The variant of DGD we use here is sometimes referred to as Combine-then-Adapt (CTA) strategy, since agents first calculate their average (combine) then evaluate and apply the gradient step (adapt) [21]. Other distributed optimization algorithms can be used to solve the minimization problem in (2) [22], [23], [24]. However, these algorithms do not produce iterates that are equivalent to the FJ model. For example, the original algorithm in [9] follows an Adapt-then-Combine (ATC) strategy, where the gradient is evaluated at the value before averaging, i.e., at $x_i(t)$. As we discuss in Section V-B, the update rule of the ATC strategy does not correspond to the FJ model.

IV. Analysis of the FJ Model

It is well-known that the DGD with fixed stepsizes does not converge to a consensus in general [25]. The following result shows that analyzing the behavior of DGD with fixed stepsizes yields a necessary condition for consensus with the FJ model.

Proposition 1: Suppose Assumption 1 holds true, and consider the DGD update rule in (4) whose weights satisfy $\sum_{j\in\mathcal{V}} w_{ij}(t) = 1$ where $w_{ij}(t) \geq 0$ all $i \in \mathcal{V}$ and for all $t \geq 0$. Suppose further that both the stepsizes and the iterates of the algorithm converge, i.e., $\lim_{t\to\infty} \lambda_i(t) = \bar{\lambda}_i$ and $\lim_{t\to\infty} x_i(t) = \bar{x}$ for some $\bar{x} \in \mathbb{R}$ and $\bar{\lambda}_i \in [0, 1]$ for all $i \in \mathcal{V}$. Then, we have $\bar{\lambda}_i \nabla f_i(\bar{x}) = 0$ for all $i \in \mathcal{V}$. Moreover, if Assumption 2 holds true and the agents start from distinct initial values $x_i(0) \neq x_j(0)$ for all $i, j \in \mathcal{V}$, then the limit condition $\bar{\lambda}_i > 0$ holds for at most one agent in $\mathcal{V}$.

Proof: Taking limits on both sides of (4) and using Assumption 1 to interchange limit and gradient yield

$$\begin{aligned} \bar{x} &= \sum_{j\in\mathcal{V}} w_{ij}(t)\bar{x} - \bar{\lambda}_i \nabla f_i\left(\sum_{j\in\mathcal{V}} w_{ij}(t)\bar{x}\right) \\ \bar{x} &= \bar{x} + \bar{\lambda}_i \nabla f_i(\bar{x}). \end{aligned}$$

Therefore, we have $\bar{\lambda}_i \nabla f_i(\bar{x}) = 0$. When we choose $f_i(x) = \frac{1}{2}(x - x_i(0))^2$ with different $x_i(0)$ for each agent, this condition implies that either $\nabla f_i(\bar{x}) = 0$ or $\bar{\lambda}_i = 0$ for each agent. Since we can have $\bar{x} = x_i(0)$ for at most one agent due to distinct initial values, every other agent must have $\bar{\lambda}_i = 0$. ■

Proposition 1 uses fixed-point analysis from optimization theory to show that a necessary condition for consensus in the FJ dynamics is that at most one agent is stubborn in the limit. Similar results are established in the FJ literature with constant stubbornness parameters rather than at the limit, e.g., [2].

In the rest of this section, we primarily provide sufficiency conditions or existence guarantees for convergence and convergence rate, for the homogeneous (coordinated) stepsize case where $\lambda_i(t) = \lambda(t)$ for all $i \in \mathcal{V}$. Note that these results are not intended to be exhaustive as we do not cover all possible optimization approaches that could be applied to the FJ model.

A. Assumptions

In the optimization literature, stepsizes are often times shared (coordinated) across all the agents. This corresponds to having the same stubbornness parameter for all agents in the FJ model, which we refer to as uniform stubbornness. The uniform stubbornness case is also studied in the FJ literature [5], [16].[1] Therefore, in our results, we consider the uniform stubbornness case where we have $\lambda_i(t) = \lambda(t)$ for all $i \in \mathcal{V}$. For the ease of exposition, we present our results for doubly stochastic weight-matrices $W(t)$. However, these results can easily be extended to the cases with only row-stochastic matrices over static graphs. In these cases, the iterates converges to $\sum_{i\in\mathcal{V}} \phi_i x_i(0)$ where ϕ is the Perron vector of the static weight matrix W corresponding to eigenvalue 1 such that $\phi^\intercal W = \phi^\intercal$, instead of the average point $\frac{1}{n}\sum_{i\in\mathcal{V}} x_i(0)$ (See [26, Th. 1] for this kind of analysis).

Assumption 3 (Doubly Stochastic Weights): Assume that the weights $w_{ij}(t)$ are compatible with graph $\mathcal{G}_t$, i.e., $w_{ij}(t) > 0$ if and only if $(j, i) \in \mathcal{E}_t$. Since $(i, i) \in \mathcal{E}_t$ for all $i \in \mathcal{V}$, it follows that $w_{ii}(t) > 0$ always holds. Moreover, the weight matrices $W(t)$ are doubly stochastic, i.e., for all $t \geq 0$, we have $\sum_{i\in\mathcal{V}} w_{ij}(t) = 1$ for all $j \in \mathcal{V}$ and $\sum_{j\in\mathcal{V}} w_{ij}(t) = 1$ for all $i \in \mathcal{V}$.

Assumption 4 (Lower Bound on Non-Zero Weights): Assume that for all agents $i, j \in \mathcal{V}$ and any time $t \geq 0$, non-zero weights are always bounded below by a constant $0 < w < 1$, i.e., $w_{ij}(t) \geq w$ if $w_{ij}(t) > 0$.

Assumption 5 (Infinitely Often Connectedness): There exist time-intervals defined by the infinite sequence $t_0, t_1, \ldots$ with $t_0 = 0$ and $0 < t_{s+1} - t_s \leq T$ for all $s \in \mathbb{Z}$ and some positive integer $T \in \mathbb{Z}^+$ such that the union of graphs in the interval $[t_s, t_{s+1})$ $\bigcup_{k=t_s}^{t_{s+1}-1} G_k$ is strongly connected.

Asssumption 3 ensures that agents can reach consensus on the average of their initial values while following the connectivity conditions imposed by the graph. Asssumption 4 and Asssumption 5 together ensure that agents interact with each other infinitely often, and the impact of these interactions does not diminish over time. Now, we can state our first result.

B. Convergence Guarantees

Proposition 2: Let Assumptions 2–5 hold true. Assume that the stepsizes $\lambda(t)$ satisfy $\sum_{t=0}^{\infty} \lambda(t) = +\infty$ and $\sum_{t=0}^{\infty} \lambda(t)^2 < +\infty$. Then, the sequence $\{x_i(t)\}$ generated by the iterates (5) converges to $\frac{1}{n}\sum_{i\in\mathcal{V}} x_i(0)$ for all $i \in \mathcal{V}$.

Proof: We show this result by using [26, Th. 2] and verifying that the constraint set $\mathcal{X}$ satisfies its assumptions. Note that the algorithm studied in [26] is the PGD in (5). As we show in the proof of Theorem 1 and under Assumption 2, the sequence $\{x_i(t)\}$ always stays within the set $\mathcal{X}$ and the projection operator is the identity. Moreover, $\mathcal{X}$ is a compact and convex set. It only remains to show that $\mathcal{X}$ has a non-empty interior. Note that the consensus point $x_i(0) = x_j(0)$ for all $i, j \in \mathcal{V}$ is a fixed point of (5). Therefore, if agents start at consensus, i.e., $x_i(0) = x_j(0)$ for all $i, j \in \mathcal{V}$, then they always stay at consensus. If there exists at least a pair of agents $i, j \in \mathcal{V}$ such that $x_i(0) \neq x_j(0)$, then, $\mathcal{X}$ has a non-empty interior. By [26, Th. 2], we obtain the claim. ■

This result extends the convergence guarantees in [16] to time-varying uniformly strongly connected graphs from fixed graphs. It is worth remarking that the lenses of DGD allow us to establish convergence properties of FJ dynamics with relatively low effort, while the ad-hoc analysis in [16] is rather convoluted even with constant communication graphs. Proposition 2 proves that the stepsizes $\lambda_i(t)$ need not remain within the interval [0, 1] at all times, unlike in the classical FJ dynamics (1). However, it requires that $\lambda_i(t) \to 0$ as $t \to \infty$, implying that the stepsizes must eventually lie within [0, 1]. Furthermore, if $\lambda_i(t) \approx 1$ at any time t, the update rule in (1) effectively resets the dynamics since in that case $x_i(t + 1) \approx x_i(0)$.

With our next result, we characterize the convergence rate for a specific sequence of stepsizes $\lambda(t)$.

Proposition 3: Let Assumptions 2 and 3 hold true. Let the communication graph $\mathcal{G}_t \equiv \mathcal{G}$ be fixed, strongly connected, and let the matrix $W(t) = W$ be fixed. Choose $\lambda(t) = \min\{1, \frac{4}{(t+2)}\}$. Then, the sequence $\{x_i(t)\}$ generated by (5) converges to $\frac{1}{n}\sum_{i\in\mathcal{V}} x_i(0)$ for all $i \in \mathcal{V}$ at a rate $\mathcal{O}(1/T)$.

Proof: The proof follows from [27, Th. 1] with similar arguments to the proof of Proposition 2. ■

Propositions 2 and 3 show that convergence to the average is guaranteed with non-negative, monotonically non-increasing, square summable $\lambda(t)$. However, [16, Theorem 1] shows that, in the case of the FJ model, $\lim_{t\to\infty} \lambda(t) = 0$ for $\lambda(t) \in [0, 1]$ is both a necessary and sufficient condition for convergence and the summability conditions of $\lambda(t)$ in Proposition 2 are not necessary. To understand this more, we look at the cost functions in the optimization formulation of the FJ model. DGD is a consensus based optimization algorithm. With general cost functions f_i, it uses the variable υ_i from (4) to ensure that the agents reach consensus, and uses the gradients of the local cost functions to guide this consensus toward the optimal solution x^*. For the FJ model, x^* is the point to which the process would converge to without the gradient updates, i.e., if we have no stubbornness, $\lambda(t) = 0$ uniformly across all agents. Thus, the gradient step does not alter where the process converges. We formalize this intuition with the next result.

Proposition 4: Assume that the communication graphs $\mathcal{G}_t$ are fixed, strongly connected, and the matrices $W(t) = W$ are fixed over time. Additionally, assume that W is row-stochastic with the Perron vector ϕ such that $\phi^\intercal W = \phi^\intercal$. Let $x(t)$ denote the state vector with elements $x_i(t)$, $i \in \mathcal{V}$. Then, the weighted average $\phi^\intercal x(0) \in \mathbb{R}$ is an invariant of the iterates in (4), i.e., $\phi^\intercal x(t) = \phi^\intercal x(0)$ for all $t \geq 0$.

Proof: We will use induction to prove this.

Base case $t = 1$:

$$\phi^\intercal x(1) = \phi^\intercal \Big[\lambda(0)x(0) + \big(1 - \lambda(0)\big)Wx(0)\Big] = \phi^\intercal x(0).$$

Inductive hypothesis: $\phi^\intercal x(t) = \phi^\intercal x(0)$ at iteration t.

Inductive step:

$$\begin{aligned}\phi^\intercal x(t+1) &= \phi^\intercal \Big[\lambda(t)x(0) + (1 - \lambda(t))Wx(t)\Big] \\ &= \lambda(t)\phi^\intercal x(0) + (1 - \lambda(t))\phi^\intercal x(t) \\ &= \lambda(t)\phi^\intercal x(0) + (1 - \lambda(t))\phi^\intercal x(0) = \phi^\intercal x(0).\end{aligned}$$

It follows that $\phi^\intercal x(t) = \phi^\intercal x(0)$ for all $t \geq 0$. ■

Proposition 4 implies that the gradient step in (4) does not affect the convergence point as long as convergence happens. This result in turn implies that $\lim_{t\to\infty} \lambda(t) = 0$ guarantees convergence with fixed graphs as shown in [16, Th. 1].

[1] In [5], stubborn agents share the same stubbornness parameter but there are some possibly non-stubborn agents with $\lambda_i(t) = 0$.

Nonetheless, the gradient step becomes particularly useful in recovering from some disturbances to the consensus process. We will explore this in more detail in the next section.

V. Discussion

A. Impact of Noise and Adversaries

In the previous section, our analysis focused on what happens to the FJ dynamics in the absence of external perturbations. In this part, we discuss the impact of noise and adversarial perturbations. Both noisy communications and stochastic gradients are well studied in the optimization literature, whereas noise is typically neglected in opinion dynamics. We reformulate the DGD dynamics in (4) to incorporate additive communication noise and gradient noise:

$$\upsilon_i(t+1) = \sum_{j\in\mathcal{V}} w_{ij}(t)x_j(t) + \sum_{j\in\mathcal{V}\setminus\{i\}} w_{ij}(t)\xi_{ij}(t) \tag{7}$$

$$x_i(t+1) = \upsilon_i(t+1) - \lambda_i(t)\tilde{\nabla} f_i(\upsilon_i(t+1)), \tag{8}$$

where $\tilde{\nabla} f_i(\upsilon_i(t+1)) = \nabla f_i(\upsilon_i(t+1)) + \epsilon_i(t)$. Here, $\xi_{ij}(t)$ is the communication noise in the channel from agent j to i and $\epsilon_i(t)$ is the gradient noise. This setting is commonly studied in the optimization literature, e.g., [28]. Under Assumption 2, after a few steps of algebraic manipulation we can get:

$$x_i(t+1) = \lambda_i(t)(x_i(0) + \epsilon_i(t)) + (1-\lambda_i(t))\left(\sum_{j\in\mathcal{V}} w_{ij}(t)x_j(t) + \sum_{j\in\mathcal{V}\setminus\{i\}} w_{ij}(t)\xi_{ij}(t)\right). \tag{9}$$

Here, $\epsilon_i(t)$ can be interpreted as having noisy access to initial opinions. Notice that allowing for $\lambda_i(t) \to 0$ as we studied in the last section amplifies the impact of the communication noise while reducing the influence of the gradient noise $\epsilon_i(t)$. Without the gradient noise $\epsilon_i(t)$, the dynamics in (9) is similar to the noisy FJ dynamics studied in [6], [13]. Therefore, results from the DGD literature that address communication and stochastic noise, such as [28], can be leveraged to derive theoretical insights into noisy FJ dynamics, including models like [6].

The noise model in Eq. (7) is general enough to capture the impact of adversaries, a widely studied topic in distributed optimization. In particular, if the additive noise term $\xi_{ij}(t)$ is not random but deliberately chosen by an adversary, the dynamics capture scenarios with malicious agents, such as those studied in [7]. While adversarial behavior is difficult to analyze directly in the FJ framework, the optimization perspective help us gain further insights. For example, Eq. (8) shows that the gradients "correct" disturbances caused by malicious agents by steering the state $x_i(t)$ toward the local minima of $f_i(x)$. Moreover, we can apply results from the optimization literature, such as those in [27], to derive new convergence guarantees for FJ dynamics with adversaries, as in the setting studied in [7].

B. Adapt-Then-Combine Strategies

In this section, we consider a more commonly studied DGD variant proposed in [9] that follows the Adapt-then-Combine (ACT) strategy. In this version, agents follow the update rule:

$$x_i(t+1) = \sum_{j\in\mathcal{V}} w_{ij}(t)x_j(t) - \lambda_i(t)\nabla f_i(x_i(t)). \tag{10}$$

Notice that the gradient is evaluated at $x_i(t)$ instead of $\sum_{j\in\mathcal{V}} w_{ij}(t)x_j(t)$ as in the CTA strategy in (4). By employing this algorithm to solve the minimization problem in (2) with quadratic costs $f_i(t) = \frac{1}{2}(x_i(t) - x_i(0))^2$ and following steps similar to the proof of Theorem 1, we obtain:

$$x_i(t+1) = \lambda_i(t)x_i(0) + (w_{ii}(t) - \lambda_i(t))x_i(t) + \sum_{j\in\mathcal{V}\setminus\{i\}} w_{ij}(t)x_j(t). \tag{11}$$

This update rule is similar to the time-varying FJ model in (1), except that the weights assigned to $x_i(t)$ are now $w_{ii}(t) - \lambda_i(t)$ instead of $(1 - \lambda_i(t))w_{ii}(t)$ and the average of the neighbors values $\sum_{j\in\mathcal{V}\setminus\{i\}} w_{ij}(t)x_j(t)$ is not multiplied by $(1 - \lambda_i(t))$. Therefore, agents can potentially assign negative weights to themselves (i.e., when $w_{ii}(t) - \lambda_i(t) < 0$), depending on their choice of $w_{ii}(t)$ and the stubbornness $\lambda_i(t)$. This model still achieves convergence with diminishing uniform stepsizes as shown in [9]. Moreover, this model has been analyzed in various settings, and we believe that similar techniques can be applied to analyze the DGD variant using the CTA strategy in (4), which is equivalent to the FJ iterates.

VI. Numerical Studies

To showcase our analysis, we conduct two experiments over an Erdos-Rényi graph with $n = 100$ agents and edge probability $p = 0.6$. The graph was generated using the `networkx` package in Python with all the random seeds set to 42. The matrix W is computed with each agent assigning equal weights to its neighbors. The stubbornness parameters are identical for all agents at each timestep. The agent's initial values are drawn from $\mathcal{U}([0, 1])$. In both experiments, we plot the mean absolute distance of the agents from $\phi^\intercal x(0)$ as an error metric, where ϕ is the Perron-vector of the weight matrix W.

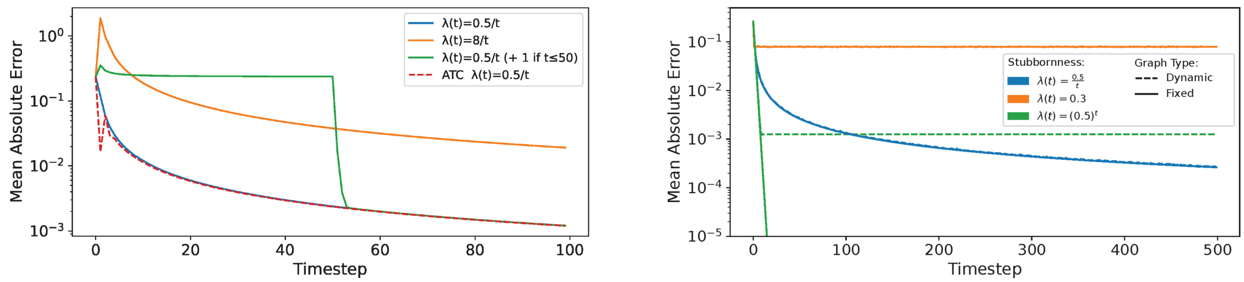
A. ATC and Unconventional Stepsizes

In this part, we study the impact of stepsizes starting from $\lambda(0) > 1$ and the ATC update rule in (10) on a static graph. In Figure 1(a), we observe that $\lambda(0) > 1$ causes the agents to deviate further from $\phi^\intercal x(0)$. However, as the stepsize decays, the dynamics still converge as predicted by Proposition 2, at a rate consistent with Proposition 3. In the case of the ATC update, we see that the gradient is zero at the first step, leading to a sharp decrease in error due to the pure consensus update. After this initial step, both ATC and CTA schemes converge to the same values.

B. Convergence in Time-Varying Graphs

In the second experiment, we study the FJ model on time-varying graphs. While the underlying graph is the same as in Section VI-A, here the edges intermittently fail with some probability at each timestep. For half of the edges the failure probability is $p_\mathrm{f} = 0.2$, and for the other half it is $p_\mathrm{f} = 0.5$. The graphs generated in the experiment satisfy Asssumption 5. Agents assign equal weights to their neighbors and themselves at each timestep, ensuring that $W(t)$ is row-stochastic.

Figure 1(b) shows the average distance of agents from $\phi^\intercal x(0)$ for these time-varying graphs. For comparison, we included the corresponding curves for the same graph and



(a) ATC and the effect of $\lambda(t) > 0$.

(b) Effect of static vs time-varying graphs.

Fig. 1. Average distance of $n = 100$ agents running FJ from $\phi^{\mathsf{T}} x(0)$ against time.

stepsizes, without the edges failing. The experiment aligns with our optimization-based insights, demonstrating that a stubbornness parameter decaying as $\frac{0.5}{t}$ converges to $\phi^{\mathsf{T}} x(0)$. Additionally, we highlight that when stubbornness decays geometrically, the summability of the sequence results in insufficient weight on the gradient-step, preventing it from counteracting imbalances introduced by the graph topology. Consequently, the process does not converge to $\phi^{\mathsf{T}} x(0)$ under time-varying graphs. On the other end of the spectrum, the constant stepsizes do not allow the effect of the gradient steps to vanish. Therefore, while the process does not converge to consensus with a fixed stepsize, it still recovers the fixed point from our first experiment, despite the time-varying topology.

VII. Conclusion

In this letter we formally established the equivalence between Combine-then-Adapt Distributed Gradient Descent with quadratic costs and the Friedkin-Johnsen model. We leveraged analysis tools in the distributed optimization literature to derive new results about the Friedkin-Johnsen model with coordinated and diminishing stubbornness parameters on fixed and time-varying graphs. Also, we discussed how to include noise and use other algorithms. Finally, we validated our analysis in simulations, providing insight to multi-agent control.

References

[1] N. E. Friedkin and E. C. Johnsen, "Social influence networks and opinion change," *Adv. Group Process.*, vol. 16, no. 1, pp. 1–29, 1999.

[2] A. V. Proskurnikov, R. Tempo, M. Cao, and N. E. Friedkin, "Opinion evolution in time-varying social influence networks with prejudiced agents," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 11896–11901, 2017.

[3] M. H. DeGroot, "Reaching a consensus," *J. Am. Stat. Assoc.*, vol. 69, no. 345, pp. 118–121, 1974.

[4] S. E. Parsegov, A. V. Proskurnikov, R. Tempo, and N. E. Friedkin, "Novel multidimensional models of opinion dynamics in social networks," *IEEE Trans. Autom. Control*, vol. 62, no. 5, pp. 2270–2285, May 2017.

[5] L. Wang, Y. Xing, and K. H. Johansson, "On final opinions of the Friedkin-Johnsen model over random graphs with partially stubborn community," in *Proc. IEEE Conf. Decis. Control*, 2025, pp. 4562–4567.

[6] S. Stern and G. Livan, "The impact of noise and topology on opinion dynamics in social networks," *Roy. Soc. Open Sci.*, vol. 8, no. 4, 2021, Art. no. 201943.

[7] L. Ballotta and M. Yemini, "The role of confidence for trust-based resilient consensus," in *Proc. Am. Control Conf.*, 2024, pp. 2822–2829.

[8] L. Ballotta, G. Como, J. S. Shamma, and L. Schenato, "Can competition outperform collaboration? the role of misbehaving agents," *IEEE Trans. Autom. Control*, vol. 69, no. 4, pp. 2308–2323, Apr. 2024.

[9] A. Nedić and A. Ozdaglar, "Distributed subgradient methods for multi-agent optimization," *IEEE Trans. Autom. Control*, vol. 54, no. 1, pp. 48–61, Jan. 2009.

[10] J. Ghaderi and R. Srikant, "Opinion dynamics in social networks with stubborn agents: Equilibrium and convergence rate," *Automatica*, vol. 50, no. 12, pp. 3209–3215, Dec. 2014.

[11] L. Wang, C. Bernardo, Y. Hong, F. Vasca, G. Shi, and C. Altafini, "Consensus in concatenated opinion dynamics with stubborn agents," *IEEE Trans. Autom. Control*, vol. 68, no. 7, pp. 4008–4023, Jul. 2023.

[12] Q. Zhou and Z. Wu, "Multidimensional Friedkin-Johnsen model with increasing stubbornness in social networks," *Inf. Sci.*, vol. 600, pp. 170–188, Jul. 2022.

[13] E. Konovalchikova and Y. A. Dorofeeva, "Opinion dynamics models with noise," in *Proc. CEUR Workshop*, vol. 2792, Apr. 2020, pp. 176–190.

[14] L. Damonte, G. Como, and F. Fagnani, "Systemic risk and network intervention," *IFAC-PapersOnLine*, vol. 53, no. 2, pp. 2856–2861, 2020.

[15] D. Bindel, J. Kleinberg, and S. Oren, "How bad is forming your own opinion?" *Games Economic Behavior*, vol. 92, pp. 248–265, Jul. 2015.

[16] L. Ballotta, Á. Vékássy, S. Gil, and M. Yemini, "Friedkin-Johnsen model with diminishing competition," *IEEE Control Syst. Lett.*, vol. 8, pp. 2679–2684, 2024.

[17] A. Halder, "DeGroot–friedkin map in opinion dynamics is mirror descent," *IEEE Control Syst. Lett.*, vol. 3, no. 2, pp. 463–468, Apr. 2019.

[18] A. Nedić, A. Ozdaglar, and P. A. Parrilo, "Constrained consensus and optimization in multi-agent networks," *IEEE Trans. Autom. Control*, vol. 55, no. 4, pp. 922–938, Apr. 2010.

[19] N. E. Friedkin and E. C. Johnsen, "Social influence and opinions," *J. Math. Sociol.*, vol. 15, nos. 3–4, pp. 193–206, 1990.

[20] R. Olfati-Saber, J. A. Fax, and R. M. Murray, "Consensus and cooperation in networked multi-agent systems," *Proc. IEEE*, vol. 95, no. 1, pp. 215–233, Jan. 2007.

[21] A. H. Sayed, "Adaptation, learning, and optimization over networks," *Found. Trends® Mach. Learn.*, vol. 7, nos. 4–5, pp. 311–801, Jul. 2014.

[22] A. Nedić and A. Olshevsky, "Distributed optimization over time-varying directed graphs," *IEEE Trans. Autom. Control*, vol. 60, no. 3, pp. 601–615, Mar. 2014.

[23] A. Nedić, A. Olshevsky, and W. Shi, "Achieving geometric convergence for distributed optimization over time-varying graphs," *SIAM J. Optim.*, vol. 27, no. 4, pp. 2597–2633, 2017.

[24] R. Xin and U. A. Khan, "A linear algorithm for optimization over directed graphs with geometric convergence," *IEEE Control Syst. Lett.*, vol. 2, no. 3, pp. 315–320, Jul. 2018.

[25] W. Shi, Q. Ling, G. Wu, and W. Yin, "Extra: An exact first-order algorithm for decentralized consensus optimization," *SIAM J. Optim.*, vol. 25, no. 2, pp. 944–966, 2015.

[26] P. Lin, W. Ren, and Y. Song, "Distributed multi-agent optimization subject to nonidentical constraints and communication delays," *Automatica*, vol. 65, pp. 120–131, Mar. 2016.

[27] M. Yemini, A. Nedić, S. Gil, and A. J. Goldsmith, "Resilience to malicious activity in distributed optimization for cyberphysical systems," in *Proc. IEEE Conf. Decis. Control*, 2022, pp. 4185–4192.

[28] K. Srivastava and A. Nedic, "Distributed asynchronous constrained stochastic optimization," *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 4, pp. 772–790, Aug. 2011.