Proximal dynamics in multi-agent network games

Grammatico, Sergio

DOI
10.1109/TCNS.2017.2754358

Publication date
2017

Document Version
Accepted author manuscript

Published in
IEEE Transactions on Control of Network Systems

Citation (APA)
Grammatico, S. (2017). Proximal dynamics in multi-agent network games. *IEEE Transactions on Control of Network Systems*, *5 (2018)*(4), 1707-1716. https://doi.org/10.1109/TCNS.2017.2754358

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCNS.2017.2754358, IEEE Transactions on Control of Network Systems

Proximal dynamics in Multi-Agent Network Games

Sergio Grammatico

***Abstract*—We consider dynamics and protocols for agents seeking an equilibrium in a network game with proximal quadratic cost coupling. We adopt an operator theoretic perspective to show global convergence to a network equilibrium, under the assumption of convex cost functions with proximal quadratic couplings, time-invariant and time-varying communication graph along with convex local constraints, and time-invariant communication graph along with convex local constraints and separable convex coupling constraints. We show that proximal dynamics generalize opinion dynamics in social networks and are applicable to distributed tertiary control in power networks.**

I. INTRODUCTION

Motivation: Distributed decision making in networks populated by rational agents is currently an active field of research across several areas, such as network systems and control, signal processing, computational game theory and operations research. Application domains are in fact numerous and include power systems [1], [2], demand side management [3], [4], [5], network congestion control [6], [7], social networks [8], [9], consensus and flocking [10], [11], robotic and sensor networks [12], [13].

Advantageously, distributed computation and communication setups allow each agent to keep its own data private and exchange information with selected agents only. Typically, in networked multi-agent systems, the state (or decision) variables of each agent evolve as a result of *local decision making*, e.g. local optimization subject to private constraints, and *distributed communication* with some other agents, according to a communication graph. It then follows naturally that the aim of the agents is to reach a collective equilibrium state, where no agent can benefit from updating its state variables.

Literature overview: In this paper, we study network games with proximal, hence quadratic, cost coupling between neighboring agents, that are related to the literature of distributed multi-agent equilibrium seeking in network games and distributed multi-agent optimization.

Network games among agents with convex compact local constraints have been considered in [14] under the assumption of strongly convex quadratic cost functions and time-invariant communication graph; in [15] [16], under the assumption of differentiable cost functions with Lipschitz continuous gradient, strictly monotone pseudo-gradient game mapping (hence strictly convex cost functions), and undirected, possibly time-varying, communication graph. Multi-agent games with convex compact local and also *coupling* constraints have been considered in [17] under the assumption of strongly convex twice differentiable cost functions with bounded gradients, with strictly increasing congestion cost term.

The author is with the Delft Center for Systems and Control, TU Delft, The Netherlands. E-mail address: s.grammatico@tudelft.nl.

Whenever the communication graph is a complete graph with uniform weights, network games reduce to aggregative games, studied e.g. in [15], [18], [19] and [4], the latter under the assumption of strongly convex quadratic cost functions and time-invariant communication graph. Incentive mechanisms for agents playing aggregative games with strongly convex cost functions, convex local and *coupling constraints* have first been studied in [20], and more generally in [21], both with time-invariant communication graph.

Multi-agent convex constrained optimization has been considered in [22], under the assumption of uniformly bounded subgradients, and either homogeneous constraint sets or time-invariant, complete communication graph with uniform weights; in [23] under the assumption of differentiable cost functions with Lipschitz continuous and uniformly bounded gradients; and in [24]. We note that in [23], [24], convergence is proven for agent dynamics with *vanishing* step sizes, which slows down the convergence rate and prevents the protocols to be translated into a continuous-time counterpart, as usual for example in power systems [1], [2], [25], [26].

In general, the theory of generalized (quasi-) variational inequalities [27], [28] and their solution algorithms are applicable to both game equilibrium seeking, under the assumption of convex differentiable cost functions [28, §10, §12], [29, §12], [30, Part II], and to convex optimization [31, §25]. However, the presence of a structured, possibly time-varying, communication graph in multi-agent network games and multi-agent optimization generates the need to design distributed computation and structured information exchange.

Paper contribution: We develop a mathematical framework for multi-agent network games with proximal quadratic cost couplings and show global convergence of multi-agent proximal dynamics and protocols. Thus, we adopt an operator theoretic perspective, which is new in the area of multi-agent and network systems, and allows us to be the first to address network games with non-smooth objective functions, time-varying communication and coupling constraints. Technically, our main contributions are summarized next.

- We consider time-invariant (§II, III) and time-varying (§IV) dynamics and protocols for agents playing network games with convex cost functions and proximal couplings, convex local constraints and separable convex coupling constraints (§V).
- We show that proximal dynamics in multi-agent network games are fixed point iterations with specific structure.
- We derive a distributed protocol for multi-agent network games with separable convex coupling constraints (§V).
- We exploit operator theory (§VI) to show global convergence of some classes of multi-agent dynamics (§VII).
- We show that proximal dynamics in multi-agent network games generalize opinion dynamics in social networks

(§VIII-A) and are applicable to distributed tertiary control in power networks (§VIII-B).

In Section IX, we conclude the paper and provide an outlook on some open research avenues.

Notation and basic definitions: $\mathbb{R}$, $\mathbb{R}_{>0}$, $\mathbb{R}_{\geq 0}$ respectively denote the set of real, positive, and non-negative real numbers; $\overline{\mathbb{R}} := \mathbb{R} \cup \{\infty\}$; $\mathbb{N}$ denotes the set of natural numbers; for $a, b \in \mathbb{N}$, $a \leq b$, $\mathbb{N}[a, b] := [a, b] \cap \mathbb{N}$. $A^\top$ denotes the transpose of A. Given vectors $x_1, \ldots, x_N \in \mathbb{R}^n$, $\boldsymbol{x} := \left[x^1; \ldots; x^N\right]$ denotes $\left[x_1^\top, \ldots, x_N^\top\right]^\top \in \mathbb{R}^{nN}$. I denotes the identity matrix; $\mathbf{0}$ ($\mathbf{1}$) denotes a matrix/vector with all elements equal to 0 (1); to improve clarity, we may add the dimension of these matrices/vectors as subscript. $A \otimes B$ denotes the Kronecker product between matrices A and B. $\|A\|$ denotes the maximum singular value of matrix A. $\mathrm{Id} : \mathbb{R}^n \to \mathbb{R}^n$ denotes the identity operator. $\iota_{\mathcal{S}} : \mathbb{R}^n \to \{0, \infty\}$ denotes the indicator function for the set $\mathcal{S} \subseteq \mathbb{R}^n$, i.e., $\iota_{\mathcal{S}}(x) = 0$ if $x \in \mathcal{S}$, ∞ otherwise.

For a function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$, $\mathrm{dom}(f) := \{x \in \mathbb{R}^n \mid f(x) < \infty\}$; $\partial f : \mathrm{dom}(f) \rightrightarrows \mathbb{R}^n$ denotes its subdifferential set-valued mapping, defined as $\partial f(x) := \{v \in \mathbb{R}^n \mid f(z) \geq f(x) + v^\top (z - x) \text{ for all } z \in \mathrm{dom}(f)\}$; $\mathrm{prox}_f : \mathbb{R}^n \to \mathrm{dom}(f)$ denotes the proximal mapping, defined as $\mathrm{prox}_f(x) := \mathrm{argmin}_{y \in \mathbb{R}^n} f(y) + \frac{1}{2}\|x - y\|^2$.

For a set-valued mapping $\mathcal{A} : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$, $\mathrm{fix}(\mathcal{A}) := \{x \in \mathbb{R}^n \mid x \in \mathcal{A}(x)\}$, $\mathrm{zer}(\mathcal{A}) := \{x \in \mathbb{R}^n \mid 0 \in \mathcal{A}(x)\}$ denote the graph, the sets of fixed point and zeros, respectively.

II. Multi-agent network games with local constraints

We consider N noncooperative agents, where each agent $i \in \mathbb{N}[1, N]$ has state variable $x^i \in \mathcal{X}^i \subseteq \mathbb{R}^n$. We assume that the agents interact through a directed communication graph with $N \times N$ adjacency matrix

$$P := [\, a_{i,j} \,] = \begin{bmatrix} a_{1,1} & \cdots & a_{1,N} \\ \vdots & \ddots & \vdots \\ a_{N,1} & \cdots & a_{N,N} \end{bmatrix}, \tag{1}$$

where $a_{i,j} \in [0, 1]$ is the weight of the communication from agent j to agent i, and $a_{i,j} = 0$ implies no communication from agent j to i. We attach to each agent i a local cost function J^i, and assume that the agents are seeking a collective equilibrium state as defined next.

Definition 1: Network equilibrium. A collective vector $\bar{\boldsymbol{x}} = \left[\bar{x}^1; \bar{x}^2; \cdots; \bar{x}^N\right]$ is a NetWork Equilibrium (NWE) if ($\forall i \in \mathbb{N}[1, N]$)

$$\bar{x}^i \in \underset{y \in \mathcal{X}^i}{\mathrm{argmin}}\; J^i\left(y,\, \textstyle\sum_{j=1}^N a_{i,j}\, \bar{x}^j\right). \tag{2}$$

□

Remark 1: Nash and Wardrop-like equilibria. The network equilibrium concept in Definition 1 reduces to a Nash equilibrium or to a Wardrop-like equilibrium (Wardrop equilibrium with finite number of players) in special cases. In (1), if $a_{i,i} = 0$ for all i, then an NWE reduces to the network version of Nash equilibrium as in [9], [14], [16]. We note that if (2) in Definition 1 is replaced by $\bar{x}^i \in \mathrm{argmin}_{y \in \mathcal{X}^i} J^i\left(y,\, a_{i,i}\, y + \sum_{j \neq i}^N a_{i,j}\, \bar{x}^j\right) =: \mathrm{argmin}_{y \in \mathcal{X}^i} \tilde{J}^i\left(y,\, \bar{\boldsymbol{x}}^{-i}\right)$, then an NWE is a Nash equilibrium for the network game with cost functions $\{\tilde{J}^i\}_{i=1}^N$. If $a_{i,i} > 0$ for all i, then an NWE reduces to a network version of a Wardrop-like equilibrium. □

Throughout the paper, we assume that the local constraint sets are compact and convex, and that the local cost functions are convex, analogously to [15, Assumption 1], [16, Assumption 2], but not necessarily strictly convex.

Standing Assumption 1: Compact, convex sets. For all $i \in \mathbb{N}[1, N]$, the set $\mathcal{X}^i \subset \mathbb{R}^n$ is nonempty, compact and convex. □

Standing Assumption 2: Convex cost functions. For all $i \in \mathbb{N}[1, N]$, the function $J^i : \mathbb{R}^n \times \mathbb{R}^n \to \overline{\mathbb{R}}$ is defined as ($\forall y, z \in \mathbb{R}^n$)

$$J^i(y, z) := f_\circ^i(y) + \iota_{\mathcal{X}^i}(y) + \tfrac{1}{2}\|y - z\|^2, \tag{3}$$

where $f^i := f_\circ^i + \iota_{\mathcal{X}^i} : \mathbb{R}^n \to \overline{\mathbb{R}}$ is lower semi-continuous and convex. □

The addend $\frac{1}{2}\|y - z\|^2$ in (3) is a proximal term used to penalize the distance between the local state and the weighted average of the states of the neighboring agents. We show later in Section VI-C that Standing Assumption 2 implies that the pseudo-gradient game mapping is monotone, not necessarily strictly monotone as in [15, Assumption 2], [16, Assumption 3]. We note that in Standing Assumption 2 the local cost functions need not to be differentiable, nor their gradient need to be Lipschitz continuous and bounded as instead assumed in [15, Assumption 3], [16, Assumption 3]. We implicitly assume that each agent can read the variables of its neighboring agents that interfere with its cost function, namely, we assume that the communication graph is the interference graph.

Standing Assumptions 1, 2 ensure the existence of an NWE. We note in fact that, for all i, $\mathrm{argmin}_{y \in \mathbb{R}^n} J^i(y, z^i) = \mathrm{prox}_{f^i}(z^i)$, where at an NWE as in Definition 1 we have that $z^i = \sum_{j=1}^N a_{i,j}\, x^j$. Now, let us group together the proximal operators and define the mapping $\mathbf{prox}_{\boldsymbol{f}} : \mathbb{R}^{nN} \to (\mathcal{X}^1 \times \cdots \times \mathcal{X}^N) \subset \mathbb{R}^{nN}$ as

$$\mathbf{prox}_{\boldsymbol{f}} := \mathrm{diag}\left(\mathrm{prox}_{f^1}, \ldots, \mathrm{prox}_{f^N}\right) \tag{4}$$

to represent the local optimization of the agents, and let us define the $(n\,N) \times (n\,N)$ matrix

$$\boldsymbol{A} := P \otimes I_n \tag{5}$$

to represent the distributed communication between neighboring agents. Then, it follows that the existence of an NWE can be shown via a fixed point argument.

Lemma 1: Network equilibrium as fixed point. A collective vector $\bar{\boldsymbol{x}} = \left[\bar{x}^1; \cdots; \bar{x}^N\right]$ is a network equilibrium for the game in (2) if and only if $\bar{\boldsymbol{x}} \in \mathrm{fix}(\mathbf{prox} \circ \boldsymbol{A})$. □

Proof: It follows directly from Definition 1. ■

Proposition 1: Existence of network equilibrium. There exists a network equilibrium for the game in (2). □

Proof: The mapping $\mathbf{prox} \circ \boldsymbol{A}$ is continuous and takes values on a compact set, hence $\mathrm{fix}(\mathbf{prox} \circ \boldsymbol{A}) \neq \varnothing$ [32, Theorem 4.1.5 (b)]. The proof then follows by Lemma 1. ■

We note that uniqueness of an NWE does not necessarily hold, see for instance [29, Example 12.4]. We refer to [29,

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCNS.2017.2754358, IEEE Transactions on Control of Network Systems

§12.4] for sufficient conditions that ensure uniqueness of Nash equilibria in noncooperative games.

III. Time-invariant multi-agent dynamics

A. Distributed Banach dynamics

We first analyze simple proximal dynamics, that are ($\forall i \in \mathbb{N}[1,N]$, $\forall k \in \mathbb{N}$)

$$\begin{aligned} x^i(k+1) &= \underset{y \in \mathbb{R}^n}{\operatorname{argmin}} \; J^i\left(y, \, \textstyle\sum_{j=1}^N a_{i,j}\, x^j(k)\right) \\ &= \operatorname{prox}_{f^i}\left(\textstyle\sum_{j=1}^N a_{i,j}\, x^j(k)\right), \end{aligned} \tag{6}$$

and in collective compact form

$$\boldsymbol{x}(k+1) = \mathbf{prox}_{\boldsymbol{f}}\left(\boldsymbol{A}\,\boldsymbol{x}(k)\right). \tag{7}$$

We discuss later in Section VI that the proximal dynamics in (7) represent the Banach fixed point iteration [31, Equation 1.67] applied to the mapping $\mathbf{prox}_{\boldsymbol{f}} \circ \boldsymbol{A}$, hence let us call them Banach dynamics.

In this subsection, we assume that the adjacency matrix satisfies the following linear matrix inequality.

Assumption 1: Averaged adjacency matrix. The matrix P in (1) is such that

$$\begin{bmatrix} (2\eta - 1) I + (1-\eta)(P^\top + P) & P^\top \\ P & I \end{bmatrix} \succcurlyeq 0 \tag{8}$$

for some $\eta \in (0,1)$. □

We note that Assumption 1 holds true if the adjacency matrix is doubly stochastic, as assumed in [14, Remark 1], [15, Assumption 5], [16, Assumption 1], and all self loops are present, as assumed in [22, Assumption 2].

Proposition 2: Doubly stochastic adjacency matrix with self loops. If the matrix $P = [a_{i,j}]$ in (1) is doubly stochastic, i.e., ($\forall i,j \in \mathbb{N}[1,N]$) $a_{i,j} \in \mathbb{R}_{\geq 0}$, $\sum_{j=1}^N a_{i,j} = \sum_{i=1}^N a_{i,j} = 1$, and such that $\min_{i \in \mathbb{N}[1,N]} a_{i,i} =: \underline{a} > 0$, then it satisfies Assumption 1. □

Proof: See §VI-B. ■

We can now show global convergence of the Banach dynamics to an NWE.

Theorem 1: Global convergence of distributed Banach dynamics. If Assumption 1 holds, then the sequence $(\boldsymbol{x}(k))_{k=0}^\infty$ defined as in (7) converges, for any initial condition, to a network equilibrium $\bar{\boldsymbol{x}}$, and ($\forall k \in \mathbb{N}$)

$$\begin{aligned} &\operatorname{dist}\left(\boldsymbol{x}(k), \operatorname{fix}\left(\mathbf{prox}_{\boldsymbol{f}} \circ \boldsymbol{A}\right)\right) \\ &\qquad \leq \frac{1}{(1-\eta)\,(k+1)} \left\| \boldsymbol{x}(0) - \bar{\boldsymbol{x}} \right\|^2 . \end{aligned} \tag{9}$$

□

B. Distributed Krasnoselskij dynamics

Whenever the communication matrix is doubly stochastic but not all self loops are present, the Banach dynamics cannot ensure convergence in general, as illustrated in the following example.

Example 1: Non-convergence of Banach dynamics. The Banach dynamics for the game with $N = 2$ agents, $n = 1$, $f^1 = f^2 = 0$, $\mathcal{X}^1 = \mathcal{X}^2 = [-1, 1]$, $P = \left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right]$ and for $x^1(0) = -x^2(0) = 1$ evolve as $x^1(k) = -x^2(k) = -x^1(k+1) = x^2(k+1) = 1$ for all $k \in \mathbb{N}$, that is, $\boldsymbol{x}(k)$ oscillates persistently, hence does not converge. □

Instead, global convergence to an NWE can be achieved via averaged proximal dynamics. Specifically, for some $\alpha \in (0,1)$, we consider ($\forall i \in \mathbb{N}[1,N]$, $\forall k \in \mathbb{N}$)

$$x^i(k+1) = (1-\alpha)\, x^i(k) + \alpha \operatorname{prox}_{f^i}\left(\textstyle\sum_{j=1}^N a_{i,j}\, x^j(k)\right), \tag{10}$$

so that the collective dynamics read as

$$\boldsymbol{x}(k+1) = (1-\alpha)\,\boldsymbol{x}(k) + \alpha\, \mathbf{prox}_{\boldsymbol{f}}\left(\boldsymbol{A}\,\boldsymbol{x}(k)\right). \tag{11}$$

Analogously to the Banach dynamics in (7), we discuss later in Section VI that the dynamics in (11) represent the Krasnoselskij fixed point iteration [31, Equation 5.12 (fixed step size)] applied to the mapping $\mathbf{prox}_{\boldsymbol{f}} \circ \boldsymbol{A}$, hence let us call them Krasnoselskij dynamics.

We show next global convergence of the distributed Krasnoselskij dynamics to an NWE, under the assumption that the adjacency matrix has norm at most unitary.

Assumption 2: Nonexpansive adjacency matrix. The matrix P in (1) is such that $\|P\| \leq 1$, i.e., (8) holds with $\eta = 1$. □

Proposition 3: Doubly stochastic adjacency matrix. If the matrix $P = [a_{i,j}]$ in (1) is doubly stochastic, i.e., ($\forall i,j \in \mathbb{N}[1,N]$) $a_{i,j} \in \mathbb{R}_{\geq 0}$, $\sum_{j=1}^N a_{i,j} = \sum_{i=1}^N a_{i,j} = 1$, then it satisfies Assumption 2. □

Proof: Since P is doubly stochastic, we have that $\|P\|_1 = \|P\|_\infty = 1$, hence by Hölder's inequality, $\|P\| \leq \sqrt{\|P\|_1 \, \|P\|_\infty} = 1$. ■

Theorem 2: Global convergence of distributed Krasnoselskij dynamics. If Assumption 2 holds, then the sequence $(\boldsymbol{x}(k))_{k=0}^\infty$ defined as in (11) converges, for any initial condition, to a network equilibrium $\bar{\boldsymbol{x}}$, and ($\forall k \in \mathbb{N}$)

$$\begin{aligned} &\operatorname{dist}\left(\boldsymbol{x}(k), \operatorname{fix}\left(\mathbf{prox}_{\boldsymbol{f}} \circ \boldsymbol{A}\right)\right) \\ &\qquad \leq \frac{1}{\alpha\,(1-\alpha)\,(k+1)} \left\| \boldsymbol{x}(0) - \bar{\boldsymbol{x}} \right\|^2 . \end{aligned} \tag{12}$$

□

IV. Time-varying multi-agent dynamics

In this section, we extend the setup in Section II and the convergence results in Section III to the case in which the communication between agents is time dependent. Thus, we consider a time-varying adjacency matrix

$$P(k) := [\, a_{i,j}(k) \,] \tag{13}$$

and, analogously to (5), we define $\boldsymbol{A}(k) := P(k) \otimes I_n$.

In the time-varying case, the notion of NWE and its existence are unclear in general. Thus, let us formulate the following existence assumption.

Assumption 3: Existence of persistent network equilibrium. There exists $\bar{k} \in \mathbb{N}$ such that

$$\mathcal{E} := \textstyle\bigcap_{k \geq \bar{k}} \operatorname{fix}\left(\mathbf{prox}_{\boldsymbol{f}} \circ \boldsymbol{A}(k)\right) \neq \varnothing. \tag{14}$$

□

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCNS.2017.2754358, IEEE Transactions on Control of Network Systems

A. Time-varying Banach and Krasnoselskij dynamics

We are ready to generalize Theorems 1 and 2 in Section III to the time-varying case. Namely, we show global convergence of the time-varying Banach dynamics ($\forall k \in \mathbb{N}$)

$$\boldsymbol{x}(k+1) = \mathbf{prox}_{\boldsymbol{f}}\left(\boldsymbol{A}(k)\,\boldsymbol{x}(k)\right) \tag{15}$$

and the time-varying Krasnoselskij dynamics ($\forall k \in \mathbb{N}$)

$$\boldsymbol{x}(k+1) = (1-\alpha)\,\boldsymbol{x}(k) + \alpha\,\mathbf{prox}_{\boldsymbol{f}}\left(\boldsymbol{A}(k)\,\boldsymbol{x}(k)\right), \tag{16}$$

where $\alpha \in (0,1)$.

For the time-varying Banach dynamics, we assume that the adjacency matrix is persistently averaged, analogously to Assumption 1, while for the time-varying Krasnoselskij dynamics, we assume that the adjacency matrix is persistently nonexpansive, analogously to Assumption 2.

Assumption 4: Persistently averaged adjacency matrix. There exists $\bar{k} \in \mathbb{N}$ such that, for all $k \geq \bar{k}$, the matrix $P(k)$ in (13) is such that

$$\begin{bmatrix} (2\eta-1)I + (1-\eta)(P(k)^\top + P(k)) & P(k)^\top \\ P(k) & I \end{bmatrix} \succcurlyeq 0 \tag{17}$$

for some $\eta \in (0,1)$. □

Assumption 5: Persistently nonexpansive adjacency matrix. There exists $\bar{k} \in \mathbb{N}$ such that, for all $k \geq \bar{k}$, the matrix $P(k)$ in (13) is such that $\|P(k)\| \leq 1$, i.e., (17) holds with $\eta = 1$. □

It follows from Propositions 2, 3 that if the adjacency matrix is persistently doubly stochastic, then Assumption 5 holds true; if in addition, the adjacency matrix has all self loops persistent, e.g. as in [22, Assumptions 2, 3], then Assumption 4 is satisfied. With persistently averaged or nonexpansive adjacency matrix, global convergence of the time-varying Banach and Krasnoselkij dynamics, respectively, holds as shown in the next statements.

Theorem 3: Global convergence of distributed time-varying Banach dynamics. If Assumptions 3, 4 hold, then the sequence $(\boldsymbol{x}(k))_{k=0}^{\infty}$ defined as in (15) converges, for any initial condition, to some vector $\bar{\boldsymbol{x}} = \left[\bar{x}^1 ; \cdots ; \bar{x}^N\right]$, and ($\forall k \in \mathbb{N}$) (9) holds. □

Theorem 4: Global convergence of distributed time-varying Krasnoselskij dynamics. If Assumptions 3, 5 hold and $\alpha \in (0,1)$, then the sequence $(\boldsymbol{x}(k))_{k=0}^{\infty}$ defined as in (16) converges, for any initial condition, to some vector $\bar{\boldsymbol{x}} = \left[\bar{x}^1 ; \cdots ; \bar{x}^N\right]$, and ($\forall k \in \mathbb{N}$) (12) holds. □

We emphasize that the convergence to a specific NWE does depend on the time-varying communication graph. Whenever some communication graph with adjacency matrix $\bar{P}$ that can generate a persistent network equilibrium recurs infinitely often, then we can show that the agents reach an NWE for the game in (2) with adjacency matrix $\bar{P}$.

Assumption 6: Feasible recurrent adjacency matrix. It holds that $\liminf_{k\to\infty} \left\|P(k) - \bar{P}\right\| = 0$, where the matrix $\bar{P}$ satisfies Assumption 2, and that $\mathrm{fix}\left(\mathbf{prox}_{\boldsymbol{f}} \circ \left(\bar{P} \otimes I_n\right)\right) \subseteq \mathcal{E}$, with $\mathcal{E}$ defined as in (14). □

Corollary 1: Global convergence of distributed time-varying dynamics. Let Assumptions 3, 6 hold. If either Assumption 5 holds and $\alpha \in (0,1)$, or Assumption 4 holds and $\alpha \in (0,1]$, then the sequence $(\boldsymbol{x}(k))_{k=0}^{\infty}$ defined as in (16) converges, for any initial condition, to $\bar{\boldsymbol{x}} \in \mathrm{fix}\left(\mathbf{prox}_{\boldsymbol{f}} \circ \bar{\boldsymbol{A}}\right)$, where $\bar{\boldsymbol{A}} := \bar{P} \otimes I_n$. □

V. Multi-agent network games with local and coupling constraints

In this section, we extend the setup in Section II to the case in which the agents are subject to both local and coupling constraints, the latter of the kind $g(\boldsymbol{x}) \leq 0$.

Namely, for each agent $i \in \mathbb{N}[1,N]$, we consider a joint local and coupling constraint set that depends on the other agents, that is,

$$\begin{aligned} \tilde{\mathcal{X}}^i(\boldsymbol{x}^{-i}) \quad &:= \quad \left\{ y \in \mathcal{X}^i \,|\, g\left(\left[\cdots ; x^{i-1} ; y ; x^{i+1} ; \cdots\right]\right) \leq 0 \right\} \\ &=: \quad \left\{ y \in \mathcal{X}^i \,|\, g^i\left(y, \boldsymbol{x}^{-i}\right) \leq 0 \right\}. \end{aligned} \tag{18}$$

Let us then generalize Definition 1 to the case of network games with coupling constraints.

Definition 2: Generalized network equilibrium. A collective vector $\bar{\boldsymbol{x}} = \left[\bar{x}^1 ; \bar{x}^2 ; \cdots ; \bar{x}^N\right]$ is a Generalized NetWork Equilibrium (GNWE) if ($\forall i \in \mathbb{N}[1,N]$)

$$\bar{x}^i \in \underset{y \in \tilde{\mathcal{X}}^i(\bar{\boldsymbol{x}}^{-i})}{\operatorname{argmin}} \; J^i\left(y, \textstyle\sum_{j=1}^N a_{i,j}\,\bar{x}^j\right). \tag{19}$$

□

Remarkably, we can transform a generalized network game into an auxiliary (extended) network game [30, §3] and then, based on the latter, design a protocol that ensures convergence to a GNWE. With this aim, we start with the definition of the following auxiliary NWE.

Definition 3: Extended network equilibrium. The pair $\left(\bar{\boldsymbol{x}}, \bar{\lambda}\right)$, with $\bar{\boldsymbol{x}} = \left[\bar{x}^1 ; \bar{x}^2 ; \cdots , \bar{x}^N\right]$, is an Extended NetWork Equilibrium (ENWE) if ($\forall i \in \mathbb{N}[1,N]$)

$$\bar{x}^i \in \underset{y \in \mathcal{X}^i}{\operatorname{argmin}} \; J^i\left(y, \bar{\boldsymbol{x}}^{-i}\right) + \bar{\lambda}^\top g^i\left(y, \bar{\boldsymbol{x}}^{-i}\right) \tag{20}$$

$$\bar{\lambda} \in \underset{\xi \in \mathbb{R}^M_{\geq 0}}{\operatorname{argmin}} \; -\xi^\top g(\bar{\boldsymbol{x}}). \tag{21}$$

□

For the existence of an equilibrium, we need to assume that the dual variables are bounded, which is implied by a standard constraint qualification, e.g. the Slater or the Mangasarian–Fromovitz constraint qualifications [29, p. 346, pp. 447–448].

Standing Assumption 3: Bounded dual variables. There exists $\hat{\lambda} \in \mathbb{R}_{>0}$, with $[0, \hat{\lambda}]^M =: \mathcal{L} \subset \mathbb{R}^M_{\geq 0}$, such that the system of inclusions in (20), (21) is equivalent to

$$\bar{x}^i \in \underset{y \in \mathcal{X}^i}{\operatorname{argmin}} \; J^i\left(y, \bar{\boldsymbol{x}}^{-i}\right) + \bar{\lambda}^\top g^i\left(y, \bar{\boldsymbol{x}}^{-i}\right) \tag{22}$$

$$\bar{\lambda} \in \underset{\xi \in \mathcal{L}}{\operatorname{argmin}} \; -\xi^\top g(\bar{\boldsymbol{x}}). \tag{23}$$

□

As usual in generalized games [30, Part II, §3], let us also consider separable convex coupling constraints that are affine.

Remark 2: Separable convex coupling constraints. A game with separable convex coupling constraints, $g(\boldsymbol{x}) = \sum_{i=1}^{N} g^i(x^i) \leq 0$, can be reformulated with affine coupling constraints, by letting, for all i, $g^i(x^i) \leq y^i$ be an additional local constraint in the augmented local variable $[x^i\,;\,y^i]$, and $\sum_{i=1}^{N} y^i \leq 0$ be the affine coupling constraint. □

Standing Assumption 4: Affine coupling constraints. The mapping $g : (\mathbb{R}^n)^N \to \mathbb{R}^M$ in (18) is defined as

$$g(\boldsymbol{x}) := C\,\boldsymbol{x} + c = \left[C^1 \mid \cdots \mid C^N\right] \boldsymbol{x} + c, \tag{24}$$

for some matrices $C^1, \ldots, C^N \in \mathbb{R}^{M \times n}$ and vector $c \in \mathbb{R}^M$. □

Analogously to an NWE, also an ENWE in (20)–(21) can be characterized as a fixed point, and in turn related to a GNWE in (19). To show that, first, we exploit a basic property of the proximal mapping [31, Proposition 12.28], namely that the inclusion in (23) holds if and only if

$$\begin{aligned} \bar\lambda &= \underset{\xi \in \mathcal{L}}{\operatorname{argmin}} \; -\xi^\top g(\bar{\boldsymbol{x}}) + \tfrac{1}{2}\left\|\xi - \bar\lambda\right\|^2 \\ &= \operatorname{prox}_{\iota_{\mathcal{L}}}\left(\bar\lambda + g(\bar{\boldsymbol{x}})\right) \\ &= \operatorname{proj}_{\mathcal{L}}\left(\bar\lambda + g(\bar{\boldsymbol{x}})\right). \end{aligned} \tag{25}$$

Then, to combine (20) for all $i \in \mathbb{N}[1, N]$ and (25) together in compact form, let us define the mapping $\boldsymbol{\mathcal{F}} : \mathbb{R}^{nN} \times \mathbb{R}^M \to (\mathcal{X}^1 \times \cdots \times \mathcal{X}^N) \times \mathcal{L} \subset \mathbb{R}^{nN} \times \mathbb{R}^M$ as

$$\boldsymbol{\mathcal{F}} := \operatorname{diag}\left(\mathbf{prox}_{\boldsymbol{f}},\, \operatorname{proj}_{\mathcal{L}}\right) \tag{26}$$

and the mapping $\boldsymbol{\mathcal{G}} : \mathbb{R}^{nN} \times \mathbb{R}^M \to \mathbb{R}^{nN} \times \mathbb{R}^M$ as

$$\boldsymbol{\mathcal{G}}(\cdot) := \boldsymbol{G}\cdot + \begin{bmatrix}\mathbf{0}\\ c\end{bmatrix} := \begin{bmatrix}\boldsymbol{A} & -C^\top \\ C & I\end{bmatrix} \cdot + \begin{bmatrix}\mathbf{0}\\ c\end{bmatrix}. \tag{27}$$

Thus, it follows from Lemma 1 that the pair $\left(\bar{\boldsymbol{x}}, \bar\lambda\right)$ is an ENWE if and only if $[\bar{\boldsymbol{x}}\,;\,\lambda]$ is a fixed point of $\boldsymbol{\mathcal{F}} \circ \boldsymbol{\mathcal{G}}$. Note in fact that $[\bar{\boldsymbol{x}}\,;\,\lambda] = (\boldsymbol{\mathcal{F}} \circ \boldsymbol{\mathcal{G}})\left([\bar{\boldsymbol{x}}\,;\,\lambda]\right)$ if and only if $\bar{\boldsymbol{x}} = \mathbf{prox}_{\boldsymbol{f}}\left(\boldsymbol{A}\bar{\boldsymbol{x}} - g(\bar{\boldsymbol{x}})^\top\bar\lambda\right)$ as it follows from (20) and $\bar\lambda = \operatorname{proj}_{\mathcal{L}}\left(\bar\lambda + g(\bar{\boldsymbol{x}})\right)$ as in (25).

We are now ready to formalize that an ENWE in (20)–(21) generates a GNWE as defined in (19).

Lemma 2: Generalized network equilibrium from fixed point. A collective vector $\bar{\boldsymbol{x}} = \left[\bar x^1\,;\ldots;\,\bar x^N\right]$ is a generalized network equilibrium for the game in (19) if $\left[\bar{\boldsymbol{x}}\,;\,\bar\lambda\right] \in \operatorname{fix}(\boldsymbol{\mathcal{F}} \circ \boldsymbol{\mathcal{G}})$, for some $\bar\lambda \in \mathcal{L}$. □

Proof: It follows from [33, Theorem 3.1], analogously to [19, Theorem 1]. ■

Proposition 4: Existence of generalized network equilibrium. There exists a generalized network equilibrium for the game in (19). □

Proof: The mapping $\boldsymbol{\mathcal{F}} \circ \boldsymbol{\mathcal{G}}$ is continuous and valued on a compact set, hence $\operatorname{fix}(\boldsymbol{\mathcal{F}} \circ \boldsymbol{\mathcal{G}}) \neq \varnothing$ [32, Theorem 4.1.5 (b)]. The proof then follows by Lemma 2. ■

Lemma 2 allows us the reformulate the GNWE problem as an NWE problem. However, the affine mapping $\boldsymbol{\mathcal{G}}$ in (27) does not inherit from the linear mapping $\boldsymbol{A}\,\cdot$ the properties that are sufficient for the convergence of the proximal dynamics studied in Section II. Indeed, the extended proximal dynamics $[\boldsymbol{x}(k+1)\,;\,\lambda(k+1)] = \boldsymbol{\mathcal{F}}\left(\boldsymbol{\mathcal{G}}\left([\boldsymbol{x}(k)\,;\,\lambda(k)]\right)\right)$ may fail to converge, since $\|\boldsymbol{G}\| > 1$ in general.

A. Distributed Tseng protocol

To design a distributed protocol that ensures global convergence to a GNWE for the game in (19), we reformulate the fixed point problem arising in Lemma 2 into an equivalent zero finding problem. Technically, we exploit the following equivalence result.

Lemma 3 (from [31, Proposition 25.1 (iv)]*): Fixed points as zeros.* $\operatorname{fix}(\boldsymbol{\mathcal{F}} \circ \boldsymbol{\mathcal{G}}) = \operatorname{zer}(\boldsymbol{\mathcal{J}_{\mathcal{F}}} + \operatorname{Id} - \boldsymbol{\mathcal{G}})$, where $\boldsymbol{\mathcal{J}_{\mathcal{F}}} := \operatorname{diag}\left(\partial f^1, \ldots, \partial f^N, \partial\iota_{\mathcal{L}}\right)$. □

In view of Lemma 3, splitting methods are applicable for the equilibrium seeking. Inspired by the Tseng splitting [31, §25.4], we derive the following forward-backward-forward distributed protocol:

$$\begin{bmatrix}\tilde{\boldsymbol{x}}(k)\\ \tilde\lambda(k)\end{bmatrix} = \left((1-\alpha)\mathrm{Id} + \alpha\,\boldsymbol{\mathcal{G}}\right)\begin{bmatrix}\boldsymbol{x}(k)\\ \lambda(k)\end{bmatrix} \tag{28}$$

$$\begin{aligned} \boldsymbol{x}(k{+}1/2) &= \mathbf{prox}_{\alpha\boldsymbol{f}}\left(\tilde{\boldsymbol{x}}(k)\right) \\ \lambda(k{+}1/2) &= \operatorname{proj}_{\mathcal{L}}(\tilde\lambda(k)) \end{aligned} \tag{29}$$

$$\begin{bmatrix}\tilde{\boldsymbol{x}}(k{+}1/2)\\ \tilde\lambda(k{+}1/2)\end{bmatrix} = \left((1-\alpha)\mathrm{Id} + \alpha\,\boldsymbol{\mathcal{G}}\right)\begin{bmatrix}\boldsymbol{x}(k{+}1/2)\\ \lambda(k{+}1/2)\end{bmatrix} \tag{30}$$

$$\begin{bmatrix}\boldsymbol{x}(k+1)\\ \lambda(k+1)\end{bmatrix} = \begin{bmatrix}\tilde{\boldsymbol{x}}(k+1/2)\\ \tilde\lambda(k+1/2)\end{bmatrix} + \alpha\left(\mathrm{Id} - \boldsymbol{\mathcal{G}}\right)\begin{bmatrix}\boldsymbol{x}(k)\\ \lambda(k)\end{bmatrix}. \tag{31}$$

We note that the update of the dual variable $\lambda(k)$ shall be carried out by an (additional) agent that has full information on the coupling constraint, namely on the quantity $g\left(\boldsymbol{x}(k)\right)$. The protocol consists of four steps as summarized next:

(28) the agents exchange state information via the mapping $\boldsymbol{\mathcal{G}}$ and then average the new information with their current state;

(29) the agents perform their local optimization in parallel via the weighted proximal mappings $\operatorname{prox}_{\alpha f^i}$ and $\operatorname{prox}_{\alpha\iota_{\mathcal{L}}} = \operatorname{proj}_{\mathcal{L}}$, that is, for all $i \in \mathbb{N}[1, N]$,

$$x^i(k{+}1/2) = \operatorname{prox}_{\alpha f^i}\left(\tilde x^i(k)\right),\ \lambda(k{+}1/2) = \operatorname{proj}_{\mathcal{L}}\left(\tilde\lambda(k)\right);$$

(30) analogously to (29), the agents exchange state information via the mapping $\boldsymbol{\mathcal{G}}$ and then average the new information with their current state;

(31) the agents update their states based on the outcome of the first step in (28) and of the third step in (30).

We conclude the section by establishing global convergence of the distributed Tseng protocol in (28)–(31).

Theorem 5: Global convergence of distributed Tseng protocol. Let $\alpha \in (0, 1/\|\boldsymbol{G}\|)$. Then the sequence $(\boldsymbol{x}(k))_{k=0}^{\infty}$ defined as in (28)–(31) converges, for any initial condition, to a generalized network equilibrium for the game in (19). □

Finally, note that for the global convergence of the distributed Tseng protocol, in Theorem 5, we do not assume that the communication matrix P is averaged or nonexpansive.

VI. An operator theoretic perspective to multi-agent network games

In the following, we adopt an operator theoretic perspective to multi-agent dynamics. Specifically, in view of the compact

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCNS.2017.2754358, IEEE Transactions on Control of Network Systems

notations that describe the collective multi-agent dynamics, namely the Banach dynamics in (7), the Krasnoselskij dynamics in (11), the time-varying dynamics in (15) and (16), and the forward-backward-forward protocol in (28)–(31), we give to them the interpretation of fixed point iterations with special structure. Thus, in the next subsection, let us review the definitions and results that are necessary for the convergence proofs in Section VII.

A. Operator theoretic definitions and results

Definition 4: Nonexpansive mapping. A mapping $\mathcal{T} : \mathbb{R}^n \to \mathbb{R}^n$ is nonexpansive (NE) if $(\forall x, y \in \mathrm{dom}(\mathcal{T}))$

$$\|\mathcal{T}(x) - \mathcal{T}(y)\| \leq \|x - y\| . \tag{32}$$

□

Definition 5: Averaged mapping. A mapping $\mathcal{T} : \mathbb{R}^n \to \mathbb{R}^n$ is η-averaged (η-AVG), with $\eta \in (0, 1)$, if $(\forall x, y \in \mathrm{dom}(\mathcal{T}))$

$$\begin{aligned} &\|\mathcal{T}(x) - \mathcal{T}(y)\|^2 \\ &\leq \|x - y\|^2 - \tfrac{1-\eta}{\eta} \|(\mathrm{Id} - \mathcal{T})(x) - (\mathrm{Id} - \mathcal{T})(y)\|^2 \end{aligned} \tag{33}$$

or equivalently if there exists a nonexpansive mapping $\mathcal{B} : \mathbb{R}^n \to \mathbb{R}^n$ such that $\mathcal{T} = (1 - \eta)\mathrm{Id} + \eta\mathcal{B}$. □

Note that if $\mathcal{T}$ is $\bar{\eta}$-AVG, then it is η-AVG for all $\eta \in (0, \bar{\eta}]$, and that the inequality in (33) with $\eta = 1$ is equivalent to that in (32), hence AVG mappings are NE.

The convergence proofs for the main statements in the paper are based on the following technical results.

Lemma 4 ([31, Proposition 5.15]*): Banach iteration*. Assume that: (i) $\mathrm{fix}(\mathcal{T}) \neq \varnothing$; (ii) $\mathcal{T} : \mathbb{R}^n \to \mathbb{R}^n$ is η-averaged. Then the Banach iteration

$$z(k + 1) := \mathcal{T}(z(k)) \tag{34}$$

is such that $\lim_{k\to\infty} z(k) = \bar{z} \in \mathrm{fix}(\mathcal{T})$. □

Lemma 5 ([31, Theorem 5.14]*): Krasnoselskij iteration*. Assume that: (i) $\mathrm{fix}(\mathcal{T}) \neq \varnothing$; (ii) $\mathcal{T} : \mathbb{R}^n \to \mathbb{R}^n$ is nonexpansive; (iii) $\alpha \in (0, 1)$. Then the Krasnoselskij iteration

$$z(k + 1) := (1 - \alpha)\, z(k) + \alpha\, \mathcal{T}(z(k)) \tag{35}$$

is such that $\lim_{k\to\infty} z(k) = \bar{z} \in \mathrm{fix}(\mathcal{T})$. □

Lemma 6: Time-varying Banach–Krasnoselskij iteration. Assume that: (i) $\exists \bar{k} \in \mathbb{N}$ s.t. $\bigcap_{k\geq\bar{k}} \mathrm{fix}(\mathcal{T}_k) \neq \varnothing$; (ii) $\exists \eta \in (0, 1)$ s.t. $(\forall k \in \mathbb{N})$ $\mathcal{T}_k : \mathbb{R}^n \to \mathbb{R}^n$ is η-averaged; (iii) $\alpha \in (0, 1]$. Then the iteration

$$z(k + 1) := (1 - \alpha)\, z(k) + \alpha\, \mathcal{T}_k(z(k)) \tag{36}$$

is such that $\lim_{k\to\infty} z(k) = \bar{z} \in \bigcap_{k\geq\bar{k}} \mathrm{fix}(\mathcal{T}_k)$. □

Proof: It follows from [34, Proposition 3.4 (iii)], since every cluster point of the sequence $(z(k))_{k=\bar{k}}^{\infty}$ is in $\bigcap_{k\geq\bar{k}} \mathrm{fix}(\mathcal{T}_k)$. ■

Definition 6: Monotone mapping. A set-valued mapping $\mathcal{A} : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is (strictly) *monotone* (MON) if

$$(u - v)^\top (x - y) \geq 0\, (> 0)$$

for all $x \neq y \in \mathrm{dom}(\mathcal{A})$, $(u, v) \in \mathcal{A}(x) \times \mathcal{A}(y)$. $\mathcal{A}$ is ϵ-*strongly monotone*, with $\epsilon \in \mathbb{R}_{>0}$, if $\mathcal{A} - \epsilon\,\mathrm{Id}$ is monotone. $\mathcal{A}$ is *maximally monotone* if $(\forall (x, u) \in \mathrm{gph}(\mathcal{A}))$ $(x, u) \in \mathrm{gph}(\mathcal{A}) \Leftrightarrow (\forall (y, v) \in \mathrm{gph}(\mathcal{A}))$ $(u - v)^\top (x - y) \geq 0$. □

Definition 7: Resolvent operator. The resolvent operator of a set-valued mapping $\mathcal{A} : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is the mapping $\mathcal{J}_{\mathcal{A}} : \mathbb{R}^n \to \mathbb{R}^n$ defined as $\mathcal{J}_{\mathcal{A}} := (\mathrm{Id} + \mathcal{A})^{-1}$. □

Lemma 7 ([31, Theorem 25.10, Remark 25.10]*): Tseng splitting algorithm*. Assume that: (i) $\mathrm{zer}(\mathcal{A} + \mathcal{B}) \neq \varnothing$; (ii) $\mathcal{A} : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is maximally monotone; (iii) $\mathcal{B} : \mathbb{R}^n \to \mathbb{R}^n$ is ℓ-Lipschitz continuous and monotone; (iv) $\alpha \in (0, 1/\ell)$. Then the iteration

$$\begin{aligned} \tilde{z}(k) &= z(k) - \alpha\, \mathcal{B}(z(k)) \\ z(k+1/2) &= \mathcal{J}_{\alpha\mathcal{A}}(\tilde{z}(k)) \\ \tilde{z}(k+1/2) &= z(k+1/2) - \alpha\, \mathcal{B}(z(k+1/2)) \\ z(k + 1) &= \tilde{z}(k+1/2) + \alpha\, \mathcal{B}(z(k)) \end{aligned}$$

is such that $\lim_{k\to\infty} z(k) = \bar{z} \in \mathrm{zer}(\mathcal{A} + \mathcal{B})$. □

B. Application to multi-agent networks

In this subsection, we analyse the mappings that arise in multi-agent network games under an operator theoretic lens, to provide some preliminary results for the main proofs in Section VII.

Lemma 8: If the matrix P satisfies Assumption 1, then the mappings $P\,\cdot$ and $\boldsymbol{A}\,\cdot = (P \otimes I_n)\,\cdot$ are η-averaged. □

Proof: By Schur complement, the linear matrix inequality in (8) is equivalent to $P^\top P \preccurlyeq (2\eta - 1)I + (1 - \eta)(P^\top + P)$. Therefore, the linear mapping $P\cdot$ is η-AVG by Definition 5, as well as $\boldsymbol{A}\cdot$. ■

Lemma 9: If the matrix P satisfies Assumption 2, then the mappings $P\,\cdot$ and $\boldsymbol{A}\,\cdot = (P \otimes I_n)\,\cdot$ are nonexpansive, and $(\forall \alpha \in (0, 1))$ the mappings $((1 - \alpha)I_N + \alpha P)\,\cdot$ and $((1 - \alpha)I_{nN} + \alpha \boldsymbol{A})\,\cdot$ are α-averaged. □

Proof: $\|P\| \leq 1$ implies that $P\,\cdot$ and $\boldsymbol{A}\,\cdot$ are NE by Definition 4. Then $((1 - \alpha)I_N + \alpha P)\,\cdot$ and $((1 - \alpha)I_{nN} + \alpha \boldsymbol{A})\,\cdot$ are α-AVG by Definition 5. ■

We can now prove Proposition 2.

Proof (Proposition 2): Let $\eta := 1 - \underline{a}$, where $\underline{a} := \min_{i\in\mathbb{N}[1,N]} a_{i,i} \in (0, 1)$. By Definition 5, $P\cdot$ is η-AVG if and only if $P = (1 - \eta)I_N + \eta B$ for some matrix $B = [b_{i,j}]$ such that $B\cdot$ is NE. Specifially, $(\forall i)$ $b_{i,i} = (a_{i,i} - (1 - \eta))/\eta$, and $(\forall j \neq i)$ $b_{i,j} = a_{i,j}/\eta$. Thus, $\sum_{j=1}^{N} b_{i,j} = \frac{1}{\eta}(a_{i,1} + \ldots + a_{i,N} - 1 + \eta) = 1$, and $\sum_{i=1}^{N} b_{i,j} = \frac{1}{\eta}(a_{1,j} + \ldots + a_{N,j} - 1 + \eta) = 1$. Next, since $\underline{a} \in (0, 1)$, $\eta \in (0, 1)$ and $(\forall i)$ $b_{i,i} \geq 0$. Therefore, B is doubly stochastic, and by Proposition 3 and Lemma 9, $B\cdot$ is NE. Since $P\cdot$ is η-AVG, then $\boldsymbol{A}\cdot$ is η-AVG as well due to Definition 5. ■

Lemma 10: The mappings $\mathbf{prox}_{\boldsymbol{f}}$ in (4) and $\boldsymbol{\mathcal{F}}$ in (26) are $\frac{1}{2}$-averaged, hence strictly monotone. □

Proof: It follows from [31, Proposition 12.27, Example 20.5] and Definition 5. ■

Lemma 11: The mapping $\mathrm{Id} - \boldsymbol{\mathcal{G}}$ from (27) is monotone. □

Proof: By [4, Lemma 3], $\mathrm{Id} - \boldsymbol{\mathcal{G}}$ is MON if and only if $2\,I - (\boldsymbol{G}^\top + \boldsymbol{G}) \succcurlyeq 0$, which is equivalent to $2I_N - (P + P^\top) \succcurlyeq 0$. The latter holds due to the Gershgorin circle theorem. ■

C. Non-strict monotonicity of the pseudo-gradient game mapping

To conclude the section, we show and discuss the monotonicity of the pseudo-gradient game mapping, and refer to solution algorithms for monotone variational inequalities.

Let us consider the Nash game associated with the network game in (2), that is, the game with best responses

$$x^i \in \underset{y \in \tilde{\mathcal{X}}^i(\boldsymbol{x}^{-i})}{\operatorname{argmin}} \; J^i\left(y,\, a_{i,i}\, y + \textstyle\sum_{j \neq i}^N a_{i,j}\, x^j\right), \tag{37}$$

for all $i \in \mathbb{N}[1, N]$. The pseudo-gradient game mapping $\Theta : (\mathbb{R}^n)^N \to (\mathbb{R}^n)^N$ is defined as the matrix of partial subgradients [27, p. 197], i.e., $\Theta = [\theta_{i,j}]$ where $(\forall i, j \in \mathbb{N}[1, N])$

$$\theta_{i,j}(\boldsymbol{x}) := \partial_{x^j} J^i\left(x^i,\, \textstyle\sum_{h=1}^N a_{i,h}\, x^h\right),$$

hence in view of J^i in (3), we have

$$\Theta = \operatorname{diag}\left((\partial f^i)_{i=1}^N\right) + ((I_N - P) \otimes I_n)\,\cdot\,. \tag{38}$$

Proposition 5: Monotone pseudo-gradient game mapping. The mapping Θ in (38) is monotone. □

Proof: It is sufficient to show that Θ is the sum of two MON mappings [31, p. 351]. First, the mapping $\operatorname{diag}\left((\partial f^i)_{i=1}^N\right)$ is MON [31, Example 20.3, Proposition 20.3]. Then, it follows from [4, Lemma 3] that the mapping $((I_N - P) \otimes I_n)\,\cdot$ is MON if and only if $I_N - P + (I_N - P)^\top = 2I_N - (P + P^\top) \succcurlyeq 0$, which holds true due to the Gershgorin circle theorem. ■

Since the pseudo-gradient game mapping Θ is monotone, solution algorithms for monotone variational inequalities [28, §12], [30, Part II, §2–4] are applicable to derive Nash equilibrium seeking dynamics, under the assumption of convex differentiable cost functions, convex local constraints and separable, convex differentiable coupling constraints.

While in the literature the pseudo-gradient game mapping is typically assumed to be strictly MON [15, Assumption 2], [16, Assumption 3], the mapping Θ in (38) is MON, but not strictly/strongly MON, nor cocoercive (that is, $\frac{1}{2}$-AVG under positive scaling). Therefore, projected gradient dynamics [28, §12] cannot ensure convergence in general, see [28, Example 12.1.3] for an example with non-convergent dynamics.

VII. Main proofs

Proof of Theorem 1: By Lemma 8, $\boldsymbol{A}\cdot$ in (5) is η-AVG and by Lemma 10, the mapping $\mathbf{prox}_{\boldsymbol{f}}$ in (4) is $\frac{1}{2}$-AVG. We note that the composition of AVG mappings is an AVG mapping itself [34, Proposition 2.5], specifically $\frac{1}{2-\eta}$-AVG. Then convergence follows by Lemma 4. Since $\mathbf{prox}_{\boldsymbol{f}} \circ \boldsymbol{A}$ is $\frac{1}{2-\eta}$-AVG, $(\forall \bar{\boldsymbol{x}} \in \operatorname{fix}\left(\mathbf{prox}_{\boldsymbol{f}} \circ \boldsymbol{A}\right))$, $\left\|\boldsymbol{x}(h+1) - \bar{\boldsymbol{x}}\right\|^2 \leq \left\|\boldsymbol{x}(h) - \bar{\boldsymbol{x}}\right\|^2 - (1 - \eta)\left\|\boldsymbol{x}(h+1) - \boldsymbol{x}(h)\right\|^2$. If we sum over h, then we have that $(k+1)\left\|\boldsymbol{x}(k+1) - \boldsymbol{x}(k)\right\|^2 \leq \sum_{h=0}^k \left\|\boldsymbol{x}(h+1) - \boldsymbol{x}(h)\right\|^2 \leq \frac{1}{1-\eta}\left\|\boldsymbol{x}(0) - \bar{\boldsymbol{x}}\right\|^2$. The inequality in (9) then follows since $\operatorname{dist}\left(\boldsymbol{x}(k), \operatorname{fix}\left(\mathbf{prox}_{\boldsymbol{f}} \circ \boldsymbol{A}\right)\right) \leq \left\|\left(\mathbf{prox}_{\boldsymbol{f}} \circ \boldsymbol{A}\right)(\boldsymbol{x}(k)) - \boldsymbol{x}(k)\right\|^2 = \left\|\boldsymbol{x}(k+1) - \boldsymbol{x}(k)\right\|^2$. ■

Proof of Theorem 2: By Lemma 9, the mapping $\boldsymbol{A}\cdot$ in (5) is NE. Thus, the mapping $\mathbf{prox}_{\boldsymbol{f}} \circ \boldsymbol{A}$ is the composition of two NE mappings, hence it is NE itself, and convergence holds by Lemma 5. Analogously to the proof of Theorem 1, we exploit the fact that the mapping $(1-\alpha)\mathrm{Id} + \alpha\left(\mathbf{prox}_{\boldsymbol{f}} \circ \boldsymbol{A}\right)$ is α-averaged, that is, $(\forall \bar{\boldsymbol{x}} \in \operatorname{fix}\left(\mathbf{prox}_{\boldsymbol{f}} \circ \boldsymbol{A}\right))$, $\left\|\boldsymbol{x}(h+1) - \bar{\boldsymbol{x}}\right\|^2 \leq \left\|\boldsymbol{x}(h) - \bar{\boldsymbol{x}}\right\|^2 - \frac{1-\alpha}{\alpha}\left\|\boldsymbol{x}(h+1) - \boldsymbol{x}(h)\right\|^2$. If we sum over h, then we have that $(k+1)\left\|\boldsymbol{x}(k+1) - \boldsymbol{x}(k)\right\|^2 \leq \sum_{h=0}^k \left\|\boldsymbol{x}(h+1) - \boldsymbol{x}(h)\right\|^2 \leq \frac{\alpha}{1-\alpha}\left\|\boldsymbol{x}(0) - \bar{\boldsymbol{x}}\right\|^2$. The inequality in (12) then follows since $\operatorname{dist}\left(\boldsymbol{x}(k), \operatorname{fix}\left(\mathbf{prox}_{\boldsymbol{f}} \circ \boldsymbol{A}\right)\right) \leq \left\|\left(\mathbf{prox}_{\boldsymbol{f}} \circ \boldsymbol{A}\right)(\boldsymbol{x}(k)) - \boldsymbol{x}(k)\right\|^2 = \alpha^2\left\|\boldsymbol{x}(k+1) - \boldsymbol{x}(k)\right\|^2$. ■

Proof of Theorem 3: Convergence follows by applying Lemma 6 with $\mathcal{T}_k := \mathbf{prox}_{\boldsymbol{f}}\left(\boldsymbol{A}(k)\,\cdot\right)$. Then the proof is analogous to the proof of Theorem 1, since, for all $k \in \mathbb{N}$, $\mathcal{T}_k$ is $\frac{1}{2-\eta}$-AVG by Definition 5 and [34, Proposition 2.5]. ■

Proof of Theorem 4: Convergence follows by applying Lemma 6 with $\mathcal{T}_k := (1-\alpha)\mathrm{Id} + \alpha\,\mathbf{prox}_{\boldsymbol{f}}\left(\boldsymbol{A}(k)\,\cdot\right)$. Then the proof is analogous to the proof of Theorem 2, since, for all $k \in \mathbb{N}$, $\mathcal{T}_k$ is α-AVG. ■

Proof of Corollary 1: It follows from Theorems 3, 4 that $\boldsymbol{x}(k) \to \bar{\boldsymbol{x}}$ and, since $\alpha > 0$, that $\mathbf{prox}_{\boldsymbol{f}}\left(\boldsymbol{A}(k)\,\bar{\boldsymbol{x}}(k)\right) \to \bar{\boldsymbol{x}}$. Due to Assumption 6, there exists a subsequence indexed by $h \in \mathcal{H} \subseteq \mathbb{N}$, $\{a_{i,j}(h)\}_{h \in \mathcal{H}}^\infty$, such that $\lim_{h\to\infty,\, h \in \mathcal{H}} \boldsymbol{A}(h) = \bar{\boldsymbol{A}}$. Thus, we have that $\lim_{h\to\infty,\, h\in\mathcal{H}} \boldsymbol{A}(h)\,\boldsymbol{x}(h) = \bar{\boldsymbol{A}}\,\bar{\boldsymbol{x}}$. We conclude that $\bar{\boldsymbol{x}} = \lim_{k\to\infty} \mathbf{prox}_{\boldsymbol{f}}\left(\boldsymbol{A}(k)\,\boldsymbol{x}(k)\right) = \lim_{h\to\infty,\, h\in\mathcal{H}} \mathbf{prox}_{\boldsymbol{f}}\left(\boldsymbol{A}(h)\,\boldsymbol{x}(h)\right) = \mathbf{prox}_{\boldsymbol{f}}\left(\bar{\boldsymbol{A}}\,\bar{\boldsymbol{x}}\right)$. ■

Proof of Theorem 5: It follows by applying Lemma 7 with $\mathcal{A} := \boldsymbol{\mathcal{J}}_{\boldsymbol{\mathcal{F}}}$ and $\mathcal{B} := \mathrm{Id} - \boldsymbol{\mathcal{G}}$. $\mathcal{A}$ is $\frac{1}{2}$-AVG and MON by Lemma 10; $\mathcal{B}$ is affine, hence Lipschitz continuous, and MON by Lemma 11. ■

VIII. Applications

A. Opinion dynamics in social networks

Opinion dynamics in social networks have been modeled in the context of multi-agent network games [8], [9], [35]. In this subsection, we build upon this literature and conceive opinion dynamics as multi-agent proximal dynamics, possibly multi-dimensional, interdependent, locally constrained, with possibly time-varying social interactions.

We consider N agents, where each agent $i \in \mathbb{N}[1, N]$ has a vector of opinions on $n \geq 1$ topics. Specifically, for all i, we consider $x^i \in \mathcal{X}^i \subseteq [0, 1]^n$, where $x^i_\tau = 0$ represents the most negative opinion of agent i on the topic $\tau \in \mathbb{N}[1, n]$, $x^i_\tau = 1$ represents the most positive opinion of agent i on the topic τ, and $x^i \in \mathcal{X}^i$ can represent limitations on the opinion of agent i on individual topics and also across the n topics.

For each agent i, we consider an initial opinion vector $x^i(0) \in \mathcal{X}^i$ and an ideal opinion $\hat{x}^i \in \mathcal{X}^i$, which is the opinion

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCNS.2017.2754358, IEEE Transactions on Control of Network Systems

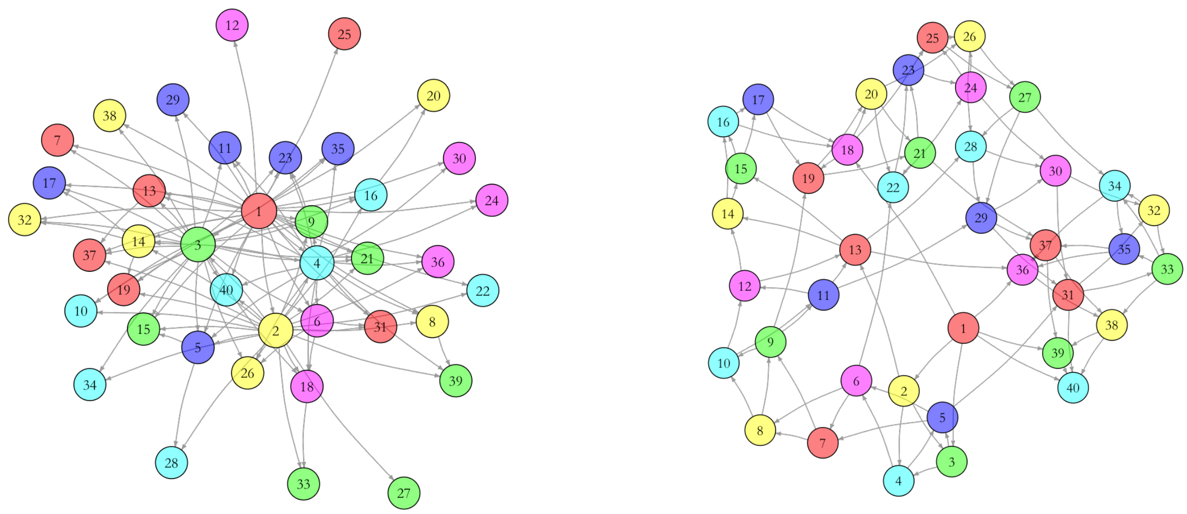

Fig. 1. Samples of Barabási–Albert scale-free (left) and Watts–Strogatz small-world (right) networks (with Fruchterman-Reingold layout) used in the numerical experiments on social networks. Self loops are not shown for ease of visualization.

the agent would reach without interactions with other agents, and which need not to be equal to $x^i(0)$. Then, we model the opinion dynamics as ($\forall i \in \mathbb{N}[1,N]$, $\forall k \in \mathbb{N}$)

$$\begin{aligned} &x^i(k+1) \\ &= \underset{y \in \mathcal{X}^i}{\operatorname{argmin}}\ f_\circ^i(y - \hat{x}^i) + \tfrac{1}{2}\left\| y - \textstyle\sum_{j=1}^N a_{i,j}(k)\, x^j(k) \right\|^2 \\ &= \operatorname{prox}_{f_\circ^i(\,\cdot\, - \hat{x}^i) + \iota_{\mathcal{X}^i}}\left(\textstyle\sum_{j=1}^N a_{i,j}(k)\, x^j(k) \right). \end{aligned} \tag{39}$$

The cost function in (39) has two addends: $f_\circ^i(y - \hat{x}^i)$ penalizes the deviation of the local opinion from the ideal opinion $\hat{x}^i$, namely, it weights the stubbornness of agent i, while the term $\frac{1}{2}\left\| y - \sum_{j=1}^N a_{i,j}(k)\, x^j(k) \right\|^2$ penalizes the deviation of the local opinion from the weighted average among the opinions of the neighboring agents, possibly including the current local opinion $x^i(k)$ as a memory effect if $a_{i,i}(k) > 0$. Therefore, in the extreme case that $f_\circ^i = 0$ and $\mathcal{X}^i = [0,1]^n$, agent i is typically referred as follower, while in the other extreme case that $a_{i,j}(k) = 0$ for all $j \neq i$ and $k \in \mathbb{N}$, or $\mathcal{X}^i = \{\hat{x}^i\}$, agent i can be referred as fully stubborn.

If we assume that ($\forall i \in \mathbb{N}[1,N]$) $f_\circ^i$ is lower semi-continuous, convex and positive semi-definite, $\mathcal{X}^i$ is nonempty, compact and convex, $P(k) = [a_{i,j}(k)]$ is doubly stochastic for all $k \in \mathbb{N}$, and that a persistent network equilibrium exists, then convergence to an NWE follows from the results in Section IV for the time-varying communication case, or in Section II for the time-invariant one.

Remark 3: Generalized opinion dynamics. The multi-agent network game model in (39) reduces to the model in [35] if $f_\circ^i(y - \hat{x}^i) = \frac{1}{2}\left\| y - x^i(0) \right\|_{Q_i}^2$ for some $Q_i \succcurlyeq 0$, $x^i(0) = \hat{x}^i$, $P(k) = [a_{i,j}(k)] = [a_{i,j}] = P$ for all k; also, in [35] it holds that $\mathcal{X}^i = \mathbb{R}^n$ for all i. If $n = 1$, $\mathcal{X}^i = [0,1]$ and $f_\circ^i(y - \hat{x}^i) = \theta^i \frac{1}{2}\left(y - x^i(0)\right)^2$ for some $\theta^i \geq 0$, then the model in (39) reduces to the Friedkin–Johnsen and De Groot models in [36], [37], [8]; furthermore, in [37] it is assumed that all agents are followers, and in [8] it is assumed that there exists i such that $\theta^i > 0$ and that $P(k) = P = [a_{i,j}] = [a_{j,i}] = P^\top$ for all k. □

Next, we investigate numerically and illustrate opinion dynamics on directed time-invariant graphs that have the topology of Barabási–Albert and Watts-Strogatz scale-free networks, with the addition of self loops of random weight - Figure 1 shows two examples of such networks. We consider row stochastic, but not doubly stochastic, weighted adjacency matrices to explore the behavior of opinion dynamics beyond our theoretical guarantees.

We run several numerical experiments to compare the distributed Banach dynamics in (7) on Barabási–Albert and Watts-Strogatz networks for different numbers of agents, $N \in \mathbb{N}[10,40]$, and $n = 2$ topics. For each experiment, for all $i \in \mathbb{N}[1,N]$, we sample an ideal opinion $\hat{x}^i$ and an initial one $x^i(0)$ with uniform distribution from the set $\{[0;0],[0;1],[1;0],[1;1]\}$ and the set $[0,1]^2$, respectively; we impose a polyhedral constraint set $\mathcal{X}^i = \{y \in [0,1]^n \mid \mathbf{1}^\top y \leq \theta_1^i\}$, and a piecewise-affine convex cost function $f_\circ^i(y) = \max\{0, \theta_2^i \left\| y - \hat{x}^i \right\|_1 - \theta_3^i\}$, where the triple $\left(\theta_1^i, \theta_2^i, \theta_3^i\right)$ is sampled uniformly in the set $[0,2] \times [1/2,1] \times [0,1/2]$.

Experiments illustrated in Figure 2 suggest that the convergence speed is not much affected by the network size.

B. Distributed tertiary control in power networks

Tertiary control in power networks has been considered as the solution to the optimal economic dispatch problem in terms of maximal social welfare [1], [2], [25], [26].

In this subsection, we formulate the problem as a multi-agent network game, where the agents (generators, flexible storage and loads) communicate among neighbors to reach a network equilibrium.

We assume that each agent $i \in \mathbb{N}[1,N]$ can decide on its variable $x^i = u^i$, which is the controllable power injection $u^i \in \mathcal{X}^i := [\underline{u}_i, \overline{u}^i]$ at bus i, to minimize its operating cost $f_\circ^i(u^i) := \frac{\theta^i}{2}\left(u^i\right)^2$, for some $\theta^i \in [0,1/2]$.

We consider the typical security constraints $|\delta^i - \delta^j| \leq \gamma := \pi/4$ that limit the power flow on each branch $(i,j) \in \mathcal{E}$

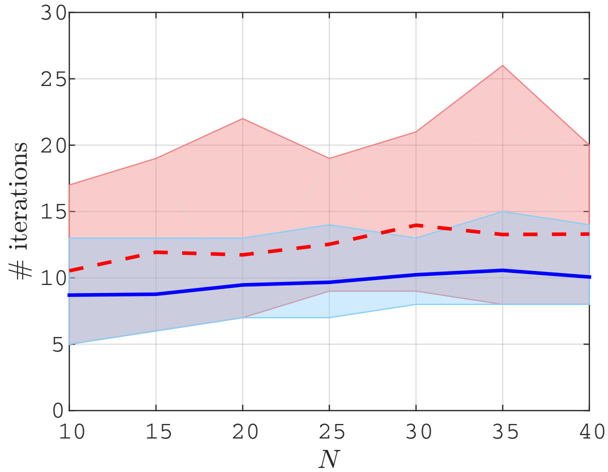

Fig. 2. Average number of iterations required for convergence ($\|\boldsymbol{x}(k+1) - \boldsymbol{x}(k)\|_\infty \leq 10^{-4}$) versus network size N, for Barabási–Albert (blue, solid line) and Watts–Strogatz (red, dashed line) networks, together with max-min intervals (shaded areas).

[1, §II-D], where δ^i is the phase angle deviation at bus i. In compact form, we have $-\boldsymbol{\gamma} \leq \boldsymbol{E}\boldsymbol{\delta} \leq \boldsymbol{\gamma}$, for some matrix $\boldsymbol{E} \in \{-1, 1\}^{|\mathcal{E}| \times N}$, where $\boldsymbol{\delta} := [\delta^1 ; \ldots ; \delta^N]$ and $\boldsymbol{\gamma} := \gamma\, \mathbf{1}_{|\mathcal{E}|}$.

As in [1, Equation 30], we relate the phase angles and the power injections via the linearized DC injection equation $u^i + \hat{u}^i = \sum_{j=1}^N \beta_{i,j} (\delta^i - \delta^j)$ for all i, where $\hat{u}^i$ is the nominal injection for agent i and $\beta_{i,j}$ is the effective susceptance of the edge (i, j). In compact form, we can write $\boldsymbol{B}\boldsymbol{\delta} = \boldsymbol{u} + \hat{\boldsymbol{u}}$, for some matrix $\boldsymbol{B} \in \mathbb{R}^{N \times N}$, $\boldsymbol{u} := [u^1 ; \ldots ; u^N]$ and $\hat{\boldsymbol{u}} := [\hat{u}^1 ; \ldots ; \hat{u}^N]$, hence derive the approximation $\boldsymbol{\delta} \simeq \boldsymbol{B}^\dagger (\boldsymbol{u} + \hat{\boldsymbol{u}})$. In addition, we want the net power balance to be zero or relatively small, e.g. not higher than some threshold $\epsilon \geq 0$, for the stabilization of the network frequency [2, §2]. The approximated power network constraints then read as the set of affine coupling constraints

$$\begin{array}{rcl} -\boldsymbol{\gamma} \leq & \boldsymbol{E}\,\boldsymbol{B}^\dagger (\boldsymbol{u} + \hat{\boldsymbol{u}}) & \leq \boldsymbol{\gamma} \\ 0 \leq & \mathbf{1}_N^\top (\boldsymbol{u} + \hat{\boldsymbol{u}}) & \leq \epsilon. \end{array}$$

We run some numerical experiments on the IEEE New England test power network, shown in Figure 3, with susceptance parameters obtained from the Power Systems Toolbox. We set one undirected communication link between generators 1 and 7, namely, nodes 39 and 36 of the physical graph, and simulate a scenario where the imbalance is generated as follows: the largest power injection, at node 39, is reduced by 20% and the largest load demand, at node 1, is increased by 20%, compared to their nominal values.

We compare the power imbalance under the Banach dynamics (for the mapping $\boldsymbol{\mathcal{F}} \circ \boldsymbol{\mathcal{G}}$) from (7), the Krasnoselskij dynamics from (11) and the Tseng protocol dynamics from (28)–(31) in several experiments. In our numerical experience, the Banach dynamics have the fastest convergence, although their global convergence is not supported theoretically. Figure 4 shows a representative simulation.

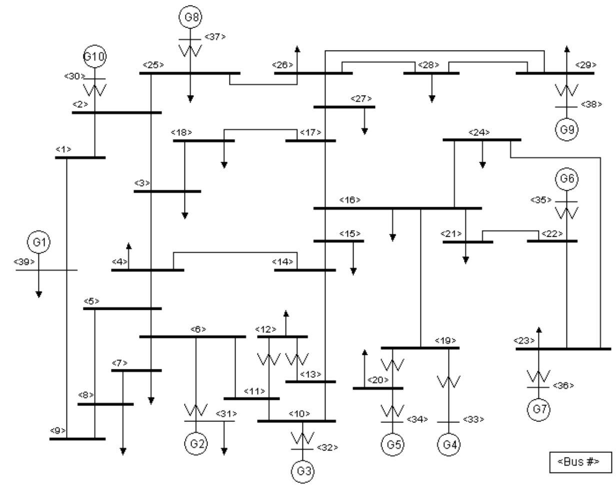

Fig. 3. IEEE New England test power network.

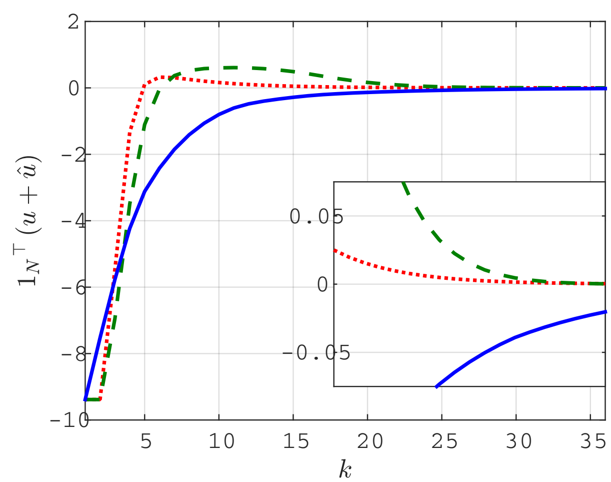

Fig. 4. Power imbalance regulated via the Banach dynamics (red dotted line), the Krasnoselskij dynamics (green dashed line) and the Tseng protocol dynamics (blue solid line).

IX. Conclusion and Outlook

Global convergence of selected classes of equilibrium seeking proximal dynamics hold in multi-agent network games, under the assumption of convex cost functions with proximal quadratic coupling, time-invariant and time-varying communication along with convex local constraints, time-invariant communication along with convex local constraints and separable coupling constraints.

More generally, equilibrium seeking dynamics for multi-agent network games with the simultaneous presence of convex cost functions, time-varying communication and convex (possibly non-differentiable) coupling constraints are currently unexplored. The analysis of equilibrium seeking dynamics for multi-agent network games with communication graph that is possibly different from the interference graph and with coupling constraints would be a relevant extension to this paper. We have modeled the information exchange between agents via a linear mapping and assumed non-expansiveness,

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCNS.2017.2754358, IEEE Transactions on Control of Network Systems

which holds true if the adjacency matrix is doubly stochastic. However, we believe that our convergence results can be extended to multi-agent network games with row-stochastic adjacency matrix, and also, under appropriate regularity assumptions, to nonlinear information exchange mappings.

References

[1] F. Dörfler, J. Simpson-Porco, and F. Bullo, "Breaking the hierarchy: Distributed control and economic optimality in microgrids," *IEEE Trans. on Control of Network Systems*, vol. 3, no. 3, pp. 241–253, 2016.

[2] F. Dörfler and S. Grammatico, "Gather-and-broadcast frequency control in power systems," *Automatica*, vol. 79, pp. 296–305, 2017.

[3] A.-H. Mohsenian-Rad, V. Wong, J. Jatskevich, R. Schober, and A. Leon-Garcia, "Autonomous demand-side management based on game-theoretic energy consumption scheduling for the future smart grid," *IEEE Trans. on Smart Grid*, vol. 1, no. 3, pp. 320–331, 2010.

[4] S. Grammatico, F. Parise, M. Colombino, and J. Lygeros, "Decentralized convergence to Nash equilibria in constrained deterministic mean field control," *IEEE Trans. on Automatic Control*, vol. 61, no. 11, pp. 3315–3329, 2016.

[5] S. Grammatico, "Exponentially convergent decentralized charging control for large populations of plug-in electric vehicles," in *Proc. of the IEEE Conf. on Decision and Control*, Las Vegas, USA, 2016.

[6] R. Jaina and J. Walrand, "An efficient Nash-implementation mechanism for network resource allocation," *Automatica*, vol. 46, pp. 1276–1283, 2010.

[7] J. Barrera and A. Garcia, "Dynamic incentives for congestion control," *IEEE Trans. on Automatic Control*, vol. 60, no. 2, pp. 299–310, 2015.

[8] J. Ghaderi and R. Srikant, "Opinion dynamics in social networks with stubborn agents: Equilibrium and convergence rate," *Automatica*, vol. 50, pp. 3209–3215, 2014.

[9] S. R. Etesami and T. Başar, "Game-theoretic analysis of the hegselmann-krause model for opinion dynamics in finite dimensions," *IEEE Trans. on Automatic Control*, vol. 60, no. 7, pp. 1886–1897, 2015.

[10] R. Olfati-Saber and R. Murray, "Consensus problems in networks of agents with switching topology and time-delays," *IEEE Trans. on Automatic Control*, vol. 49, no. 9, pp. 1520–1533, 2004.

[11] R. Olfati-Saber, "Flocking for multi-agent dynamic systems: Algorithms and theory," *IEEE Trans. on Automatic Control*, vol. 51, no. 3, pp. 401–420, 2006.

[12] S. Martínez, F. Bullo, J. Cortés, and E. Frazzoli, "On synchronous robotic networks – Part i: Models, tasks, and complexity," *IEEE Trans. on Automatic Control*, vol. 52, pp. 2199–2213, 2007.

[13] M. Stanković, K. Johansson, and D. Stipanović, "Distributed seeking of Nash equilibria with applications to mobile sensor networks," *IEEE Trans. on Automatic Control*, vol. 57, no. 4, pp. 904–919, 2012.

[14] F. Parise, B. Gentile, S. Grammatico, and J. Lygeros, "Network aggregative games: Distributed convergence to Nash equilibria," in *Proc. of the IEEE Conference on Decision and Control*, Osaka, Japan, 2015, pp. 2295–2300.

[15] J. Koshal, A. Nedić, and U. Shanbhag, "Distributed algorithms for aggregative games on graphs," *Operations Research*, vol. 64, no. 3, pp. 680–704, 2016.

[16] F. Salehisadaghiani and L. Pavel, "Distributed Nash equilibrium seeking: A gossip-based algorithm," *Automatica*, vol. 72, pp. 209–216, 2016.

[17] H. Yin, U. Shanbhag, and P. Mehta, "Nash equilibrium problems with scaled congestion costs and shared constraints," *IEEE Trans. on Automatic Control*, vol. 56, no. 7, pp. 1702–1708, 2011.

[18] S. Grammatico, "Aggregative control of competitive agents with coupled quadratic costs and shared constraints," in *Proc. of the IEEE Conf. on Decision and Control*, Las Vegas, USA, 2016.

[19] G. Belgioioso and S. Grammatico, "Semi-decentralized Nash equilibrium seeking in aggregative games with coupling constraints and non-differentiable cost functions," *IEEE Control Systems Letters*, vol. 1, no. 2, pp. 400–405, 2017.

[20] S. Grammatico, "An incentive mechanism for agents playing competitive aggregative games," in *Proc. of the Int. Conf. on Network Games, Control and Optimization*, Avignon, France, November 2016, pp. 113–122.

[21] ——, "Dynamic control of agents playing aggregative games with coupling constraints," *IEEE Trans. on Automatic Control*, vol. 62, no. 9, pp. 4537–4548, 2017.

[22] A. Nedić, A. Ozdaglar, and P. Parrillo, "Constrained consensus and optimization in multi-agent networks," *IEEE Trans. on Automatic Control*, vol. 55, no. 4, pp. 922–938, 2010.

[23] S. Lee and A. Nedić, "Distributed random projection algorithm for convex optimization," *IEEE Journal of Selected Topics in Signal Processing*, vol. 7, no. 2, pp. 221–229, 2013.

[24] A. Falsone, K. Margellos, S. Garatti, and M. Prandini, "Dual decomposition for multi-agent distributed optimization with coupling constraints," *Automatica*, vol. 84, pp. 149–158, 2017.

[25] C. Zhao, U. Topcu, N. Li, and S. Low, "Design and stability of load-side primary frequency control in power systems," *IEEE Trans. on Automatic Control*, vol. 59, no. 5, pp. 1177–1189, 2014.

[26] C. Zhao, E. Mallada, and F. Dörfler, "Distributed frequency control for stability and economic dispatch in power networks," in *Proc. of the IEEE American Control Conference*, 2015, pp. 2359–2364.

[27] F. Facchinei and C. Kanzow, "Generalized Nash equilibrium problems," *A Quarterly Journal of Operations Research, Springer*, vol. 5, pp. 173–210, 2007.

[28] F. Facchinei and J. Pang, *Finite-dimensional variational inequalities and complementarity problems*. Springer Verlag, 2003.

[29] D. Palomar and Y. Eldar, *Convex optimization in signal processing and communication*. Cambridge University Press, 2010.

[30] R. Cominetti, F. Facchinei, and J. Lasserre, *Modern optimization modelling techniques*. Birkhäuser, 2010.

[31] H. H. Bauschke and P. L. Combettes, *Convex analysis and monotone operator theory in Hilbert spaces*. Springer, 2010.

[32] D. R. Smart, *Fixed point theorems*. Cambridge University Press Archive, 1974.

[33] A. Auslender and M. Teboulle, "Lagrangian duality and related multiplier methods for variational inequality problems," *SIAM Journal on Optimization*, vol. 10, no. 4, pp. 1097–1115, 2000.

[34] P. L. Combettes and I. Yamada, "Compositions and convex combinations of averaged nonexpansive operators," *Journal of Mathematical Analysis and Applications*, pp. 55–70, 2015.

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

[36] N. E. Friedkin and E. C. Johnsen, "Social influence networks and opinion change," *Advances in group processes*, vol. 16, no. 1, pp. 1–29, 1999.

[37] M. H. DeGroot, "Reaching a consensus," *Journal of the American Statistical Association*, vol. 69, no. 345, pp. 118–121, 1974.

Sergio Grammatico (M'16) was born in Marsala, Italy, in 1987. He received the B.Sc. degree in Computer Engineering, the M.Sc. and Ph.D. degrees in Automation Engineering, in 2008, 2009 and 2013, respectively, all from the University of Pisa, Italy. He also received a M.Sc. degree in Engineering Science from the Sant'Anna School of Advanced Studies, the Italian Graduate School of Excellence in Applied Sciences, Pisa, Italy, in 2011. He visited the Department of Mathematics, University of Hawai'i at Manoa in 2010 and 2011, and the Department of Electrical and Computer Engineering at UC Santa Barbara in 2012. From 2013 to 2015, he was a post-doctoral Research Fellow in the Automatic Control Laboratory, ETH Zurich, Zurich, Switzerland. From 2015 to 2017, he was an Assistant Professor in the Control Systems group at the Department of Electrical Engineering, TU Eindhoven, The Netherlands. Since 2017, he is an Assistant Professor at the Delft Center for Systems and Control, TU Delft, The Netherlands. His research interests include large-scale multi-agent systems, stochastic and game-theoretic control. Sergio Grammatico was awarded "TAC Outstanding Reviewer" in 2014 and 2013 by the Editorial Board of the IEEE TRANSACTIONS ON AUTOMATIC CONTROL and was the recipient of the Best Paper Award at the INTERNATIONAL CONFERENCE ON NETWORK GAMES, CONTROL AND OPTIMIZATION 2016.