Nash Equilibrium Seeking with Non-doubly Stochastic Communication Weight Matrix

Farzad Salehisadaghiani and Lacra Pavel(✉)

Department of Electrical and Computer Engineering, University of Toronto,
10 King's College Road, Toronto, ON M5S 3G4, Canada
farzad.salehisadaghiani@mail.utoronto.ca, pavel@control.utoronto.ca

Abstract. A distributed Nash equilibrium seeking algorithm is presented for networked games. We assume an incomplete information available to each player about the other players' actions. The players communicate over a strongly connected digraph to send/receive the estimates of the other players' actions to/from the other local players according to a gossip communication protocol. Due to asymmetric information exchange between the players, a non-doubly (row) stochastic weight matrix is defined. We show that, due to the non-doubly stochastic property, there is no exact convergence. Then, we present an almost sure convergence proof of the algorithm to a Nash equilibrium of the game. Moreover, we extend the algorithm for graphical games in which all players' cost functions are only dependent on the local neighboring players over an interference digraph. We design an assumption on the communication digraph such that the players are able to update all the estimates of the players who interfere with their cost functions. It is shown that the communication digraph needs to be a superset of a transitive reduction of the interference digraph. Finally, we verify the efficacy of the algorithm via a simulation on a social media behavioral case.

1 Introduction

The problem of finding a Nash equilibrium (NE) of a networked game has recently drawn many attentions. The players who participate in this game aim to minimize their own cost functions selfishly by making decision in response to other players' actions. Each player in the network has only access to local information of the neighbors. Due to the imperfect information available to players, they maintain an estimate of the other players' actions and communicate over a communication graph in order to exchange the estimates with local neighbors. Using this information, players update their actions and estimates.

In many algorithms in the context of NE seeking problems, it is assumed that the communications between the players are symmetric in the sense that the players who are in communication can exchange their information altogether and update their estimates at the same time. This, in general, leads to a doubly

This work was supported by an NSERC Discovery Grant.

L. Duan et al. (Eds.): GameNets 2017, LNICST 212, pp. 3–15, 2017.
DOI: 10.1007/978-3-319-67540-4_1

stochastic communication weight matrix which preserves the global average of the estimates over time. However, there are many real-world applications in which symmetric communication is not of interest or may be an undesired feature in applications such as sensor network.

Literature review. Our work is related to the literature on Nash games and distributed NE seeking algorithms [1,4,11,16,17]. A distributed algorithm is proposed in [18] to compute a generalized NE of the game for a complete communication graph. In [7], an algorithm is provided to find an NE of *aggregative games* for a partial communication graph but complete interference graph. This algorithm is extended by [12] for a more general class of games in which the players' cost functions does not necessarily depend on the aggregate of players' actions. It is further generalized for the partial interference graph in [13]. For a *two-network zero-sum game* [5] considers a distributed algorithm for NE seeking. To find distributed algorithms for games with local-agent utility functions, a methodology is presented in [8] based on state-based *potential games*.

Gossip-based communication has been widely used in synchronous and asynchronous algorithms in consensus and distributed optimization problems [2,3,9]. In [9], a gossip algorithm is designed for a distributed broadcast-based optimization problem. An almost-sure convergence is provided due to the non-doubly stochasticity of the communication matrix. In [2], a broadcast gossip algorithm is studied to compute the average of the initial measurements which is proved to converge almost surely to a consensus.

Contributions. We propose an asynchronous gossip-based algorithm to find an NE of a distributed game over a communication digraph. We assume that players send/receive information to/from their local out/in-neighbors over a strongly connected communication digraph. Players update their own actions and estimates based on the received information. We prove an almost sure convergence of the algorithm to the NE of the game. *Unlike in the undirected case* [12,13], *herein we cannot exploit the doubly stochastic property for the communication weight matrix due to asymmetric information exchange. Non-doubly stochastic property leads to have total average of the players' estimates not preserved over time. This was one of the critical steps in the convergence proof in* [12,13].

Moreover, we extend the algorithm for graphical games in which the players' cost functions may be interfered by any subset of players' (not necessarily all the players') actions. The locality of cost functions is specified by an interference digraph which marks the pair of players who interfere one with another. In order to have a convergent algorithm, we design an assumption on the communication digraph by which there exists a lower bound on the communication digraph which is a transitive reduction of the interference digraph. By this assumption, it is proved that all the players are able to exchange and update all the estimates of the actions interfering with their cost functions.

The proofs are omitted due to space limitations, and are available in [14].

2 Problem Statement: Game with a Complete Interference Digraph

Consider a multi-player game in a network with a set of players V. The interference of players' actions on the cost functions is represented by a complete *interference digraph* $G(V, E)$, with E marking the pair of players that interfere one with another. Note that for a complete digraph every pair of distinct nodes is connected by a pair of unique edges (one in each direction).

The game is denoted by $\mathcal{G}(V, \Omega_i, J_i)$ and defined over

- $V = \{1, \ldots, N\}$: Set of players,
- $\Omega_i \subset \mathbb{R}$: Action set of player i, $\forall i \in V$ with $\Omega = \prod_{i \in V} \Omega_i \subset \mathbb{R}^N$ the action set of all players,
- $J_i : \Omega \to \mathbb{R}$: Cost function of player i, $\forall i \in V$,

In the following we define a few notations for players' actions.

- $x = (x_i, x_{-i}) \in \Omega$: All players actions,
- $x_i \in \Omega_i$: Player i's action, $\forall i \in V$ and $x_{-i} \in \Omega_{-i} := \prod_{j \in V \setminus \{i\}} \Omega_j$: All other players' actions except i.

The game is defined as a set of N simultaneous optimization problems as follows:

$$\begin{cases} \underset{y_i}{\text{minimize}} & J_i(y_i, x_{-i}) \\ \text{subject to} & y_i \in \Omega_i \end{cases} \quad \forall i \in V. \tag{1}$$

Each problem is run by an individual player and its solution is dependent on the solution of the other problems. The objective is to find an NE of this game which is defined as follows:

Definition 1. *Consider an N-player game $\mathcal{G}(V, \Omega_i, J_i)$, each player i minimizing the cost function $J_i : \Omega \to \mathbb{R}$. A vector $x^* = (x_i^*, x_{-i}^*) \in \Omega$ is called an NE of this game if*

$$J_i(x_i^*, x_{-i}^*) \leq J_i(x_i, x_{-i}^*) \quad \forall x_i \in \Omega_i, \ \forall i \in V. \tag{2}$$

We state a few assumptions for the existence and the uniqueness of an NE.

Assumption 1. *For every $i \in V$,*

- *Ω_i is non-empty, compact and convex,*
- *$J_i(x_i, x_{-i})$ is C^1 in x_i, continuous in x and convex in x_i for every x_{-i}.*

The compactness of Ω implies that $\forall i \in V$ and $x \in \Omega$,

$$\|\nabla_{x_i} J_i(x)\| \leq C, \quad \text{for some } C > 0. \tag{3}$$

Let $F : \Omega \to \mathbb{R}^N$, $F(x) := [\nabla_{x_i} J_i(x)]_{i \in V}$ be the pseudo-gradient vector of the cost functions (game map).

Assumption 2. *F is strictly monotone,* $(F(x)-F(y))^T(x-y)>0 \quad \forall x,y\in \Omega,\ x\neq y$.

Assumption 3. *$\nabla_{x_i}J_i(x_i,u)$ is Lipschitz continuous in x_i, for every fixed $u\in \Omega_{-i}$ and for every $i\in V$, i.e., there exists $\sigma_i>0$ such that*

$$\|\nabla_{x_i}J_i(x_i,u)-\nabla_{x_i}J_i(y_i,u)\|\leq \sigma_i\|x_i-y_i\| \quad \forall x_i,y_i\in \Omega_i.$$

Moreover, $\nabla_{x_i}J_i(x_i,u)$ is Lipschitz continuous in u with a Lipschitz constant $L_i>0$ for every fixed $x_i\in\Omega_i$, $\forall i\in V$.

In game (1), the only information available to each player i is J_i and Ω. Thus, each player maintains an estimate of the other players actions and exchanges those estimates with the neighbors to update them. A *communication digraph* $G_C(V,E_C)$ is defined where $E_C\subseteq V\times V$ denotes the set of communication links between the players. $(i,j)\in E_C$ if and only if player i sends his information to player j. Note that $(i,j)\in E_C$ does not necessarily imply $(j,i)\in E_C$. The set of in-neighbors of player i in G_C, denoted by $N_C^{\text{in}}(i)$, is defined as $N_C^{\text{in}}(i):=\{j\in V|(j,i)\in E_C\}$. The following assumption on G_C is used.

Assumption 4. *G_C is strongly connected.*

Our objective is to find an algorithm for computing an NE of $\mathcal{G}(V,\Omega_i,J_i)$ using only imperfect information over the communication digraph $G_C(V,E_C)$.

3 Asynchronous Gossip-Based Algorithm

We propose a projected gradient-based algorithm using an asynchronous gossip-based method as in [12]. The algorithm is inspired by [12] except that the communications are supposed to be directed in a sense that the information exchange is considered over a directed path. Our challenge here is to deal with the asymmetric communications between the players. This makes the convergence proof dependent on a *non-doubly stochastic weight matrix*, whose properties need to be investigated and proved. The algorithm is elaborated as follows:

1- ***Initialization Step:*** Each player i maintains an initial *temporary* estimate $\tilde{x}^i(0)\in\Omega$ for all players. Let $\tilde{x}^i_j(0)\in\Omega_j\subset\mathbb{R}$ be player i's initial temporary estimate of player j's action, for $i,j\in V$.

2- ***Gossiping Step:*** At iteration k, player i_k becomes active uniformly at random and selects a communication in-neighbor indexed by $j_k\in N_C^{\text{in}}(i_k)$ uniformly at random. Let $\tilde{x}^i(k)\in\Omega\subset\mathbb{R}^N$ be player i's temporary estimate at iteration k. Then player j_k sends his temporary estimate $\tilde{x}^{j_k}(k)$ to player i_k. After receiving the information, player i_k constructs his final estimate of all players. Let $\hat{x}^i_j(k)\in\Omega_j\subset\mathbb{R}$ be player i's final estimate of player j's action, for $i,j\in V$. The final estimates are computed as in the following:

1. Players i_k's final estimate:

$$\begin{cases} \hat{x}^{i_k}_{i_k}(k) = \tilde{x}^{i_k}_{i_k}(k) \\ \hat{x}^{i_k}_{-i_k}(k) = \frac{\tilde{x}^{i_k}_{-i_k}(k)+\tilde{x}^{j_k}_{-i_k}(k)}{2}. \end{cases} \tag{4}$$

   Note that $\tilde{x}^i_i(k) = x_i(k)$ for all $i \in V$.
2. For all other players $i \neq i_k$, the temporary estimate is maintained, i.e.,

$$\hat{x}^i(k) = \tilde{x}^i(k), \quad \forall i \neq i_k. \tag{5}$$

We use communication weight matrix $W(k) := [w_{ij}(k)]_{i,j\in V}$ to obtain a compact form of the gossip protocol. $W(k)$ is *a non-doubly (row) stochastic weight matrix* defined as follows:

$$W(k) = I_N - \frac{e_{i_k}(e_{i_k} - e_{j_k})^T}{2}, \tag{6}$$

where $e_i \in \mathbb{R}^N$ is a unit vector. Note that $W(k)$ is different from the doubly stochastic one used in [12]. The non-doubly (row) stochasticity of $W(k)$ is translated into:

$$W(k)\mathbf{1}_N = \mathbf{1}_N,\ \mathbf{1}^T_N W(k) \neq \mathbf{1}^T_N. \tag{7}$$

Let $\bar{x}(k) = [\bar{x}^1(k), \ldots, \bar{x}^N(k)]^T \in \Omega^N$ be an intermediary variable such that

$$\bar{x}(k) = (W(k) \otimes I_N)\tilde{x}(k), \tag{8}$$

where $\tilde{x}(k) = [\tilde{x}^1(k), \ldots, \tilde{x}^N(k)]^T \in \Omega^N$ is the overall temporary estimate at k. Using (6) one can combine (4) and (5) in a compact form of $\hat{x}^{i_k}_{-i_k}(k) = \bar{x}^{i_k}_{-i_k}(k)$ and $\hat{x}^i(k) = \bar{x}^i(k)$ for $\forall i \neq i_k$.

3- ***Local Step:*** At this moment all the players update their actions according to a projected gradient-based method. Let $\bar{x}^i = (\bar{x}^i_i, \bar{x}^i_{-i}) \in \Omega$, $\forall i \in V$ with $\bar{x}^i_i \in \Omega_i$ be the intermediary variable associated to player i. Because of imperfect information available to player i, he uses $\bar{x}^i_{-i}(k)$ and updates his action as follows: if $i = i_k$,

$$x_i(k+1) = T_{\Omega_i}[x_i(k) - \alpha_{k,i}\nabla_{x_i} J_i(x_i(k), \bar{x}^i_{-i}(k))], \tag{9}$$

otherwise, $x_i(k+1) = x_i(k)$. In (9), $T_{\Omega_i} : \mathbb{R} \to \Omega_i$ is an Euclidean projection and $\alpha_{k,i}$ are diminishing step sizes such that $\sum_{k=1}^{\infty} \alpha^2_{k,i} < \infty$, $\sum_{k=1}^{\infty} \alpha_{k,i} = \infty \ \forall i \in V$. The players use their updated actions to update their temporary estimates as follows:

$$\tilde{x}^i(k+1) = \bar{x}^i(k) + (x_i(k+1) - \bar{x}^i_i(k))e_i, \quad \forall i \in V. \tag{10}$$

At this point, the players are ready to begin a new iteration from step 2. We elaborate on the non-doubly stochasticity of $W(k)$ from two perspectives.

1. **Design:** By the row (non-doubly) stochastic property of $W(k)$, the temporary estimates remain at consensus subspace once they reach there. This can be verified by (8) when $\tilde{x}(k) = \mathbf{1}_N \otimes \boldsymbol{\alpha}$ for an $N \times 1$ vector $\boldsymbol{\alpha}$, since,

$$\bar{x}(k) = (W(k) \otimes I_N)(\mathbf{1}_N \otimes \boldsymbol{\alpha}) = \mathbf{1}_N \otimes \boldsymbol{\alpha}. \tag{11}$$

   Equations (9), (10) and (11) imply that the consensus is maintained. On the other hand $W(k)$ *is not column-stochastic* which is a critical property used in [12]. This implies that the average of temporary estimates is not equal to the average of $\bar{x}$. Indeed by (8),

$$\frac{1}{N}(\mathbf{1}_N^T \otimes I_N)\bar{x}(k) = \frac{1}{N}(\mathbf{1}_N^T \otimes I_N)(W(k) \otimes I_N)\tilde{x}(k) \neq \frac{1}{N}(\mathbf{1}_N^T \otimes I_N)\tilde{x}(k). \tag{12}$$

   Equations (9), (10) *and* (12) *imply that the average of temporary estimates is not preserved for the next iteration. Thus, it is infeasible to obtain an exact convergence to the average consensus* [2]. Instead, we show an almost sure (a.s.) convergence of the temporary estimates to an average consensus[1].
2. **Convergence Proof:** $\lambda_{\max}(W(k)^T W(k))$ is a key parameter in the proof (as in [9,12]). Unlike [12], the non-doubly stochastic property of $W(k)^T W(k)$ ends up in having $\lambda_{\max}(W(k)^T W(k)) > 1$. We resolve this issue in Lemma 1.

4 Convergence for Diminishing Step Sizes

In this section we prove convergence of the algorithm for diminishing step sizes. Consider a memory in which the history of the decision making is recorded. Let $\mathcal{M}_k$ denote the *sigma-field* generated by the history up to time $k-1$ with

$$\mathcal{M}_0 = \{\tilde{x}^i(0),\ i \in V\}.\ \mathcal{M}_k = \mathcal{M}_0 \cup \Big\{(i_l, j_l); 1 \leq l \leq k-1\Big\}, \quad \forall k \geq 2. \tag{13}$$

As explained in the design challenge in Sect. 3, we consider a.s. convergence. Convergence is shown in two parts. First, we prove a.s. convergence of the temporary estimate vectors $\tilde{x}^i$, to an average consensus, proved to be the vectors' average. Then we prove a.s. convergence of players' actions toward an NE.

Let $\tilde{x}(k)$ be the overall temporary estimate vector. The average of all temporary estimates at $T(k)$ is defined as follows:

$$Z(k) = \frac{1}{N}(\mathbf{1}_N^T \otimes I_N)\tilde{x}(k). \tag{14}$$

As mentioned in Sect. 3, the major difference between the proposed algorithm and the one in [12] is in using a non-doubly stochastic weight matrix $W(k)$ which was a key step. The following lemma is used to overcome these challenges.

[1] The same objective is followed by [9] to find a broadcast gossip algorithm (with non-doubly stochastic weight matrix) in the area of distributed optimization. However, in the proof of Lemma 2 ([9] page 1348) which is mainly dedicated to this discussion, the doubly stochasticity of $W(k)$ is used right after Eq. (22) which violates the main assumption on $W(k)$.

Lemma 1. *Let* $Q(k) = (W(k) - \frac{1}{N}\mathbf{1}_N\mathbf{1}_N^T W(k)) \otimes I_N$ *and* $W(k)$ *be a non-doubly (row) stochastic weight matrix defined in* (6) *which satisfies* (7). *Let also* $\gamma = \lambda_{\max}\big(\mathbb{E}[Q(k)^T Q(k)]\big)$. *Then* $\gamma < 1$.

Proof. See [14].

Theorem 1. *Let* $\tilde{x}(k)$ *be the stack vector with all temporary estimates of the players and* $Z(k)$ *be its average as in* (14). *Let also* $\alpha_{k,max} = \max_{i\in V}\alpha_{k,i}$. *Then under Assumptions 1, 4, the following hold.*

(i) $\sum_{k=0}^{\infty}\alpha_{k,max}\|\tilde{x}(k) - (\mathbf{1}_N \otimes I_N)Z(k)\| < \infty$ *a.s.,*
(ii) $\sum_{k=0}^{\infty}\|\tilde{x}(k) - (\mathbf{1}_N \otimes I_N)Z(k)\|^2 < \infty$ *a.s.*

Proof. The proof follows as in the proof of Theorem 1 in [12], but the critical step here is in using Lemma 1.

Corollary 1. *For the players' actions* $x(k)$ *and* $\bar{x}(k)$, *the following terms hold a.s. under Assumptions 1–4.*

(i) $\sum_{k=0}^{\infty}\alpha_{k,max}\|x(k) - Z(k)\| < \infty$ *a.s.,* *(ii)* $\sum_{k=0}^{\infty}\|x(k) - Z(k)\|^2 < \infty$ *a.s.,*
(iii) $\sum_{k=0}^{\infty}\mathbb{E}\Big[\|\bar{x}(k) - (\mathbf{1}_N \otimes I_N)Z(k)\|^2\Big|\mathcal{M}_k\Big] < \infty$ *a.s.*

Proof. See [14].

Theorem 2. *Let* $x(k)$ *and* x^* *be the players' actions and the NE of* $\mathcal{G}$, *respectively. Under Assumptions 3–4, the sequence* $\{x(k)\}$ *generated by the algorithm converges to* x^*, *almost surely.*

Proof. The proof is similar to the proof of Theorem 2 in [12] based on Theorem 1. Theorem 2 verifies that the actions of the players converge a.s. toward the NE using the fact that the actions converge to a consensus subspace (Corollary 1).

5 Game with a Partial Interference Digraph

We extend the game defined in Sect. 2 to the case with partially coupled cost functions, such that the cost functions may be interfered by the actions of any subset of players. The game is denoted by $\mathcal{G}(V, G_I, \Omega_i, J_i)$ where $G_I(V, E_I)$ is an interference digraph with E_I marking the players whose actions interfere with the other players' cost functions. We denote by $N_I^{\text{in}}(i) := \{j \in V | (j,i) \in E_I\}$, the set of in-neighbors of player i in G_I whose actions affect J_i and $\tilde{N}_I^{\text{in}}(i) := N_I^{\text{in}}(i) \cup \{i\}$.

The following assumption is considered for G_I.

Assumption 5. G_I *is strongly connected.*

The cost function of player i, J_i, $\forall i \in V$, is defined over $\Omega^i \to \mathbb{R}$ where $\Omega^i = \prod_{j\in\tilde{N}_I^{\text{in}}(i)}\Omega_j \subset \mathbb{R}^{|\tilde{N}_I^{\text{in}}(i)|}$ is the action set of players interfering with the cost function of player i. A few notations for players' actions are given:

- $x^i = (x_i, x^i_{-i}) \in \Omega^i$: All players' actions which interfere with J_i,
- $x^i_{-i} \in \Omega^i_{-i} := \prod_{j \in N_I^{\text{in}}(i)} \Omega_j$: Other players' actions which interfere with J_i.

Given x^i_{-i}, each player i aims to minimize his own cost function selfishly,

$$\begin{cases} \underset{y_i}{\text{minimize}} & J_i(y_i, x^i_{-i}) \\ \text{subject to} & y_i \in \Omega_i \end{cases} \quad \forall i \in V. \tag{15}$$

Known parameters to player i are as follows: (1) Cost function of player i, J_i and (2) Action set Ω^i. Note that this game setup is similar to the one in [13] except for a directed G_C used for asymmetric communications. Our first objective is to design an assumption on G_C such that all required information is communicated by the players after sufficiently many iterations. In other words, we ensure that player i, $\forall i \in V$ receives information from all the players whose actions interfere with his cost function.

Definition 2. *Transitive reduction: A digraph H is a transitive reduction of G which is obtained as follows: for all three vertices i, j, l in G such that edges (i, j), (j, l) are in G, (i, l) is removed from G.*

Note that transitive reduction is different from *maximal triangle-free spanning subgraph* which is used in Assumption 2 in [13].

Assumption 6. *The following holds for the communication graph G_C:*

- *$G_{TR} \subseteq G_C \subseteq G_I$, where G_{TR} is a transitive reduction of G_I.*

Lemma 2. *Let G_I and G_C satisfying Assumptions 5, 6. Then, $\forall i \in V$,*

$$\bigcup_{j \in N_C^{in}(i)} \left(N_I^{in}(i) \cap \tilde{N}_I^{in}(j) \right) = N_I^{in}(i). \tag{16}$$

Proof. See [14].

Remark 1. (16) verifies that using Assumptions 5, 6 the first objective is satisfied.

The assumptions for existence and uniqueness of an NE are Assumptions 1–3 with the cost functions adapted to G_I. Our second objective is to find an algorithm for computing an NE of $\mathcal{G}(V, G_I, \Omega_i, J_i)$ over $G_C(V, E_C)$ with partially coupled cost functions as described by the directed graph $G_I(V, E_I)$.

6 Asynchronous Gossip-Based Algorithm Adapted to G_I

The structure of the algorithm is similar to the one in Sect. 3. The steps are elaborated in the following:

1- ***Initialization Step:***
   – $\tilde{x}^i(0) \in \Omega^i$: Player i's initial temporary estimate.
2- ***Gossiping Step:***
   – $\tilde{x}^i_j(k) \in \Omega_j \subset \mathbb{R}$: Player i's temporary estimate of player j's action at k.
   – $\hat{x}^i_j(k) \in \Omega_j \subset \mathbb{R}$: Player i's final estimate of player j's action at k, for $i \in V$, $j \in \tilde{N}^{\text{in}}_I(i)$.
   – Final estimate construction:

$$\hat{x}^{i_k}_l(k) = \begin{cases} \frac{\tilde{x}^{i_k}_l(k)+\tilde{x}^{j_k}_l(k)}{2}, & l \in (N^{\text{in}}_I(i_k) \cap \tilde{N}^{\text{in}}_I(j_k)) \\ \tilde{x}^{i_k}_l(k), & l \in \tilde{N}^{\text{in}}_I(i_k) \backslash (N^{\text{in}}_I(i_k) \cap \tilde{N}^{\text{in}}_I(j_k)). \end{cases} \tag{17}$$

For

$$i \neq i_k,\ j \in \tilde{N}^{\text{in}}_I(i), \hat{x}^i_j(k) = \tilde{x}^i_j(k). \tag{18}$$

We suggest a compact form for gossip protocol by using $W^I(k)$.
Let for player i,

$$W^I(k) := I_m - \sum_{l \in (\tilde{N}^{\text{in}}_I(i_k) \cap \tilde{N}^{\text{in}}_I(j_k))} \frac{e_{s_{i_k l}}(e_{s_{i_k l}} - e_{s_{j_k l}})^T}{2}, \tag{19}$$

where $e_i \in \mathbb{R}^m$ is a unit vector. Note that $W^I(k)$ is different from the doubly stochastic one used in [13]. See [14] for the design of s_{ij} which is an index of the estimate vector element associated with player i's estimate of player j's action.
   – $\tilde{x}(k) := \left[\tilde{x}^{1^T}, \ldots, \tilde{x}^{N^T}\right]^T$: Stack vector of all temporary estimates,
   – $\bar{x}(k) := W^I(k)\tilde{x}(k)$: Intermediary variable.

Using the intermediary variable, one can construct the final estimates as follows:

$$\hat{x}^i_{-i}(k) = [\bar{x}_{s_{ij}}(k)]_{j \in N^{\text{in}}_I(i)}. \tag{20}$$

3- ***Local Step:*** Player i updates his action as follows: If $i = i_k$, $x_i(k+1) = T_{\Omega_i}\Big[x_i(k) - \alpha_{k,i} \nabla_{x_i} J_i\big(x_i(k), [\bar{x}_{s_{ij}}(k)]_{j \in N^{\text{in}}_I(i)}\big)\Big]$, otherwise,

$$x_i(k+1) = x_i(k), \tag{21}$$

Then he updates his temporary estimates:

$$\tilde{x}^i_j(k+1) = \begin{cases} \bar{x}_{s_{ij}}(k), & \text{if } j \neq i \\ x_i(k+1), & \text{if } j = i. \end{cases} \tag{22}$$

At this point, the players are ready to begin a new iteration from step 2.

7 Convergence of the Algorithm Adapted to G_I

Similar to Sect. 4, the convergence proof is split into two steps:

1. First, we prove a.s. convergence of $\tilde{x}(k) \subset \mathbb{R}^m$ to an average consensus which is shown to be the augmented average of all temporary estimate vectors. Let
   - $m_i^{\text{out}} := \deg_{G_I}^{\text{out}}(i) + 1$, where $\deg_{G_I}^{\text{out}}(i)$ is the out-degree of vertex i in G_I,
   - $1./\mathbf{m}^{\text{out}} := [\frac{1}{m_1^{\text{out}}}, \ldots, \frac{1}{m_N^{\text{out}}}]^T$,
   -
   $$H := [\sum_{i:1\in N_I^{\text{in}}(i)} e_{s_{i1}}, \ldots, \sum_{i:N\in N_I^{\text{in}}(i)} e_{s_{iN}}] \in \mathbb{R}^{m\times N}, \quad (23)$$

   where $i : j \in N_I^{\text{in}}(i)$ is all i's such that $j \in N_I^{\text{in}}(i)$. The *augmented average* of all temporary estimates is denoted by $Z^I(k) \in \mathbb{R}^m$ and defined as follows:

   $$Z^I(k) := H\text{diag}(1./\mathbf{m}^{\text{out}})H^T\tilde{x}(k) \in \mathbb{R}^m. \quad (24)$$

2. Secondly, we prove almost sure convergence of the players actions to an NE.

The proof depends on some key properties of W^I and H given in Lemmas 3, 4.

Lemma 3. *Let $W^I(k)$ and H be defined in (19) and (23). Then, $W^I(k)H = H$. This can be interpreted as the generalized row stochastic property of $W^I(k)$.*

Proof. See [14].

Lemma 4. *Let $Q^I(k) := W^I(k) - H\,diag(1./\boldsymbol{m}^{out})H^TW^I(k)$ and $\gamma^I = \lambda_{\max}\big(\mathbb{E}[Q^I(k)^TQ^I(k)]\big)$. Then $\gamma^I < 1$.*

Proof. See [14].

Theorem 3. *Let $\tilde{x}(k)$ be the stack vector with all temporary estimates of the players and $Z^I(k)$ be its average as in (24). Let also $\alpha_{k,max} = \max_{i\in V}\alpha_{k,i}$. Then under Assumptions 1′, 5, 6, the following hold.*
(i) $\sum_{k=0}^{\infty}\alpha_{k,max}\|\tilde{x}(k) - Z^I(k)\| < \infty$ a.s., (ii) $\sum_{k=0}^{\infty}\|\tilde{x}(k) - Z^I(k)\|^2 < \infty$ a.s.

Proof. The proof uses Lemmas 3, 4 and is similar to the proof Theorem 1 in [15].

Corollary 2. *Let $z^I(k) := diag(1./\boldsymbol{m}^{out})H^T\tilde{x}(k) \in \mathbb{R}^N$ be the average of all players' temporary estimates. Under Assumptions 1′, 5, 6 the following hold for players' actions $x(k)$ and $\bar{x}(k)$:*

(i) $\sum_{k=0}^{\infty}\alpha_{k,max}\|x(k) - z^I(k)\| < \infty$ a.s., (ii) $\sum_{k=0}^{\infty}\|x(k) - z^I(k)\|^2 < \infty$ a.s.,
(iii) $\sum_{k=0}^{\infty}\mathbb{E}\Big[\|\bar{x}(k) - Z^I(k)\|^2\Big|\mathcal{M}_k\Big] < \infty$ a.s.

Proof. See [14].

Theorem 4. *Let $x(k)$ and x^* be all players' actions and the NE of $\mathcal{G}$, respectively. Under Assumptions 1′–3′, 5, 6, the sequence $\{x(k)\}$ generated by the algorithm converges to x^*, almost surely.*

Proof. The proof uses Theorem 3 and is similar to the proof of Theorem 2 in [15].

8 Simulation Results

8.1 Social Media Behavior

In this example we aim to investigate social networking media for users' behavior. In such media like Facebook, Twitter and Instagram users are allowed to follow (or be friend with) the other users and post statuses, photos and videos or also share links and events. Depending on the type of social media, the way of communication is defined. For instance, in Instagram, friendship is defined unidirectional in a sense that either side could be only a follower and/or being followed. Recently, researchers at Microsoft have been studying the behavioral attitude of the users of Facebook as a giant and global network [10]. This study can be useful in many areas e.g. business (posting advertisements) and politics (posting for the purpose of presidential election campaign). Generating new status usually comes with the cost for the users such that if there is no benefit in posting status, the users don't bother to generate new ones. In any social media drawing others' attention is one of the most important motivation/stimulation to post status [6]. Our objective is to find the optimal rate of posting status for each user to draw more attention in his network. In the following, we make an information/attention model of a generic social media [6] and define a communication between users (G_C) and an interference graph between them (G_I).

Consider a social media network of N users. Each user i produces x_i unit of information that the followers can see in their news feeds. The users' communication network is defined by a strongly connected digraph G_C in which ⓘ→ⓙ means j is a follower of i or j receives x_i in his news feed. We also assume a strongly connected interference digraph G_I to show the influence of the users on the others. We assume that each user i's cost function is not only affected by the users he follows, but also by the users that his followers follow. The cost function of user i is denoted by J_i and consists of three parts: (1) $C_i(x_i) := h_i x_i$, $h_i > 0$ which is a cost that user i pays to produce x_i unit of information. (2) $f_i^1(x) := L_i \sqrt{\sum_{j \in N_C^{\text{in}}(i)} q_{ji} x_j}$, $L_i > 0$ which is a differentiable, increasing and concave utility function of user i from receiving information from his news feed with $f_i^1(\mathbf{0}) = 0$ and q_{ji} represents follower i's interest in user j's information and L_i is a user-specific parameter. (3) $f_i^2(x) := \sum_{l: i \in N_C^{\text{in}}(l)} L_l \left(\sqrt{\sum_{j \in N_C^{\text{in}}(l)} q_{jl} x_j} - \sqrt{\sum_{j \in N_C^{\text{in}}(l) \setminus \{i\}} q_{jl} x_j} \right)$ which is an incremental utility function that each user obtains from receiving attention in his network with $f_i^2(x)|_{x_i=0} = 0$. Specifically, this function targets the amount of attention that each follower pays to the information of other users in his news feed. The total cost function for user i is then $J_i(x) = C_i(x_i) - f_i^1(x) - f_i^2(x)$. For this example, we consider 5 users in the social media whose network of followers G_C is given in Fig. 1(a). From G_C and taking J_i into account, one can construct G_I (Fig. 1(b)) in a way that the interferences among users are specified. Note that this is a reverse process of the one discussed in Sect. 5 because G_C is given as the network of followers and G_I is constructed from G_C. For the particular networks in Fig. 1(a, b), Assumptions 5, 6 hold. We then employ the algorithm

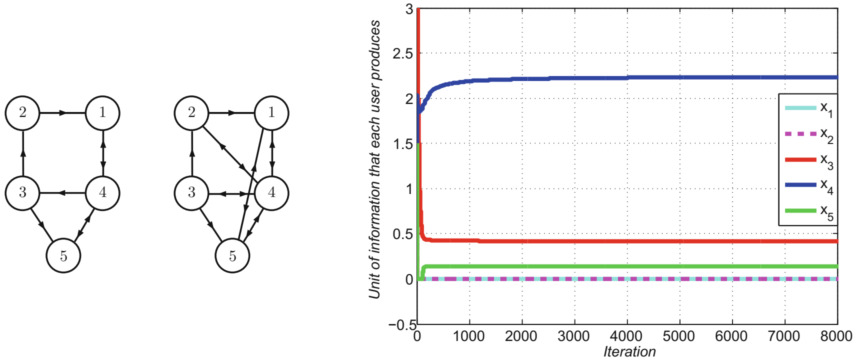

Fig. 1. (a) G_C (b) G_I (c) Convergence of the unit of information that each user produces to a NE over G_C.

in Sect. 6 to find an NE of this game for $h_i = 2$ and $L_i = 1.5$ $\forall i \in V$, and $q_{41} = q_{45} = 1.75$, $q_{32} = q_{43} = 2$ and the rest of $q_{ij} = 1$. The result is shown in Fig. 1(c). To analyze the NE $x^* = [0, 0, 0.42, 2.24, 0.14]^T$, one can realize from G_C that user 4 has 3 followers (users 1, 3 and 5), user 3 has 2 followers (users 2 and 5) and the rest has only 1 follower. Then, it is straightforward to predict that users 4 and 3 could draw more attentions and produce more information.

References

1. Alpcan, T., Başar, T.: Distributed algorithms for Nash equilibria of flow control games. In: Nowak, A.S., Szajowski, K. (eds.) Advances in Dynamic Games, pp. 473–498. Springer, Boston (2005). doi:10.1007/0-8176-4429-6_26
2. Aysal, T.C., Yildiz, M.E., Sarwate, A.D., Scaglione, A.: Broadcast gossip algorithms for consensus. Trans. Sig. Process. **57**(7), 2748–2761 (2009)
3. Fagnani, F., Zampieri, S.: Randomized consensus algorithms over large scale networks. IEEE J. Sel. Areas Commun. **26**(4), 634–649 (2008)
4. Frihauf, P., Krstic, M., Basar, T.: Nash equilibrium seeking in noncooperative games. IEEE Trans. Autom. Control **57**(5), 1192–1207 (2012)
5. Gharesifard, B., Cortes, J.: Distributed convergence to Nash equilibria in two-network zero-sum games. Automatica **49**(6), 1683–1692 (2013)
6. Goel, A., Ronaghi, F.: A game-theoretic model of attention in social networks. In: Bonato, A., Janssen, J. (eds.) WAW 2012. LNCS, vol. 7323, pp. 78–92. Springer, Heidelberg (2012). doi:10.1007/978-3-642-30541-2_7
7. Koshal, J., Nedic, A., Shanbhag, U.V.: A gossip algorithm for aggregative games on graphs. In: IEEE 51st Conference on Decision and Control, pp. 4840–4845 (2012)
8. Li, N., Marden, J.R.: Designing games for distributed optimization. IEEE J. Sel. Top. Sig. Process. **7**(2), 230–242 (2013)
9. Nedic, A.: Asynchronous broadcast-based convex optimization over a network. IEEE Trans. Autom. Control **56**(6), 1337–1351 (2011)
10. Olson, P.: Microsoft uses facebook as giant 'lab' to study game theory. Forbes (2011). http://www.forbes.com/sites/parmyolson/2011/10/12/microsoft-uses-facebook-as-giant-lab-to-study-game-theory

11. Pan, Y., Pavel, L.: Games with coupled propagated constraints in optical networks with multi-link topologies. Automatica **45**(4), 871–880 (2009)
12. Salehisadaghiani, F., Pavel, L.: Distributed Nash equilibrium seeking: a gossip-based algorithm. Automatica **72**, 209–216 (2016)
13. Salehisadaghiani, F., Pavel, L.: Distributed Nash equilibrium seeking by gossip in games on graphs. In: 2016 IEEE 55th Conference on Decision and Control (CDC), pp. 6111–6116. IEEE (2016)
14. Salehisadaghiani, F., Pavel, L.: Nash equilibrium seeking with non-doubly stochastic communication weight matrix. arXiv preprint (2016). arXiv:1612.07179
15. Salehisadaghiani, F., Pavel, L.: A distributed Nash equilibrium seeking in networked graphical games. arXiv preprint (2017). arXiv:1703.09765
16. Salehisadaghiani, F., Pavel, L.: Distributed Nash equilibrium seeking via the alternating direction method of multipliers. IFAC World Congress (to appear, 2017)
17. Yin, H., Shanbhag, U.V., Mehta, P.G.: Nash equilibrium problems with scaled congestion costs and shared constraints. IEEE Trans. Autom. Control **56**(7), 1702–1708 (2011)
18. Zhu, M., Frazzoli, E.: Distributed robust adaptive equilibrium computation for generalized convex games. Automatica **63**, 82–91 (2016)