

# A Gossip Algorithm for Aggregative Games on Graphs

Jayash Koshal, Angelia Nedić and Uday V. Shanbhag

**Abstract**—We consider a class of games, termed as aggregative games, being played over a distributed multi-agent networked system. In an aggregative game, an agent’s objective function is coupled through a function of the aggregate of all agents decisions. Every agent maintains an estimate of the aggregate and agents exchange this information over a connected network. We study the gossip-based distributed algorithm for information exchange and computation of equilibrium decisions of agents over the network. Our primary emphasis is on proving the convergence of the algorithm under an assumption of a diminishing (agent-specific) stepsize sequence. Under standard conditions, we establish the almost-sure convergence of the algorithm to an equilibrium point. Finally, we present numerical results to assess the performance of the gossip algorithm for aggregative games.

## I. INTRODUCTION

An aggregative game is a game where each agent’s objective is affected by its own action and the aggregate of the actions taken by all agents [1]. Nash-Cournot games represent one instance of such games; here firms make quantity bids that fetch a price based aggregate quantity sold, implying that the payoff function of any player is a function of aggregate quantity sold [2], [3]. This paper considers an aggregative game where agents are coupled by the aggregate of all agent decisions. The goal of this paper is to study a gossip algorithm for information exchange and equilibrium computation when the agents do not have a full access to the aggregate of their decisions. Instead, each agent may communicate only locally with his neighbors. The neighbor relations are captured by a network modeling the connectivity structure among the agents. Under this limited communication regime, it is not possible for any agent to have the knowledge of true aggregate and, thus, the computation of an equilibrium point is more involving.

In this work we propose a gossip-based algorithm which allows agents to build estimates and compute an equilibrium point of aggregative games. Importantly, this computation can be carried out in a distributed and asynchronous manner. In a gossip-based algorithm, an

agent randomly wakes up, selects a neighboring agent and performs a mutual averaging operation. As such, the algorithm is inherently distributed and asynchronous and allows the agents to use uncoordinated step sizes. When the network is connected, the algorithm converges to an equilibrium point almost surely for a (random) diminishing stepsize rule independently selected by each agent. The novelty of this work is in our examination of gossip-based algorithms for computation of a Nash equilibrium point while the majority of preceding efforts have been applied towards the solution of feasibility and optimization problems. Succinctly, the main contributions of this work lies in a gossip algorithm for aggregative Nash games and proving that it produces sequences that converge almost-surely to the unique equilibrium.

Our work in this paper is related to the literature on aggregative games [4], [5], and [6]. Distributed algorithms for computing equilibria have received significant attention recently [7], [8], [9], [10], [11]. Our work considers a class of aggregative Nash games for which regularized stochastic approximation schemes were developed for more generalized stochastic counterparts in [12]. More generally, the work in this paper is related to the gossip-based algorithms introduced in [13] [14]. Since then these algorithms have received significant attention from a wide range of applications. Distributed gossip-based schemes have been studied for distributed averaging with quantization constraints are proposed in [15], [16], and in optimization over network [17], [18], [19], [20]

The paper is organized as follows. In section II, we describe the problem of interest, propose a gossip-based algorithm, and state our assumptions. In section III, we study the convergence of the algorithm under a diminishing stepsize requirement. We present some numerical results in section IV and, finally, conclude in section V.

Throughout this paper, we view vectors as columns. We write  $x^T$  to denote the transpose of a vector  $x$ , and  $x^T y$  to denote the inner product of vectors  $x$  and  $y$ . We use  $\|x\| = \sqrt{x^T x}$  to denote the Euclidean norm of a vector  $x$ . We use  $\Pi_K$  to denote the Euclidean projection operator on a set  $K$ , i.e.  $\Pi_K(x) \triangleq \operatorname{argmin}_{z \in K} \|x - z\|$ . The expectation of a random variable  $Y$  is denoted by

Department of Industrial and Enterprise Systems Engineering, University of Illinois, Urbana IL 61801, Email: {koshal1,angelia,udaybag}@illinois.edu. Nedić and Shanbhag gratefully acknowledge the NSF support of this work through grant NSF CMMI 09-48905 EAGER ARRA.

$\mathbb{E}[Y]$  and *a.s.* denotes *almost surely*.

## II. PROBLEM FORMULATION AND BACKGROUND

In this section we introduce an aggregative game and set up the equilibrium problem of our interest. We also discuss the connectivity graph for the agents and present our gossip-based algorithm.

### A. Aggregative game

Let  $\mathcal{N} = \{1, \dots, N\}$  be a set of  $N$  agents indexed by  $1, \dots, N$ . The problem of interest requires agent  $i$  solving the following parameterized problem:

$$\begin{aligned} & \text{minimize} && f_i(x_i, \bar{x}) \\ & \text{subject to} && x_i \in K_i, \end{aligned} \quad (1)$$

where  $K_i \subseteq \mathbb{R}$  and the function  $f_i$  is of the form

$$f_i(x_i, \bar{x}) \equiv c_i(x_i) - x_i U(\bar{x}).$$

The term  $c_i(x_i)$  is the agent specific cost and  $U(\bar{x})$  captures the utility associated with aggregate output  $\bar{x}$ , which is given by

$$\bar{x} = \sum_{j=1}^N x_j = x_i + \sum_{\substack{j=1 \\ j \neq i}}^N x_j = x_i + \bar{x}_{-i} \quad \text{with } x_j \in K_j.$$

The cost  $c_i(x_i)$  and the set  $K_i$  are known only to agent  $i$ , while the function  $U(y)$  is known by all agents. We let  $\bar{K}$  denote the Minkowski sum of the sets  $K_i$ :

$$\bar{K} = \sum_{i=1}^N K_i,$$

and we make the following assumptions on the constraint sets  $K_i$  and the functions  $f_i$ .

*Assumption 1:* For  $i = 1, \dots, N$ , the set  $K_i \subseteq \mathbb{R}$  is compact and convex and the function  $f_i(x_i, y)$  is continuously differentiable in  $(x_i, y)$  over some open set containing the set  $K_i \times \bar{K}$ . Furthermore,  $f_i(x_i, \bar{x})$  is convex in  $x_i$  over the set  $K_i$  for every  $\bar{x} \in \bar{K}$ .

Under Assumption 1, the equilibrium conditions of the game in (1) can be characterized by a scalar variational inequality problem denoted by  $\text{VI}(K, F)$  where

$$F(x) \triangleq \begin{pmatrix} \nabla_{x_1} f_1(x_1, \bar{x}) \\ \vdots \\ \nabla_{x_N} f_N(x_N, \bar{x}) \end{pmatrix}, \quad K = \prod_{i=1}^N K_i, \quad (2)$$

with  $x \triangleq (x_1, \dots, x_N)^T$ ,  $\bar{x} \triangleq \sum_{i=1}^N x_i$ , and  $x_i \in K_i$  for all  $i$ . Note that, by Assumption 1, the set  $K$  is a compact and convex set in  $\mathbb{R}^N$ . The mapping  $F$  maps  $x \in K$  to  $F(x) \in \mathbb{R}^N$ . We use  $F_i$  to denote the  $i$ -th coordinate of the mapping  $F$ . To emphasize the particular form that  $F_i$

has, we will write  $F_i(x_i, \bar{x})$  for the value  $F_i(x)$ . Thus, with this notation, we have

$$F(x) \triangleq \begin{pmatrix} F_1(x_1, \bar{x}) \\ \vdots \\ F_N(x_N, \bar{x}) \end{pmatrix}, \quad (3)$$

where  $F_i(x_i, \bar{x}) = \nabla_{x_i} f_i(x_i, \bar{x})$ . We now state our assumptions on the mapping  $F_i$ , the  $i$ th coordinate mapping of  $F$  in (2).

*Assumption 2:* For each  $F_i(x_i, \bar{x})$  the following hold:

- (i) The mapping  $F_i(x_i, \bar{x})$  is Lipschitz continuous in  $x_i$ , for every fixed  $\bar{x} \in \bar{K}$  i.e., for some  $L_i > 0$  and for all  $x_i, y_i \in K_i$ ,

$$\|F_i(x_i, \bar{x}) - F_i(y_i, \bar{x})\| \leq L_i \|x_i - y_i\|;$$

- (ii) The mapping  $F_i(x_i, u)$  is Lipschitz continuous in  $u$ , for every fixed  $x_i \in K_i$  i.e., for some  $L_{-i} > 0$  and for all  $u, z \in \bar{K}$ ,

$$\|F_i(x_i, u) - F_i(x_i, z)\| \leq L_{-i} \|u - z\|.$$

It is relatively straightforward to see that such assumptions hold with little difficulty when considering Nash-Cournot games with affine price functions.

As a consequence of Assumption 2 it can be seen that the gradient map  $F$  is also Lipschitz over  $K$ , i.e.,

$$\|F(x) - F(y)\| \leq L \|x - y\| \quad \text{for all } x, y \in K, \quad (4)$$

where  $L = \sqrt{2 \sum_{i=1}^N (L_i^2 + N L_{-i}^2)}$ .

We assume that no central entity that can provide the aggregate  $\bar{x}$  of agent decisions  $x_i$  at any time. However, we allow agents to be connected in some manner so they can communicate locally with their neighbors. Under this restriction, *we are interested in designing an algorithm for computing an equilibrium of aggregative game* (1).

### B. Gossip algorithm

We propose a gossip-based algorithm for computing an equilibrium of aggregative game (1). In our algorithm, each agent maintains an estimate of the true aggregate  $\bar{x}$  while taking decisions within his constraint set that minimize his objective function. To estimate the aggregate  $\bar{x}$  of decisions, each agent communicates occasionally with his neighbors at which instance the exchange of their estimates takes place.

We model the agent connectivity structure by an undirected graph  $\mathcal{G}(\mathcal{N}, \mathcal{E})$ , with node  $i \in \mathcal{N}$  being agent  $i$  and  $\mathcal{E}$  being the set of undirected edges among the agents. When  $\{i, j\} \in \mathcal{E}$ , the agents  $i$  and  $j$  can talk to each other and exchange some information. Let  $\mathcal{N}_i$

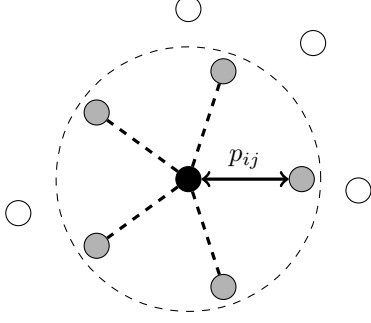


Fig. 1. A depiction of gossip communication

denote the set of agents who are immediate neighbors of agent  $i$ , i.e.,  $\mathcal{N}_i = \{j \mid \{i, j\} \in \mathcal{E}\}$ .

For the agents connectivity graph  $\mathcal{G}(\mathcal{N}, \mathcal{E})$  we use the following assumption.

*Assumption 3:* The undirected graph  $\mathcal{G}(\mathcal{N}, \mathcal{E})$  is connected.

We use a gossip protocol to model agent exchange of the estimates of the aggregate  $\bar{x}$ , which is varying in time as the agents decisions are changing. In the gossip-based information exchange, each agent is assumed to have a local clock which ticks according to a Poisson process with rate 1. Consequently, the agents' clocks process can be equivalently modeled as a single (virtual) clock which ticks according to Poisson process with rate  $N$ . Let  $Z^k$  denote  $k$ th tick time of the global Poisson process. We discretize the time so that time  $k$  corresponds to the time-slot  $[Z^{k-1}, Z^k)$ . At time  $k$ , we assume that every agent  $i$  has his iterate  $x_i^k$  and an estimate  $v_i^k$  of the current aggregate. We assume that only one agent wakes up at time  $k$  and we denote the index of such an agent by  $I^k$ . We let agent  $I^k$  contact agent  $J^k$  with probability  $p_{ij}$ , where  $J^k$  is an immediate neighbor of agent  $I^k$ , i.e.,  $J^k \in \mathcal{N}_{I^k}$ . At this time,  $I^k$  and  $J^k$  exchange their estimates  $v_{I^k}^k$  and  $v_{J^k}^k$  of the aggregate variable and compute intermittent variables as follows:

$$\hat{v}_i^k = \frac{v_{I^k}^k + v_{J^k}^k}{2} \quad \text{for } i \in \{I^k, J^k\}, \quad (5)$$

whereas other agents do nothing, i.e.,

$$\hat{v}_i^k = v_i^k \quad \text{for } i \notin \{I^k, J^k\}.$$

To capture this step, we define the weight matrix  $W(k)$ :

$$W(k) = I - \frac{1}{2}(e_{I^k} - e_{J^k})(e_{I^k} - e_{J^k})^T, \quad (6)$$

where  $e_i$  is  $N$ -dimensional vector with  $i$ th entry equal to 1 and 0 otherwise. By using  $W(k)$  we can rewrite

the intermittent estimate computation, as follows:

$$\hat{v}_i^k = \sum_{j=1}^N [W(k)]_{ij} v_j^k \quad \text{and} \quad v_i^0 = x_i^0 \quad \forall i = 1, \dots, N,$$

where  $x_i^0 \in K_i$  are initial agent decisions.

Using this aligned aggregate estimate  $\hat{v}_i^k$  and his own iterate  $x_i^k$ , agent  $i$  updates its iterate and average estimate as follows:

$$x_i^{k+1} = (\Pi_{K_i} [x_i^k - \alpha_{k,i} F_i(x_i^k, N\hat{v}_i^k)] - x_i^k) \mathbb{1}_{\{i \in \{I^k, J^k\}\}} + x_i^k \quad (7)$$

$$v_i^{k+1} = \hat{v}_i^k + x_i^{k+1} - x_i^k, \quad (8)$$

where,  $\mathbb{1}_{\{i \in \{I^k, J^k\}\}}$  denotes the indicator of the event  $\{i \in \{I^k, J^k\}\}$ ,  $\alpha_{k,i}$  is the stepsize for agent  $i$  and  $F_i(x_i, y) = \nabla_{x_i} f_i(x_i, y)$ .

Note that only agents  $i \in \{I^k, J^k\}$  update. If  $i \notin \{I^k, J^k\}$ , then  $\mathbb{1}_{\{i \in \{I^k, J^k\}\}} = 0$  and hence  $x_i^{k+1} = x_i^k$  and  $v_i^{k+1} = v_i^k$ . Letting  $p_i$  denote the probability of the event that agent  $i$  updates, i.e.  $\{i \in \{I^k, J^k\}\}$ , we have

$$p_i = \frac{1}{N} \left( 1 + \sum_{j \in \mathcal{N}_i} p_{ji} \right) \quad \text{for all } i \in \mathcal{N},$$

where  $p_{ji} > 0$  is the probability that agent  $i$  is contacted by his neighbor.

To take account for the history, we introduce  $\mathcal{F}_k$  to denote the  $\sigma$ -algebra generated by the entire history up to  $k$ . More precisely

$$\mathcal{F}_k = \mathcal{F}_0 \cup \{I^l, J^l; 1 \leq l \leq k\} \quad \text{for all } k \geq 1,$$

with  $\mathcal{F}_0 = \{x_i^0, i \in \mathcal{N}\}$ .

The stepsize update rule is defined by the frequency of agent updates, i.e.  $\alpha_{k,i} = \frac{1}{\Gamma_k(i)}$ , where  $\Gamma_k(i)$  denotes the number of updates that agent  $i$  has executed up to time  $k$  inclusively. Some long term estimates for  $\alpha_{k,i}$  that we use in our analysis are given in the following lemma (cf. [20], Lemma 3).

*Lemma 1:* Let Assumption 3 hold and let  $p_{\min} = \min_{i \in \mathcal{N}} p_i$ . Let  $\alpha_{k,i} = 1/\Gamma_k(i)$  for all  $k$  and  $i$ . Also let  $q \in (0, 1/2)$ . Then there is a large enough  $\tilde{k} = \tilde{k}(q, N)$  such that for all  $k \geq \tilde{k}$  and  $i \in \mathcal{N}$  a.s.,

$$\alpha_{k,i} \leq \frac{2}{kp_i}, \alpha_{k,i}^2 \leq \frac{4N^2}{k^2 p_{\min}}, \left| \alpha_{k,i} - \frac{1}{kp_i} \right| \leq \frac{2}{k^{3/2-q} p_{\min}^2}.$$

In our analysis we also use a well-known result on supermartingale convergence, which we provide in the following lemma for convenience. The result can be found in [21], Lemma 11, page 50.

*Lemma 2:* Let  $V_k, u_k, \beta_k$  and  $\gamma_k$  be non-negative random variables adapted to  $\sigma$ -algebra  $\mathcal{F}_k$ . If almost surely  $\sum_{k=0}^{\infty} u_k < \infty$ ,  $\sum_{k=0}^{\infty} \beta_k < \infty$ , and

$$\mathbb{E}[V_{k+1} | \mathcal{F}_k] \leq (1 + u_k)V_k - \gamma_k + \beta_k \quad \text{for all } k \geq 0,$$

then almost surely  $V_k$  converges and  $\sum_{k=0}^{\infty} \gamma_k < \infty$ .

### III. CONVERGENCE THEORY

In this section we establish the convergence of the gossip algorithm for the aggregative game in (1). We consider an agent specific diminishing stepsize of the form  $\alpha_{k,i} = \frac{1}{\Gamma_k(i)}$ . First we state some results which will allow us to claim the convergence of the algorithm. These results involve the average  $y^k$  of the estimates  $v_i^k, i \in \mathcal{N}$ , given by

$$y^k = \frac{1}{N} \mathbf{1}^T v^k \quad \text{for all } k \geq 0, \quad (9)$$

where  $\mathbf{1}$  is the vector in  $\mathbb{R}^N$  with all entries equal to 1.

We have the following result.

*Lemma 3:* Let  $\hat{v}^k \triangleq (\hat{v}_1^k, \dots, \hat{v}_N^k)$  and  $v^k \triangleq (v_1^k, \dots, v_N^k)$ , where  $\hat{v}_i^k$  and  $v^k$  are given respectively by (5) and (8). Then, we have

$$\sum_{k=1}^{\infty} \mathbb{E}[\|\hat{v}^k - y^k \mathbf{1}\|^2 | \mathcal{F}_{k-1}] < \infty \quad \text{a.s.}$$

*Proof:* We first bound the term  $\sum_{i=1}^N \|v_i^k - y^k\|^2$ . Using the weight matrix  $W(k)$  we can write the update of  $v^{k+1}$  in (8), as follows

$$v^{k+1} = W(k)v^k + x^{k+1} - x^k,$$

where  $v^k = (v_1^k, \dots, v_N^k)^T$  and  $x^k = (x_1^k, \dots, x_N^k)^T$ . From the decision update rule of (7), we have

$$v^{k+1} = W(k)v^k + \zeta^{k+1},$$

where  $\zeta^{k+1} = (\zeta_1^{k+1}, \dots, \zeta_N^{k+1})$  and  $\zeta_i^{k+1} = (\Pi_{K_i}[x_i^k - \alpha_{k,i} F_i(x_i^k, N\hat{v}_i^k)] - x_i^k) \mathbb{1}_{\{i \in \{I^k, J^k\}\}}$ . Since  $y^k$  is the average of entries of  $v^k$ , i.e. we have  $y^{k+1} = \frac{1}{N} (\mathbf{1}^T W(k)v^k + \mathbf{1}^T \zeta^{k+1})$  which leads us to

$$\begin{aligned} v^{k+1} - y^{k+1} \mathbf{1} &= \left( W(k) - \frac{1}{N} \mathbf{1} \mathbf{1}^T W(k) \right) v^k \\ &\quad + \left( I - \frac{1}{N} \mathbf{1} \mathbf{1}^T \right) \zeta^{k+1}, \end{aligned}$$

where  $I$  is the identity matrix. By construction  $W(k)$  is a stochastic matrix i.e.  $W(k)\mathbf{1} = \mathbf{1}$ . Thus  $(W(k) - \frac{1}{N} \mathbf{1} \mathbf{1}^T W(k)) \mathbf{1} = 0$ , implying that

$$\left( W(k) - \frac{1}{N} \mathbf{1} \mathbf{1}^T W(k) \right) y^k \mathbf{1} = 0.$$

Thus we have

$$v^{k+1} - y^{k+1} \mathbf{1} = D(k)(v^k - y^k \mathbf{1}) + M\zeta^{k+1}.$$

where  $D(k) = (W(k) - \frac{1}{N} \mathbf{1} \mathbf{1}^T W(k))$  and  $M = I - \frac{1}{N} \mathbf{1} \mathbf{1}^T$ . By taking the norm and the conditional expectation w.r.t.  $\mathcal{F}_k$  we obtain

$$\begin{aligned} \mathbb{E}[\|v^{k+1} - y^{k+1} \mathbf{1}\| | \mathcal{F}_k] &\leq \underbrace{\mathbb{E}[\|D(k)(v^k - y^k \mathbf{1})\| | \mathcal{F}_k]}_{\text{Term 1}} \\ &\quad + \underbrace{\mathbb{E}[\|M\zeta^{k+1}\| | \mathcal{F}_k]}_{\text{Term 2}}. \end{aligned}$$

We first estimate **Term 1**. Let  $\lambda = \|\mathbb{E}[D(k)^T D(k)]\| = \mathbb{E}[\|D(k)\|^2]$ . Then

$$\text{Term 1} \leq \sqrt{\lambda} \|v^k - y^k \mathbf{1}\|,$$

where we use the fact that  $v^k$  is adapted to  $\mathcal{F}_k$ . To estimate **Term 2**, we employ the spectral properties of  $M$  to obtain the following inequality:

$$\begin{aligned} \|M\zeta^{k+1}\|^2 &\leq \|\zeta^{k+1}\|^2 \\ &= \sum_{i \in \{I^k, J^k\}} \|\Pi_{K_i}[x_i^k - \alpha_{k,i} F_i(x_i^k, N\hat{v}_i^k)] - x_i^k\|^2. \end{aligned}$$

Using the non-expansive property of the projection operator and  $\alpha_{k,i}^2 \leq \frac{4N^2}{k^2 p_{\min}^2}$  (cf. Lemma 1), we have for all  $k$  large enough

$$\|(I - \frac{1}{N} \mathbf{1} \mathbf{1}^T) \zeta^{k+1}\|^2 \leq \frac{4N^2}{k^2 p_{\min}^2} \sum_{i \in \{I^k, J^k\}} \|F_i(x_i^k, N\hat{v}_i^k)\|^2.$$

By Assumption 1, the gradients are uniformly bounded, i.e.,  $\|F_i(x_i, y)\| \leq C$  for all  $i$  and some  $C > 0$ . This and  $|\{I^k, J^k\}| = 2$  imply

$$\|(I - \frac{1}{N} \mathbf{1} \mathbf{1}^T) \zeta^{k+1}\|^2 \leq \frac{8N^2}{k^2 p_{\min}^2} C^2.$$

By Jensen's inequality it follows that

$$\text{Term 2} \leq \sqrt{\mathbb{E}[\|(I - \frac{1}{N} \mathbf{1} \mathbf{1}^T) \zeta^{k+1}\|^2 | \mathcal{F}_k]} = \frac{\sqrt{2}N}{k p_{\min}} C.$$

Combining the estimate of **Term 1** and **Term 2** we find

$$\mathbb{E}[\|v^{k+1} - y^{k+1} \mathbf{1}\| | \mathcal{F}_k] \leq \sqrt{\lambda} \|v^k - y^k \mathbf{1}\| + \frac{\sqrt{2}N}{k p_{\min}} C,$$

which after some algebra leads to

$$\begin{aligned} &\frac{1}{k+1} \mathbb{E}[\|v^{k+1} - y^{k+1} \mathbf{1}\| | \mathcal{F}_k] \\ &\leq \frac{1}{k} \|v^k - y^k \mathbf{1}\| - \frac{1 - \sqrt{\lambda}}{k} \|v^k - y^k \mathbf{1}\| + \frac{\sqrt{2}N}{k^2 p_{\min}} C. \end{aligned}$$

Since  $\lambda \in (0, 1)$  we have  $1 - \sqrt{\lambda} > 0$ . Thus by supermartingale convergence of Lemma 2, we have

$$\sum_k \frac{1}{k} \|v^k - y^k \mathbf{1}\| < \infty \quad \text{a.s.} \quad (10)$$

We further have

$$\begin{aligned}
& \mathbb{E}[\|v^{k+1} - y^{k+1}\mathbf{1}\|^2 \mid \mathcal{F}_k] \\
& \leq \mathbb{E}[\|D(k)(v^k - y^k\mathbf{1})\|^2 \mid \mathcal{F}_k] + \mathbb{E}[\|M\zeta^{k+1}\|^2 \mid \mathcal{F}_k] \\
& + 2\sqrt{\mathbb{E}[\|D(k)(v^k - y^k\mathbf{1})\|^2 \mid \mathcal{F}_k]}\sqrt{\mathbb{E}[\|M\zeta^{k+1}\|^2 \mid \mathcal{F}_k]} \\
& \leq \lambda\|v^k - y^k\mathbf{1}\|^2 + \frac{8N^2C^2}{k^2p_{\min}^2} + \frac{2\sqrt{2\lambda}NC}{kp_{\min}}\|v^k - y^k\mathbf{1}\| \\
& = \|v^k - y^k\mathbf{1}\|^2 - (1-\lambda)\|v^k - y^k\mathbf{1}\|^2 + \frac{8N^2C^2}{k^2p_{\min}^2} \\
& + \frac{2\sqrt{2\lambda}NC}{kp_{\min}}\|v^k - y^k\mathbf{1}\|.
\end{aligned}$$

By Lemma 2 (applied with an index shift) and Eq. (10), we find  $\sum_{k=1}^{\infty} \|v^k - y^k\mathbf{1}\|^2 < \infty$  a.s. This relation, combined with the doubly stochastic nature of  $W(k)$  and convexity of the Euclidean norm yields

$$\sum_{k=1}^{\infty} \mathbb{E}[\|\hat{v}^k - y^k\mathbf{1}\|^2 \mid \mathcal{F}_{k-1}] < \infty \quad \text{a.s.}$$

As we will see,  $y^k$  will play a key role in establishing the convergence of projection algorithm. One important property of  $y^k$  is that we have  $y^k = \frac{1}{N} \sum_{j=1}^N x_j^k$  for all  $k \geq 0$ . That is  $y^k$  not only captures the average belief of the agents in the network but also represents the true average information, as seen in the following lemma.

*Lemma 4:* For  $y^k$  of (9) we have  $y^k = \frac{1}{N} \sum_{i=1}^N x_i^k$  for all  $k \geq 0$ .

*Proof:* It suffices to show that for all  $k \geq 0$ ,

$$\sum_{j=1}^N v_j^k = \sum_{j=1}^N x_j^k. \quad (11)$$

We show this by induction on  $k$ . Note that for  $k = 0$  relation (11) holds trivially, as we have  $v_j^0 = x_j^0$ . Assuming relation (11) holds for  $k - 1$ , as the induction step, we have

$$\begin{aligned}
\sum_{j=1}^N v_j & = \sum_{j=1}^N (\hat{v}_j^{k-1} + x_j^k - x_j^{k-1}) \\
& = \sum_{j=1}^N \sum_{i=1}^N [W(k)]_{ji} v_i^{k-1} + \sum_{j=1}^N (x_j^k - x_j^{k-1}) \\
& = \sum_{i=1}^N v_i^{k-1} + \sum_{j=1}^N (x_j^k - x_j^{k-1}) = \sum_{j=1}^N x_j^k,
\end{aligned}$$

where we use the fact that  $W(k)$  is doubly stochastic and in the last equality, we use the induction hypothesis. Thus, it follows that  $y^k = \frac{1}{N} \sum_{j=1}^N v_j^k = \frac{1}{N} \sum_{j=1}^N x_j^k$ .  $\blacksquare$

We now show the convergence of the algorithm. We have the following result, where  $K^*$  denotes the set of Nash equilibria for aggregative game in (1).

*Proposition 1:* Let Assumptions 1–3 hold. Also let  $K^*$  be nonempty, and let the mapping  $F$  be strictly monotone over the set  $K$ , in the following sense:

$$(F(x) - F(y))^T(x - y) > 0 \quad \text{for all } x, y \in K,$$

where the equality holds only for  $x = y$ . Then, the sequence  $\{x^k\}$  generated by the method (7) with stepsize  $\alpha_{k,i} = \frac{1}{\Gamma_k(i)}$  converges to the (unique)  $x^*$  of the game almost surely.

*Proof:* Since the map is strictly monotone, there is a unique Nash equilibrium  $x^* \in K$  (see Theorem 2.3.3. in [22]). Then, by the definition of  $x_i^{k+1}$  we have

$$\begin{aligned}
& \|x_i^{k+1} - x_i^*\|^2 \\
& = \left\| \left( \Pi_{K_i} [x_i^k - \alpha_{k,i} F_i(x_i^k, N\hat{v}_i^k)] - x_i^k \right) \mathbb{1}_{\{i \in \{I^k, J^k\}\}} \right. \\
& \quad \left. + x_i^k - x_i^* \right\|^2.
\end{aligned}$$

Using  $x_i^* = \Pi_{K_i} [x_i^* - \alpha_{k,i} F_i(x_i^*, \bar{x}^*)]$  and the non-expansive property of the projection operator, we have for  $i \in \{I^{k+1}, J^{k+1}\}$ ,

$$\begin{aligned}
& \|x_i^{k+1} - x_i^*\|^2 \\
& \leq \|x_i^k - \alpha_{k,i} F_i(x_i^k, N\hat{v}_i^k) - x_i^k + x_i^* + \alpha_{k,i} F_i(x_i^*, \bar{x}^*)\|^2 \\
& = \|x_i^k - x_i^*\|^2 + \alpha_{k,i}^2 \|F_i(x_i^k, N\hat{v}_i^k) - F_i(x_i^*, \bar{x}^*)\|^2 \\
& \quad - 2\alpha_{k,i} (F_i(x_i^k, N\hat{v}_i^k) - F_i(x_i^*, \bar{x}^*))^T (x_i^k - x_i^*).
\end{aligned}$$

By Lemma 1, for  $k$  large enough we have

$$\alpha_{k,i} \leq \frac{2}{kp_i}, \alpha_{k,i}^2 \leq \frac{4N^2}{k^2 p_{\min}}, \left| \alpha_{k,i} - \frac{1}{kp_i} \right| \leq \frac{2}{k^{3/2-q} p_{\min}^2}.$$

Writing  $\alpha_{k,i} = \left( \alpha_{k,i} - \frac{1}{kp_i} \right) + \frac{1}{kp_i}$ , we obtain

$$\begin{aligned}
& \|x_i^{k+1} - x_i^*\|^2 \\
& \leq \|x_i^k - x_i^*\|^2 + \frac{4N^2}{k^2 p_{\min}} \|F_i(x_i^k, N\hat{v}_i^k) - F_i(x_i^*, \bar{x}^*)\|^2 \\
& \quad + \frac{4}{k^{3/2-q} p_{\min}^2} |(F_i(x_i^k, N\hat{v}_i^k) - F_i(x_i^*, \bar{x}^*))^T (x_i^k - x_i^*)| \\
& \quad - \frac{2}{kp_i} (F_i(x_i^k, N\hat{v}_i^k) - F_i(x_i^*, \bar{x}^*))^T (x_i^k - x_i^*). \quad (12)
\end{aligned}$$

We now approximate the inner product term by adding and subtracting  $F_i(x_i^k, Ny^k)$  and using  $y^k = \sum_{i=1}^N x_i^k$  (cf. Lemma 4), as follows:

$$\begin{aligned}
& (F_i(x_i^k, N\hat{v}_i^k) - F_i(x_i^*, \bar{x}^*))^T (x_i^k - x_i^*) \\
& = (F_i(x_i^k, N\hat{v}_i^k) - F_i(x_i^k, Ny^k))^T (x_i^k - x_i^*) \\
& \quad + (F_i(x_i^k, Ny^k) - F_i(x_i^*, \bar{x}^*))^T (x_i^k - x_i^*).
\end{aligned}$$

Similar operations on the square norm term yield

$$\begin{aligned} & \|F_i(x_i^k, N\hat{v}_i^k) - F_i(x_i^*, \bar{x}^*)\|^2 \\ & \leq 2\|F_i(x_i^k, N\hat{v}_i^k) - F_i(x_i^k, Ny^k)\|^2 \\ & \quad + 2\|F_i(x_i^k, \bar{x}^k) - F_i(x_i^*, \bar{x}^*)\|^2, \end{aligned}$$

where we use  $(a + b)^2 \leq 2(a^2 + b^2)$ . Substituting the preceding two bounds in (12), after rearranging the terms, we obtain

$$\begin{aligned} & \|x_i^{k+1} - x_i^*\|^2 \\ & \leq \|x_i^k - x_i^*\|^2 + \frac{8N^2}{k^2 p_{\min}} \|F_i(x_i^k, N\hat{v}_i^k) - F_i(x_i^k, Ny^k)\|^2 \\ & \quad - \frac{2}{kp_i} (F_i(x_i^k, N\hat{v}_i^k) - F_i(x_i^k, Ny^k))^T (x_i^k - x_i^*) \\ & \quad - \frac{2}{kp_i} (F_i(x_i^k, \bar{x}^k) - F_i(x_i^*, \bar{x}^*))^T (x_i^k - x_i^*) \\ & \quad + \frac{4}{k^{3/2-q} p_{\min}^2} |(F_i(x_i^k, N\hat{v}_i^k) - F_i(x_i^k, Ny^k))^T (x_i^k - x_i^*)| \\ & \quad + \frac{4}{k^{3/2-q} p_{\min}^2} |(F_i(x_i^k, \bar{x}^k) - F_i(x_i^*, \bar{x}^*))^T (x_i^k - x_i^*)| \\ & \quad + \frac{8N^2}{k^2 p_{\min}} \|F_i(x_i^k, \bar{x}^k) - F_i(x_i^*, \bar{x}^*)\|^2. \end{aligned}$$

Applying  $\pm 2a^T b \leq \|a\|^2 + \|b\|^2$ , to the following three terms

$$\begin{aligned} & - \frac{2}{kp_i} (F_i(x_i^k, N\hat{v}_i^k) - F_i(x_i^k, Ny^k))^T (x_i^k - x_i^*), \\ & \frac{4}{k^{3/2-q} p_{\min}^2} |(F_i(x_i^k, N\hat{v}_i^k) - F_i(x_i^k, Ny^k))^T (x_i^k - x_i^*)| \\ & \frac{4}{k^{3/2-q} p_{\min}^2} |(F_i(x_i^k, \bar{x}^k) - F_i(x_i^*, \bar{x}^*))^T (x_i^k - x_i^*)| \end{aligned}$$

modifies our estimate as follows:

$$\begin{aligned} & \|x_i^{k+1} - x_i^*\|^2 \\ & \leq \left(1 + \frac{1}{k^2 p_i^2}\right) \|x_i^k - x_i^*\|^2 \\ & \quad + \|F_i(x_i^k, N\hat{v}_i^k) - F_i(x_i^k, Ny^k)\|^2 \\ & \quad - \frac{2}{kp_i} (F_i(x_i^k, \bar{x}^k) - F_i(x_i^*, \bar{x}^*))^T (x_i^k - x_i^*) \\ & \quad + \frac{2}{k^{3/2-q} p_{\min}^2} \|F_i(x_i^k, N\hat{v}_i^k) - F_i(x_i^k, y^k)\|^2 \\ & \quad + \frac{2}{k^{3/2-q} p_{\min}^2} \|x_i^k - x_i^*\|^2 + \frac{2}{k^{3/2-q} p_{\min}^2} \|x_i^k - x_i^*\|^2 \\ & \quad + \frac{2}{k^{3/2-q} p_{\min}^2} \|F_i(x_i^k, \bar{x}^k) - F_i(x_i^*, \bar{x}^*)\|^2 \\ & \quad + \frac{8N^2}{k^2 p_{\min}} \|F_i(x_i^k, N\hat{v}_i^k) - F_i(x_i^k, Ny^k)\|^2 \\ & \quad + \frac{8N^2}{k^2 p_{\min}} \|F_i(x_i^k, \bar{x}^k) - F_i(x_i^*, \bar{x}^*)\|^2. \end{aligned}$$

Thus, by combining the common terms and using the Lipschitz property of the mapping  $F_i$  (Assumption 2 (ii)) we obtain for  $i \in \{I^k, J^k\}$ ,

$$\begin{aligned} & \|x_i^{k+1} - x_i^*\|^2 \leq \left(1 + \frac{1}{k^2 p_i^2} + \frac{4}{k^{3/2-q} p_{\min}^2}\right) \|x_i^k - x_i^*\|^2 \\ & \quad - \frac{2}{kp_i} (F_i(x_i^k, \bar{x}^k) - F_i(x_i^*, \bar{x}^*))^T (x_i^k - x_i^*) \\ & \quad + \left(1 + \frac{2}{k^{3/2-q} p_{\min}^2} + \frac{8N^2}{k^2 p_{\min}}\right) L_{-i}^2 N^2 \|\hat{v}_i^k - y^k\|^2 \\ & \quad + \left(\frac{2}{k^{3/2-q} p_{\min}^2} + \frac{8N^2}{k^2 p_{\min}}\right) \|F_i(x_i^k, \bar{x}^k) - F_i(x_i^*, \bar{x}^*)\|^2. \end{aligned}$$

From the fact that  $x_i^{k+1} = x_i^k$  when  $i \notin \{I^k, J^k\}$ , it follows that  $\|x_i^{k+1} - x_i^*\|^2 = \|x_i^k - x_i^*\|^2$ . We combine these two cases with the fact that agent  $i$  updates with probability  $p_i$ , where we bound  $p_i$  by the maximum value  $p_{\max} = \max_i p_i$ , and thus obtain for all  $i \in \mathcal{N}$ ,

$$\begin{aligned} & \mathbb{E}[\|x_i^{k+1} - x_i^*\|^2 \mid \mathcal{F}_{k-1}] \\ & \leq \left(1 + \frac{1}{k^2 p_i} + \frac{4p_i}{k^{3/2-q} p_{\min}^2}\right) \|x_i^k - x_i^*\|^2 \\ & \quad - \frac{2}{k} (F_i(x_i^k, \bar{x}^k) - F_i(x_i^*, \bar{x}^*))^T (x_i^k - x_i^*) \\ & \quad + p_{\max} \left(1 + \frac{2}{k^{3/2-q} p_{\min}^2} + \frac{8N^2}{k^2 p_{\min}}\right) L_{-i}^2 N^2 \\ & \quad \times \mathbb{E}[\|\hat{v}_i^k - y^k\|^2 \mid \mathcal{F}_{k-1}] \\ & \quad + \left(\frac{2p_{\max}}{k^{3/2-q} p_{\min}^2} + \frac{8N^2 p_{\max}}{k^2 p_{\min}}\right) \|F_i(x_i^k, \bar{x}^k) - F_i(x_i^*, \bar{x}^*)\|^2. \end{aligned}$$

Summing over all  $i = 1, \dots, N$  and using the notation given in (3), we further obtain

$$\begin{aligned} & \mathbb{E}[\|x^{k+1} - x^*\|^2 \mid \mathcal{F}_{k-1}] \\ & \leq \left(1 + \frac{N}{k^2 p_{\min}} + \frac{4p_{\max}}{k^{3/2-q} p_{\min}^2}\right) \|x^k - x^*\|^2 \\ & \quad - \frac{2}{k} (F(x^k, \bar{x}^k) - F(x^*, \bar{x}^*))^T (x^k - x^*) \\ & \quad + p_{\max} \left(1 + \frac{2}{k^{3/2-q} p_{\min}^2} + \frac{8N^2}{k^2 p_{\min}}\right) \\ & \quad \times \sum_{i=1}^N L_{-i}^2 N^2 \mathbb{E}[\|\hat{v}_i^k - y^k\|^2 \mid \mathcal{F}_{k-1}] \\ & \quad + \left(\frac{2p_{\max}}{k^{3/2-q} p_{\min}^2} + \frac{8N^2 p_{\max}}{k^2 p_{\min}}\right) \|F(x^k) - F(x^*)\|^2. \end{aligned}$$

Using Lipschitz continuity of the mapping  $F$  with constant  $L$  (see Eq. (4)) and gathering the common terms,

we have for all  $k \geq \tilde{k}$  with  $\tilde{k}$  large enough,

$$\begin{aligned} & \mathbb{E}[\|x^{k+1} - x^*\|^2 \mid \mathcal{F}_{k-1}] \\ &= \left(1 + \frac{N + 8N^2L^2p_{\max}}{k^2p_{\min}} + \frac{4p_{\max} + 2L^2p_{\max}}{k^{3/2-q}p_{\min}^2}\right) \|x^k - x^*\|^2 \\ & \quad - \frac{2}{k} (F(x^k, \bar{x}^k) - F(x^*, \bar{x}^*))^T (x^k - x^*) \\ & \quad + p_{\max} \left(1 + \frac{2}{k^{3/2-q}p_{\min}^2} + \frac{8N^2}{k^2p_{\min}}\right) \sum_{i=1}^N L_{-i}^2 N^2 \\ & \quad \times \sum_{i=1}^N \mathbb{E}[\|\hat{v}_i^k - y^k\|^2 \mid \mathcal{F}_{k-1}]. \end{aligned}$$

We now apply the supermartingale convergence lemma (to a delayed sequence). For  $q \in (0, 1/2)$  we have

$$\sum_{k=1}^{\infty} \frac{N + 8N^2L^2p_{\max}}{k^2p_{\min}} + \frac{4p_{\max} + 2L^2p_{\max}}{k^{3/2-q}p_{\min}^2} < \infty.$$

Further from Lemma 3 it follows that

$$\sum_{k=1}^{\infty} \sum_{i=1}^N \mathbb{E}[\|\hat{v}_i^k - y^k\|^2 \mid \mathcal{F}_{k-1}] < \infty.$$

By Lemma 2 we conclude that

$$\{\|x^k - x^*\|\} \text{ converges a.s.}, \quad (13)$$

$$\sum_k \frac{2}{k} (F(x^k, \bar{x}^k) - F(x^*, \bar{x}^*))^T (x^k - x^*) < \infty \text{ a.s.} \quad (14)$$

Since  $\{x^k\} \subset K$  and  $K$  is compact (Assumption 1),  $\{x^k\}$  has accumulation points in  $K$ . By (14) and  $\sum 1/k = \infty$  it follows that  $(x^k - x^*)^T (F(x^k) - F(x^*)) \rightarrow 0$  along a subsequence, say  $\{x^{k_\ell}\}$ , *a.s.* This and the strict monotonicity of  $F$  imply that  $\{x^{k_\ell}\} \rightarrow x^*$  as  $\ell \rightarrow \infty$  *a.s.* By relation (13), the entire sequence converges to  $x^*$  *a.s.* ■

#### IV. NUMERICS

In this section, we examine the behavior of the gossip algorithm on a class of aggregative games. In particular, we consider a class of Nash-Cournot games in which all players have identical cost functions. This leads to a symmetric Nash equilibrium in which the player's decisions are identical. As a consequence, player  $i$ 's optimization problem can be expressed as:

$$\begin{aligned} & \text{minimize} && c_i(x_i) - x_i U(\bar{x}) \\ & \text{subject to} && x_i \in K_i, \end{aligned}$$

where  $x_i$  denotes player  $i$ 's decision,  $\bar{x}$  denotes the aggregate of all the players' decisions ( $\bar{x} = \sum_{i=1}^N x_i$ ),  $c_i(x_i)$  represents the cost function and  $U(\bar{x})$  captures

the inverse demand function. Different forms of cost functions and inverse demand functions are presented in Table I. Furthermore,  $K_i = [0, 250]$  for  $i = 1, \dots, N$ . We investigate a setting with 20 players. Table I also presents the Nash equilibrium decision of each player. We terminate the algorithm after a fixed number of

TABLE I  
MODEL AND NASH EQUILIBRIUM DECISION

Model and functional form			
Model no.	cost $c_i(x_i)$	demand $U(\bar{x})$	N.E.
1	$56x_i$	$256 - \bar{x}$	9.52
2	$10x_i + 10$	$7.36e7 + 10 - \bar{x}^3$	20.00
3	$56x_i$	$8300 - \bar{x}^{3/2}$	19.37
4	$40x_i + 1.5x_i^2$	$1000 - \bar{x}$	40.00

iterations  $\tilde{k}$  and report the mean error after termination for a sample of size 10. The error is defined as

$$\text{err}_{\tilde{k}} := \|x^{\tilde{k}} - x^*\|_{\infty} = \max_{1 \leq i \leq n} \{|x_i^{\tilde{k}} - x_i^*|\}.$$

We assume that the connectivity graph linking the players is given by a cycle. Table II reports our experimental results. Figures 2-3 demonstrate players' decisions and aggregate estimate trajectory for four players chosen at random. The y-axis in the plots denotes the decision/aggregate estimate level and the x-axis denotes the iteration number. Since the game is symmetric, other players' decisions are also expected to follow a similar behavior (and they indeed do). Also it is apparent that players agree much earlier on the aggregate than converging on the optimal decision.

#### V. SUMMARY AND CONCLUSIONS

This paper focuses on a class of Nash games in which user interactions are seen through the aggregate sum of all players' actions. These agents are a part of network with limited connectivity which only allows for restricted local communication. We propose a gossip-based distributed algorithm which abides the information exchange restriction for computation of equilibrium point. Our algorithm allows for asynchronous implementation with distributed architecture. The contribution of our work can broadly be summarized as: (1) the

TABLE II  
MEAN SAMPLE ERROR IN PLAYERS' DECISIONS FOR VARIOUS  $\tilde{k}$

Model No.	Mean error for 10 samples		
	$\tilde{k} = 10^3$	$\tilde{k} = 10^4$	$\tilde{k} = 10^5$
1	4.82	1.11	2.56e-1
2	10.6	2.39	8.99e-2
3	9.51	2.13	4.13e-2
4	9.42	2.85	3.00e-6

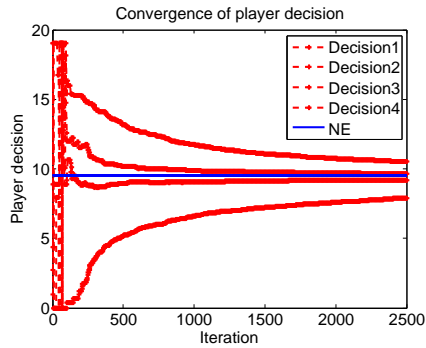


Fig. 2. Convergence of player's decision for model 1.

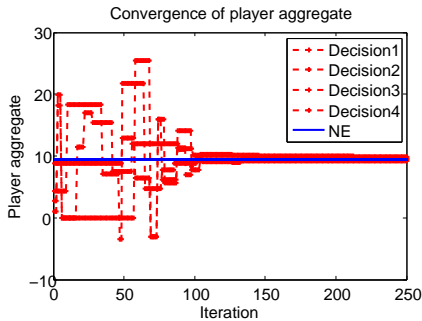


Fig. 3. Convergence of player's estimate of aggregate for model 1.

development of an asynchronous distributed algorithm for aggregative games over graphs; and (2) the establishment of the convergence of the algorithm (with agent specific stepsizes) to an equilibrium point. We also provide illustrative numerical results that support our theoretical findings.

## REFERENCES

- [1] M. K. Jensen, "Aggregative games and best-reply potentials," *Econom. Theory*, vol. 43, no. 1, pp. 45–66, 2010.
- [2] A. Cournot, "Recherches sur les principes mathématiques de la théorie des richesses," 1838.
- [3] D. Fudenberg and J. Tirole, *Game Theory*. MIT Press, 1991.
- [4] W. Novshek, "On the existence of cournot equilibrium," *Review of Economic Studies*, vol. 52, no. 1, pp. 85–98, 1985.
- [5] M. K. Jensen, "Aggregative games," Tech. Rep., 2006.
- [6] D. Martimort and L. Stole, "Aggregate representations of aggregate games," Tech. Rep., 2011.
- [7] T. Alpcan and T. Başar, "A game-theoretic framework for congestion control in general topology networks," in *41th IEEE Conf. Decision and Control*, Las Vegas, NV, 2002.
- [8] —, "Distributed algorithms for Nash equilibria of flow control games," in *Advances in Dynamic Games*, ser. Annals of the International Society of Dynamic Games, vol. 7. Birkhäuser Boston, 2003, pp. 473–498.
- [9] A. Rantzer, "Using game theory for distributed control engineering," Department of Automatic Control, Lund University, Sweden, Tech. Rep. ISRN LUTFD2/TFRT--7620--SE, 2008.
- [10] L. Pavel, "An extension of duality to a game-theoretic framework," *Automatica*, vol. 43, pp. 226–237, 2007.
- [11] Y. Pan and L. Pavel, "Games with coupled propagated constraints in optical network with multi-link topologies," *Automatica*, vol. 45, pp. 871–880, 2009.
- [12] J. Koshal, A. Nedić, and U. V. Shanbhag, "Single timescale regularized stochastic approximation schemes for monotone nash games under uncertainty," *Proceedings of the IEEE Conference on Decision and Control (CDC)*, 2010.
- [13] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah, "Randomized gossip algorithms," *IEEE Transactions on Information Theory*, 2006.
- [14] J. Tsitsiklis, "Problems in decentralized decision making and computation," Ph.D. dissertation, Massachusetts Institute of Technology, 1985.
- [15] A. Kashyap, T. Başar, and R. Srikant, "Quantized consensus," in *2006 IEEE International Symposium on Information Theory*, 2006.
- [16] J. Lavaei and R. Murray, "Quantized consensus by means of gossip algorithm," *IEEE Transactions on Automatic Control*, 2012.
- [17] S. S. Ram, A. Nedić, and V. Veeravalli, "Asynchronous gossip algorithms for stochastic optimization: Constant stepsize analysis," in *Recent Advances in Optimization and its Applications in Engineering*, M. Diehl, F. Glineur, E. Jarlebring, and W. Michiels, Eds. Volume of the 14th Belgian-French-German Conference on Optimization (BFG), 2010, pp. 51–60.
- [18] —, "Asynchronous gossip algorithms for stochastic optimization," 2009, pp. 3581–3586.
- [19] K. Srivastava and A. Nedić, "Distributed asynchronous constrained stochastic optimization," *IEEE Journal of Selected Topics in Signal Processing*, vol. 5, no. 4, pp. 772–790, 2011.
- [20] A. Nedić, "Asynchronous broadcast-based convex optimization over a network," *IEEE Transactions on Automatic Control*, vol. 56, no. 6, pp. 1337–1351, 2011.
- [21] B. Polyak, *Introduction to Optimisation*. New York: Optimization Software, Inc., 1987.
- [22] F. Facchinei and J.-S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems*. Springer-Verlag Inc., New York, 2003, vol. 1 and 2.