Proceedings of the
47th IEEE Conference on Decision and Control
Cancun, Mexico, Dec. 9-11, 2008

Noncooperative Dynamic Games for Inventory Applications: a consensus approach

D. Bauso, L. Giarré, and R. Pesenti

Abstract— **We focus on a finite horizon noncooperative dynamic game where the stage cost of a single player associated to a decision is a monotonically nonincreasing function of the total number of players making the same decision. For the single-stage version of the game, we characterize Nash equilibria and derive a consensus protocol that makes the players converge to the unique Pareto optimal Nash equilibrium. Such an equilibrium guarantees the interests of the players and is also social optimal in the set of Nash equilibria. For the multi-stage version of the game, we present an algorithm that converges to Nash equilibria, unfortunately not necessarily Pareto optimal. The algorithm returns a sequence of joint decisions, each one obtained from the previous one by an unilateral improvement on the part of a single player. The sequence with which the players act is chosen a priori and may influence the Nash equilibrium to which the path converges. We also specialize the game to a multi-retailer inventory system, where competing retailers aim at coordinating their supply strategies in order to minimize their local costs.**

Keywords: Game Theory, Inventory, Consensus Protocols, Dynamic Programming.

I. Introduction

Since the seminal paper of [2] where fictitious play and dynamic games were revisited in a control theoretic perspective, the connection between consensus and game theory is a subject of research. *Consensus protocols* are distributed control policies based on neighbors' state feedback that allow the coordination of multi-agent systems. According to the usual meaning of consensus, the system state must converge to an equilibrium point with all equal components in finite time or asymptotically. Two recent surveys on consensus [14], [16] report in details the main contribution of the past few years on consensus. Consensus problems have been recently largely studied, but the literature on the connection between consensus and game theory is not so extensive. In particular, in [10] the relationship between cooperative control problems, such as the consensus problem, and game theoretic methods, has been established. The effectiveness of using game theoretic approaches for controlling multi-agent systems is presented in [1]. In [6] and [7], it has been shown that the consensus protocol design is the solution of individual optimizations performed by the agents. This notion suggested a game theoretic interpretation of consensus problems as mechanism design problems. Under this perspective a supervisor entails the agents to reach a consensus by imposing individual objectives. We proved that the supervisor can choose objectives so that rational agents have a unique optimal protocol, and reach consensus asymptotically on the desired group decision value. Here, we consider a finite horizon noncooperative game [4] where the stage cost of the ith player associated to a decision is a monotonically nonincreasing function of the total number of players making the same decision. The paper is organized as follows. In Section II, we introduce the game and we recall some properties of the existence of Nash equilibria and of at least one Pareto optimal Nash equilibrium, proved in [5], by recasting the game within the framework of potential games [17]. In Section III and IV, we show that stronger results are obtained if the horizon reduces to a single stage. We find all Nash equilibria and in particular a Pareto optimal one that is social optimal in the set of all Nash equilibria, as it minimizes the sum of the players' costs. We also define a consensus protocol [6], [14], [15], [16] that makes the players converge to the Pareto optimal Nash equilibrium. We do this in agreement with a large body of literature on evolutionary game theory and fictitious play (see e.g., the book [9] and [18]) that centers around the convergence to refined Nash equilibria, that is, Nash equilibria that meet special properties. Social and Pareto optimality are just properties characterizing the Nash equilibria to which the dynamics induced by the consensus protocols converges. In Section V, we come back to the multi-stage game and we modify the above protocol to derive a so called best response path algorithm that makes the players converge to a Nash equilibrium. This algorithm is based on the property of potential games establishing that any best response path converges to a Nash equilibrium [17], [18]. A best response path is a sequence of joint decisions, each one obtained from the previous one by an unilateral improvement on the part of a single player. In Section VI, we specialize the game to a multi-inventory application [3], [11], [12], [13].

II. Noncooperative dynamic game and Nash equilibria

We deal with a discrete time finite horizon noncooperative game which presents all the ingredients typical of an inventory application. However, we deal with the game in its general form in order to emphasize what characteristics make the results of this paper hold.

Consider a set of n players $\Gamma = \{1, \ldots, n\}$ and let N be the horizon length. For each $i \in \Gamma$ and each stage $k = 0, \ldots, N$, let $x_i^k \in X_i^k \subseteq \mathbb{Z}$ be a discrete time state and $u_i^k \in U_i^k \subseteq \mathbb{N}$ be a decision. Here, we have denoted by X_i^k

D. Bauso is with DINFO, Università di Palermo, 90128 Palermo, Italy, E-Mail: dario.bauso@unipa.it

Corresponding author L. Giarré is with DIAS, Università di Palermo, 90128 Palermo, Italy, E-Mail: giarre@unipa.it

R. Pesenti is with DMA, Università di Venezia, 30123 Venezia, Italy, E-Mail: pesenti@unive.it

978-1-4244-3124-3/08/$25.00 ©2008 IEEE

and U_i^k the set of feasible states and decisions at stage k and by $\mathbb{Z}$, $\mathbb{N}$ the set of integers and non negative integers (zero included), respectively. Let $u_{-i}^k = \{u_j^k\}_{j\in\Gamma, j\neq i}$ be the vector of the decisions of players $j \neq i$ at stage k. Also, define $u^k = \{u_i^k\}_{i\in\Gamma}$, $\mathbf{u}_i = \{u_i^0, \dots, u_i^N\}$ and $\mathbf{u}_{-i} = \{u_{-i}^0, \dots, u_{-i}^N\}$. Let the following finite horizon noncooperative game be given: for each player $i \in \Gamma$,

$$\hat{J}_i(x_i^0, \mathbf{u}_i, \mathbf{u}_{-i}) = \sum_{k=0}^{N} g_i(x_i^k, u_i^k, a(u^k)) \quad (1)$$

$$x_i^{k+1} = \Xi(x_i^k, u_i^k), \quad k = 0, \dots, N-1, \quad (2)$$

where equation (1) is the cost function, obtained as sum over the horizon of a stage cost $g_i(x_i^k, u_i^k, a(u^k))$ and equation (2) is the state dynamics with $\Xi(.,.)$ being a generic nonlinear function, possibly time variant and player specific, but such that $\lim_{u_i^k\to+\infty} \Xi(x_i^k, u_i^k) = +\infty$, for all $x_i^k \in \mathbb{Z}$. The stage cost $g_i(x_i^k, u_i^k, a(u^k))$ is of type

$$g_i(x_i^k, u_i^k, a(u^k)) = \delta(u_i^k)\psi(a(u^k)) + \gamma(x_i^k, u_i^k), \quad (3)$$

where: function $\delta(u_i^k)$ is equal to one if $u_i^k > 0$ (we say that the ith player is *active*), and zero otherwise; function $a(u^k)$ returns the *number of active players (at time k)*, $a(u^k) = \sum_{j=1}^{n} \delta(u_j^k)$; function $\psi(a(u^k))$ is positive and strictly decreasing on $a(.)$; function $\gamma(x_i^k, u_i^k)$ is coercive, non negative and independent of $a(.)$. Henceforth, for the short of notation, we write a^k to mean $a(u^k)$. Also we denote by $\mathbf{u} = [\mathbf{u}_1, \dots, \mathbf{u}_n]$ a generic solution of the game (in the following we also use the notation $[\mathbf{u}_i, \mathbf{u}_{-i}]$ to mean $\mathbf{u}$). Finally, we define $J_i(x_i^0, \mathbf{u}_{-i}) = \min_{\mathbf{u}_i} \hat{J}_i(x_i^0, \mathbf{u}_i, \mathbf{u}_{-i})$.

In [5] we proved the existence of Nash equilibria by exploiting the well-known result in [17] asserting that a noncooperative game always admits a pure Nash Equilibrium if a *potential function* exists. A potential function is a function $\Phi(x^0, \mathbf{u})$ such that, if $\hat{\mathbf{u}} = [\hat{\mathbf{u}}_i, \hat{\mathbf{u}}_{-i}]$ is a solution obtained from an unilateral deviation from $\mathbf{u}$ on the part of a generic player i (hence $\mathbf{u}_i \neq \hat{\mathbf{u}}_i$, but $\mathbf{u}_{-i} = \hat{\mathbf{u}}_{-i}$), the difference induced to the potential function $\Delta\Phi = \Phi(x^0, [\hat{\mathbf{u}}_i, \hat{\mathbf{u}}_{-i}]) - \Phi(x^0, [\mathbf{u}_i, \mathbf{u}_{-i}])$ is equal to, or at least proportional to, the difference in the cost for player i, that is, $\Delta\hat{J}_i = \hat{J}_i(x_i^0, \hat{\mathbf{u}}_i, \hat{\mathbf{u}}_{-i}) - \hat{J}_i(x_i^0, \mathbf{u}_i, \mathbf{u}_{-i})$. We recall hereafter the main results proved in [5].

Theorem 1: Game (1)-(2) is a potential game.

As a consequence, by the results in [17], we can state the following corollary.

Corollary 1: Game (1)-(2) admits at least one Nash equilibrium.

Let us now characterize a generic Nash equilibrium $\mathbf{u}^* = [\mathbf{u}_i^*, \mathbf{u}_{-i}^*]$ where $\mathbf{u}_i^* = \{u_i^{0*}, \dots, u_i^{N*}\}$ and $\mathbf{u}_{-i}^* = \{u_{-i}^{0*}, \dots, u_{-i}^{N*}\}$. In particular, we consider the ith player and study the unilateral improvements by fixing the decisions of all other players over the horizon $\mathbf{u}_{-i}^*$. We denote by $\mathbf{a}^{k*} = \{a^{k*}, \dots, a^{N*}\}$ with $a^{\hat{k}*} = \sum_{j=1, j\neq i}^{n} \delta(u_j^{\hat{k}*}) + \delta(u_i^{\hat{k}})$ for $\hat{k} = k, \dots, N$. The vector $\mathbf{a}^{k*}$ collects the number of active players from stage k to N as a function of $\{u_i^k, \dots, u_i^N\}$ and for fixed $\{u_{-i}^{k*}, \dots, u_{-i}^{N*}\}$. By applying the dynamic programming approach to (1)-(2), we can define

$$J_i^N(x_i^N, \mathbf{a}^{N*}) = 0, \quad (4)$$

$$J_i^k(x_i^k, \mathbf{a}^{k*}) = \min_{u_i^k\in U_i^k} [g_i(x_i^k, u_i^k, a^{k*}) + J_i^{k+1}(x_i^{k+1}, \mathbf{a}^{k+1*})] \quad . \quad (5)$$

Then, $J_i(x_i^0, \mathbf{u}_{-i}^*)$ is equal to $J_i^0(x_i^0, \mathbf{a}^{0*})$. In solving (4)-(5), we can do as if a^{k*} was independent of u_i^k. Actually, we can substitute a^{k*} by $\tilde{a}^k = \sum_{j=1, j\neq i}^{n} \delta(u_j^{k*}) + 1$, for $k = 0, \dots, N$. We can do such a substitution as it turns out that $g_i(x_i^k, u_i^k, a^{k*}) = g_i(x_i^k, u_i^k, \tilde{a}^k)$. To see why the latter equality holds true, observe that the stage cost $g_i(x_i^k, u_i^k, a^{k*})$ depends on a^{k*} only through the term, $\delta(u_i^k)\psi(a^{k*})$, which is different from zero only when $\delta(u_i^k) = 1$, that is when $a^{k*} = a^{k*} - \delta(u_i^k) + 1 = \tilde{a}^k$. It follows that the best response for player i must be a solution of equation (5), i.e.,

$$\begin{aligned} u_i^{k*} &= \arg\min_{u_i^k\in U_i^k} \\ &(\delta(u_i^k)\psi(a^{k*}) + \gamma(x_i^k, u_i^k) + J_i^{k+1}(x_i^{k+1}, \mathbf{a}^{k+1*})) \\ &= \arg\min_{u_i^k\in U_i^k} \\ &(\delta(u_i^k)\psi(\tilde{a}^k) + \gamma(x_i^k, u_i^k) + J_i^{k+1}(x_i^{k+1}, \tilde{\mathbf{a}}^{k+1})) \end{aligned} \quad (6)$$

where we define $\tilde{\mathbf{a}}^k = \{\tilde{a}^k, \dots, \tilde{a}^N\}$ for $k = 0, \dots, N$. The above equation may present multiple solutions, [19]. However, the values assumed by u_i^{k*} depends on the other player decisions only in terms of the number of active players. With this in mind, we can derive that given two equilibria $\hat{\mathbf{u}}$ and $\tilde{\mathbf{u}}$, if $\delta(\hat{u}_i^k) = \delta(\tilde{u}_i^k)$ for all $i \in \Gamma$ and for all $k = 0, \dots, N-1$, then the two equilibria are equivalent, that is $\hat{J}_i(x_i^0, \hat{\mathbf{u}}_i, \hat{\mathbf{u}}_{-i}) = \hat{J}_i(x_i^0, \tilde{\mathbf{u}}_i, \tilde{\mathbf{u}}_{-i})$ for all $i \in \Gamma$. In the following, in case of multiple solutions, we choose u_i^{k*} as the lowest among the possible scalar values that satisfy (6). In this way we guarantee the uniqueness of the best response and we can describe the equilibria indifferently in term of either $\mathbf{u}^*$ or $\mathbf{a}^0$ given their bijective correspondence. Needless to say that the players can choose any other criterium that guarantees the uniqueness of the best response in (6) without compromising the validity of the results.

Still in [5] we prove that Nash equilibria are finite in number and as a consequence of this we establish the following result.

Theorem 2: At least a Nash equilibrium is Pareto optimal.

III. Single stage game

We now consider a finite horizon noncooperative game consisting in a single stage game with payoffs (in all the equations of this subsection we drop the dependence on k)

$$\hat{J}_i(x_i, u_i, u_{-i}) = \delta(u_i)\psi(a(u)) + \gamma(x_i, u_i), \quad (7)$$

where all the variables and functions have the same definitions and properties of the original game. Game (7) is trivially obtained from the original game by imposing $N = 0$.

For each $i \in \Gamma$, let $l : \mathbb{Z} \to \mathbb{N}$, increasing function of x_i, be given. Henceforth, we simply use the notation l_i to mean $l(x_i)$, i.e., the value of the function for fixed x_i. Note that in the single stage game and once fixed the scenario (x_i fixed),

x_i becomes a known parameter (the initial inventory) and therefore we can omit dependence of $l(x_i)$ on x_i.

Definition 1: A *threshold strategy* is any function $\tilde{u}(.) : \mathbb{N} \times \mathbb{N} \longrightarrow \mathbb{R}$ such that $\tilde{u}(a, l_i)$ assumes a positive value if $a \geq l_i$ and is null otherwise. In this case l_i is said *threshold*. The above threshold strategy says that player i is active only if the number of active players a is greater than or equal to threshold l_i. Let us now characterize a Nash equilibrium, $u^* = [u_1^*, \ldots, u_n^*]$, for the single stage game, where u_i^* is the best response of player i. Again, denote by $a^* = \sum_{j=1, j\neq i}^n \delta(u_j^*) + \delta(u_i)$ the vector collecting the number of active players as a function of u_i and for fixed u_{-i}^*. Condition (6) becomes

$$u_i^* = \arg\min_{u_i \in U_i} [\delta(u_i)\psi(a^*) + \gamma(x_i, u_i)], \tag{8}$$

and in case of multiple solutions we choose u_i^* as the lowest among the possible scalar values that satisfy the above equation. Note that in (8) we can replace a^* by $\tilde{a} = \sum_{j=1, j\neq i}^n \delta(u_j^*) + 1$ and use the same trick explained for the solution of (4)-(5).

Lemma 1: At a Nash equilibrium $u^* = [u_1^*, \ldots, u_n^*]$, the best response u_i^* of each player i is a threshold strategy $u_i^* = \tilde{u}(a^*, l_i)$ with threshold

$$l_i = \min\{\mu \in \{1, \ldots, n\} : \psi(\mu) < \gamma(x_i, 0)\}. \tag{9}$$

Proof: Let us first prove that the best response u_i^* of player i is a threshold strategy. On this purpose, for each player i, and for any number of active players $\beta \geq \alpha$, let ζ_α and ζ_β be the best responses for $a^* = \alpha$ and $a^* = \beta$ respectively (they solve (8) with $a^* = \alpha$ and $a^* = \beta$). We show that if $\zeta_\alpha > 0$ (it means $\delta(\zeta_\alpha) = 1$, the ith player is active) then $\zeta_\beta > 0$. To see this observe that $\zeta_\alpha > 0$ only if

$$\psi(\alpha) + \gamma(x_i, \zeta_\alpha) \leq \gamma(x_i, 0).$$

As $\psi(.)$ is a positive function, to have $\zeta_\beta > 0$ it suffices to prove that

$$\psi(\beta) + \gamma(x_i, \zeta_\beta) \leq \gamma(x_i, 0).$$

Note that the rhs of the above two inequalities are equal as they do not depend on the number of active players. Then we can show that the latter inequality holds as

$$\begin{aligned} &\psi(\beta) + \gamma(x_i, \zeta_\beta) \leq \\ &\psi(\beta) + \gamma(x_i, \zeta_\alpha) \leq \psi(\alpha) + \gamma(x_i, \zeta_\alpha) \\ &\leq \gamma(x_i^k, 0), \end{aligned} \tag{10}$$

where the first inequality is due to the optimality of ζ_β and the second inequality is due to the monotonicity of ψ on the number of active players. Then, we have proved that $u_i^* = \tilde{u}(a^*, l_i)$.

Now, to see that the threshold is as in (9) observe that it must also hold $\psi(\alpha) + \gamma(x_i, u_i^*) < \gamma(x_i, 0)$ for all $\alpha \geq l_i$ and $\psi(\alpha) + \gamma(x_i, u_i^*) \geq \gamma(x_i, 0)$ for all $\alpha < l_i$. But the latter conditions hold if and only if the value of l_i is as in (9). ■

As in (5), the best response u_i^* defined in the above lemma depends on other players course of action u_{-i}^* only through a^*. In the next theorem we characterize the unique Pareto optimal Nash equilibrium. To this aim, let us relate Nash equilibria to subsets of players as follows. Without loss of generality, assume that the players are indexed increasingly on their thresholds, i.e., $l_1 \leq l_2 \leq \ldots \leq l_n$. Define *compatible set* any set of consecutive players $C = \{1, \ldots, r\}$ such that $l_r \leq r$. Any player of a compatible set C benefits from being active if all the other players in C are active. Observe that for any Nash equilibrium $u^* = [u_1^*, \ldots, u_n^*]$ there exists a compatible set C such that $\delta(u_i^*) = 1$ if and only if $i \in C$. Indeed, let $\hat{i} = \max\{i : \delta(u_i^*) = 1\}$, then $\delta(u_i^*) = 1$ for all $i \in \Gamma$ such that $i < \hat{i}$ since $l_i \leq l_{\hat{i}}$. Now, consider the *maximal compatible set* $\overline{C} = \{1, \ldots, \bar{\lambda}\}$ where

$$\bar{\lambda} = \arg\max_{\lambda} \{\lambda \in \{1, \ldots, n\} : l_\lambda \leq \lambda\}.$$

Note that $\overline{C}$ may be empty and that, by maximality of $\overline{C}$, $l_i > \bar{\lambda} + 1$ for all players $i \notin \overline{C}$.

Lemma 2: There always exists a Nash equilibrium $u^* = [u_1^*, \ldots, u_n^*]$ such that $\delta(u_i^*) = 1$ if and only if $i \in \overline{C}$

Proof: The solution u^* describes the case where the active players are the only players in $\overline{C}$ and therefore the number of active players is $\bar{\lambda}$. Then, no players $i \in \overline{C}$ benefit by unilaterally deciding of becoming non active as $l_i \leq \bar{\lambda}$ and also no players $j \notin \overline{C}$ benefit by deciding of becoming active as $l_j > \bar{\lambda} + 1$. ■

Theorem 3: Let u^* be the Nash equilibrium associated to the maximal compatible set $\overline{C}$, i.e.,

$$\delta(u_i^*) = \begin{cases} 1 & \text{if } i \in \overline{C} \\ 0 & \text{otherwise} \end{cases}.$$

If $\psi(\bar{\lambda}) + \gamma(x_i, u_i^*) \neq \gamma(x_i, 0)$ for all $i \in \overline{C}$, then

- *Pareto optimality.* The Nash equilibrium u^* is Pareto optimal;
- *Uniqueness.* The Nash equilibrium u^* is the unique Pareto optimal Nash equilibrium.
- *Social optimality*. The Nash equilibrium u^* is social optimal in the set of all Nash equilibria.

Proof: Pareto optimality. We show that the Nash equilibrium $u^* = [u_1^*, \ldots, u_n^*]$ is Pareto optimal since any other vector of strategies $u = [u_1, \ldots, u_n]$ induces a worse payoff for at least one player. In the Nash equilibrium u^*, each $i \in \overline{C}$ gets a payoff $\hat{J}_i(x_i, u_i^*, u_{-i}^*) = \psi(\bar{\lambda}) + \gamma(x_i, u_i^*) < \gamma(x_i, 0)$, each $i \notin \overline{C}$ gets a payoff $\hat{J}_i(x_i, 0, u_{-i}^*) = \gamma(x_i, 0) < \psi(\bar{\lambda} + 1) + \gamma(x_i, u_i)$ for all $u_i > 0$. Now, consider the vector of strategies u. Define $D = \{i \in \overline{C} : \delta(u_i) = 0\}$ as the set of players with $l_i \leq \bar{\lambda}$ that are not active in u and $E = \{i \notin \overline{C} : \delta(u_i) = 1\}$ as the set of players with $l_i > \bar{\lambda} + 1$ that are active in u. Let us denote by ν and η the cardinality of D and E respectively. Trivially, $D \cup E \neq \emptyset$ as $u \neq u^*$. We deal with $E \neq \emptyset$ and $E = \emptyset$ separately.

If $E \neq \emptyset$ and $D = \emptyset$, each player $i \in E$ gets a payoff $\hat{J}_i(x_i, u_i, u_{-i}) = \psi(\bar{\lambda} + \eta) + \gamma(x_i, u_i)$ strictly greater than $\hat{J}_i(x_i, 0, u_{-i}^*) = \gamma(x_i, 0)$ as $\overline{C}$ is the maximal compatible set. The latter condition trivially holds also when $D \neq \emptyset$ since, in this case, each player $i \in E$ incurs in a higher payoff $\hat{J}_i(x_i, u_i, u_{-i}) = \psi(\bar{\lambda} + \eta - \nu) + \gamma(x_i, u_i)$.

If $E = \emptyset$, then $D \neq \emptyset$, and each player $i \in \overline{C} \setminus D$, if exists, gets a payoff $\hat{J}_i(x_i, u_i, u_{-i}) = \psi(\bar{\lambda} - \nu) + \gamma(x_i, u_i) > \hat{J}_i(x_i, u_i^*, u_{-i}^*) = \psi(\bar{\lambda}) + \gamma(x_i, u_i^*)$. At the same time, each player $i \in D$ gets a payoff $\hat{J}_i(x_i, 0, u_{-i}) = \gamma(x_i, 0) > \hat{J}_i(x_i, u_i^*, u_{-i}^*) = \psi(\bar{\lambda}) + \gamma(x_i, u_i^*)$. Finally, each $i \in \Gamma \setminus \overline{C}$ gets a payoff $\hat{J}_i(x_i, 0, u_{-i}) = \gamma(x_i, 0) = \hat{J}_i(x_i, 0, u_{-i}^*)$.

Uniqueness and social optimality. We prove the uniqueness and the social optimality of the Pareto optimal Nash Equilibrium by showing that it dominates all the other equilibria. Consider a generic Nash equilibrium u associated to a compatible set C, say λ its cardinality, different from $\overline{C}$. Since $\overline{C}$ is maximal then $C \subset \overline{C}$. Then, each $i \in C$, if exists, gets a payoff $\hat{J}_i(x_i, u_i, u_{-i}) = \psi(\lambda) + \gamma(x_i, u_i) > \hat{J}_i(x_i, u_i^*, u_{-i}^*) = \psi(\bar{\lambda}) + \gamma(x_i, u_i^*)$; analogously, each $i \in \overline{C} \setminus C$ gets a payoff $\hat{J}_i(x_i, u_i, u_{-i}) = \gamma(x_i, 0) > \hat{J}_i(x_i, u_i^*, u_{-i}^*) = \psi(\bar{\lambda}) + \gamma(x_i, u_i^*)$; finally, each player $i \in \Gamma \setminus \overline{C}$, gets a payoff $\hat{J}_i(x_i, u_i, u_{-i}) = \gamma(x_i, 0) = \hat{J}_i(x_i, u_i^*, u_{-i}^*)$. Then, in any generic Nash equilibrium each player has a payoff not better than the one associated to u^*. ■

Observe that if and only if $\psi(\bar{\lambda}) + \gamma(x_i, u_i^*) = \gamma(x_i, 0)$ for all i, there exist two Pareto optimal Nash equilibria with equal payoff. They are associated respectively to the maximal compatible set $\overline{C}$ and to the empty set. Henceforth, we will call *Pareto optimal Nash equilibrium* only the equilibrium u^* associated to the maximal compatible set $\overline{C}$. Also, observe that there is no other Nash equilibrium with a higher number of active players than the Pareto optimal Nash equilibrium. Let us finally note that the minimizer of the sum of players' costs, say it *social optimum*, is in general not an equilibrium. However, if we restrict the minimization within the set of Nash equilibria, then the social optimum is on the Pareto optimal Nash equilibrium as it has been shown in the above theorem. Restricting the minimization within the set of Nash equilibria makes sense as the players participate to a noncooperative game, then any solution that is not an equilibrium is of no interest.

IV. Consensus Problem

With focus on the single stage game (7), we now introduce a protocol that makes the players strategies converge to the Pareto optimal Nash equilibrium characterized in Theorem 3.

For all players $i \in \Gamma$, let us refer to $\hat{a}_i$ as their estimate of a in the assumption that each player may exchange information only with a subset of neighbor players. In this sense, the set Γ induces an undirected connected graph $G = (\Gamma, E)$ whose edgeset E includes all non oriented couples (i, j) of players that exchange information with each other. Also, define the neighborhood of player i the set $N_i = \{j : (i, j) \in E\} \cup \{i\}$. Let $z_i(\tau) \in \mathbb{R}$ be a continuous time variable describing the transmitted information for $\tau \geq 0$ and let T be a sufficiently large time interval. The information flow is managed through a *distributed* protocol $\Pi = \{(f_i, \phi_i) : \text{ for all } i \in \Gamma\}$

$$\dot{z}_i(\tau) = f_i(z_j(\tau) \text{ for all } j \in N_i), \; 0 \leq \tau \leq T, \quad (11)$$

$$\hat{a}_i(\tau) = \phi_i(z_i(\tau)) \quad (12)$$

$$u_i^* = \tilde{u}(\hat{a}_{i,ss}, l_i) \quad (13)$$

where $f_i : \mathbb{R}^n \to \mathbb{R}$ describes the dynamics of the transmitted information of the ith node as a function of the information both available at the node itself and transmitted by the other nodes, as in (11); $\phi_i : \mathbb{R} \to \mathbb{R}$ estimates, based on current information, the aggregate info, as in (12).

The protocol receives as input x_i and z_j for all $j \in N_i$ and must be initialized at a pre-defined value $z_i(0)$. The value of x_i is used in (13) to compute l_i according to (9). The protocol uses the estimate $\hat{a}_{i,ss}$ to return as output the best response u_i^* as in (13), where $\hat{a}_{i,ss}$ represents the steady state value assumed by $\hat{a}_i(\tau)$, namely

$$\hat{a}_{i,ss} = \lim_{\tau \to T^-} \hat{a}_i(kT + \tau), \quad \text{for all } i \in \Gamma. \quad (14)$$

In the rest of this section, we present a distributed protocol $\Pi = \{(f_i, \phi_i) : \text{ for all } i \in \Gamma\}$ proposed by the authors in [8], such that the steady state estimate coincides with the current number of active players and with $\bar{\lambda}$, i.e., $\hat{a}_{i,ss} = a = \sum_{i \in \Gamma} \delta(u_i) = \bar{\lambda}$. Actually, the latter condition is sufficient for the convergence to the Pareto optimal Nash equilibrium of Theorem 3.

Assume that the transmitted information $z_i(\tau)$ is the current estimate of the percentage of active players. For instance, $z_i(\tau) = 0.2$ means that the ith player estimates only a twenty percent of active players. Then, given the percentage of active players $z_i(\tau)$, the estimate of the number of active players is simply

$$\hat{a}_i(\tau) = \phi(z_i(\tau)) = n z_i(\tau).$$

The protocol starts by assuming that all the players are active. This corresponds to initialize the transmitted states $z_i(0) = 1$ or which is the same the estimates $\hat{a}_i(0) = n$ for all $i \in \Gamma$.

Then, each player averages its estimate on-line on the basis of neighbors' estimates. If we denote by $z(\tau) = \{z_i(\tau)\}_{i \in \Gamma}$, the averaging process can be described by

$$f_i(z(\tau)) = -L_{i\bullet} z(\tau) - \Delta(t - t_i)$$

where $L_{i\bullet}$ is the ith row of the Laplacian matrix (see, e.g., [14], [18] for details), and $\Delta(t - t_i)$ is an impulse signal due to which $z_i(t_i^-)$ switches to a lower value $z_i(t_i^+)$. Such a switch has the meaning of a correction term acting at any time t_i where the estimate $\hat{a}_i(t_i)$ crosses from above the threshold l_i and consequently the ith player is no longer willing to be active. Impulses may be activated only after the transient evolution of $\dot{z}_i(\tau)$ has expired. We assume that this occurs after t_f time units, where t_f is an estimate of the worst case possible settling time of the protocol dynamics. A standard result in graph theory is that the settling time decreases as the number of edges in the network increases. Actually, the speed of convergence depends on the second smallest in magnitude eigenvalue of the Laplacian (known as Fiedler eigenvalue) in the sense that the higher (in magnitude) the Fiedler eigenvalue the faster the convergence [15]. In the light of the above consideration, t_i is the first sampled time rt_f, with $r = 0, 1, \ldots$ where function $\delta(\tilde{u}(\hat{a}_i(rt_f), l_i))$ reaches zero, namely

$$t_i = \arg\min_{r \in \mathbb{N}} rt_f \quad (15)$$
$$s.t. \quad \delta(\tilde{u}(\hat{a}_i(rt_f), l_i)) = 0. \quad (16)$$

Note that there may exist players characterized by $l_i > n$, for which $t_i = 0$, and players that never satisfy condition (16), for which $t_i = T$. Observe that, as players are indexed by increasing thresholds, it must also hold $T \geq t_1 \geq t_2 \geq \ldots \geq t_n \geq t_{n+1} = 0$. Furthermore, note that the evolution of the sampled values $z(rt_f)$ for $r = 0, 1, \ldots$ is monotonically decreasing which implies that the impulse may be activated only one time for each player (once you exit the group you are no longer allowed to rejoin it).

Theorem 4: It holds $\hat{a}_{i,ss} = a = \sum_{i \in \Gamma} \delta(u_i) = \bar{\lambda}$ for all $i \in \Gamma$.

Proof: With in mind the values t_i as in (15), let us set $t_{n+1} = 0$, $t_0 = T$ and consider the sequence of increasing discrete times $t_{n+1}, t_n, \ldots, t_{j+1}, t_j, \ldots, t_0$. Also denote recursively by $M(t_j) = \{i \in \mathcal{A}(t_j) : l_i > |\mathcal{A}(t_j)|\}$, where $\mathcal{A}(t_j) = \Gamma \setminus \bigcup_{k=j+1}^{n+1} M(t_k)$, and $\mathcal{A}(t_{n+1}) = \Gamma$. Roughly speaking, $\mathcal{A}(t_j)$ is the set of players that are willing to be active at time t_j whereas $M(t_j)$ is the set of players that are no longer willing to be active from time t_j on. Then the evolution of $\hat{a}_i(\tau)$ follows the discrete time dynamics

$$\hat{a}_i(t_{j-1}) = \hat{a}_i(t_j) - |M(t_j)|, \quad \text{for all } i \in \Gamma.$$

The above dynamics is monotonic decreasing and converges at the first time t_j where $\mathcal{A}(t_j)$ is a compatible set. To see this, note that if $\mathcal{A}(t_j)$ is compatible then $M(t_j) = \emptyset$, and therefore

$$\hat{a}_i(T) = \ldots = \hat{a}_i(t_{j-1}) = \hat{a}_i(t_j), \quad \text{for all } i \in \Gamma.$$

The above equation implies that $t_{j-1} = t_{j-2} = \ldots = T$, which means that condition (16) is never met for player $j-1$, if exists, and for all its predecessors, if any. In the extreme case, we may have $\mathcal{A}(t_j) = \ldots = \mathcal{A}(t_1) = \emptyset$ which means $t_j < T$ for all $j \in \Gamma$ and also that condition (16) is met for all players $j \in \Gamma$. We have then proved that the above dynamics converges when $\mathcal{A}(t_j)$ is compatible. It is left to show that the compatible set $\mathcal{A}(t_j)$ is the maximal one, namely, $\mathcal{A}(t_j) = \overline{C}$. We show this, by proving that if $\mathcal{A}(t_k) \supseteq \overline{C}$ then $\mathcal{A}(t_{k-1}) \supseteq \overline{C}$ for all $k = j+1, \ldots, n+1$. By contradiction, if $\mathcal{A}(t_{k-1}) \not\supseteq \overline{C}$, there must exist a player $i \in M(t_k)$ such that $l_i \leq |\overline{C}| \leq |\mathcal{A}(t_k)|$ but the latter fact is not possible from the definition of $M(t_k)$. We conclude the proof by observing that $\bigcap_{k=j+1}^{n+1} M(t_k) = \emptyset$ and consequently

$$\hat{a}_i(t_j) = n - \textstyle\sum_{k=j+1}^{n+1} |M(t_k)| = |\Gamma \setminus \bigcup_{k=j+1}^{n+1} M(t_k)| = |\mathcal{A}(t_j)| = |\overline{C}| = \bar{\lambda}.$$

■

V. A BEST RESPONSE PATH ALGORITHM

We have shown that the game (1)-(2) is a potential game as it always admits a potential function (see Theorem 1). Potential games have the strong property that any best response path converges to a Nash equilibrium. By best response path we intend a sequence of joint decisions $\mathbf{u}(0) \to \mathbf{u}(1) \to \ldots$ where $\mathbf{u}(j) = \{\mathbf{u}_1(j) \ldots \mathbf{u}_n(j)\}$ and $\mathbf{u}_i(j)$ is the vector of decisions (over the horizon) of player i at iteration j. Define a function $\sigma : \mathbb{N} \to \Gamma$, which returns a player for each iteration j of the sequence, i.e., $\sigma(1) = 2$, $\sigma(2) = 5 \ldots$ means that at iteration 1, only player 2 updates its decision, whereas at iteration 2, only player 5 updates its decision. By updating a decision we simply mean replacing the current decision by the best response. It may happen that the current decision is already the best response and then the updated decision coincides with the current decision. Now, each joint decision $\mathbf{u}(j+1)$ is obtained from $\mathbf{u}(j)$ by an unilateral improvement on the part of player $i = \sigma(j)$, i.e., $\mathbf{u}(j+1) = [\mathbf{u}_i^*, \mathbf{u}_{-i}(j)]$ and $\mathbf{u}_i^* = \{u_i^{0*}, \ldots, u_i^{N*}\}$ is the solution of (6) for fixed $\mathbf{u}_{-i}(j+1) = \mathbf{u}_{-i}(j)$.

More precisely, at iteration j, let the current decision be $\mathbf{u}(j) = \{\mathbf{u}_1(j), \ldots, \mathbf{u}_n(j)\}$ with $\mathbf{u}_i(j) = \{u_i^0(j), \ldots, u_i^N(j)\}$ for $i = 1, \ldots, n$. To solve (6) player $i = \sigma(j)$ needs to estimate the number of active players over the horizon. This is possible by modifying the protocol presented in the previous section. For fixed $\mathbf{u}(j)$, denote the vector of decisions at time k by $u^k(j) = \{u_i^k(j)\}_{i \in \Gamma}$, then the protocol $\Pi = \{(f_i, \phi_i) : \text{for all } i \in \Gamma\}$, where

$$f_i^k(z(\tau)) = -L_{i\bullet} z^k(\tau), \quad z_i^k(0) = \delta(u_i^k(j)) \quad (17)$$
$$\hat{a}_i^k(\tau) = \phi(z_i^k(\tau)) = n z_i^k(\tau). \quad (18)$$

is such that $\hat{a}_{i,ss}^k = a(u^k(j))$. Remind that $a(u^k(j))$ is the number of active players at stage k given the decision vector $u^k(j)$. Repeating the same argument for $k = 0, \ldots, N$ (we can run the protocol in parallel) the ith player can estimate the number of active players over the horizon $\mathbf{a}^0(j)$ associated to the current decision $\mathbf{u}(j)$, namely, $\mathbf{a}^0(j) = \{a(u^0(j)), \ldots, a(u^N(j))\}$ with $a(u^k(j)) = \sum_{i \in \Gamma} \delta(u_i^k(j))$. In the light of the above comments, we show below the pseudo code of an algorithm that, for a given function $\sigma(.)$, returns a best response path and consequently converges to a Nash equilibrium. Let $\mathbf{u}_i(j)$ be the solution (decisions of player i) at iteration j, then

$j = 0$; `WHILE` `not converging`
$\{i = \sigma(j)$, `compute` $\mathbf{a}^0(j)$ `from (17)-(18)`
`using current` $\mathbf{u}(j)$
`update` $\mathbf{u}_i(j+1) = \mathbf{u}_i^*$
`solution of (6)`
`based on` $\mathbf{a}^0(j)$,
$j := j + 1\}$

The algorithm eventually converges to a Nash equilibrium which depends on the chosen function $\sigma(.)$. However, the choice of any generic function $\sigma(.)$ do not compromise the convergence of the algorithm. The number of iterations is at most 2^{nN}. Actually, the best response for player i does not depend on the value of u_{-i}, but only on the number of active players. Also, the algorithm can be stopped if no players have changed their decisions in the last n iterations.

w_1	4	8	6	5	7	8	4	5	6	8
w_2	0	0	1	7	8	0	6	2	1	4
w_3	0	3	2	0	3	1	1	3	3	0

TABLE I

In the next section we use the above algorithm in a multi-inventory application.

VI. MULTI-INVENTORY APPLICATION

Each player $i \in \Gamma$ is a retailer, the state $x_i^k \in \mathbb{Z}$ is the ith inventory, $u_i^k \in U_i^k = \mathbb{N}$ is the ordered quantity. Let $w_i^k \in \mathbb{N}$ be a deterministic demand, the inventory dynamics is

$$x_i^{k+1} = x_i^k + u_i^k - w_i^k. \tag{19}$$

Let c be the purchase cost per stock unit, h the penalty on holding, p the penalty on shortage, and K_i^k the transportation cost charged to the ith retailer that replenishes at stage k. Also, let us make the common assumption that $c - p < 0$. The stage cost for the ith retailer is

$$\begin{aligned} &g_i(x_i^k, u_i^k, a^k) = \\ &\underbrace{K_i^k}_{\psi(a^k(u^k))} \ \delta(u_i^k) + \underbrace{cu_i^k + p\max(0, -x_i^{k+1}) + h\max(0, x_i^{k+1})}_{\gamma(x_i^k, u_i^k)}. \end{aligned} \tag{20}$$

Here $\psi(a^k(u^k))$ is monotone since the active retailers may share the same truck for their supplies and so the more they are, the less each of them pays for the transportation.

Example 1: Consider three retailers and parameters $K = 24$, $p = 8$, $h = 1$, $c = 2$. Retailers face a deterministic demand over the horizon of ten stages (see Table I). The initial state is $x^0 = [0 \quad 0 \quad 0]$. Let us run the algorithm of the previous section in order to obtain a best response path. The retailers, at the first iteration, do not consider the possibility of sharing the transportation cost. No communication occurs among the retailers and they replenish in a fully uncoordinated fashion as displayed in Fig. 1, left column. The absence of coordination is evident as retailer 1 replenishes on days $0, 2, 5$ and 8 (top-left), retailer 2 on day 3 and 6 (middle-left), while retailer 3 on days 1 and 7 (bottom-left). At a second iteration, the 3rd retailer ($\sigma(2) = 3$) estimates the number of active players over the horizon by running the protocol (17)-(18) and finds its best response by solving (6) and finds its best response. The same argument is repeated at the successive iterations letting the retailers unilaterally improving their payoffs one after the other. The algorithm converges in six iterations. The supply decisions at Nash equilibrium are displayed in Fig. 1, right column. Here you can notice that retailers 1 and 3 replenish on day 1, retailers 1, 2 and 3 replenish on day 3 and 7, and retailer 1 and 2 on day 5.

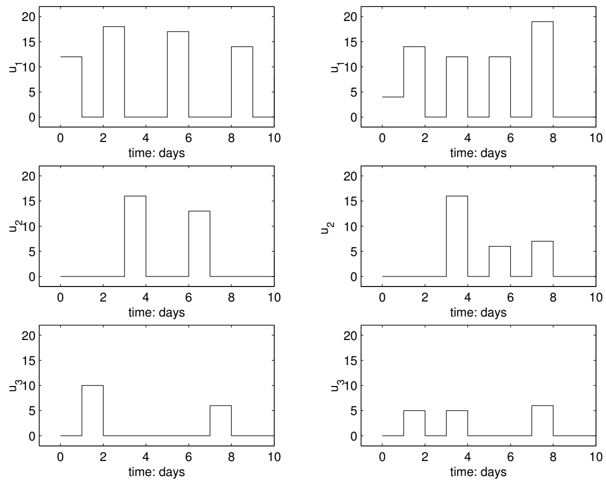

Fig. 1. Uncoordinated (left column) and coordinated (right column) supply strategies.

REFERENCES

[1] G. Arslan, J.R. Marden and J.S. Shamma, "Autonomous vehicle-target assignment: A game theoretical formulation", *ASME Journal of Dynamic Systems, Measurement, and Control*, special issue on "Analysis and Control of Multi-Agent Dynamic Systems", 584–596, 2007.

[2] G. Arslan and J. S. Shamma, "Distributed Convergence to Nash Equilibria with Local Measurements", *Proceedings of the IEEE CDC*, The Bahamas, 2004, pp. 1538-1543.

[3] Axsäter, S., "A Framework for Decentralized Multi-Echelon Inventory Control", *IIE Transactions*, vol. 33, no. 1, 2001, pp. 91–97.

[4] Basar, T. and Olsder, G.J., *Dynamic Noncooperative Game Theory*, Academic Press, 2nd edtn, London, 1995.

[5] Bauso D., Giarré, L. and Pesenti, R., "Consensus in Noncooperative Dynamic Games: a Multi-Retailer Inventory Application", to be published in *IEEE Transaction on Automatic Control*, 2008.

[6] Bauso, D., Giarré, L., and Pesenti, R., "Nonlinear Protocols for the Optimal Distributed Consensus in Networks of Dynamic Agents", *Systems and Control Letters*, 55(11), 2006, pp. 918–928.

[7] D. Bauso, L. Giarré and R. Pesenti. "Mechanism Design for Optimal Consensus Problems" *IEEE Conference on Decision and Control*, San Diego, Ca, 2006, pp. 3381-3386.

[8] Bauso, D., Giarré, L. and Pesenti, R., "Distributed Consensus Protocols for Coordinating Buyers", in *Proc. of the IEEE Conference on Decision and Control*, Maui, Hawaii, Dec. 2003, vol. 1, pp. 588–592.

[9] Fudenberg, D., Levine, D.K., *The Theory of Learning in Games*, MIT Press, 1998.

[10] J.R. Marden, G, Arslan and J.S. Shamma, "Connections between Cooperative Control and Potential Games illustrated on the Consensus Problem. *European Control Conference*, 2007.

[11] Meca, A., García Jurado, I., and Borm, P., "Cooperation and Competition in Inventory Games", *Mathematical Methods of Operation Research*, vol. 57, no. 3, 2003, pp. 481–493.

[12] Meca, A., Timmer, J., Garccía Jurado, I., and Borm, P., "Inventory Games", *European Journal of Operational Research*, vol. 156/1 pp. 127–139, 2004.

[13] Nagarajan, M., and Sósic, G., "Game-theoretic analysis of cooperation among supply chain agents: Review and extensions", *European Journal of Operational Research* (in press), 2006.

[14] Olfati-Saber, R., J. A. Fax, R. M. Murray, "Consensus and Cooperation in Networked Multi-Agent Systems," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 215-233, Jan. 2007.

[15] Olfati-Saber, R., and R. M. Murray, "Consensus Problems in Networks of Agents with Switching Topology and Time-Delays", *IEEE Transactions on Automatic Control*, vol. 49, no. 9, 2004, pp. 1520–1533.

[16] W. Ren, R. W. Beard, and E. Atkins, "Information Consensus in Multivehicle Cooperative Control: Collective Group Behavior through Local Interaction," IEEE Control Systems Magazine, Vol. 27, Issue 2, 71-82, 2005.

[17] Rosenthal, R. W., "A Class of Games Possessing Pure-Strategy Nash Equilibria", *Int. J. Game Theory* 2, 1973, pp. 65-67.

[18] Shamma, J. S. and G. Arslan, "Dynamic Fictitious Play, Dynamic Gradient Play, and Distributed Convergence to Nash Equilibria", *IEEE Transactions on Automatic Control*, vol. 50, no. 3, 2005, pp. 312–327.

[19] Young, H. P., *Individual Strategy and Social Structure: An Evolutionary Theory of Institutions*, Princeton University Press, 1988.