49th IEEE Conference on Decision and Control
December 15-17, 2010
Hilton Atlanta Hotel, Atlanta, GA, USA

Distributed Seeking of Nash Equilibria in Mobile Sensor Networks

Miloš S. Stanković, Karl Henrik Johansson and Dušan M. Stipanović

Abstract— **In this paper we consider the problem of distributed convergence to a Nash equilibrium based on minimal information about the underlying noncooperative game. We assume that the players/agents generate their actions based only on measurements of local cost functions, which are corrupted with additive noise. Structural parameters of their own and other players' costs, as well as the actions of the other players are unknown. Furthermore, we assume that the agents may have dynamics: their actions can not be changed instantaneously. We propose a method based on a stochastic extremum seeking algorithm with sinusoidal perturbations and we prove its convergence, with probability one, to a Nash equilibrium. We discuss how the proposed algorithm can be adopted for solving coordination problems in mobile sensor networks, taking into account specific motion dynamics of the sensors. The local cost functions can be designed such that some specific overall goal is achieved. We give an example in which each agent/sensor needs to fulfill a locally defined goal, while maintaining connectivity with neighboring agents. The proposed algorithms are illustrated through simulations.**

I. Introduction

Problems of distributed, multi-agent optimization, coordination, estimation and control have been in the focus of significant research in past years. Depending on the problem setup and the available resources, agents may have access to different measurements, different *a priori* information, such as system models and sensor characteristics, and different inter-agent communication channels. One approach to these problems is *game theoretic*, since the agents can be treated as players in a game. In this way, a decentralized optimization or coordination problem can be formulated as a noncooperative game, where the players are selfishly trying to optimize their local cost functions, based on locally available information. Depending on the structure of the game, and the design of the local cost functions, the Nash equilibria of the underlying game can have different properties and they may or may not correspond to the optimal solution to some global optimization problem [1]–[7].

The focus of this paper is on the problem of *learning in games*, or designing the algorithms that converge to a Nash equilibrium. The majority of the existing literature in this area is focused on the model-based algorithms; that is, the algorithm is designed having in mind a specific form of the players cost functions. Furthermore, it is usually assumed that the players can observe the actions of the other players. In this way, the algorithms can be designed on the basis of the "best response" strategy. For example, in [8], convergence properties have been analyzed for such a class of infinite, convex games. For the games with finite action sets, where the players can use mixed strategies, the convergence of the underlying best response algorithm, called fictitious play and its modifications have been analyzed intensively (see [9] and references therein). The recently proposed algorithms in [10] and [11] deal with an information structure similar to the one imposed in this paper, but require synchronization between the agents, and the convergence is proved only for a special class of games (weakly acyclic or potential games). Also, non of the mentioned approaches deals with the dynamic nature of the players while also taking into account the measurement noise. A similar approach to the one proposed in this paper, applied to deterministic games in markets, has appeared independently in [12].

On the other hand, the *extremum seeking* algorithms have received significant attention recently for dealing with nonmodel-based online optimization problems involving dynamical systems. The basic algorithm, based on introducing sinusoidal perturbations, has been treated in [13]. In [14] and [15] a time varying version of the algorithm has been introduced, whose convergence, with probability one, has been proved in the presence of measurement noise. It has been demonstrated how this technique can be applied to autonomous vehicles source seeking in deterministic environments [16], or optimal positioning in stochastic environments [14], [17].

In this paper we propose an algorithm for distributed seeking of a pure Nash strategy in infinite games where the players are generating their actions, based solely on the measurements of their local cost functions, whose analytical form is unknown. Furthermore, similarly as in the extremum seeking problems, it is assumed that the agents may have some local dynamics, so that their inputs are filtered through stable filters before affecting the cost functions; hence, the actions can not be changed instantaneously. Also, the local measurements of cost functions are not available directly, they are filtered through a stable filter, and corrupted with measurement noise. The proposed algorithm is based on the time-varying extremum seeking scheme with sinusoidal perturbations, under stochastic noise, analyzed in [15]. We formulate necessary conditions regarding the structure of the players' cost functions and regarding the parameters of the proposed distributed scheme, under which we prove almost sure (a.s.) convergence to a Nash equilibrium.

The proposed Nash equilibrium seeking algorithm is ap-

Miloš S. Stanković and Karl Henrik Johansson are with the ACCESS Linnaeus Center, School of Electrical Engineering, Royal Institute of Technology, 100-44 Stockholm, Sweden; E-mail: milsta@kth.se, kallej@kth.se

Dušan M. Stipanović is with the Department of Industrial and Enterprise Systems Engineering and the Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA; E-mail: dusan@illinois.edu

pealing for dealing with distributed coordination and optimization problems within *mobile sensor networks*, having in mind that it does not require a detailed model of the problem, that it is possible to include agents' dynamics, and that it copes with a stochastic environment. We give an example on how to design local cost functions such that the agents can fulfill some locally defined objectives, while maintaining connectivity with the neighboring agents. The proposed framework can also tackle more general distributed optimization problems within mobile sensor networks, formulated as noncooperative games (*e.g.*, coverage control, target assignment, randevouz). The existing literature in the area of mobile sensor networks, in which game theory or distributed optimization have been applied are based on specific sensing models, do not take robot dynamics into consideration or are only focused on specific scenarios (*e.g.*, [18], [19], [11], [20] and references therein).

In Section II the problem setup and the algorithm description is given. Section III is devoted to the convergence analysis of the algorithm. In Section IV we discuss possible applications within mobile sensor networks where the vehicle dynamics can include single or double integration. Simulation results for the network of three agents are shown and discussed in Section V.

II. A Nash Equilibrium Seeking Algorithm

We consider a scenario in which N agents are noncooperatively minimizing their local cost functions by only updating their local actions, based on their current local information. We assume that the actions of the players belong to infinite spaces ($\mathcal{R}^{m_i}, i = 1, ..., N$). Hence, we are dealing with a noncooperative static game with infinite strategy spaces of the players where the optimality is characterized by a (pure) Nash equilibrium; a point from which neither agent have incentive to deviate [5].

We assume that the information that each player has about the underlying game is restricted solely to the measurements of its local cost function, which are, in addition, filtered through an unknown stable filter and corrupted with a measurement noise. The players do not have any direct information about either the underlying structure of the game or the actions of the other players. Motivated by the fact that the formulated nonmodel based information structure resembles the one in the extremum seeking problems, we propose an algorithm based on sinusoidal perturbations, depicted in Fig. 1.

Without loss of generality, we will assume that each agent's strategy space is two dimensional, $u_i = (x_i, y_i) \in \mathcal{R}^2$, having in mind that we will apply this methodology to vehicle coordination problems in the plane. The framework can be extended to multi-dimensional strategy spaces in a straightforward way. Furthermore, we assume that the agents may have some local dynamics, so that their actions are filtered through a stable filter, having the transfer function matrix $F_i(z)$, before affecting the measured cost function $J_i(u_i, u_{-i})$, where u_i denotes the action of agent i, while u_{-i} denotes the actions of all the other agents. In general,

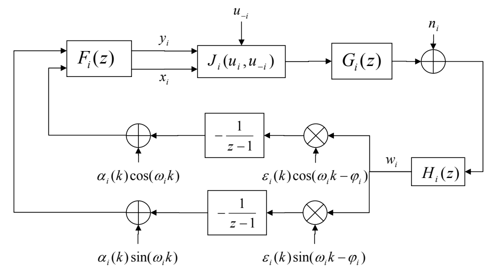

Fig. 1. Nash equilibrium seeking scheme

each cost function J_i doesn't necessarily have to depend on the actions of all the other players. So, let us define neighbor sets $\mathcal{N}_i$, $i = 1, ..., N$, whose elements are indices of the agents whose actions affect the i-th agent's cost function.

As shown in Fig. 1, each agent implements a local extremum seeking loop. The estimation of the gradient of the local cost function is performed by inserting a sinusoidal perturbation, with frequency ω_i, which, by passing through the function J_i, is being modulated by its local slope. The estimate of the slope (found by the multiplication/demodulation with the same frequency sinusoid) is then used to move (by integration) in the opposite direction. Since all the information needed to estimate the gradient is located in the amplitude of the modulated sinusoidal perturbation, the measurements are filtered by high pass filters $H_i(z)$ in order to eliminate any DC components, and, hence, to improve the overall convergence properties. Also, in order to improve the convergence properties, low pass filters can be added in the loop as part of the dynamics $G_i(z)$ and $F_i(z)$. Local decoupling between x and y dimensions is obtained using orthogonal perturbations: sine for x, and cosine for y (*cf.*, [17] or [16]). Furthermore, neighboring agents apply *different frequency* perturbations, $\omega_i \neq \omega_j$, $j \in \mathcal{N}_i$, so that decoupling between their gradient estimates is achieved.

Let us assume that the local cost functions $J_i(u_i, u_{-i})$ are continuously differentiable, strictly convex in local decision variables u_i and that $J_i(u_i, u_{-i}) \to \infty$ when $\|u_i\| \to \infty$. These conditions guarantee existence of a Nash equilibrium in pure strategies [5], [21]. Under these conditions, a necessary and sufficient condition for achieving a Nash equilibrium is that all the gradients of $J_i(u_i, u_{-i})$ with respect to u_i are equal to zero:

$$g_u(u^*) = [\nabla_1 J_1(u^*)^T, ..., \nabla_N J_N(u^*)^T]^T = 0, \quad (1)$$

where u^* denotes a Nash equilibrium and $\nabla_i J_i(u^*)$, $i = 1, ..., N$, denotes the gradient of J_i with respect to local actions u_i.

The following equations model the behavior of the system:

$$w_i(k) = H_i(z)[G_i(z)[J_i(u_i(k), u_{-i}(k))] + n_i(k)], \quad (2)$$

$$u_i(k) = u_0^i(k) + F_i(z)\left[-\frac{1}{z-1}[\xi_i(k)]\right], \quad (3)$$

$$\xi_i(k) = \varepsilon_i(k) C_i(k) w_i(k), \quad (4)$$

for $i = 1,...,N$, where $n_i(k)$ is the measurement noise of agent i, $u_0^i(k) = F_i(z)[\alpha_i(k)\cos(\omega_i k)\ \alpha_i(k)\sin(\omega_i k)]^T$ and $C_i(k) = [\cos(\omega_i k - \varphi_i)\ \sin(\omega_i k - \varphi_i)]^T$. Throughout the paper, the expression $Y(z)[x(k)]$ denotes a time domain vector whose components are obtained as the outputs of the transfer function matrix $Y(z)$ applied to the input vector $x(k)$.

For each agent, we define the tracking error as:

$$\tilde{u}_i(k) = u_i^* - u_i(k) + u_0^i(k), \tag{5}$$

where u_i^* is the i-th agent's action in the Nash equilibrium. From the above equations it is easy to obtain the following difference equation for the overall tracking error:

$$\tilde{u}(k+1) = \tilde{u}(k) + F(z)[\varepsilon(k)C(k)w(k)], \tag{6}$$

where $\tilde{u}(k) = [\tilde{u}_1(k)^T, \dots, \tilde{u}_N(k)^T]^T$,
$F(z) = \text{diag}\ \{F_1(z), \dots, F_N(z)\}$,
$\varepsilon(k) = \text{diag}\ \{\varepsilon_1(k), \dots, \varepsilon_N(k)\} \otimes I_2$,
$C(k) = \text{diag}\ \{C_1(k), \dots, C_N(k)\}$,
$w(k) = [w_1(k), \dots, w_N(k)]^T$, I_2 is 2x2 identity matrix and $\otimes$ denotes the Kronecker product.

III. Convergence Analysis

In the convergence analysis, we will assume that the following assumptions are satisfied:

(A.1) The random vectors $n(k)$ (where $n(k) = [n_1(k), \dots, n_N(k)]^T$) are mutually independent and zero mean and they satisfy

$$E\{n(k)n(k)^T\} = \Sigma(k) \leq \Gamma,\ k = 1, 2, ... \tag{7}$$

for some matrix $\Gamma \geq 0$, $||\Gamma|| < \infty$ (the notation $A \leq B$ means that the matrix $B - A$ is positive semidefinite; $||\cdot||$ denotes any matrix norm).

(A.2) The scalar sequences $\varepsilon_i(k)$ are decreasing, $\varepsilon_i(k) > 0$, $k = 1, 2, ...$, and $\lim_{k\to\infty} \varepsilon_i(k) = 0$, $i = 1, ..., N$.

(A.3) The scalar sequences $\alpha_i(k)$ are decreasing, $\alpha_i(k) > 0$, $k = 1, 2, ...$, and $\lim_{k\to\infty} \alpha_i(k) = 0$, $i = 1, ..., N$.

(A.4) $\sum_{k=1}^{\infty} \varepsilon_i(k)\alpha_i(k) = \infty$, $i = 1, ..., N$.

(A.5) $\sum_{k=1}^{\infty} \varepsilon_i(k)\varepsilon_j(k) < \infty$ for all $i = 1, ..., N$ and $j \in \mathcal{N}_i \cup \{i\}$.

(A.6) $\varepsilon_i(k)\alpha_i(k) = O(\varepsilon_j(k)\alpha_j(k))$ for all $i, j = 1, ..., N$.

According to (A.6), $\varepsilon_i(k)\alpha_i(k)$ can be written as

$$\varepsilon_i(k)\alpha_i(k) = \min_j[\varepsilon_j(k)\alpha_j(k)](c_i + o(\varepsilon_i(k)\alpha_i(k))), \tag{8}$$

for each $i = 1, ..., N$ and for some constants $c_i > 0$.

(A.7) $\tilde{u}(k) \in B$ a.s. for all $k = 1, 2, ...$, where B is a compact connected subset of $\mathcal{R}^{2N}$ containing the origin. $J_i(u)$, $u = [u_1^T\ \dots\ u_N^T]^T$, $i = 1, ..., N$, are analytic in an open set B_u, containing u^*, which is related to set B in such a way that for any point $\tilde{u} \in B$, $u^* - \tilde{u} + u_0(k) \in B_u$, for all $k = 1, 2, ...$ (in accordance with (5), where $u_0(k) = [u_0^1(k)^T, ..., u_0^N(k)^T]^T$).

(A.8) There exists a continuously differentiable Lyapunov function $V(\tilde{u})$ such that $V(0) = 0$ and

$$-g^T(\tilde{u})K^T\nabla_{\tilde{u}}V(\tilde{u}) < 0, \tag{9}$$

for all $\tilde{u} \neq 0$, $\tilde{u} \in B$, where $g(\tilde{u}) = g_u(u^* - \tilde{u})$, $K = \text{diag}\{c_1K_1, ..., c_NK_N\}$,
$K_i = F_i(1)\begin{bmatrix} \text{Re}\{\theta_i\} & \text{Im}\{\theta_i\} \\ -\text{Im}\{\theta_i\} & \text{Re}\{\theta_i\} \end{bmatrix}$, $\theta_i = e^{j\varphi_i}F_i(e^{j\omega_i})G_i(e^{j\omega_i})H_i(e^{j\omega_i})$, and $\nabla_{\tilde{u}}V(\tilde{u})$ denotes the gradient of $V(\tilde{u})$.

(A.9) $\omega_i \in (0, \pi)$ and $\omega_i \neq \omega_j$ for all $i = 1, ..., N$ and $j \in \mathcal{N}_i$.

Observe here that Assumption (A.8), besides stability of our algorithm, also ensures uniqueness of the Nash equilibrium u^* (see also [21] where stability and uniqueness have been ensured with the, so called, strict diagonal convexity condition). The following theorem deals with the asymptotic behavior of the algorithm.

Theorem 1. Consider the multi-agent system with Nash equilibrium seeking scheme defined in (2)-(4) and shown in Fig. 1. Let the Assumptions (A.1)–(A.9) be satisfied. Then the actions $u(k) = [u_1(k)^T\ \dots\ u_N(k)^T]^T$ of the players converge to the Nash equilibrium u^* almost surely.

Proof. Recall that the tracking error for each agent satisfies:

$$\tilde{u}_i(k+1) - \tilde{u}_i(k) = F_i(z)[\varepsilon_i(k)C_i(k)H_i(z)[G_i(z)[J_i(u_i(k), u_{-i}(k))] + n_i(k)]]. \tag{10}$$

Since we have assumed that the functions $J_i(u_i, u_{-i})$ are analytic in the region B_u containing u^* (Assumption (A.7)) one can obtain their Taylor series expansion around the Nash equilibrium point u^*, and by using (5) write it as the sum of three terms defined below:

$$J_i(u_i(k), u_{-i}(k)) = L_i(k) + D_i(k) + d_i(k). \tag{11}$$

The first term $L_i(k)$ is linear with respect to the perturbation signal u_0^i; therefore, it is essential for achieving the adequate approximation of the gradient of the cost function (since it will be demodulated by the multiplication with $C_i(k)$). It is given by:

$$L_i(k) = u_0^i(k)^T\nabla_i J_i(u^* - \tilde{u}). \tag{12}$$

The term $d_i(k)$ in (11) contains the deterministic input terms (not depending on any $\tilde{u}_i$, $i = 1, ..., N$) and $D_i(k)$ contains all the remaining terms. Now we focus on the term $F_i(z)[\varepsilon_i(k)C_i(k)M_i(z)[L_i(k)]]$ obtained from (10) and (11), where $M_i(z) = H_i(z)G_i(z)$, since it is essential for achieving the contraction of the tracking error. By plugging (12) into (11) and then into (10), applying a modulation lemma (*e.g.*, Lemma 2 from [22]), and taking into account multiplication with $C_i(k)$, after some algebra, one obtains the following equation:

$$C_i(k)M_i(z)[L_i(k)] = Q_i(z)\left[A_i(k)\nabla_i J_i(u^* - \tilde{u})\right] + S_i(k)P_i(z)\left[A_i(k)\nabla_i J_i(u^* - \tilde{u})\right], \tag{13}$$

where $Q_i(z) = \begin{bmatrix} Q_i^1(z) & Q_i^2(z) \\ Q_i^2(z) & -Q_i^1(z) \end{bmatrix}$, $Q_i^1(z) = -\text{Re}\{e^{j\varphi_i}M_i(e^{j\omega_i}z)\}$, $Q_i^2(z) = \text{Im}\{e^{j\varphi_i}M_i(e^{j\omega_i}z)\}$,

$A_i(k) = \begin{bmatrix} \alpha_1^i(k) & \alpha_2^i(k) \\ \alpha_2^i(k) & -\alpha_1^i(k) \end{bmatrix}$, $\alpha_1^i(k) = \mathrm{Re}\{F_i(e^{j\omega_i}z)[\alpha_i(k)]\}$, $\alpha_2^i(k) = \mathrm{Im}\{F_i(e^{j\omega_i}z)[\alpha_i(k)]\}$, $P_i(z) = \begin{bmatrix} P_i^1(z) & P_i^2(z) \\ P_i^2(z) & -P_i^1(z) \end{bmatrix}$, $P_i^1(z) = -\mathrm{Re}\{M_i(e^{j\omega_i}z)\}$, $P_i^2(z) = \mathrm{Im}\{M_i(e^{j\omega_i}z)\}$ and $S_i(k) = \begin{bmatrix} \cos(2\omega_i k - \varphi_i) & \sin(2\omega_i k - \varphi_i) \\ \sin(2\omega_i k - \varphi_i) & -\cos(2\omega_i k - \varphi_i) \end{bmatrix}$.

Following the methodology developed in [15], after plugging (13) into (10), we decompose the first term in the following way:

$$F_i(z)\left[\varepsilon_i(k)Q_i(z)[A_i(k)\nabla_i J_i(u^* - \tilde{u})]\right] = \varepsilon_i(k)B_i(z)[A_i(k)\nabla_i J_i(u^* - \tilde{u})] + \delta_i^1(k), \quad (14)$$

where $B_i(z) = F_i(z)Q_i(z)$ and it can be shown [15] that $||\sum_{k=1}^{\infty}\delta_i^1(k)|| < \infty$ (a.s.), $i = 1, ..., N$. By applying similar decompositions (whose goal is to extract the essential term allowing the contraction of the algorithm, while proving that all the other terms are summable a.s.) it can be shown that the whole term (14) can be put in the form $-\varepsilon_i(k)\alpha_i(k)K_i\nabla_i J_i(u^* - \tilde{u}) + \delta_i(k)$, where $K_i = -B_i(1)A_f^i(1)$, $A_f^i(z) = \begin{bmatrix} \mathrm{Re}\{F_i(e^{j\omega_i}z)\} & \mathrm{Im}\{F_i(e^{j\omega_i}z)\} \\ \mathrm{Im}\{F_i(e^{j\omega_i}z)\} & -\mathrm{Re}\{F_i(e^{j\omega_i}z)\} \end{bmatrix}$ (compare with $A_i(k)$) and $\delta_i(k)$ contains all the summable terms, so that $||\sum_{k=1}^{\infty}\delta_i(k)|| < \infty$ (a.s.). It is easy to derive that the matrices K_i have the form as defined in Assumption (A.8).

Finally, by plugging this term into (13) and then back into the tracking equations (10), and by using (11) and (A.6), we obtain the tracking equation for the whole system:

$$\tilde{u}(k+1) = \tilde{u}(k) - \rho(k)Kg(\tilde{u}(k)) + \phi(k) + F(z)[\varepsilon(k)\pi(k)], \quad (15)$$

where K is as given in (A.8), $\phi(k) = \delta(k) + \varepsilon(k)S(k)P(z)[A(k)g(\tilde{u}(k))]$, $\pi(k) = C(k)M(z)[d(k) + D(k)] + C(k)H(z)[n(k)]$, $\alpha(k) = \mathrm{diag}\{\alpha_1(k), ..., \alpha_N(k)\} \otimes I_2$, $\delta(k) = [\delta_1(k)^T, ..., \delta_N(k)^T]^T$, $P(z) = \mathrm{diag}\{P_1(z), ..., P_N(z)\}$, $M(z) = \mathrm{diag}\{M_1(z), ..., M_N(z)\}$, $S(k) = \mathrm{diag}\{S_1(k), ..., S_N(k)\}$, $A(k) = \mathrm{diag}\{A_1(k), ..., A_N(k)\}$, $D(k) = [D_1(k), ..., D_N(k)]^T$, $d(k) = [d_1(k), ..., d_N(k)]^T$, $H(z) = \mathrm{diag}\{H_1(z), ..., H_N(z)\}$, $\rho(k) = \varepsilon(k)\alpha(k)$ and where we have incorporated the summable terms $\rho(k)[o(\varepsilon_i(k)\alpha_i(k)))]$ (according to (A.5)) in the term $\delta(k)$.

Now it is obvious that the recursive equation (15) is actually the Robbins-Monro algorithm to which we can directly apply Theorem 2.2.3 from [23], having in mind that, by Assumption (A.8), there exists a Lyapunov function $V(\tilde{u})$ that satisfies conditions of this theorem. Therefore, $\tilde{u} \to 0$ a.s. if the "error" term satisfies

$$\sum_{k=1}^{\infty}\{\phi(k) + F(z)[\varepsilon(k)\pi(k)]\} \text{ converges (a.s.).} \quad (16)$$

Having in mind that the filter $F(z)$ is linear and asymptotically stable, we can switch the summation and filtering in (16); hence it is enough to show that $\sum_{k=1}^{\infty}\phi(k) + \varepsilon(k)\pi(k)$ converges (a.s.). We have already shown that $\delta(k)$ is summable a.s.

Furthermore, all the terms in $\varepsilon(k)S(k)P(z)[A(k)g(\tilde{u}(k))]$ and in $\varepsilon(k)C(k)M(z)[D(k)]$, actually contain a sinusoidal signal multiplied with filtered terms having the following forms $\chi_1^{n_1}\chi_2^{n_2}\cdots\chi_N^{n_N}$, where χ_i denotes either x_i or y_i scalar coordinate and $n_i \in \{0, 1, 2, ...\}$ for all $i \in \{1, ..., N\}$. Therefore, by applying the same methodology exposed in [15] for proving convergence of similar sums, it can be shown that the sum of all these terms converges a.s. under the assumptions (A.2), (A.3), (A.5) and (A.7). It is important to observe that the terms in $D_i(k)$ containing the j-th perturbation u_0^j will be multiplied with a *different frequency* sinusoid contained in $C_i(k)$ (Assumption (A.9)). By converting this product of sinusoids into a summation, these terms will end up having the same, summable, form and can be treated in the same way as the other terms.

For the deterministic input terms, contained in $\varepsilon_i(k)C_i(k)M_i(z)[d_i(k)]$ it is obvious that all of them will have the form $c_0\varepsilon_i(k)\alpha_i(k)\sin(\omega(i,j)k + \phi(i,j))$, where $\omega(i,j)$ and $\phi(i,j)$ depend on ω_i, ω_j, φ_i and φ_j for $j \in \mathcal{N}_i$ and c_0 is a constant. This form is summable according to results in [15].

Finally, the stochastic input terms $\varepsilon_i(k)C_i(k)H_i(z)[n_i(k)]$, which are independent sequences filtered through stable filters $H_i(z)$ and multiplied with $\varepsilon_i(k)$, can be treated using the results from [24], which deal with stochastic approximation algorithms with colored noise. Under the adopted assumptions (A.2), (A.3) and (A.5) it can be shown that these terms satisfy the necessary conditions to be summable a.s. (see also [15] where a similar problem is treated in an analogous way).

Therefore, we have shown that (16) is satisfied, which proves the theorem. ■

Remark 1. If the underlying game is a *potential game* [4], the vector (1) will be equal to the gradient of the potential function. It is obvious that, in this case, we can choose this potential function (shifted such that $V(0) = 0$ and assuming its strict convexity) as a Lyapunov function, so that the condition (9) will always be satisfied if K is positive definite. Therefore, in this case, Assumption (A.8) can be replaced with the simple condition: $-\frac{\pi}{2} < \varphi_i + \mathrm{Arg}\left\{F_i(e^{j\omega_i})H_i(e^{j\omega_i})G_i(e^{j\omega_i})\right\} < \frac{\pi}{2}$, $F_i(1) > 0$, $i = 1, ..., N$, which ensures the positive definiteness of K. In fact, this condition ensures that the phase shift of the sinusoidal perturbation, induced by the filters $F_i(z)$, $H_i(z)$ and $G_i(z)$, is close enough to the phase shift $-\varphi_i$ of the multiplying sinusoids.

Remark 2. In the case of quadratic cost functions Assumption (A.8) has the following direct interpretation in terms of a Jacobian matrix stability. Assume that the cost functions are given by

$$J_i(u_i, u_{-i}) = u_i^T R_{ii}^i u_i + u_i^T r_i + k_i + \sum_{j\in\mathcal{N}_i} u_i^T R_{ij} u_j + u_j^T R_{jj}^i u_j + u_j^T r_j^i, \quad (17)$$

where $R_{ii}^i \in \mathcal{R}^{2\times 2}$, $R_{ii}^i > 0$, $R_{ij} \in \mathcal{R}^{2\times 2}$,$R_{jj}^i \in \mathcal{R}^{2\times 2}$, $r_i \in \mathcal{R}^2$, $r_i^j \in \mathcal{R}^2$. The condition (1) becomes now $Ru^* = -r$, where $r = [r_1^T, ..., r_N^T]^T$ and

$$R = \begin{bmatrix} 2R_{11}^1 & R_{12} & \dots & R_{1N} \\ R_{21} & 2R_{22}^2 & \dots & R_{2N} \\ \vdots & \vdots & & \vdots \\ R_{N1} & R_{N2} & \dots & 2R_{NN}^N \end{bmatrix}. \qquad (18)$$

It is obvious that, in this case $g(\tilde{u}) = R\tilde{u}$. Therefore, we can choose a quadratic Lyapunov function $V(\tilde{u}) = \tilde{u}^T P \tilde{u}$, where $P > 0$ is chosen such that the condition (9) is satisfied. Such a matrix $P > 0$ will always exist if the matrix KR is stable (Hurwitz). If we assume that the matrix R is stable and *strictly diagonally dominant*, then the stability of the whole matrix KR is ensured for all positive definite and diagonal matrices K. From the definition of matrix K it is obvious that it will be positive definite and diagonal if $\varphi_i + \mathrm{Arg}\left\{F_i(e^{j\omega_i})H_i(e^{j\omega_i})G_i(e^{j\omega_i})\right\} = 0$, $F_i(1) > 0$, $i = 1, ..., N$. ■

Remark 3. The boundedness assumption (A.7) is a standard assumption for convergence analysis of many stochastic approximation algorithms (see, *e.g.*, [25], [26], [23]). However, in practice it might be hard to check the boundedness *a priori*. In order to ensure that this condition is satisfied, the algorithm can be modified by introducing truncation, or projection into some prespecified set S, containing the Nash equilibrium, whenever the estimate $u(k)$ leaves the predefined region B_u, containing the set S. Based on the results from [23], Theorem 1.4.1 it can be shown that the number of truncations can only be finite, which means that for large enough k the algorithm simply reduces to the one without truncations, but with guaranteed boundedness of $\tilde{u}(k)$ by the algorithm construction. ■

Remark 4. Assumptions (A.2)–(A.5) are standard assumptions on the step size in recursive, stochastic and deterministic (sub)gradient algorithms (see *e.g.* [27], [23], [26]). They aim at reducing the effect of measurement noise; however, in order to achieve convergence of the algorithm, they need to converge to zero slow enough so that (A.4) is satisfied. A straightforward way of satisfying Assumptions (A.2)–(A.6) is by simply taking $\varepsilon_i(k) = e_i k^{-m_\varepsilon}$ and $\alpha_i(k) = a_i k^{-m_\alpha}$ where $0.5 < m_\varepsilon < 1$, $0 < m_\alpha < 0.5$, $m_\varepsilon + m_\alpha \leq 1$ and e_i and a_i can account for asynchronicity between the agents (note that for this case, $c_i = e_i a_i$, according to the above notation). ■

IV. Application to Mobile Sensor Networks

It is straightforward to modify the scheme from Fig. 1 and adapt it for problems involving self-organizing networks of autonomous vehicles (mobile sensors), where the vehicles, treated as players in a game, are seeking positions corresponding to a Nash equilibrium. In Fig. 2 a scheme involving force actuated vehicles (double integrators) is shown (see also [17], [16] where similar schemes are proposed for single agent systems). The discrete-time integrator from Fig. 1 is now contained in the vehicle dynamics, and moved in front of the perturbing signal, whose phase now needs to be adjusted to compensate for the integrators phase shift. Also, a discrete-time differentiator is added in order to compensate double integration of the vehicle. Therefore, it can be shown [17] that the overall equivalent discrete-time scheme is actually the scheme from Fig. 1, with the input filters having the transfer function $F_i(z) = 1 + z^{-1}$. Furthermore, the vehicles may have any additional stable dynamics, since they can be incorporated in $F_i(z)$ and $G_i(z)$.

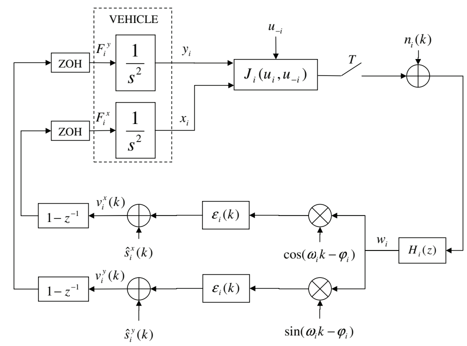

Fig. 2. Nash equilibrium seeking scheme for force actuated vehicles

The formulated framework for Nash equilibrium seeking is general, and allows the cost function to be designed in order to reach some specific goal. For example, the game can be designed in such a way that a Nash equilibrium corresponds to some overall (centralized) goal or a Pareto optimal point. In this scenario the proposed algorithm gives a solution to online decentralized optimization or coordination problems in multi-agent networks based only on the measurements of local costs. However, in general, achieving a social (centralized) goal is not an easy task, since the agents act selfishly, so that some artificial cooperation structure needs to be imposed on agents behavior (by proper design of the agents' cost functions) in order to enforce convergence to an efficient equilibrium (see, *e.g.*, [6], [2], [3], [7]). In what follows, we are going to present a simple example of how to construct the cost functions in order to deal with some typical problems in mobile sensor networks.

Broadening the scenarios analyzed in [14], [17], where an agent is either searching for a source of some signal with unknown distribution, or positioning itself to an optimal sensing point for some estimation task, in our interconnected problem setting, the local cost functions of the agents can be designed to achieve the mentioned local goals, while keeping good connection with selected neighboring agents. This can be important in distributed estimation where the local estimators are communicating with each other in order to improve the overall performance. Therefore the "interconnection" term in the cost functions can correspond to the variance of the agents intercommunication noise, or it can be the reciprocal value of the signal power received from the neighbors. In the latter case, assuming that the signal power

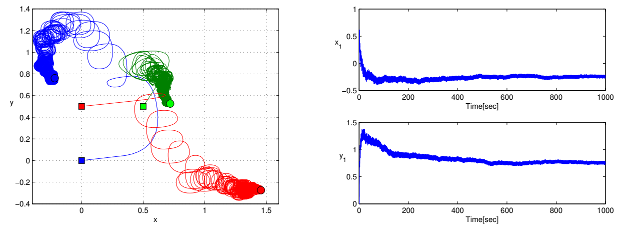

Fig. 3. Trajectories of the vehicles and coordinates of the first vehicle

is inversely proportional to the distance between the agents, *i.e.*, $P(u_i, u_j) \sim 1/||u_i - u_j||^2$, and taking its reciprocal value as the interconnection term which is to be minimized, we can define quadratic cost functions as

$$J_i(u_i, u_{-i}) = u_i^T r_{ii} u_i + u_i^T r_i + k_i + \sum_{j \in \mathcal{N}_i} m_{ij} ||u_i - u_j||^2, \qquad (19)$$

where $r_{ii} > 0$, $||\cdot||$ is the Euclidian norm and the coefficients m_{ij} are selected *a priori*, reflecting the importance of the signal received from the j-th agent. Therefore, the elements of the matrix R in (18) are $R_{ij} = \text{diag}\{-2m_{ij}, -2m_{ij}\}$, $R_{ii}^i = r_{ii} - \frac{1}{2}\sum_{j \in \mathcal{N}_i} R_{ij}$. It is straightforward to check that the matrix $-R$ is strictly diagonally dominant and stable. According to Remark 2, this game will always admit a unique Nash equilibrium and the condition (A.8) is satisfied for any diagonal positive definite K.

V. Example

In this example we illustrate the algorithm proposed in Fig. 2 for a network of three force actuated vehicles, where the cost functions are given by (19) with $r_{11} = r_{22} = r_{33} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $r_1 = [2 \ \ -2]^T$, $r_2 = [-2 \ \ -2]^T$, $r_3 = [-4 \ \ 2]^T$, $k_1 = 3$, $k_2 = 3$, $k_3 = 6$, $m_{12} = m_{21} = m_{23} = m_{32} = 1$ and $m_{13} = m_{31} = 0$. Hence, by solving (1) we obtain that the unique Nash equilibrium is the point $u^* = [-0.125 \ \ 0.75 \ \ 0.75 \ \ 0.5 \ \ 1.375 \ \ -0.25]$. For the other system parameters we assume the following values: the noise covariance matrix (7) is $\Sigma(k) = \text{diag}\{0.1 \ 0.1 \ 0.1\}$, $\varphi_i = -\pi/4$, $T = 0.1$, $H_i(z) = \frac{z-1}{z+0.07}$ (high pass filters), $\varepsilon_i(k) = 1.5k^{-0.65}$, $\alpha_i(k) = 0.4k^{-0.25}$, for $i = 1, 2, 3$, and $\omega_1 = \omega_3 = 0.5\pi$, $\omega_2 = 0.7\pi$. We are allowed to pick the same frequencies for players 1 and 3 since they are not interconnected. Trajectories of the vehicles and time response for the first vehicle are shown in Fig. 3, for the initial conditions $u_1(1) = [0 \ \ 0]^T$, $u_2(1) = [0.5 \ \ 0.5]^T$, $u_3(1) = [0 \ \ 0.5]^T$. The time responses for the other two vehicles are similar. The convergence to the Nash equilibrium is evident.

Acknowledgments

This work was supported by the Knut and Alice Wallenberg Foundation, the Swedish Research Council, the Swedish Strategic Research Foundation, and the EU project FeedNetBack.

References

[1] J. Nash, "Non-cooperative games," *The Annals of Mathematics*, vol. 54, no. 2, pp. 286–295, 1951.

[2] A. W. Starr and Y. C. Ho, "Further properties of nonzero-sum differential games," *Journal of Optimization Theory and Applications*, vol. 3, no. 4, pp. 207–219, 1969.

[3] K. Tumer and D. Wolpert, *Collectives and the design of complex systems*. New York: Springer-Verlag, 2004.

[4] D. Monderer and L. S. Shapley, "Potential games," *Games and economic behavior*, no. 14, p. 124143, 1996.

[5] T. Basar and G. J. Olsder, *Dynamic Noncooperative Game Theory*, 2nd ed. Philadelphia: SIAM, 1999.

[6] P. Dubey, "Inefficiency of Nash equilibria," *Mathematics of Operations Research*, vol. 11, no. 1, pp. 1–8, Feb. 1986.

[7] J. R. Marden and A. Wierman, "Distributed welfare games," Submitted to Operations Research, 2008.

[8] S. Li and T. Basar, "Distributed algorithms for the computation of noncooperative equilibria," *Automatica*, vol. 23, pp. 523–533, 1987.

[9] J. S. Shamma and G. Arslan, "Dynamic fictitious play, dynamic gradient play, and distributed convergence to Nash equilibria," *Automatic Control, IEEE Transactions on*, vol. 50, no. 3, pp. 312–327, March 2005.

[10] J. R. Marden, H. P. Young, G. Arslan, and J. S. Shamma, "Payoff based dynamics for multi-player weakly acyclic games," *SIAM Journal on Control and Optimization*, vol. 48, no. 1, pp. 373–396, 2009.

[11] M. Zhu and S. Martinez, "Distributed coverage games for mobile visual sensors (I) : Reaching the set of Nash equilibria," in *Proc. IEEE Conf. Decision and Control*, 2009.

[12] M. Krstic, P. Frihauf, J. Krieger, and T. Basar, "Nash equilibrium seeking with finitely- and infinitely-many players," Submitted to 8th IFAC Symposium on Nonlinear Control Systems, 2010.

[13] K. B. Ariyur and M. Krstić, *Real Time Optimization by Extremum Seeking*. Hoboken, NJ: Wiley, 2003.

[14] M. S. Stanković and D. M. Stipanović, "Stochastic extremum seeking with applications to mobile sensor networks," in *Proc. American Control Conference*, 2009, pp. 5622–5627.

[15] ——, "Extremum seeking under stochastic noise and applications to mobile sensors," *Automatica*, vol. 46, pp. 1243–1251, 2010.

[16] C. Zhang, A. Siranosian, and M. Krstić, "Extremum seeking for moderately unstable systems and for autonomous vehicle target tracking without position measurements," *Automatica*, vol. 43, pp. 1832–1839, 2007.

[17] M. S. Stanković and D. M. Stipanović, "Discrete time extremum seeking by autonomous vehicles in a stochastic environment," in *Proc. IEEE Conf. Decision and Control*, 2009.

[18] G. M. Hoffmann and C. J. Tomlin, "Mobile sensor network control using mutual information methods and particle filters," *Automatic Control, IEEE Transactions on*, vol. 55, no. 1, pp. 32–47, Jan. 2010.

[19] A. Ganguli, S. Susca, S. Martinez, F. Bullo, and J. Cortes, "On collective motion in sensor networks: sample problems and distributed algorithms," in *Proc. 44th IEEE Conference on Decision and Control*, Dec. 2005, pp. 4239 – 4244.

[20] C. G. Cassandras and W. Li, "Sensor networks and cooperative control," *European Journal of Control*, vol. 11, no. 4-5, pp. 436–463, 2005.

[21] J. B. Rosen, "Existence and uniqueness of equilibrium points for concave n-person games," *Econometrica*, vol. 33, no. 3, pp. 520–534, 1965.

[22] J. Y. Choi, M. Krstić, K. B. Ariyur, and J. S. Lee, "Extremum seeking control for discrete-time systems," *IEEE Trans. Autom. Control*, vol. 47, pp. 318– 323, 2002.

[23] H. F. Chen, *Stochastic approximation and its applications*. Kluwer Academic Publisher, 2003.

[24] A. S. Poznyak and D. O. Chikin, "Asymptotic properties of the stochastic approximation procedure with dependent noise," *Automation and Remote Control*, no. 12, pp. 78–93, 1984.

[25] L. Ljung, "Analysis of recursive stochastic algorithms," *IEEE Trans. Autom. Control*, vol. 22, 1977.

[26] H. J. Kushner and D. S. Clark, *Stochastic approximation methods for contrained and unconstrained systems*. New York: Springer, 1978.

[27] B. T. Polyak and Y. Z. Tsypkin, "Pseudogradient algorithms of adaptation and learning," *Automation and Remote Control*, no. 3, pp. 45–68, 1973.