Distributed Seeking of Nash Equilibria With Applications to Mobile Sensor Networks

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

***Abstract*—We consider the problem of distributed convergence to a Nash equilibrium in a noncooperative game where the players generate their actions based only on online measurements of their individual cost functions, corrupted with additive measurement noise. Exact analytical forms and/or parameters of the cost functions, as well as the current actions of the players may be unknown. Additionally, the players' actions are subject to linear dynamic constraints. We propose an algorithm based on discrete-time stochastic extremum seeking using sinusoidal perturbations and prove its almost sure convergence to a Nash equilibrium. We show how the proposed algorithm can be applied to solving coordination problems in mobile sensor networks, where motion dynamics of the players can be modeled as: 1) single integrators (velocity-actuated vehicles), 2) double integrators (force-actuated vehicles), and 3) unicycles (a kinematic model with nonholonomic constraints). Examples are given in which the cost functions are selected such that the problems of connectivity control, formation control, rendezvous and coverage control are solved in an adaptive and distributed way. The methodology is illustrated through simulations.**

***Index Terms*—Convergence, extremum seeking, learning, mobile sensor networks, multi-agent control, Nash equilibrium, noncooperative games, stochastic optimization.**

I. Introduction

PROBLEMS of distributed, multi-agent optimization, coordination, estimation, and control have been 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. A possible approach to these problems is *game theoretic*, where one formulates a noncooperative game with players/agents selfishly trying to optimize their individual cost functions by using locally available information. Depending on the structure of the game, its Nash equilibria can have different properties and they may or may not correspond to the optimal solution of some global optimization problem [2]–[10].

The focus of this paper is on the problem of *learning in games*, or designing algorithms that converge to a Nash equilibrium. The majority of the existing literature in this area is focused on algorithms based on a detailed model of the underlying game; that is, an algorithm is designed based on a specific form of the players' cost functions and properties. 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 a best/better response strategy. For example, in [11], convergence properties have been analyzed for such a class of infinite, convex games. For 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 [12] and references therein). The recently proposed algorithms in [13] and [14] 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 [4]) with finite action sets. In [15] an interactive learning procedure based on trial and error is proposed. A similar approach to the one proposed in this paper, applied to quadratic games in markets, has appeared independently in [16] and extended to non-quadratic cost functions in [17], but the authors provided only local stability analysis of a simpler one-dimensional scheme under strong conditions. Also, none of the mentioned approaches can deal with the measurement noise while taking into account specific dynamics of the players.

Extremum seeking algorithms have received significant attention recently in adaptive online optimization problems involving dynamical systems. The basic algorithm, based on introducing sinusoidal perturbations, has been treated in, e.g., [18]–[20]. In [21] and [22] a time-varying version of the algorithm has been introduced, whose almost sure convergence has been proved in the presence of measurement noise. It has been demonstrated how this technique can be applied to autonomous vehicles seeking a target in deterministic environments [23], or optimal positioning in stochastic environments [22], [24].

In this paper, we propose a discrete-time algorithm for *distributed* seeking of a pure Nash strategy in infinite games where the players generate their actions based solely on the measurements of their individual cost functions, whose detailed analytical forms may be unknown. Furthermore, similarly as in the extremum seeking problems, we assume that the players may have some (possibly unknown) linear *dynamics* , filtering the players'

Manuscript received August 02, 2010; revised January 25, 2011; accepted October 24, 2011. Date of publication November 03, 2011; date of current version March 28, 2012. This work was supported in part by the Knut and Alice Wallenberg Foundation, in part by the Swedish Research Council, in part by the Swedish Strategic Research Foundation, and in part by the EU project FeedNetBack. Parts of this work have been presented in a preliminary form in [1]. Recommended by Associate Editor J. Cortes.

M. S. Stanković and K. H. Johansson are with the ACCESS Linnaeus Center, School of Electrical Engineering, KTH Royal Institute of Technology, 100-44 Stockholm, Sweden (e-mail: milsta@kth.se; kallej@kth.se).

D. 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).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TAC.2011.2174678

actions before affecting their cost functions. Also, the local measurements of individual 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 algorithm with sinusoidal perturbations, under stochastic noise, analyzed in [22]. We formulate sufficient conditions on the structure of the cost functions and on the parameters of the proposed scheme, under which we prove almost sure convergence to a Nash equilibrium. Some important special cases (potential games and quadratic games), and possible extensions and limitations are treated in details.

The proposed algorithm can be applied to general noncooperative games in which there is a need for *adaptive* learning of Nash equilibria; that is, when the players may not know the exact parameters of the cost functions and the current players' actions and properties, or even if they do not know the exact analytical forms of the costs, but are able to obtain (noisy) values of their individual costs at each time instance. The players may also have some dynamics which imply certain *constraints on their capability of changing actions*. These properties make the algorithm appealing for distributed coordination and optimization problems related to *mobile sensor networks*. These networks are usually deployed in some only partially known or unknown and noisy environments, lacking any communication infrastructure and without some omniscient central or fusion node. The information that the agents have about the environment as well as about the actions/properties of the other agents is limited only to certain local sensing and local low bandwidth communications.

In certain scenarios, the individual cost for each agent can be expressed as a sum of a locally defined goal, depending only on the individual agent's position/action (such as the one treated in [22], [24], or [23]) and a "collective" goal, depending on the positions/actions of the other agents. This is typical in *connectivity control* problems [25]–[28]. By applying our scheme to such problems the agents can adaptively achieve a compromise between these locally defined goals, and a "collective" goal of maintaining connectivity with the neighboring agents, without detailed inter-agent communications and without position measurements. We show that a special case of this scenario leads to an adaptive solution of a *formation control* problem and a *rendezvous* or *consensus* problem (see, e.g., [29], [30], and references therein).

In some cases a cooperative control task to be performed by a mobile network is expressed using a global cost function which may depend on all the environmental parameters and properties/positions of all the agents. In these cases, to deal with the agents' lack of global information it is possible to design individual cost functions which would capture the given information structure constraints. The design should be done such that local optimizations by individual agents (Nash equilibrium) leads to an optimum of a global objective [3], [4], [9], [10]. Even though the cost functions are designed, they usually depend on some parameters/functions (e.g., environmental conditions, individual agents' properties) which are unknown *a priori* and must be *learned* (implicitly or explicitly) in the process of convergence to a Nash equilibrium. A typical example for this scenario is a *coverage control* problem as defined in [31] and formulated as a potential game in [32]–[34]. Using our methodology, this problem can be solved under much more realistic assumptions on the information available to the agents. Specifically, in our setting, the agents do not need any (absolute or relative) position measurements and do not need *a priori* knowledge about the distribution of the events to be detected (environmental density function) and about the specific detection capabilities of the individual agents.

The existing literature in the area of mobile sensor networks which treats similar problems is mostly focused on specific scenarios requiring detailed sensing/comunication models as well as the model of the environment, without taking specific mobile robot dynamics into consideration (e.g., [14], [25]–[28], [30]–[32], [35]–[38] and references therein).

The rest of the paper is organized as follows. In Section II the problem setup and the algorithm description are given. Section III is devoted to the convergence analysis, where we prove that the algorithm converges almost surely (a.s.) to a Nash equilibrium. We discuss possible extensions, limitations and special cases in Section IV. In Section V, we give detailed analysis of some concrete applications of the proposed algorithm to mobile sensor networks (connectivity control, formation control, rendezvous and coverage control), where the dynamics of the mobile robots can be modeled as single integrators, double integrators or unicycles. Simulation results for networks of three agents are shown and discussed.

II. Nash Equilibrium Seeking Algorithm

We consider a scenario in which N agents are noncooperatively minimizing their individual cost functions by updating their local actions, based only on their current local information. We assume that the actions of the players belong to $\mathbb{R}^{m_i}$, $m_i \in \{1, 2, \ldots\}$, $i = 1, \ldots, N$. Hence, we are dealing with a noncooperative static game with infinite action spaces where the optimality is characterized by a (pure) Nash equilibrium; a point from which neither agent has incentive to deviate [2], [5]. We assume that the information that each player has about the underlying game is restricted solely to discrete-time measurements of its individual cost function, which are, in addition, filtered through (possibly unknown) stable, linear and time invariant (LTI) filter with transfer function $G_i(z)$ and corrupted with a measurement noise. The players may not have any direct information about either the underlying structure of the game (exact analytical forms of the cost functions) or the actions/properties of the players. Without loss of generality, let us assume that each agent's action space is two-dimensional, $u_i = (x_i, y_i) \in \mathbb{R}^2$, since we will apply this methodology to vehicles' coordination problems in the plane. The framework can be extended to multidimensional action spaces in a straightforward way, as discussed later in Remark 5. Furthermore, the agents may have some local dynamics, so that their actions are filtered through (possibly unknown) stable LTI filter, having the transfer function matrix $F_i(z)$, before affecting the measured cost function $J_i(u_i, u_{-i}) : \mathbb{R}^{2N} \longrightarrow \mathbb{R}$, where u_i denotes the action of agent i, while u_{-i} denotes the actions of all the other agents. In general, each cost function J_i does not necessarily have to depend on the actions of all the other players. So, let us define time-invariant neighbor sets $\mathcal{N}_i$, $i = 1, \ldots, N$, whose

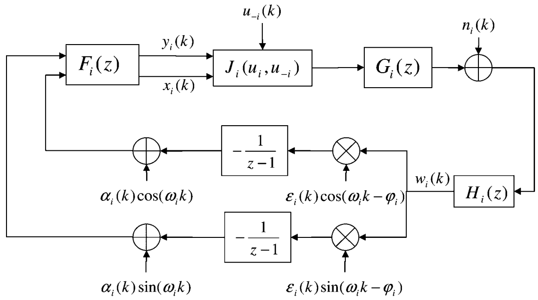

Fig. 1. Nash equilibrium seeking scheme.

elements are indices of the agents whose actions affect the ith agent's cost function.

Motivated by the fact that the formulated information structure relates to extremum seeking problems, we propose an algorithm based on sinusoidal perturbations, depicted in Fig. 1, where each agent implements a local, discrete-time (denoted by k), extremum seeking loop. The estimation of the gradient of the individual cost function is performed by inserting a sinusoidal perturbation, with frequency ω_i and with positive, deterministic, time-varying amplitude $\alpha_i(k)$, which, by passing through the function J_i, is being modulated by its local slope. The estimate of the slope is found by the multiplication/demodulation using the sinusoid with the same frequency and with positive, deterministic, time-varying amplitude $\varepsilon_i(k)$. This slope estimate is then used to move in the opposite direction (by the negative integration block: $-\frac{1}{z-1}$). Since all the information needed to estimate the gradient is located in the amplitude of the modulated sinusoidal perturbation, the measurements are filtered by *washout* filters $H_i(z)$ to eliminate any DC components, and, hence, to improve the overall convergence properties. Also, to improve the convergence properties, low-pass filters can be added in the loop as part of the dynamics $F_i(z)$. Local decoupling between x and y dimensions is obtained using orthogonal perturbations: cosine for x and sine for y (cf., [24] or [23]). 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.

The following equations model the behavior of the proposed Nash equilibrium seeking algorithm:

$$w_i(k) = H_i(z)[G_i(z)[J_i(u_i(k), u_{-i}(k))] + n_i(k)] \quad (1)$$

$$\xi_i(k) = \varepsilon_i(k) C_i(k) w_i(k) \quad (2)$$

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

for $i = 1, \ldots, N$, where $n_i(k)$ is the measurement noise of agent i

$$u_0^i(k) = F_i(z)[\alpha_i(k)\cos(\omega_i k),\ \alpha_i(k)\sin(\omega_i k)]^T \quad (4)$$

$$C_i(k) = [\cos(\omega_i k - \varphi_i),\ \sin(\omega_i k - \varphi_i)]^T. \quad (5)$$

Throughout the paper, the expression $Y(z)[x(k)]$ denotes a time domain vector obtained as the output of LTI system with the transfer function matrix $Y(z)$, with the input vector $x(k)$, and with some arbitrary finite initial condition.

In what follows we are going to introduce assumptions needed for proving the convergence of the algorithm to a Nash equilibrium.

Assumption On the Measurement Noise:

(A.1) The random vectors $n(k)$ (where $n(k) = [n_1(k), \ldots, n_N(k)]^T$) are measurable with respect to a flow of σ-algebras $\mathcal{F}_k \subseteq \mathcal{F}_{k+1}$, mutually independent and zero mean, and they satisfy

$$E\{n(k)n(k)^T\} = \Sigma(k) \leq \Gamma, k = 1, 2, \ldots \quad (6)$$

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

Assumptions On the Parameters of the Algorithm:

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

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

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

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

(A.6) $\sum_{k=1}^{\infty} \varepsilon_i(k)\alpha_i(k)\alpha_j(k) < \infty$ for all $i = 1, \ldots, N$ and $j \in \mathcal{N}_i$.

(A.7) $\varepsilon_i(k)\alpha_i(k) = O(\varepsilon_j(k)\alpha_j(k))$ when $k \to \infty$, for all $i, j = 1, \ldots, N$.

(A.8) $\omega_i = a_i\pi$, $a_i \in (0, 1)$ is a rational number, and $\omega_i \neq \omega_j$ for all $i = 1, \ldots, N$ and $j \in \mathcal{N}_i$.

According to (A.7), $\varepsilon_i(k)\alpha_i(k)$ can be written as

$$\varepsilon_i(k)\alpha_i(k) = \min_j\{\varepsilon_j(k)\alpha_j(k)\}(c_i + o(\varepsilon_i(k)\alpha_i(k))) \quad (7)$$

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

Assumptions (A.2)–(A.6) are standard assumptions on the step size in recursive, stochastic and deterministic (sub)gradient and extremum seeking algorithms (see, e.g., [22], and [39]–[42]). They aim at reducing the effect of measurement noise; however, to achieve convergence of the algorithm, these parameters need to converge to zero slow enough so that (A.4) is satisfied. A straightforward way of satisfying Assumptions (A.2)–(A.7) is by simply taking $\varepsilon_i(k) = e_i k^{-m_\varepsilon}$ and $\alpha_i(k) = a_i k^{-m_\alpha}$ where $0.5 < m_\varepsilon < 1$, $0 < m_\alpha < 0.5$, $m_\varepsilon + m_\alpha = 1$, and $e_i > 0$ and $a_i > 0$ can account for asynchronicity between the agents.

Assumption on the Existence of a Nash Equilibrium:

(A.9) The individual cost functions $J_i(u_i, u_{-i})$ are continuously differentiable and strictly convex in local decision variables u_i, and there exists a Nash equilibrium, i.e., a point u^* for which the following holds:

$$[\nabla_1 J_1(u^*)^T, \ldots, \nabla_N J_N(u^*)^T]^T = 0 \quad (8)$$

where $\nabla_i J_i(\cdot)$, $i = 1, \ldots, N$, denotes the gradient of J_i with respect to local actions u_i.

Due to strict convexity in local decision variables, (8) is a necessary and sufficient condition for achieving a Nash equilibrium [5], [43].

Before stating the last three assumptions which ensure the stability of the algorithm, let us define the tracking error for each agent as

$$\tilde{u}_i(k) = u_i(k) - u_i^* - u_0^i(k) \tag{9}$$

where u_i^* is the ith agent's action in a Nash equilibrium. By stacking together the individual two-dimensional vectors we define $\tilde{u}(k) = [\tilde{u}_1(k)^T, \ldots, \tilde{u}_N(k)^T]^T \in \mathbb{R}^{2N}$ and $u_0(k) = [u_0^1(k)^T, \ldots, u_0^N(k)^T]^T \in \mathbb{R}^{2N}$.

Assumptions Related to the Stability of the Algorithm:

(A.10) The LTI systems with transfer functions $F_i(z)$, $G_i(z)$ and $H_i(z)$, $i = 1, \ldots, N$, are asymptotically stable.

(A.11) $\tilde{u}(k) \in B$ a.s. for all $k = 1, 2, \ldots$, where B is an open ball in $\mathbb{R}^{2N}$ containing the origin and with bounded radius. $J_i(u)$, $i = 1, \ldots, N$, are analytic in an open ball 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 $k = 1, 2, \ldots$ [in accordance with (9)].

(A.12) There exists a continuously differentiable Lyapunov function $V(\tilde{u}) : \mathbb{R}^{2N} \to \mathbb{R}$ such that

$$-g(\tilde{u})^T K^T \nabla_{\tilde{u}} V(\tilde{u}) < 0, \text{ for all } \tilde{u} \neq 0,\ \tilde{u} \in B \tag{10}$$

where $g(\tilde{u}) = [\nabla_1 J_1(u^* + \tilde{u})^T, \ldots, \nabla_N J_N(u^* + \tilde{u})^T]^T$, $K = \text{diag}\{c_1 K_1, \ldots, c_N K_N\}$, $K_i = F_i(1)\begin{bmatrix} \text{Re}\{\theta_i\} & \text{Im}\{\theta_i\} \\ -\text{Im}\{\theta_i\} & \text{Re}\{\theta_i\} \end{bmatrix}$, $\theta_i = e^{j\varphi_i} F_i(e^{j\omega_i}) \times G_i(e^{j\omega_i}) H_i(e^{j\omega_i})$, $\nabla_{\tilde{u}} V(\tilde{u})$ denotes the gradient of $V(\tilde{u})$, and $\text{diag}\{\cdot\}$ denotes the corresponding block diagonal matrix.

The boundedness Assumption (A.11) might be hard to check *a priori*, but it can be guaranteed by introducing truncations or projections of the players' actions to a prespecified set, once they leave a predefined region B_u, containing u^*, as discussed later in Remark 1. Assumption (A.7) ensures that the matrix K defined in (A.12) is constant. This assumption can be removed if we modify (A.12) as commented in Remark 4.

Assumption (A.12), besides stability of our algorithm, also ensures uniqueness of the Nash equilibrium u^* (see also [43] where stability and uniqueness are ensured with a strong condition called strict diagonal convexity). It will be evident in the sequel (see Remark 2 in the next section) that this assumption can be relaxed by allowing existence of multiple (possibly infinite number of) Nash equilibria as long as an appropriate Lyapunov function exists. In the case of quadratic cost functions, Assumption (A.12) can be directly related to a matrix stability condition, involving Jacobian of the vector function $g(\tilde{u})$. If the underlying game is a *potential game* [4], a natural choice for the Lyapunov function is the potential function. These two important special cases have been analyzed in detail in Section IV.

III. Convergence Analysis

Before stating the main convergence theorem, let us introduce a few lemmas that will be used in the proof. In Lemma 1 conditions for the convergence of the standard Robbins–Monro stochastic approximation algorithm with *state-dependent noise* are formulated (see, e.g., [40, Theorem 2.2.3]). In Lemma 2 [44, Lemma 2], conditions are introduced for a.s. convergence of a stochastic process defined as a sum of a weighted correlated noise sequence. These conditions are formulated in terms of statistical properties of the given sequence, which in fact specify a class of noise with sufficiently slowly increasing second moment and sufficiently fast decreasing correlations. Lemma 3 and Lemma 4 are useful in the analysis of filtered, uniformly bounded sequences whose difference tends to zero with rate defined in conditions (A.2)–(A.5). In Lemma 5 we prove convergence of sums of sinusoidal signals modulated with fast enough vanishing signals, which will frequently appear in the proof of the main theorem. Lemma 6 is a simple modulation lemma useful in dealing with general filtered modulated sinusoidal signals. The proofs of Lemmas 3, 4, and 5 are given in Appendix.

Lemma 1 [40, Theorem 2.2.3][1]: Consider the following recursive (Robbins–Monro) algorithm:

$$u(k+1) = u(k) + \rho(k)(f(u(k)) + \xi(k)) \tag{11}$$

where $u(k) \in \mathbb{R}^N$, $\rho(k)$ is a predefined deterministic sequence of real numbers, $f(u)$ is a vector function $\mathbb{R}^N \to \mathbb{R}^N$, and $\xi(k) \in \mathbb{R}^N$ is the "observation error" term which can depend on $u(k)$. Assume that the following assumptions are satisfied:

(B.1) $\rho(k) > 0$, $k = 1, 2, \ldots$, $\lim_{k \to \infty} \rho(k) = 0$ and $\sum_{k=1}^{\infty} \rho(k) = \infty$.

(B.2) There exists a continuously differentiable Lyapunov function $V(u) : \mathbb{R}^N \to \mathbb{R}$ such that $f(u)^T \nabla_u V(u) < 0$ for all $u \notin U^*$, where U^* is the set of zeros of $f(u)$ (i.e., $f(u) = 0$ for every $u \in U^*$), and $V(U^*) \triangleq \{V(u) : u \in U^*\}$ is nowhere dense.

(B.3) $\|u(k)\|$ is uniformly bounded a.s. for all $k = 1, 2, \ldots$.

(B.4) $\|\sum_{k=1}^{\infty} \rho(k)\xi(k)\| < \infty$ a.s.

(B.5) $f(u)$ is continuous.

Then, $\text{dist}(u(k), U^*) \longrightarrow 0$ (a.s.) as $k \to \infty$, where $\text{dist}(u, U^*) \triangleq \inf_v \{\|u - v\| : v \in U^*\}$ and $\|\cdot\|$ denotes Euclidean norm.

Lemma 2 [44, Lemma 2]: Let $\zeta(k)$ be a sequence of random variables that is measurable with respect to a flow of σ-algebras $\mathcal{F}_k$ and such that $E\{\zeta(k)\} = 0$ and $E\{\zeta(k)^2\} = \sigma_\zeta(k)^2 < \infty$ and $\varepsilon(k)$ be a deterministic positive sequence of real numbers. If the following conditions are satisfied:

(C.1) $r(k) \triangleq \sum_{j=k+1}^{\infty} \varepsilon(j)\Psi_{j,k} \to 0,\ k \to \infty$;

(C.2) $\sum_{k=1}^{\infty} \varepsilon(k)^2 \sigma_\zeta(k)^2 < \infty$;

(C.3) $\sum_{k=1}^{\infty} \varepsilon(k)\sigma_\zeta(k) r(k) < \infty$;

where $\Psi_{j,k} = \|E\{\zeta(j)|\mathcal{F}_k\}\|_2$ with $j > k$, $\|\cdot\|_2 = (E\{|\cdot|^2\})^{1/2}$, then $|\sum_{k=1}^{\infty} \varepsilon(k)\zeta(k)| < \infty$ a.s.

Lemma 3: Assume that $\varepsilon(k)$ is a sequence of real numbers which satisfies (A.2)–(A.5), $H(z)$ is the transfer function matrix of a stable LTI system and $x(k)$ is a uniformly bounded vector sequence. Then the following equation holds:

$$H(z)[\varepsilon(k)x(k)] = \varepsilon(k)H(z)[x(k)] + \delta(k) \tag{12}$$

[1]This lemma is a special case of the cited theorem since Assumptions (B.2), (B.4), and (B.5) are stronger than in [40], but they are sufficient for our convergence analysis. That (B.4) implies the original Assumption (A2.2.3) of the cited theorem follows directly from, e.g., [40, Theorem 2.4.1 ii)].

where $\delta(k)$ is a summable vector sequence, i.e., $\|\sum_{k=1}^{\infty}\delta(k)\|<\infty$.

Lemma 4: Assume that $\varepsilon(k)$ is a sequence of real numbers which satisfies (A.2)–(A.5), $H(z)$ is the transfer function matrix of a stable LTI system and $x(k)$ is a bounded vector sequence which satisfies

$$x(k+1)-x(k)=\varepsilon(k)y(k),\ k=1,2,\dots \tag{13}$$

where $y(k)$ is a uniformly bounded vector sequence. Then the following equation holds:

$$H(z)[x(k)]=H(1)x(k)+\delta(k) \tag{14}$$

where $\|\sum_{k=1}^{\infty}\varepsilon(k)\delta(k)\|<\infty$.

Lemma 5: Assume that $\varepsilon_i(k)$, $i=1,\dots,N$, are sequences of real numbers which satisfy (A.2)–(A.7), and that bounded scalar sequences $x_i(k)$, $i=1,\dots,N$, satisfy

$$x_i(k+1)-x_i(k)=\varepsilon_i(k)y_i(k),\ k=1,2,\dots \tag{15}$$

where $y_i(k)$ are uniformly bounded sequences. Then $|\sum_{k=1}^{\infty}x_1(k)^{n_1}x_2(k)^{n_2}\cdots x_N(k)^{n_N}\varepsilon_1(k)\cos(\omega k+\varphi)|<\infty$, for every fixed $n_1,n_2,\dots,n_N\in\{0,1,2,\dots\}$, where $\omega=a\pi$, a is a rational number, and φ is a constant.

Lemma 6 [19, Lemma 2]: If the transfer functions $G(z)$ and $H(z)$ are stable, the following statement is true for any real ϕ and ω and any uniformly bounded scalar sequence $x(k)$:

$$\begin{aligned}G(z)&[(H(z)[\cos(\omega k-\phi)])x(k)]\\&=\mathrm{Re}\{e^{j(\omega k-\phi)}H(e^{j\omega})G(e^{j\omega}z)[x(k)]\}+\epsilon^{-k}\end{aligned} \tag{16}$$

where ϵ^{-k} denotes exponentially decaying terms.

Now we are in a position to prove the following main convergence theorem:

Theorem 1: Consider the Nash equilibrium seeking algorithm defined in (1)–(5) and shown in Fig. 1. Let Assumptions (A.1)–(A.12) 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^* a.s.

Proof: By substituting (3) into (9) we obtain

$$\tilde{u}_i(k)=-u_i^*-\frac{1}{z-1}F_i(z)[\xi_i(k)] \tag{17}$$

which can be written as a difference equation

$$\tilde{u}_i(k+1)=\tilde{u}_i(k)-F_i(z)[\xi_i(k)]. \tag{18}$$

After plugging (2) and (1) into (18), we obtain for each agent

$$\begin{aligned}\tilde{u}_i(k+1)-\tilde{u}_i(k)=&-F_i(z)[\varepsilon_i(k)C_i(k)\\&\times H_i(z)[G_i(z)[J_i(u_i(k),u_{-i}(k))]+n_i(k)]].\end{aligned} \tag{19}$$

Since we have assumed that the functions $J_i(u_i,u_{-i})$ are analytic in the region B_u containing u^* [Assumption (A.11)] one can write their Taylor series expansion around the Nash equilibrium point u^*:

$$\begin{aligned}J_i(u_i,u_{-i})=&J_i(u^*)+\nabla_iJ_i(u^*)^T(u_i-u_i^*)\\&+\nabla_{-i}J_i(u^*)^T(u_{-i}-u_{-i}^*)\\&+\frac{1}{2!}((u_i-u_i^*)^T\nabla_{ii}^2J_i(u^*)(u_i-u_i^*)\\&+2(u_i-u_i^*)^T\nabla_{i,-i}^2J_i(u^*)(u_{-i}-u_{-i}^*)\\&+(u_{-i}-u_{-i}^*)^T\nabla_{-i,-i}^2J_i(u^*)(u_{-i}-u_{-i}^*))\\&+\cdots\end{aligned} \tag{20}$$

where $\nabla_iJ_i(u^*)$ denotes the gradient of J_i at u^* with respect to the ith player actions, $\nabla_{-i}J_i(u^*)$ denotes the gradient at u^* with respect to the actions of all the other players and $\nabla_{ii}^2J_i(u^*)$, $\nabla_{i,-i}^2J_i(u^*)$ and $\nabla_{-i,-i}^2J_i(u^*)$ denote their corresponding Jacobians at point u^*. By substituting (9) into (20), $J_i(u_i(k),u_{-i}(k))$ can be written as a sum of three terms:

$$J_i(u_i(k),u_{-i}(k))=L_i(k)+D_i(k)+d_i(k) \tag{21}$$

which will be defined one by one. The first term $L_i(k)$ contains the terms that are linear with respect to the perturbation signal $u_0^i(k)$; therefore, it is essential for achieving an 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

$$\begin{aligned}L_i(k)=&u_0^i(k)^T(\nabla_iJ_i(u^*)+\nabla_{ii}^2J_i(u^*)\tilde{u}_i\\&+\nabla_{i,-i}^2J_i(u^*)\tilde{u}_{-i}+\cdots)\\=&u_0^i(k)^T\nabla_iJ_i(u^*+\tilde{u})\end{aligned} \tag{22}$$

where the last equality follows after calculating the gradient of (20) with respect to u_i. Term $d_i(k)$ in (21) contains the deterministic input terms (not depending on any $\tilde{u}_i$, $i=1,\dots,N$):

$$\begin{aligned}d_i(k)=&J_i(u^*)+\nabla_{-i}J_i(u^*)^Tu_0^{-i}+\frac{1}{2}(u_0^i)^T\nabla_{ii}^2J_i(u^*)u_0^i\\&+(u_0^i)^T\nabla_{i,-i}^2J_i(u^*)u_0^{-i}\\&+\frac{1}{2}(u_0^{-i})^T\nabla_{-i,-i}^2J_i(u^*)u_0^{-i}+\cdots.\end{aligned} \tag{23}$$

Term $D_i(k)$ in (21) contains the remaining terms:

$$\begin{aligned}D_i(k)=&\nabla_{-i}J_i(u^*)^T\tilde{u}_{-i}+\frac{1}{2}\tilde{u}_i^T\nabla_{ii}^2J_i(u^*)\tilde{u}_i\\&+\tilde{u}_i^T\nabla_{i,-i}^2J_i(u^*)(u_0^{-i}+\tilde{u}_{-i})\\&+((u_0^{-i})^T+\frac{1}{2}\tilde{u}_{-i}^T)\nabla_{-i,-i}^2J_i(u^*)\tilde{u}_{-i}+\cdots.\end{aligned} \tag{24}$$

By applying Lemma 6 to $u_0^i(k)$ (given in (4)) we obtain

$$u_0^i(k)=\begin{bmatrix}\alpha_1^i(k)\cos(\omega_ik)-\alpha_2^i(k)\sin(\omega_ik)\\\alpha_1^i(k)\sin(\omega_ik)+\alpha_2^i(k)\cos(\omega_ik)\end{bmatrix}+\epsilon_{i0}^{-k} \tag{25}$$

where $\alpha_1^i(k)=\mathrm{Re}\{F_i(e^{j\omega_i}z)[\alpha_i(k)]\}$, $\alpha_2^i(k)=\mathrm{Im}\{F_i(e^{j\omega_i}z)[\alpha_i(k)]\}$ and ϵ_{i0}^{-k} denotes exponentially decaying terms of appropriate dimension.

Now we focus on the essential term for achieving the contraction of the tracking error, which is obtained at the right-hand side of (19) after plugging (21), and is given by $-F_i(z)[\varepsilon_i(k)C_i(k)M_i(z)[L_i(k)]]$, where $M_i(z) = H_i(z)G_i(z)$. By plugging (25) into (22) and then into this term, we can again apply Lemma 6 to all the obtained terms because they all contain a modulated sinusoidal signal being filtered by $M_i(z)$. Since the obtained signal is then multiplied (demodulated) by $C_i(k)$, after some algebra, one obtains the following equation:

$$C_i(k)M_i(z)[L_i(k)] = Q_i(z)\left[A_i(k)\nabla_i J_i(u^* + \tilde{u})\right] + S_i(k)P_i(z)\left[A_i(k)\nabla_i J_i(u^* + \tilde{u})\right] + \epsilon_i^{-k} \quad (26)$$

where $Q_i(z) = \begin{bmatrix} Q_i^1(z) & Q_i^2(z) \\ Q_i^2(z) & -Q_i^1(z) \end{bmatrix}$, $Q_i^1(z) = \mathrm{Re}\{e^{j\varphi_i} M_i(e^{j\omega_i}z)\}$, $Q_i^2(z) = -\mathrm{Im}\{e^{j\varphi_i} M_i(e^{j\omega_i}z)\}$, $A_i(k) = \begin{bmatrix} \alpha_1^i(k) & \alpha_2^i(k) \\ \alpha_2^i(k) & -\alpha_1^i(k) \end{bmatrix}$, $P_i(z) = \begin{bmatrix} P_i^1(z) & P_i^2(z) \\ P_i^2(z) & -P_i^1(z) \end{bmatrix}$, $P_i^1(z) = -\mathrm{Re}\{M_i(e^{j\omega_i}z)\}$, $P_i^2(z) = \mathrm{Im}\{M_i(e^{j\omega_i}z)\}$, $S_i(k) = \begin{bmatrix} \cos(2\omega_i k - \varphi_i) & \sin(2\omega_i k - \varphi_i) \\ \sin(2\omega_i k - \varphi_i) & -\cos(2\omega_i k - \varphi_i) \end{bmatrix}$, and ϵ_i^{-k} denotes exponentially decaying terms of appropriate dimension.

Now we take the first term on the right-hand side of (26) and apply Lemma 3, to obtain

$$-F_i(z)\left[\varepsilon_i(k)Q_i(z)[A_i(k)\nabla_i J_i(u^* + \tilde{u})]\right] = -\varepsilon_i(k)B_i(z)[A_i(k)\nabla_i J_i(u^* + \tilde{u})] + \delta_i^1(k) \quad (27)$$

where $B_i(z) = F_i(z)Q_i(z)$ and $\|\sum_{k=1}^{\infty}\delta_i^1(k)\| < \infty$ (a.s.), $i = 1, \ldots, N$.

By further applying Lemma 3 and then Lemma 4 to the first term on the right-hand side of (27) one obtains

$$-F_i(z)\left[\varepsilon_i(k)Q_i(z)[A_i(k)\nabla_i J_i(u^* + \tilde{u})]\right] = -\varepsilon_i(k)\alpha_i(k)K_i\nabla_i J_i(u^* + \tilde{u}) + \delta_i(k) \quad (28)$$

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 $K_i = F_i(1)\begin{bmatrix} \mathrm{Re}\{\theta_i\} & \mathrm{Im}\{\theta_i\} \\ -\mathrm{Im}\{\theta_i\} & \mathrm{Re}\{\theta_i\} \end{bmatrix}$, where $\theta_i = e^{j\varphi_i}F_i(e^{j\omega_i})G_i(e^{j\omega_i})H_i(e^{j\omega_i})$, as given in (A.12). Finally, coming back to the individual tracking equation (19), by using (28), (26) and (21), we obtain the tracking equation for the whole system:

$$\tilde{u}(k+1) = \tilde{u}(k) - \varepsilon(k)\alpha(k)K_0 g(\tilde{u}(k)) + \delta(k) + F(z)[\varepsilon(k)\pi(k)] \quad (29)$$

where $\varepsilon(k) = \mathrm{diag}\{\varepsilon_1(k), \ldots, \varepsilon_N(k)\} \otimes I_2$, $\alpha(k) = \mathrm{diag}\{\alpha_1(k), \ldots, \alpha_N(k)\} \otimes I_2$, $K_0 = \mathrm{diag}\{K_1, \ldots, K_N\}$, $g(\tilde{u}(k)) = [\nabla_1 J_1(u^* + \tilde{u}(k))^T, \ldots, \nabla_N J_N(u^* + \tilde{u}(k))^T]^T$, $\delta(k) = [\delta_1(k)^T, \ldots, \delta_N(k)^T]^T$, $F(z) = \mathrm{diag}\{F_1(z), \ldots, F_N(z)\}$

$$\pi(k) = -S(k)P(z)[A(k)g(\tilde{u}(k))] - C(k)\left(M(z)[d(k) + D(k)] + H(z)[n(k)]\right) \quad (30)$$

$S(k) = \mathrm{diag}\{S_1(k), \ldots, S_N(k)\}$, $P(z) = \mathrm{diag}\{P_1(z), \ldots, P_N(z)\}$, $A(k) = \mathrm{diag}\{A_1(k), \ldots, A_N(k)\}$, $C(k) = \mathrm{diag}\{C_1(k), \ldots, C_N(k)\}$, $M(z) = \mathrm{diag}\{M_1(z), \ldots, M_N(z)\}$, $d(k) = [d_1(k), \ldots, d_N(k)]^T$, $D(k) = [D_1(k), \ldots, D_N(k)]^T$, $H(z) = \mathrm{diag}\{H_1(z), \ldots, H_N(z)\}$, I_2 is 2×2 identity matrix, $\otimes$ denotes the Kronecker product, and we have incorporated exponentially decaying terms ϵ_i^{-k} in $\delta_i(k)$.

Because of Assumption (A.7), for each $i = 1, \ldots, N$ we can write $\varepsilon_i(k)\alpha_i(k) = \rho(k)(c_i + o(\varepsilon_i(k)\alpha_i(k)))$, where $\rho(k) = \min_j \varepsilon_j(k)\alpha_j(k)$ and $c_i > 0$, so that, after plugging it in the second term on the right-hand side of (29), we obtain

$$\tilde{u}(k+1) = \tilde{u}(k) - \rho(k)Kg(\tilde{u}(k)) + \delta(k) + F(z)[\varepsilon(k)\pi(k)] \quad (31)$$

where K is as given in (A.12) and where we have incorporated the summable terms $\rho(k)(o(\varepsilon_i(k)\alpha_i(k)))$ [according to (A.5)] in $\delta(k)$.

Now it is obvious that the recursive equation (31) is actually the Robbins–Monro algorithm (11), where $u(k)$ is replaced by $\tilde{u}(k)$, $f(u)$ replaced by $-Kg(\tilde{u})$ [which has a unique zero for $\tilde{u} = 0$, according to (A.9) and (A.12)], and having the error term $\rho(k)\xi(k)$ equal to $\delta(k) + F(z)[\varepsilon(k)\pi(k)]$ which contains "structural" perturbation terms (depending on $\tilde{u}(k)$), deterministic input terms, and a stochastic input term (depending on $n(k)$). Therefore, we can apply Lemma 1 since by Assumption (A.12), there exists a Lyapunov function $V(\tilde{u})$ that satisfies condition (B.2) of the lemma [note that set U^* is a singleton, i.e., $U^* = \{0\}$ because of (A.12)]. Therefore, $\tilde{u} \to 0$ a.s. if the "observation error" satisfies (B.4), i.e., if

$$\sum_{k=1}^{\infty}\{\delta(k) + F(z)[\varepsilon(k)\pi(k)]\} \text{ converges (a.s.).} \quad (32)$$

Since the filter $F(z)$ is linear and asymptotically stable, we can switch the summation and filtering in the second term in (32); hence, it is sufficient to show that $\sum_{k=1}^{\infty}\delta(k)$ and $\sum_{k=1}^{\infty}\varepsilon(k)\pi(k)$ converge (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 $\varepsilon(k)C(k)M(z)[d(k) + D(k)]$ [obtained using (30)] can only have one of the following two forms:

1) $\chi(k)\varepsilon_i(k)\alpha_i(k)\alpha_j(k)$ where $\chi(k)$ is a bounded scalar sequence possibly *not containing* a sinusoidal signal. These terms can only originate from the higher order terms in $\varepsilon(k)C(k)M(z)[D(k) + d(k)]$ for which the perfect matching of the multiples, sums or differences of frequencies of multiplying sinusoids happen, e.g., if $\omega_i = 2\omega_j$ for some $i \in \{1, \ldots, N\}$ and $j \in \mathcal{N}_i$. These terms are summable due to Assumption (A.6).

2) A vanishing 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 $n_i \in \{0,1,2,\ldots\}$ for all $i \in \{1,\ldots,N\}$. Also, each $\chi_i(k)$ satisfies (15), with $\chi_i(k)$ in place of $x_i(k)$, so that one can apply Lemma 4 and Lemma 5 and conclude that they are summable. It is important to observe that some of the sinusoidal signals multiplying the above terms, will originate from the terms in $D_i(k)$ or $d_i(k)$ containing the jth perturbation u_0^j, $j \in \mathcal{N}_i$, multiplied with a different frequency sinusoid contained in $C_i(k)$ [Assumption (A.8)]. By converting these products of sinusoids into summations, these terms will end up having the above-mentioned, summable forms.

Therefore, $\sum_{k=1}^{\infty}\varepsilon(k)S(k)P(z)[A(k)g(\tilde{u}(k))]$ and $\sum_{k=1}^{\infty}\varepsilon(k)C(k)M(z)[d(k)+D(k)]$ converge a.s.

Finally, we are left to show that the stochastic input terms $\varepsilon_i(k)C_i(k)H_i(z)[n_i(k)]$, which are independent sequences [by (A.1)] filtered through stable filters $H_i(z)$ and multiplied with $C_i(k)\varepsilon_i(k)$, are summable a.s. We will treat these terms using Lemma 2. Namely, we need to show that they satisfy conditions (C.1)–(C.3), for $\zeta_i(k) = s_i(k)H_i(z)[n_i(k)]$, $i = 1,\ldots,N$, where $s_i(k)$ denotes either $\cos(\omega_i k - \varphi_i)$ or $\sin(\omega_i k - \varphi_i)$. Following the approach presented in [22], for condition (C.1) we have, for $j > k$

$$\begin{aligned} E\{\zeta_i(j)|\mathcal{F}_k\} &= E\left\{\sum_{p=1}^{j} l_i(j-p)n_i(p)|\mathcal{F}_k\right\} s_i(k) \\ &= s_i(k)\sum_{p=1}^{k} l_i(j-p)n_i(p) \text{ (a.s.)} \end{aligned} \tag{33}$$

where we used the fact that $E\{n_i(s)|\mathcal{F}_k\} = 0$ a.s. for $s > k$, $E\{n_i(s)|\mathcal{F}_k\} = n_i(s)$ a.s. for $s \le k$ [Assumption (A.1)], and $\{l_i(j)\}$ is the impulse response sequence of $H_i(z)$. Furthermore, from (33), we have

$$\begin{aligned} r_i(k) &= \sum_{j=k+1}^{\infty}\varepsilon_i(j)|s_i(k)|\left(E\left\{\left(\sum_{p=1}^{k} l_i(j-p)n_i(p)\right)^2\right\}\right)^{1/2} \\ &\le K'\sum_{j=k+1}^{\infty}\varepsilon_i(j)\left(\sum_{p=1}^{k} l_i(j-p)^2\right)^{1/2} \\ &\le K''\varepsilon_i(k+1)\sum_{j=k+1}^{\infty}\left(\sum_{p=1}^{k} l_i(j-p)^2\right)^{1/2} \end{aligned} \tag{34}$$

for some positive constants K' and K'', where we used (A.1), the fact that $E\{n_i(p)n_i(j)\} = 0$ for $p \neq j$ and $E\{n_i(p)n_i(j)\} = \sigma_i^2(j)$ for $p = j$ [$\sigma_i^2(j)$ is the ith diagonal element of $\Sigma(j)$ in (6)], together with the fact that $\varepsilon_i(k)$ is a decreasing sequence. The last term in (34) goes to zero when $k \to \infty$ because $\varepsilon_i(k) \to 0$ and $\sum_{j=k+1}^{\infty}(\sum_{p=1}^{k} l_i(j-p)^2)^{1/2} \le M' < \infty$, since the washout filter $H_i(z)$ is exponentially stable. Therefore, the condition (C.1) is satisfied. Condition (C.2) follows directly from Assumptions (A.1) and (A.5). To prove condition (C.3) we have

$$\sum_{k=1}^{\infty}\varepsilon_i(k)E\left\{(s_i(k)H_i(z)[n_i(k)])^2\right\}r_i(k) \le M''\sum_{k=1}^{\infty}\varepsilon_i(k)^2 < \infty \tag{35}$$

for some $M'' > 0$, where we used (34) and Assumptions (A.1), (A.2), and (A.5).

Therefore, we have shown that the sum in (32) converges a.s., which proves that $\tilde{u}(k)$ converges to zero a.s. This proves the theorem, having in mind the tracking error definition (9) and Assumption (A.3). ■

Remark 1 (on the Boundedness Assumption): In the proof of the theorem we have frequently used the boundedness assumption (A.11). This is a standard assumption for convergence analysis of stochastic approximation algorithms (see, e.g., [40]–[42]). However, in practice it might be hard to check the boundedness *a priori*. To ensure that this condition is satisfied, the algorithm can be modified by introducing truncation, or projection into some prespecified ball 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 [40], Theorem 1.4.1, for the convergence of this truncated algorithm it is sufficient that $V(u_S - u^*) < \inf_{u\in\partial B_u} V(u - u^*)$, for all the points u_S in the projection set S, where $V(\tilde{u})$ is the Lyapunov function defined in (A.12) and ∂B_u denotes the boundary of B_u. This means that the value of the Lyapunov function $V(u - u^*)$ evaluated at any point in the projection set S should be less than the smallest value evaluated on the boundary of B_u. Obviously, this is not a restrictive condition due to the nature of the Lyapunov function. Under this assumption, it has been shown in [40] 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 now with guaranteed boundedness of $\tilde{u}(k)$ by the algorithm construction. Since the Nash equilibrium is not known *a priori* the set S must be chosen conservatively so that it is guaranteed that the equilibrium is in its interior.

Remark 2 (Non-Unique Nash Equilibrium): For clarity of presentation we assumed that the Nash equilibrium is unique, i.e., we assumed that (10) holds for all $\tilde{u} \neq 0$. However, Lemma 1 allows that the set of zeros of function $f(u)$ is not just a singleton, so that (A.12) can be easily relaxed such that (10) holds for all $\tilde{u} \notin U^*$, where $\tilde{u}$ is now defined as in (9) but with respect to any Nash equilibrium (which we denote here by u_0^*), and U^* is the set of all points $\tilde{u}^*$ for which $u_0^* + \tilde{u}^*$ satisfies (8). Under this relaxed assumption, Lemma 1 can still be directly applied to equation (31) (assuming that the technical condition that $V(U^*)$ is nowhere dense is satisfied). Hence, the algorithm will converge to a set of Nash equilibria, provided that the appropriate Lyapunov function exists. This is an important generalization since it allows many practical applications (see Section V-D2 where an application to robotic networks is presented in which the set of Nash equilibria forms a linear subspace).

Having in mind the generality of the cost functions J_i, it might be the case that there exist multiple separated (locally) stable Nash equilibria (or separated sets of equilibria) [5]. This means that for each one of them (or for each separated set) there exists a different Lyapunov function but applied to a different domain [ball B_u in (A.11)]. Therefore, the algorithm will converge to an equilibrium which belongs to the ball B_u in which the algorithm is initialized. Note that the set B_u is in this case analogous to the region of attraction of an equilibrium in standard Lyapunov stability theory.

Remark 3: From the analysis of the deterministic input term $d(k)$, given in (23), it can be concluded that convergence can be achieved even without the washout filters $H_i(z)$ since the DC value in $d_i(k)$ will be multiplied by $\varepsilon_i(k)C_i(k)$ [see (30) and (29)] resulting in the summable term, by Lemma 5. However, it is beneficial to include these filters, since this DC gain is unknown and can be very large so that in the initial iterations it can cause large fluctuations. Also, for this deterministic input term to converge faster, it is beneficial that $\alpha_i(k), i = 1, \ldots, N$, decay faster [condition (A.6)], but slow enough so that (A.4) is satisfied.

Remark 4: Assumption (A.7) ensures that the variables c_i, $i = 1, \ldots, N$, in the matrix K in the main recursion (31), are constant. If we remove this assumption, then, in general, according to (7), these variables can diverge to infinity and we will have a time-dependent matrix $K(k)$. Therefore, if we remove (A.7), we need to make Assumption (A.12) stronger, i.e., instead of (10) we may assume that the following holds:

$$\sup_k \{-g(\tilde{u})^T K(k)^T \nabla_u V(\tilde{u})\} < 0 \tag{36}$$

for all $\tilde{u} \neq 0$, $\tilde{u} \in B$. This condition follows directly from [40, Theorem 2.8.1], which is actually an extension of Lemma 1 for the case of time-varying function $f(u)$ in (11).

Remark 5 (Multi-Dimensional Action Spaces): So far we have focused on the case of two-dimensional agents' action spaces, since we are going to consider coordination problems in the plane. The proposed methodology and the proof of convergence can be easily extended to the multidimensional action spaces for each agent, i.e., $u_i \in \mathbb{R}^{m_i}$, with m_i being any natural number. Indeed, in this case we can allow each agent to implement a sinusoid of different frequency for each component of their local action spaces. It is easy to conclude that, for this case all the results will still hold, with the only difference that the matrix K will now always be diagonal, and positive definite for

$$-\frac{\pi}{2} < \varphi_i + \operatorname{Arg}\left\{F_i(e^{j\omega_i})H_i(e^{j\omega_i})G_i(e^{j\omega_i})\right\} < \frac{\pi}{2} \tag{37}$$

assuming that $F_i(1) > 0$ and $F_i(e^{j\omega_i})H_i(e^{j\omega_i})G_i(e^{j\omega_i}) \neq 0$, $i = 1, \ldots, N$. If the agents use the same frequency for at most two components of the action spaces (with orthogonal phase shifts), as in the 2-D case shown in Fig. 1, then K will be diagonal only if

$$\varphi_i + \operatorname{Arg}\left\{F_i(e^{j\omega_i})H_i(e^{j\omega_i})G_i(e^{j\omega_i})\right\} = 0. \tag{38}$$

Otherwise, it will be block diagonal with 2×2 antisymmetric diagonal blocks, as defined in (A.12), and positive definite under (37).

IV. Discussion

1) Potential Games: If the underlying game is a potential game [4], the vector in (8) will be equal to the gradient of the potential function. Denoting the potential function with $U(u)$ and assuming that it has a unique minimum u^* in B_u (which is also a Nash equilibrium), we can choose the Lyapunov function $V(\tilde{u}) = U(\tilde{u} + u^*)$ (shifted potential function such that $\tilde{u} = 0$ corresponds to the minimum), so that the condition (10) will always be satisfied if K is positive definite (since $g(\tilde{u}) = \nabla_{\tilde{u}} V(\tilde{u})$). Therefore, in this case, Assumption (A.12) can be replaced with the condition (37), which guarantees 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. The case when there exist multiple equilibria (e.g., if $U(u)$ is positive semidefinite) can be treated similarly, as commented in Remark 2.

2) Quadratic Cost Functions: In the case of quadratic cost functions there is a direct interpretation of the stability condition in terms of a Jacobian matrix stability. Assume that the cost functions are given by

$$\begin{aligned} J_i(u_i, u_{-i}) = {} & u_i^T R_{ii}^i u_i + u_i^T r_i + k_i \\ & + \sum_{j \in \mathcal{N}_i} \left(u_i^T R_{ij} u_j + u_j^T R_{jj}^i u_j + u_j^T r_j^i\right) \end{aligned} \tag{39}$$

where $R_{ii}^i \in \mathbb{R}^{m_i \times m_i}$, $R_{ii}^i > 0$, $R_{ij} \in \mathbb{R}^{m_i \times m_j}$, $R_{jj}^i \in \mathbb{R}^{m_j \times m_j}$, $r_i \in \mathbb{R}^{m_i}$, $r_i^j \in \mathbb{R}^{m_j}$. Condition (8) becomes now

$$2R_{ii}^i u_i + \sum_{j \in \mathcal{N}_i} R_{ij} u_j + r_i = 0, \; i = 1, \ldots, N \tag{40}$$

which can be written as

$$Ru = -r \tag{41}$$

where $r = [r_1^T, \ldots, r_N^T]^T$ and

$$R = \begin{bmatrix} 2R_{11}^1 & R_{12} & \ldots & R_{1N} \\ R_{21} & 2R_{22}^2 & \ldots & R_{2N} \\ \vdots & \vdots & & \vdots \\ R_{N1} & R_{N2} & \ldots & 2R_{NN}^N \end{bmatrix} \tag{42}$$

where we assume that $R_{ij} = 0$ if $j \notin \mathcal{N}_i$. Therefore, the game admits a Nash equilibrium if and only if the system (41) has a solution. If the matrix R is invertible the system admits a unique Nash equilibrium given by $u^* = -R^{-1}r$. From (40) and (41) it is easy to derive that $g(\tilde{u}) = R\tilde{u}$ so that 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 (10) 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 one can deduce that it will be positive definite and diagonal under condition (38). Therefore, strict diagonal dominance of R together with condition (38) ensures stability, independently of the locally chosen parameters of the proposed algorithm. Furthermore, a diagonal form of the matrix K can always be ensured by applying sinusoidal perturbations with different frequencies for the x and y coordinates of each agent, as commented in Remark 5 in the context of the multi-dimensional case. In this case, condition (38) can be relaxed to (37). Also, if matrix $-R$ is stable and symmetric (implying that the underlying game is a potential game), for the stability of matrix $-KR$ it is sufficient that K is positive definite, ensured by (37).

The case of quadratic cost functions is important, since it represents a second order approximation of other types of nonlinear cost functions around the equilibrium point. The constant matrix R obtained above can be replaced by the Jacobian of the vector $g(\tilde{u})$ evaluated at the point $\tilde{u} = 0$. Then, the above global stability analysis for quadratic case can be applied for obtaining local stability conditions for general nonlinearities.

3) On the Vanishing Gains: Theoretically, for Assumptions (A.2)–(A.6) to be satisfied it is not required that the gains $\varepsilon_i(k)$ and $\varepsilon_j(k)$ and $\alpha_i(k)$ and $\alpha_j(k)$ are synchronized among the agents. Specifically, if $\varepsilon_i(k) = e_i k^{-m_\varepsilon}$ and $\alpha_i(k) = a_i k^{-m_\alpha}$, it is not required that $e_i = e_j$ and $a_i = a_j$, for $i, j \in \{1, \ldots, N\}$. However, if the asynchronicity among the agents is high so that the gains of some agents have already reached low enough values such that their further changes are negligible, while the gains of the other agents have not, the algorithm will practically never exactly reach the Nash equilibrium. This problem is related to the problem of slow convergence of stochastic approximation algorithms (see, e.g., [40], [45] or [41]). In order to deal with it, we can relax Assumptions (A.2)–(A.5) and define positive lower bounds for the time varying coefficients $\alpha_i(k)$ and $\varepsilon_i(k)$, at the expense of not being able to completely eliminate the noise influence. In this way, the algorithm could also *track* the position of the Nash equilibrium in the cases when it has some constant drift and is slowly changing in time. The lower bounds should be chosen in such a way as to achieve a compromise between the tracking capabilities of the algorithm, the convergence rate and the noise immunity. In this case the convergence analysis would require a quantification of the asymptotic expected value of the Lyapunov function $V(\tilde{u})$ (as was done in, e.g., [46] for similar iterative stochastic schemes), which is analogous to boundedness of solutions in classical, deterministic, stability theory. Also, if it is possible to neglect the measurement noise influence, for large enough k we can assume that the amplitudes are approximately constant, so that the scheme in this case reduces to deterministic Nash equilibrium seeking with constant amplitudes, for which global practical stability (in continuous time) have been analyzed in [33], [34] and for which some local stability results are presented in [17]. Furthermore, it is possible to apply adaptive procedures for selecting the gains $\varepsilon_i(k)$ and $\alpha_i(k)$ based on the observations of the noisy cost functions, by following similar principles presented in, e.g., [47].

4) Selection of the Perturbations Frequencies: Assumption (A.8) ensures decoupling of the agents' gradient estimates (by ensuring summability of all the terms in (23) and(24) after multiplication with the sinusoids of different frequency). Note that it is only necessary that the frequencies of *neighboring* agents are different. This ensures scalability, since the assigned frequencies can be repeated for the agents which do not affect each others' cost functions.

V. Applications To Mobile Sensor Networks

In order to apply the scheme depicted in Fig. 1 to the problems involving self-organizing networks of autonomous vehicles, with local sensory measurements (mobile sensor networks), we need to introduce continuous-time blocks that will model dynamics of the vehicles. In this problem setting, the vehicles are treated as players in a game, that are seeking positions corresponding to a Nash equilibrium. We are going to propose schemes for three frequently used models of autonomous vehicles in practice: velocity-actuated vehicles (single integrators), force-actuated vehicles (double integrators), and nonholonomic unicycles. Then we will apply these schemes to some typical problems in mobile robotic sensor networks: connectivity control, formation control, rendezvous and coverage control.

A. Velocity-Actuated Vehicles

In this subsection, we assume that the players of a Nash game are velocity-actuated autonomous vehicles moving in a plane. Hence, we model them as point masses such that

$$\dot{x}_i = v_i^x, \; \dot{y}_i = v_i^y \tag{43}$$

where $i = 1, \ldots, N$, $(x_i, y_i) = u_i$ are the positions of the vehicles and v_i^x and v_i^y are the velocity inputs. We will consider the proposed discrete-time Nash equilibrium seeking algorithm connected to (43), as shown in Fig. 2. The main difference, compared to the scheme in Fig. 1, besides its hybrid dynamics (ZOH denotes zero-order-hold blocks, and T is the sampling period), is that the integrators are moved in front of the perturbing signal, whose phase now needs to be adjusted to compensate for the integrators phase shift. Therefore, the perturbing signals $\hat{s}_i^x(k)$ and $\hat{s}_i^y(k)$ will have the following forms:

$$\hat{s}_i^x(k) = \alpha_i(k)\cos(\omega_i k) - \alpha_i(k-1)\cos(\omega_i(k-1)) \tag{44}$$

$$\hat{s}_i^y(k) = \alpha_i(k)\sin(\omega_i k) - \alpha_i(k-1)\sin(\omega_i(k-1)) \tag{45}$$

for all $i = 1, \ldots, N$. These signals can easily be mapped to the vehicle output, so we simply obtain

$$\begin{aligned} x_i(k) &= T\sum_{j=1}^{k} v_i^x(j) \qquad (46)\\ &= T\Big(\alpha_i(k)\cos(\omega_i k) \\ &\quad - \sum_{j=1}^{k}\varepsilon(j)\cos(\omega_i j - \varphi_i)w_i(j)\Big), \\ y_i(k) &= T\sum_{j=1}^{k} v_i^y(j) \\ &= T\Big(\alpha_i(k)\sin(\omega_i k) \\ &\quad - \sum_{j=1}^{k}\varepsilon(j)\sin(\omega_i j - \varphi_i)w_i(j)\Big). \qquad (47) \end{aligned}$$

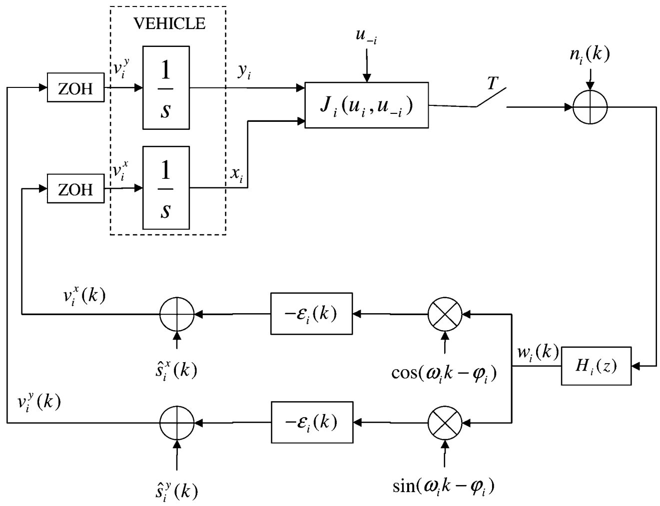

Fig. 2. Nash equilibrium seeking scheme for velocity-actuated vehicles.

Therefore, these mappings are the same as the corresponding mappings in Fig. 1, except for the multiplication with T which can be incorporated in the filter $F_i(z)$. We conclude that the equivalent overall discrete-time scheme corresponding to Fig. 2 is the scheme in Fig. 1 with $F_i(z) = T, i = 1, \ldots, N$. Hence, all the results from the previous section can be applied immediately. We summarize in the following corollary.

Corollary 1: Consider the system of networked velocity-actuated vehicles with Nash equilibrium seeking scheme defined in Fig. 2 where the perturbation signals $\hat{s}_i^x(k)$ and $\hat{s}_i^y(k)$ are defined in (44) and (45). Let Assumptions (A.1)–(A.12) be satisfied, with $F_i(z) = T$ and $G_i(z) = 1$. Then the positions $u(k) = [u_1(k)^T, \ldots, u_N(k)^T]^T$ of the vehicles converge to the Nash equilibrium u^* a.s.

B. Force-Actuated Vehicles

In Fig. 3, a scheme involving force-actuated vehicles (double integrators) is shown. The discrete-time integrator from Fig. 1 is again contained in the vehicle dynamics and moved in front of the perturbing signal. However, because of the vehicle's double integration, a discrete-time differentiator is needed to compensate one integration. Therefore, the perturbing signals are the same as in the single integrator case, given by (44) and (45). The equations modeling the behavior of the scheme are similar to the ones for the scheme in Fig. 2. The only difference is that we have now

$$x_i(k) = (1 - z^{-1})^2 \mathcal{Z}\{\mathcal{L}^{-1}\{\frac{1}{s^3}\}|_{t=kT}\}[v_i^x(k)] \quad (48)$$

$$y_i(k) = (1 - z^{-1})^2 \mathcal{Z}\{\mathcal{L}^{-1}\{\frac{1}{s^3}\}|_{t=kT}\}[v_i^y(k)] \quad (49)$$

where $\mathcal{Z}\{\cdot\}$ denotes the z-transform, $\mathcal{L}^{-1}\{\cdot\}$ the inverse Laplace transform, so that $x_i(k)$ and $y_i(k)$ are the equivalent discrete-time positions of the vehicles. By the following calculation:

$$\begin{aligned}(1 - z^{-1})^2 \mathcal{Z}\left\{\mathcal{L}^{-1}\{\frac{1}{s^3}\}|_{t=kT}\right\} &= \frac{1}{2}T^2(1 - z^{-1})\frac{z+1}{(z-1)^2} \\ &= \frac{1}{2}T^2(1 + z^{-1})\frac{1}{z-1} \quad (50)\end{aligned}$$

we conclude that the overall equivalent discrete-time scheme corresponding to Fig. 3 is the one in Fig. 1 with the input filters having the 2×2 diagonal transfer function matrices $F_i(z) = \mathrm{diag}\{\frac{1}{2}T^2(1 + z^{-1}), \frac{1}{2}T^2(1 + z^{-1})\}$. Therefore, we can formulate the following corollary:

Corollary 2: Consider the system of networked force-actuated vehicles with Nash equilibrium seeking scheme defined in Fig. 3 where the perturbation signals $\hat{s}_i^x(k)$ and $\hat{s}_i^y(k)$ are defined in (44) and (45). Let Assumptions (A.1)–(A.12) be satisfied, with $F_i(z) = \mathrm{diag}\{\frac{1}{2}T^2(1 + z^{-1}), \frac{1}{2}T^2(1 + z^{-1})\}$ and $G_i(z) = 1$. Then the positions $u(k) = [u_1(k)^T, \ldots, u_N(k)^T]^T$ of the vehicles converge to the Nash equilibrium u^* a.s.

C. Unicycles

Finally, we are going to consider the case in which the mobile robots are modeled as unicycles, having the sensors collocated at the centers of the vehicles. The equations of motion of the vehicles/sensors are

$$\dot{x}_i = v_i \cos\theta_i, \ \dot{y}_i = v_i \sin\theta_i, \ \dot{\theta}_i = \Omega_i^0 \quad (51)$$

where $(x_i, y_i) = u_i$ are the coordinates of the centers of the vehicles, θ_i their orientations and v_i, Ω_i^0 are the forward and angular velocity inputs, respectively, and $i = 1, \ldots, N$. For this vehicle model, because of the inherent nonholonomic constraints, the scheme from Fig. 1 cannot be applied directly, as in the case of single and double integrators. In this case, we are instead going to apply a scalar feedback, for each agent: we adjust only the forward velocity input v_i, keeping the angular velocity Ω_i^0 constant. Similar schemes have been effectively applied in [24] and [48] for single agent extremum seeking problems. The whole scheme containing both the vehicle and the discrete-time control algorithm is represented in Fig. 4. Our immediate concern is the mapping of the continuous-time unicycle variables to their discrete-time equivalents. It is straightforward to show that we have (see also [24])

$$x_i(k) = 2\frac{\sin\frac{\omega_i^0}{2}}{\Omega_i^0}\sum_{j=0}^{k-1} v_i(j)\cos\left(\omega_i^0(j + \frac{1}{2})\right) \quad (52)$$

$$y_i(k) = 2\frac{\sin\frac{\omega_i^0}{2}}{\Omega_i^0}\sum_{j=0}^{k-1} v_i(j)\sin\left(\omega_i^0(j + \frac{1}{2})\right) \quad (53)$$

where $\omega_i^0 = T\Omega_i^0$. Assuming that the perturbation signal is given by

$$\hat{s}_i(k) = -\alpha_i(k)\sin\left(\omega_i(k + \frac{1}{2})\right) \quad (54)$$

we obtain that its maps to the cost function inputs, $s_i^x(k) = 2\frac{\sin\frac{\omega_i^0}{2}}{\Omega_i^0}\sum_{j=0}^{k-1}\hat{s}_i(j)\cos(\omega_i^0(j + \frac{1}{2}))$ and $s_i^y(k) = 2\frac{\sin\frac{\omega_i^0}{2}}{\Omega_i^0}\sum_{j=0}^{k-1}\hat{s}_i(j)\sin(\omega_i^0(j + \frac{1}{2}))$, are

$$\begin{aligned}s_i^x(k) = \frac{\sin\frac{\omega_i^0}{2}}{\Omega_i^0}\sum_{j=0}^{k-1}\alpha_i(j)(\sin((\omega_i + \omega_i^0)(j + \frac{1}{2})) \\ + \sin((\omega_i - \omega_i^0)(j + \frac{1}{2}))) \quad (55)\end{aligned}$$

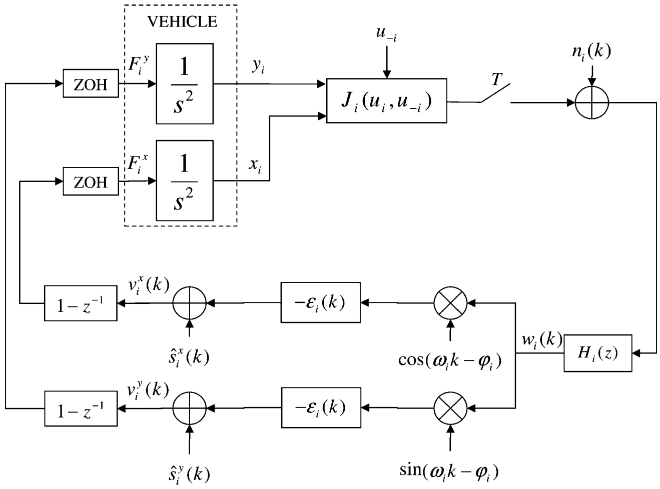

Fig. 3. Nash equilibrium seeking scheme for force-actuated vehicles.

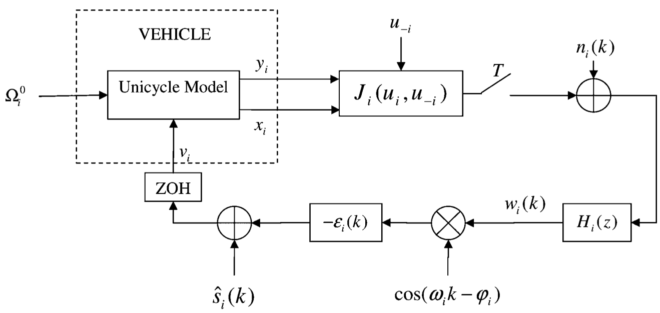

Fig. 4. Nash equilibrium seeking scheme for unicycles.

$$s_i^y(k) = \frac{\sin\frac{\omega_i^0}{2}}{\Omega_i^0}\sum_{j=0}^{k-1}\alpha_i(j)(-\cos((\omega_i+\omega_i^0)(j+\frac{1}{2})) + \cos((\omega_i-\omega_i^0)(j+\frac{1}{2}))). \quad (56)$$

After further applying standard trigonometric transformations and by doing the discrete-time integration we obtain

$$s_i^x(k) = \kappa_1^i\alpha_i(k)\cos((\omega_i+\omega_i^0)k) + \kappa_2^i\alpha_i(k)\cos((\omega_i^0-\omega_i)k) + \delta_i^x(k) \quad (57)$$

$$s_i^y(k) = \kappa_1^i\alpha_i(k)\sin((\omega_i+\omega_i^0)k) + \kappa_2^i\alpha_i(k)\sin((\omega_i^0-\omega_i)k) + \delta_i^y(k) \quad (58)$$

where $\kappa_1^i = \frac{\sin\frac{\omega_i^0}{2}}{2\Omega_i^0\sin\frac{\omega_i^0+\omega_i}{2}}$, $\kappa_2^i = \frac{\sin\frac{\omega_i^0}{2}}{2\Omega_i^0\sin\frac{\omega_i-\omega_i^0}{2}}$ and $\delta_i^x(k)$ and $\delta_i^y(k)$ are $O(\alpha_i(k)-\alpha_i(k-1))$ and, therefore, absolutely summable due to (A.2). Now, we define the "perturbation signal" $u_0^i(k) = [s_i^x(k),\ s_i^y(k)]^T$ [compare to $u_0^i(k)$ in (4)], and consider the tracking error given by(9). Therefore, from (52) and (53) we obtain the following recursive tracking error equation, analogous to (19), for each agent

$$\tilde{u}_i(k+1) = \tilde{u}_i(k) - 2\frac{\sin\frac{\omega_i^0}{2}}{\Omega_i^0}C_i(k)\varepsilon_i(k)\cos(\omega_i k - \varphi_i)w_i(k) \quad (59)$$

where $C_i(k) = [\cos(\omega_i^0(k+\frac{1}{2})),\ \sin(\omega_i^0(k+\frac{1}{2}))]^T$ and $w_i(k) = H_i(z)[J_i(u_i(k), u_{-i}(k)) + n_i(k)]$.

By converting the products of sinusoids in (59) into sums, and by proceeding in an analogous way to the proof of Theorem 1 [the derivation after equation (19)], we obtain the same tracking equation (31) but with a slightly different matrix K. Namely, from (59), (57), and (58) it can be seen that the blocks K_i of matrix K will be the sums of two terms, $K_i = \frac{\sin\frac{\omega_i^0}{2}}{\Omega_i^0}(K_1^i + K_2^i)$, because of the presence of two "perturbing" signals, and two "demodulating" signals, originating from conversion of the products of sinusoids in (59) into sums with frequencies $\omega_i^0+\omega_i$ and $\omega_i^0-\omega_i$ and with corresponding phase shifts. Therefore, we replace Assumption (A.12) with (A.12'), which differs only in the definition of K_i:

(A.12') There exists a continuously differentiable Lyapunov function $V(\tilde{u}) : \mathbb{R}^{2N} \rightarrow \mathbb{R}$ such that

$$-g(\tilde{u})^T K^T \nabla_{\tilde{u}} V(\tilde{u}) < 0 \quad (60)$$

for all $\tilde{u} \neq 0$, $\tilde{u} \in B$, where $K = \mathrm{diag}\{c_1K_1, \ldots, c_NK_N\}$, $K_i = \frac{\sin\frac{\omega_i^0}{2}}{\Omega_i^0}(K_1^i + K_2^i)$, $K_p^i = \kappa_p^i\begin{bmatrix}\mathrm{Re}\{\theta_p^i\} & \mathrm{Im}\{\theta_p^i\}\\ -\mathrm{Im}\{\theta_p^i\} & \mathrm{Re}\{\theta_p^i\}\end{bmatrix}$ for $p \in \{1,2\}$, $\theta_1^i = e^{j\varphi_i}e^{-j\frac{\omega_i^0}{2}}H_i(e^{j(\omega_i^0+\omega_i)})$, and $\theta_2^i = e^{-j\varphi_i}e^{-j\frac{\omega_i^0}{2}}H_i(e^{j(\omega_i^0-\omega_i)})$. The other terms in the overall tracking equation will remain the same, except for the structural differences in the "perturbation" terms inside $\pi(k)$ in (31), which will remain summable a.s. if Assumption (A.8) is replaced by:

(A.8') $\omega_i = a_i\pi$, $\omega_i^0 = a_i^0\pi$: $a_i, a_i^0 \in (0,1)$ are rational numbers, $\omega_i \neq \omega_i^0$, and for all permutations $(\omega_1^p, \omega_2^p, \omega_3^p, \omega_4^p)$ of set $\{\omega_i, \omega_j, \omega_i^0, \omega_j^0\}$, $i = 1, \ldots, N$ and $j \in \mathcal{N}_i$, the following holds: $|\omega_1^p - \omega_2^p| \neq |\omega_3^p - \omega_4^p|$, $\omega_1^p + \omega_2^p \neq |\omega_3^p - \omega_4^p|$, $\omega_1^p + \omega_2^p \neq \omega_3^p + \omega_4^p$, $2\pi - (\omega_1^p + \omega_2^p) \neq |\omega_3^p - \omega_4^p|$ and $2\pi - (\omega_1^p + \omega_2^p) \neq \omega_3^p + \omega_4^p$.

Thus, we have proved the following theorem:

Theorem 2: Consider the system of networked unicycles with Nash equilibrium seeking scheme defined in Fig. 4 where the perturbation signal $\hat{s}_i(k)$ is defined in (54). Let Assumptions (A.1)–(A.7), (A.8'), (A.9)–(A.11), and (A.12') be satisfied. Then the positions $u(k) = [u_1(k)^T, \ldots, u_N(k)^T]^T$ of the vehicles converge to the Nash equilibrium u^* a.s.

D. Applications

In mobile sensor networks the information that the agents have about the environment as well as about the actions and properties of the other agents is typically limited to certain local sensing and local low bandwidth communications. Therefore, the proposed schemes can be effectively applied here, due to their adaptive nature and because the problem is approached in the framework of noncooperative games, which can effectively capture distributed information structure constraints. The problem of designing individual cost functions in such a way that a Nash equilibrium corresponds to some global goal or a Pareto optimal point has been treated extensively in the existing literature (see, e.g., [3], [6], [9], [10]). In general, achieving a social (centralized) goal is not an easy task in noncooperative scenarios. The agents are acting selfishly (locally) and the cooperation is to be imposed by proper design of the agents' cost functions. In what follows, we are going to present some examples of how to select the agents' costs such that, by applying the proposed Nash equilibrium seeking algorithm, some typical problems in mobile sensor networks can be solved in an *adaptive* and *distributed* way.

1) Connectivity Control: Connectivity control in mobile robotic networks has been analyzed intensively in the existing literature. In many practical applications these networks are designed for achieving some primary objective, assuming that the overall connectivity is preserved or kept above some threshold level. In these situations, connectivity preserving can be considered as a secondary objective (see, e.g., [25]–[28] and references therein). In what follows we propose an approach related to [25], [27], where the authors use potential functions for locally preserving existing links in the network while performing some primary objective. However, our approach does not require any direct inter-agent or absolute position measurements and can be applied to the robots with any motion dynamics mentioned in previous subsections.

Broadening the scenarios analyzed in [22] and [24], where an agent is either searching for a source of some signal with unknown distribution, or positioning itself to an optimal sensing point for an estimation task, in our interconnected problem setting, the individual costs of the agents can be designed to achieve a compromise between the mentioned "local" goals, and a "collective" goal of keeping good connections with selected neighboring agents. This can be important in, for example, distributed estimation where the local estimators are communicating with each other to improve the overall performance (see, e.g., [49]). Hence, the cost functions for agents can be written as sums $J_i(u_i, u_{-i}) = \mathcal{L}_i(u_i) + \mathcal{C}_i(u_i, u_{-i})$ where $\mathcal{L}_i(u_i)$ corresponds to a "local" goal (depending only on the decision u_i of agent i) and $\mathcal{C}_i(u_i, u_{-i})$ is an interconnection term defining a "collective" goal. The former one 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, which can be directly measured. In the latter case, assuming that the signal power is inversely proportional to the squared distance between the agents, i.e., $\mathcal{P}(u_i, u_j) \sim 1/\|u_i - u_j\|^2$, and taking its reciprocal value as the interconnection term which is to be minimized, we can define quadratic cost functions as

$$J_i(u_i, u_{-i}) = u_i^T r_{ii} u_i + u_i^T r_i + k_i + \sum_{j \in \mathcal{N}_i} m_{ij} \|u_i - u_j\|^2 \quad (61)$$

where we assumed that local goals $\mathcal{L}_i(u_i)$ are strictly convex quadratic functions, i.e., that $r_{ii} > 0$, $\|\cdot\|$ is the Euclidian norm and the coefficients m_{ij} are selected *a priori*, reflecting the importance of the signal received from the jth agent. Also, we assume that the communication topology is fixed, i.e., that the sets $\mathcal{N}_i$ are time-invariant for all i. Therefore, the elements of the matrix R in (42) 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. Hence, in this case the game will always admit a unique Nash equilibrium and the condition (A.12) is satisfied for any diagonal positive definite K [see conditions (37) and (38)]. Therefore, when deployed, the mobile agents do not need to know the parameters of the cost functions (61): they only need to measure the "local" costs and the power of a signal received from the neighbors.

2) Formation Control and Rendezvous: Consider the following cost functions $J_i(u_i, u_{-i}) = \sum_{j \in \mathcal{N}_i} m_{ij} \|u_i - u_j - r_{ij}\|^2$ which correspond to a formation control problem (see, e.g., [30]), where r_{ij} are the desired vectors of inter-agent distances, and m_{ij} are positive constants. Note that this can be considered as a special case of the connectivity control costs (61) where the quadratic term corresponding to the "local" objective is zero ($r_{ii} = 0$) and where $r_i = \sum_{j \in \mathcal{N}_i} -2m_{ij} r_{ij}$. Therefore, the matrix R is not going to be strictly diagonally dominant anymore since at least one eigenvalue is 0. Hence, if the desired distances r_{ij} are feasible, i.e., if they are defined such that (41) has a solution, we will have infinite number of Nash equilibria. Furthermore, observe that, in this case, the matrix R is actually the weighted Laplacian matrix of the graph defined by the coefficients m_{ij}. Hence, by relaxing (A.12) to allow multiple equilibria as commented in Remark 2, we can again apply our Nash equilibrium seeking schemes. The set of Nash equilibria has to satisfy linear equation (41), so that (31), without the "structural perturbation" and noise terms, represents, in this case, a standard linear formation control algorithm (or a consensus algorithm with constant input term r, see, e.g., [29], [30]) with time-varying gains ($\rho(k)$), so that we can choose standard Lyapunov function applied to these problems (see, e.g., [29]). However, in order to obtain the values of the individual costs, the agents need to measure their own absolute positions and distances to the neighbors (or neighbors' absolute positions), which might not be more efficient than just using a gradient descent algorithm. Nevertheless, if we set $r_{ij} = 0$ for all $i, j \in \{1, \ldots, N\}$, and if the underlying, time-invariant graph having the Laplacian matrix R, is strongly connected [30], the agents will converge to a single point, thus achieving a *consensus* on positions, or *rendezvous*. In this case, the agents do not need absolute position measurements or direct measurements of inter-agent distances; convergence can be achieved based only on the power of a signal received from the neighbors. A similar algorithm has been analyzed in [50] where the almost sure convergence to a consensus point is proved assuming that the inter-agent state differences can be measured (with additive noise). See also [32] where the consensus algorithm was treated in the context of potential games.

3) Coverage Control: As mentioned in Section IV, if the underlying game is a potential game, the potential function can be chosen as a Lyapunov function in (A.12). Based on this result, it is possible to apply the proposed schemes to the *coverage control* problem defined in [31] and formulated as a potential game in [32], [33]. Namely, by taking the global coverage control objective function (as defined in [31]) as a potential function, it is possible to assign, so called, *Wonderful Life* individual cost functions [3], [32], [33] to each agent. It can be shown that these costs have physical meaning and that, assuming limited detection radius, their value at the current position can be obtained by only locally counting detected events and by communicating only with the close enough neighboring agents (see also [33] where a detailed analysis of this problem is presented). These cost functions actually encode a proximity based communication topology. By applying the proposed Nash equilibrium seeking schemes, we can solve the coverage control problem in a distributed way, without position measurements and without

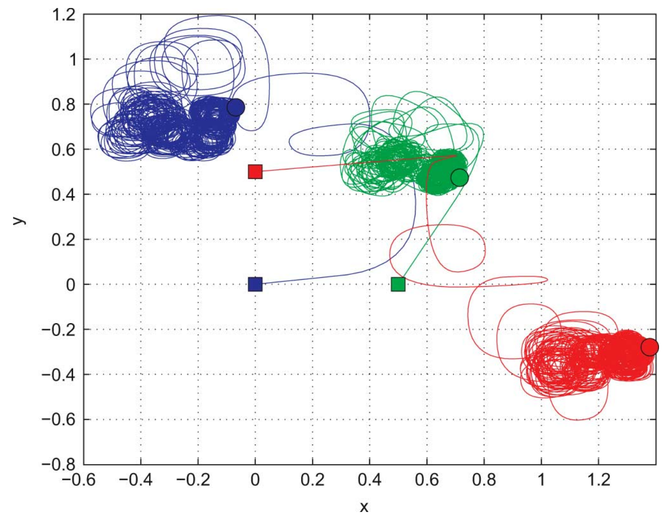

Fig. 5. Trajectories of the force-actuated vehicles.

any *a priori* knowledge about the distribution of the events to be detected and about the detection capabilities of the individual agents.

E. Simulations

Example 1: In this example we illustrate the algorithm proposed in Fig. 3 for a network of three force-actuated vehicles where the cost functions are given by (61) 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 (41) we obtain that the unique Nash equilibrium is the point $u^* = [-0.125,\ 0.75,\ 0.75,\ 0.5,\ 1.375, -0.25]^T$. It reflects the compromise between the agents' "local" objectives ($[-1,\ 1]^T$, $[1,\ 1]^T$ and $[2, -1]^T$), and the "collective" objective of maintaining the network connectivity, determined by the values of the interconnection coefficients (m_{ij}, $i, j \in \{1, 2, 3\}$, $i \neq j$). For the other system parameters we assume the following values: the noise covariance matrix (6) is $\Sigma(k) = \mathrm{diag}\{0.1,\ 0.1,\ 0.1\}$, $\varphi_i = -\pi/4$ (this phase shift is needed to compensate for the shift obtained in (50) so that the matrix K in (A.12) is positive definite), $T = 0.1$, $H_i(z) = \frac{z-1}{z+0.07}$ (washout 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 the vehicles 1 and 3 since they are not interconnected. The trajectories of the vehicles are shown in Fig. 5, and the coordinates, as a function of time, for the first vehicle are shown in Fig. 6, for the initial conditions $u_1(1) = [0,\ 0]^T$, $u_2(1) = [0.5,\ 0]^T$, $u_3(1) = [0,\ 0.5]^T$. The time responses for the other two vehicles are similar. The convergence to the Nash equilibrium while eliminating the measurement noise, is evident. The final minor fluctuations around the equilibrium point are due to the slow convergence of the perturbation amplitudes $\alpha_i(k) = 0.4k^{-0.25}$ which are always present as additive inputs. These amplitudes converge to zero much slower than the feedback signals (the outputs of the integrators in Fig. 1) which are additionally multiplied by $\varepsilon_i(k)$. This can also be seen from the definition of the tracking error (9) which converges to zero faster than the perturbation signal alone.

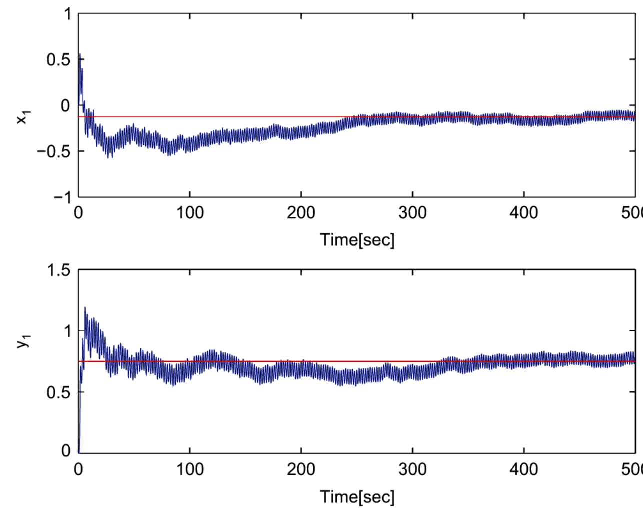

Fig. 6. Coordinates of the first vehicle. The Nash equilibrium is shown as the red line.

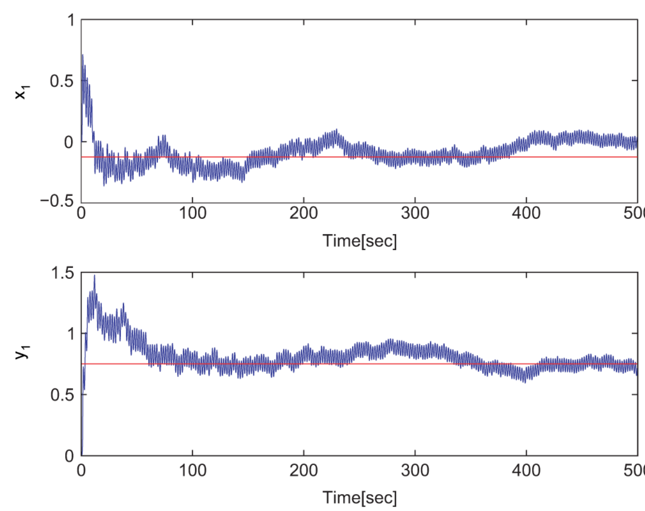

Fig. 7. Coordinates of the first vehicle: slower convergence rate. The Nash equilibrium is shown as the red line.

In Fig. 7, a time response for the first agent is depicted for the case of slower convergence of the gains $\varepsilon_i(k)$ which are here $\varepsilon_i(k) = 1.5k^{-0.55}$, $i = 1, 2, 3$ while all the other parameters are kept the same. In this case, the convergence rate to the Nash equilibrium is slower and the algorithm is more sensitive to noise, compared to the responses in Fig. 6.

Example 2: In this example, we consider the formation control game, as presented in Section V-D2, performed by three unicycles. Their goal is to reach a formation in which all mutual distances are the same, equal to 1. Therefore, we apply the scheme shown in Fig. 4 with the cost functions given in (61) and with the following parameters: $r_{11} = r_{22} = r_{33} = 0$, $r_1 = [0, -2\sqrt{3}]^T$, $r_2 = [-3, -\sqrt{3}]^T$, $r_3 = [3, -\sqrt{3}]^T$, $k_1 = k_2 = k_3 = 0$, and $m_{12} = m_{21} = m_{23} = m_{32} = m_{13} = m_{31} = 1$, whose Nash equilibrium corresponds to the formation with all the inter-vehicle distances equal to 1. The other parameters of the scheme are $\Sigma(k) = \mathrm{diag}\{0.04,\ 0.04,\ 0.04\}, T = 0.1, H_i(z) = \frac{z-1}{z+0.07}$, $\varphi_i = 0$, $\varepsilon_i(k) = k^{-0.65}$, $\alpha_i(k) = 0.6k^{-0.25}$, $\omega_i^0 = 0.14\pi$ ($\Omega_i^0 = \omega_i^0/T \approx 4.4$), for $i = 1, 2, 3$, $\omega_1 = 0.5\pi$, $\omega_2 = 0.8\pi$,

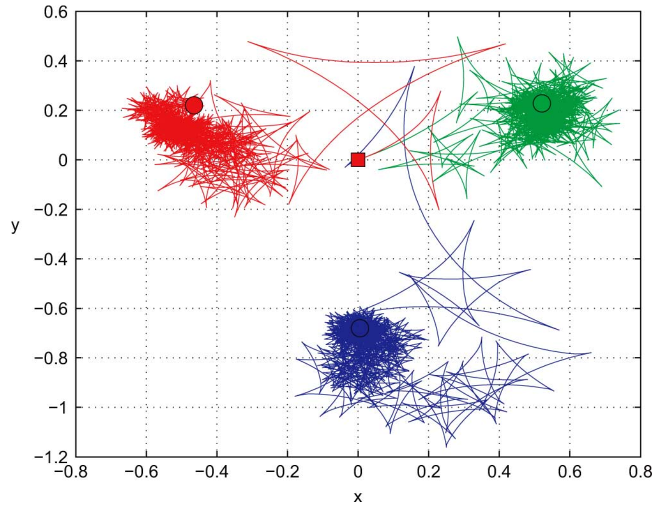

Fig. 8. Trajectories of the unicycles.

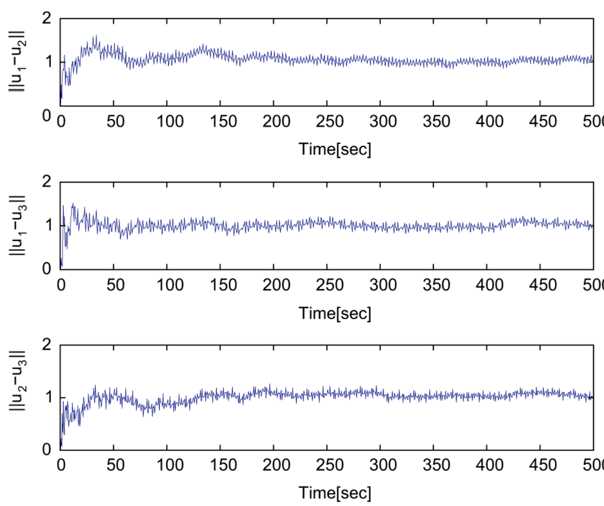

Fig. 9. Distances between the unicycles.

and $\omega_3 = 0.6\pi$. It is easy to check that all the conditions necessary for applying Theorem 2 are satisfied. Trajectories of the unicycles are shown in Fig. 8, where we have assumed that all the vehicles are initially at the origin. Trajectories are typical for vehicles with rolling without slipping condition (due to nonholonomic constraints) where spikes correspond to the points where a vehicle changes direction of motion. Distances between the vehicles, as functions of time, are shown in Fig. 9. It is evident that all the distances converge to the desired one. The center of the formation depends on the initial conditions and the noise realization. As in the previous example, the small fluctuations after the Nash equilibrium has already been reached, are due to the slow convergence of the perturbation amplitudes $\alpha_i(k)$, compared to the convergence rate of the tracking error (9).

VI. Conclusion

We have proposed a method for distributive seeking of Nash equilibria in noncooperative games, based only on measurements of the individual cost functions, corrupted by noise. The players are allowed to possess some local linear dynamics, so that their actions are filtered before affecting the measured cost functions. We have formulated conditions on the structure of the game and on the parameters of the proposed scheme, under which we proved almost sure convergence to a Nash equilibrium. It is demonstrated that the proposed method can be applied to networks of mobile robots, where the robots can have single integrator, double integrator or unicycle motion dynamics. We argue that it is desirable and, in some cases, inevitable to use the proposed algorithm for solving problems in mobile sensor networks since these networks usually operate in only partially known or unknown and unpredictable environments with distributed information structure constraints. We give examples of how to formulate problems of connectivity control, formation control, rendezvous, and coverage control as noncooperative games, which can then be directly solved using the proposed framework. The proposed schemes have been illustrated through simulations.

As a possible future research direction, one could consider extending the proposed schemes such that they can handle hard-constraints on the actions/positions of the players. In this way, the convergence to a Nash equilibrium could be guaranteed while achieving a collision and/or obstacle avoidance.

Appendix

Proof of Lemma 3: From (12) it follows that $\delta(k) = H(z)[\varepsilon(k)x(k)] - \varepsilon(k)H(z)[x(k)]$. If $h(k)$, $k = 0, 1, 2, \ldots$ is the impulse response matrix of the system $\mathcal{S}_H$ with the transfer function matrix $H(z)$, we have

$$\begin{aligned}\delta(k) = &\, h(0)(\varepsilon(k) - \varepsilon(k))x(k) \\ &+ h(1)(\varepsilon(k-1) - \varepsilon(k))x(k-1) + \cdots \\ &+ h(k-1)(\varepsilon(1) - \varepsilon(k))x(1) + \epsilon^{-k}\end{aligned} \tag{62}$$

so that

$$\delta(k) = (\varepsilon(k-1) - \varepsilon(k))y(k) + \epsilon^{-k} \tag{63}$$

where ϵ^{-k} denotes exponentially decaying terms (due to possible nonzero initial condition), and $y(k)$ can be considered as the output of a time-varying MIMO system $\mathcal{S}_\varepsilon$ with the impulse response matrix $h_\varepsilon(k,i) = h(i)\frac{\varepsilon(k-i)-\varepsilon(k)}{\varepsilon(k-1)-\varepsilon(k)}$ and input $x(k)$, i.e., $y(k) = \sum_{i=0}^{k-1} h_\varepsilon(k,i)x(k-i)$. System $\mathcal{S}_\varepsilon$ is bounded-input, bounded-output (b.i.b.o.) stable, having in mind that all the elements of $h_\varepsilon(k,i)$ are absolutely summable under the formulated assumptions [$\mathcal{S}_H$ is exponentially stable and $\varepsilon(k)$ satisfies (A.2)–(A.5)]. Therefore, $\|\sum_{k=1}^{\infty}\delta(k)\| < \infty$ since $\varepsilon(k)$ satisfies Assumption (A.2) and $x(k)$ is bounded. ■

Proof of Lemma 4: From (14) and the fact that $H(1) = \sum_{i=0}^{\infty} h(i)$, where $h(i)$ is the impulse response matrix of $H(z)$, we obtain

$$\begin{aligned}\delta(k) &= H(z)[x(k)] - H(1)x(k) \\ &= \sum_{i=0}^{k-1} h(i)(x(k-i) - x(k)) \\ &\quad + \left(\sum_{i=0}^{k-1} h(i) - \sum_{i=0}^{\infty} h(i)\right) x(k) + \epsilon^{-k}\end{aligned} \tag{64}$$

where ϵ^{-k} denotes exponentially decaying terms (due to possible nonzero initial condition). After iterating (13) i times and plugging into the first term in (64), we obtain

$$\begin{aligned}\delta(k) = &-h(1)\varepsilon(k-1)y(k-1)\\ &-h(2)(\varepsilon(k-2)y(k-2)\\ &+\varepsilon(k-1)y(k-1))-\cdots\\ &-\sum_{i=k}^{\infty}h(i)x(k)+\epsilon^{-k}.\end{aligned} \tag{65}$$

After regrouping the terms in (65), we obtain

$$\delta(k)=\sum_{j=1}^{k-1}\left(-\sum_{i=j}^{k-1}h(i)\right)\varepsilon(k-j)y(k-j)+\epsilon^{-k} \tag{66}$$

where we have incorporated $\sum_{i=k}^{\infty}h(i)x(k)$ in ϵ^{-k} due to its exponential decaying. Defining a time varying system $\mathcal{S}_\varepsilon$ with the impulse response matrix $h_\varepsilon(k,j)=h_1(k,j)\frac{\varepsilon(k-j)}{\varepsilon(k-1)}$, where $h_1(k,j)=-\sum_{i=j}^{k-1}h(i)$, we can write

$$\delta(k)=\varepsilon(k-1)y_1(k)+\epsilon^{-k} \tag{67}$$

where $y_1(k)=\sum_{j=1}^{k-1}h_\varepsilon(k,j)y(k-j)$ is the output of $\mathcal{S}_\varepsilon$ when the input is $y(k)$, which is uniformly bounded by assumption. One can easily verify that $\mathcal{S}_\varepsilon$ is b.i.b.o. stable under the adopted assumptions implying that $y_1(k)$ is uniformly bounded. Therefore, we can conclude that $\|\sum_{k=1}^{\infty}\varepsilon(k)\delta(k)\|<\infty$ since $\varepsilon(k)$ satisfies (A.5). ∎

Proof of Lemma 5: By denoting $z(k) = x_1(k)^{n_1}x_2(k)^{n_2}\cdots x_N(k)^{n_N}$, we have that

$$\begin{aligned}&\left|\sum_{k=1}^{\infty}z(k)\varepsilon_1(k)\cos(\omega k+\varphi)\right|\\ &\leq\left|\sum_{j=1}^{T_\omega}b_j\sum_{k=0}^{\infty}\varepsilon_1(j+2kT_\omega)z(j+2kT_\omega)\right.\\ &\quad\left.-\varepsilon_1(j+T_\omega+2kT_\omega)z(j+T_\omega+2kT_\omega)\right|\\ &=\left|\sum_{j=1}^{T_\omega}b_j\sum_{k=0}^{\infty}(\varepsilon_1(j+2kT_\omega)\right.\\ &\quad-\varepsilon_1(j+T_\omega+2kT_\omega))z(j+2kT_\omega)\\ &\quad-\varepsilon_1(j+T_\omega+2kT_\omega)(z(j+T_\omega+2kT_\omega)\\ &\quad\left.-z(j+2kT_\omega))\right|\end{aligned} \tag{68}$$

for some $b_j\geq 0$, $j=1,\ldots,T_\omega$, where $2T_\omega$ is the integer period of $\cos(\omega k)$. The first term in the last expression in (68) converges due to the fact that $\varepsilon_1(k)$ satisfies (A.2). For the second term we will show that $|z(k+T_\omega)-z(k)|\leq K_z\varepsilon_1(k)$, for some constant K_z. From (15), by using binomial expansion, it follows that

$$x_i(k)^{n_i}-x_i(k-1)^{n_i}=\sum_{j=1}^{n_i}\binom{n_i}{j}x_i(k)^{n_i-j}(\varepsilon_i(k)y_i(k))^j \tag{69}$$

for some $n_i\in\{1,2,\ldots\}$. Therefore, due to the boundedness of $x_i(k)$ and $y_i(k)$ it is easy to derive that $|x_i(k)^{n_i}-x_i(k-1)^{n_i}|\leq|K_{n_i}\varepsilon_i(k)v_i(k)|$, for some uniformly bounded $v_i(k)$ and some constant K_{n_i}, $i=1,\ldots,N$. From this it follows that $|z(k+T_\omega)-z(k)|\leq K_z\varepsilon_1(k)$, for some large enough K_z. This proves that the last term in (68) is summable, due to (A.5), which proves the lemma. ∎

REFERENCES

[1] M. S. Stanković, K. H. Johansson, and D. M. Stipanović, "Distributed seeking of Nash equilibria in mobile sensor networks," in *Proc. IEEE Conf. Decision Control*, 2010, pp. 5598–5603.

[2] J. Nash, "Non-cooperative games," *Ann. Math.*, vol. 54, no. 2, pp. 286–295, 1951.

[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 Econ. Behav.*, no. 14, pp. 124–143, 1996.

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

[6] P. Dubey, "Inefficiency of Nash equilibria," *Math. Operat. Res.*, vol. 11, no. 1, pp. 1–8, Feb. 1986.

[7] G. Inalhan, D. M. Stipanović, and C. J. Tomlin, "Decentralized optimization, with application to multiple aircraft coordination," in *Proc. IEEE Conf. Decision Control*, 2002, pp. 1147–1155.

[8] L. Buttyan, J.-P. Hubaux, L. Li, X.-Y. Li, T. Roughgarden, and A. Leon-Garcia, "Guest editorial non-cooperative behavior in networking," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 6, Aug. 2007.

[9] T. Roughgarden, "Intrinsic robustness of the price of anarchy," in *Proc. STOC '09: 41st ACM Symp. Theory of Comput.*, New York, pp. 513–522.

[10] N. Li and J. R. Marden, "Designing games for distributed optimization," in *Proc. IEEE Conf. Decision Control*, 2011.

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

[12] J. S. Shamma and G. Arslan, "Dynamic fictitious play, dynamic gradient play, and distributed convergence to Nash equilibria," *IEEE Trans. Autom. Control*, vol. 50, no. 3, pp. 312–327, Mar. 2005.

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

[14] M. Zhu and S. Martinez, "Distributed coverage games for mobile visual sensors (I) : Reaching the set of Nash equilibria," in *Proc. IEEE Conf. Decision Control*, 2009, pp. 169–174.

[15] H. P. Young, "Learning by trial and error," *Games Econ. Behav.*, vol. 65, no. 2, pp. 626–643, 2009.

[16] M. Krstic, P. Frihauf, J. Krieger, and T. Basar, "Nash equilibrium seeking with finitely- and infinitely-many players," in *Proc. 8th IFAC Symp. Nonlinear Control Syst.*, 2010.

[17] P. Frihauf, M. Krstic, and T. Basar, "Nash equilibrium seeking for dynamic systems with non-quadratic payoffs," in *Proc. 14th Int. Symp. Dynamic Games Applicat.*, 2010.

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

[19] J. Y. Choi, M. Krstic, K. B. Ariyur, and J. S. Lee, "Extremum seeking control for discrete-time systems," *IEEE Trans. Autom. Control*, vol. 47, no. 2, pp. 318–323, Feb. 2002.

[20] Y. Tan, D. Nešić, and I. Mareels, "On non-local stability properties of extremum seeking control," *Automatica*, vol. 42, pp. 889–903, 2006.

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

[22] M. S. Stanković and D. M. Stipanović, "Extremum seeking under stochastic noise and applications to mobile sensors," *Automatica*, vol. 46, pp. 1243–1251, 2010.

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

[24] M. S. Stanković and D. M. Stipanović, "Discrete time extremum seeking by autonomous vehicles in a stochastic environment," in *Proc. IEEE Conf. Decision Control*, 2009, pp. 4541–4546.

[25] M. Zavlanos, A. Jadbabaie, and G. Pappas, "Flocking while preserving network connectivity," in *Proc. IEEE Conf. Decision Control*, 2007, pp. 2919–2924.
[26] M. Schuresko and J. Cortes, "Distributed motion constraints for algebraic connectivity of robotic networks," *J. Intell. Robot. Syst.*, vol. 56, no. 1–2, pp. 99–126, 2009.
[27] D. V. Dimarogonas and K. H. Johansson, "Bounded control of network connectivity in multi-agent systems," *IET Control Theory Applicat.*, vol. 4, no. 8, pp. 1330–1338, 2010.
[28] P. Yang, R. Freeman, G. Gordon, K. Lynch, S. Srinivasa, and R. Sukthankar, "Decentralized estimation and control of graph connectivity for mobile sensor networks," *Automatica*, vol. 46, no. 2, pp. 390–396, 2010.
[29] M. Mesbahi and M. Egerstedt, *Graph Theoretic Methods in Multiagent Networks*. Princeton, NJ: Princeton Univ. Press, 2010.
[30] R. Olfati-Saber, J. A. Fax, and R. M. Murray, "Consensus and cooperation in networked multi-agent systems," *Proc. IEEE*, vol. 95, no. 1, pp. 215–233, Jan. 2007.
[31] C. G. Cassandras and W. Li, "Sensor networks and cooperative control," *Eur. J. Control*, vol. 11, no. 4–5, pp. 436–463, 2005.
[32] J. R. Marden, G. Arslan, and J. S. Shamma, "Cooperative control and potential games," *IEEE Trans. Syst., Man, Cybern. B*, vol. 39, no. 6, pp. 1393–1407, Dec. 2009.
[33] H.-B. Dürr, M. S. Stanković, and K. H. Johansson, "Distributed positioning of autonomous mobile sensors with application to coverage control," in *Proc. Amer. Control Conf.*, 2011, pp. 4822–4827.
[34] H.-B. Dürr, M. S. Stanković, and K. H. Johansson, "Nash equilibrium seeking in multi-vehicle systems: A Lie bracket approximation-based approach," *arXiv:1109.6129v2*, 2011.
[35] G. M. Hoffmann and C. J. Tomlin, "Mobile sensor network control using mutual information methods and particle filters," *IEEE Trans. Autom. Control*, vol. 55, no. 1, pp. 32–47, Jan. 2010.
[36] A. Ganguli, S. Susca, S. Martinez, F. Bullo, and J. Cortes, "On collective motion in sensor networks: Sample problems and distributed algorithms," in *Proc. IEEE Conf. Decision Control*, Dec. 2005, pp. 4239–4244.
[37] J. Cortes, S. Martinez, T. Karatas, and F. Bullo, "Coverage control for mobile sensing networks," *IEEE Trans. Robot. Autom.*, vol. 20, no. 2, pp. 243–255, Apr. 2004.
[38] J. Cortes, S. Martinez, and F. Bullo, "Robust rendezvous for mobile autonomous agents via proximity graphs in arbitrary dimensions," *IEEE Trans. Autom. Control*, vol. 51, no. 8, pp. 1289–1298, Aug. 2006.
[39] B. T. Polyak and Y. Z. Tsypkin, "Pseudogradient algorithms of adaptation and learning," *Autom. Remote Control*, no. 3, pp. 45–68, 1973.
[40] H. F. Chen, *Stochastic Approximation and its Applications*. Norwell, MA: Kluwer, 2003.
[41] H. J. Kushner and D. S. Clark, *Stochastic Approximation Methods for Contrained and Unconstrained Systems*. New York: Springer, 1978.
[42] L. Ljung, "Analysis of recursive stochastic algorithms," *IEEE Trans. Autom. Control*, vol. AC-22, no. 4, pp. 551–575, Aug. 1977.
[43] J. B. Rosen, "Existence and uniqueness of equilibrium points for concave n-person games," *Econometrica*, vol. 33, no. 3, pp. 520–534, 1965.
[44] A. S. Poznyak and D. O. Chikin, "Asymptotic properties of the stochastic approximation procedure with dependent noise," *Autom. Remote Control*, no. 12, pp. 78–93, 1984.
[45] B. T. Polyak, "Convergence and rate of convergence of iterative stochastic algorithms, II linear case," *Autom. Remote Control*, no. 4, pp. 101–107, 1977.
[46] B. T. Polyak, "Convergence and rate of convergence of iterative stochastic algorithms, I general case," *Autom. Remote Control*, no. 12, pp. 83–94, 1976.
[47] C. M. Bishop, *Neural Networks for Pattern Recognition*. Oxford, U.K.: Clarendon, 1995.
[48] C. Zhang, D. Arnold, N. Ghods, A. Siranosian, and M. Krstic, "Source seeking with nonholonomic unicycle without position measurement and with tuning of forward velocity," *Syst. Control Lett.*, vol. 56, pp. 245–252, 2007.
[49] S. S. Stanković, M. S. Stanković, and D. M. Stipanović, "Decentralized parameter estimation by consensus based stochastic approximation," *IEEE Trans. Autom. Control*, vol. 56, no. 3, pp. 531–543, Mar. 2011.
[50] M. Huang and J. Manton, "Stochastic consensus seeking with noisy and directed inter-agent communication: Fixed and randomly varying topologies," *IEEE Trans. Autom. Control*, vol. 55, no. 1, pp. 235–241, Jan. 2010.

Miloš S. Stanković received the B.S. and M.S. degrees from the School of Electrical Engineering, University of Belgrade, Belgrade, Serbia, in 2002 and 2006, respectively, and the Ph.D. degree in systems and entrepreneurial engineering from the University of Illinois at Urbana-Champaign (UIUC), in 2009.

He was a Research and Teaching Assistant in the Control and Decision Group, Coordinated Science Laboratory, UIUC (2006–2009). In 2009, he joined the Royal Institute of Technology (KTH), Stockholm, Sweden, as a Postdoctoral Researcher in the Automatic Control Laboratory and the ACCESS Linnaeus Centre. His research interests include decentralized control, estimation, detection and system identification, mobile sensor networks, networked control systems, dynamic game theory, optimization, and machine learning.

Karl H. Johansson received the M.Sc. and Ph.D. degrees in electrical engineering from Lund University, Lund, Sweden.

He is Director of the ACCESS Linnaeus Centre and Professor at the School of Electrical Engineering, KTH Royal Institute of Technology, Stockholm, Sweden. He has held visiting positions at UC Berkeley (1998–2000) and California Institute of Technology (2006–07). He is a Wallenberg Scholar and has held a Senior Researcher Position with the Swedish Research Council. His research interests are in networked control systems, hybrid and embedded control, and control applications in automotive, automation, and communication systems.

Dr. Johansson is the Chair of the IFAC Technical Committee on Networked Systems. He has served on the Executive Committees of several European research projects in the area of networked embedded systems. He is on the editorial board of *IET Control Theory and Applications* and the *International Journal of Robust and Nonlinear Control*, and previously of the IEEE TRANSACTIONS ON AUTOMATIC CONTROL and *Automatica*. He was awarded an Individual Grant for the Advancement of Research Leaders from the Swedish Foundation for Strategic Research in 2005. He received the triennial Young Author Prize from IFAC in 1996 and the Peccei Award from the International Institute of System Analysis, Austria, in 1993. He received Young Researcher Awards from Scania in 1996 and from Ericsson in 1998 and 1999.

Dušan M. Stipanović received the B.S. degree in electrical engineering from the University of Belgrade, Belgrade, Serbia, in 1994, and the M.S.E.E. and Ph.D. degrees (under supervision of Prof. Dragoslav Šiljak) in electrical engineering from Santa Clara University, Santa Clara, CA, in 1996 and 2000, respectively.

He was an Adjunct Lecturer and Research Associate with the Department of Electrical Engineering, Santa Clara University, from 1998 to 2001, and a Research Associate in Prof. Claire Tomlin's Hybrid Systems Laboratory, Department of Aeronautics and Astronautics, Stanford University, Stanford, CA, from 2001 to 2004. Since 2004, he has been a faculty member in the Department of Industrial and Enterprise Systems Engineering and Control and Decision Group of the Coordinated Science Laboratory at the University of Illinois at Urbana-Champaign. His research interests include decentralized control and estimation of interconnected systems with application to control of formations of vehicles and sensor networks, stability of discontinuous dynamic systems, differential game theory, and optimization with application to multiple vehicle coordination and systems safety verification.