SIAM J. CONTROL OPTIM.
Vol. 49, No. 4, pp. 1659–1679

STOCHASTIC NASH EQUILIBRIUM SEEKING FOR GAMES WITH GENERAL NONLINEAR PAYOFFS*

SHU-JUN LIU† AND MIROSLAV KRSTIĆ‡

Abstract. We introduce a multi-input stochastic extremum seeking algorithm to solve the problem of seeking Nash equilibria for a noncooperative game whose N players seek to maximize their individual payoff functions. The payoff functions are general (not necessarily quadratic), and their forms are not known to the players. Our algorithm is a nonmodel-based approach for asymptotic attainment of the Nash equilibria. Different from classical game theory algorithms, where each player employs the knowledge of the functional form of his payoff and the knowledge of the other players' actions, a player employing our algorithm measures only his own payoff values, without knowing the functional form of his or other players' payoff functions. We prove local exponential (in probability) convergence of our algorithms. For nonquadratic payoffs, the convergence is not necessarily perfect but may be biased in proportion to the third derivatives of the payoff functions and the intensity of the stochastic perturbations used in the algorithm. We quantify the size of these residual biases. Compared to the deterministic extremum seeking with sinusoidal perturbation signals, where convergence occurs only if the players use distinct frequencies, in our algorithm each player simply employs an independent ergodic stochastic probing signal in his seeking strategy, which is realistic in noncooperative games. As a special case of an N-player noncooperative game, the problem of standard multivariable optimization (when the players' payoffs coincide) for quadratic maps is also solved using our stochastic extremum seeking algorithm.

Key words. Nash equilibrium, stochastic extremum seeking, stochastic averaging

AMS subject classifications. 60H10, 93E03, 93E15, 93E35

DOI. 10.1137/100811738

1. Introduction.

Seeking Nash equilibria in continuous games is a difficult problem (see [14]). Researchers in different fields including mathematics, computer science, economics, and system engineering have interest and need for techniques for finding Nash equilibria. Most algorithms designed to achieve convergence to Nash equilibria require modeling information for the game and assume that the players can observe the actions of the other players. The first serious algorithm perhaps is [28], in which a gradient-type algorithm is studied for convex games. Distributed iterative algorithms are designed for the computation of equilibrium in [16] for a general class of nonquadratic convex Nash games. In this algorithm, the agents do not have to know each other's cost functionals and private information as well as the parameters and subjective probability distributions adopted by the others, but they have to communicate to each other their tentative decisions during each phase of computation. A strategy known as fictitious play is one such strategy that depends on the actions of the other players so that a player can devise a best response. A dynamic version of fictitious play and gradient response is developed in [31]. In [38], a syn-

*Received by the editors October 14, 2010; accepted for publication (in revised form) April 15, 2011; published electronically July 21, 2011. This research was supported by the National Natural Science Foundation of China under grant 60704016 and the Excellent Young Teacher Foundation of Southeast University 3207011203 and by grants from the U.S. National Science Foundation and the Office of Naval Research.
http://www.siam.org/journals/sicon/49-4/81173.html
†Department of Mathematics, Southeast University, Nanjing, 210096, China (sjliu@seu.edu.cn).
‡Department of Mechanical and Aerospace Engineering, University of California, San Diego, La Jolla, CA 92093-0411 (krstic@ucsd.edu).

chronous distributed learning algorithm is designed to the coverage optimization of mobile visual sensor networks. In this algorithm, players remember their own actions and utility values from the previous two times steps, and the algorithm is shown to converge in probability to the set of restricted Nash equilibria. Other diverse engineering applications of game theory include the design of communication networks in [22, 1, 3, 29], integrated structures and controls in [27], and distributed consensus protocols in [5, 23, 30]. A comprehensive treatment of static and dynamic noncooperative game theory can be found in [4].

Extremum seeking is a nonmodel-based real-time optimization approach for dynamic problems where only limited knowledge of a system is available. Since the emergence of a proof of its stability [13], extremum seeking has been an active research area both in applications [11, 17, 21, 24, 25, 37] and in further theoretical developments [2, 7, 34, 35, 36]. Based on the extremum seeking approach with sinusoidal perturbations, in [12], Nash equilibrium seeking is studied for noncooperative games with both finitely and infinitely many players. In [33], Nash games in mobile sensor networks are solved using extremum seeking. Owing to certain advantages of stochastic perturbations over the sinusoidal ones, in [19, 20], we investigated the stochastic extremum seeking algorithm for the single perturbation input case.

In this work, a multi-input stochastic extremum seeking algorithm is developed for finding Nash equilibria in N-player noncooperative games. First, to analyze the convergence of the algorithm, a multi-input stochastic averaging theory is developed. Here multi-input means multiscaled stochastic perturbation input. Most of the existing stochastic averaging theory focuses on the systems with the single-scaled stochastic perturbation input [6, 8, 15, 18] or on two-time-scales systems with slow dynamics and fast dynamics [32, 10]. There are few results on stochastic averaging for systems with multiscaled stochastic perturbation inputs. For an N-player noncooperative game, each player employs independently stochastic extremum seeking to attain a Nash equilibrium. Similar to the deterministic case [9], the key feature of our approach is that the players are not required to know the mathematical model of their payoff function or the underlying model of the game. The players need only measure their own payoff values. It is proved that under certain conditions, the actions of players converge to a neighborhood of a Nash equilibrium. The convergence result is local in the sense that convergence to any particular Nash equilibrium is assured only for initial conditions in a set around that specific stable Nash equilibrium. Moreover, convergence to a Nash equilibrium is biased in proportion to the third derivatives of the payoff functions and is dependent on the intensity of stochastic perturbation. Compared to the deterministic case, one advantage of stochastic extremum seeking is that there is no need to choose different perturbation frequencies for each player and each player only needs to choose his own perturbation process independently, which is more realistic in a practical game with adversarial players. Finally, when all players have the same quadratic payoff, the Nash equilibrium seeking problem for an N-player noncooperative game reduces to a standard multiparameter extremum seeking problem. For this special case, we design a stochastic multiparameter extremum seeking algorithm and analyze its convergence.

The paper is organized as follows: we introduce our general problem formulation in section 2, state our algorithm and convergence results in section 3, and present the convergence proof in section 4. We provide a numerical example for a two-player game in section 5. Finally, we state our extremum seeking algorithm for multiparameter quadratic static maps in section 6 and state our stochastic averaging theory for the multi-input case in the appendix.

2. Problem formulation. Consider an N-player noncooperative game where each player wishes to maximize his payoff function of the general nonlinear form. Assume the payoff function of player i is of the form

$$J_i = h_i(u_i, u_{-i}), \tag{2.1}$$

where u_i is player i's action, the action (strategy) space is the whole space $\mathbb{R}$, $u_{-i} = [u_1, \ldots, u_{i-1}, u_{i+1}, \ldots, u_N]$ represents the actions of the other players, $h_i : \mathbb{R}^N \to \mathbb{R}$ is smooth, and $i \in \{1, \ldots, N\}$.

Our algorithm is based on the following assumptions.

ASSUMPTION 2.1. *There exists at least one, possibly multiple, isolated stable Nash equilibrium* $u^* = [u_1^*, \ldots, u_N^*]$ *such that*

$$\frac{\partial h_i}{\partial u_i}(u^*) = 0, \tag{2.2}$$

$$\frac{\partial^2 h_i}{\partial u_i^2}(u^*) < 0 \tag{2.3}$$

for all $i \in \{1, \ldots, N\}$.

ASSUMPTION 2.2. *The matrix*

$$\Xi = \begin{bmatrix} \frac{\partial^2 h_1(u^*)}{\partial u_1^2} & \frac{\partial^2 h_1(u^*)}{\partial u_1 \partial u_2} & \cdots & \frac{\partial^2 h_1(u^*)}{\partial u_1 \partial u_N} \\ \frac{\partial^2 h_2(u^*)}{\partial u_1 \partial u_2} & \frac{\partial^2 h_2(u^*)}{\partial u_2^2} & \cdots & \frac{\partial^2 h_2(u^*)}{\partial u_2 \partial u_N} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 h_N(u^*)}{\partial u_1 \partial u_N} & \frac{\partial^2 h_N(u^*)}{\partial u_2 \partial u_N} & \cdots & \frac{\partial^2 h_N(u^*)}{\partial u_N^2} \end{bmatrix} \tag{2.4}$$

is strictly diagonally dominant and hence, nonsingular.

By Assumptions 2.1 and 2.2, Ξ is Hurwitz.

In our scheme, player i has no knowledge of other players' payoff h_j $(j \neq i)$ and actions u_j $(j \neq i)$. He can measure only his own payoff h_i. Our objective is to design a stochastic extremum seeking algorithm for each player to approximate Nash equilibrium.

3. Stochastic Nash equilibrium seeking algorithm. In our algorithm, each player independently employs a stochastic seeking strategy to attain the stable Nash equilibrium of the game. Player i implements the following strategy:

$$u_i(t) = \hat{u}_i(t) + a_i f_i(\eta_i(t)), \tag{3.1}$$

$$\frac{d\hat{u}_i(t)}{dt} = k_i a_i f_i(\eta_i(t)) J_i(t), \tag{3.2}$$

where for any $i = 1, \ldots, N$, $a_i > 0$ is the perturbation amplitude, $k_i > 0$ is the adaptive gain, $J_i(t)$ is the measured payoff value for player i, and f_i is a bounded smooth function that player i chooses, e.g., a sine function. $\eta_i(t), i = 1, \ldots, N$, are independent time homogeneous continuous Markov ergodic processes chosen by player i, e.g., the Ornstein–Uhlenbeck (OU) process

$$\eta_i = \frac{\sqrt{\varepsilon_i} q_i}{\varepsilon_i s + 1}[\dot{W}_i] \quad \text{or} \quad \varepsilon_i d\eta_i(t) = -\eta_i(t)dt + \sqrt{\varepsilon_i} q_i dW_i(t), \tag{3.3}$$

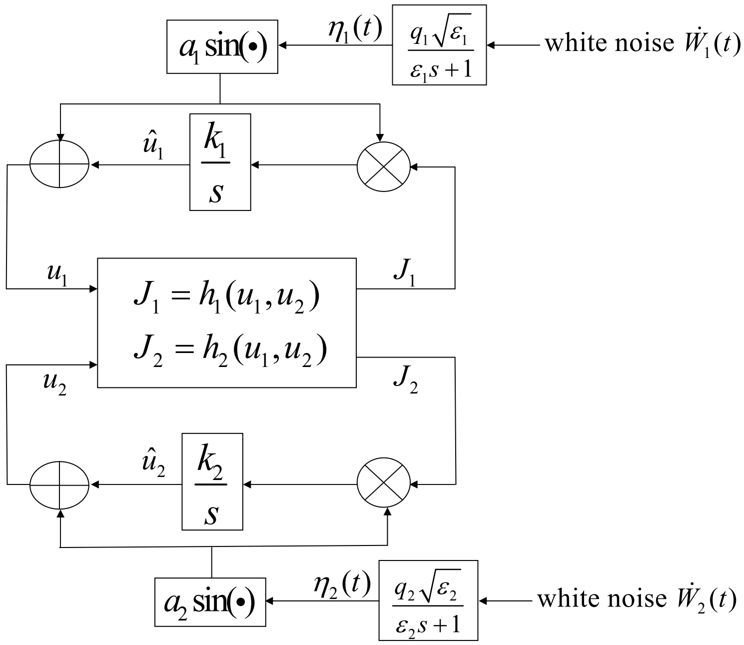

FIG. 1. *Stochastic extremum seeking scheme for a two-player noncooperative game.*

$q_i > 0$, ε_i are small parameters satisfying $0 < \max_i \varepsilon_i < \varepsilon_0$ for fixed $\varepsilon_0 > 0$, and $W_i(t)$, $i = 1, \ldots, N$, are independent 1-dimensional standard Brownian motion on a complete probability space $(\Omega, \mathcal{F}, P)$ with the sample space Ω, σ-field $\mathcal{F}$, and probability measure P.

System (3.2) is given by an ordinary differential equation with stochastic perturbation (see [8]), namely, a stochastic ordinary differential equation, and its solution can be defined for each sample path of the perturbation process $(\eta_i(t), t \geq 0)$, which is given by Ito stochastic differential equation (3.3).

Figure 1 depicts a noncooperative game played by two players implementing the stochastic extremum seeking strategy (3.1)–(3.2) to attain a Nash equilibrium.

To analyze the convergence of the algorithm, we denote the error relative to the Nash equilibrium as

$$\tilde{u}_i(t) = \hat{u}_i(t) - u_i^*. \tag{3.4}$$

Then, we obtain an error system as

$$\frac{d\tilde{u}_i(t)}{dt} = k_i \rho_i^{(1)}(t) h_i\left(u_i^* + \tilde{u}_i + \rho_i^{(1)}(t), u_{-i}^* + \tilde{u}_{-i} + \rho_{-i}^{(1)}(t)\right), \tag{3.5}$$

where $\rho_i^{(1)}(t) = a_i f_i(\eta_i(t))$, $\rho_{-i}^{(1)}(t) = [a_1 f_1(\eta_1(t)), \ldots, a_{i-1} f_{i-1}(\eta_{i-1}(t)), a_{i+1} f_{i+1}(\eta_{i+1}(t)), \ldots, a_N f_N(\eta_N(t))]$, $\tilde{u}_{-i}^* = [\tilde{u}_1^*, \ldots, \tilde{u}_{i-1}^*, \tilde{u}_{i+1}^*, \ldots, \tilde{u}_N^*]$, and $\tilde{u}_{-i} = [\tilde{u}_1, \ldots, \tilde{u}_{i-1}, \tilde{u}_{i+1}, \ldots, \tilde{u}_N]$.

If the players choose $f_i(x) = \sin x$ for all $i = 1, \ldots, N$, and η_i as independent OU processes (3.3), we have the following convergence result.

THEOREM 3.1. *Consider the error system* (3.5) *for an N-player game under Assumptions* 2.1 *and* 2.2. *Then there exists a constant $a^* > 0$ such that for*

$\max_{1\le i\le N} a_i \in (0, a^*)$ *there exist constants* $r > 0$, $c > 0$, $\gamma > 0$ *and a function* $T(\varepsilon_1) : (0, \varepsilon_0) \to \mathbb{N}$ *such that for any initial condition* $|\Lambda^{\varepsilon_1}(0)| < r$ *and any* $\delta > 0$,

$$\lim_{\varepsilon_1\to 0} \inf \left\{ t \ge 0 : |\Lambda^{\varepsilon_1}(t)| > c|\Lambda^{\varepsilon_1}(0)|e^{-\gamma t} + \delta + O(\max_i a_i^3) \right\} = \infty \ a.s. \tag{3.6}$$

and

$$\lim_{\varepsilon_1\to 0} P \left\{ |\Lambda^{\varepsilon_1}(t)| \le c|\Lambda^{\varepsilon_1}(0)|e^{-\gamma t} + \delta + O(\max_i a_i^3)\ \forall t \in [0, T(\varepsilon_1)] \right\} = 1 \tag{3.7}$$

with

$$\lim_{\varepsilon_1\to 0} T(\varepsilon_1) = \infty,$$

where

$$\Lambda^{\varepsilon_1}(t) = \left[\tilde{u}_1(t) - \sum_{j=1}^N d_{jj}^1 a_j^2, \ldots, \tilde{u}_N(t) - \sum_{j=1}^N d_{jj}^N a_j^2 \right], \tag{3.8}$$

$$\begin{bmatrix} d_{jj}^1 \\ \vdots \\ d_{jj}^{j-1} \\ d_{jj}^j \\ d_{jj}^{j+1} \\ \vdots \\ d_{jj}^N \end{bmatrix} = -\Xi^{-1} \begin{bmatrix} \frac{1}{2}G_0(q_j)\frac{\partial^3 h_1}{\partial u_1 \partial u_j^2}(u^*) \\ \vdots \\ \frac{1}{2}G_0(q_j)\frac{\partial^3 h_{j-1}}{\partial u_{j-1} \partial u_j^2}(u^*) \\ \frac{1}{6}\frac{G_1(q_j)}{G_0(q_j)}\frac{\partial^3 h_j}{\partial u_j^3}(u^*) \\ \frac{1}{2}G_0(q_j)\frac{\partial^3 h_{j+1}}{\partial u_j^2 \partial u_{j+1}}(u^*) \\ \vdots \\ \frac{1}{2}G_0(q_j)\frac{\partial^3 h_N}{\partial u_j^2 \partial u_N}(u^*) \end{bmatrix}, \tag{3.9}$$

and $G_0(q_j) = \frac{1}{2}(1 - e^{-q_j^2})$, $G_1(q_j) = \frac{3}{8} - \frac{1}{2}e^{-q_j^2} + \frac{1}{8}e^{-4q_j^2} = \frac{1}{8}(1 - e^{-q_j^2})^2(e^{-2q_j^2} + 2e^{-q_j^2} + 3)$.

Several remarks are needed in order to properly interpret Theorem 3.1. From (3.6) and the fact $|\Lambda^{\varepsilon_1}(t)| \ge \max_i |\tilde{u}_i(t) - \sum_{j=1}^N d_{jj}^i a_j^2|$, we obtain

$$\lim_{\varepsilon_1\to 0} \inf \left\{ t \ge 0 : \max_i \left\{ \left| \tilde{u}_i(t) - \sum_{j=1}^N d_{jj}^i a_j^2 \right| \right\} > c|\Lambda^{\varepsilon_1}(0)|e^{-\gamma t} + \delta + O(\max_i a_i^3) \right\} = \infty \text{ a.s.}$$

By taking all the a_i's small, $\max_i |\tilde{u}_i(t)|$ can be made arbitrarily small as $t \to \infty$.

The bias terms $\sum_{j=1}^N d_{jj}^i a_j^2$ defined by (3.9) appear complicated but have a simple physical interpretation. When the game's payoff functions are not quadratic (not symmetric), the extremum seeking algorithms, which employ zero-mean (symmetric) perturbations, will produce a bias. According to the formula (3.9), the bias depends on the third derivatives of the payoff functions, namely, on the level of asymmetry in the payoff surfaces at the Nash equilibrium. In the trivial case of a single player the interpretation is easy—extremum seeking settles on the flatter (more favorable) side of an asymmetric peak. In the case of multiple players the interpretation is more difficult, as each player contributes both to his own bias and to the other players' biases. Though difficult to intuitively interpret in the multiplayer case, the formula (3.9) is useful as it quantifies the biases.

The estimate of the region of attraction r can be conservatively taken as independent of the a_i's, for a_i's chosen sufficiently small. This fact can be seen only by going through the proof of the averaging theorem for the specific system (3.5). Hence, r is larger than the bias terms, which means that for small a_i's the algorithm reduces the distance to the Nash equilibrium for all initial conditions except for those within an $O(\max_i a_i^2)$ to the Nash equilibrium.

On the other hand, the convergence rate γ cannot be taken independently of the a_i's, because the a_i's appear as factors on the entire right-hand side of (3.5). However, by letting the k_i's increase as the a_i's decrease, independence of γ from the a_i's can be ensured.

In the rare case where the error system (3.5) may be globally Lipschitz, we obtain global convergence using the global averaging theorem in [18].

4. Proof of the algorithm convergence. We apply the multi-input stochastic averaging theory presented in the appendix to analyze the error system (3.5). First, we calculate the average system of (3.5).

Define $\chi_i(t) = \eta_i(\varepsilon_i t)$ and $B_i(t) = \frac{1}{\sqrt{\varepsilon_i}} W_i(\varepsilon_i t)$. Then by (3.3) we have

$$d\chi_i(t) = -\chi_i(t)dt + q_i dB_i(t), \tag{4.1}$$

where $[B_1(t), \ldots, B_N(t)]^T$ is an N-dimensional standard Brownian motion on the space $(\Omega, \mathcal{F}, P)$.

Thus we can rewrite the error system (3.5) as

$$\frac{d\tilde{u}_i(t)}{dt} = k_i \rho_i^{(2)}(t/\varepsilon_i) h_i \left(u_i^* + \tilde{u}_i + \rho_i^{(2)}(t/\varepsilon_i), u_{-i}^* + \tilde{u}_{-i} + \rho_{-i}^{(2)}(t/\varepsilon_{-i}) \right), \tag{4.2}$$

where $\rho_i^{(2)}(t) = a_i \sin(\chi_i(t))$, $\rho_{-i}^{(2)}(t/\varepsilon_{-i}) = [a_1 \sin(\chi_1(t/\varepsilon_1)), \ldots, a_{i-1} \sin(\chi_{i-1}(t/\varepsilon_{i-1})), a_{i+1} \sin(\chi_{i+1}(t/\varepsilon_{i+1})), \ldots, a_N \sin(\chi_N(t/\varepsilon_N))]$.

Denote

$$\varepsilon_i = \frac{\varepsilon_1}{c_i}, \quad i = 2, \ldots, N \tag{4.3}$$

for some positive real constants c_i and consider the change of variable

$$Z_1(t) = \chi_1(t),\ Z_2(t) = \chi_2(c_2 t), \ldots, Z_N(t) = \chi(c_N t). \tag{4.4}$$

Then the error system (4.2) can be transformed as one with single small parameter ε_1:

$$\frac{d\tilde{u}_i(t)}{dt} = k_i \rho_i^{(3)}(t/\varepsilon_1) h_i \left(u_i^* + \tilde{u}_i + \rho_i^{(3)}(t/\varepsilon_1), u_{-i}^* + \tilde{u}_{-i} + \rho_{-i}^{(3)}(t/\varepsilon_1) \right), \tag{4.5}$$

where $\rho_i^{(3)}(t) = a_i \sin(Z_i(t))$, $\rho_{-i}^{(3)}(t/\varepsilon_1) = [a_1 \sin(Z_1(t/\varepsilon_1)), \ldots, a_{i-1} \sin(Z_{i-1}(t/\varepsilon_1)), a_{i+1} \sin(Z_{i+1}(t/\varepsilon_1)), \ldots, a_N \sin(Z_N(t/\varepsilon_1))]$.

For $(\chi_i(t), t \geq 0)$ that is ergodic and has invariant distribution $\mu_i(dx_i) = \frac{1}{\sqrt{\pi}q_i} e^{-\frac{x_i^2}{q_i^2}} dx_i$ (see [26]), by Lemma A.2, the vector-valued process $[Z_1(t), \dots, Z_N(t)]^T$ is also ergodic with invariant distribution $(\mu_1 \times \cdots \times \mu_N)$. Thus by (A.4), we have the average error system

(4.6)
$$\frac{d\tilde{u}_i^{ave}(t)}{dt} = k_i a_i \int_{\mathbb{R}^N} \sin(x_i) h_i \left(u_i^* + \tilde{u}_i^{ave} + a_i \sin(x_i), u_{-i}^* + \tilde{u}_{-i}^{ave} + a_{-i} \sin(x_{-i})\right) \mu_1(dx_1) \times \cdots \times \mu_N(dx_N),$$

where $a_{-i} \sin(x_{-i}) = [a_1 \sin(x_1), \dots, a_{i-1} \sin(x_{i-1}),\ a_{i+1} \sin(x_{i+1}), \dots, a_N \sin(x_N)]$, and μ_i is the invariant distribution of the process $(\chi_i(t), t \geq 0)$ or $(Z_i(t), t \geq 0)$.

The equilibrium $\tilde{u}^e = [\tilde{u}_1^e, \dots, \tilde{u}_N^e]$ of (4.6) satisfies

$$0 = \int_{\mathbb{R}^N} \sin(x_i) h_i \left(u_i^* + \tilde{u}_i^e + a_i \sin(x_i), u_{-i}^* + \tilde{u}_{-i}^e + a_{-i} \sin(x_{-i})\right) \mu_1(dx_1) \times \cdots \times \mu_N(dx_N) \tag{4.7}$$

for all $i = \{1, \dots, N\}$.

To calculate the equilibrium of the average error system and analyze its stability, we postulate that $\tilde{u}^e$ has the form

$$\tilde{u}_i^e = \sum_{j=1}^{N} b_j^i a_j + \sum_{j=1}^{N} \sum_{k \geq j}^{N} d_{jk}^i a_j a_k + O(\max_i a_i^3). \tag{4.8}$$

By expanding h_i about u^* in (4.7) and substituting (4.8), the unknown coefficients b_j^i and d_{jk}^i can be determined.

The Taylor series expansion of h_i about u^* in (4.7) for an N-player game is

$$h_i(u^* + v_i, u_{-i}^* + v_{-i}) = \sum_{n_1=0}^{\infty} \cdots \sum_{n_N=0}^{\infty} \frac{v_1^{n_1} \cdots v_N^{n_N}}{n_1! \cdots n_N!} \left(\frac{\partial^{n_1+\cdots+n_N} h_i}{\partial u_1^{n_1} \cdots \partial u_N^{n_N}}\right)(u^*), \tag{4.9}$$

where $v_i = \tilde{u}_i^e + a_i \sin(x_i)$ and $v_{-i} = \tilde{u}_{-i}^e + a_{-i} \sin(x_{-i})$. Although for any $i = 1, \dots, N$, h_i may not have its Taylor series expansion only by its smoothness, here we give just the form of Taylor series expansion. In fact, we need only its third order Taylor formula.

Since the invariant distribution $\mu_i(dx_i)$ of OU process $(\chi_i(t), t \geq 0)$ is $\frac{1}{\sqrt{\pi}q_i} e^{-\frac{x_i^2}{q_i}} dx_i$,

$$\int_{\mathbb{R}} \sin^{2k+1}(x_i) \mu_i(dx_i) = \int_{-\infty}^{+\infty} \sin^{2k+1}(x_i) \frac{1}{\sqrt{\pi}q_i} e^{-\frac{x_i^2}{q_i^2}} dx_i = 0, \quad k = 0, 1, 2, \dots, \tag{4.10}$$

$$\int_{\mathbb{R}} \sin^2(x_i) \mu_i(dx_i) = \int_{-\infty}^{+\infty} \sin^2(x_i) \frac{1}{\sqrt{\pi}q_i} e^{-\frac{x_i^2}{q_i^2}} dx_i = \frac{1}{2}(1 - e^{-q_i^2}) = G_0(q_i), \tag{4.11}$$

$$(4.12)\quad \int_{\mathbb{R}} \sin^4(x_i)\mu_i(dx_i) = \int_{-\infty}^{+\infty} \sin^4(x_i)\frac{1}{\sqrt{\pi}q_i}e^{-\frac{x_i^2}{q_i^2}}dx_i$$
$$= \frac{3}{8} - \frac{1}{2}e^{-q_i^2} + \frac{1}{8}e^{-4q_i^2} = G_1(q_i),$$

$$(4.13)\quad \int_{\mathbb{R}^2} \sin(x_i)\sin(x_j)\mu_i(dx_i)\times\mu_j(dx_j)$$
$$= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \sin(x_i)\sin(x_j)\frac{1}{\sqrt{\pi}q_i}e^{-\frac{x_i^2}{q_i^2}}\frac{1}{\sqrt{\pi}q_j}e^{-\frac{x_j^2}{q_j^2}}dx_idx_j = 0,$$

$$(4.14)\quad \int_{\mathbb{R}^2} \sin^2(x_i)\sin(x_j)\mu_i(dx_i)\times\mu_j(dx_j)$$
$$= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \sin^2(x_i)\sin(x_j)\frac{1}{\sqrt{\pi}q_i}e^{-\frac{x_i^2}{q_i^2}}\frac{1}{\sqrt{\pi}q_j}e^{-\frac{x_j^2}{q_j^2}}dx_idx_j = 0,$$

$$(4.15)\quad \int_{\mathbb{R}^2} \sin^3(x_i)\sin(x_j)\mu_i(dx_i)\times\mu_j(dx_j)$$
$$= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \sin^3(x_i)\sin(x_j)\frac{1}{\sqrt{\pi}q_i}e^{-\frac{x_i^2}{q_i^2}}\frac{1}{\sqrt{\pi}q_j}e^{-\frac{x_j^2}{q_j^2}}dx_idx_j = 0,$$

$$(4.16)\quad \int_{\mathbb{R}^2} \sin^2(x_i)\sin^2(x_j)\mu_i(dx_i)\times\mu_j(dx_j)$$
$$= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \sin^2(x_i)\sin^2(x_j)\frac{1}{\sqrt{\pi}q_i}e^{-\frac{x_i^2}{q_i^2}}\frac{1}{\sqrt{\pi}q_j}e^{-\frac{x_j^2}{q_j^2}}dx_idx_j$$
$$= \frac{1}{4}(1-e^{-q_i^2})(1-e^{-q_j^2}) \triangleq G_2(q_i,q_j),$$

$$(4.17)\quad \int_{\mathbb{R}^3} \sin(x_i)\sin(x_j)\sin(x_k)\mu_i(dx_i)\times\mu_j(dx_j)\times\mu_k(dx_k)$$
$$= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \sin(x_i)\sin(x_j)\sin(x_k)\frac{1}{\sqrt{\pi}q_i}e^{-\frac{x_i^2}{q_i^2}}\frac{1}{\sqrt{\pi}q_j}e^{-\frac{x_j^2}{q_j^2}}\frac{1}{\sqrt{\pi}q_k}e^{-\frac{x_k^2}{q_k^2}}$$
$$\times dx_idx_jdx_k = 0,$$

$$(4.18)\quad \int_{\mathbb{R}^3} \sin(x_i)\sin^2(x_j)\sin(x_k)\mu_i(dx_i)\times\mu_j(dx_j)\times\mu_k(dx_k)$$
$$= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \sin(x_i)\sin^2(x_j)\sin(x_k)\frac{1}{\sqrt{\pi}q_i}e^{-\frac{x_i^2}{q_i^2}}\frac{1}{\sqrt{\pi}q_j}e^{-\frac{x_j^2}{q_j^2}}\frac{1}{\sqrt{\pi}q_k}e^{-\frac{x_k^2}{q_k^2}}$$
$$\times dx_idx_jdx_k = 0.$$

Based on the above calculations, substituting (4.9) into (4.7) and computing the average of each term gives

$$(4.19)\qquad 0 = a_i^2G_0(q_i)\tilde{u}_i^e\frac{\partial^2 h_i}{\partial u_i^2}(u^*) + a_i^2G_0(q_i)\sum_{j\neq i}^{N}\tilde{u}_j^e\frac{\partial^2 h_i}{\partial u_i\partial u_j}(u^*)$$
$$+ \left(\frac{a_i^2}{2}G_0(q_i)(\tilde{u}_i^e)^2 + \frac{a_i^4}{6}G_1(q_i)\right)\frac{\partial^3 h_i}{\partial u_i^3}(u^*)$$

$$+\ a_i^2 G_0(q_i)\tilde{u}_i^e \sum_{j\neq i}^{N} \tilde{u}_j^e \frac{\partial^3 h_i}{\partial u_i^2 \partial u_j}(u^*) + \sum_{j\neq i}^{N}\left(\frac{a_i^2}{2}G_0(q_i)(\tilde{u}_j^e)^2 + \frac{a_i^2 a_j^2}{2}G_2(q_i,q_j)\right)\frac{\partial^3 h_i}{\partial u_i \partial u_j^2}(u^*) + \sum_{j\neq i}^{N}\sum_{k>j,k\neq i}^{N} a_i^2 G_0(q_i)\tilde{u}_j^e\tilde{u}_k^e \frac{\partial^3 h_i}{\partial u_i\partial u_j\partial u_k}(u^*) + O(\max_i a_i^5),$$

or, equivalently,

$$0 = \tilde{u}_i^e \frac{\partial^2 h_i}{\partial u_i^2}(u^*) + \sum_{j\neq i}^{N}\tilde{u}_j^e \frac{\partial^2 h_i}{\partial u_i \partial u_j}(u^*) + \left(\frac{1}{2}(\tilde{u}_i^e)^2 + \frac{a_i^2}{6}\frac{G_1(q_i)}{G_0(q_i)}\right)\frac{\partial^3 h_i}{\partial u_i^3}(u^*) + \tilde{u}_i^e \sum_{j\neq i}^{N}\tilde{u}_j^e \frac{\partial^3 h_i}{\partial u_i^2 \partial u_j}(u^*) + \sum_{j\neq i}^{N}\left(\frac{1}{2}(\tilde{u}_j^e)^2 + \frac{a_j^2}{2}G_0(q_j)\right)\frac{\partial^3 h_i}{\partial u_i\partial u_j^2}(u^*) + \sum_{j\neq i}^{N}\sum_{k>j,k\neq i}^{N}\tilde{u}_j^e\tilde{u}_k^e \frac{\partial^3 h_i}{\partial u_i \partial u_j \partial u_k}(u^*) + O(\max_i a_i^3). \tag{4.20}$$

Substituting (4.8) into (4.19) and matching first order powers of a_i gives

$$\begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} = \Xi \begin{bmatrix} b_i^1 \\ \vdots \\ b_i^N \end{bmatrix}, \quad i = 1,\dots,N, \tag{4.21}$$

which implies that $b_j^i = 0$ for all i, j since Ξ is nonsingular by Assumption 2.2. Similarly, matching second order terms $a_j a_k$ $(j > k)$ and a_j^2 of a_j and substituting $b_j^i = 0$ to simplify the resulting expressions yields

$$\begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} = \Xi \begin{bmatrix} d_{jk}^1 \\ \vdots \\ d_{jk}^N \end{bmatrix}, \quad j = 1,\dots,N,\ j > k, \tag{4.22}$$

and

$$\begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} = \left(\Xi \begin{bmatrix} d_{jj}^1 \\ \vdots \\ d_{jj}^N \end{bmatrix} + \begin{bmatrix} \frac{1}{2}G_0(q_j)\frac{\partial^3 h_1}{\partial u_1 \partial u_j^2}(u^*) \\ \vdots \\ \frac{1}{2}G_0(q_j)\frac{\partial^3 h_{j-1}}{\partial u_{j-1}\partial u_j^2}(u^*) \\ \frac{1}{6}\frac{G_1(q_j)}{G_0(q_j)}\frac{\partial^3 h_j}{\partial u_j^3}(u^*) \\ \frac{1}{2}G_0(q_j))\frac{\partial^3 h_{j+1}}{\partial u_j^2 \partial u_{j+1}}(u^*) \\ \vdots \\ \frac{1}{2}G_0(q_j)\frac{\partial^3 h_N}{\partial u_j^2 \partial u_N}(u^*) \end{bmatrix} \right). \tag{4.23}$$

Thus, $d_{jk}^i = 0$ for all i, j, k when $j \neq k$, and d_{jj}^i is given by (3.9).

Therefore, by (4.8), the equilibrium of the average error system (4.6) is

$$\tilde{u}_i^e = \sum_{j=1}^{N} d_{jj}^i a_j^2 + O(\max_i a_i^3). \tag{4.24}$$

By the dominated convergence theorem, we obtain that the Jacobian $\Psi^{ave} = (\psi_{ij})_{N\times N}$ of the average error system (4.6) at $\tilde{u}^e$ has elements given by

$$\psi_{ij} = k_i \int_{\mathbb{R}^N} a_i \sin(x_i) \frac{\partial h_i}{\partial u_j} \big(u_i^* + \tilde{u}_i^e + a_i \sin(x_i), u_{-i}^* + \tilde{u}_{-i}^e \tag{4.25}$$

$$+ a_{-i} \sin(x_{-i})\big)\, \mu_1(dx_1) \times \cdots \times \mu_N(dx_N) \tag{4.26}$$

$$= k_i a_i^2 G_0(q_i) \frac{\partial^2 h_i}{\partial u_i \partial u_j}(u^*) + O\big(\max_i a_i^3\big)$$

and is Hurwitz by Assumptions 2.1 and 2.2 for sufficiently small a_i, which implies that the equilibrium (4.24) of the average error system (4.6) is locally exponentially stable. By the multi-input averaging theorem in the appendix, the theorem is proved.

5. Numerical example.

We consider two players with payoff functions

$$J_1 = -u_1^3 + 2u_1u_2 + u_1^2 - \frac{3}{4}u_1, \tag{5.1}$$

$$J_2 = 2u_1^2 u_2 - u_2^2. \tag{5.2}$$

Since J_1 is not globally concave in u_1, we restrict the action space to $\mathcal{A} = \{u_1 \geq 1/3, u_2 \geq 1/6\}$ in order to avoid the existence of maximizing actions at infinity or Nash equilibria at the boundary of the action space. (However, we do not restrict the extremum seeking algorithm to $\mathcal{A}$. Such a restriction can be imposed using parameter projection but would complicate our exposition considerably.)

The game (J_1, J_2) yields two Nash equilibria: $(u_1^{*1}, u_2^{*1}) = (0.5, 0.25)$ and $(u_1^{*2}, u_2^{*2}) = (1.5, 2.25)$. The corresponding matrices are

$$\Xi_1 = \begin{bmatrix} -1 & 2 \\ 2 & -2 \end{bmatrix} \quad \text{and} \quad \Xi_2 = \begin{bmatrix} -7 & 2 \\ 6 & -2 \end{bmatrix},$$

where Ξ_1 is nonsingular but not Hurwitz, while Ξ_2 is nonsingular and Hurwitz, and both matrices are not diagonally dominant. From the proof of the algorithm convergence, we know that diagonal dominance is only a sufficient condition for Ξ to be nonsingular and is not required in general.

The average error system for this game is

$$\frac{d\tilde{u}_1^{ave}(t)}{dt} = k_1 a_1^2 G_0(q_1)(-3\tilde{u}_1^{ave^2} - 6u_1^*\tilde{u}_1^{ave} + 2\tilde{u}_2^{ave} + 2\tilde{u}_1^{ave}) - k_1 a_1^4 G_1(q_1), \tag{5.3}$$

$$\frac{d\tilde{u}_2^{ave}(t)}{dt} = k_2 a_2^2 G_0(q_2)(-2\tilde{u}_2^{ave} + 2\tilde{u}_1^{ave^2} + 4u_1^*\tilde{u}_1^{ave}) + 2k_2 a_1^2 a_2^2 G_2(q_1, q_2), \tag{5.4}$$

where u_1^* can be u_1^{*1} or u_1^{*2}. The equilibria $(\tilde{u}_1^e, \tilde{u}_2^e)$ of this average system are

$$\tilde{u}_1^e = 1 - u_1^* \pm \sqrt{(1-u_1^*)^2 - a_1^2 \left(\frac{G_1(q_1)}{G_0(q_1)} - 2G_0(q_1)\right)}, \tag{5.5}$$

$$\tilde{u}_2^e = 2 - 2u_1^* \pm 2\sqrt{(1-u_1^*)^2 - a_1^2 \left(\frac{G_1(q_1)}{G_0(q_1)} - 2G_0(q_1)\right)} - a_1^2 \frac{G_1(q_1)}{G_0(q_1)} + 3a_1^2 G_0(q_1), \tag{5.6}$$

and their postulated form is

$$\tilde{u}_1^{e,p} = \frac{1}{2(1-u_1^*)} \left(\frac{G_1(q_1)}{G_0(q_1)} - 2G_0(q_1)\right) a_1^2 + O(\max_i a_i^3), \tag{5.7}$$

$$\tilde{u}_2^{e,p} = \left(\frac{u_1^*}{1-u_1^*} \frac{G_1(q_1)}{G_0(q_1)} + \frac{1-3u_1^*}{1-u_1^*} G_0(q_1)\right) a_1^2 + O(\max_i a_i^3). \tag{5.8}$$

The corresponding Jacobian matrices are

$$\Psi^{ave} = \begin{bmatrix} (-6\tilde{u}_1^e - 6u_1^* + 2)\gamma_1 & 2\gamma_1 \\ (2\tilde{u}_1^e + 4u_1^*)\gamma_2 & -2\gamma_2 \end{bmatrix}, \tag{5.9}$$

where $\gamma_i = k_i a_i^2 G_0(q_i)$, $i = 1, 2$, and their characteristic equation is given by $\lambda^2 + \alpha_1 \lambda + \alpha_2 = 0$, where

$$\alpha_1 = (6\tilde{u}_1^e + 6u_1^* - 2)\gamma_1 + 2\gamma_2, \tag{5.10}$$

$$\alpha_2 = (2\tilde{u}_1^e + u_1^* - 1)4\gamma_1\gamma_2. \tag{5.11}$$

Thus Ψ^{ave} is Hurwitz if and only if α_1 and α_2 are positive. For sufficiently small a_1, which makes $\tilde{u}^e \approx (0,0)$, α_1 and α_2 are positive for $u_1^* = 1.5$, but for $u_1^* = 0.5$, α_2 is not positive, which is reasonable because Ξ_1 is not Hurwitz, but Ξ_2 is Hurwitz. Thus, $(u_1^{*1}, u_2^{*1}) = (0.5, 0.25)$ is an unstable Nash equilibrium, but $(u_1^{*2}, u_2^{*2}) = (1.5, 2.25)$ is a stable Nash equilibrium. We employ the stochastic multi-input extremum seeking algorithm given in section 3 to attain this stable equilibrium.

The top picture in Figure 2 depicts the evolution of the game in the $\tilde{u}$ plane, initialized at the point $(u_1(0), u_2(0)) = (0, 3)$, i.e., at $(\tilde{u}_1(0), \tilde{u}_2(0)) = (-1.5, 0.75)$. Note that the initial condition is outside of $\mathcal{A}$. This illustrates the point that the region of attraction of the stable Nash equilibrium under the extremum seeking algorithm is not a subset of $\mathcal{A}$ but a large subset of $\mathbb{R}^2$. The parameters are chosen as $k_1 = 14$, $k_2 = 6$, $a_1 = 0.2$, $a_2 = 0.02$, $\varepsilon_1 = 0.01$, $\varepsilon_2 = 0.8$. The bottom two pictures depict the two players' actions in stochastically seeking the Nash equilibrium $(u_1^*, u_2^*) = (1.5, 2.25)$. From Figure 2, the actions of the players converge to a small neighborhood of the stable Nash equilibrium.

In the algorithm, bounded smooth functions f_i and the excitation processes $(\eta_i(t), t \geq 0)$, $i = 1, \ldots, N$, can be chosen in other forms. We can replace the bounded excitation signal $\sin(\eta_i(t)) = \sin(\chi_i(t/\varepsilon_i))$ with the signal $H^T(\check{\eta}_i(t/\varepsilon_i))$, where $\check{\eta}_i(t) = [\cos(W_i(t)), \sin(W_i(t))]^T$ is a Brownian motion on the unit circle (see [19]), and $G = [g_1, g_2]^T$ is a constant vector.

Figure 3 depicts the evolution of the game in the $\tilde{u}$ plane for games with Brownian motion on the unit circle as perturbation. The initial conditions are the same with the

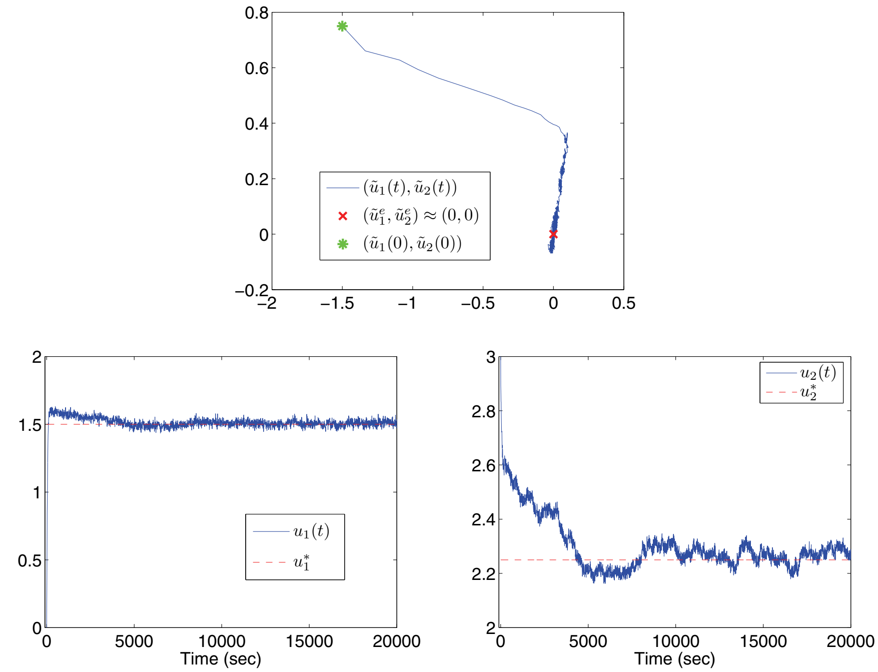

FIG. 2. *Stochastic Nash equilibrium seeking with an OU process perturbation. Top: evolution of the game in the $\tilde{u}$ plane. Bottom: two players' actions.*

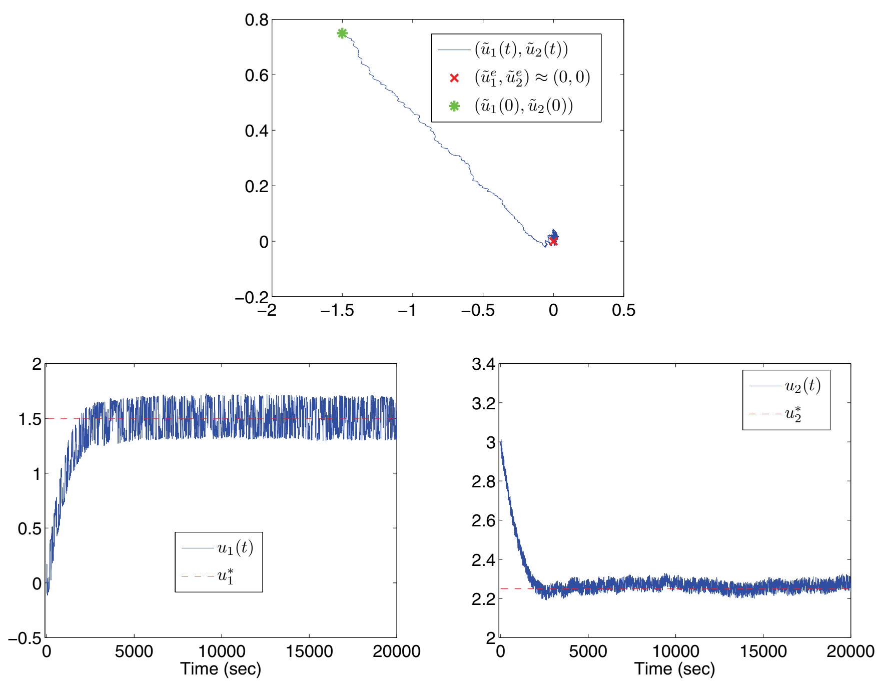

FIG. 3. *Stochastic Nash equilibrium seeking with Brownian motion on the unit circle as perturbation. Top: evolution of the game in the $\tilde{u}$ plane. Bottom: two players' actions.*

case of the OU process perturbation. The parameters are chosen as $k_1 = 5$, $k_2 = 9$, $a_1 = 0.2$, $a_2 = 0.04$, $\varepsilon_1 = 0.02$, $\varepsilon_2 = 0.02$. From Figure 3, the actions of the players also converge to a small neighborhood of the stable Nash equilibrium.

In these two simulations, a possibly different high-pass filter for each player's measurement on the payoff is used to improve the asymptotic performance but is not essential for achieving stability (see [36]), which can also be seen from the following stochastic multiparameter extremum seeking algorithm.

6. Multiparameter extremum seeking for static maps.

6.1. Stochastic multiparameter extremum seeking algorithm. Let $f(\theta)$ be a function of the form

$$f(\theta) = f^* + (\theta - \theta^*)^T P(\theta - \theta^*), \tag{6.1}$$

where $P = (p_{ij})_{l\times l} \in \mathbb{R}^{l\times l}$ is an unknown symmetric matrix, f^* is an unknown constant, $\theta = [\theta_1, \ldots, \theta_l]^T$, and $\theta^* = [\theta_1^*, \ldots, \theta_l^*]^T$. Any $C^2(\mathbb{R}^l)$ function $f(\theta)$ with an extremum at $\theta = \theta^*$ and with $\nabla^2 f \neq \mathbf{0}$ can be locally approximated by (6.1). Without loss of generality, we assume that the matrix P is positive definite.

The objective is to design an algorithm to make $|\theta - \theta^*|$ as small as possible, so that the output $y = f(\theta)$ is driven to its minimum f^*. This problem is a special case of a finite and multiplayer noncooperative game: all the players' payoffs are the same with a quadratic static map, and the corresponding matrix Ξ is $2P$. Here, we do not assume that Ξ, i.e., $2P$, is strictly diagonally dominant. For the static map, we can prove that the condition of strictly diagonal dominance is not necessary.

Denote $\hat{\theta}_j(t)$ as the estimate of the unknown optimal input θ_j^* and let

$$\tilde{\theta}_j(t) = \theta_j^* - \hat{\theta}_j(t) \tag{6.2}$$

denote the estimation error.

We use stochastic perturbation to develop a gradient estimate for every parameter. Let

$$\theta_j(t) = \hat{\theta}_j(t) + a_j \sin(\eta_j(t)), \tag{6.3}$$

where $a_j > 0$ is the perturbation amplitude and $(\eta_j(t), t \geq 0)$ is an OU process as in (3.3).

By (6.2) and (6.3), we have

$$\theta_j(t) - \theta_j^* = a_j \sin(\eta_j(t)) - \tilde{\theta}_j(t). \tag{6.4}$$

Substituting (6.4) into (6.1), we have the output

$$y(t) = f^* + (\theta(t) - \theta^*)^T P(\theta(t) - \theta^*), \tag{6.5}$$

where $\theta(t) - \theta^* = [a_1 \sin(\eta_1(t)) - \tilde{\theta}_1(t), \ldots, a_l \sin(\eta_l(t)) - \tilde{\theta}_l(t)]^T$.

We design the parameter update law as follows:

$$\frac{d\hat{\theta}_j(t)}{dt} = -k_j a_j \sin(\eta_j(t))(y(t) - \xi_i(t)), \tag{6.6}$$

$$\frac{d\xi_j(t)}{dt} = -h_j \xi_j(t) + h_j y(t), \tag{6.7}$$

$$\varepsilon_j d\eta_j(t) = -\eta_j(t)dt + \sqrt{\varepsilon_j} q_j dW_j(t), \tag{6.8}$$

where $h_j, k_j, j = 1, \dots, l$, are scalar design parameters. Different from the extremum seeking algorithm in section 3, where we excluded the standard washout filter of the output signal [13], which is not essential for convergence but helps performance, in this section we use a washout filter $\frac{s}{s+h_j}$ for each parameter, and the gradient estimation for each parameter is based on the output $\frac{s}{s+h_j}[y] = y(t) - \xi_j(t)$ of this filter.

Define $\chi_j(t) = \eta_j(\varepsilon_j t)$ and $B_j(t) = \frac{1}{\sqrt{\varepsilon_j}} W_j(\varepsilon_j t)$. Then we have

$$d\chi_j(t) = -\chi_j(t)dt + q_j dB_j(t), \tag{6.9}$$

where $B_j(t)$ is a 1-dimensional standard Brownian motion defined on the complete probability space $(\Omega, \mathcal{F}, P)$, while $[B_1(t), \dots, B_l(t)]^T$ is an l-dimensional independent standard Brownian motion on the same space.

Define the output error variable $e_j(t) = \xi_j(t) - f^*$, $j = 1, \dots, l$. Therefore, it follows from (6.2), (6.5), (6.6), and (6.7) that we have the error dynamics

$$\begin{aligned} \frac{d\tilde{\theta}_j(t)}{dt} &= -\frac{d\hat{\theta}_j(t)}{dt} \\ &= -k_j a_j \sin(\eta_j(t))((\theta(t) - \theta^*)^T P(\theta(t) - \theta^*) - e_j(t)) \\ &= -k_j a_j \sin(\chi_j(t/\varepsilon_j))((\theta(t) - \theta^*)^T P(\theta(t) - \theta^*) - e_j(t)), \end{aligned} \tag{6.10}$$

$$\begin{aligned} \frac{de_j(t)}{dt} &= h_j(y(t) - f^* - e_j(t)) \\ &= h_j((\theta(t) - \theta^*)^T P(\theta(t) - \theta^*) - e_j(t)), \quad j = 1, \dots, l. \end{aligned} \tag{6.11}$$

Denote $\tilde{\theta}(t)= [\tilde{\theta}_1(t), \dots, \tilde{\theta}_l(t)]^T$ and $e(t) = [e_1(t), \dots, e_l(t)]^T$. Then we have the following result.

THEOREM 6.1. *Consider the static map* (6.1) *under the parameter update law* (6.6)–(6.8). *Then the error system* (6.10)–(6.11) *is weak stochastic exponentially stable; i.e., there exist constants $r > 0$, $c > 0$, and $\gamma > 0$ such that for any initial condition $|\Lambda_1^{\varepsilon_1}(0)| < r$ and any $\delta > 0$*

$$\lim_{\varepsilon_1 \to 0} \inf \left\{ t \geq 0 : |\Lambda_1^{\varepsilon_1}(t)| > c|\Lambda_1^{\varepsilon_1}(0)|e^{-\gamma t} + \delta \right\} = +\infty \ a.s. \tag{6.12}$$

Moreover, there exists a function $T(\varepsilon_1) : (0, \varepsilon_0) \to \mathbb{N}$ such that

$$\lim_{\varepsilon_1 \to 0} P\left\{ \sup_{0 \leq t \leq T(\varepsilon_1)} \left\{ |\Lambda_1^{\varepsilon_1}(t)| - c|\Lambda_1^{\varepsilon_1}(0)|e^{-\gamma t} \right\} > \delta \right\} = 0 \quad \text{with} \ \lim_{\varepsilon_1 \to 0} T(\varepsilon_1) = \infty, \tag{6.13}$$

where $\Lambda_1^{\varepsilon_1}(t) = (\tilde{\theta}(t)^T, e(t)^T) - (0_{l\times l}^T, \sum_{i=1}^l p_{ii} a_i^2 G_0(q_i) I_1^T)$, $I_1 = [1, 1, \dots, 1]_{1\times l}^T$. Furthermore, (6.13) *is equivalent to*

$$\lim_{\varepsilon_1 \to 0} P\left\{ |\Lambda_1^{\varepsilon_1}(t)| \leq c|\Lambda_1^{\varepsilon_1}(0)|e^{-\gamma t} + \delta \ \forall t \in [0, T(\varepsilon_1)] \right\} = 1 \quad \text{with} \ \lim_{\varepsilon_1 \to 0} T(\varepsilon_1) = \infty. \tag{6.14}$$

6.2. Convergence analysis. We rewrite the error dynamics (6.10)–(6.11) as

$$\begin{aligned}\frac{d\tilde{\theta}_j(t)}{dt} &= k_j a_j \sin(\chi_j(t/\varepsilon_j)) \Big([a_1 \sin(\chi_1(t/\varepsilon_1)) - \tilde{\theta}_1(t), \ldots, a_l \sin(\chi_l(t/\varepsilon_l)) - \tilde{\theta}_l(t)]^T P \\ &\quad \times [a_1 \sin(\chi_1(t/\varepsilon_1)) - \tilde{\theta}_1(t), \ldots, a_l \sin(\chi_l(t/\varepsilon_l)) - \tilde{\theta}_l(t)]\Big) \\ &= k_j a_j \sin(\chi_j(t/\varepsilon_j)) \\ &\quad \times \left(\sum_{i,k=1}^{l} p_{ik} \Big(a_i \sin(\chi_i(t/\varepsilon_i)) - \tilde{\theta}_i(t)\Big) \Big(a_k \sin(\chi_k(t/\varepsilon_k)) - \tilde{\theta}_k(t) - e_j(t)\Big)\right),\end{aligned} \tag{6.15}$$

$$\begin{aligned}\frac{de_i(t)}{dt} &= h_j \left(\sum_{i,k=1}^{l} p_{ik} \Big(a_i \sin(\chi_i(t/\varepsilon_i)) - \tilde{\theta}_i(t)\Big) \Big(a_k \sin(\chi_k(t/\varepsilon_k)) - \tilde{\theta}_k(t)\Big) - e_j(t)\right), \\ &\quad j = 1, \ldots, l.\end{aligned} \tag{6.16}$$

Now we calculate the average system of the error system. Assume that

$$\varepsilon_i = \frac{\varepsilon_1}{c_i}, \quad i = 2, \ldots, l, \tag{6.17}$$

for some positive real constants c_i. Denote

$$Z_1(t) = \chi_1(t),\ Z_2(t) = \chi_2(c_2 t), \ldots, Z_l(t) = \chi(c_l t). \tag{6.18}$$

Then the error dynamics become

$$\begin{aligned}\frac{d\tilde{\theta}_j(t)}{dt} &= k_j a_j \sin(Z_j(t/\varepsilon_1)) \\ &\quad \times \left(\sum_{i,k=1}^{l} p_{ik} \Big(a_i \sin(Z_i(t/\varepsilon_1)) - \tilde{\theta}_i(t)\Big) \Big(a_k \sin(Z_k(t/\varepsilon_1)) - \tilde{\theta}_k(t) - e_j(t)\Big)\right),\end{aligned} \tag{6.19}$$

$$\begin{aligned}\frac{de_j(t)}{dt} &= h_j(y(t) - f^* - e_j(t)) \\ &= h_j \left(\sum_{i,k=1}^{l} p_{ik} \Big(a_i \sin(Z_i(t/\varepsilon_1)) - \tilde{\theta}_i(t)\Big) \Big(a_k \sin(Z_k(t/\varepsilon_1)) - \tilde{\theta}_k(t)\Big) - e_j(t)\right), \\ &\quad j = 1, \ldots, l.\end{aligned} \tag{6.20}$$

It is known that for given $j = 1, \ldots, l$, the stochastic process $(\chi_j(t), t \geq 0)$ is ergodic and has invariant distribution

$$\mu_j(dx_j) = \frac{1}{\sqrt{\pi} q_j} e^{-\frac{x_j^2}{q_j^2}} dx_j.$$

Thus by Lemma A.2, the vector-valued process $[Z_1(t), Z_2(t), \ldots, Z_l(t)]^T$ is also ergodic with invariant distribution

$$\mu_1(dx_1) \times \cdots \times \mu_l(dx_l).$$

To calculate the average system of system (6.19)–(6.20), we need to consider the terms

$$\sin(Z_j(t/\varepsilon_1))\sin(Z_i(t/\varepsilon_1))\sin(Z_k(t/\varepsilon_1)),\quad i\neq j,\ j\neq k,\ k\neq i, \tag{6.21}$$

$$\sin^3(Z_j(t/\varepsilon_1)), \tag{6.22}$$

$$\sin(Z_j(t/\varepsilon_1))\sin^2(Z_i(t/\varepsilon_1)),\quad i\neq j, \tag{6.23}$$

$$\sin^2(Z_j(t/\varepsilon_1)), \tag{6.24}$$

$$\sin(Z_j(t/\varepsilon_1))\sin(Z_i(t/\varepsilon_1)),\quad i\neq j. \tag{6.25}$$

By the integrals (4.17), (4.10), (4.14), (4.11), and (4.13), we get the following average error system:

$$\frac{d\tilde{\theta}_j^{ave}(t)}{dt} = -a_j^2k_j(1-e^{-q_j^2})\sum_{i=1}^{l}p_{ji}\tilde{\theta}_i^{ave}(t), \tag{6.26}$$

$$\frac{de_j^{ave}(t)}{dt} = h_j\left(\sum_{i=1}^{l}p_{ii}a_i^2\frac{1}{2}(1-e^{-q_i^2})+\sum_{i,k=1}^{l}\tilde{\theta}_i^{ave}\tilde{\theta}_k^{ave}-e_j^{ave}(t)\right),\ j=1,\dots,l. \tag{6.27}$$

In the matrix form, the average error system is

$$\frac{d\tilde{\theta}^{ave}(t)}{dt} = -\Pi P\tilde{\theta}^{ave}(t), \tag{6.28}$$

$$\frac{de^{ave}(t)}{dt} = H\left(\sum_{i=1}^{l}p_{ii}a_i^2G_0(q_i)I_1-e^{ave}(t)+Q(\tilde{\theta}^{ave}(t))\right), \tag{6.29}$$

where

$$\Pi = \begin{bmatrix} a_1^2k_1(1-e^{-q_1^2}) & 0 & \cdots & 0 \\ 0 & a_2^2k_2(1-e^{-q_2^2}) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_l^2k_l(1-e^{-q_l^2}) \end{bmatrix},$$

$$\tilde{\theta}^{ave}(t) = [\tilde{\theta}_1^{ave}(t),\dots,\tilde{\theta}_l^{ave}(t)]^T,$$

$$e^{ave}(t) = [e_1^{ave}(t),\dots,e_l^{ave}(t)]^T,$$

$$H = \begin{bmatrix} h_1 & 0 & \cdots & 0 \\ 0 & h_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & h_l \end{bmatrix},$$

$$Q(\tilde{\theta}^{ave}(t)) = \tilde{\theta}^{ave^T}(t)\begin{bmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{bmatrix}_{l\times l}\tilde{\theta}^{ave}(t)I_1,$$

$$I_1 = [1,1,\dots,1]_{1\times l}^T.$$

The average error system has equilibrium $(\tilde{\theta}^{e^T}, e^{e^T}) = (0_{l\times 1}^T, \sum_{i=1}^{l}p_{ii}a_i^2G_0(q_i)I_1^T)$. The corresponding Jacobi matrix at this equilibrium is

$$\Xi_1 = \begin{bmatrix} -\Pi P & 0 \\ 0 & -H \end{bmatrix}. \tag{6.30}$$

Since Π and P are positive definite, all eigenvalues of the matrix ΠP are positive; i.e., the eigenvalues of the matrix $-\Pi P$ are negative. Furthermore, from the fact $h_i > 0,\ i = 1, \ldots, l$, it follows that the matrix Ξ_1 is Hurwitz and hence the equilibrium is locally exponentially stable. Thus by Theorem A.3 in the appendix, the convergence results (6.12) and (6.14) hold. The proof is complete. $\square$

To quantify the output convergence to the extremum, for any $\varepsilon_1 > 0$, define a stopping time

$$\tau_{\varepsilon_1}^{\delta} = \inf\left\{t \geq 0 : |\Lambda_1^{\varepsilon_1}(t)| > c|\Lambda_1^{\varepsilon_1}(0)|e^{-\gamma t} + \delta\right\}.$$

Then by (6.12), we know that $\lim_{\varepsilon_1 \to 0} \tau_{\varepsilon_1}^{\delta} = \infty$ a.s. and

$$\left|\tilde{\theta}(t)\right| \leq c\,|\Lambda_1^{\varepsilon_1}(0)|\,e^{-\gamma t} + \delta \quad \forall t \leq \tau_{\varepsilon_1}^{\delta}. \tag{6.31}$$

Denote $\hat{\theta}(t) = [\hat{\theta}_1(t), \ldots, \hat{\theta}_l(t)]^T,\ a\sin(\eta(t)) = [a_1\sin(\eta_1(t)), \ldots, a_l\sin(\eta_l(t))]^T$. Then $y(t) = f(\theta^* + \tilde{\theta}(t) + a\sin(\eta(t)))$ for $\bigtriangledown f(\theta^*) = 0$, and we have

$$y(t) - f(\theta^*) = (\tilde{\theta}(t) + a\sin(\eta(t)))^T H_f(\theta^*)(\tilde{\theta}(t) + a\sin(\eta(t))) + O\left(|\tilde{\theta}(t) + a\sin(\eta(t))|^3\right),$$

where H_f is the Hessian matrix of the function f.

Thus by (6.31), it holds that

$$|y(t) - f(\theta^*)| \leq O(|a|^2) + O(\delta^2) + C\,|\Lambda_1^{\varepsilon_1}(0)|^2\,e^{-2\gamma t} \quad \forall t \leq \tau_{\varepsilon_1}^{\delta} \tag{6.32}$$

for some positive constant C, where $|a| = \sqrt{a_1^2 + a_2^2 + \cdots + a_l^2}$. Similarly, by (6.14), we have

$$\lim_{\varepsilon_1 \to 0} P\left\{|y(t) - f(\theta^*)| \leq O(|a|^2) + O(\delta^2) + C\,|\Lambda_1^{\varepsilon_1}(0)|^2\,e^{-2\gamma t} \quad \forall t \in [0, T(\varepsilon_1)]\right\} = 1, \tag{6.33}$$

where $T(\varepsilon_1)$ is a deterministic function with $\lim_{\varepsilon_1 \to 0} T(\varepsilon_1) = \infty$.

Figure 4 displays the simulation results with $f^* = 1, (\theta_1^*, \theta_2^*) = (0, 1),\ P = \left[\begin{smallmatrix}1 & 1\\ 1 & 2\end{smallmatrix}\right]$ in the static map (6.1), and $a_1 = 0.8,\ a_2 = 0.6,\ k_1 = 1.25,\ k_2 = 5/3,\ q_1 = q_2 = 1$, $\varepsilon_1 = 0.25,\ \varepsilon_2 = 0.01$ in the parameter update law (6.6)–(6.8) and initial condition $\tilde{\theta}_1(0) = 1,\ \tilde{\theta}_2(0) = -1,\ \hat{\theta}_1(0) = -1,\ \hat{\theta}_2(0) = 2$.

7. Conclusion. In this paper, we propose a multi-input stochastic extremum seeking algorithm to solve the problem of seeking Nash equilibria for an N-player nonoperative game. In our algorithm, each player independently employs his seeking strategy using only the value of his own payoff but without any information about the form of his payoff function and other players' actions. Our convergence result is local, and the convergence error is in proportion to the third derivatives of the payoff functions and is dependent on the intensity of stochastic perturbation. The advantage of our stochastic algorithm over the deterministic ones lies in that there is no requirement on different frequencies of the perturbation signal for different players. As a special case of a multiplayer noncooperative game, stochastic multiparameter extremum seeking for quadratic static maps is investigated.

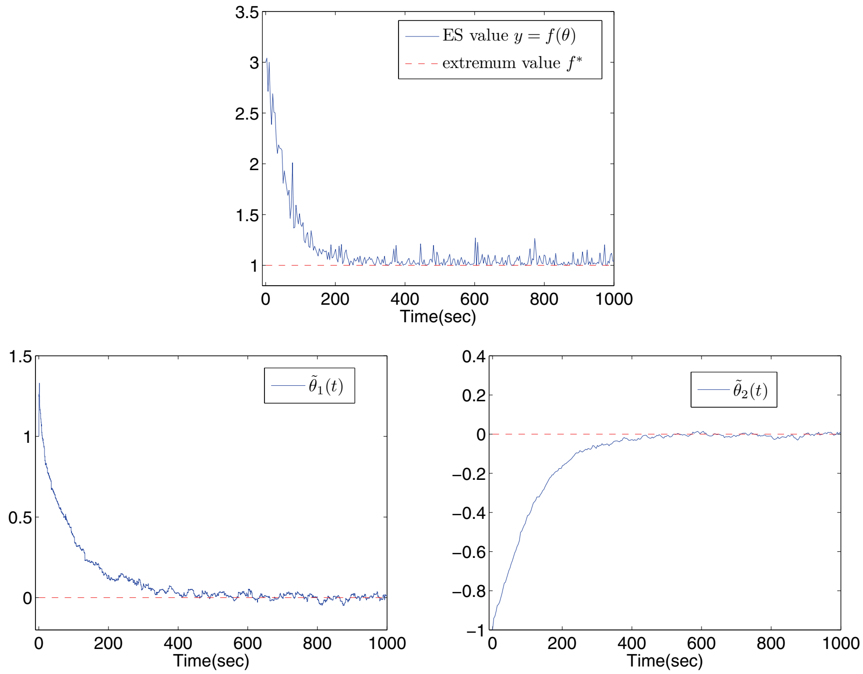

FIG. 4. *Stochastic extremum seeking with an OU process perturbation. Top: output and extremum values. Bottom: solutions of the error system.*

Appendix. Multi-input stochastic averaging. Consider the system

$$\text{(A.1)} \qquad \begin{cases} \frac{dX(t)}{dt} = a(X(t), Y_1(t/\varepsilon_1), Y_2(t/\varepsilon_2), \ldots, Y_l(t/\varepsilon_l)), \\ X(0) = x, \end{cases}$$

where $X(t) \in \mathbb{R}^n$, $Y_i(t) \in \mathbb{R}^{m_i}, 1 \le i \le l$, are time homogeneous continuous Markov processes defined on a complete probability space $(\Omega, \mathcal{F}, P)$, where Ω is the sample space, $\mathcal{F}$ is the σ-field, and P is the probability measure. The initial condition $X(0) = x$ is deterministic. ε_i, $i = 1, 2, \ldots, l$, are some small parameters in $(0, \varepsilon_0)$ with fixed $\varepsilon_0 > 0$. Let $S_{Y_i} \subset \mathbb{R}^{m_i}$ be the living space of the perturbation process $(Y_i(t), t \ge 0)$ and note that S_{Y_i} may be a proper (e.g., compact) subset of $\mathbb{R}^{m_i}$.

Assume that

$$\varepsilon_i = \frac{\varepsilon_1}{c_i}$$

for some positive real constants c_i. Denote $Z_1(t) = Y_1(t)$, $Z_2(t) = Y_2(c_2 t), \ldots, Z_l(t) = Y_l(c_l t)$. Then (A.1) becomes

$$\text{(A.2)} \qquad \begin{cases} \frac{dX(t)}{dt} = a(X(t), Z_1(t/\varepsilon_1), Z_2(t/\varepsilon_1), \ldots, Z_l(t/\varepsilon_1)), \\ X(0) = x. \end{cases}$$

About the ergodicity of the processes $(Y_i(t), t \ge 0)$ and $(Z_i(t), t \ge 0)$, we have the following lemma.

LEMMA A.1. *For $i = 1, \ldots, l$, if the process $(Y_i(t), t \geq 0)$ is ergodic with invariant distribution $\mu_i(dx_i)$ (i.e., for any x in the living space of $(Y_i(t), t \geq 0)$, we have that $\|P_i(x, t, \cdot) - \mu_i\|_{var} \to 0$ as $t \to \infty$, where $P_i(x, t, \cdot)$ is the distribution of $Y_i(t)$ when $Y_i(0) = x$, and $\| \cdot \|_{var}$ is the total variation norm), then the process $(Z_i(t), t \geq 0)$ is ergodic with the same invariant distribution $\mu_i(dx_i)$.*

Proof. Since $Z_1 \equiv Y_1$, we need only prove for $i = 2, \ldots, l$. For any $i = 2, \ldots, l$, denote by $Q_i(z_i, t, \cdot)$ the distribution of $Z_i(t)$ when $Z_i(0) = Y_i(0) = z_i$. Then, by the definition of $Z_i(t)$, we have that $Q_i(z_i, t, \cdot) = P_i(z_i, c_i t, \cdot)$, and thus $\|Q_i(z_i, t, \cdot) - \mu_i\|_{var} = \|P_i(z_i, c_i t, \cdot) - \mu_i\|_{var} \to 0$ as $t \to \infty$. The proof is complete. ∎

Denote $Z(t) = [Z_1(t)^T, Z_2(t)^T, \ldots, Z_l(t)^T]^T$. Then for the vector-valued process, we have the following result.

LEMMA A.2. *If the process $(Y_i(t), t \geq 0)$ is ergodic with invariant distribution $\mu_i(dx_i)$, and the processes $(Y_1(t), t \geq 0), \ldots, (Y_l(t), t \geq 0)$ are independent, then the process $(Z(t), t \geq 0)$ is ergodic with the invariant distribution $\mu_1(dx_1) \times \cdots \times \mu_l(dx_l)$.*

Proof. By the independence of $\{Y_1, \ldots, Y_l\}$, we can assume that the process $(Z(t), t \geq 0)$ lives in the product space of $S_{Y_1} \times \cdots \times S_{Y_l}$. Denote the distribution of $Z_i(t)$ when $Z_i(0) = z_i$, $i = 1, \ldots, l$, by $Q_i(z_i, t, \cdot)$ and the distribution of $Z(t)$ when $Z(0) = z = (z_1, \ldots, z_l)$ by $Q(z, t, \cdot)$. Then by the independence, we have that $Q(z, t, \cdot) = Q_1(z_1, t, \cdot) \times \cdots \times Q_l(z_l, t, \cdot)$. And thus by Lemma A.1, we get

$$
\begin{aligned}
&\|Q(z, t, \cdot) - \mu_1 \times \cdots \times \mu_l\|_{var} \\
&= \|Q_1(z_1, t, \cdot) \times \cdots \times Q_l(z_l, t, \cdot) - \mu_1 \times \mu_2 \times \cdots \times \mu_l\|_{var} \\
&\leq \|Q_1(z_1, t, \cdot) \times \cdots \times Q_l(z_l, t, \cdot) - \mu_1 \times Q_2(z_2, t, \cdot) \times \cdots \times Q_l(z_l, t, \cdot)\|_{var} \\
&\quad + \|\mu_1 \times Q_2(z_2, t, \cdot) \times \cdots \times Q_l(z_l, t, \cdot) - \mu_1 \times \mu_2 \times Q_3(z_3, t, \cdot) \times \cdots \times Q_l(z_l, t, \cdot)\|_{var} \\
&\quad + \cdots + \|\mu_1 \times \cdots \times \mu_{l-1} \times Q_l(z_l, t, \cdot) - \mu_1 \times \cdots \times \mu_{l-1} \times \mu_l\|_{var} \\
&\leq \|Q_1(z_1, t, \cdot) - \mu_1\|_{var} + \cdots + \|Q_l(z_l, t, \cdot) - \mu_l\|_{var} \to 0, \quad t \to \infty.
\end{aligned}
$$

The proof is complete. ∎

So we obtain the average system of system (A.2) as follows:

$$
\text{(A.3)} \qquad \frac{d\bar{X}(t)}{dt} = \bar{a}(\bar{X}(t)), \qquad \bar{X}_0 = x,
$$

where

$$
\text{(A.4)} \qquad \bar{a}(x) = \int_{S_{Y_1} \times \cdots \times S_{Y_l}} a(x, z_1, \ldots, z_l)\mu_1(dz_1) \times \cdots \times \mu_l(dz_l).
$$

To obtain the multi-input stochastic averaging theorem, we consider the following assumptions.

ASSUMPTION A.1. *The vector field $a(x, y_1, y_2, \ldots, y_l)$ is a continuous function of $(x, y_1, y_2, \ldots, y_l)$, and for any $x \in \mathbb{R}^n$, it is a bounded function of $y = [y_1^T, y_2^T, \ldots, y_l^T]^T$. Further, it satisfies the locally Lipschitz condition in $x \in \mathbb{R}^n$ uniformly in $y \in S_{Y_1} \times S_{Y_2} \times \cdots \times S_{Y_l}$; i.e., for any compact subset $D \subset \mathbb{R}^n$, there is a constant k_D such that for all $x_1, x_2 \in D$ and all $y \in S_{Y_1} \times S_{Y_2} \times \cdots \times S_{Y_l}$, $|a(x_1, y) - a(x_2, y)| \leq k_D\ |x_1 - x_2|$.*

ASSUMPTION A.2. *The perturbation processes $(Y_i(t), t \geq 0)$, $i = 1, \ldots, l$, are ergodic with invariant distribution μ_i, respectively, and independent.*

By the same method as in our work [19] and [20] for the single input stochastic averaging theorem, we obtain the following multi-input averaging theorem.

THEOREM A.3. *Consider system* (A.1) *under Assumptions* A.1 *and* A.2. *If the equilibrium* $\bar{X}(t) \equiv 0$ *of the average system* (A.3) *is locally exponentially stable, then the following statements hold:*

(i) *the solution of system* (A.1) *is weakly stochastic exponentially stable under random perturbation; i.e., there exist constants* $r > 0$, $c > 0$, *and* $\gamma > 0$ *such that for any initial condition* $x \in \{\check{x} \in \mathbb{R}^n : |\check{x}| < r\}$ *and any* $\delta > 0$, *the solution of system* (A.1) *satisfies*

$$\lim_{\varepsilon_1 \to 0} \inf \left\{ t \geq 0 : |X(t)| > c|x|e^{-\gamma t} + \delta \right\} = +\infty \ a.s. \tag{A.5}$$

(ii) *Moreover, there exists a function* $T(\varepsilon_1) : (0, \varepsilon_0) \to \mathbb{N}$ *such that*

$$\lim_{\varepsilon_1 \to 0} P \left\{ \sup_{0 \leq t \leq T(\varepsilon_1)} \left\{ |X(t)| - c|x|e^{-\gamma t} \right\} > \delta \right\} = 0 \quad \text{with } \lim_{\varepsilon_1 \to 0} T(\varepsilon_1) = \infty. \tag{A.6}$$

Furthermore, (A.6) *is equivalent to*

$$\lim_{\varepsilon_1 \to 0} P \left\{ |X(t)| \leq c|x|e^{-\gamma t} + \delta \quad \forall t \in [0, T(\varepsilon_1)] \right\} = 1 \quad \text{with } \lim_{\varepsilon_1 \to 0} T(\varepsilon_1) = \infty. \tag{A.7}$$

REFERENCES

[1] E. ALTMAN, T. BAŞAR, AND R. SRIKANT, *Nash equilibria for combined flow control and routing in networks: Asymptotic behavior for a large number of users*, IEEE Trans. Automat. Control, 47 (2002), pp. 917–930.

[2] K. B. ARIYUR AND M. KRSTIĆ, *Real-Time Optimization by Extremum Seeking Control*, Wiley-Interscience, Hoboken, NJ, 2003.

[3] T. BAŞAR, *Control and game-theoretic tools for communication networks*, Appl. Comput. Math., 6 (2007), pp. 104–125.

[4] T. BAŞAR AND G. J. OLSDER, *Dynamic Noncooperative Game Theory*, 2nd ed., Classics Appl. Math. 23, SIAM, Philadelphia, 1999.

[5] D. BAUSO, L. GIARRE, AND R. PESENTI, *Consensus in noncooperative dynamic games: A multiretailer inventory application*, IEEE Trans. Automat. Control, 53 (2008), pp. 998–1003.

[6] G. BLANKENSHIP AND G. C. PAPANICOLAOU, *Stability and control of stochastic systems with wide-band noise disturbances*. I, SIAM J. Appl. Math., 34 (1978), pp. 437–476.

[7] J.-Y. CHOI, M. KRSTIĆ, K. B. ARIYUR, AND J. S. LEE, *Extremum seeking control for discrete time systems*, IEEE Trans. Automat. Control, 47 (2002), pp. 318–323.

[8] M. I. FREIDLIN AND A. D. WENTZELL, *Random Perturbations of Dynamical Systems*, Grundlehren Math. Wiss. 260, Springer-Verlag, New York, 1984.

[9] P. FRIHAUF, M. KRSTIĆ, AND T. BAŞAR, *Nash equilibrium seeking for games with non-quadratic payoffs*, in Proceedings of the 2010 IEEE Conference on Decision and Control, 2010, pp. 881–886.

[10] R. Z. KHASMINSKII AND G. YIN, *On averaging principles: An asymptotic expansion approach*, SIAM J. Math. Anal., 35 (2004), pp. 1534–1560.

[11] R. KING, R. BECKER, G. FEUERBACH, L. HENNING, R. PETZ, W. NITSCHE, O. LEMKE, AND W. NEISE, *Adaptive flow control using slope seeking*, in Proceedings of the 14th IEEE Mediterranean Conference on Control and Automation, 2006, pp. 1–6.

[12] M. KRSTIĆ, P. FRIHAUF, J. KRIEGER, AND T. BAŞAR, *Nash Equilibrium Seeking with Finitely- and Infinitely-Many Players*, in Proceedings of the 8th IFAC Symposium on Nonlinear Control Systems, 2010.

[13] M. KRSTIĆ AND H. H. WANG, *Stability of extremum seeking feedback for general nonlinear dynamic systems*, Automatica J. IFAC, 36 (2000), pp. 595–601.

[14] B. J. KUBICA AND A. WOZNIAK, *An interval method for seeking the Nash equilibrium of non-cooperative games*, Lecture Notes in Comput. Sci. 6068, Springer-Verlag, Berlin, 2010, pp. 446–455.

[15] H. J. Kushner and K. M. Ramachandran, *Nearly optimal singular controls for wideband noise driven systems*, SIAM J. Control Optim., 26 (1988), pp. 569–591.

[16] S. Li and T. Başar, *Distributed algorithms for the computation of noncooperative equilibria*, Automatica J. IFAC, 23 (1987), pp. 523–533.

[17] Y. Li, M. A. Rotea, G. T.-C. Chiu, L. G. Mongeau, and I.-S. Paek, *Extremum seeking control of a tunable thermoacoustic cooler*, IEEE Trans. Control Syst. Tech., 36 (2005), pp. 527–536.

[18] S.-J. Liu and M. Krstić, *Continuous-time stochastic averaging on the infinite interval for locally Lipschitz systems*, SIAM J. Control Optim., 48 (2010), pp. 3589–3622.

[19] S.-J. Liu and M. Krstić, *Stochastic averaging in continuous time and its applications to extremum seeking*, IEEE Trans. Automat. Control, 55 (2010), pp. 2235–2250.

[20] S.-J. Liu and M. Krstić, *Stochastic source seeking for nonholonomic unicycle*, Automatica J. IFAC, 46 (2010), pp. 1443–1453.

[21] L. Luo and E. Schuster, *Mixing enhancement in 2D magnetohydrodynamic channel flow by extremum seeking boundary control*, in Proceedings of the 2009 American Control Conference, 2009, pp. 1530–1535.

[22] A. B. MacKenzie and S. B. Wicker, *Game theory and the design of self-configuring, adaptive wireless networks*, IEEE Commun. Mag., 39 (2001), pp. 126–131.

[23] J. R. Marden, G. Arslon, and J. S. Shamma,*Cooperative control and potential games*, IEEE Trans. Syst. Man Cybernet.: Part B: Cybernetics, 39 (2009), pp. 1393–1407.

[24] W. H. Moase, C. Manzie, and M. J. Brear, *Newton-like extremum-seeking Part* I: *Theory*, in Proceedings of the 48th IEEE Conference on Decision and Control and 28th Chinese Control Conference, 2009, pp. 3840–3844.

[25] W. H. Moase, C. Manzie, and M. J. Brear, *Newton-like extremum-seeking Part* II: *Simulation and experiments*, in Proceedings of the 48th IEEE Conference on Decision and Control and 28th Chinese Control Conference, 2009, pp. 3845–3850.

[26] E. Pardoux and A. Yu. Veretennikov, *On the Poisson equation and diffusion approximation.* I, Ann. Probab., 29 (2001), pp. 1061–1085.

[27] S. S. Rao, V. B. Venkayya, and, N. S. Khot, *Game theory approach for the integrated design of structures and controls*, AIAA J., 26 (1988), pp. 463–469.

[28] J. B. Rosen, *Existence and uniqueness of equilibrium points for concave N-person games*, Econometrica, 33 (1965), pp. 520–534.

[29] G. Scutari, D. P. Palomar, and S. Barbarossa, *The MIMO iterative waterfilling algorithm*, IEEE Trans. Signal Process., 57 (2009), pp. 1917–1935.

[30] E. Semsar-Kazerooni and K. Khorasani, *Multi-agent team cooperation: A game theory approach*, Automatica J. IFAC, 45 (2009), pp. 2205–2213.

[31] J. S. Shamma and G. Arslan, *Dynamic fictitious play, dynamic gradient play, and distributed convergence to Nash equilibria*, IEEE Trans. Automat. Control, 53 (2005), pp. 312–327.

[32] A. V. Skorokhod, *Asymptotic Methods in the Theory of Stochastic Differential Equations*, Transl. Math. Monogr. 78, AMS, Providence, RI, 1989.

[33] M. S. Stanković, K. H. Johansson, and D. M. Stipanović, *Distributed seeking of Nash equilibrium in mobile sensor networks*, in Proceedings of the 2010 IEEE Conference on Decision and Control, 2010, pp. 5598–5603.

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

[35] M. S. Stanković and D. M. Stipanović, *Discrete time extremum seeking by autonomous vehicles in a stochastic environment*, in Proceedings of the 48th IEEE Conference on Decision and Control and 28th Chinese Control Conference, 2009, pp. 4541–4546.

[36] Y. Tan, D. Nešić, and I. Mareels, *On non-local stability properties of extremum seeking control*, Automatica J. IFAC, 42 (2006), pp. 889–903.

[37] W. Wehner and E. Schuster, *Stabilization of neoclassical tearing modes in tokamak fusion plasmas via extremum seeking*, in Proceedings of the Third IEEE Multi-conference on Systems and Control (MSC 2009), 2009, pp. 853–860.

[38] M. Zhu and S. Martinez, *Distributed coverage games for mobile visual sensor networks*, SIAM J. Control Optim., submitted.