Nash Equilibrium Seeking in Noncooperative Games

Paul Frihauf, *Student Member, IEEE*, Miroslav Krstic, *Fellow, IEEE*, and Tamer Başar, *Fellow, IEEE*

***Abstract*—We introduce a non-model based approach for locally stable convergence to Nash equilibria in static, noncooperative games with N players. In classical game theory algorithms, each player employs the knowledge of the functional form of his payoff and the knowledge of the other players' actions, whereas in the proposed algorithm, the players need to measure only their own payoff values. This strategy is based on the extremum seeking approach, which has previously been developed for standard optimization problems and employs sinusoidal perturbations to estimate the gradient. We consider static games with quadratic payoff functions before generalizing our results to games with non-quadratic payoff functions that are the output of a dynamic system. Specifically, we consider general nonlinear differential equations with N inputs and N outputs, where in the steady state, the output signals represent the payoff functions of a noncooperative game played by the steady-state values of the input signals. We employ the standard local averaging theory and obtain local convergence results for both quadratic payoffs, where the actual convergence is semi-global, and non-quadratic payoffs, where the potential existence of multiple Nash equilibria precludes semi-global convergence. Our convergence conditions coincide with conditions that arise in model-based Nash equilibrium seeking. However, in our framework the user is not meant to check these conditions because the payoff functions are presumed to be unknown. For non-quadratic payoffs, convergence to a Nash equilibrium is not perfect, but is biased in proportion to the perturbation amplitudes and the higher derivatives of the payoff functions. We quantify the size of these residual biases and confirm their existence numerically in an example noncooperative game. In this example, we present the first application of extremum seeking with projection to ensure that the players' actions remain in a given closed and bounded action set.**

***Index Terms*—Extremum seeking, learning, Nash equilibria, noncooperative games.**

I. Introduction

We study the problem of solving noncooperative games with N players in real time by employing a non-model-based approach where the players determine their actions using only their own measured payoff values. By utilizing deterministic extremum seeking with sinusoidal perturbations, the players attain their Nash strategies without the need for any model information. We analyze both static games and games with dynamics, where the players' actions serve as inputs to a general, stable nonlinear dynamic system whose outputs are the players' payoff values. In the latter scenario, the dynamic system evolves on a faster time scale compared to the time scale of the players' strategies, resulting in a static game being played at the steady state of the dynamic system. In these scenarios, the players possess either quadratic or non-quadratic payoff functions, which may result in multiple, isolated Nash equilibria. We quantify the convergence bias relative to the Nash equilibria that results from payoff function terms of higher order than quadratic.

1) Literature Review: Most algorithms designed to achieve convergence to Nash equilibria require modeling information of the game and assume the players can observe the actions of the other players. Two classical examples are the best response and fictitious play strategies, where each player chooses the action that maximizes its payoff given the actions of the other players. The Cournot adjustment, a myopic best response strategy, was first studied by Cournot [1] and refers to the scenario where a firm in a duopoly adjusts its output to maximize its payoff based on the known output of its competitor. The strategy known as fictitious play (employed in finite games), where a player devises a best response based on the history of the other players' actions was introduced in [2] in the context of mixed-strategy Nash equilibria in matrix games. In [3], a gradient-based learning strategy where a player updates its action according to the gradient of its payoff function was developed. Stability of general player adjustments was studied in [4] under the assumption that a player's response mapping is a contraction, which is ensured by a diagonal dominance condition for games with quadratic payoff functions. Distributed iterative algorithms for the computation of equilibria in a general class of non-quadratic convex Nash games were designed in [5], and conditions for the contraction of general nonlinear operators were obtained to achieve convergence.

In more recent work, a dynamic version of fictitious play and gradient response, which also includes an entropy term, is developed in [6] and is shown to converge to a mixed-strategy Nash equilibrium in cases that previously developed algorithms did not converge. In [7], a synchronous distributed learning algorithm, where players remember their own actions and utility values from the previous two times steps, is shown to converge in probability to a set of restricted Nash equilibria. An approach that is similar to our Nash seeking method (found in [8], [9] and in this paper) is studied in [10] to solve coordination problems in mobile sensor networks. Additional results on learning in games can be found in [11]–[15]. Some diverse engineering applications of game theory include the design of communication net-

Manuscript received August 30, 2010; revised May 15, 2011; accepted September 23, 2011. Date of publication October 25, 2011; date of current version April 19, 2012.This work was supported by the Department of Defense (DoD), Air Force Office of Scientific Research (AFOSR), National Defense Science and Engineering Graduate (NDSEG) Fellowship, 32 CFR 168a, and by grants from National Science Foundation, Department of Energy (DOE), and AFOSR. Recommended by Associate Editor J. Grizzle.

P. Frihauf and M. Krstic are with the Department of Mechanical and Aerospace Engineering University of California at San Diego, La Jolla, CA 92093 (e-mail: pfrihauf@ucsd.edu; krstic@ucsd.edu).

T. Başar is with the Coordinated Science Laboratory, University of Illinois, Urbana, IL 61801-2307 USA (e-mail: basar1@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.2011.2173412

0018-9286/$26.00 © 2011 IEEE

works in [16]–[20], integrated structures and controls in [21], and distributed consensus protocols in [22]–[24]. A comprehensive treatment of static and dynamic noncooperative game theory can be found in [25].

The results of this work extend the extremum seeking method [26]–[33], originally developed for standard optimization problems. Many works have used extremum seeking, which performs non-model based gradient estimation, for a variety of applications, such as steering vehicles toward a source in GPS-denied environments [34]–[36], optimizing the control of homogeneous charge-compression ignition (HCCI) engines [37] and nonisothermal continuously stirred tank reactors [38], reducing the impact velocity of an electromechanical valve actuator [39], and controlling flow separation [40] and Tokamak plasmas [41].

2) Contributions: We develop a Nash seeking strategy to stably attain a Nash equilibrium in noncooperative N-player games. The key feature of our approach is that the players do not need to know the mathematical model of their payoff functions or of the underlying model of the game. They only need to measure their own payoff values when determining their respective time-varying actions, which classifies this learning strategy as *radically uncoupled* according to the terminology of [12].

As noted earlier, a game may possess multiple, isolated Nash equilibria if the players have non-quadratic payoff functions. Hence, we pursue local convergence results because, in non-quadratic problems, global results come under strong restrictions. We do, however, make the connection to semi-global practical asymptotic stability when the players have quadratic payoff functions. For non-quadratic payoff functions, we show that the convergence is biased in proportion to the amplitudes of the perturbation signals and the third derivatives of the payoff functions, and confirm this convergence bias in a numerical example. In this example, we impose an action set $U \subset \mathbb{R}^N$ for the players and implement a Nash seeking strategy with projection to ensure that the players' actions remain in U. This example is the first such application of extremum seeking with projection.

3) Organization: We motivate our Nash seeking strategy with a duopoly price game in Section II and prove convergence to the Nash equilibrium in N-player games with quadratic payoff functions in Section III. Also in Section III, we show that the players converge to their best response strategies when a subset of players have a fixed action. In Section IV, we provide our most general result for N-player games with non-quadratic payoff functions and a dynamic mapping from the players' actions to their payoff values. A detailed numerical example is provided in Section V before concluding with Section VI.

II. Two-Player Game

To introduce our Nash seeking algorithm, we first consider a specific two-player noncooperative game, which for example, may represent two firms competing for profit in a duopoly market structure. Common duopoly examples include the soft drink companies, Coca-Cola and Pepsi, and the commercial aircraft companies, Boeing and Airbus. We present a duopoly price game in this section for motivational purposes before proving convergence to the Nash equilibrium when N players

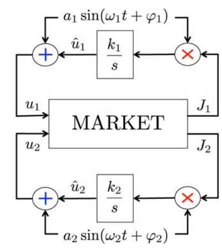


Fig. 1. Deterministic Nash seeking schemes applied by players in a duopoly market structure.

with quadratic payoff functions employ our Nash seeking strategy in Section III.

Let players P1 and P2 represent two firms that produce the same good, have dominant control over a market, and compete for profit by setting their prices u_1 and u_2, respectively. The profit of each firm is the product of the number of units sold and the profit per unit, which is the difference between the sale price and the marginal or manufacturing cost of the product. In mathematical terms, the profits are modeled by

$$J_i(t) = s_i(t)\left(u_i(t) - m_i\right) \tag{1}$$

where s_i is the number of sales, m_i the marginal cost, and $i \in \{1, 2\}$ for P1 and P2. Intuitively, the profit of each firm will be low if it either sets the price very low, since the profit per unit sold will be low, or if it sets the price too high, since then consumers will buy the other firm's product. The maximum profit is to be expected to lie somewhere in the middle of the price range, and it crucially depends on the price level set by the other firm.

To model the market behavior, we assume a simple, but quite realistic model, where for whatever reason, the consumer prefers the product of P1, but is willing to buy the product of P2 if its price u_2 is sufficiently lower than the price u_1. Hence, we model the sales for each firm as

$$s_1(t) = S_d - s_2(t) \tag{2}$$
$$s_2(t) = \frac{1}{p}\left(u_1(t) - u_2(t)\right) \tag{3}$$

where the total consumer demand S_d is held fixed for simplicity, the preference of the consumer for P1 is quantified by $p > 0$, and the inequalities $u_1 > u_2$ and $(u_1 - u_2)/p < S_d$ are assumed to hold.

Substituting (2) and (3) into(1) yields expressions for the profits $J_1(u_1, u_2)$ and $J_2(u_1, u_2)$ that are both quadratic functions of the prices u_1 and u_2, namely,

$$J_1 = \frac{-u_1^2 + u_1 u_2 + (m_1 + S_d p)u_1 - m_1 u_2 - S_d p m_1}{p} \tag{4}$$
$$J_2 = \frac{-u_2^2 + u_1 u_2 - m_2 u_1 + m_2 u_2}{p} \tag{5}$$

and thus, the Nash equilibrium is easily determined to be

$$u_1^* = \frac{1}{3}(2m_1 + m_2 + 2S_d p) \tag{6}$$
$$u_2^* = \frac{1}{3}(m_1 + 2m_2 + S_d p). \tag{7}$$

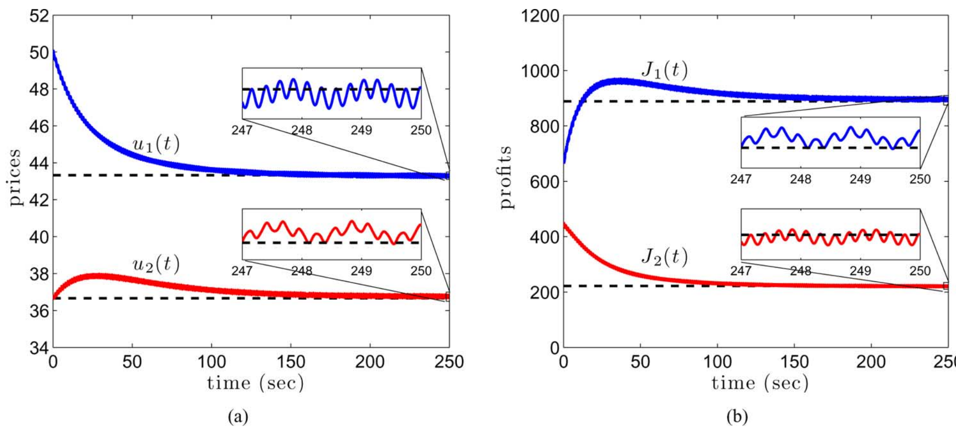


Fig. 2. (a) Price and (b) profit time histories for P1 and P2 when implementing the Nash seeking scheme (9)–(10). The dashed lines denote the values at the Nash equilibrium.

To make sure the constraints $u_1 > u_2$, $(u_1 - u_2)/p < S_d$ are satisfied by the Nash equilibrium, we assume that $m_1 - m_2$ lies in the interval $(-S_d p, 2S_d p)$. If $m_1 = m_2$, this condition is automatically satisfied.

For completeness, we provide here the definition of a Nash equilibrium $u^* = [u_1^*, \ldots, u_N^*]^T$ in an N-player game:

$$J_i(u_i^*, u_{-i}^*) \geq J_i(u_i, u_{-i}^*) \quad \forall u_i \in U_i,\ i \in \{1, \ldots, N\} \tag{8}$$

where J_i is the payoff function of player i, u_i its action, U_i its action set, and u_{-i} denotes the actions of the other players. Hence, no player has an incentive to unilaterally deviate its action from u^*. In the duopoly example, $U_1 = U_2 = \mathbb{R}_+$, where $\mathbb{R}_+$ denotes the set of positive real numbers.

To attain the Nash strategies (6)–(7) without any knowledge of modeling information, such as the consumer's preference p, the total demand S_d, or the other firm's marginal cost or price, the firms implement a non-model based real-time optimization strategy, e.g., deterministic extremum seeking with sinusoidal perturbations, to set their price levels. Specifically, P1 and P2 set their prices, u_1 and u_2 respectively, according to the time-varying strategy (Fig. 1):

$$\dot{\hat{u}}_i(t) = k_i \mu_i(t) J_i(t) \tag{9}$$

$$u_i(t) = \hat{u}_i(t) + \mu_i(t) \tag{10}$$

where $\mu_i(t) = a_i \sin(\omega_i t + \varphi_i)$, $k_i, a_i, \omega_i > 0$, and $i \in \{1, 2\}$. Further, the frequencies are of the form

$$\omega_i = \omega \bar{\omega}_i \tag{11}$$

where ω is a positive, real number and $\bar{\omega}_i$ is a positive, rational number. This form is convenient for the convergence analysis performed in Section III. The resulting pricing, sales, and profit transients when the players implement (9)–(10) are shown in Fig. 2 for a simulation with $S_d = 100, p = 0.2, m_1 = m_2 = 30$, $a_1 = 0.075$, $a_2 = 0.05$, $k_1 = 2$, $k_2 = 5$, $\omega_1 = 26.75$ rad/s, $\omega_2 = 22$ rad/s, and $u_1(0) = \hat{u}_1(0) = 50$, $u_2(0) = \hat{u}_2(0) = u_2^* = 110/3$. From Fig. 2(a), we see that convergence to the Nash equilibrium is not trivially achieved since $u_2(0) = u_2^*$ and yet $u_2(t)$ increases initially before decreasing to u_2^* due to the overall system dynamics of (9)–(10) with(4)–(5). In these simulations (and those shown in Section III-D), we utilize a washout (high pass) filter in each player's extremum seeking loop [28], but we do not include this filter in the theoretical derivations since it is not needed to derive our stability results. Its inclusion would substantially lengthen the presentation due to the doubling of the state dimension and obfuscate the stability results. The washout filter removes the DC component from signal $J_i(t)$, which, while not necessary, typically improves performance.

In contrast, the firms are also guaranteed to converge to the Nash equilibrium when employing the standard parallel action update scheme [25, Proposition 4.1]

$$u_1^{(k+1)} = \frac{1}{2}\left(u_2^{(k)} + m_1 + S_d p\right) \tag{12}$$

$$u_2^{(k+1)} = \frac{1}{2}\left(u_1^{(k)} + m_2\right) \tag{13}$$

which requires each firm to know both its own marginal cost and the other firm's price at the previous step of the iteration, and also requires P1 to know the total demand S_d and the consumer preference parameter p. In essence, P1 must know nearly all the relevant modeling information. When using the extremum seeking algorithm (9)–(10), the firms only need to measure the value of their own payoff functions, J_1 and J_2. Convergence of (12)–(13) is global, whereas the convergence of the Nash seeking strategy for this example can be proved to be semi-global, following [29], or locally, by applying the theory of averaging [43]. We establish local results in this paper since we consider non-quadratic payoff functions in Section IV. We do, however, state a non-local result for static games with quadratic payoff functions using the theory found in [42]. For detailed analysis of the non-local convergence of extremum seeking controllers applied to general convex systems, the reader is referred to [29].

III. N-Player Games With Quadratic Payoff Functions

We now generalize the duopoly example in Section II to static noncooperative games with N players that wish to maximize their quadratic payoff functions. We prove convergence to a neighborhood of the Nash equilibrium when the players employ the Nash seeking strategy (9)–(10).

A. General Quadratic Games

First, we consider games with general quadratic payoff functions. Specifically, the payoff function of player i is of the form

$$J_i(t) = \frac{1}{2}\sum_{j=1}^{N}\sum_{k=1}^{N} D^i_{jk} u_j(t) u_k(t) + \sum_{j=1}^{N} d^i_j u_j(t) + c^i \quad (14)$$

where the action of player j is $u_j \in U_j = \mathbb{R}$, D^i_{jk}, d^i_j, and c^i are constants, $D^i_{ii} < 0$, and $D^i_{jk} = D^i_{kj}$. Quadratic games of this form are studied in [25, Sec. 4.6] where Proposition 4.6 states that the N-player game with payoff functions (14) admits a Nash equilibrium $u^* = [u^*_1, \ldots, u^*_N]^T$ if and only if

$$D^i_{ii} u^*_i + \sum_{j \neq i} D^i_{ij} u^*_j + d^i_i = 0, \quad i \in \{1, \ldots, N\} \quad (15)$$

admits a solution. Rewritten in matrix form, we have $Du^* = -d$, where

$$D \triangleq \begin{bmatrix} D^1_{11} & D^1_{12} & \cdots & D^1_{1N} \\ D^2_{21} & D^2_{22} & & \\ \vdots & & \ddots & \\ D^N_{N1} & & & D^N_{NN} \end{bmatrix}, \quad d \triangleq \begin{bmatrix} d^1_1 \\ d^2_2 \\ \vdots \\ d^N_N \end{bmatrix} \quad (16)$$

and u^* is unique if D is invertible. We make the following stronger assumption concerning this matrix:

Assumption 3.1: The matrix D given by (16) is strictly diagonally dominant, i.e.,

$$\sum_{j \neq i}^{N} \left|D^i_{ij}\right| < \left|D^i_{ii}\right|, \quad i \in \{1, \ldots, N\}. \quad (17)$$

By Assumption 3.1, the Nash equilibrium u^* exists and is unique since strictly diagonally dominant matrices are nonsingular by the Levy–Desplanques theorem [44], [45]. To attain u^* stably in real time, without any modeling information, each player i employs the extremum seeking strategy (9)–(10).

Theorem 1: Consider the system (9)–(10) with (14) under Assumption 3.1 for an N-player game, where $\omega_i \neq \omega_j$, $2\omega_i \neq \omega_j$, and $\omega_i \neq \omega_j + \omega_k$ for all distinct $i, j, k \in \{1, \ldots, N\}$, and where ω_i/ω_j is rational for all $i, j \in \{1, \ldots, N\}$. There exist $\omega^*, M, m > 0$ such that for all $\omega > \omega^*$, if $|\Delta(0)|$ is sufficiently small, then for all $t \geq 0$,

$$|\Delta(t)| \leq M e^{-mt} |\Delta(0)| + O\left(\frac{1}{\omega} + \max_i a_i\right) \quad (18)$$

where $\Delta(t) = [u_1(t) - u^*_1, \ldots, u_N(t) - u^*_N]^T$ and $|\cdot|$ denotes the Euclidean norm.

Proof: Denote the relative Nash equilibrium error as

$$\begin{aligned} \tilde{u}_i(t) &= u_i(t) - \mu_i(t) - u^*_i \\ &= \hat{u}_i(t) - u^*_i. \end{aligned} \quad (19)$$

By substituting (14) into (9)–(10), we get the error system

$$\begin{aligned} \dot{\tilde{u}}_i(t) = {} & k_i \mu_i(t) \\ & \times \left(\frac{1}{2}\sum_{j=1}^{N}\sum_{k=1}^{N} D^i_{jk} \left(\tilde{u}_j(t) + u^*_j + \mu_j(t)\right)\right. \\ & \times \left(\tilde{u}_k(t) + u^*_k + \mu_k(t)\right) \\ & \left. + \sum_{j=1}^{N} d^i_j \left(\tilde{u}_j(t) + u^*_j + \mu_j(t)\right) + c^i\right). \end{aligned} \quad (20)$$

Let $\tau = \omega t$ where ω is the positive, real number in (11). Rewriting(20) in the time scale τ and rearranging terms yields

$$\begin{aligned} \frac{d\tilde{u}_i(\tau)}{d\tau} = {} & \frac{k_i \mu_i(\tau)}{2\omega} \\ & \times \left[\sum_{j=1}^{N}\sum_{k=1}^{N} D^i_{jk}\left(\tilde{u}_j(\tau) + u^*_j\right)\left(\tilde{u}_k(\tau) + u^*_k\right)\right. \\ & + 2\sum_{j=1}^{N}\sum_{k=1}^{N} D^i_{jk}\left(\tilde{u}_j(\tau) + u^*_j\right)\mu_k(\tau) \\ & + \sum_{j=1}^{N}\sum_{k=1}^{N} D^i_{jk}\mu_j(\tau)\mu_k(\tau) \\ & \left. + 2\sum_{j=1}^{N} d^i_j\left(\tilde{u}_j(\tau) + u^*_j + \mu_j(\tau)\right) + 2c^i\right] \\ = {} & \frac{1}{\omega} f_i\left(\tau, \tilde{u}_1, \ldots, \tilde{u}_N, \frac{1}{\omega}\right) \end{aligned} \quad (21)$$

where $\mu_i(\tau) = a_i \sin(\bar{\omega}_i \tau + \varphi_i)$ and $\bar{\omega}_i$ is a rational number. Hence, the error system (21) is periodic with period $T = 2\pi \times \mathrm{LCM}\{1/\bar{\omega}_1, \ldots, 1/\bar{\omega}_N\}$, where LCM denotes the least common multiple. With $1/\omega$ as a small parameter, (21) admits the application of the averaging theory [43] for stability analysis. The average error system can be shown to be

$$\begin{aligned} \frac{d}{d\tau}\tilde{u}^{\mathrm{ave}}_i &= \frac{1}{\omega}\left(\frac{1}{T}\int_0^T f_i(\tau, \tilde{u}^{\mathrm{ave}}_1, \ldots, \tilde{u}^{\mathrm{ave}}_N, 0)\, d\tau\right) \\ &= \frac{1}{2\omega} k_i a_i^2 \sum_{j=1}^{N} D^i_{ij} \tilde{u}^{\mathrm{ave}}_j(\tau) \end{aligned} \quad (22)$$

which in matrix form is $d\tilde{u}^{\mathrm{ave}}/d\tau = A\tilde{u}^{\mathrm{ave}}$, where

$$A = \frac{1}{2\omega}\begin{bmatrix} \kappa_1 D^1_{11} & \kappa_1 D^1_{12} & \cdots & \kappa_1 D^1_{1N} \\ \kappa_2 D^2_{21} & \kappa_2 D^2_{22} & & \\ \vdots & & \ddots & \\ \kappa_N D^N_{N1} & & & \kappa_N D^N_{NN} \end{bmatrix} \quad (23)$$

and $\kappa_i = k_i a_i^2, 1 \in \{1, \ldots, N\}$. (The details of computing (22), which require that $\omega_i \neq \omega_j$, $2\omega_i \neq \omega_j$, and $\omega_i \neq \omega_j + \omega_k$ for all distinct $i, j, k \in \{1, \ldots, N\}$, are shown in Appendix A.)

From the Gershgorin Circle Theorem [45, Theorem 6.1.1], we have $\lambda(A) \subseteq \bigcup_{i=1}^{N} \rho_i$, where $\lambda(A)$ denotes the spectrum of A and ρ_i is a Gershgorin disc:

$$\rho_i = \frac{k_i a_i^2}{2\omega}\left\{z \in \mathbb{C} \;\middle|\; \left|z - D^i_{ii}\right| < \sum_{j \neq i} |D^i_{ij}|\right\}. \quad (24)$$

Since $D^i_{ii} < 0$ and D is strictly diagonally dominant, the union of the Gershgorin discs lies strictly in the left half of the complex plane, and we conclude that $\mathrm{Re}\{\lambda\} < 0$ for all $\lambda \in \lambda(A)$. Thus, given any matrix $Q = Q^T > 0$, there exists a matrix $P = P^T > 0$ satisfying the Lyapunov equation $PA + A^T P = -Q$.

Using $V(\tau) = (\tilde{u}^{\text{ave}}(\tau))^T P \tilde{u}^{\text{ave}}(\tau)$ as a Lyapunov function, we obtain

$$\dot{V} = -(\tilde{u}^{\text{ave}})^T Q \tilde{u}^{\text{ave}} \leq -\lambda_{\min}(Q) \left|\tilde{u}^{\text{ave}}\right|^2. \tag{25}$$

Noting that V satisfies the bounds, $\lambda_{\min}(P) \left|\tilde{u}^{\text{ave}}(\tau)\right|^2 \leq V(\tau) \leq \lambda_{\max}(P) \left|\tilde{u}^{\text{ave}}(\tau)\right|^2$, and applying the Comparison Lemma [43] gives

$$\left|\tilde{u}^{\text{ave}}(\tau)\right| \leq M e^{-m\tau/\omega} \left|\tilde{u}^{\text{ave}}(0)\right|, \tag{26}$$

$$\text{where } M = \sqrt{\frac{\lambda_{\max}(P)}{\lambda_{\min}(P)}},\ m = \frac{\omega \lambda_{\min}(Q)}{2\lambda_{\max}(P)}. \tag{27}$$

From (26) and [43, Th. 10.4], we obtain $|\tilde{u}(\tau)| \leq M e^{-m\tau/\omega} |\tilde{u}(0)| + O(1/\omega)$, provided $\tilde{u}(0)$ is sufficiently close to $\tilde{u}^{\text{ave}}(0)$. Reverting to the time scale t and noting that $u_i(t) - u_i^* = \tilde{u}_i(t) + \mu_i(t) = \tilde{u}_i(t) + O(\max_i a_i)$ completes the proof. ■

From the proof, we see that the convergence result holds if (23) is Hurwitz, which does not require the strict diagonal dominance assumption. However, Assumption 3.1 allows convergence to hold for k_i, $a_i > 0$, whereas merely assuming (23) is Hurwitz would create a potentially intricate dependence on the unknown game model and the selection of the parameters k_i, a_i. Also, while we have considered only the case where the action variables of the players are scalars, the results equally apply to the vector case, namely $u_i \in \mathbb{R}^n$, by simply considering each different component of a player's action variable to be controlled by a different (virtual) player. In this case, the payoff functions of all virtual players corresponding to player i will be the same.

Even though u^* is unique for quadratic payoffs, Theorem 1 is local due to our use of standard local averaging theory. From the theory in [42], we have the following non-local result:

Corollary 1: Consider N players with quadratic payoff functions (14) that implement the Nash seeking strategy (9)–(10) with frequencies satisfying the inequalities stated in Theorem 1. Then, the Nash equilibrium u^* is semi-globally practically asymptotically stable.

Proof: In the proof of Theorem 1, the average error system (22) is shown to be globally asymptotically stable. By [42, Theorem 2], with the error system (21) satisfying the theorem's conditions, the origin of (21) is semi-globally practically asymptotically stable. ■

For more details on semi-global convergence with extremum seeking controllers, the reader is referred to [29].

B. Symmetric Quadratic Games

If the matrix D is symmetric, we can develop a more precise expression for the convergence rate in Theorem 1. Specifically, we assume the following.

Assumption 3.2: $D^i_{ij} = D^j_{ji}$ for all $i, j \in \{1, \ldots, N\}$.

Under Assumptions 3.1 and 3.2, D is a negative definite symmetric matrix. Noncooperative games satisfying 3.2 are a class of games known as potential games [46, Th. 4.5]. In potential games, the maximization of the players' payoff functions J_i corresponds to the maximization of a global function $\Phi(u_1, \ldots, u_N)$, known as the potential function.

Corollary 2: Consider the system (9)–(10) with (14) under Assumptions 3.1 and 3.2 for an N-player game, where $\omega_i \neq \omega_j$, $2\omega_i \neq \omega_j$, and $\omega_i \neq \omega_j + \omega_k$ for all distinct i, j, $k \in \{1, \ldots, N\}$, and where ω_i/ω_j is rational for all i, $j \in \{1, \ldots, N\}$. The convergence properties of Theorem 1 hold with

$$M = \sqrt{\frac{\max_i\{k_i a_i^2\}}{\min_i\{k_i a_i^2\}}} \tag{28}$$

$$m = \frac{1}{2} \min_i\{k_i a_i^2\} \min_i \left\{-D^i_{ii} - \sum_{j\neq i}^{N} |D^i_{ij}|\right\}. \tag{29}$$

Proof: From the proof of Theorem 1, given any matrix $Q = Q^T > 0$, there exists a matrix $P = P^T > 0$ satisfying the Lyapunov equation $PA + A^T P = -Q$ since A, given by (23), is Hurwitz. Under Assumption 3.2, we select $Q = -D$ and obtain $P = \operatorname{diag}\left(\omega/k_1 a_1^2, \ldots, \omega/k_N a_N^2\right)$. Then, we directly have

$$\lambda_{\min}(P) = \frac{\omega}{\max_i\{k_i a_i^2\}},$$
$$\lambda_{\max}(P) = \frac{\omega}{\min_i\{k_i a_i^2\}} \tag{30}$$

and using the Gershgorin Circle Theorem [45, Th. 6.1.1], we obtain the bound

$$\lambda_{\min}(Q) \geq \min_i \left\{-D^i_{ii} - \sum_{j\neq i} |D^i_{ij}|\right\} \tag{31}$$

where we note that $D^i_{ii} < 0$. From (27),(30), and(31), we obtain the result. ■

Of note, the coefficient M in Corollary 2 is determined completely by the extremum seeking parameters k_i, a_i while the convergence rate m depends on both k_i, a_i, and the unknown game parameters D^i_{ij}. Thus, the convergence rate cannot be fully known since it depends on the unknown structure of the game.

C. Duopoly Price Game

Revisiting the duopoly example in Section II, we see that the profit functions (4) and (5) have the form (14). Moreover, this game satisfies both Assumptions 3.1 and 3.2 since

$$D = \frac{1}{p}\begin{bmatrix} -2 & 1 \\ 1 & -2 \end{bmatrix}.$$

By Theorem 1 and Corollary 2, the firms converge to a neighborhood of (u_1^*, u_2^*) (6)–(7) according to

$$\left|\begin{bmatrix} u_1(t) - u_1^* \\ u_2(t) - u_2^* \end{bmatrix}\right| \leq M e^{-mt} \left|\begin{bmatrix} u_1(0) - u_1^* \\ u_2(0) - u_2^* \end{bmatrix}\right| + O\left(\frac{1}{\omega} + \max(a_1, a_2)\right)$$

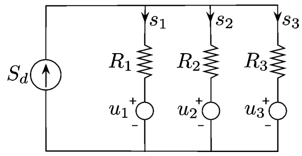


Fig. 3. Model of sales s_1, s_2, s_3 in a three-firm oligopoly with prices u_1, u_2, u_3 and total consumer demand S_d. The desirability of product i is proportional to $1/R_i$.

where M is given by (28) and $m = \min(k_1 a_1^2, k_2 a_2^2)/2p$ since $\lambda_{\min}(Q) = 1/p$.

D. Oligopoly Price Game

Consider a static noncooperative game with N firms in an oligopoly market structure that compete to maximize their profits by setting the price u_i of their product. As in the duopoly example in Section II, let m_i be the marginal cost of player i, s_i its sales volume, and J_i its profit, given by (1). The simplified sales model for the duopoly (2)–(3), where one firm has a clear advantage over the other in terms of consumer preference is no longer appropriate. Instead, we model the sales volume s_i as

$$s_i(t) = \frac{R_{\|}}{R_i}\left(S_d - \frac{u_i(t)}{\bar{R}_i} + \sum_{j\neq i}^{N} \frac{u_j(t)}{R_j}\right) \tag{32}$$

where S_d is the total consumer demand, $R_i > 0$ for all i, $1/R_{\|} = \left(\sum_{k=1}^{N} 1/R_k\right)$, and $1/\bar{R}_i = \left(\sum_{k\neq i}^{N} 1/R_k\right)$.

The sales model (32) is motivated by an analogous electric circuit, shown in Fig. 3, where S_d is an ideal current generator, u_i are ideal voltage generators, and most importantly, the resistors R_i represent the "resistance" that consumers have toward buying product i. This resistance may be due to quality or brand image considerations—the most desirable products have the lowest R_i. The sales in (32) are inversely proportional to R_i and grow as u_i decreases and as u_j, $j \neq i$, increases. The profit (1), in electrical analogy, corresponds to the power absorbed by the $u_i - m_i$ portion of the voltage generator i.

Proposition 1: There exists ω^* such that, for all $\omega > \omega^*$, the system (9)–(10) with (1) and (32) for the oligopoly price game, where $\omega_i \neq \omega_j$, $2\omega_i \neq \omega_j$ for all distinct $i, j = \{1, \ldots, N\}$ and where ω_i/ω_j is rational for all $i, j \in \{1, \ldots, N\}$, exponentially converges to a neighborhood of the Nash equilibrium:

$$u_i^* = \frac{\Pi R_i}{2R_i + \bar{R}_i}\left(\bar{R}_i S_d + m_i + \sum_{j=1}^{N} \frac{m_j \bar{R}_i - m_i \bar{R}_j}{2R_j + \bar{R}_j}\right) \tag{33}$$

where $\Pi^{-1} = 1 - \sum_{j=1}^{N} \bar{R}_j/(2R_j + \bar{R}_j) > 0$. Namely, if $|\Delta(0)|$ is sufficiently small, then for all $t \geq 0$

$$|\Delta(t)| \leq M e^{-mt}|\Delta(0)| + O\left(\frac{1}{\omega} + \max_i(a_i)\right) \tag{34}$$

where Δ is defined in Theorem 1, M is given by (28), $m = R_{\|}\min_i\{k_i a_i^2\}/(2\max_i\{R_i\Gamma_i\})$ and $\Gamma_i = \min_{j\in\{1,\ldots,N\}, j\neq i} R_j$.

Proof: Substituting (32) into (1) yields payoff functions

$$J_i(t) = \frac{R_{\|}}{R_i}\left(-\frac{u_i^2}{\bar{R}_i} + u_i \sum_{j\neq i}^{N}\frac{u_j}{R_j} + \left(\frac{m_i}{\bar{R}_i} + S_d\right)u_i - m_i \sum_{j\neq i}^{N}\frac{u_j}{R_j} - S_d m_i\right)$$

that are of the form (14). Therefore, the Nash equilibrium u^* satisfies $Du^* = -d$, where D and d are given by (16) and have elements

$$D_{ij}^i = \begin{cases} -\frac{2R_{\|}}{R_i \bar{R}_i}, & \text{if } i = j \\ \frac{R_{\|}}{R_i R_j}, & \text{if } i \neq j \end{cases}$$
$$d_j^j = \frac{m_j R_{\|}}{R_j \bar{R}_j} + \frac{S R_{\|}}{R_j}$$

for $i, j \in \{1, \ldots, N\}$. Assumption 3.1 is satisfied by D since

$$\sum_{j\neq i}^{N}\left|\frac{R_{\|}}{R_i R_j}\right| = \frac{R_{\|}}{R_i \bar{R}_i} < \left|-\frac{2R_{\|}}{R_i \bar{R}_i}\right|, \quad i \in \{1, \ldots, N\}.$$

Thus, the Nash equilibrium of this game exists, is unique, and can be shown to be (33). (The various parameters here are assumed to be selected such that u_i^* is positive for all i.) Moreover, $D_{ij}^i = D_{ji}^j$, so D is a negative definite symmetric matrix, satisfying Assumption 3.2. Then, by Theorem 1 and Corollary 2, we have the convergence bound (34) and obtain m by computing

$$\lambda_{\min}(Q) \geq \min_i\left\{\frac{2R_{\|}}{R_i \bar{R}_i} - \sum_{j\neq i}\left|\frac{R_{\|}}{R_i R_j}\right|\right\} = \frac{R_{\|}}{\max_i\{R_i \bar{R}_i\}}$$

where $Q = -D$, and by noting that $\max_i\{R_i \bar{R}_i\} < \max_i\{R_i \Gamma_i\}$. Finally, the error system (20) for this game does not contain any terms with the product $\mu_i(t)\mu_j(t)\mu_k(t)$, so the requirement that $\omega_i \neq \omega_j + \omega_k$ for all distinct i, j, $k \in \{1, \ldots, N\}$ does not arise when computing the average error system. ■

The resulting pricing, sales, and profit transients when four firms implement (9)–(10) are shown in Fig. 4 for a simulation with game parameters: $S_d = 100$, $R_1 = 0.15$, $R_2 = 0.3$, $R_3 = .6$, $R_4 = 1$, $m_1 = 30$, $m_2 = 30$, $m_3 = 25$, $m_4 = 20$; extremum seeking parameters: $a_1 = a_2 = a_3 = a_4 = 0.05$, $k_1 = 6$, $k_2 = 18$, $k_3 = 10$, $k_4 = 24$, $\omega_1 = 30$, $\omega_2 = 24$, $\omega_3 = 44$, $\omega_4 = 36$; and initial conditions: $u_1(0) = \hat{u}_1(0) = 52$, $u_2(0) = \hat{u}_2(0) = u_2^* = 40.93$, $u_3(0) = \hat{u}_3(0) = 33.5$, $u_4(0) = \hat{u}_4(0) = u_4^* = 35.09$.

E. N-Player Games With Quadratic Payoff Functions and Stubborn Players

An interesting scenario to consider is when not all the players utilize the Nash seeking strategy (9)–(10). Without loss of generality, we assume that players $i \in \{1, \ldots, n\}$ implement

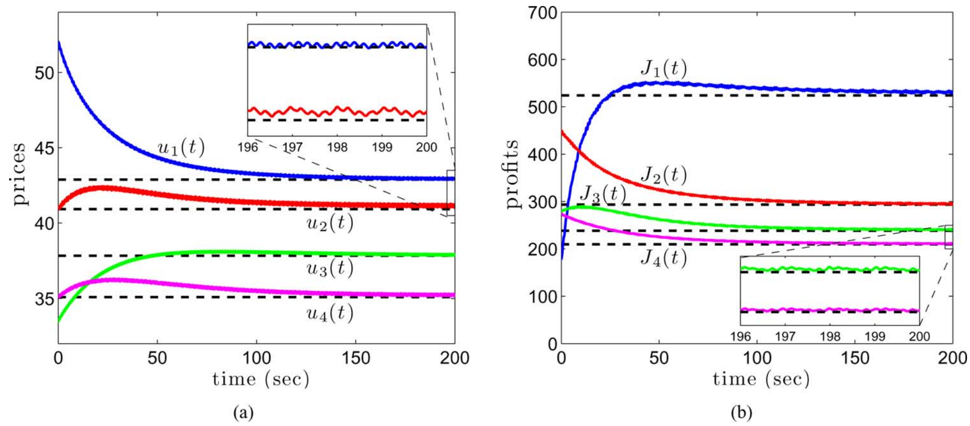


Fig. 4. (a) Price and (b) profit time histories of firms P1, P2, P3, and P4 when implementing the Nash seeking scheme (9)–(10). The dashed lines denote the values at the Nash equilibrium.

(9)–(10) while players $j \in \{n+1, \ldots, N\}$ are stubborn and use fixed actions

$$u_j(t) \equiv \bar{u}_j, \quad j \in \{n+1, \ldots, N\}. \tag{35}$$

To discuss the convergence of the Nash seeking players, we introduce their reaction curves [25, Def. 4.3], which are the players' best response strategy given the actions of the other players. When the players have quadratic payoff functions (14), their reaction curves are

$$l_i(u_{-i}) = -\frac{1}{D^i_{ii}} \left(\sum_{j \neq i}^{N} D^i_{ij} u_j + d^i_i \right) \tag{36}$$

where u_{-i} denotes the actions of the other $N-1$ players. In the presence of stubborn players, the remaining players' unique best response $u^{\mathrm{br}} = [u^{\mathrm{br}}_1, \ldots, u^{\mathrm{br}}_n]^T$ is given by

$$\begin{bmatrix} u^{\mathrm{br}}_1 \\ \vdots \\ u^{\mathrm{br}}_n \end{bmatrix} = -\begin{bmatrix} D^1_{11} & \cdots & D^1_{1n} \\ \vdots & \ddots & \\ D^n_{n1} & & D^n_{nn} \end{bmatrix}^{-1} \times \left(\begin{bmatrix} D^1_{1,n+1} & \cdots & D^1_{1N} \\ \vdots & \ddots & \\ D^n_{n,n+1} & & D^n_{nN} \end{bmatrix} \begin{bmatrix} \bar{u}_{n+1} \\ \vdots \\ \bar{u}_N \end{bmatrix} + \begin{bmatrix} d^1_1 \\ \vdots \\ d^n_n \end{bmatrix} \right). \tag{37}$$

When there are no stubborn players, $u^{\mathrm{br}} = u^*$.

Theorem 2: Consider the N-player game with (14) under Assumption 3.1, where players $i \in \{1, \ldots, n\}$ implement the strategy (9)–(10), with $\omega_i \neq \omega_j$, $2\omega_i \neq \omega_j$, and $\omega_i \neq \omega_j + \omega_k$ for all distinct i, j, $k \in \{1, \ldots, n\}$, and stubborn players $j \in \{n+1, \ldots N\}$ implement (35). There exist $\omega^*, M, m > 0$ such that for all $\omega > \omega^*$, if $|\Omega(0)|$ is sufficiently small, then for all $t \geq 0$

$$|\Omega(t)| \leq M e^{-mt} |\Omega(0)| + O\left(\frac{1}{\omega} + \max_i a_i \right) \tag{38}$$

where $\Omega(t) = [u_1(t) - u^{\mathrm{br}}_1, \ldots, u_n(t) - u^{\mathrm{br}}_n]^T$ and u^{br} is given by (37).

Proof: Following the proof of Theorem 1 for the n Nash seeking players, one can obtain

$$\frac{d}{d\tau} \begin{bmatrix} \tilde{u}^{\mathrm{ave}}_1 \\ \vdots \\ \tilde{u}^{\mathrm{ave}}_n \end{bmatrix} = \frac{1}{2\omega} \begin{bmatrix} \kappa_1 D^1_{11} & \cdots & \kappa_1 D^1_{1n} \\ \vdots & \ddots & \\ \kappa_n D^n_{n1} & & \kappa_n D^n_{nn} \end{bmatrix} \begin{bmatrix} \tilde{u}^{\mathrm{ave}}_1 \\ \vdots \\ \tilde{u}^{\mathrm{ave}}_n \end{bmatrix} + \frac{1}{2\omega} \begin{bmatrix} \kappa_1 D^1_{1,n+1} & \cdots & \kappa_1 D^1_{1N} \\ \vdots & \ddots & \\ \kappa_n D^n_{n,n+1} & & \kappa_n D^n_{nN} \end{bmatrix} \begin{bmatrix} \tilde{v}_{n+1} \\ \vdots \\ \tilde{v}_N \end{bmatrix} \tag{39}$$

where $\kappa_i = k_i a_i^2, i \in \{1, \ldots, n\}$ and the stubborn players' error relative to the Nash equilibrium is denoted by $\tilde{v}_j = \bar{u}_j - u^*_j, j \in \{n+1, \ldots, N\}$. The unique, exponentially stable equilibrium of (39) is

$$\begin{bmatrix} \tilde{u}^{\mathrm{e}}_1 \\ \vdots \\ \tilde{u}^{\mathrm{e}}_n \end{bmatrix} = -\begin{bmatrix} D^1_{11} & \cdots & D^1_{1n} \\ \vdots & \ddots & \\ D^n_{n1} & & D^n_{nn} \end{bmatrix}^{-1} \times \begin{bmatrix} D^1_{1,n+1} & \cdots & D^1_{1N} \\ \vdots & \ddots & \\ D^n_{n,n+1} & & D^n_{nN} \end{bmatrix} \begin{bmatrix} \tilde{v}_{n+1} \\ \vdots \\ \tilde{v}_N \end{bmatrix} \tag{40}$$

and so from the proof of Theorem 1 and [43, Th. 10.4], we have

$$|u(t) - \tilde{u}^{\mathrm{e}} - u^*| \leq M e^{-mt} |u(0) - \tilde{u}^{\mathrm{e}} - u^*| + O\left(\frac{1}{\omega} + \max_i a_i \right) \tag{41}$$

for players $i \in \{1, \ldots, n\}$. What remains to be shown is that $\tilde{u}^{\mathrm{e}}_i + u^*_i$ is the best response of player i.

From (35) and (41), the players' actions converge on average to $u^{\mathrm{ss}} = [\tilde{u}^{\mathrm{e}}_1 + u^*_1, \ldots, \tilde{u}^{\mathrm{e}}_n + u^*_n, \bar{u}_{n+1}, \ldots, \bar{u}_N]^T$. At u^{ss}, the best response of player $i \in \{1, \ldots, n\}$ is

$$\begin{aligned} l_i(u^{\mathrm{ss}}_{-i}) &= -\frac{1}{D^i_{ii}} \left(\sum_{j \neq i}^{n} D^i_{ij} (\tilde{u}^{\mathrm{e}}_j + u^*_j) + \sum_{j=n+1}^{N} D^i_{ij} \bar{u}_j + d^i_i \right) \\ &= -\frac{1}{D^i_{ii}} \left(\sum_{j \neq i}^{n} D^i_{ij} \tilde{u}^{\mathrm{e}}_j + \sum_{j=n+1}^{N} D^i_{ij} \tilde{v}_j \right) \\ &\quad - \frac{1}{D^i_{ii}} \left(\sum_{j \neq i}^{N} D^i_{ij} u^*_j + d^i_i \right) \end{aligned} \tag{42}$$

where we have substituted $\bar{u}_j = \tilde{v}_j + u_j^*$. Noting (36) and using (40) to substitute for $\tilde{v}_j$ yields

$$\begin{aligned} l_i(\tilde{u}_{-i}^{\mathrm{ss}}) &= -\frac{1}{D_{ii}^i}\left(\sum_{j\neq i}^{n} D_{ij}^i \tilde{u}_j^{\mathrm{e}} - \sum_{j=1}^{n} D_{ij}^i \tilde{u}_j^{\mathrm{e}}\right) + u_i^* \\ &= \tilde{u}_i^{\mathrm{e}} + u_i^*. \end{aligned} \tag{43}$$

Hence, the n Nash seeking players converge to their best response actions u^{br}. ■

To provide more insight into this result, we note that the average response of player i is

$$\begin{aligned} \frac{d}{d\tau}\hat{u}_i^{\mathrm{ave}}(\tau) &= \frac{k_i a_i^2}{2\omega}\sum_{j=1}^{N} D_{ij}^i \hat{u}_j^{\mathrm{ave}}(\tau) + d_i^i \\ &= \frac{k_i a_i^2}{2\omega}|D_{ii}^i|\left(l_i(\hat{u}_{-i}^{\mathrm{ave}}(\tau) - \hat{u}_i^{\mathrm{ave}}(\tau)\right) \end{aligned} \tag{44}$$

which is the scaled, continuous-time best response for games with quadratic payoff functions. Continuous-time best response dynamics are studied for more general scenarios in [47], [48].

IV. N-Player Games With Non-Quadratic Payoff Functions

Now consider a more general noncooperative game with N players and a dynamic mapping from the players' actions u_i to their payoff values J_i. Each player attempts to maximize the steady-state value of its payoff. Specifically, we consider a general nonlinear model

$$\dot{x} = f(x,u) \tag{45}$$

$$J_i = h_i(x), \quad i \in \{1,\dots,N\} \tag{46}$$

where $x \in \mathbb{R}^n$ is the state, $u \in \mathbb{R}^N$ is a vector of the players' actions, u_i is the action of player i, $J_i \in \mathbb{R}$ its payoff value, $f : \mathbb{R}^n \times \mathbb{R}^N \to \mathbb{R}^n$ and $h_i : \mathbb{R}^n \to \mathbb{R}$ are smooth, and h_i is a possibly non-quadratic function. Oligopoly games may possess nonlinear demand and cost functions [50], which motivate the inclusion of the dynamic system (45) in the game structure and the consideration of non-quadratic payoff functions. For this scenario, we pursue local convergence results since the payoff functions h_i may be non-quadratic and multiple, isolated Nash equilibria may exist. If the payoff functions are quadratic, semi-global practical stability can be achieved following the results of [29].

We make the following assumptions about this N-player game.

Assumption 4.1: There exists a smooth function $l : \mathbb{R}^N \to \mathbb{R}^n$ such that

$$f(x,u) = 0 \quad \text{if and only if} \quad x = l(u). \tag{47}$$

Assumption 4.2: For each $u \in \mathbb{R}^N$, the equilibrium $x = l(u)$ of (45) is locally exponentially stable.

Hence, we assume that for all actions, the nonlinear dynamic system is locally exponentially stable. We can relax the requirement that this assumption holds for each $u \in \mathbb{R}^N$ as we need to be only concerned with the action sets of the players, namely, $u \in U = U_1 \times \cdots \times U_N \subset \mathbb{R}^N$, and we do this in Section V for our numerical example. For notational convenience, we use this more restrictive case.

The following assumptions are central to our Nash seeking scheme as they ensure that at least one stable Nash equilibrium exists at steady state.

Assumption 4.3: There exists at least one, possibly multiple, isolated stable Nash equilibria $u^* = [u_1^*,\dots,u_N^*]$ such that, for all $i \in \{1,\dots,N\}$

$$\begin{aligned} \frac{\partial(h_i \circ l)}{\partial u_i}(u^*) &= 0 \\ \frac{\partial^2(h_i \circ l)}{\partial u_i^2}(u^*) &< 0. \end{aligned} \tag{48}$$

Assumption 4.4: The matrix

$$\Lambda = \begin{bmatrix} \frac{\partial^2(h_1\circ l)(u^*)}{\partial u_1^2} & \frac{\partial^2(h_1\circ l)(u^*)}{\partial u_1 \partial u_2} & \cdots & \frac{\partial^2(h_1\circ l)(u^*)}{\partial u_1 \partial u_N} \\ \frac{\partial^2(h_2\circ l)(u^*)}{\partial u_1 \partial u_2} & \frac{\partial^2(h_2\circ l)(u^*)}{\partial u_2^2} & & \\ \vdots & & \ddots & \\ \frac{\partial^2(h_N\circ l)(u^*)}{\partial u_1 \partial u_N} & & & \frac{\partial^2(h_N\circ l)(u^*)}{\partial u_N^2} \end{bmatrix} \tag{49}$$

is strictly diagonally dominant and hence, nonsingular.

By Assumptions 4.3 and 4.4, Λ is Hurwitz.

As with the games considered in Sections II and III, each player converges to a neighborhood of u^* by implementing the extremum seeking strategy (9)–(10) to evolve its action u_i according to the measured value of its payoff J_i. Unlike the previous games, however, we select the parameters $k_i = \varepsilon\omega K_i = O(\varepsilon\omega)$, where ε, ω are small, positive constants and ω is related to the players' frequencies by (11). Intuitively, ω is small since the players' actions should evolve more slowly than the dynamic system, creating an overall system with two time scales. In contrast, our earlier analysis assumed $1/\omega$ to be small, which can be seen as the limiting case where the dynamic system is infinitely fast and allows ω to be large.

Formulating the error system clarifies why these parameter selections are made. The error relative to the Nash equilibrium is denoted by (19), which in the time scale $\tau = \omega t$, leads to

$$\omega\frac{dx}{d\tau} = f(x, u^* + \tilde{u} + \mu(\tau)) \tag{50}$$

$$\frac{d\tilde{u}_i}{d\tau} = \varepsilon K_i \mu_i(\tau) h_i(x), \quad i \in \{1,\dots,N\} \tag{51}$$

where $\tilde{u} = [\tilde{u}_1,\dots,\tilde{u}_N]$, $\mu(\tau) = [\mu_1(\tau),\dots,\mu_N(\tau)]$, and $\mu_i(\tau) = a_i \sin(\bar{\omega}_i \tau + \varphi_i)$. The system (50)–(51) is in the standard singular perturbation form with ω as a small parameter. Since ε is also small, we analyze (50)–(51) using the averaging theory for the quasi-steady state of (50), followed by the use of the singular perturbation theory for the full system.

A. Averaging Analysis

For the averaging analysis, we first "freeze" x in (50) at its quasi-steady state $x = l(u^* + \tilde{u} + \mu(\tau))$, which we substitute into (51) to obtain the "reduced system"

$$\frac{d\tilde{u}_i}{d\tau} = \varepsilon K_i \mu_i(\tau)(h_i \circ l)(u^* + \tilde{u} + \mu(\tau)). \tag{52}$$

This system is in the form to apply averaging theory [43] and leads to the following result.

Theorem 3: Consider the system (52) for an N-player game under Assumptions 4.3 and 4.4, where $\bar{\omega}_i \neq \bar{\omega}_j$, $\bar{\omega}_i \neq \bar{\omega}_j + \bar{\omega}_k$, $2\bar{\omega}_i \neq \bar{\omega}_j + \bar{\omega}_k$, and $\bar{\omega}_i \neq 2\bar{\omega}_j + \bar{\omega}_k$ for all distinct i, j, $k \in \{1, \ldots, N\}$ and $\bar{\omega}_i$ is rational for all $i \in \{1, \ldots, N\}$. There exist parameters $\Xi, \xi > 0$ and ε^*, a^* such that, for all $\varepsilon \in (0, \varepsilon^*)$ and $a_i \in (0, a^*)$, if $|\Theta(0)|$ is sufficiently small, then for all $\tau \geq 0$

$$|\Theta(\tau)| \leq \Xi e^{-\xi\tau}|\Theta(0)| + O\left(\varepsilon + \max_i a_i^3\right) \tag{53}$$

where $\Theta(\tau) = \left[\tilde{u}_1(\tau) - \sum_{j=1}^N c_{jj}^1 a_j^2, \ldots, \tilde{u}_N(\tau) - \sum_{j=1}^N c_{jj}^N a_j^2\right]^T$, and

$$\mathbf{c}_j = -\frac{1}{4}\Lambda^{-1}\mathbf{g}_j \tag{54}$$

$$\mathbf{c}_j \triangleq \left[c_{jj}^1, \ldots, c_{jj}^{j-1}, c_{jj}^j, c_{jj}^{j+1}, \ldots, c_{jj}^N\right]^T \tag{55}$$

$$\mathbf{g}_j \triangleq \left[g_j^1, \ldots, g_j^{j-1}, g_j^j, g_j^{j+1}, \ldots, g_j^N\right]^T \tag{56}$$

$$g_j^i = \begin{cases} \frac{\partial^3 (h_k \circ l)(u^*)}{\partial u_k \partial u_j^2}, & \text{if } i \neq j \\ \frac{1}{2}\frac{\partial^3 (h_j \circ l)(u^*)}{\partial u_j^3}, & \text{if } i = j. \end{cases} \tag{57}$$

Proof: As already noted, the form of (52) allows for the application of averaging theory, which yields the average error system

$$\frac{d\tilde{u}_i^{\text{ave}}}{d\tau} = \varepsilon\left(\frac{K_i}{T}\int_0^T \mu_i(\tau)(h_i \circ l)(u^* + \tilde{u}^{\text{ave}} + \mu(\tau))\, d\tau\right). \tag{58}$$

The equilibrium $\tilde{u}^{\text{e}} = [\tilde{u}_1^{\text{e}}, \ldots, \tilde{u}_N^{\text{e}}]$ of (58) satisfies

$$0 = \frac{1}{T}\int_0^T \mu_i(\tau)(h_i \circ l)(u^* + \tilde{u}^{\text{e}} + \mu(\tau))\, d\tau \tag{59}$$

for all $i \in \{1, \ldots, N\}$, and we postulate that $\tilde{u}^{\text{e}}$ has the form

$$\tilde{u}_i^{\text{e}} = \sum_{j=1}^N b_j^i a_j + \sum_{j=1}^N \sum_{k \geq j}^N c_{jk}^i a_j a_k + O\left(\max_i a_i^3\right). \tag{60}$$

By approximating $(h_i \circ l)$ about u^* in (59) with a Taylor polynomial and substituting (60), the unknown coefficients b_j^i and c_{jk}^i can be determined.

To capture the effect of higher order derivatives on average error system's equilibrium, we use the Taylor polynomial approximation [49], which requires $(h_i \circ l)$ to be $k+1$ times differentiable

$$\begin{aligned}
&(h_i \circ l)(u^* + \tilde{u}^{\text{e}} + \mu(\tau)) \\
&\quad = \sum_{|\alpha|=0}^{k} \frac{D^\alpha (h_i \circ l)(u^*)}{\alpha!}(\tilde{u}^{\text{e}} + \mu(\tau))^\alpha \\
&\qquad + \sum_{|\alpha|=k+1} \frac{D^\alpha (h_i \circ l)(\zeta)}{\alpha!}(\tilde{u}^{\text{e}} + \mu(\tau))^\alpha \\
&\quad = \sum_{|\alpha|=0}^{k} \frac{D^\alpha (h_i \circ l)(u^*)}{\alpha!}(\tilde{u}^{\text{e}} + \mu(\tau))^\alpha \\
&\qquad + O\left(\max_i a_i^{k+1}\right)
\end{aligned} \tag{61}$$

where ζ is a point on the line segment that connects the points u^* and $u^* + \tilde{u}^{\text{e}} + \mu(\tau)$. In (61), we have used multi-index notation, namely, $\alpha = (\alpha_1, \ldots, \alpha_N)$, $|\alpha| = \alpha_1 + \cdots + \alpha_N$, $\alpha! = \alpha_1! \cdots \alpha_N!$, $u^\alpha = u_1^{\alpha_1} \cdots u_N^{\alpha_N}$, and $D^\alpha (h_i \circ l) = \partial^{|\alpha|}(h_i \circ l)/\partial u_1^{\alpha_1} \cdots \partial u_N^{\alpha_N}$. The second term on the last line of (61) follows by substituting the postulated form of $\tilde{u}^{\text{e}}$(60).

For this analysis, we choose $k = 3$ to capture the effect of the third order derivative on the system as a representative case. Higher order estimates of the bias can be pursued if the third order derivative is zero. Substituting (61) into (59) and computing the average of each term gives

$$\begin{aligned}
0 = &\frac{a_i^2}{2} \\
&\times \Bigg[\tilde{u}_i^{\text{e}} \frac{\partial^2 (h_i \circ l)(u^*)}{\partial u_i^2} + \sum_{j \neq i}^N \tilde{u}_j^{\text{e}} \frac{\partial^2 (h_i \circ l)(u^*)}{\partial u_i \partial u_j} \\
&\quad + \left(\frac{1}{2}(\tilde{u}_i^{\text{e}})^2 + \frac{a_i^2}{8}\right)\frac{\partial^3 (h_i \circ l)(u^*)}{\partial u_i^3} \\
&\quad + \tilde{u}_i^{\text{e}} \sum_{j \neq i}^N \tilde{u}_j^{\text{e}} \frac{\partial^3 (h_i \circ l)(u^*)}{\partial u_i^2 \partial u_j} \\
&\quad + \sum_{j \neq i}^N \left(\frac{1}{2}(\tilde{u}_j^{\text{e}})^2 + \frac{a_j^2}{4}\right)\frac{\partial^3 (h_i \circ l)(u^*)}{\partial u_i \partial u_j^2} \\
&\quad + \sum_{j \neq i}^N \sum_{\substack{k > j \\ k \neq i}}^N \tilde{u}_j^{\text{e}} \tilde{u}_k^{\text{e}} \frac{\partial^3 (h_i \circ l)(u^*)}{\partial u_i \partial u_j \partial u_k}\Bigg] \\
&+ O(\max_i a_i^5)
\end{aligned} \tag{62}$$

where we have noted (48), utilized (60), and computed the integrals shown in Appendices I and II. Substituting (60) into (62) and matching first order powers of a_i gives

$$\begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} = a_1 \Lambda \begin{bmatrix} b_1^1 \\ \vdots \\ b_1^N \end{bmatrix} + \cdots + a_N \Lambda \begin{bmatrix} b_1^1 \\ \vdots \\ b_N^N \end{bmatrix}$$

which implies that $b_j^i = 0$ for all i, j since Λ is nonsingular by Assumption 4.4. Similarly, matching second-order terms of a_i, and substituting $b_j^i = 0$ to simplify the resulting expressions, yields

$$\begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} = \sum_{j=1}^{N}\sum_{k>j}^{N} a_j a_k \Lambda \begin{bmatrix} c_{jk}^1 \\ \vdots \\ c_{jk}^N \end{bmatrix} + \sum_{j=1}^{N} a_j^2 \left(\Lambda \begin{bmatrix} c_{jj}^1 \\ \vdots \\ c_{jj}^N \end{bmatrix} + \frac{1}{4}\mathbf{g}_j \right)$$

where $\mathbf{g}_j$ is defined in (56). Thus, $c_{jk}^i = 0$ for all i, j, k when $j \neq k$, and c_{jj}^i is given by (55) for $i, j \in \{1, \ldots, N\}$. The equilibrium of the average system is then

$$\tilde{u}_i^{\mathrm{e}} = \sum_{j=1}^{N} c_{jj}^i a_j^2 + O\left(\max_i a_i^3\right). \tag{63}$$

By again utilizing a Taylor polynomial approximation, one can show that the Jacobian $\Psi^{\mathrm{ave}} = [\psi_{i,j}]_{N\times N}$ of (58) at $\tilde{u}^{\mathrm{e}}$ has elements given by

$$\begin{aligned} \psi_{i,j} &= \varepsilon \frac{K_i}{T} \int_0^T \mu_i(\tau) \frac{\partial (h_i \circ l)}{\partial u_j}(u^* + \tilde{u}^{\mathrm{e}} + \mu(\tau))\, d\tau \\ &= \frac{1}{2}\varepsilon K_i a_i^2 \frac{\partial^2 (h_i \circ l)}{\partial u_i \partial u_j}(u^*) + O\left(\varepsilon \max_i a_i^3\right). \end{aligned} \tag{64}$$

By Assumptions 4.3 and 4.4, Ψ^{ave} is Hurwitz for sufficiently small a_i, which implies that the equilibrium (63) of the average error system (58) is locally exponentially stable, i.e., there exist constants Ξ, $\xi > 0$ such that $|\tilde{u}^{\mathrm{ave}}(\tau) - \tilde{u}^{\mathrm{e}}| \leq \Xi e^{-\xi\tau} |\tilde{u}^{\mathrm{ave}}(0) - \tilde{u}^{\mathrm{e}}|$, which with [43, Th. 10.4] implies

$$|\tilde{u}(\tau) - \tilde{u}^{\mathrm{e}}| \leq \Xi e^{-\xi\tau} |\tilde{u}(0) - \tilde{u}^{\mathrm{e}}| + O(\varepsilon) \tag{65}$$

provided $\tilde{u}(0)$ is sufficiently close to $\tilde{u}^{\mathrm{ave}}(0)$. Defining $\Theta(\tau)$ as in Theorem 3 completes the proof. ■

From Theorem 3, we see that u of reduced system (52) converges to a region that is biased away from the Nash equilibrium u^*. This bias is in proportion to the perturbation magnitudes a_i and the third derivatives of the payoff functions, which are captured by the coefficients c_{jj}^i. Specifically, $\hat{u}_i$ of the reduced system converges to $u_i^* + \sum_{j=1}^{N} c_{jj}^i a_j^2 + O(\varepsilon + \max_i a_i^3)$ as $t \to \infty$.

Theorem 3 can be viewed as a generalization of the Theorem 1, but with a focus on the error system to highlight the effect of the payoff functions' non-quadratic terms on the players' convergence. This emphasis is needed because u_i of the reduced system converges to an $O(\varepsilon + \max_i a_i)$-neighborhood of u_i^*, as in the quadratic payoff case, since $u_i = \hat{u}_i + \mu_i$.

B. Singular Perturbation Analysis

We analyze the full system (50)–(51) in the time scale $\tau = \omega t$ using singular perturbation theory [43]. In Section IV-A, we analyzed the reduced model (52) and now, must study the boundary layer model to state our convergence result. First, however, we translate the equilibrium of the reduced model to the origin by defining $z_i = \tilde{u}_i - \tilde{u}_i^p$, where by [43, Th. 10.4], $\tilde{u}_i^p$ is a unique, exponentially stable, T-periodic solution $\tilde{u}^p = [\tilde{u}_1^p, \ldots, \tilde{u}_N^p]$ such that

$$\frac{d\tilde{u}_i^p}{d\tau} = \varepsilon K_i \mu_i(\tau)(h_i \circ l)(u^* + \tilde{u}^p + \mu(\tau)). \tag{66}$$

In this new coordinate system, we have for $i \in \{1, \ldots, N\}$

$$\begin{aligned} \frac{dz_i}{d\tau} &= \varepsilon K_i \mu_i(\tau) \\ &\quad \times [h_i(x) - (h_i \circ l)(u^* + \tilde{u}^p + \mu(\tau))] \qquad (67) \\ \omega \frac{dx}{d\tau} &= f(x, u^* + z + \tilde{u}^p + \mu(\tau)) \qquad (68) \end{aligned}$$

which from Assumption 4.1 has the quasi-steady state $x = l(u^* + z + \tilde{u}^p + \mu(\tau))$, and consequently, the reduced model in the new coordinates is

$$\begin{aligned} \frac{dz_i}{d\tau} = \varepsilon K_i \mu_i(\tau) & [(h_i \circ l)(u^* + z + \tilde{u}^p + \mu(\tau)) \\ & - (h_i \circ l)(u^* + \tilde{u}^p + \mu(\tau))] \end{aligned} \tag{69}$$

which has an equilibrium at $z = 0$ that is exponentially stable for sufficiently small a_i.

To formulate the boundary layer model, let $y = x - l(u^* + z + \tilde{u}^p + \mu(\tau))$, and then in the time scale $t = \tau/\omega$, we have

$$\begin{aligned} \frac{dy}{dt} &= f(y + l(u^* + z + \tilde{u}^p + \mu(\tau)), u^* + z + \tilde{u}^p + \mu(\tau)) \\ &= f(y + l(u), u) \end{aligned} \tag{70}$$

where $u = u^* + \tilde{u} + \mu(\tau)$ should be viewed as a parameter independent of the time variable t. Since $f(l(u), u) = 0$, $y = 0$ is an equilibrium of (70) and is exponentially stable by Assumption 4.2.

With ω as a singular perturbation parameter, we apply Tikhonov's Theorem on the Infinite Interval [43, Th. 11.2] to (67)–(68), which requires the origin to be an exponentially stable equilibrium point of both the reduced model (69) and the boundary layer model (70) and leads to the following:

- the solution $z(\tau)$ of (67) is $O(\omega)$-close to the solution $\bar{z}(\tau)$ of the reduced model (69), so
- the solution $\tilde{u}(\tau)$ of (51) converges exponentially to an $O(\omega)$-neighborhood of the T-periodic solution $\tilde{u}^p(\tau)$, and
- the T-periodic solution $\tilde{u}^p(\tau)$ is $O(\varepsilon)$-close to the equilibrium $\tilde{u}^{\mathrm{e}}$.

Hence, as $t \to \infty$, $\tilde{u}(\tau)$ converges to an $O(\omega + \varepsilon)$-neighborhood of $\tilde{u}^{\mathrm{e}} = \left[\sum_{j=1}^{N} c_{jj}^1 a_j^2, \ldots, \sum_{j=1}^{N} c_{jj}^N a_j^2\right] + O(\max_i a_i^3)$. Since $u(\tau) - u^* = \tilde{u}(\tau) + \mu(\tau) = \tilde{u}(\tau) + O(\max_i a_i)$, $u(\tau)$ converges to an $O(\omega + \varepsilon + \max_i a_i)$-neighborhood of u^*.

Also from Tikhonov's Theorem on the Infinite Interval, the solution $x(\tau)$ of (68), which is the same as the solution of (50), satisfies

$$x(\tau) - l(u^* + \tilde{u}^{\mathrm{r}}(\tau) + \mu(\tau)) - y(t) = O(\omega) \tag{71}$$

where $\tilde{u}^{\mathrm{r}}(\tau)$ is the solution of the reduced model (52) and $y(t)$ is the solution of the boundary layer model (70). Rearranging terms and subtracting $l(u^*)$ from both sides yields

$$x(\tau) - l(u^*) = O(\omega) + l(u^* + \tilde{u}^{\mathrm{r}}(\tau) + \mu(\tau)) - l(u^*) + y(t). \tag{72}$$

After noting that

- $\tilde{u}^{\mathrm{r}}(\tau)$ converges exponentially to $\tilde{u}^p(\tau)$, which is $O(\varepsilon)$-close to the equilibrium $\tilde{u}^{\mathrm{e}}$,
- $\mu(\tau)$ is $O(\max_i a_i)$, and
- $y(t)$ is exponentially decaying,

we conclude that $x(\tau) - l(u^*)$ exponentially converges to an $O(\omega + \varepsilon + \max_i a_i)$-neighborhood of the origin. Thus, $J_i = h_i(x)$ exponentially converges to an $O(\omega+\varepsilon+\max_i a_i)$-neighborhood of the payoff value $(h_i \circ l)(u^*)$.

We summarize with the following theorem.

Theorem 4: Consider the system (45)–(46) with (9)–(10) for an N-player game under Assumptions 4.1–4.4, where $\omega_i \neq \omega_j$, $\omega_i \neq \omega_j+\omega_k$, $2\omega_i \neq \omega_j+\omega_k$, and $\omega_i \neq 2\omega_j+\omega_k$ for all distinct $i, j, k \in \{1,\ldots,N\}$ and where ω_i/ω_j is rational for all $i, j \in \{1,\ldots,N\}$. There exists $\omega^* > 0$ and for any $\omega \in (0,\omega^*)$ there exist ε^*, $a^* > 0$ such that for the given ω and any $\varepsilon \in (0,\varepsilon^*)$ and $\max_i a_i \in (0,a^*)$, the solution $(x(t), u_1(t),\ldots,u_N(t))$ converges exponentially to an $O(\omega+\varepsilon+\max_i a_i)$-neighborhood of the point $(l(u^*), u_1^*,\ldots,u_N^*)$, provided the initial conditions are sufficiently close to this point.

V. Numerical Example With Non-Quadratic Payoff Functions

For an example non-quadratic game with players that employ the extremum seeking strategy (9)–(10), we consider the system

$$\begin{aligned}
\dot{x}_1 &= -4x_1 + x_1x_2 + u_1\\
\dot{x}_2 &= -4x_2 + u_2\\
J_1 &= -16x_1^2 + 8x_1^2x_2 - x_1^2x_2^2 - 6x_1x_2^2\\
&\quad + \left(24 + \frac{5}{32}\right)x_1x_2 - \frac{5}{8}x_1\\
J_2 &= -64x_2^3 + 48x_1x_2 - 12x_1x_2^2
\end{aligned}$$

whose equilibrium state is given by $(\bar{x}_1, \bar{x}_2) = (4u_1/(16-u_2), u_2/4)$. The Jacobian at the equilibrium $(\bar{x}_1,\bar{x}_2)$ is Hurwitz if $u_2 < 16$. Thus, the $(\bar{x}_1,\bar{x}_2)$ is locally exponentially stable, but not for all $(u_1,u_2) \in \mathbb{R}^2$, violating Assumption 4.2. However, as noted earlier, this restrictive requirement of local exponential stability for all $u \in \mathbb{R}^N$ was done merely for notational convenience, and we actually only require this assumption to hold for the players' action sets. Hence, in this example, we restrict the players' actions to the set $U = \{(u_1,u_2) \in \mathbb{R}^2 \mid u_1, u_2 \geq 0, u_2 < 16\}$.

At $x = \bar{x}$, the payoff functions are

$$\begin{aligned}
J_1 &= -u_1^2 + \frac{3}{2}u_1u_2 - \frac{5}{32}u_1\\
J_2 &= -u_2^3 + 3u_1u_2
\end{aligned} \tag{73}$$

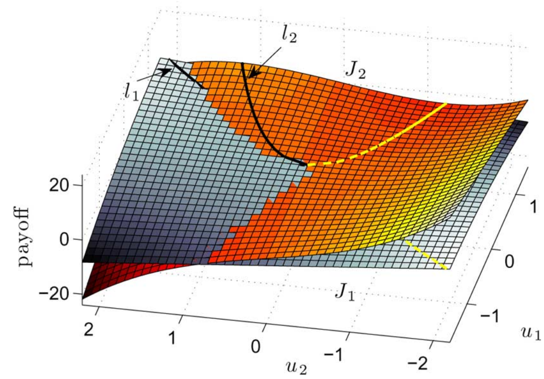


Fig. 5. Steady-state payoff function surfaces (73) with both their associated reaction curves (black) (74)–(75), which lie in the action set U, and the extremals $\partial J_i/\partial u_i = 0$ (dashed yellow), which lie outside of U, superimposed.

which have the associated reaction curves, defined on the action set U, shown as follows:

$$l_1(u_2) = \begin{cases} \frac{3}{4}u_2 - \frac{5}{64}, & \text{if } \frac{5}{48} \leq u_2 < 16\\ 0, & \text{if } 0 \leq u_2 < \frac{5}{48}\end{cases} \tag{74}$$

$$l_2(u_1) = \sqrt{u_1}. \tag{75}$$

Fig. 5 depicts each player's payoff surface with both the reaction curves and the extremals $\partial J_i/\partial u_i = 0$ that lie outside of U superimposed. These reaction curves have two intersections in the interior of U [Fig. 6(a)], which correspond to two Nash equilibria: $(u_1^*, u_2^*) = (25/64, 5/8)$ and $(v_1^*, v_2^*) = (1/64, 1/8)$. At u^*, Assumptions 4.3 and 4.4 are satisfied, implying the stability of u^*, whereas at v^*, these assumptions are violated—as they must be—since further analysis will show that v^* is an unstable Nash equilibrium. The reaction curves also intersect on the action set boundary ∂U at $(0,0)$, which means this game has a Nash equilibrium on the boundary that the players may seek depending on the game's initial conditions. To ensure the players remain in the action set, we employ a modified Nash seeking strategy that utilizes projection [51]. For this example, we will first present simulation results for Nash seeking in the interior of U before considering the Nash equilibrium on the boundary.

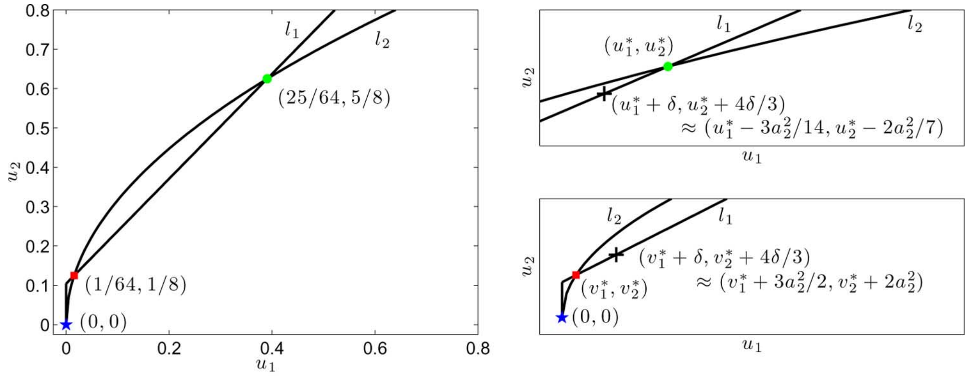
A. Nash Seeking in the Interior

For the Nash equilibrium points in the interior of U, we want to determine their stability and to quantify the convergence bias due to the non-quadratic payoff function. First, we compute the average error system for the reduced model according to (58), obtaining

$$\frac{d\tilde{u}_1^{\mathrm{ave}}}{d\tau} = \varepsilon K_1 a_1^2\left(-\tilde{u}_1^{\mathrm{ave}} + \frac{3}{4}\tilde{u}_2^{\mathrm{ave}}\right) \tag{76}$$

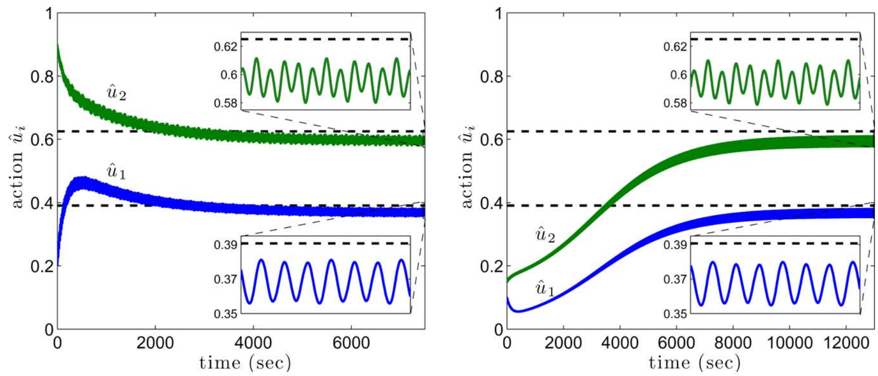
$$\frac{d\tilde{u}_2^{\mathrm{ave}}}{d\tau} = \varepsilon K_2 a_2^2\left(\frac{3}{2}\tilde{u}_1^{\mathrm{ave}} - 3\eta^*\tilde{u}_2^{\mathrm{ave}} - \frac{3}{2}(\tilde{u}_2^{\mathrm{ave}})^2 - \frac{3}{8}a_2^2\right) \tag{77}$$

where $\eta^* = u_2^*$ or v_2^*. The equilibria of (76)–(77) are given by $(\tilde{u}_1^{\mathrm{e}}(\eta^*), \tilde{u}_2^{\mathrm{e}}(\eta^*)) = \big((9 - 24\eta^*)/32 \pm 3\sqrt{(3-8\eta^*)^2 - 16a_2^2}/32,\ 4\tilde{u}_1^{\mathrm{e}}(\eta^*)/3\big)$. While the system appears to have four equilibria, two for each value of η^*, two equilibria correspond to the difference between the two



(a) (b)

Fig. 6. (a) Reaction curves associated with the steady-state payoff functions (73) with the stable Nash equilibrium (green circle), the unstable Nash equilibrium (red square), and the Nash equilibrium on the boundary (blue star), and (b) the convergence bias $(+)$ relative to each interior Nash equilibrium, due to the non-quadratic payoff function of player 2 and lies along the reaction curve of player 1.



(a) (b)

Fig. 7. Time history of the two-player game initialized at (a) $(u_1, u_2) = (0.2, 0.9)$, and (b) $(u_1, u_2) = (0.1, 0.15)$.

Nash equilibrium points, meaning that in actuality, only two equilibria exist and can be written as

$$(\tilde{u}_1^{\rm e}(\eta^*), \tilde{u}_2^{\rm e}(\eta^*)) = \left(\delta(\eta^*), \frac{4\delta(\eta^*)}{3}\right) \tag{78}$$

where $\delta(\eta^*) = (9 - 24\eta^*)/32 - \mathrm{sgn}(3 - 8\eta^*)3\sqrt{(3-8\eta^*)^2 - 16a_2^2}/32$ is the smallest value of $\tilde{u}_1^{\rm e}(\eta^*)$ for each Nash equilibrium. In the sequel, we omit the dependency of $\tilde{u}_1^{\rm e}$, $\tilde{u}_2^{\rm e}$, and δ on η^* for conciseness. As seen in Fig. 6(b), the equilibrium is biased away from the reaction curve of player 2 but still lies on the reaction curve l_1 of player 1 since only the payoff function for player 2 is non-quadratic.

The Jacobian $\Psi^{\rm ave}$ of the average error system for the reduced model is

$$\Psi^{\rm ave} = \begin{bmatrix} -\kappa_1 & \frac{3}{4}\kappa_1 \\ \frac{3}{2}\kappa_2 & -\kappa_2\left(4\delta + 3\eta^*\right) \end{bmatrix}$$

when evaluated at $(\tilde{u}_1^{\rm e}, \tilde{u}_2^{\rm e}) = (\delta, 4\delta/3)$, where $\kappa_1 = \varepsilon K_1 a_1^2$ and $\kappa_2 = \varepsilon K_2 a_2^2$. Its characteristic equation is

$$\lambda^2 + \underbrace{\left(\kappa_1 + \kappa_2\left(4\delta + 3\eta^*\right)\right)}_{\zeta_1}\lambda + \underbrace{\kappa_1\kappa_2\left(4\delta + 3\eta^* - \frac{9}{8}\right)}_{\zeta_2} = 0.$$

Thus, $\Psi^{\rm ave}$ is Hurwitz if and only if ζ_1 and ζ_2 are positive. For sufficiently small a_2 so that $\delta \approx 0$, $\zeta_1, \zeta_2 > 0$ when $\eta^* = u_2^* = 5/8$, and $\zeta_2 < 0$ when $\eta^* = v_2^* = 1/8$, which implies u^* is a stable Nash equilibrium and v^* is unstable. Closer analysis shows that the Jacobian associated with v^* has two real eigenvalues—one that is negative and one that is positive.

For the simulations, we select $k_1 = k_2 = 1.75$, $a_1 = 0.06$, $a_2 = 0.03$, $\omega_1 = 1.2$, and $\omega_2 = 2$, where the parameters are chosen to be small, in particular the perturbation frequencies ω_i, since the perturbation must occur at a time scale that is slower than the fast time scale of the nonlinear system. Fig. 7(a) and (b) depicts the evolution of the players' actions $\hat{u}_1$ and $\hat{u}_2$ initialized at $(u_1(0), u_2(0)) = (\hat{u}_1(0), \hat{u}_2(0)) = (0.2, 0.9)$ and $(0.1, 0.15)$. The state (x_1, x_2) is initialized at the origin in both cases. We show $\hat{u}_i$ instead of u_i to better illustrate the convergence of the players' actions to a neighborhood near—but biased away from—the Nash strategies since u_i contains the additive signal $\mu_i(t)$. In both scenarios, the average actions of both player 1 and player 2 lie below u^*, which is consistent with the equilibrium of the reduced model's average error system (78).

The slow initial convergence in Fig. 7(b) can be explained by examining the phase portrait of the average of the reduced model $\hat{u}$-system. The initial condition $(0.1, 0.15)$ lies near the unstable equilibrium, causing the slow initial convergence seen in Fig. 7(b). In Fig. 8, the stable and unstable interior Nash equilibria (green circle and red square), the Nash equilibrium on ∂U (blue star), and the eigenvectors (magenta arrows) associated with each interior Nash equilibrium are denoted. The

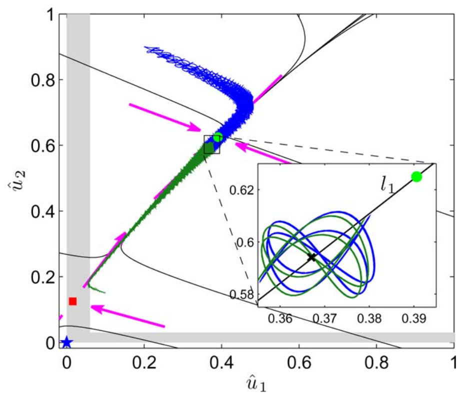


Fig. 8. Phase portrait of the $\hat{u}$-reduced model average system, with the players' $\hat{u}$-trajectories superimposed for each initial condition. The stable and unstable Nash equilibrium points (green circle and red square), their eigenvectors (magenta arrows), and Nash equilibrium on ∂U (blue star) are denoted. The shaded region is the set difference $U \setminus \hat{U}$. If $\hat{u}_i \in \hat{U}$, then $u_i \in U$. The zoomed-in area in shows the players' convergence to an almost-periodic orbit that is biased away from the stable Nash equilibrium and centered on the reaction curve l_1. The average actions of the players are marked by $\times$.

eigenvectors are, in general, neither tangent nor perpendicular to the reaction curves at the Nash equilibrium points. The phase portrait and orientation of the eigenvectors do depend on the values of a_i and k_i, but the stability of each Nash equilibrium is invariant to their values, provided these parameters are positive and small as shown in Sections III and IV. Since Fig. 8 is a phase portrait of the $\hat{u}^{\rm ave}$-reduced system, we include the shaded regions to show that $\hat{u}_i$ must be restricted to the set $\hat{U} = \{(\hat{u}_1, \hat{u}_2) \in \mathbb{R}^2 \mid a_1 \leq \hat{u}_1, a_2 \leq \hat{u}_2 \leq 16 - a_2\}$ to ensure that u_i lies in U. Hence, the shaded region's outer edge is ∂U, the inner edge is $\partial \hat{U}$, and the shaded interior is the set difference $U \setminus \hat{U}$.

When the players' initial conditions lie in the stable region of the phase portrait, the players' actions remain in U and converge to a neighborhood of the stable Nash equilibrium in the interior of U. The zoomed-in area highlights the convergence of the trajectories to an almost-periodic orbit that is biased away from the Nash equilibrium u^*. The depicted orbit consists of the last 300 s of each trajectory with the players' average actions denoted by $\times$. These average actions lie on the reaction curve l_1 as predicted by (78).

B. Nash Seeking With Projection

To ensure that a player's action remains in the action set U, we employ a modified extremum seeking strategy that utilizes projection. Namely, the players implement the following strategy:

$$\dot{\hat{u}}_i(t) = k_i \mathrm{Proj}\{\mu_i(t) J_i(t); a_i\} \tag{79}$$
$$u_i(t) = \hat{u}_i(t) + \mu_i(t) \tag{80}$$

where $\mu_i(t)$ is defined as in (9)–(10) and

$$\mathrm{Proj}\{\phi; \chi\} = \begin{cases} 0, & \text{if } \hat{u}_i \leq \chi \text{ and } \phi < 0 \\ \phi, & \text{otherwise.} \end{cases} \tag{81}$$

Lemma 1: The following properties hold for the projection operator (81).

1) For all $(\phi, \hat{u}_i) \in \mathbb{R}^2$, $|\mathrm{Proj}\{\phi; \chi\}| \leq |\phi|$.
2) With $\hat{u}_i(0) \geq \chi$, the update law (79) with the projection operator (81) guarantees that $\hat{u}_i$ is maintained in the projection set, namely, $\hat{u}_i(t) \in [\chi, \infty)$ for all $t \geq 0$.
3) For all $(\phi, \hat{u}_i) \in \mathbb{R}^2$ and $u_i^* \geq \chi$, the following holds: $(u_i^* - \hat{u}_i)(\phi - \mathrm{Proj}\{\phi; \chi\}) \leq 0$.

Proof: Points 1 and 2 are immediate. Point 3 follows from the calculation

$$\begin{aligned} &(u_i^* - \hat{u}_i)(\phi - \mathrm{Proj}\{\phi; \chi\}) \\ &= \begin{cases} (u_i^* - \hat{u}_i)\phi, & \text{if } \hat{u}_i \leq \chi \text{ and } \phi < 0 \\ 0, & \text{otherwise} \end{cases} \end{aligned} \tag{82}$$

since $u_i^* - \hat{u}_i \geq 0$ when $u_i \leq \chi$. ∎

Several challenges remain in the development of convergence proofs for Nash seeking players with projection, such as computing the average error system and defining the precise notion of stability for equilibria that exist on the action space boundaries. To alleviate concerns of existence and uniqueness of solutions that arise from the projection operator being discontinuous, a Lipschitz continuous version of projection [51] can be used.

Fig. 9 depicts the trajectories $\hat{u}_1$ and $\hat{u}_2$ when the players employ (79)–(80) with the same parameters as in Section V-A and the initial condition $(u_1(0), u_2(0)) = (\hat{u}_1(0), \hat{u}_2(0)) = (0.2, 0.03)$. For this initial condition, the players reach $\partial \hat{U}$ where $\hat{u}_1 = a_1 = 0.06$ while $\hat{u}_2$ increases until the trajectory reaches the eigenvector and $\hat{u}_1$ is pushed away from the boundary. From this point, the players follow a trajectory that is similar to the one depicted in Fig. 7(b).

While the modified strategy (79)–(80) ensures that the players' actions lie in U, it effectively prevents the players from converging to a neighborhood of any equilibria that lie in $U \setminus \hat{U}$. To address this limitation, we propose the modified projection strategy

$$\dot{\hat{u}}_i(t) = k_i \mathrm{Proj}\{\mu_i(t) J_i(t); 0\} \tag{83}$$
$$u_i(t) = \hat{u}_i(t) + \min\left(1, \frac{|\hat{u}_i(t)|}{a_i}\right) \mu_i(t) \tag{84}$$

where $\mathrm{Proj}\{\phi; \chi\}$ is defined as in (81). The strategy (83)–(84) is equivalent to (79)–(80) when $\chi = a_i$ since $\mathrm{Proj}\{\mu_i(t) J_i(t); a_i\}$ ensures that $\min(1, |\hat{u}_i|/a_i) = 1$.

In Fig. 10, the players employ (83)–(84) for the same scenario shown in Fig. 9. With this strategy, the players approach the boundary $u_1 = \hat{u}_1 = 0$, where, unlike at the boundary $\hat{u}_1 = 0.06$, the vector field causes $\hat{u}_2$ to decrease toward the Nash equilibrium at $(0, 0)$. This Nash equilibrium is attainable only by allowing $\hat{u}_i$ to enter $U \setminus \hat{U}$, but the convergence rate slows near the boundary since $\min(a_i, |\hat{u}_i|)/a_i \to 0$ as $\hat{u}_i$ approaches ∂U and due to the nearby unstable Nash equilibrium (Fig. 10). If the nonnegative action requirement were relaxed to allow the perturbation to leave U, i.e., we restrict u_i to the set $\bar{U} = \{(u_1, u_2) \in \mathbb{R}^2 \mid -a_1 \leq u_1, -a_2 \leq u_2 < 16\}$, then the point $(0, 0)$ would be attainable using (79)–(80) with $\chi = 0$. This strategy would exhibit a faster convergence rate along ∂U than if the players used (83)–(84), but convergence would still be slow due to the presence of the unstable Nash equilibrium.

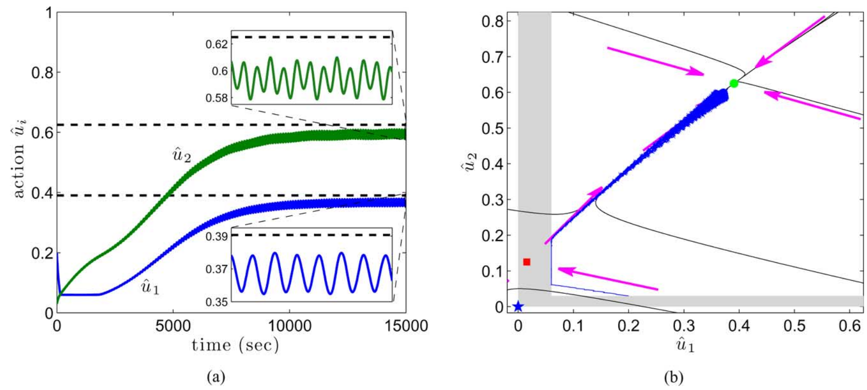


Fig. 9. (a) Time history of the two-player game initialized at $(0.2, 0.03)$ when implementing the modified Nash seeking strategy (79)–(80), and (b) planar $\hat{u}$-trajectories superimposed on the $\hat{u}$-reduced model average system phase portrait. The stable and unstable Nash equilibrium points (green circle and red square), their eigenvectors (magenta arrows), and the Nash equilibrium on ∂U (blue star) are denoted. The shaded region is the set difference $U \setminus \hat{U}$.

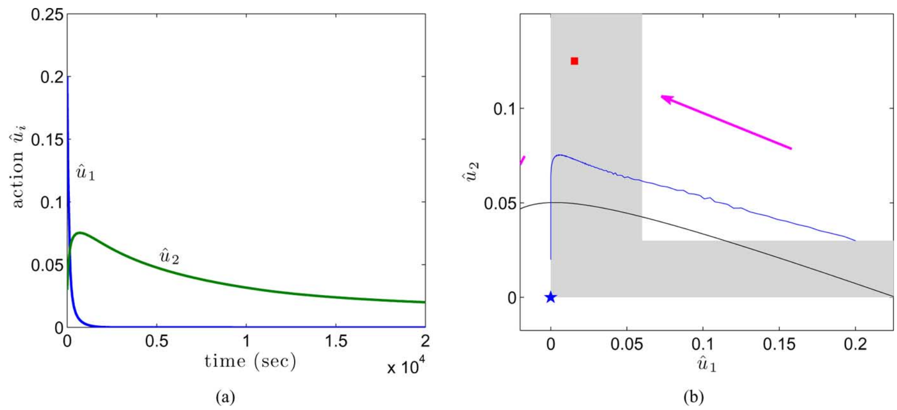


Fig. 10. (a) Time history of the two-player game initialized at $(0.2, 0.03)$ when implementing the modified Nash seeking strategy (83)–(84), and (b) planar $\hat{u}$-trajectories superimposed on the $\hat{u}$-reduced model average system phase portrait. The unstable Nash equilibrium (red square), its eigenvectors (magenta arrows), and the Nash equilibrium on ∂U (blue star) are denoted. The shaded region is the set difference $U \setminus \hat{U}$.

VI. Conclusion

We have introduced a non-model based approach for convergence to the Nash equilibria of static, noncooperative games with N players, where the players measure only their own payoff values. Our convergence results include games with non-quadratic functions and games where the players' actions serve as inputs to a general, stable nonlinear differential equation whose outputs are the players' payoffs. For non-quadratic payoff functions, the convergence is biased in proportion to both payoff functions' higher derivatives and the perturbation signals' amplitudes. When the players' actions are restricted to a closed and bounded action set $U \subset \mathbb{R}^N$, extremum seeking with projection ensures that their actions remain in U.

A player cannot determine *a priori* if its initial action is sufficiently close to u^* to guarantee convergence since u^* is unknown and because a quantitative estimate of the region of attraction for u^* is very difficult to determine. However, for quadratic payoffs, convergence to u^* is semi-global. If we assume the players possess a good estimate of u^*, based on either partial information of the game or historical data, this learning strategy remains attractive since it allows players to improve their initial actions by measuring only their payoff values and does not require the estimation of potentially highly uncertain parameters, e.g., competitors' prices or consumer demand. Also, due to the dynamic nature of this algorithm, the players can track movements in u^* should the game's model change smoothly over time in a manner that is unknown to the players.

Appendix A

Several integrals over the period $T = 2\pi \times \mathrm{LCM}\{1/\bar{\omega}_1, \ldots, 1/\bar{\omega}_N\}$, where LCM denotes the least common multiple, must be computed to obtain the average error system (22) and equilibrium (62) for the average error system (58). First, we detail the necessary calculations to compute the average error system for games with quadratic payoffs, and then compute the additional terms that arise for games with non-quadratic payoffs in Appendix B. To evaluate the integrals, we use the following trigonometric identities to simplify the integrands: $\cos^2\theta = (1+\cos 2\theta)/2$, $\sin^2\theta = (1-\cos 2\theta)/2$, $\sin^3\theta = (3\sin\theta - \sin 3\theta)/4$, $\cos\theta\cos\varphi = (\cos(\theta-\varphi)+\cos(\theta+\varphi))/2$, $\cos\theta\sin\varphi = (\sin(\theta+\varphi)-\sin(\theta-\varphi))/2$, $\sin\theta\sin\varphi = (\cos(\theta-\varphi)-\cos(\theta+\varphi))/2$.

To obtain (22), we compute the average of each term in the error system (21) and assume that $\omega_i \neq \omega_j$, $2\omega_i \neq \omega_j$, and $\omega_i \neq \omega_j + \omega_k$ for all distinct $i, j, k \in \{1, \ldots, N\}$. The averages of the first and last terms of (21) are zero since

$$\frac{1}{T}\int_0^T \mu_i(\tau)\,d\tau = a_i \frac{\cos\varphi_i - \cos(\bar{\omega}_i T + \varphi_i)}{\bar{\omega}_i T} = 0. \quad (85)$$

The average of the second term of (21) depends on whether or not $i = k$. If $i \neq k$

$$\frac{1}{T}\int_0^T \mu_i(\tau)\mu_k(\tau)\,d\tau = 0 \tag{86}$$

since $\omega_i \neq \omega_k$, and if $i = k$

$$\frac{1}{T}\int_0^T \mu_i^2(\tau)\,d\tau = \frac{a_i^2}{2} \tag{87}$$

which with (86) implies

$$\frac{1}{T}\int_0^T \frac{k_i}{\omega}\sum_{j=1}^N\sum_{k=1}^N D_{jk}^i\left(\tilde{u}_j(\tau)+u_j^*\right)\mu_i(\tau)\mu_k(\tau)\,d\tau = \frac{k_i a_i^2}{2\omega}\sum_{j=1}^N D_{ij}^i\left(\tilde{u}_j(\tau)+u_j^*\right). \tag{88}$$

Similarly, the average of the fourth term of (21) is

$$\frac{1}{T}\int_0^T \frac{k_i}{\omega}\sum_{j=1}^N d_j^i\left(\tilde{u}_j(\tau)+u_j^*+\mu_j(\tau)\right)\mu_i(\tau)\,d\tau = \frac{k_i a_i^2}{2\omega}d_i^i. \tag{89}$$

The average of the third term requires computing the following three integrals:

$$(1/T)\int_0^T \mu_i^3(\tau)\,d\tau = 0,$$
$$(1/T)\int_0^T \mu_i^2(\tau)\mu_j(\tau)\,d\tau = 0,$$
$$(1/T)\int_0^T \mu_i(\tau)\mu_j(\tau)\mu_k(\tau)\,d\tau = 0.$$

Appendix B

For games with non-quadratic payoffs, we assume the players' frequencies satisfy the requirements for games with quadratic payoff functions and also $3\omega_i \neq \omega_j$, $\omega_i \neq 2\omega_j + \omega_k$, $2\omega_i \neq \omega_j + \omega_k$, where $i, j, k \in \{1,\ldots,N\}$ are distinct. Then, in addition to the integrals computed in Appendix A, the following integrals are computed to obtain (62):

$$(1/T)\int_0^T \mu_i^4(\tau)\,d\tau = 3a_i^4/8,$$
$$(1/T)\int_0^T \mu_i^2(\tau)\mu_j^2(\tau)\,d\tau = a_i^2 a_j^2/4,$$
$$(1/T)\int_0^T \mu_i^3(\tau)\mu_j(\tau)\,d\tau = 0,$$

and

$$(1/T)\int_0^T \mu_i(\tau)\mu_j^2(\tau)\mu_k(\tau)\,d\tau = 0.$$

References

[1] A. Cournot, *Recherches sur les Principes Mathématiques de la Théorie des Richesses*. Paris, France: Hachette, 1838.

[2] G. W. Brown, "Iterative solutions of games by fictitious play," in *Activity Analysis of Production and Allocation*, T. C. Koopmans, Ed. New York: Wiley, 1951, pp. 374–376.

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

[4] D. G. Luenberger, "Complete stability of noncooperative games," *J. Optim. Theory Appl.*, vol. 25, no. 4, pp. 485–505, 1978.

[5] S. Li and T. Başar, "Distributed algorithms for the computation of noncooperative equilibria," *Automatica*, vol. 23, no. 4, pp. 523–533, 1987.

[6] J. S. Shamma and G. Arslan, "Dynamic fictitious play, dynamic gradient play, and distributed convergence to Nash equilibria," *IEEE Trans. Autom. Control*, vol. 53, no. 3, pp. 312–327, 2005.

[7] M. Zhu and S. Martínez, "Distributed coverage games for mobile visual sensor networks," *SIAM J. Control Optim.*, Jan. 2010 [Online]. Available: http://arxiv.org/abs/1002.0367, submitted for publication

[8] M. Krstic, P. Frihauf, J. Krieger, and T. Başar, "Nash equilibrium seeking with finitely- and infinitely-many players," in *Proc. 8th IFAC Symp. Nonlinear Control Syst.*, Bologna, Italy, Sep. 2010, pp. 1086–1091.

[9] P. Frihauf, M. Krstic, and T. Başar, "Nash equilibrium seeking for games with non-quadratic payoffs," in *Proc. IEEE Conf. Decision Control*, Atlanta, GA, Dec. 2010, pp. 881–886.

[10] 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*, Atlanta, GA, Dec. 2010, pp. 5598–5603.

[11] A. Jafari, A. Greenwald, D. Gondek, and G. Ercal, "On no-regret learning, fictitious play, and Nash equilibrium," in *Proc. 18th Int. Conf. Mach. Learn.*, 2001, pp. 226–233.

[12] D. P. Foster and H. P. Young, "Regret testing: Learning to play Nash equilibrium without knowing you have an opponent," *Theoret. Econ.*, vol. 1, no. 3, pp. 341–367, 2006.

[13] R. Sharma and M. Gopal, "Synergizing reinforcement learning and game theory—A new direction for control," *Applied Soft Comput.*, vol. 10, no. 3, pp. 675–688, 2010.

[14] D. Fudenberg and D. K. Levine, *The Theory of Learning in Games*. Cambridge, MA: The MIT Press, 1998.

[15] N. Cesa-Bianchi and G. Lugosi, *Prediction, Learning, and Games*. New York: Cambridge Univ. Press, 2006.

[16] Q. Zhu, H. Tembine, and T. Başar, "Heterogeneous learning in zero-sum stochastic games with incomplete information," in *Proc. IEEE Conf. Decision Control*, Atlanta, GA, Dec. 2010, pp. 219–224.

[17] A. B. MacKenzie and S. B. Wicker, "Game theory and the design of self-configuring, adaptive wireless networks," *IEEE Commun. Mag.*, vol. 39, no. 11, pp. 126–131, 2001.

[18] 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. Autom. Control*, vol. 47, no. 6, pp. 917–930, 2002.

[19] T. Başar, "Control and game-theoretic tools for communication networks (overview)," *Appl. Comput. Math.*, vol. 6, no. 2, pp. 104–125, 2007.

[20] G. Scutari, D. P. Palomar, and S. Barbarossa, "The MIMO iterative waterfilling algorithm," *IEEE Trans. Signal Process.*, vol. 57, no. 5, pp. 1917–1935, May 2009.

[21] S. S. Rao, V. B. Venkayya, and N. S. Khot, "Game theory approach for the integrated design of structures and controls," *AIAA J.*, vol. 26, no. 4, pp. 463–469, 1988.

[22] D. Bauso, L. Giarré, and R. Pesenti, "Consensus in noncooperative dynamic games: A multiretailer inventory application," *IEEE Trans. Autom. Control*, vol. 53, no. 4, pp. 998–1003, 2008.

[23] J. R. Marden, G. Arslan, and J. S. Shamma, "Cooperative control and potential games," *IEEE Trans. Syst., Man, Cybern.—Part B: Cybern.*, vol. 39, no. 6, pp. 1393–1407, Jun. 2009.

[24] E. Semsar-Kazerooni and K. Khorasani, "Multi-agent team cooperation: A game theory approach," *Automatica*, vol. 45, no. 10, pp. 2205–2213, 2009.

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

[26] H.-H. Wang and M. Krstic, "Extremum seeking for limit cycle minimization," *IEEE Trans. Autom. Control*, vol. 45, no. 12, pp. 2432–2437, Dec. 2000.

[27] 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.

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

[29] Y. Tan, D. Nešić, and I. Mareels, "On non-local stability properties of extremum seeking control," *Automatica*, vol. 42, no. 6, pp. 889–903, 2006.

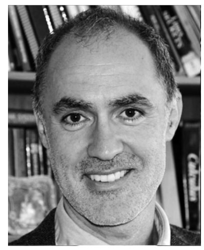
[30] L. Luo and E. Schuster, "Mixing enhancement in 2D magnetohydrodynamic channel flow by extremum seeking boundary control," in *Proc. Amer. Control Conf.*, St. Louis, MO, Jun. 2009, pp. 1530–1535.

[31] W. H. Moase, C. Manzie, and M. J. Brear, “Newton-like extremum-seeking part I: Theory,” in *Proc. IEEE Conf. Decision Control*, Shanghai, China, Dec. 2009, pp. 3839–3844.
[32] M. S. Stanković and D. M. Stipanović, “Extremum seeking under stochastic noise and applications to mobile sensors,” *Automatica*, vol. 46, pp. 1243–1251, 2010.
[33] D. Nešić, Y. Tan, W. H. Moase, and C. Manzie, “A unifying approach to extremum seeking: Adaptive schemes based on estimation of derivatives,” in *Proc. IEEE Conf. Decision Control*, Atlanta, GA, Dec. 2010, pp. 4625–4630.
[34] 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.
[35] J. Cochran and M. Krstic, “Nonholonomic source seeking with tuning of angular velocity,” *IEEE Trans. Autom. Control*, vol. 54, no. 4, pp. 717–731, Apr. 2009.
[36] J. Cochran, E. Kanso, S. D. Kelly, H. Xiong, and M. Krstic, “Source seeking for two nonholonomic models of fish locomotion,” *IEEE Trans. Robot.*, vol. 25, no. 5, pp. 1166–1176, Oct. 2009.
[37] N. J. Killingsworth, S. M. Aceves, D. L. Flowers, F. Espinosa-Loza, and M. Krstic, “HCCI engine combustion-timing control: Optimizing gains and fuel consumption via extremum seeking,” *IEEE Trans. Control Syst. Technol.*, vol. 17, no. 6, pp. 1350–1361, Nov. 2009.
[38] M. Guay, M. Perrier, and D. Dochain, “Adaptive extremum seeking control of nonisothermal continuous stirred reactors,” *Chem. Eng. Sci.*, vol. 60, pp. 3671–3681, 2005.
[39] K. Peterson and A. Stefanopoulou, “Extremum seeking control for soft landing of an electromechanical valve actuator,” *Automatica*, vol. 29, pp. 1063–1069, 2004.
[40] R. Becker, R. King, R. Petz, and W. Nitsche, “Adaptive closed-loop separation control on a high-lift configuration using extremum seeking,” *AIAA J.*, vol. 45, no. 6, pp. 1382–1392, 2007.
[41] D. Carnevale, A. Astolfi, C. Centioli, S. Podda, V. Vitale, and L. Zaccarian, “A new extremum seeking technique and its application to maximize RF heating on FTU,” *Fusing Eng. Design*, vol. 84, pp. 554–558, 2009.
[42] A. R. Teel, J. Peuteman, and D. Aeyels, “Semi-global practical asymptotic stability and averaging,” *Syst. Control Lett.*, vol. 37, no. 5, pp. 329–334, 1999.
[43] H. K. Khalil, *Nonlinear Systems*, 3rd ed. Upper Saddle River, NJ: Prentice Hall, 2002.
[44] O. Taussky, “A recurring theorem on determinants,” *Amer. Math. Monthly*, vol. 56, no. 10, pp. 672–676, 1949.
[45] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1985.
[46] D. Monderer and L. S. Shapley, “Potential games,” *Games Econ. Behav.*, vol. 14, no. 1, pp. 124–143, 1996.
[47] E. Hopkins, “A note on best response dynamics,” *Games Econ. Behav.*, vol. 29, no. 1–2, pp. 138–150, 1999.
[48] J. Hofbauer, “Best response dynamics for continuous zero-sum games,” *Discrete Contin. Dynam. Syst.: Ser. B*, vol. 6, no. 1, pp. 215–224, 2006.
[49] T. M. Apostol, *Mathematical Analysis*, 2nd ed. Reading, MA: Addison-Wesley, 1974.
[50] A. K. Naimzada and L. Sbragia, “Oligopoly games with nonlinear demand and cost functions: Two boundedly rational adjustment processes,” *Chaos, Solitons, Fractals*, vol. 29, pp. 707–722, 2006.
[51] M. Krstic, I. Kanellakopoulos, and P. Kokotović, *Nonlinear and Adaptive Control Design*. New York: Wiley-Interscience, 1995.

Paul Frihauf (S’09) received the B.S. and M.S. degrees in systems science and engineering from Washington University, St. Louis, MO, in 2005. He is currently working toward the Ph.D. degree in dynamic systems and control from the Mechanical and Aerospace Engineering Department, University of California at San Diego, La Jolla.

He was an Industrial Engineering Intern at The Boeing Company in 2003, a Summer Research Assistant at MIT Lincoln Laboratory in 2004, Associate Professional Staff in the Guidance, Navigation, and Control Group, The Johns Hopkins University Applied Physics Laboratory, from 2005 to 2007, and an Algorithms Intern at Cymer in 2011. His research interests include noncooperative games, multi-agent control, and the control of distributed parameter systems.

Mr. Frihauf is a member of SIAM. He is a National Defense Science and Engineering Graduate (NDSEG) Fellow and an ARCS scholar.



Miroslav Krstic (S’92–M’95–SM’99–F’02) received the Ph.D. degree from the University of California (UC), Santa Barbara, in 1994.

He was an Assistant Professor at the University of Maryland, College Park, until 1997. He is the Daniel L. Alspach Professor and the founding Director of the Cymer Center for Control Systems and Dynamics (CCSD) at UC San Diego, La Jolla. He is a coauthor of eight books: *Nonlinear and Adaptive Control Design* (Wiley, 1995), *Stabilization of Nonlinear Uncertain Systems* (Springer, 1998), *Flow Control by Feedback* (Springer, 2002), *Real-time Optimization by Extremum Seeking Control* (Wiley, 2003), *Control of Turbulent and Magnetohydrodynamic Channel Flows* (Birkhauser, 2007), *Boundary Control of PDEs: A Course on Backstepping Designs* (SIAM, 2008), *Delay Compensation for Nonlinear, Adaptive, and PDE Systems* (Birkhauser, 2009), and *Adaptive Control of Parabolic PDEs* (Princeton Univ. Press, 2009).

Dr. Krstic is a Fellow of IFAC and has received the Axelby and Schuck Paper Prizes, NSF Career, ONR Young Investigator, and PECASE Award. He has held the appointment of Springer Distinguished Visiting Professor of Mechanical Engineering at UC Berkeley.

Tamer Başar (S’71–M’73–SM’79–F’83) received the B.S.E.E. degree from Robert College, Istanbul, Turkey, and the M.S., M.Phil, and Ph.D. degrees from Yale University, New Haven, CT.

He is with the University of Illinois at Urbana-Champaign (UIUC), where he holds the positions of Swanlund Endowed Chair; Center for Advanced Study Professor of Electrical and Computer Engineering; Research Professor at the Coordinated Science Laboratory; and Research Professor at the Information Trust Institute. He joined UIUC in 1981 after holding positions at Harvard University and Marmara Research Institute (Turkey). He has published extensively in systems, control, communications, and dynamic games, and has current research interests in modeling and control of communication networks; control over heterogeneous networks; estimation and control with limited sensing and transmission; resource allocation, management and pricing in networks; cooperative and noncooperative game theory for networks; mobile and distributed computing; and security and trust in computer systems. He is currently the Editor-in-Chief of *Automatica*, the Editor of the Birkhäuser Series on *Systems & Control*, the Editor of the Birkhäuser Series on *Static & Dynamic Game Theory*, Managing Editor of the *Annals* of the International Society of Dynamic Games (ISDG), and member of editorial and advisory boards of several international journals in control, wireless networks, and applied mathematics.

Dr. Başar has received several awards and recognitions over the years, among which are the Medal of Science of Turkey (1993); Distinguished Member Award (1993), Axelby Outstanding Paper Award (1995), and Bode Lecture Prize (2004) of the IEEE Control Systems Society (CSS); Millennium Medal of IEEE (2000); Tau Beta Pi Drucker Eminent Faculty Award of UIUC (2004); the Outstanding Service Award (2005) and the Giorgio Quazza Medal (2005) of the International Federation of Automatic Control (IFAC); the Richard Bellman Control Heritage Award (2006) of the American Automatic Control Council (AACC); honorary doctorate (Doctor Honoris Causa) from Doğuş University (Istanbul; 2007); honorary professorship from Northeastern University (Shenyang; 2008); and Isaacs Award of ISDG (2010). He is a member of the U.S. National Academy of Engineering, a member of the European Academy of Sciences, a Fellow of IFAC, a council member of IFAC, a past president of CSS, a past (founding) president of ISDG, and the current President of AACC.