

- [6] D. Liberzon and A. S. Morse, "Basic problems in stability and design of switched systems," *IEEE Contr. Syst. Mag.*, vol. 19, pp. 59–70, Oct. 1999.
- [7] A. P. Molchanov and Ye. S. Pyatnitskiy, "Criteria of asymptotic stability of differential and difference inclusions encountered in control theory," *Syst. Control Lett.*, vol. 13, no. 1, pp. 59–64, 1989.
- [8] H. Radjavi and P. Rosenthal, *Simultaneous Triangularization*. New York: Springer-Verlag, 1999.
- [9] A. V. Savkin and R. J. Evans, *Hybrid Dynamical Systems: Controller and Sensor Switching Problems*. Boston, MA: Birkhäuser, 2002.
- [10] D. P. Standford, "Stability for a multi-rate sampled-data system," *SIAM J. Control Optim.*, vol. 17, pp. 390–399, 1979.
- [11] D. P. Stanford and J. M. Urbano, "Some convergence properties of matrix sets," *SIAM J. Matrix Anal. Appl.*, vol. 14, no. 4, pp. 1132–1140, 1994.
- [12] Z. Sun and S. S. Ge, "Dynamic output feedback stabilization of a class of switched linear systems," *IEEE Trans. Circuits Syst. I*, vol. 50, pp. 1111–1115, Aug. 2003.
- [13] Z. Sun and D. Zheng, "On reachability and stabilization of switched linear control systems," *IEEE Trans. Automat. Contr.*, vol. 46, pp. 291–295, Feb. 2001.
- [14] J. Tokarzewski, "Stability of periodically switched linear systems and the switching frequency," *Int. J. Syst. Sci.*, vol. 18, no. 4, pp. 697–726, 1987.
- [15] M. A. Wicks, P. Peleties, and R. A. DeCarlo, "Switched controller synthesis for the quadratic stabilization of a pair of unstable linear systems," *Eur. J. Control*, vol. 4, no. 2, pp. 140–147, 1998.
- [16] H. Ye, A. N. Michel, and L. Hou, "Stability theory for hybrid dynamical systems," *IEEE Trans. Automat. Contr.*, vol. 43, pp. 461–474, Apr. 1998.
- [17] X. Xu and P. Antsaklis, "Stabilization of second-order LTI switched systems," *Int. J. Control*, vol. 73, no. 14, pp. 1261–1279, 2000.

Unified Convergence Proofs of Continuous-Time Fictitious Play

Jeff S. Shamma and Gurdal Arslan

Abstract—We consider a continuous-time version of fictitious play (FP), in which interacting players evolve their strategies in reaction to their opponents' actions without knowledge of their opponents' utilities. It is known that FP need not converge, but that convergence is possible in certain special cases including zero-sum games, identical interest games, and two-player/two-move games. We provide a unified proof of convergence in all of these cases by showing that a Lyapunov function previously introduced for zero-sum games also can establish stability in the other special cases. We go on to consider a two-player game in which only one player has two-moves and use properties of planar dynamical systems to establish convergence.

Index Terms—Fictitious play, game theory, Nash equilibrium.

I. OVERVIEW

The procedure of fictitious play (FP) was introduced in 1951 [5], [18] as a mechanism to compute Nash equilibria in matrix games. In

Manuscript received October 17, 2003. Recommended by Associate Editor C. D. Charalambous. This work was supported by the Air Force Office of Scientific Research/MURI under Grant F49620-01-1-0361 and by summer support by the UF Graduate Engineering Research Center.

The authors are with the Department of Mechanical and Aerospace Engineering, University of California, Los Angeles, CA 90095-1597 USA (e-mail:shamma@seas.ucla.edu; garslan@seas.ucla.edu).

Digital Object Identifier 10.1109/TAC.2004.831143

FP, game players repeatedly use strategies that are best responses to the historical averages, or *empirical frequencies*, of opponents. These empirical frequencies and, hence, player strategies may or may not converge. The important implication of convergence of empirical frequencies is that it implies a convergence to a Nash equilibrium.

There is a substantial body of literature on FP [9]. A selected timeline of results that establish convergence for special cases of games is as follows: 1951, two player zero-sum games [18]; 1961, two player two move games [16]; 1993, noisy two player two move games with a unique Nash equilibrium [8]; 1996, multiplayer games with identical player utilities [17]; 1999, noisy two-player/two-move games with countable Nash equilibria [2]; and two player games in which one player has only two moves [3]. A convergence *counterexample* due to Shapley in 1964 has two players with three moves each [20]. A 1993 counterexample due to Jordan has three players with two moves each [12]. Nonconvergence issues are also discussed in [6], [7], [10], [13], and [21].

In this note, we consider a continuous-time form of FP and provide a unified proof of convergence of empirical frequencies for the special cases of zero-sum games, identical interest games, and two-player/two-move games. The proofs are unified in the sense that they all employ an energy function that has the natural interpretation as an "opportunity for improvement." This energy function was used as a Lyapunov function in [11] for zero-sum games. We show that the same energy function can establish convergence for all of the above cases, in some cases by a Lyapunov argument and in other cases by an integrability argument.

We go on to consider games in which one of two players has only two moves. We provide an alternative proof that exploits some simple properties of planar dynamical systems.

The remainder of this note is organized as follows. Section II sets up the problem of continuous time FP. Section III contains convergence proofs for zero-sum, identical interest, and two-player/two-move games. Section IV discusses games in which one of two players has only two moves. Finally, Section V has some concluding remarks.

Notation

- For $i \in \{1, 2, \dots, n\}$, $-i$ denotes the complementary set $\{1, \dots, i-1, i+1, \dots, n\}$.

- Boldface $\mathbf{1}$ denotes the vector $\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$ of appropriate dimension.

- $\Delta(n)$ denotes the simplex in \mathcal{R}^n , i.e.,

$$\left\{ s \in \mathcal{R}^n \mid s \geq 0 \text{ componentwise, and } \mathbf{1}^T s = 1 \right\}$$

- $\text{Int}(\Delta(n))$ denotes the set of interior points of a simplex, i.e., $s > 0$ componentwise.

- $\mathbf{v}_i \in \Delta(n)$ denotes the i th vertex of the simplex $\Delta(n)$, i.e., the vector whose i th term equals 1 and remaining terms equal 0.

- $\mathcal{H} : \text{Int}(\Delta(n)) \rightarrow \mathcal{R}$ denotes the entropy function

$$\mathcal{H}(s) = -s^T \log(s)$$

- $\sigma : \mathcal{R}^n \rightarrow \text{Int}(\Delta(n))$ denotes the "logit" or "soft-max" function

$$(\sigma(x))_i = \frac{e^{x_i}}{e^{x_1} + \dots + e^{x_n}}.$$

This function is continuously differentiable. The Jacobian matrix of partial derivatives, denoted $\nabla\sigma(\cdot)$, is

$$\nabla\sigma(x) = \text{diag}(\sigma(x)) - \sigma(x)\sigma^T(x)$$

where $\text{diag}(\sigma(x))$ denotes the diagonal square matrix with elements taken from $\sigma(x)$.

II. FICTITIOUS PLAY SETUP

A. Static Game

We consider a two player game with players \mathcal{P}_1 and \mathcal{P}_2 , each with positive integer dimensions m_1 and m_2 , respectively. Each player, \mathcal{P}_i , selects a strategy, $p_i \in \Delta(m_i)$, and receives a real-valued reward according to the utility function $\mathcal{U}_i(p_i, p_{-i})$. These utility functions take the form

$$\begin{aligned}\mathcal{U}_1(p_1, p_2) &= p_1^T M_1 p_2 + \tau \mathcal{H}(p_1) \\ \mathcal{U}_2(p_2, p_1) &= p_2^T M_2 p_1 + \tau \mathcal{H}(p_2)\end{aligned}$$

characterized by matrices M_i of appropriate dimension and $\tau > 0$.

The standard interpretation is that p_i represent probabilistic strategies. Each player selects an integer action $a_i \in \{1, \dots, m_i\}$ according to the probability distribution p_i . The reward to player \mathcal{P}_i is

$$\mathbf{v}_{a_i}^T M_i \mathbf{v}_{a_{-i}} + \tau \mathcal{H}(p_i)$$

i.e., the reward to player \mathcal{P}_i is the element of M_i in the a_i^{th} row and a_{-i}^{th} column, plus the weighted entropy of its strategy. For a given strategy pair, (p_1, p_2) , the utilities represent the expected rewards

$$\mathcal{U}_i(p_i, p_{-i}) = E \left[\mathbf{v}_{a_i}^T M_i \mathbf{v}_{a_{-i}} \right] + \tau \mathcal{H}(p_i).$$

Define the *best response* mappings

$$\beta_i : \Delta(m_{-i}) \rightarrow \Delta(m_i)$$

by

$$\beta_i(p_{-i}) = \arg \max_{p_i \in \Delta(m_i)} \mathcal{U}_i(p_i, p_{-i}).$$

The best response turns out to be the logit or soft-max function (see the **Notation** section)

$$\beta_i(p_{-i}) = \sigma \left(\frac{M_i p_{-i}}{\tau} \right).$$

A Nash equilibrium is a pair $(p_1^*, p_2^*) \in \Delta(m_1) \times \Delta(m_2)$ such that for all $p_i \in \Delta(m_i)$

$$\mathcal{U}_i(p_i, p_{-i}^*) \leq \mathcal{U}_i(p_i^*, p_{-i}^*), \quad i \in \{1, 2\} \quad (1)$$

i.e., each player has no incentive to deviate from an equilibrium strategy provided that the other player maintains an equilibrium strategy. In terms of the best response mappings, a Nash equilibrium is any pair (p_1^*, p_2^*) such that

$$p_i^* = \beta_i(p_{-i}^*), \quad i \in \{1, 2\}.$$

B. Discrete-Time Fictitious Play

Now, suppose that the game is repeated at every time $k \in \{0, 1, 2, \dots\}$. In particular, we are interested in an ‘‘evolutionary’’ version of the game in which the strategies at time k , denoted by $p_i(k)$, are selected in response to the entire prior history of an opponent’s actions.

Toward this end, let $a_i(k)$ denote the action of player \mathcal{P}_i at time k , chosen according to the probability distribution $p_i(k)$, and let $\mathbf{v}_{a_i(k)} \in \Delta(m_i)$ denote the corresponding simplex vertex. The *empirical frequency*, $q_i(k)$, of player \mathcal{P}_i is defined as the running average of the actions of player \mathcal{P}_i , which can be computed by the recursion

$$q_i(k+1) = q_i(k) + \frac{1}{k+1} (\mathbf{v}_{a_i(k)} - q_i(k)).$$

In discrete-time FP, the strategy of player \mathcal{P}_i at time k is the optimal response to the running average of the opponent’s actions, i.e.,

$$p_i(k) = \beta_i(q_{-i}(k)).$$

The case where $\tau = 0$ corresponds to classical FP. Setting τ positive rewards randomization, thereby imposing in so-called mixed strategies. As τ approaches zero, the best response mappings approximate selecting the maximal element since the probability of selecting a maximal element approaches one when the maximal element is unique. The game with τ positive then can be viewed as a smoothed version of the matrix game [8] in which rewards are subject to random perturbations.

C. Continuous-Time Fictitious Play

Now, consider the continuous-time dynamics

$$\begin{aligned}\dot{q}_1(t) &= \beta_1(q_2(t)) - q_1(t) \\ \dot{q}_2(t) &= \beta_2(q_1(t)) - q_2(t).\end{aligned} \quad (2)$$

We will call these equations *continuous-time FP*. These are the dynamics obtained by viewing discrete-time FP as stochastic approximation iterations and applying associated ordinary differential equation (ODE) analysis methods [1], [14].

III. CONVERGENCE PROOFS FOR ZERO-SUM, IDENTICAL INTEREST, AND TWO-MOVE GAMES

We will derive a unified framework which establishes convergence of (2) to a Nash equilibrium of the static game (1) in the aforementioned special cases of zero-sum, identical interest, and two-move games.

Zero-sum and identical interest here refer to the portion of the utility other than the weighted entropy. In other words, zero-sum means that for all p_1 and p_2

$$p_1^T M_1 p_2 = -p_2^T M_2 p_1$$

and identical interest means that

$$p_1^T M_1 p_2 = p_2^T M_2 p_1.$$

Strictly speaking, the inclusion of the entropy term does not result in a zero-sum or identical interest game, but we will use these terms nonetheless.

Define the function $V_1 : \Delta(m_1) \times \Delta(m_2) \rightarrow [0, \infty)$ as

$$\begin{aligned}V_1(q_1, q_2) &= \max_{s \in \Delta(m_1)} \mathcal{U}_1(s, q_2) - \mathcal{U}_1(q_1, q_2) \\ &= (\beta_1(q_2) - q_1)^T M_1 q_2 + \tau (\mathcal{H}(\beta_1(q_2)) - \mathcal{H}(q_1)).\end{aligned}$$

Similarly define

$$V_2(q_2, q_1) = \max_{s \in \Delta(m_2)} \mathcal{U}_2(s, q_1) - \mathcal{U}_2(q_2, q_1).$$

Each V_i has the natural interpretation as the maximum possible reward improvement to player \mathcal{P}_i by using the best response to q_{-i} rather than the specified q_i . Note that by definition

$$V_i(q_i, q_{-i}) \geq 0$$

with equality if and only if

$$q_i = \beta_i(q_{-i}).$$

These functions were used in [11] for zero-sum games, i.e., $M_1 = -M_2^T$, through a Lyapunov argument using $V_1 + V_2$ to show that the continuous-time empirical frequencies converge to a Nash equilibrium.

We will show that the *same* functions can be used to establish convergence to a Nash equilibrium in the case of identical interest games and in the case of two-move games. The identical interest case will not be a Lyapunov argument. Rather, we will show that the sum, $V_1 + V_2$, is integrable. For two-move games, we will show that an appropriately scaled sum, $\alpha_1 V_1 + \alpha_2 V_2$, is either a Lyapunov function or is integrable.

The following lemma reveals a special structure for the derivatives of the V_i along trajectories of continuous-time FP (2).

Lemma 3.1: Define

$$\tilde{V}_i(t) = V_i(q_i(t), q_{-i}(t))$$

along solutions of continuous-time FP (2). Then

$$\begin{aligned} \dot{\tilde{V}}_1 &\leq -\tilde{V}_1 + \dot{q}_1^T M_1 \dot{q}_2 \\ \dot{\tilde{V}}_2 &\leq -\tilde{V}_2 + \dot{q}_2^T M_2 \dot{q}_1. \end{aligned}$$

The proof uses the following lemma.

Lemma 3.2 [4, Lemma 3.3.1]: Let $F(x, u)$ be a continuously differentiable function of $x \in \mathcal{R}^n$ and $u \in \mathcal{R}^m$. Let U be a convex subset of \mathcal{R}^m . Assume that $\mu^*(x)$ is a continuously differentiable function such that for all x

$$\mu^*(x) = \arg \max_{u \in U} F(x, u).$$

Then

$$\nabla_x (\max_{u \in U} F(x, u)) = \nabla_x F(x, \mu^*(x)).$$

Proof: (Lemma 3.1) By definition

$$\begin{aligned} \dot{\tilde{V}}_1 &= -(M_1 q_2)^T (\beta_1(q_2) - q_1) - \tau \frac{d}{dt} \mathcal{H}(q_1(t)) \\ &\quad + (\beta_1(q_2) - q_1)^T M_1 \dot{q}_2 \\ &= -\tilde{V}_1 + \tau (\mathcal{H}(\beta_1(q_2)) - \mathcal{H}(q_1)) - \tau \frac{d}{dt} \mathcal{H}(q_1(t)) \\ &\quad + \dot{q}_1^T M_1 \dot{q}_2 \end{aligned}$$

where we used Lemma 3.2 to show that

$$\nabla_{q_2} \max_{s \in \Delta(m_2)} (s^T M_1 q_2 + \tau \mathcal{H}(s)) = \beta_1^T(q_2) M_1.$$

The lemma follows by noting that concavity of $\mathcal{H}(\cdot)$ implies that [19, Th. 25.1]

$$\begin{aligned} \mathcal{H}(\beta_1(q_2)) - \mathcal{H}(q_1) &\leq \nabla \mathcal{H}(q_1) (\beta_1(q_2) - q_1) \\ &= \frac{d}{dt} \mathcal{H}(q_1(t)). \end{aligned}$$

Similar statements apply for \tilde{V}_2 .

A. Zero-Sum and Identical Interest Games

Theorem 3.1: Assume that either

$$M_1 = -M_2^T$$

or

$$M_1 = M_2^T.$$

Then solutions of continuous-time FP (2) satisfy

$$\begin{aligned} \lim_{t \rightarrow \infty} (q_1(t) - \beta_1(q_2(t))) &= 0 \\ \lim_{t \rightarrow \infty} (q_2(t) - \beta_2(q_1(t))) &= 0. \end{aligned}$$

Proof: Define

$$V_{12}(t) = \tilde{V}_1(t) + \tilde{V}_2(t).$$

Zero-sum (see also [11]): $M_1 = -M_2^T$

In case $M_1 = -M_2^T$, summing the expressions for $\dot{\tilde{V}}_i$ in Lemma 3.1 results in a cancellation of terms, thereby producing

$$\dot{V}_{12} + V_{12} \leq 0.$$

Since

$$V_{12} \geq 0$$

with equality only at an equilibrium point of (2), the theorem follows from standard Lyapunov arguments.

Identical interest: $M_1 = M_2^T$.

By definition

$$\begin{aligned} \tilde{V}_1 - \tau (\mathcal{H}(\beta_1(q_2)) - \mathcal{H}(q_1)) &= \dot{q}_1^T M_1 q_2 \\ \tilde{V}_2 - \tau (\mathcal{H}(\beta_2(q_1)) - \mathcal{H}(q_2)) &= \dot{q}_2^T M_2 q_1. \end{aligned}$$

Therefore

$$\begin{aligned} V_{12} - \tau (\mathcal{H}(\beta_1(q_2)) - \mathcal{H}(q_1)) - \tau (\mathcal{H}(\beta_2(q_1)) - \mathcal{H}(q_2)) \\ = \dot{q}_1^T M_1 q_2 + \dot{q}_2^T M_2 q_1. \end{aligned}$$

Under the assumption $M_1 = M_2^T = M$

$$\frac{d}{dt} (q_1^T(t) M q_2(t)) = \dot{q}_1^T M_1 q_2 + \dot{q}_2^T M_2 q_1.$$

Therefore

$$\begin{aligned} V_{12} &= \tau (\mathcal{H}(\beta_1(q_2)) - \mathcal{H}(q_1)) + \tau (\mathcal{H}(\beta_2(q_1)) - \mathcal{H}(q_2)) \\ &\quad + \frac{d}{dt} (q_1(t)^T M q_2(t)). \end{aligned}$$

By concavity of $\mathcal{H}(\cdot)$ and [19, Th. 25.1]

$$\begin{aligned} \tau (\mathcal{H}(\beta_1(q_2)) - \mathcal{H}(q_1)) &\leq \tau \frac{d}{dt} \mathcal{H}(q_1(t)) \\ \tau (\mathcal{H}(\beta_2(q_1)) - \mathcal{H}(q_2)) &\leq \tau \frac{d}{dt} \mathcal{H}(q_2(t)) \end{aligned}$$

which implies that for any $T > 0$

$$\int_0^T V_{12} \leq \left(q_1^T(t) M q_2(t) + \tau \mathcal{H}(q_1(t)) + \tau \mathcal{H}(q_2(t)) \right) \Big|_{t=0}^T.$$

The integrand is positive, and $T > 0$ is arbitrary. Furthermore, one can show that \tilde{V}_{12} is bounded. Therefore, $V_{12}(t)$ asymptotically approaches zero as desired. ■

We comment that the integrability argument previously mentioned can be viewed a version of the discrete-time argument in [17], but applied to a smoothed game (i.e., $\tau > 0$) in continuous-time.

B. Two-Move Games

We now consider the case in which each player in the original static game has two moves, i.e., $m_1 = m_2 = 2$.

Continuous-time FP dynamics (2) involve differences of probability distributions. Since these distributions live on the simplex, their elements sum to one. Therefore, the sum of the elements of the *difference*

of two distributions must equal zero, i.e., differences of distributions must lie in the subspace spanned by the vector

$$N = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Using this fact, we see that necessarily

$$\dot{q}_1(t) = \beta_1(q_2(t)) - q_1(t) = N w_1(t)$$

for an appropriately defined scalar variable $w_1(t)$. Similarly

$$\dot{q}_2(t) = N w_2(t).$$

This observation will be the key to proving the desired results in the two-move game. Two separate cases will emerge

$$(N^T M_1 N)(N^T M_2 N) < 0$$

or

$$(N^T M_1 N)(N^T M_2 N) > 0.$$

In the first case, the proof will follow the same Lyapunov argument of the zero-sum proof. In the second case, the proof will follow the integrability argument of the identical interest proof. [17] suggests a link between the proofs for zero-sum, identical interest, and two-player two-move games. Namely, it states that nondegenerate two-player two-move games are best-response equivalent in mixed-strategies to an appropriate zero-sum or identical interest game, and since FP relies on best responses, this equivalence establishes convergence. The present approach does not utilize this equivalence, but does exploit the present zero-sum and identical interest proofs by establishing a direct link in terms of the constructed storage functions V_i .

Theorem 3.2: Assume that $m_1 = m_2 = 2$ and

$$(N^T M_1 N)(N^T M_2 N) \neq 0.$$

Then, solutions of continuous-time FP (2) satisfy

$$\begin{aligned} \lim_{t \rightarrow \infty} (q_1(t) - \beta_1(q_2(t))) &= 0 \\ \lim_{t \rightarrow \infty} (q_2(t) - \beta_2(q_1(t))) &= 0. \end{aligned}$$

Proof: First suppose that $N^T M_1 N$ and $N^T M_2 N$ have opposite signs. Then there exist positive scalars, α_1 and α_2 , such that

$$(N^T M_1 N)\alpha_1 + (N^T M_2 N)\alpha_2 = 0. \quad (3)$$

From Lemma 3.1

$$\begin{aligned} \dot{\tilde{V}}_1 + \tilde{V}_1 &\leq \dot{q}_1^T M_1 \dot{q}_2 \\ \dot{\tilde{V}}_2 + \tilde{V}_2 &\leq \dot{q}_2^T M_2 \dot{q}_1. \end{aligned}$$

Since $m_1 = m_2 = 2$,

$$\begin{aligned} \dot{\tilde{V}}_1 + \tilde{V}_1 &\leq N^T M_1 N w_1 w_2 \\ \dot{\tilde{V}}_2 + \tilde{V}_2 &\leq N^T M_2 N w_1 w_2. \end{aligned}$$

Scaling these equations by α_1 and α_2 , respectively, and summing them results in a cancellation of the $N^T M_i N$ terms and leads to

$$\dot{V}_{12} + V_{12} \leq 0$$

where V_{12} is now defined as

$$V_{12} = \alpha_1 \tilde{V}_1 + \alpha_2 \tilde{V}_2$$

with α_i calculated in (3). As in the zero-sum case,

$$V_{12} \geq 0$$

with equality only at an equilibrium point of continuous-time FP (2). Standard Lyapunov arguments imply the desired result.

Now, suppose that $N^T M_1 N$ and $N^T M_2 N$ have the same sign. Then, there exist positive scalars, α_1 and α_2 , such that

$$(N^T M_1 N)\alpha_1 = (N^T M_2 N)\alpha_2. \quad (4)$$

By definition

$$\begin{aligned} \tilde{V}_1 - \tau(\mathcal{H}(\beta_1(q_2)) - \mathcal{H}(q_1)) &= \dot{q}_1^T M_1 q_2 \\ \tilde{V}_2 - \tau(\mathcal{H}(\beta_2(q_1)) - \mathcal{H}(q_2)) &= \dot{q}_2^T M_2 q_1. \end{aligned}$$

Since $(1/2)NN^T$ is a projection matrix and $\dot{q}_i = N w_i$

$$\begin{aligned} \dot{q}_1^T M_1 q_2 &= \frac{(\dot{q}_1^T NN^T M_1 q_2)}{2} \\ \dot{q}_2^T M_2 q_1 &= \frac{(\dot{q}_2^T NN^T M_2 q_1)}{2}. \end{aligned}$$

Define

$$f(t) = \frac{(\dot{q}_1^T(t) NN^T M_1 q_2(t)\alpha_1 + \dot{q}_2^T(t) NN^T M_2 q_1(t)\alpha_2)}{2}.$$

Then

$$\begin{aligned} 2\dot{f} &= \dot{q}_1^T NN^T M_1 q_2 \alpha_1 + \dot{q}_2^T NN^T M_2 q_1 \alpha_2 \\ &\quad + q_1^T NN^T M_1 \dot{q}_2 \alpha_1 + q_2^T NN^T M_2 \dot{q}_1 \alpha_2 \\ &= 2(\tilde{V}_1 - \tau(\mathcal{H}(\beta_1(q_2)) - \mathcal{H}(q_1)))\alpha_1 \\ &\quad + 2(\tilde{V}_2 - \tau(\mathcal{H}(\beta_2(q_1)) - \mathcal{H}(q_2)))\alpha_2 \\ &\quad + q_1^T NN^T M_1 N w_2 \alpha_1 + q_2^T NN^T M_2 N w_1 \alpha_2. \end{aligned}$$

Define κ as

$$\kappa = (N^T M_1 N)\alpha_1 = (N^T M_2 N)\alpha_2.$$

Then

$$\begin{aligned} & q_1^T NN^T M_1 N w_2 \alpha_1 + q_2^T NN^T M_2 N w_1 \alpha_2 \\ &= (q_1^T N w_2 + q_2^T N w_1) \kappa \\ &= (q_1^T \dot{q}_2 + q_2^T \dot{q}_1) \kappa \\ &= \kappa \frac{d}{dt} (q_1^T q_2). \end{aligned}$$

Finally, define

$$V_{12} = \alpha_1 \tilde{V}_1 + \alpha_2 \tilde{V}_2$$

with α_i calculated in (4). Then, similarly to the identical interest case

$$\begin{aligned} & 2 \left(V_{12} - \alpha_1 \tau(\mathcal{H}(\beta_1(q_2)) - \mathcal{H}(q_1)) - \alpha_2 \tau(\mathcal{H}(\beta_2(q_1)) - \mathcal{H}(q_2)) \right) \\ &= 2\dot{f} - \kappa \frac{d}{dt} (q_1^T q_2) \end{aligned}$$

which, using concavity of $\mathcal{H}(\cdot)$ and [19, Th. 25.1], implies that

$$\begin{aligned} V_{12} &\leq 2\dot{f} - \kappa \frac{d}{dt} (q_1^T q_2) + \alpha_1 \tau \frac{d}{dt} \mathcal{H}(q_1(t)) \\ &\quad + \alpha_2 \tau \frac{d}{dt} \mathcal{H}(q_2(t)). \end{aligned}$$

Therefore

$$\int_0^\infty V_{12}(t) < \infty$$

which again leads to the desired result. ■

IV. TWO PLAYER GAMES WITH ONE PLAYER RESTRICTED TO TWO MOVES

Section III used energy arguments, either Lyapunov or integrability, with the same energy functions that represent the “opportunity for improvement.” Reference [2] uses properties of planar dynamical systems to establish convergence for two-player/two-move games. Reference [3] considers games in which only *one* of the players has two moves, and uses a relatively extended argument to establish convergence by eliminating the possibility of so-called Shapley polygons. In this section, we also consider games in which only one player has two moves, but we will apply properties of planar dynamical systems to provide an alternative proof.

Suppose that player \mathcal{P}_1 has only two moves, i.e., $m_1 = 2$. Following [3], we will introduce a change of variables that will lead to planar dynamics that describe the evolution of the player \mathcal{P}_1 's strategy. Define

$$\phi : \mathcal{R} \rightarrow \text{Int}(\Delta(2))$$

as

$$\phi(\delta) = \begin{pmatrix} \frac{e^\delta}{e^\delta + 1} \\ \frac{1}{e^\delta + 1} \end{pmatrix}.$$

Then, it is straightforward to show that the two-dimensional softmax function, $\sigma : \mathcal{R}^2 \rightarrow \text{Int}(\Delta(2))$ can be written as

$$\sigma \left(\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \right) = \phi(v_1 - v_2)$$

i.e., $\sigma(\cdot)$ only depends on the *difference* of v_1 and v_2 .

We will exploit this equivalence as follows. Define

$$N = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

and define the scalar

$$w_2 = \frac{1}{\tau} N^T M_1 q_2.$$

Then, a subset of the continuous FP dynamics can be written as

$$\begin{aligned} \dot{q}_1 &= \beta_1(q_2) - q_1 = \phi(w_2) - q_1 \\ \dot{w}_2 &= \frac{1}{\tau} N^T \left(M_1 \beta_2(q_1) - q_2 \right) \\ &= \frac{1}{\tau} N^T M_1 \beta_2(q_1) - w_2. \end{aligned}$$

Since q_1 evolves in the simplex interior, the scalar $w_1(t)$ is uniquely defined by

$$q_1(t) = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix} + N w_1(t).$$

Furthermore, \dot{w}_1 satisfies

$$\dot{w}_1(t) = \frac{1}{2} N^T \dot{q}_1(t).$$

The result of introducing w_1 and w_2 is that a subset of the continuous FP dynamics can be expressed completely in terms of w_1 and w_2 , namely

$$\begin{aligned} \dot{w}_1(t) &= \frac{1}{2} N^T \phi(w_2(t)) - w_1(t) \\ \dot{w}_2(t) &= \frac{1}{\tau} N^T M_1 \beta_2 \left(\begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix} + N w_1(t) \right) - w_2(t). \end{aligned} \quad (5)$$

Theorem 4.1: Assume a finite number of Nash equilibria satisfying (1). Assume further that $m_1 = 2$. Then, solutions of continuous-time FP (2) satisfy

$$\begin{aligned} \lim_{t \rightarrow \infty} (q_1(t) - \beta_1(q_2(t))) &= 0 \\ \lim_{t \rightarrow \infty} (q_2(t) - \beta_2(q_1(t))) &= 0. \end{aligned}$$

Proof: Equation (5) is planar dynamics that describe the evolution of $q_1(t)$. These dynamics form area contracting flow, due to the negative divergence of the right-hand-side. Furthermore, solutions evolve over a bounded rectangular set. A suitable modification of Bendixson's criterion [15] leads to the conclusion that the only ω -limit points are equilibria. In the original coordinates, this implies that $q_1(t)$ converges, and hence so does $q_2(t)$. ■

V. CONCLUDING REMARKS

This note has provided unified energy based convergence proofs for several special cases of games under FP. These proofs of convergence of continuous-time FP, in themselves, do not immediately guarantee the almost sure convergence of discrete-time FP. Additional arguments are needed to establish that the deterministic continuous-time limits completely capture the stochastic discrete-time limits. Such issues are discussed in general in [1] and specifically for FP in [2].

REFERENCES

- [1] M. Benaïm and M. W. Hirsch, “A dynamical systems approach to stochastic approximation,” *SIAM J. Control Optim.*, vol. 34, pp. 437–472, 1996.
- [2] —, “Mixed equilibria and dynamical systems arising from fictitious play in perturbed games,” *Games Econ. Behavior*, vol. 29, pp. 36–72, 1999.
- [3] U. Berger. Fictitious play in $2 \times n$ games. presented at Economics Working Paper Archive at WUSTL. [Online]. Available: <http://econwpa.wustl.edu/eprints/game/papers/0303/0303009.abs>
- [4] D. P. Bertsekas, *Dynamic Programming and Optimal Control*. Belmont, MA: Athena Scientific, 1995.
- [5] 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.
- [6] G. Ellison and D. Fudenberg, “Learning purified mixed equilibria,” *J. Econ. Theory*, vol. 90, pp. 83–115, 2000.
- [7] D. P. Foster and H. P. Young, “On the nonconvergence of fictitious play in coordination games,” *Games Econ. Behavior*, vol. 25, pp. 79–96.
- [8] D. Fudenberg and D. Kreps, “Learning mixed equilibria,” *Games Econ. Behavior*, vol. 5, pp. 320–367, 1993.
- [9] D. Fudenberg and D. K. Levine, *The Theory of Learning in Games*. Cambridge, MA: MIT Press, 1998.
- [10] S. Hart and A. Mas-Colell, “Uncoupled dynamics do not lead to nash equilibrium,” *Amer. Econ. Rev.*, vol. 93, no. 5, pp. 1830–1836, 2003.
- [11] J. Hofbauer and W. Sandholm, “On the global convergence of stochastic fictitious play,” *Econometrica*, vol. 70, pp. 2265–2294, 2002.
- [12] J. Jordan, “Three problems in game theory,” *Games Econ. Behavior*, vol. 5, pp. 368–386, 1993.
- [13] V. Krishna and T. Sjöström, “On the convergence of fictitious play,” *Math. Oper. Res.*, vol. 23, no. 2, pp. 479–511, 1998.
- [14] H. J. Kushner and G. G. Yin, *Stochastic Approximation Algorithms and Applications*. New York: Springer-Verlag, 1997.
- [15] C. C. McCluskey and J. S. Muldowney, “Stability implications of Bendixson's criterion,” *SIAM Rev.*, vol. 40, pp. 931–934, 1998.

- [16] K. Miyasawa, "On the convergence of learning processes in a 2×2 non-zero-sum two person game," Economic Research Program, Princeton Univ., Princeton, NJ, Tech. Rep. 33, 1961.
- [17] D. Monderer and L. S. Shapley, "Fictitious play property for games with identical interests," *J. Econ. Theory*, vol. 68, pp. 258–265, 1996.
- [18] J. Robinson, "An iterative method of solving a game," *Ann. Math.*, vol. 54, pp. 296–301, 1951.
- [19] R. T. Rockafellar, *Convex Analysis*. Princeton, NJ: Princeton Univ. Press, 1996.
- [20] L. S. Shapley, "Some Topics in Two-person Games," in *Advances in Game Theory*, L. S. Shapley, M. Dresher, and A. W. Tucker, Eds. Princeton, NJ: Princeton Univ. Press, 1964, pp. 1–29.
- [21] A. Vasin, "On stability of mixed equilibria," *Nonlinear Anal.*, vol. 38, pp. 793–802, 1999.

The Effect of Regularization on Variance Error

Brett Ninness and Håkan Hjalmarsson

Abstract—This note addresses the problem of quantifying the effect of noise induced error (so called "variance error") in system estimates found via a regularised cost criterion. It builds on recent work by the authors in which expressions for nonregularised criteria are derived which are exact for finite model order. Those new expressions were established to be very different to previous quantifications that are widely used but based on asymptotic in model order arguments. A key purpose of this note is to expose a rapprochement between these new finite model order, and the pre-existing asymptotic model order quantifications. In so doing, a further new result is established. Namely, that variance error in the frequency domain is dependent on the choice of the point about which regularization is affected.

Index Terms—Orthonormal bases, parameter estimation, system identification, variance error.

I. INTRODUCTION

When performing system identification via the widely used prediction-error method with a quadratic criterion [1], [2], then a seminal result is that under open-loop conditions the noise-induced error, as measured by the variability of the ensuing frequency response estimate $G(e^{j\omega}, \hat{\theta}_N^n)$, may be quantified via the following approximation [1], [3]–[5]:

$$\text{Var} \left\{ G(e^{j\omega}, \hat{\theta}_N^n) \right\} \approx \frac{m}{N} \frac{\Phi_v(\omega)}{\Phi_u(\omega)}. \quad (1)$$

Here, Φ_v and Φ_u are, respectively, the measurement noise and input excitation power spectral densities, and $\hat{\theta}_N^n$ is the prediction error estimate based on N observed data points of a vector $\theta^n \in \mathbf{R}^n$ that parameterizes a model structure $G(q, \theta^n)$ for which (essentially) the model order $m = \dim \theta^n / (2^d)$ where d is the number of denominator polynomials to be estimated in the model structure.

Manuscript received May 23, 2003; revised February 3, 2004. Recommended by Associate Editor E. Bai. This work was supported by the Australian Research Council. Part of this work was completed while the authors were visiting S3-Automatic Control, The Royal Institute of Technology, Stockholm, Sweden.

B. Ninness is with the School of Electrical Engineering and Computer Science, University of Newcastle, Newcastle 2308, Australia (e-mail: brett@ee.newcastle.edu.au).

H. Hjalmarsson is with the Department of Sensors, Signals and Systems (Automatic Control), The Royal Institute of Technology, S-100 44 Stockholm, Sweden (e-mail: hakan.hjalmarsson@s3.e.kth.se).

Digital Object Identifier 10.1109/TAC.2004.831089

A fundamental aspect of the approximation (1) is that it is derived by taking the limiting value of the variance as model order m tends to infinity, and then employing that limiting value as an approximation for finite m .

Motivated by the desire to improve the accuracy of variance error quantifications, [6] and [7] have derived new expressions that are exact for finite model order (although they are still based on limiting arguments with respect to observed data length N).

As discussed in [6], there can be very large discrepancies between the new quantifications derived for finite-model order [6], and the approximation (1); [6] illustrates orders of magnitude difference on a simple example.

A key purpose of this note is to address this issue and provide a rapprochement between the results. The approach taken here is to derive new quantifications that are exact for finite model order. Although finite, this order may also be arbitrarily large, provided an appropriate regularised criterion is used to ensure that at the arbitrarily large model order, the limiting (in N) estimate is uniquely defined.

Essentially, via this strategy, the work here establishes that when the regularising point (in parameter space) implies that any pole zero cancellations in the estimated model are constrained to be at the origin, then as model order m increases, the "exact" (for finite-model order) variance expression becomes arbitrarily close to the well known approximation (1). However, when the pole zero cancellations are not at the origin, the rapprochement is lost. This fact exposes the further new result that variance error (in the frequency domain) is dependent on the point about which regularization is imposed.

As overview of the organization of this note, Section II makes concrete the estimation algorithms and model structures being considered. Certain key ideas, notation and definitions are also introduced. Section III presents the main technical results, which are new variance error quantifications that are novel in that they do not depend on asymptotic in model order arguments, yet they still apply for model orders possibly greater than that of an underlying true system. Section IV discusses the ramifications and practical consequences of these results and, in particular, uses them to argue a rapprochement between new finite model order expressions [6] and pre-existing asymptotic model order approximations [3]. Section V provides concluding remarks and comments about prospective future studies.

II. PROBLEM FORMULATION

In what follows, it is assumed that the relationship between an observed input data record $\{u_t\}$ and output data record $\{y_t\}$ obeys

$$S : y_t = G(q)u_t + v_t \quad v_t = H(q)e_t \quad (2)$$

and that this is modeled according to

$$\mathcal{M} : y_t = G(q, \theta^n)u_t + H(q, \theta^n)e_t \quad (3)$$

where the "dynamics model" $G(q, \theta^n)$ and the "noise model" $H(q, \theta^n)$ are jointly parameterized by a vector $\theta^n \in \mathbf{R}^n$ and are of the rational forms $(A(q, \theta^n) - D(q, \theta^n))$ that follow are all polynomials in the backward shift operator q^{-1})

$$G(q, \theta^n) = \frac{B(q, \theta^n)}{A(q, \theta^n)} \quad H(q, \theta^n) = \frac{C(q, \theta^n)}{D(q, \theta^n)} \quad (4)$$

while $\{e_t\}$ in (3) is a zero-mean white noise sequence that satisfies $\mathbf{E}\{e_t^2\} = \sigma^2$, $\mathbf{E}\{|e_t|^8\} < \infty$.

The postulated relationship (3) can encompass a range of model structures such as FIR, ARMAX, "Output-Error," and "Box-Jenkins" [1], [2], [8]. For all these cases, since $H(q, \theta^n)$ is also constrained to be