

A Fictitious Play Approach to Large-Scale Optimization

Theodore J. Lambert III

Mathematics Department, Truckee Meadows Community College, 7000 Dandini Boulevard, Vista B200, Reno, Nevada 89512,
lamber12@tmcc.nevada.edu

Marina A. Epelman, Robert L. Smith

Department of Industrial and Operations Engineering, University of Michigan, 1205 Beal Avenue, Ann Arbor, Michigan 48109
{mepelman@umich.edu, rlsmith@umich.edu}

In this paper, we investigate the properties of the sampled version of the fictitious play algorithm, familiar from game theory, for games with identical payoffs, and propose a heuristic based on fictitious play as a solution procedure for discrete optimization problems of the form $\max\{u(y): y = (y^1, \dots, y^n) \in \mathcal{Y}^1 \times \dots \times \mathcal{Y}^n\}$, i.e., in which the feasible region is a Cartesian product of finite sets \mathcal{Y}^i , $i \in N = \{1, \dots, n\}$. The contributions of this paper are twofold. In the first part of the paper, we broaden the existing results on convergence properties of the fictitious play algorithm on games with identical payoffs to include an approximate fictitious play algorithm that allows for errors in players' best replies. Moreover, we introduce sampling-based approximate fictitious play that possesses the above convergence properties, and at the same time provides a computationally efficient method for implementing fictitious play. In the second part of the paper, we motivate the use of algorithms based on sampled fictitious play to solve optimization problems in the above form with particular focus on the problems in which the objective function $u(\cdot)$ comes from a "black box," such as a simulation model, where significant computational effort is required for each function evaluation.

Subject classifications: games/group decisions: noncooperative; programming: algorithms/heuristic; transportation: route choice.

Area of review: Optimization.

History: Received December 2002; revision received January 2004; accepted March 2004.

1. Introduction

In this paper, we investigate the properties of the sampled version of the fictitious play algorithm—familiar from game theory—for games with identical payoffs, and propose a heuristic based on fictitious play as a solution procedure for optimization problems of the form

$$\max\{u(y): y \in \mathcal{Y}\},$$

where the feasible region \mathcal{Y} is a Cartesian product of finite sets \mathcal{Y}^i , $i \in N = \{1, \dots, n\}$: $\mathcal{Y} = \mathcal{Y}^1 \times \dots \times \mathcal{Y}^n$. In view of this representation, this problem can be restated as

$$\max\{u(y): y = (y^1, \dots, y^n) \in \mathcal{Y}^1 \times \dots \times \mathcal{Y}^n\}. \quad (1)$$

The contributions of this paper are twofold. In the first part of the paper, we broaden the existing results on convergence properties of the fictitious play algorithm on games with identical payoffs to include approximate fictitious play, which allows for errors in the players' best replies. Moreover, we introduce sampling-based approximate fictitious play that retains the above convergence properties, and at the same time provides a computationally efficient method for implementing fictitious play. In the second part

of the paper, we illustrate the usefulness of the algorithm based on sampled fictitious play to solve problems of the form (1). We will be particularly interested in the instances of (1) in which the objective function $u(\cdot)$ does not possess a closed-form expression, but instead comes from, in effect, a "black box," such as a simulation model, and thus requires significant computational effort for each function evaluation. We do not make any assumptions on the objective function, or the problem in general, other than that the variables, or, more broadly, decisions, can be partitioned in such a way that the feasible region becomes a Cartesian product, as above.

Fictitious play, first introduced by Brown (1951) and Robinson (1951), is an iterative procedure applied to a finite noncooperative game in which each of the players is faced with selecting a strategy from a finite set of available strategies. At every iteration of fictitious play, each player chooses a strategy that is a best reply (with respect to that player's payoff function) to the other players' strategies, assuming that they will be chosen based on the empirical probability distribution induced by the historical frequency of their decisions in all previous iterations. Thus, fictitious play was designed to mimic the behavior of the players learning each others' strategies. Whether this procedure,

proposed as an algorithmic way of finding mixed equilibria of the underlying game, indeed converges to an equilibrium solution has been studied under various assumptions. In particular, Monderer and Shapley (1996) demonstrated that fictitious play does converge to equilibrium for games of identical interest, i.e., in which all players have identical payoff functions.

Fictitious play for a game with identical payoffs can be viewed as similar in spirit to coordinate-search algorithms for optimization, in that the single-variable (i.e., player) optimizations performed at each iteration are the main tool of the algorithm. However, in the absence of additional assumptions on the payoff function, these single-player optimizations can become prohibitively expensive—in the extreme, requiring evaluations of the payoff for all combinations of strategies. This effectively would result in a pure enumeration and evaluation of all feasible solutions. In this paper, we describe a way to sidestep these difficulties by replacing exact best-reply computations with an approximation based on sampling. We show that the resulting error of approximation disappears in the limit as the number of iterations grows, and thus convergence to equilibrium of the resulting sampled fictitious play algorithm is maintained for games with identical payoffs.

In the second part of the paper, we discuss the merits of using sampled fictitious play as a heuristic for solving optimization problems of the form (1), specifically in cases when the objective function $u(\cdot)$ possesses no special properties and requires extensive computational effort for each function call. Such complex general problems arising in a variety of applications present formidable challenges, and a vast array of optimization heuristics already exists; we will not attempt to review them here. Some of these are based on naturally self-optimizing systems. For example, simulated annealing, which as its name implies comes from an analogy to the annealing process of metals (Kirkpatrick et al. 1983), is used to solve discrete optimization problems. The framework for genetic algorithms, which can be thought of as evolutionary processes, was first proposed by Holland (1975) and Jong (1975), and has also been successfully applied to discrete optimization (Goldberg 1989). The metaphors, of course, do not by themselves establish the value or effectiveness of the algorithms, although they did motivate in part their development. With the success of these algorithms motivated by “natural” optimization methods, a natural direction might be to examine a “behavioral” optimization method, such as fictitious play. With fictitious play we animate the components of a system to play a non-cooperative game with identical payoffs, representing the performance of the system.

To be more specific, we can view the optimization problem (1) as a game Γ in which n players, each player $i \in N$ having the finite set of strategies \mathcal{Y}^i , have identical payoffs; i.e., $u^1(y) = \dots = u^n(y) = u(y)$. Note that every optimal solution of (1) corresponds to an assignment of pure strategies that is an equilibrium of Γ . Conversely, every

pure-strategy equilibrium of Γ is a *local optimum* of (1) in the sense that no component y^i can improve the objective function by unilaterally changing to a different pure or mixed strategy. In the second part of this paper, we will use sampled fictitious play as an algorithmic device to obtain a Nash equilibrium of the mixed extension of Γ , and as an optimization heuristic to guide the search of the feasible set of (1).

The practical potential of a fictitious play approach to optimization has already been demonstrated in Garcia et al. (2000). Their use of fictitious play to find system-optimal routings in a large-scale dynamic traffic network yielded substantial reductions in computational requirements at no deterioration in quality of solution obtained. However, in their implementation, simplifications were made in the fictitious play algorithm to ensure tractability. In the first part of this paper we, in effect, establish a rigorous foundation for a version of fictitious play close in spirit to that employed by them, which can be efficiently implemented as a heuristic for such complex optimization problems. Additionally, we elaborate on properties of optimization problems of the form (1) that support sampled fictitious play as a viable heuristic for their solution. As we discuss in §5, problems in which we are able to solve single-variable optimization subproblems (i.e., compute best replies) efficiently stand to benefit most from a fictitious play paradigm. Additionally, in problems for which evaluations of the objective function require significant computational effort, parallel computation of best replies allows for a significant reduction in the running time of the algorithm, as illustrated in the dynamic traffic routing problem above, and a protocol selection problem for a situation awareness application, which we also summarize (see also Lambert and Wang 2003).

This paper is organized as follows: In §2, we review the notions of a game in strategic form, a Nash equilibrium, and an approximate Nash equilibrium, as well as provide several preliminary results. We end the section with the definition of ε_t -fictitious play, which provides a structured way of studying approximate implementations of the fictitious play algorithm as well as a way to measure the quality of such approximations. Section 3 introduces *sampled fictitious play* as means for practical implementation of the fictitious play algorithm. This section contains the main theoretical results of this paper. In particular, Theorem 4 gives conditions under which an ε_t -fictitious play process converges in beliefs to equilibrium in finite games with identical payoffs. This result is then used in Theorem 5 to show that every *sampled fictitious play process*, with sample sizes growing sufficiently fast, converges in beliefs to equilibrium with probability 1 in finite games with identical payoffs. Section 4 addresses the question of uniqueness of a limit point of the sequence generated by the algorithm. Section 5 motivates and discusses the use of fictitious play as an optimization tool, and in particular, §5.2 presents two computational examples. Section 6 contains concluding remarks.

2. Games in Strategic Form and Fictitious Play

Following Monderer and Shapley (1996), let Γ be a finite game in strategic form with the set of players $N = \{1, 2, \dots, n\}$. We denote the finite set of strategies of player $i \in N$ by \mathcal{Y}^i , and let $\mathcal{Y} = \mathcal{Y}^1 \times \mathcal{Y}^2 \times \dots \times \mathcal{Y}^n$. The *payoff*, or *utility*, *function* of player $i \in N$ is $u^i: \mathcal{Y} \rightarrow \mathbb{R}$, where \mathbb{R} denotes the set of real numbers.

For $i \in N$, let Δ^i be the set of *mixed strategies* of player i . That is,

$$\Delta^i = \left\{ f^i: \mathcal{Y}^i \rightarrow [0, 1]: \sum_{y^i \in \mathcal{Y}^i} f^i(y^i) = 1 \right\}.$$

Each $f^i \in \Delta^i$ can be viewed as an assignment of probabilities, or *beliefs*, to the elements of \mathcal{Y}^i ; in particular, with a slight abuse of notation we identify the pure strategy $y^i \in \mathcal{Y}^i$ with the extreme point of Δ^i which assigns probability 1 to y^i . Set $\Delta = \Delta^1 \times \Delta^2 \times \dots \times \Delta^n$.

For $i \in N$, we extend u^i to be the payoff function of player i in the mixed extension of Γ . That is, for any $f \in \Delta$,

$$u^i(f) = u^i(f^1, f^2, \dots, f^n) = \sum_{y \in \mathcal{Y}} u^i(y^1, y^2, \dots, y^n) f^1(y^1) f^2(y^2) \dots f^n(y^n), \quad (2)$$

which can be interpreted as the expected payoff of player i with the players choosing their strategies according to the probability distributions characterized by the beliefs f^1, \dots, f^n . Note that we have assumed players choose their strategies independently.

Let $g \in \Delta$ and let $\varepsilon \geq 0$. We say that g is an ε -*equilibrium* if for each $i \in N$,

$$u^i(g) \geq u^i(f^i, g^{-i}) - \varepsilon \quad \forall f^i \in \Delta^i,$$

where $(f^i, g^{-i}) = (g^1, \dots, g^{i-1}, f^i, g^{i+1}, \dots, g^n)$. A *Nash equilibrium* is a 0-equilibrium, and will be simply referred to as an *equilibrium*.

Denote by K_ε the set of all ε -equilibria, and by K the set of all equilibria of Γ , and denote by $\|\cdot\|$ the Euclidean norm on the Euclidean space that may be viewed as containing Δ . For $\delta > 0$, set

$$B_\delta(K) = \left\{ g \in \Delta: \min_{f \in K} \|g - f\| < \delta \right\}.$$

A *belief path* is a sequence $(f(t))_{t=1}^\infty$ in Δ . We say that a belief path $(f(t))_{t=1}^\infty$ *converges to equilibrium* if each accumulation point of $(f(t))_{t=1}^\infty$ is an equilibrium; that is, if for every $\delta > 0$, there exists an integer T such that $f(t) \in B_\delta(K)$ for all $t \geq T$.

Two propositions below contain claims useful in studying convergence properties of belief paths. These claims were made without proof in Monderer and Shapley (1996); for completeness, we provide our proofs in the appendix.

PROPOSITION 1. *Suppose that the belief path $(f(t))_{t=1}^\infty$ is such that for every $\varepsilon > 0$, there exists an integer T such that $f(t) \in K_\varepsilon$ for every $t \geq T$. Then, the belief path $(f(t))_{t=1}^\infty$ converges to equilibrium.*

We say that the belief path $(f(t))_{t=1}^\infty$ *converges to equilibrium in the Cesaro mean* if

$$\lim_{T \rightarrow \infty} \frac{\text{card}(\{1 \leq t \leq T: f(t) \notin B_\delta(K)\})}{T} = 0 \quad \text{for every } \delta > 0, \quad (3)$$

where $\text{card}(A)$ is the cardinality of the set A . Clearly, convergence to equilibrium implies convergence to equilibrium in the Cesaro mean. The following proposition gives a sufficient condition for a belief path converging to equilibrium in the Cesaro mean.

PROPOSITION 2. *Suppose that the belief path $(f(t))_{t=1}^\infty$ is such that for every $\varepsilon > 0$,*

$$\lim_{T \rightarrow \infty} \frac{\text{card}(\{1 \leq t \leq T: f(t) \notin K_\varepsilon\})}{T} = 0. \quad (4)$$

Then, the belief path converges to equilibrium in the Cesaro mean.

A *path* in \mathcal{Y} is a sequence $(y(t))_{t=1}^\infty$ of elements of \mathcal{Y} . To each path $(y(t))_{t=1}^\infty$, we naturally associate a belief path $(f_y(t))_{t=1}^\infty$ by letting

$$f_y(t) = \frac{1}{t} \sum_{s=1}^t y(s) \quad \text{for every } t \geq 1. \quad (5)$$

In (5) the $y(s)$ s should be viewed as elements of Δ . Note that for any integer $k \geq 0$,

$$f_y(t+k) = f_y(t) + \frac{1}{t+k} \sum_{s=t+1}^{t+k} (y(s) - f_y(t)). \quad (6)$$

We say that a path $(y(t))_{t=1}^\infty$ *converges in beliefs to equilibrium (in the Cesaro mean)* if the associated belief path $(f_y(t))_{t=1}^\infty$ defined by (5) converges to equilibrium (in the Cesaro mean). The following lemma establishes the equivalence of these two types of convergence.

LEMMA 3 (MONDERER AND SHAPLEY 1996). *For every game in strategic form, a path $(y(t))_{t=1}^\infty$ converges in beliefs to equilibrium if and only if it converges in beliefs to equilibrium in the Cesaro mean.*

This lemma was stated, with a sketch of a proof, by Monderer and Shapley (1996) for a subset of paths that are *fictitious play processes*, to be defined shortly. The above statement emphasizes that the result can be established for any path $(y(t))_{t=1}^\infty$. For completeness, we provide a detailed proof in the appendix.

We now formally define a fictitious play process. For $i \in N$ and for $f \in \Delta$, let

$$v^i(f) = \max\{u^i(g^i, f^{-i}): g^i \in \Delta^i\}.$$

That is, $v^i(f)$ is the payoff of player i 's best response to the other players' strategies f^{-i} . Note from the definition of $u^i(f)$ that $v^i(f)$ can always be attained by an extreme point of Δ^i ; i.e., $\max\{u^i(g^i, f^{-i}): g^i \in \Delta^i\} = \max\{u^i(y^i, f^{-i}): y^i \in \mathcal{Y}^i\}$. A path $(y(t))_{t=1}^\infty$ is a *fictitious play process* if for every $i \in N$,

$$u^i(y^i(t+1), f_y^{-i}(t)) = v^i(f_y(t)) \quad \text{for every } t \geq 1; \quad (7)$$

i.e., $y^i(t+1)$ is a best response of player i to the mixed strategies of the other players, as represented by the beliefs $f_y^{-i}(t)$.

For a sequence of vectors $(\varepsilon_t)_{t=1}^\infty$, where $\varepsilon_t \in \mathbb{R}^n$ and $\varepsilon_t \geq 0$ for all t , we say that a path $(y(t))_{t=1}^\infty$ is an ε_t -fictitious play process if for every $i \in N$,

$$u^i(y^i(t+1), f_y^{-i}(t)) \geq v^i(f_y(t)) - \varepsilon_t^i \quad \text{for every } t \geq 2. \quad (8)$$

The ε_t^i 's can be viewed as player errors in their respective optimizations (i.e., computations of best replies). A standard fictitious play process can be characterized as an ε_t -fictitious play process with $\varepsilon_t = 0$ for every $t \geq 1$; i.e., the players make no errors.

As above, we say that a (ε_t) -fictitious play process $(y(t))_{t=1}^\infty$ converges in beliefs to equilibrium (in the Cesaro mean) if the associated belief path converges to equilibrium (in the Cesaro mean).

3. Approximate Fictitious Play and Sampled Fictitious Play Algorithms

In this section, we focus our attention on games with *identical payoffs* (also referred to as games with identical interests, or common utility games), i.e., games in strategic form in which all the players share a common payoff function $u^1(\cdot) = \dots = u^n(\cdot) = u(\cdot)$. Monderer and Shapley (1996) demonstrated that for every finite game with identical payoffs, every fictitious play process converges to equilibrium. Based on this result, a straightforward algorithmic approach for computing equilibria of any such game can be proposed in the following form:

Fictitious Play Algorithm

Initialization: Set $t = 1$ and select $y(1) \in \mathcal{Y} = \mathcal{Y}^1 \times \mathcal{Y}^2 \times \dots \times \mathcal{Y}^n$ arbitrarily; set $f_y(1) = y(1)$.

Iteration $t \geq 1$: Given $f_y(t)$, find

$$y^i(t+1) \in \arg \max_{y^i \in \mathcal{Y}^i} u^i(y^i, f_y^{-i}(t)), \quad i = 1, \dots, n. \quad (9)$$

Set

$$f_y(t+1) = f_y(t) + \frac{1}{t+1}(y(t+1) - f_y(t))$$

and increment t by 1.

However, a straightforward implementation of the above algorithm will usually be impractical. Performing the n individual player optimizations (9) in the iterative step of the algorithm can be prohibitively expensive—in the extreme, if the vector $f_y(t)$ is strictly positive, computing $u(y^i, f_y(t)^{-i})$ can require evaluating the payoff function $u(\cdot)$ for *all* combinations of pure strategies! We propose to sidestep these difficulties by replacing exact computations in (9) with an approximation based on sampling. We will later show that the resulting error of approximation disappears in the limit as the number of iterations grows.

3.1. Approximate Fictitious Play

We begin by showing that an ε_t -fictitious play process will converge to a Nash equilibrium in finite games with identical payoffs, if the sequence $(\varepsilon_t)_{t=1}^\infty$ converges to zero fast enough. The following theorem is an extension of a theorem first proved by Monderer and Shapley (1996). The original result applied to fictitious play processes (i.e., ε_t -fictitious play processes with $\varepsilon_t = 0$) in finite games with identical payoffs, and thus is a special case of our Theorem 4. The proof of Theorem 4 is in the same spirit as the proof of Monderer and Shapley.

THEOREM 4. *Let Γ be a finite game in strategic form with identical payoff functions $u^1(\cdot) = \dots = u^n(\cdot) = u(\cdot)$. Then, any ε_t -fictitious play process with $\max_i \{\varepsilon_t^i\} = O(1/t^\alpha)$ for some $\alpha > 0$, converges in beliefs to equilibrium.*

PROOF. Because the payoff functions in Γ are identical, the payoff functions u^i in the mixed extension of Γ defined by (2) are also identical, and will be denoted by u . Note that u is a multilinear function; in particular, for any player $i \in N$, if $f \in \Delta$, $g^i \in \Delta^i$, and $\lambda \in [0, 1]$, then $f^i + \lambda(g^i - f^i) \in \Delta^i$ and

$$\begin{aligned} u(f^i + \lambda(g^i - f^i), f^{-i}) \\ = u(f^i, f^{-i}) + \lambda(u(g^i, f^{-i}) - u(f^i, f^{-i})). \end{aligned} \quad (10)$$

Let $(f_y(t))_{t=1}^\infty$ be the belief process associated with an ε_t -fictitious play process $(y(t))_{t=1}^\infty$. Also, let $L = \max_{f \in \Delta} |u(f)|$.

Equation (6) implies that

$$\begin{aligned} u(f_y(t+1)) = u \left(f_y^1(t) + \frac{1}{t+1}(y^1(t+1) - f_y^1(t)), \right. \\ \left. f_y^2(t) + \frac{1}{t+1}(y^2(t+1) - f_y^2(t)), \dots, \right. \\ \left. f_y^n(t) + \frac{1}{t+1}(y^n(t+1) - f_y^n(t)) \right). \end{aligned}$$

Applying the multilinearity property (10) repeatedly, we obtain

$$\begin{aligned} u(f_y(t+1)) \\ = u(f_y(t)) + \frac{1}{t+1} \sum_{i=1}^n [u(y^i(t+1), f_y^{-i}(t)) - u(f_y(t))] \end{aligned}$$

+ $\left((3^n - (2n + 1)) \right)$ terms of the form
 $\pm \frac{1}{(t+1)^j} u(g)$ with $j \geq 2, g \in \Delta$.

Each of the latter terms is bounded below by $-(1/(t+1)^2)L$, so by (8),

$$u(f_y(t+1)) - u(f_y(t)) \geq \frac{1}{t+1} \sum_{i=1}^n [v^i(f_y(t)) - u(f_y(t)) - \varepsilon_i^i] - \frac{3^n L}{(t+1)^2}. \quad (11)$$

For $t \geq 1$, set

$$a_t = \sum_{i=1}^n [v^i(f_y(t)) - u(f_y(t))]. \quad (12)$$

Note that each term in the above summation is nonnegative, and, in particular, $a_t \geq 0$ for every $t \geq 1$. By (11),

$$\frac{a_t}{t} \leq \frac{t+1}{t} (u(f_y(t+1)) - u(f_y(t))) + 3^n L \frac{1}{t(t+1)} + \frac{n \max_i \{\varepsilon_i^i\}}{t}. \quad (13)$$

We will establish that $\sum_{i=1}^{\infty} a_i/t = \lim_{T \rightarrow \infty} \sum_{i=1}^T a_i/t < \infty$. First, note that

$$\sum_{i=1}^{\infty} \frac{1}{t(t+1)} = \sum_{i=1}^{\infty} \left(\frac{1}{t} - \frac{1}{t+1} \right) = 1.$$

Next, consider

$$\begin{aligned} & \sum_{i=1}^T \frac{t+1}{t} (u(f_y(t+1)) - u(f_y(t))) \\ &= \sum_{i=1}^{T-1} \frac{1}{t(t+1)} u(f_y(t+1)) + \frac{T+1}{T} u(f_y(T+1)) - 2u(f_y(1)), \end{aligned}$$

and hence,

$$\sum_{i=1}^{\infty} \frac{t+1}{t} (u(f_y(t+1)) - u(f_y(t))) \leq 4L.$$

Finally, because $\max_i \{\varepsilon_i^i\} \rightarrow 0$ at an asymptotic order of $1/t^\alpha$ for some $\alpha > 0$, as $t \rightarrow \infty$,

$$\sum_{i=1}^{\infty} \frac{n \max_i \{\varepsilon_i^i\}}{t} < \infty,$$

because

$$\sum_{i=1}^{\infty} \frac{1}{t^{\alpha+1}} \leq 1 + \int_2^{\infty} \frac{1}{t^{\alpha+1}} dt < \infty \quad \text{for } \alpha > 0.$$

Combining these bounds with (13), we conclude that

$$\sum_{i=1}^{\infty} \frac{a_i}{t} < \infty. \quad (14)$$

We next prove that

$$\lim_{T \rightarrow \infty} \frac{a_1 + a_2 + \dots + a_T}{T} = 0. \quad (15)$$

For $T \geq 1$, set $b_T = \sum_{i=T}^{\infty} (a_i/t)$. By (14), $\lim_{T \rightarrow \infty} b_T = 0$. Therefore,

$$\lim_{T \rightarrow \infty} \frac{b_1 + b_2 + \dots + b_T}{T} = 0,$$

which, via algebraic manipulation, implies

$$\begin{aligned} & \lim_{T \rightarrow \infty} \left(\frac{a_1 + a_2 + \dots + a_T}{T} + b_{T+1} \right) \\ &= \lim_{T \rightarrow \infty} \frac{b_1 + b_2 + \dots + b_T}{T} = 0. \end{aligned}$$

Because $\lim_{T \rightarrow \infty} b_{T+1} = 0$, (15) follows.

Let $\varepsilon > 0$ be chosen arbitrarily. From (12), we conclude that if $f_y(t) \notin K_\varepsilon$, then $a_t \geq \varepsilon$. Therefore,

$$\text{card}(\{1 \leq t \leq T: f_y(t) \notin K_\varepsilon\}) \leq \frac{1}{\varepsilon} (a_1 + a_2 + \dots + a_T),$$

and so

$$\begin{aligned} & \lim_{T \rightarrow \infty} \frac{\text{card}(\{1 \leq t \leq T: f_y(t) \notin K_\varepsilon\})}{T} \\ & \leq \frac{1}{\varepsilon} \lim_{T \rightarrow \infty} \frac{a_1 + a_2 + \dots + a_T}{T} = 0. \end{aligned}$$

Therefore, (4) is satisfied, and the proof follows from Proposition 2 and Lemma 3. \square

The previous theorem guarantees convergence to the set of Nash equilibria as long as $\max_i \{\varepsilon_i^i\} \rightarrow 0$ at an asymptotic order of $1/t^\alpha$ for some $\alpha > 0$. This means that the players need not start out as “good optimizers” as long as they sufficiently improve over time. In the following subsection, we propose a mechanism based on statistical sampling that achieves this desired error reduction.

3.2. Sampled Fictitious Play

Recall that the components of the belief vector $f_y(t)$ can be viewed as components of a probability distribution over the pure strategies of the players. From this perspective, the one-dimensional optimization (9) can be viewed as

$$\max_{y^i \in \mathcal{Y}^i} \{E_{Y^{-i}}(u(y^i, Y^{-i}))\},$$

where Y^{-i} is a random vector whose components, Y^j , $j \neq i$, have probability distribution described by $f_y^j(t)$, respectively.

We define the function $\bar{U}_k^i(\cdot, f_y^{-i}(t)): \mathcal{Y}^i \rightarrow \mathbb{R}$ by

$$\bar{U}_k^i(y^i, f_y^{-i}(t)) = \sum_{j=1}^k \frac{u^i(y^i, Y_j^{-i}(t))}{k}, \quad (16)$$

where $Y_j^{-i}(t)$, $j = 1, \dots, k$, are i.i.d. random vectors with the distribution given by $f_y^{-i}(t)$. Then, $\bar{U}_k^i(y^i, f_y^{-i}(t))$ can be viewed as a sample mean (with sample size k) of player i 's payoff when using strategy y^i .

A typical iteration of the sampled fictitious play algorithm is as follows. At the beginning of iteration t , a vector of beliefs $f_y(t) \in \Delta$ is given. For the sample size $k_t \geq 1$, we draw i.i.d. random samples $Y_j(t)$, $j = 1, \dots, k_t$, from the distribution given by $f_y(t)$, and let $\bar{u}_{k_t}^i(y^i, f_y^{-i}(t))$ denote the resulting realization of $\bar{U}_{k_t}^i(y^i, f_y^{-i}(t))$ for $i = 1, \dots, n$. Each player chooses a "best response" based on sample means (instead of the expectations); i.e., $y^i(t+1)$ is chosen so that $y^i(t+1) \in \arg \max_{y^i \in \mathcal{Y}^i} \{\bar{u}_{k_t}^i(y^i, f_y^{-i}(t))\}$, and the belief vector is updated. Note that the sampled mixed strategy will, by design, have a limited number of positive components, making the computation of best replies an easier task.

The formal statement of the sampled fictitious play algorithm is presented below.

Sampled Fictitious Play Algorithm

Initialization: Set $t = 1$ and select $y(1) \in \mathcal{Y} = \mathcal{Y}^1 \times \mathcal{Y}^2 \times \dots \times \mathcal{Y}^n$ arbitrarily; set $f_y(1) = y(1)$.

Iteration $t \geq 1$: Given $f_y(t)$, select a sample size $k_t \geq 1$, and draw i.i.d. random samples $Y_j(t)$, $j = 1, \dots, k_t$, from the distribution given by $f_y(t)$. Using the above sample, find

$$y^i(t+1) \in \arg \max_{y^i \in \mathcal{Y}^i} \{\bar{u}_{k_t}^i(y^i, f_y^{-i}(t))\}, \quad i = 1, \dots, n, \quad (17)$$

where $\bar{u}_{k_t}^i(y^i, f_y^{-i}(t))$ is the realization of $\bar{U}_{k_t}^i(y^i, f_y^{-i}(t))$ as defined by (16). Set

$$f_y(t+1) = f_y(t) + \frac{1}{t+1}(y(t+1) - f_y(t)),$$

increment t by 1.

In the analysis of the above algorithm, the following consideration must be kept in mind: If $y(t)$ were defined by (9), $(y(t))_{t=1}^\infty$ would be a deterministic sequence, assuming a deterministic tie-breaking rule for the single-player optimizations, as would be the associated belief path $(f_y(t))_{t=1}^\infty$. However, with sampling, $(y(t))_{t=1}^\infty$ is now a stochastic process, and we will denote the associated (stochastic) belief process as $(F_y(t))_{t=1}^\infty$. We will call the stochastic process $(y(t))_{t=1}^\infty$ a *sampled fictitious play process*. The reader should note that in Equation (17) the sample mean was generated conditional on $F_y^{-i}(t) = f_y^{-i}(t)$.

The next theorem shows that it suffices to choose our sample sizes $k_t = \lceil Ct^\beta \rceil$ with $\beta > 1/2$ and $C > 0$, where $\lceil Ct^\beta \rceil$ is the smallest integer greater than or equal to Ct^β , to ensure convergence of a sampled fictitious play process to equilibrium in a finite game with identical payoffs.

THEOREM 5. *Let Γ be a finite game in strategic form with identical payoffs. Then, any sampled fictitious play process $y(t)$ with sample sizes $k_t = \lceil Ct^\beta \rceil$ for $\beta > 1/2$ and $C > 0$ converges in beliefs to equilibrium with probability 1.*

PROOF. Let $(y(t))_{t=1}^\infty$ be a sampled fictitious play process with sample sizes $k_t = \lceil Ct^\beta \rceil$, and $(F_y(t))_{t=1}^\infty$ be the associated belief process. We begin by establishing a bound on the random variable

$$\bar{U}_{k_t}^i(y^i, F_y^{-i}(t)) - u(y^i, F_y^{-i}(t))$$

for an arbitrary $y^i \in \mathcal{Y}^i$. Fix t , let $y^i \in \mathcal{Y}^i$, and define

$$X_j(t) = u(y^i, Y_j^{-i}(t)) - u(y^i, F_y^{-i}(t)), \quad j = 1, \dots, k_t,$$

where the $Y_j^{-i}(t)$ are random vectors with distribution $F_y^{-i}(t)$. The $X_j(t)$ s are not independent random variables as they are described by functions of the same random variable $F_y^{-i}(t)$; also, the $Y_j^{-i}(t)$ s are dependent both on $F_y^{-i}(t)$ and on each other. However, for a fixed value of t and conditional on $F_y^{-i}(t) = f_y^{-i}(t)$, vectors $Y_j^{-i}(t)$, $j = 1, \dots, k_t$, are i.i.d. with distribution $f_y^{-i}(t)$, and $u(y^i, f_y^{-i}(t))$ is a constant. With such conditioning, therefore, $X_1(t), \dots, X_{k_t}(t)$ are i.i.d. random variables with mean 0 (by the law of the unconscious statistician).

Let $L = \max_{f \in \Delta} |u(f)|$ and let A denote the event that $F_y^{-i}(t) = f_y^{-i}(t)$. Then, $E[X_j(t)^4 | A] \leq (2L)^4$ and $E[X_j(t)^2 | A] \leq (2L)^2$ for all $j = 1, \dots, k_t$.

Let $S(t) = \sum_{j=1}^{k_t} X_j(t)$. Then,

$$\begin{aligned} E[S(t)^4 | A] &= k_t E[X_1(t)^4 | A] + \binom{4}{2} \binom{k_t}{2} E[X_1(t)^2 X_2(t)^2 | A] \\ &\leq k_t (2L)^4 + (3k_t^2 - 3k_t) E[X_1^2 | A] E[X_2^2 | A] \leq 3(2L)^4 k_t^2. \end{aligned}$$

Neither of the above bounds depended on $f_y^{-i}(t)$; therefore, unconditionally,

$$E \left[\frac{S(t)^4}{k_t^4} \right] \leq \frac{3(2L)^4 k_t^2}{k_t^4} = \frac{3(2L)^4}{\lceil Ct^\beta \rceil^2} \leq \frac{3(2L)^4}{C^2 t^{2\beta}}. \quad (18)$$

By the Markov inequality, for any $\delta > 0$,

$$P \left\{ \frac{S(t)^4}{k_t^4} > \delta t^{0.5-\beta} \right\} \leq \frac{E[S(t)^4/k_t^4]}{\delta t^{0.5-\beta}}.$$

Combining this with (18),

$$\begin{aligned} \sum_{t=1}^\infty P \left\{ \frac{S(t)^4}{k_t^4} > \delta t^{0.5-\beta} \right\} &\leq \sum_{t=1}^\infty \frac{E[S(t)^4/k_t^4]}{\delta t^{0.5-\beta}} \\ &\leq \sum_{t=1}^\infty \frac{3(2L)^4}{C^2 \delta t^{\beta+0.5}} < \infty. \end{aligned}$$

By the Borel-Cantelli lemma (which does not require independence of events; see Ross 1996), we have, with probability 1, $(S(t)^4/k_t^4) \rightarrow 0$ at an asymptotic order of $1/t^{\beta-0.5}$. The previous argument implies that, with probability 1, $(S(t)/k_t) \rightarrow 0$ at an asymptotic order of $1/t^\alpha$, where $\alpha = (\beta - 0.5)^{0.25}$. Because

$$\frac{S(t)}{k_t} = \bar{U}_{k_t}^i(y^i, F_y^{-i}(t)) - u(y^i, F_y^{-i}(t)),$$

we conclude that $\bar{U}_{k_i}^i(y^i, F_t^{-i}) - u(y^i, F_y^{-i}(t))$ converges to 0 at an asymptotic rate of $1/t^\alpha$, where $\alpha = (\beta - 0.5)^{0.25}$, for any $y^i \in \mathcal{Y}^i$, with probability 1.

For the sampled fictitious play process $(y(t))_{t=1}^\infty$, define

$$\varepsilon_t^i = v^i(F_y(t)) - u(y^i(t+1), F_y^{-i}(t)) \geq 0$$

for $i \in N$ and $t \geq 1$. (As before, the vectors $(\varepsilon_t)_t$ can be interpreted as the errors in the players' responses at iteration t as reflected by the utility function. In this case, the errors are brought on by the players optimizing the sample means of their payoff functions instead of the true payoff functions, and hence form a stochastic process.) We will show that, with probability 1, for any player $i \in N$, $\varepsilon_t^i \rightarrow 0$ as $t \rightarrow \infty$ at an asymptotic order of $t^{-\alpha}$, with $\alpha = (\beta - 0.5)^{0.25} > 0$. By Theorem 4, this would imply that the path $(y(t))_{t=1}^\infty$ converges in beliefs to equilibrium with probability 1.

Let $\tilde{y}^i \in \arg \max_{y^i \in \mathcal{Y}^i} u^i(y^i, f_y^{-i}(t))$. Conditioning on the event $F_y^{-i}(t) = f_y^{-i}(t)$,

$$\begin{aligned} 0 \leq \varepsilon_t^i &= v^i(f_y(t)) - u(y^i(t+1), f_y^{-i}(t)) \\ &= u(\tilde{y}^i, f_y^{-i}(t)) - u(y^i(t+1), f_y^{-i}(t)) \\ &= u(\tilde{y}^i, f_y^{-i}(t)) - \bar{u}_{k_i}^i(\tilde{y}^i, f_y^{-i}(t)) \\ &\quad + (\bar{u}_{k_i}^i(\tilde{y}^i, f_y^{-i}(t)) - \bar{u}_{k_i}^i(y^i(t+1), f_y^{-i}(t))) \\ &\quad + (\bar{u}_{k_i}^i(y^i(t+1), f_y^{-i}(t)) - u(y^i(t+1), f_y^{-i}(t))) \\ &\leq (u(\tilde{y}^i, f_y^{-i}(t)) - \bar{u}_{k_i}^i(\tilde{y}^i, f_y^{-i}(t))) \\ &\quad + (\bar{u}_{k_i}^i(y^i(t+1), f_y^{-i}(t)) - u(y^i(t+1), f_y^{-i}(t))), \end{aligned}$$

where the last inequality follows because $y^i(t+1)$ is chosen to maximize $\bar{u}_{k_i}^i(y^i, f_y^{-i}(t))$. The above bound did not depend on a particular realization $f_y^{-i}(t)$ of $F_y^{-i}(t)$; therefore, we have unconditionally

$$\begin{aligned} 0 \leq \varepsilon_t^i &\leq (u(\tilde{y}^i, F_y^{-i}(t)) - \bar{U}_{k_i}^i(\tilde{y}^i, F_y^{-i}(t))) \\ &\quad + (\bar{U}_{k_i}^i(y^i(t+1), F_y^{-i}(t)) - u(y^i(t+1), F_y^{-i}(t))). \end{aligned}$$

Applying the derived asymptotic rate of convergence to the two terms of the above bound, we conclude that, with probability 1, ε_t^i converges to 0 at an asymptotic order of $1/t^\alpha$, where $\alpha = (\beta - 0.5)^{0.25}$, establishing the desired result. \square

As a final note, observe that because the size of the sample taken increases with the iteration count, the number of positive components in the sampled mixed strategies may increase. However, as the probability distribution from which the sample is taken approaches its limit, the resulting sampled mixed strategies will tend to have no more positive components than the limiting equilibrium strategy.

4. Limiting Behavior of Belief Paths

Theorems 4 and 5 only stipulate that belief paths converge to equilibrium; i.e., all of their accumulation points

are contained in the set of Nash equilibria. In an algorithmic setting, however, it is desirable to have a guarantee of uniqueness of the limit point of the sequence of iterates (for example, having this guarantee typically leads to simpler and more reasonable stopping rules). In this section, we derive conditions under which a belief path associated with an arbitrary path that converges in beliefs to equilibrium has a *unique* limit point that is a Nash equilibrium. Note that the discussion in this section is applicable to any finite game in strategic form (not just a game with identical payoffs) and arbitrary paths converging to equilibrium in beliefs. The application of these results to fictitious play processes, or appropriately designed approximate and sampled fictitious play processes, in games with identical payoffs is immediate in view of Theorems 4 and 5.

Recall that a set S in a metric space is *perfect* if it is closed and every point of S is an accumulation point of S , i.e., if S is its accumulation points (see, for example, Munkres 1975).

LEMMA 6. *Let Γ be a finite n -player game in strategic form. The set F of limit points of the belief path associated with any path $(y(t))_{t=1}^\infty \subset \mathcal{Y}$ is a connected subset of Δ under the Euclidean norm. Moreover, F is either a singleton or a perfect set.*

PROOF. Let $(y(t))_{t=1}^\infty$ be an arbitrary path, and $(f_y(t))_{t=1}^\infty$ be the associated belief process. Let F be the set of all limit points of $(f_y(t))_{t=1}^\infty$.

Suppose that F is disconnected; i.e., there exist two disjoint open sets A and B such that

$$(A \cap F) \neq \emptyset, \quad (B \cap F) \neq \emptyset, \quad \text{and} \quad F \subset A \cup B. \quad (19)$$

Let $C = \Delta \setminus (A \cup B)$. Then, C is a compact subset of Δ and $C \neq \emptyset$ (otherwise $A \cup B \supseteq \Delta$, which would imply that Δ is disconnected). We will show that C must contain a limit point of $(f_y(t))_{t=1}^\infty$, resulting in a contradiction.

By definition, both A and B contain an infinite number of points of the sequence $(f_y(t))_{t=1}^\infty$. In particular, the sequence $(f_y(t))_{t=1}^\infty$ leaves the set A infinitely often. More precisely, there exists an infinite subsequence $(f_y(t_i))_{i=1}^\infty$ such that $f_y(t_i) \in A$, but $f_y(t_i + 1) \notin A$; i.e., either $f_y(t_i + 1) \in B$ or $f_y(t_i + 1) \in C$. Note that if $f_y(t_i + 1) \in B$, then because $A \cap B = \emptyset$, the line segment connecting $f_y(t_i)$ and $f_y(t_i + 1)$ contains at least one point $\tilde{f}_y(t_i + 1) \notin A \cup B$. By the convexity of the set Δ , $\tilde{f}_y(t_i + 1) \in \Delta$, and hence $\tilde{f}_y(t_i + 1) \in C$.

Consider the following infinite sequence: For each i , let

$$g_i = \begin{cases} f_y(t_i + 1) & \text{if } f_y(t_i + 1) \in C, \\ \tilde{f}_y(t_i + 1) & \text{otherwise,} \end{cases}$$

where $\tilde{f}_y(t_i + 1)$ is an arbitrary point constructed as above. Note that $(g_i)_{i=1}^\infty \subset C$, and hence (passing to a subsequence if necessary) $\lim_{i \rightarrow \infty} g_i = g \in C$. We will show that g is a limit point of $(f_y(t))_{t=1}^\infty$, leading to a contradiction.

By construction,

$$\begin{aligned} \|f_y(t_i) - g_i\| &\leq \|f_y(t_i) - f_y(t_i + 1)\| \\ &= \frac{1}{1 + t_i} \|y(t_i + 1) - f_y(t_i)\| < \frac{M}{t_i}, \end{aligned}$$

where $M = \max_{f, g \in \Delta} \|f - g\|$. Therefore,

$$\begin{aligned} \|f_y(t_i) - g\| &\leq \|f_y(t_i) - g_i\| + \|g_i - g\| \\ &< \frac{M}{t_i} + \|g_i - g\| \rightarrow 0 \quad \text{as } i \rightarrow \infty, \end{aligned}$$

establishing the desired contradiction and thus proving the first statement of the lemma.

To establish the second claim, assume that the set F contains more than one element. Suppose that there exists $d \in F$, which is not an accumulation point of F . Then, for some $\delta > 0$, $B_\delta(d) \cap F = \{d\}$ (here $B_\delta(d)$ denotes the open ball of radius δ centered at point d). Let $A = B_{\delta/2}(d)$, and let B be defined as the complement of the closure of $B_{\delta/2}(d)$. Then, the sets A and B are disjoint open sets satisfying conditions (19). This implies that the set F is disconnected, resulting in a contradiction—thus establishing the second claim of the lemma. \square

The following is an immediate corollary of the previous lemma.

COROLLARY 7. *Let Γ be a finite n -player game in strategic form. The image under the payoff function $u^i(f)$ of the set of limit points of the belief path associated with any path $(y(t))_{t=1}^\infty \in \Delta$ is a closed interval in \mathbb{R} .*

PROOF. Because $u^i(f)$ is a continuous function, the image under $u^i(f)$ of a perfect connected set or singleton is a perfect connected set or singleton (Munkres 1975), which in \mathbb{R} is a closed interval. Combining this observation with Lemma 6 provides the desired result. \square

Using Lemma 6, we are able to give sufficient conditions under which the belief path associated with a path will converge to a single point in the set of Nash equilibria.

THEOREM 8. *Let Γ be a finite n -player game in strategic form. If the set $K = K(\Gamma)$ of Nash equilibria has a finite set of accumulation points, then for any path converging in beliefs to equilibrium, the associated belief path will converge to a single point in the set of Nash equilibria.*

PROOF. By Lemma 6, for any path $(y(t))_{t=1}^\infty \subset \mathcal{Y}$, the set F of limit points of the associated belief path is either a singleton or a perfect set. The case when F is a singleton is trivial. If F is a perfect set, every element of F is an accumulation point of F , and hence F is an infinite set. On the other hand, $F \subseteq K$, and hence every point of F is an accumulation point of K . Because K has only finitely many accumulation points, we conclude that F must be a singleton. \square

It has been established in game theory literature (see, for example, Wilson 1971, Rosenmuller 1971, Harsanyi 1973) that, apart from certain degenerate cases, the number of equilibrium points in any finite game is finite and odd, suggesting that the assertion of the theorem is satisfied in most games.

5. Fictitious Play and Optimization

5.1. Fictitious Play Paradigm in Optimization

Having defined and investigated the properties of the fictitious play algorithm and its variants for games with identical payoffs, we will now discuss their use as an optimization heuristic. Recall that we are considering the following optimization problem:

$$\max\{u(y): y = (y^1, \dots, y^n) \in \mathcal{Y}^1 \times \dots \times \mathcal{Y}^n\}, \quad (20)$$

where the \mathcal{Y}^i , $i \in N = \{1, \dots, n\}$ are finite sets.

We can associate with the optimization problem (20) an n -player game Γ with the players $i \in N$ having the finite set of strategies \mathcal{Y}^i and sharing identical payoffs, i.e., $u^1(y) = \dots = u^n(y) = u(y)$. We can also define the mixed extension of Γ and the set of its Nash equilibria. Note that every optimal solution of (20) corresponds to an assignment of pure strategies that is a Nash equilibrium of Γ . Conversely, every pure-strategy equilibrium of Γ is a local optimum of (20) in the sense that no component y^i can improve the objective function by unilaterally changing to a different pure or mixed strategy.

The above observation suggests the use of fictitious play as an optimization heuristic for solving (20): An appropriately designed version of fictitious play can be used as an algorithmic way to obtain a Nash equilibrium of the mixed extension of Γ . The obtained equilibrium will serve as a surrogate for the optimal solution. In fact, an equilibrium is, by definition, a playerwise maximum of the mixed extension of the payoff function. Thus, by finding a Nash equilibrium we find a local optimum of the mixed extension of (20), in the above sense, and if a pure-strategy equilibrium is found, then it is a coordinate-wise optimum in the usual sense.

In the remainder of this section, we will use the terms “variable” and “player,” and “solution” and “strategy” interchangeably, in view of the above association.

Certainly not every problem of the form (20) can be successfully approached by the (sampled) fictitious play algorithm. As usual, before applying the algorithm to a specific problem, it is worthwhile to consider whether the problem properties indicate that this heuristic will be a viable solution approach. One possible consideration is the fact that computation of best replies is the main component of the algorithm. Therefore, the fictitious play heuristic is an especially promising method when this computation can be done efficiently, for example, when it can be formulated as a well-solved subproblem.

Another promising scenario is if the number of strategies available to each player is relatively small, and it is the large number of players and/or lack of favorable assumptions on the structural properties of the function $u(\cdot)$ that make the problem difficult to solve. In this case, best reply for each player can be computed by simply enumerating and comparing all possible strategies available to that player. We noted in §3.2 that computing best replies to the *sample* of the other players' strategies is still relatively easy, even in this setting. Moreover, after the strategies of the players have been sampled, the computation (17) of best replies for each player can be done *in parallel*. The “wall-clock” time of this computation can thus be reduced to that of one best reply, if a sufficient number of processors is available.

As we have mentioned already, fictitious play can be likened to coordinate search. It should be noted, however, that while the coordinate search algorithm only updates the value of one variable per iteration, the sampled fictitious play algorithm updates the mixed strategies of all the players (using the same amount of wall-clock time, if parallel computation is employed), and thus has the potential of achieving greater progress at each iteration. Recall that we are particularly interested in applying the algorithm to problems in which each function call requires that a simulation, or other time-consuming computation, is performed. Therefore, updating all the variables simultaneously, and making the required function calls in parallel, is a beneficial feature of the sampled fictitious play algorithm.

Finally, the quality and usability of the solution returned by the algorithm need to be addressed. The limit points of the sequence generated by the algorithm as it is stated in this paper are only guaranteed to be mixed-strategy equilibria of the identical payoff game that we associate with the optimization problem (20), whose feasible solutions are precisely pure strategies. While the pure-strategy response structure of fictitious play favors pure-strategy limiting behavior, simple (but, perhaps, artificial) examples of games for which fictitious play converges to mixed-strategy equilibria can be readily constructed. In some applications, a mixed-strategy solution has a fairly natural interpretation as a feasible solution of the underlying optimization problem (see the following subsection for one such example); for such applications, sampled fictitious play would be a natural solution approach.

In the majority of applications, however, a pure-strategy solution is required. Additionally, in theory, the equilibrium found by the algorithm is not guaranteed to have an objective value close to the optimum. There exist modifications to the fictitious play algorithm that guarantee convergence to a pure-strategy equilibrium corresponding to an optimal solution of (20) in the limit (see Lambert 2002). As a simpler alternative, to address both of these issues when necessary, one can use a technique common to many global optimization heuristics. Namely, the algorithm can be modified to keep track of, and output at termination, the

best pure-strategy solution encountered in the course of the algorithm. The fictitious play framework of the algorithm then serves as a search-guiding tool, allowing for efficient exploration of promising solutions. The following proposition implies that the value of the mixed extension of $u(\cdot)$ at the equilibrium point found by the algorithm is in fact a lower bound on the value of the best-encountered pure strategy.

PROPOSITION 9. *Let Γ be a finite n -player game in strategic form with identical payoffs $u(\cdot)$. Given a mixed-strategy equilibrium $f = (f^1, \dots, f^n) \in \Delta$, there exists $y = (y^1, \dots, y^n) \in \mathcal{Y}$ such that $f^i(y^i) > 0$, $i = 1, \dots, n$ and $u(y) \geq u(f)$.*

The proof relies on the interpretation of the payoff of a mixed strategy as a convex combination of payoffs of pure strategies, and is left to the reader.

In the following subsection, we briefly illustrate the use of sampled fictitious play through two computational experiments. These examples illustrate some of the properties that make this algorithm suitable for specific problems, as well as provide an indication that it is competitive with existing algorithms in terms of solution quality and computational efficiency.

5.2. Computational Experience with Sampled Fictitious Play

5.2.1. Dynamic Traffic Routing Assignment. In Garcia et al. (2000), a version of our sampled fictitious play algorithm was proposed as a way to approach the problem of finding system-optimal vehicle routes in dynamic traffic networks. Their paper predated our work, and served as motivation for considering approximate and sampled versions of the fictitious play algorithm. The problem considered in the paper possesses many of the properties that make fictitious play an effective heuristic, as discussed above, and the empirical results obtained illustrate the quality of solutions that the methods can provide.

We begin by formally introducing the particular version of the dynamic traffic assignment problem considered. Let $N = \{1, 2, \dots, n\}$ be the index set of the vehicles that wish to travel through the given road network. Every vehicle $i \in N$ is endowed with origin and destination locations in the network and an origin departure time, and has a finite set of routes \mathcal{Y}^i (i.e., sequences of road segments, or links, joining origins with destinations) which it can take; let $\mathcal{Y} = \prod_{i \in N} \mathcal{Y}^i$. Cardinality of the set \mathcal{Y}^i for each $i \in N$ depends on the topology of the road network in question, but typically grows exponentially with the number of intersections. The time it takes each vehicle to traverse a route is influenced by the traffic congestion on each road link included in the route during the time the vehicle is travelling along the link, and thus depends on the choices of routes made by the other vehicles in the network. We denote by $a_i: \mathcal{Y} \rightarrow \mathbb{R}$, $i \in N$ the *assignment* mappings, where $a_i(y)$ is the total travel time for vehicle i if vehicles

follow the routes specified by $y \in \mathcal{Y}$. The objective of the dynamic traffic assignment problem is to find a system-optimal assignment of routes, which is attained by minimizing the average trip time experienced in the network, i.e., the function $u(y) = \sum_{i \in N} (a_i(y)/n)$, over all choices of $y = (y^1, \dots, y^n) \in \mathcal{Y}^1 \times \dots \times \mathcal{Y}^n$; the problem is therefore of the form (20).

The system-optimal traffic assignment problem with flow-dependent costs has been studied extensively; see, for example, Potts and Oliver (1972) for an early survey. However, a crucial distinguishing characteristic of this version of the problem is the dynamic, time-dependent nature of the congestion on the links in the network. The problem of finding a system-optimal assignment of routes in a dynamic network is the subject of a great deal of research in the field of intelligent transportation systems; in particular, several versions of a pertinent subproblem of finding a shortest path in a dynamic network has been successfully addressed (see, for example, Kaufman and Smith 1993 and Chabini 1998). Specifically, for given routing decisions of all other vehicles in the network, the dynamic (i.e., time-dependent) shortest path for one vehicle can be efficiently computed if the network conditions are such that a single vehicle will have little effect on congestion, and if some mild assumptions of the assignment mappings are satisfied.

Following the discussion in §5.1, fictitious play is an attractive method of approaching the dynamic traffic assignment problem, with the individual vehicles acting as players. First, the problem fits neatly in the format of (20). Second, given the routing choices of the other vehicles in the network and the corresponding values of the assignment mappings a_i , $i \in N$, the best routing choice of a particular vehicle $i \in N$ can be efficiently computed using a dynamic shortest path algorithm. In particular, each vehicle is assigned to its feasible route minimizing the increase of total system travel time, corresponding to the best-reply computations with respect to the mixed extension of utility function $u(\cdot)$ above (these computations can be done in parallel, although this was not implemented in the actual test performed). Third, the numerical procedures for evaluating the assignment mappings for a given choice of routes, while available for both commercial and academic users, usually require significant computational effort. Finally, mixed strategies found by the algorithm can be easily and naturally addressed in problem instances in which many vehicles share the same origins, destinations, and departure times. Indeed, the frequencies of the mixed strategy can be viewed as representing the percentages of the vehicles taking particular routes (i.e., using particular pure strategies) among the platoon of those sharing origins, destinations, and departure times.

The latter consideration also motivated the use of sampling in the implementation of the algorithm in the aforementioned paper (although no rigorous claims on the limiting behavior of the algorithm were made). The authors chose to take one sample at each iteration. Although the

theoretical results proven in this paper are not applicable in this setting, it can again be argued that the large number of vehicles sharing the same origins, destinations, and departure times results in close replication of the historic frequencies of routes taken by the vehicles in each such group because the number of samples taken for each origin-destination pair is actually rather large.

Although no a priori guarantee on the quality of the solution found by the (sampled) fictitious play algorithm can be made at this point, the computational experiments in Garcia et al. (2000) indicate that the solution found was quite good, as measured by the criterion of overall system performance. In the experiment, performed for the traffic network of Troy, MI, approximately 16,500 vehicles were simulated to “flow” through the network according to approximations of actual travel patterns observed in Troy. The road network considered included 200 intersections and 50 zone centroids that can serve as origins or destinations, and 529 links. The majority of the vehicles were presumed to follow the “free-flow” fastest paths, i.e., paths from their origins to their destinations that are the fastest when no traffic congestion is present on the network. A fraction of the vehicles, however, were “guided” to follow the routes found for them by the above implementation of the fictitious play algorithm. The authors investigated the impact such (limited) route guidance had on the individual travel times of the vehicles as well as the system average travel times. The results and performance of the sampled fictitious play algorithm were compared with those of the heretofore best method *SAVaNT* (Kaufman et al. 1998). When the guided vehicles constituted 5% of the total number of vehicles, their average travel times were approximately 10% better than those of the unguided vehicles, with the systemwide average travel time of 17.22 minutes. It took 20 iterations of sampled fictitious play to obtain the above result, compared to the 68 iterations of *SAVaNT*, requiring more c.p.u. time than fictitious play, resulting in an average trip time of 17.39 minutes. Moreover, guiding 25% of the vehicles not only decreased the average travel time for a guided vehicle to 8.72 minutes, but also reduced the system-average travel time to 8.82 minutes. (That is, the ability of a quarter of the vehicles in the network to choose their routes intelligently sufficiently reduced the overall network congestion to allow for faster travel times of the unguided vehicles.) It is remarkable that the fictitious play algorithm obtained results nearly identical to the output of *SAVaNT* using considerably fewer iterations and less c.p.u. time, even without exploring the possible parallel implementation of best-response calculations.

5.2.2. Mobile Units Situation Awareness. In their paper “Fictitious Play Approach to a Mobile Unit Situation Awareness Problem,” Lambert and Wang (2003) demonstrate the effectiveness of the fictitious play approach to optimization on a large-scale situation awareness simulation developed for the Multidisciplinary University

Research Initiative (MURI) on “Low-Energy Electronic Design for Mobile Platforms.” In this problem, a number of mobile units desire to keep track of each other’s locations over the course of a military exercise lasting a prespecified period of time. The MURI project as a whole was concerned with every aspect of the operation—from the design of the components of devices that units use to broadcast their location information, to deciding on the communication protocol. It was the latter aspect of the project that presented an optimization problem that was addressed by fictitious play.

The units’ communication devices operate with batteries, and thus the units have a limited amount of energy. The transmission of information by a unit requires energy, as does the processing of any received signal. The probability that a unit will receive a transmission broadcast by another unit depends on the distance between them, the power used to transmit, and the power used to receive the signal. The process of designing the communication protocol consisted of selecting values for the following set of parameters: T —the time between transmissions of position information, q —the probability of retransmitting (to compensate for possible “lost” transmissions), E_{ct} —the energy used in transmission, and E_{cr} —the energy used in receiving. It was assumed that all units will use the same values for the above parameters. The objective was to minimize, over all choices of T , q , E_{ct} , and E_{cr} , the sum of the mean absolute errors of each unit’s estimates of the positions of the other units over the course of the exercise. The nature of the dependence of the objective function on these parameters is complex, with no analytic expression of the function available. Thus, to evaluate the objective function for any choice of values of the parameters, a complex simulation of the process down to the electronic components of the communication devices was required, taking two to five minutes of computer time. This feature of the problem motivated the use of sampled fictitious play as an optimization heuristic in this problem. Its performance was compared to both pure random search and simulated annealing by Lambert and Wang (2003).

In the implementation of sampled fictitious play, the four parameters were selected as the players, with a discretization of 10 points over each parameter’s range of values as the strategy space for the respective player. The resulting discretized problem fits into the framework of (20), because every combination of parameter values was considered feasible (if a unit used all its allotted power before the end of the exercise, it simply stopped transmitting and receiving information, and thus updating its estimates of the other units’ locations). At each iteration, a sample of the strategies of the four players was taken. Due to the complex nature of dependence of the objective function on each of the parameters, the only available method to compute, for each player, a best reply to the above sample was to evaluate (via simulation) the objective function for each of the 10 pure strategies, with the other players choosing

their sampled strategies. However, the implementation took advantage of the opportunity to perform the best-reply computations in parallel.

Unlike in the dynamic traffic assignment problem, this problem did not possess a natural interpretation for a mixed strategy. Therefore, the sampled fictitious play algorithm was used, as proposed in §5.1, as a search guidance heuristic, where the output returned was the best pure-strategy combination encountered in the course of the algorithm.

Several problem instances, with various initial battery capacity levels, were considered. The results found by Stark et al. (2002) using simulated annealing, as well as results using pure random search, were used as a benchmark against which to compare the solutions found by sampled fictitious play. The sampled fictitious play algorithm was allowed to run for 30 iterations, so that the number of unique function calls made by the algorithm approximately matched those of the other two algorithms; however, the wall-clock running time of the sampled fictitious play algorithm was lower by as much as a factor of six, due to the parallelization. They found that sampled fictitious play and simulated annealing returned solutions of comparable quality, while both outperformed pure random search by as much as 25%. Moreover, sampled fictitious play, unlike simulated annealing, could gracefully scale to allow for unit-dependent communication protocols without significant increase in computation time.

6. Conclusions

In this paper, we broadened the existing results on convergence properties of the fictitious play algorithm on games with identical payoffs to include an approximate fictitious play algorithm that allows for errors in players’ best replies. Moreover, we introduced the sampled fictitious play algorithm, which possesses the above convergence properties, and at the same time provides a computationally efficient method for implementing fictitious play.

In the second part of the paper, we introduced the use of algorithms based on sampled fictitious play to solve optimization problems in the form (1) and argued when this approach is likely to be beneficial, with particular focus on the problems in which the objective function $u(\cdot)$ comes from a “black box,” such as a simulation model, where significant computational effort is required for each function evaluation, and no assumptions on the objective function structure are made.

Computational experiments evaluating the performance of algorithms based on the fictitious play paradigm are ongoing. We cited two examples, a dynamic traffic routing problem and a situation awareness application, to demonstrate the potential of these algorithms in complex applied problems. Other applications are being considered, as well as extensions of the algorithms to optimization problems of more general form.

Appendix

PROOF OF PROPOSITION 1. First note the following fact: Given $\delta > 0$, there exists an $\varepsilon > 0$ such that for each $\bar{f} \in \Delta$ with $\bar{f} \notin B_\delta(K)$, $u^i(f^i, \bar{f}^{-i}) - u^i(\bar{f}) > \varepsilon$ for some $i \in N$ and some $f \in \Delta$. Simply take

$$\varepsilon = \frac{1}{2} \min_{\bar{f} \in \Delta \setminus B_\delta(K)} \max_{i \in N} \max_{f^i \in \Delta^i} (u^i(f^i, \bar{f}^{-i}) - u^i(\bar{f})), \tag{21}$$

which is well defined and positive because $\Delta \setminus B_\delta(K)$ is a compact set, N is finite, and Δ^i is compact and u^i is continuous for all $i \in N$.

Now suppose that the belief path $(f(t))_{t=1}^\infty$ does not converge to equilibrium. Then, for some $\delta > 0$, for any T there exists $t \geq T$ such that $f(t) \notin B_\delta(K)$. Then, for $\varepsilon > 0$ defined as in (21), for any T there is a $t \geq T$ for which $f(t)$ is not an ε -equilibrium, which contradicts our hypothesis. Hence, the belief path $(f(t))_{t=1}^\infty$ converges to equilibrium. \square

PROOF OF PROPOSITION 2. Assume that (4) holds, but there exists $\delta > 0$ such that

$$\limsup_{T \rightarrow \infty} \frac{\text{card}(\{1 \leq t \leq T: f(t) \notin B_\delta(K)\})}{T} = c > 0.$$

Define ε as in (21). Then, for each $\bar{f} \in \Delta$ with $\bar{f} \notin B_\delta(K)$, $u^i(f^i, \bar{f}^{-i}) - u^i(\bar{f}) > \varepsilon$ for some $i \in N$ and some $f \in \Delta$. That is,

$$\bar{f} \notin K_\varepsilon \quad \text{if } \bar{f} \notin B_\delta(K).$$

Therefore,

$$\begin{aligned} &\text{card}(\{1 \leq t \leq T: f(t) \notin K_\varepsilon\}) \\ &\geq \text{card}(\{1 \leq t \leq T: f(t) \notin B_\delta(K)\}), \end{aligned}$$

and so

$$\begin{aligned} &\limsup_{T \rightarrow \infty} \frac{\text{card}(\{1 \leq t \leq T: f(t) \notin K_\varepsilon\})}{T} \\ &\geq \limsup_{T \rightarrow \infty} \frac{\text{card}(\{1 \leq t \leq T: f(t) \notin B_\delta(K)\})}{T} = c > 0, \end{aligned}$$

which contradicts (4). \square

PROOF OF LEMMA 3. Let $(y(t))_{t=1}^\infty$ be a path, and let $(f_y(t))_{t=1}^\infty$ be the associated belief path. We have already noted that for any belief process, convergence to equilibrium implies convergence to equilibrium in Cesaro mean.

To establish the converse, suppose that $(y(t))_{t=1}^\infty$ converges in belief to equilibrium in Cesaro mean. Let $\delta > 0$ be chosen arbitrarily. Let

$$M = \max_{f, g \in \Delta} \|f - g\|.$$

Choose $\eta \in (0, \delta/(2\delta + M))$. By (3), there exists an integer T_0 such that for every $T \geq T_0$,

$$\text{card}(\{1 \leq t \leq T: f_y(t) \notin B_\delta(K)\}) < \eta T. \tag{22}$$

We will show that for every $T \geq T_0$, $f_y(T) \in B_{2\delta}(K)$.

Suppose that $T \geq T_0$, but $f_y(T) \notin B_{2\delta}(K)$. Then, for any integer $t \in [T, T + \lfloor T\delta/(\delta + M) \rfloor]$, $f_y(t) \notin B_\delta(K)$ (here, $\lfloor x \rfloor$ is the greatest integer less than or equal to x). Indeed, if for some t in the above interval $f_y(t) \in B_\delta(K)$, then there exists $f \in K$ such that $\|f_y(t) - f\| < \delta$, which implies

$$\begin{aligned} 2\delta &\leq \|f_y(T) - f\| \leq \|f_y(t) - f_y(T)\| + \|f_y(t) - f\| \\ &< \|f_y(t) - f_y(T)\| + \delta; \end{aligned}$$

i.e., $\|f_y(t) - f_y(T)\| > \delta$. On the other hand, using expression (6),

$$\begin{aligned} \|f_y(t) - f_y(T)\| &= \frac{1}{t} \left\| \sum_{s=T+1}^t (y(s) - f_y(T)) \right\| \\ &\leq \frac{1}{t} \sum_{s=T+1}^t \|y(s) - f_y(T)\| \\ &\leq \frac{t-T}{t} \cdot M \leq \frac{\delta}{M+\delta} \cdot M \leq \delta, \end{aligned}$$

resulting in a contradiction. Hence, for any integer $t \in [T, T + \lfloor T\delta/(\delta + M) \rfloor]$, $f_y(t) \notin B_\delta(K)$.

We conclude that

$$\begin{aligned} &\text{card}\left(\left\{1 \leq t \leq T + \left\lfloor \frac{T\delta}{\delta + M} \right\rfloor : f(t) \notin B_\delta(K)\right\}\right) \\ &\geq 1 + \left\lfloor \frac{T\delta}{\delta + M} \right\rfloor > \eta \left(T + \left\lfloor \frac{T\delta}{\delta + M} \right\rfloor\right), \end{aligned}$$

contradicting (22). Therefore, $f_y(T) \in B_{2\delta}(K)$ for every $T \geq T_0$. Because δ was chosen arbitrarily, convergence follows. \square

Acknowledgments

This research was supported in part by the National Science Foundation under grants DMI-9713723, DMI-9900267, and DMI-0217283; MURI grant ARO DAAH04-96-1-0377; ASSERT grant ARO DAAG55-98-1-0155; GM Collaborative Research Laboratory in Advanced Vehicle Manufacturing at UM Grant; and by the Horace H. Rackham School of Graduate Studies Faculty Grant.

References

Bazaraa, M. S., H. D. Sherali, C. M. Shetty. 1993. *Nonlinear Programming: Theory and Algorithms*. John Wiley and Sons, New York.

Brown, G. W. 1951. Iterative solution of games by fictitious play. *Activity Analysis of Production and Allocation*. Wiley, New York.

Chabini, I. 1998. Discrete dynamic shortest-path problems in transportation applications: Complexity and algorithms with optimal run time. *Transportation Res. Record* **1645** 170–175.

Garcia, A., D. Reaume, R. L. Smith. 2000. Fictitious play for finding system optimal routings in dynamic traffic networks. *Transportation Res. B* **34** 146–157.

- Goldberg, D. E. 1989. *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison-Wesley, Reading, MA.
- Harsanyi, J. C. 1973. Oddness of the number of equilibrium points: A new proof. *Internat. J. Game Theory* **2** 235–250.
- Holland, J. H. 1975. *Adaptation in Natural and Artificial Systems*. The University of Michigan Press, Ann Arbor, MI.
- Jong, K. A. D. 1975. An analysis of the behavior of a class of genetic adaptive systems. Ph.D. thesis, University of Michigan, Ann Arbor, MI.
- Kaufman, D. E., R. L. Smith. 1993. Fastest paths in time-dependent networks for intelligent vehicle-highway systems application. *IVHS J.* **1** 1–11.
- Kaufman, D. E., R. L. Smith, K. Wunderlich. 1998. Dynamic user-equilibrium properties of fixed-points in iterative routing-assignment methods. *Transportation Res. C* **6** 1–16.
- Kirkpatrick, S., C. D. Gelatt Jr., M. P. Vecchi. 1983. Optimization by simulated annealing. *Science* **20** 671–680.
- Lambert, T. J. 2002. Deterministic and stochastic systems optimization. Ph.D. thesis, University of Michigan, Ann Arbor, MI.
- Lambert, T. J., H. Wang. 2003. Fictitious play approach to a mobile unit situation awareness problem. Technical report, University of Michigan, Ann Arbor, MI.
- Monderer, D., L. Shapley. 1996. Fictitious play property for games with identical interests. *J. Econom. Theory* **68** 258–265.
- Munkres, J. R. 1975. *Topology: A First Course*. Prentice-Hall, Englewood Cliffs, NJ.
- Potts, R. B., R. M. Oliver. 1972. *Flows in Transportation Networks*. Academic Press, New York.
- Robinson, J. 1951. An iterative method of solving a game. *Ann. Math.* **54** 296–301.
- Rosenmüller, J. 1971. On a generalization of the Lemke-Howson algorithm to noncooperative N -person games. *SIAM J. Appl. Math.* **21** 73–79.
- Ross, S. M. 1996. *Stochastic Processes*. John Wiley and Sons, New York.
- Stark, W. E., H. Wang, A. Worthen, S. Lafortune, D. Teneketzis. 2002. Low-energy wireless communication network design. *IEEE Wireless Comm.* **9** 60–72.
- Wilson, R. 1971. Computing equilibria of N -person game. *SIAM J. Appl. Math.* **21** 80–87.