

# DISTRIBUTED LEARNING IN LARGE-SCALE MULTI-AGENT GAMES: A MODIFIED FICTITIOUS PLAY APPROACH

BRIAN SWENSON<sup>†</sup>, SOUMMYA KAR<sup>†</sup> AND JOÃO XAVIER<sup>\*</sup>

**Abstract**—The paper concerns the development of distributed equilibria learning strategies in large-scale multi-agent games with repeated plays. With inter-agent information exchange being restricted to a preassigned communication graph, the paper presents a modified version of the fictitious play algorithm that relies only on local neighborhood information exchange for agent policy update. Under the assumption of identical agent utility functions that are permutation invariant, the proposed distributed algorithm leads to convergence of the networked-averaged empirical play histories to a subset of the Nash equilibria, designated as the consensus equilibria. Applications of the proposed distributed framework to strategy design problems encountered in large-scale traffic networks are discussed.

## I. INTRODUCTION

We present a distributed algorithm for learning Nash equilibria in large-scale repeated games. The algorithm, which we call modified fictitious play (MFP), is based on the well known fictitious play (FP) algorithm. MFP is an adapted form of FP which seeks to reduce the communication and computational burdens which make FP egregiously impractical for large-scale applications. We show that in a certain class of games, MFP converges in average empirical frequency to a subset of the Nash equilibria which we call the consensus equilibria.

FP was originally introduced as a way of computing Nash equilibria in repeated two player games [2]. Since it was introduced, FP has been the subject of much research to determine the class of games for which it does in fact converge<sup>1</sup>. While it has been shown that FP does not converge for all games, there are many useful games for which convergence has been proven. In particular, in [4] it was shown that FP can be used to compute equilibria in identical interest games with an arbitrarily large number of players. However, the algorithm is limited in practicality due to the intense computational requirements of FP in large games.

Concomitant to research in large-scale games is the study of distributed convergence to Nash equilibrium. In a non-distributed setting it is generally assumed that players have unrestricted ability to observe the actions of all other players. In a large-scale game it may be impractical or impossible for players to have such universal knowledge. A distributed

setting generally implies a restriction on the ability of players to observe the actions of others. It may be restricted to a neighborhood of the player, or in the extreme case, a player may only be permitted to observe their own actions [1],[9]. Some researchers have looked at adapting the FP algorithm to a distributed setting. In Joint Strategy Fictitious Play [6], the information tracking burden is reduced by requiring that players only know the predicted utilities for each action. In [9], a method is presented where players need only observe their personal utility received at each stage. In this paper, we take a slightly different approach to the problem. We investigate the question: what if all players have access to only the *average* empirical distribution of all players?

Our main contribution is a distributed learning algorithm which converges in terms of average empirical frequency to a subset of the mixed strategy Nash equilibria which we call the consensus equilibria. Moreover, this convergence to the set of consensus equilibria guarantees that each agent obtains an accurate estimate of the limiting equilibrium strategy. Previous work has not explicitly considered the scenario where information gathering occurs through a preassigned communication graph. We consider a scenario where players engaged in a repeated game are able to exchange information once per iteration of the game using a preassigned communication graph  $G$  in order to estimate the average empirical distribution. The graph  $G$  must be connected but may still be sparse. Convergence results are proven for games with identical utility functions and can be extended to the larger class of games known as potential games [7] with the restriction that the potential function be permutation invariant. We emphasize that the mode of convergence used in this paper, convergence in *average empirical frequency*, is different from the more commonplace convergence in empirical frequency.<sup>2</sup>

The concept of a consensus equilibrium is closely related to that of a symmetric equilibrium. The existence of symmetric equilibrium in finite normal form games was first proven by Nash [8] in the same work where the concept of Nash equilibrium was originally presented. In general, a symmetric equilibrium is a Nash equilibrium which is invariant under automorphisms of the game. A consensus equilibrium, on the other hand, is a Nash equilibrium in which all players use an identical strategy. In the case of a symmetric game, the two concepts coincide. Symmetric games can be useful, especially when considering large-scale scenarios. As noted in [3], the inherent symmetry can allow for ‘algorithmic shortcuts leading to significantly more effective or efficient solution procedures.’ MFP may be considered as a tool which takes advantage of the

This work was supported by the FCT project [PEst-OE/EEI/LA0009/2011] through the Carnegie-Mellon/Portugal Program managed by ICTI from FCT and partially supported by FCT Grant CMU-PT/SIA/0026/2009.

<sup>†</sup>Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA 15213, USA (brianswe@andrew.cmu.edu and soumyak@andrew.cmu.edu)

<sup>\*</sup>Institute for Systems and Robotics(ISR), Instituto Superior Tecnico (IST), Technical University of Lisbon, Portugal(jxavier@isr.ist.utl.pt)

<sup>1</sup>In reference to FP or one of its variants, we use the term convergence to mean that the empirical frequency distribution of a FP process asymptotically converges to the set of Nash equilibria, a notion to be made precise in section II.

<sup>2</sup>We prove our results for the general case of arbitrary tie breaking rules. By making stronger assumptions on the tie breaking rules it may be possible to guarantee other modes of convergence.

symmetry in certain classes of games to provide an efficient method for finding consensus (symmetric) equilibria. As such, MFP is specifically tailored to finding consensus equilibria, and it remains to be seen to what extent the scope of MFP can be enlarged for learning symmetric equilibria in more general classes of games where the two concepts do not precisely coincide.

The rest of the paper is organized as follows. Section II sets up notation to be used in the subsequent development. The set of consensus equilibria is defined and the classical (centralized) FP algorithm is reviewed in the same section. Section III introduces MFP as a low-information-overhead repeated-play alternative to FP for learning consensus equilibria in multi-agent games. A fully distributed implementation of the proposed MFP, the distributed MFP, in multi-agent scenarios in which agent information dynamics is restricted to communication over a preassigned sparse communication network, is presented and analyzed in section IV. In section VI we demonstrate an application of distributed MFP in a traffic routing scenario, whereas some generalizations are discussed in section V. Finally, section VII concludes the paper.

## II. SETUP AND NOTATION

### A. Notation

Let  $\Gamma$  be an  $n$ -player game in normal form with the set of players  $N = \{1, 2, \dots, n\}$ . The set of pure strategies for player  $i$  is given by  $Y_i$  and the set of joint pure strategies for all players is  $Y = Y_1 \times Y_2 \times \dots \times Y_n$ . The utility of each  $y \in Y$  for player  $i$  is given by  $u_i(y) : Y \rightarrow \mathbb{R}$ . This paper will deal mostly with games with identical utility functions such that  $u_i(y) = u_j(y) \forall i, j$ , in such cases we drop the subscript and write  $u(y) = u_i(y) \forall i$ .

The set of mixed strategies<sup>3</sup> for player  $i$  is given by  $\Delta_i$  and the set of joint mixed strategies for all players is given by  $\Delta = \Delta_1 \times \Delta_2 \times \dots \times \Delta_n$ . The mixed utility for  $f \in \Delta$  is given by  $U_i(f) = \sum_{y \in Y} u_i(y) f_1(y_1) f_2(y_2) \dots f_n(y_n)$ . For convenience the notation  $U_i(f)$  will often be written as  $U_i(f_i, f_{-i})$  where  $f_i \in \Delta_i$  is the mixed strategy for player  $i$  and  $f_{-i}$  indicates the joint mixed strategy for all players other than  $i$ . This is merely a notational convenience and it should be remembered that  $U(f) = U(f_i, f_{-i})$ .

Let  $\{\Upsilon_i(t, y_i)\}_{t=0}^{\infty}$  be a sequence of actions for player  $i$ . Viewed as a mixed strategy,  $\Upsilon_i(t)$  for each time  $t$ , may be represented in terms of a Dirac or delta distribution over the set of possible actions, i.e.,

$$\Upsilon_i(t, y_i) = \begin{cases} 1, & \text{if } y_i \text{ is the action taken at time } t \\ 0, & \text{otherwise.} \end{cases} \quad (1)$$

In order to simplify notation we usually write this as  $\Upsilon_i(t)$ , leaving out the second argument when it is clear from the context. Let  $\{\Upsilon(t)\}_{t=0}^{\infty}$  be the associated sequence of actions in the joint mixed strategy space. Let  $f_i(t)$  be the normalized histogram (empirical distribution) of the actions of player  $i$  up

to time  $t$ , i.e.,

$$f_i(t) = \frac{1}{t} \sum_{k=1}^t \Upsilon_i(k).$$

Similarly,

$$f(t) = \frac{1}{t} \sum_{k=1}^t \Upsilon(k)$$

is the joint empirical distribution corresponding to the joint actions of the players up to time  $t$ . The sequence of distributions  $\{f(t)\}_{t=0}^{\infty}$  is often called a belief sequence.

A mixed strategy  $f$  is a Nash equilibrium of  $\Gamma$  if  $U(f) \geq U(g_i, f_{-i}) \forall g_i \in \Delta_i$ . We define the set of Nash equilibria as

$$K = \{f : U_i(f) \geq U_i(g_i, f_{-i}) \forall g_i \in \Delta_i\}, \quad (2)$$

and the set of consensus equilibria as

$$C = \{f : U_i(f) \geq U_i(g_i, f_{-i}) \forall g_i \in \Delta_i, \\ f_1 = f_2 = \dots = f_n\}. \quad (3)$$

Clearly, the set  $C$  of consensus equilibria defined above is a subset of the set  $K$  of Nash equilibria. The set of  $\varepsilon$  Nash equilibria is given by

$$K_\varepsilon = \{f : U_i(f_i, f_{-i}) + \varepsilon \geq U_{-i}(g_i, f_i) \forall g_i \in \Delta_i\}, \quad (4)$$

The distance of a distribution  $f \in \Delta$  from  $C$  is given by  $d(f, C) = \inf\{\|f - g^*\| : g^* \in C\}$ .

Unless stated otherwise, we will restrict attention to games with identical permutation-invariant utilities; formally, we assume:

**A. 1.** All players use the same strategy space.

**A. 2.** The players' utility functions are identical and permutation invariant

Let

$$\bar{f}(t) = \frac{1}{n} \sum_{i=1}^n f_i(t)$$

be the average empirical distribution. With a slight abuse of notation, let  $\bar{f}_n(t) = (\bar{f}(t), \bar{f}(t), \dots, \bar{f}(t)) \in \Delta$  denote the mixed strategy where all players use the empirical average as their individual strategy. Both  $\bar{f}(t)$  and  $\bar{f}_n(t)$  will be extremely important in the exposition of MFP.

Note that under these assumptions, the set of consensus equilibria is known to be nonempty [3].

### B. Fictitious Play

In fictitious play the best response of player  $i$  at time  $t$  is given by

$$v_i(f(t)) = \max_{g_i(y_i) \in \Delta_i} U(g_i, f_{-i}(t)) \quad (5)$$

$$= \max_{g_i(y_i) \in \Delta_i} \sum_{y \in Y} u(y) g_i(y_i) f_{-i}(t, y_{-i}) \quad (6)$$

In words this means that at time  $t$  each player chooses a best response by assuming that the empirical distributions of the other players accurately represent their respective mixed

<sup>3</sup>A mixed strategy  $f_i$  is a probability distribution over the set  $Y_i$ . Player  $i$  chooses an action by sampling  $f_i$ .

strategies. A fictitious play process is a sequence  $\{\Upsilon(t)\}_{t=0}^{\infty}$  such that,

$$v_i(f(t)) = U(\Upsilon_i(t+1), f_{-i}(t))$$

A fictitious play process is said to converge in empirical frequencies if  $\lim_{t \rightarrow \infty} d(f(t), K) = 0$ . In [4] it was shown that a fictitious play process converges in the above sense for games satisfying A. 1 - A. 2.

### III. MODIFIED FICTITIOUS PLAY

In theory, FP is a powerful tool for learning equilibria in  $n$ -player games. However, when it comes to practical implementation, it quickly becomes infeasible as the game size grows. The difficulty of implementing FP in large-scale distributed games can be understood by analyzing the FP best response calculation given in (6). The two major problems with FP are:

- i) Each player must have access to the marginal empirical frequency distributions of the other  $n - 1$  players in order to compute a best response. It is impractical for each agent to track the actions of all other agents.
- ii) The computational complexity of computing the mixed utility given an  $n$  dimensional pdf grows exponentially with the size of the game.

Consider a modification of the FP best response given in (6), such that each player calculates a best response by assuming that all other players are independently playing the average empirical distribution.

$$\tilde{v}_i(\bar{f}(t)) = \max_{g_i(y_i) \in \Delta_i} \sum_{y \in Y} u(y) g_i(y_i) \bar{f}_{-i}(t, y_{-i}) \quad (7)$$

where

$$\bar{f}_{-i}(t, y_{-i}) = \bar{f}(t, y_1) \dots \bar{f}(t, y_{i-1}) \bar{f}(t, y_{i+1}) \dots \bar{f}(t, y_n),$$

is the joint distribution where the other  $-i$  players independently use  $\bar{f}(t)$ . In words, this means that player  $i$  chooses a best response by assuming that all other players are independently using the mixed strategy given by  $\bar{f}(t)$ . Using such a best response rule, there would be no need for each player to track the empirical distribution of all other players; rather, each player need only have access to the average empirical distribution,  $\bar{f}(t)$ . This is the fundamental idea of MFP. The information tracking problem of FP is mitigated by requiring that players only track  $\bar{f}(t)$ . The second problem of computational complexity is resolved in a less direct manner. By exploiting the symmetry inherent in the MFP best response calculation, the computation can often be greatly simplified. For example, in the distributed traffic routing scenario of section VI, the complexity of the best response calculation is reduced to constant time complexity, in terms of the number of players.

In an ideal MFP process, the best response calculation is given by (7). We consider a more general case where players do not have access  $\bar{f}(t)$  directly. Instead, player  $i$  has access to  $\hat{f}_i(t)$ , an approximation of  $\bar{f}(t)$ . Let  $\varepsilon_{max}(t) \geq \max_{i \in N} \|\bar{f}(t) - \hat{f}_i(t)\|_{\infty}$  be the maximum error in any players approximation

of  $\bar{f}(t)$ . We make the following assumption about the decay rate of the error,

$$\mathbf{A. 3.} \quad \varepsilon_{max}(t) = O\left(\frac{\log t}{t^r}\right), r > 0$$

A sequence of actions  $\{\Upsilon(t)\}_{t=0}^{\infty}$  is a *modified fictitious play process* if

$$\begin{aligned} v_i^m(\bar{f}(t)) &= \max_{g_i(y_i) \in \Delta_i} \sum_y u(y) g_i(y_i) \hat{f}_i(t, y_{-i}) \\ &= \sum_y u(y) \Upsilon_i(t+1, y_i) \hat{f}_i(t, y_{-i}). \end{aligned}$$

That is, each player best responds to  $\hat{f}_i(t)$  his personal estimate of  $\bar{f}(t)$ . We use the superscript in  $v_i^m()$  to indicate the MFP best response, as opposed to the traditional FP best response of (6).

**Theorem 1.** *Let  $\{\Upsilon_i(t)\}_{t=1}^{\infty}$  be a MFP process such that A. 1 - 3 hold. Then  $d(\bar{f}_n(t), C) \rightarrow 0$  as  $t \rightarrow \infty$ .*

The proof of theorem 1 follows along essentially the same lines as the proof in [4] for convergence of classical FP in games with identical interests, though with obvious modifications<sup>4</sup>. We emphasize that this result shows that the  $n$ -tuple of the average empirical distribution converges to  $C$ ,  $d(\bar{f}_n(t), C) \rightarrow 0$ . This is not the same as the more traditional definition of convergence in empirical frequencies,

$$d(f(t), C) \rightarrow 0 \text{ as } t \rightarrow \infty. \quad (8)$$

In fact, it is not difficult to construct a counter example that violates (8) given our assumptions.<sup>5</sup> The practical meaning of the theorem is that it allows player  $i$  to learn a mixed strategy  $\hat{f}_i(t)$  which is part of a Nash equilibrium strategy.

### IV. DISTRIBUTED MFP

#### A. Distributed Problem Formulation

The result given in theorem 1 is powerful in that it guarantees convergence to a Nash equilibrium given only an estimate of average of all players empirical distributions  $\bar{f}(t)$ , rather than all of the individual empirical distributions  $f(t)$ . In a distributed setting, not only is it impractical for players to track  $f(t)$  precisely, but it is impractical to even have precise knowledge of  $\bar{f}(t)$ . We show that if players are able to exchange information through a connected communication graph in order to estimate  $\bar{f}(t)$ , then a modified fictitious play process will converge to the set of consensus equilibria. Consider the following problem setup.

Players are engaged in an  $n$ -player repeated game. Players are able to exchange information with neighboring agents once per iteration of the game via a sparse but connected communication graph,  $G = (V, E)$  where each node represents a player. In the distributed problem formulation we maintain assumptions 1 and 2 given in section II for MFP and we

<sup>4</sup>The proofs in this paper are omitted due to space constraints. The complete proofs can be found in the forthcoming longer version of this paper.

<sup>5</sup>A counterexample can be constructed by taking advantage of the arbitrary tie breaking rules. By constraining the tie breaking rules it may be possible to ensure convergence in empirical frequencies, as mentioned in footnote 2.

add an additional assumption pertaining to the communication network.

**A. 4.** *The ancillary communication network can be modeled by a connected graph.*

In order to compactly describe the distributed MFP algorithm, we recast some of the previous definitions of the paper in terms of vectors and matrices. Let  $\Upsilon_i(t) \in \mathbb{R}^{1 \times m}$  be a delta function representing player  $i$ 's action at time  $t$ . Let  $f_i(t) = \frac{1}{t} \sum_{s=1}^t \Upsilon_i(s)$ ,  $f_i(t) \in \mathbb{R}^{1 \times m}$  be the empirical distribution for player  $i$ . Let

$$f(t) = \begin{pmatrix} f_1(t) \\ f_2(t) \\ \vdots \\ f_n(t) \end{pmatrix} \in \mathbb{R}^{n \times m}. \quad (9)$$

Let  $\bar{f}(t) \in \mathbb{R}^{1 \times m}$  be the average empirical distribution,

$$\bar{f}(t) = \frac{1}{n} \sum_{i=1}^n f_i(t). \quad (10)$$

Let  $\hat{f}_i(t) \in \mathbb{R}^{n \times m}$  be player  $i$ 's estimate of  $\bar{f}(t)$ . Let

$$\hat{f}(t) = \begin{pmatrix} \hat{f}_1(t) \\ \hat{f}_2(t) \\ \vdots \\ \hat{f}_n(t) \end{pmatrix}. \quad (11)$$

### B. Distributed MFP algorithm

#### Initialize

i) At time  $t = 0$ , players have not taken any actions. The empirical distribution for all players is given by  $f_i(0) = 0$ . The average empirical distribution is also given by  $f(0) = 0$ . Players initialize their estimate of the average empirical distribution to be  $\hat{f}_i(t) = 0$ . At time  $t = 1$  players choose an arbitrary initial action  $\Upsilon_i(1)$ .  $f_i(1)$  is updated to reflect the action taken.

$$f_i(t+1) = f_i(t) + \frac{1}{t+1} (\Upsilon_i(t+1) - f_i(t)) \quad (12)$$

#### Iterate

ii) Each player computes a new estimate of the average using the following update rule:

$$\hat{f}_i(t+1) = \sum_{j \in N_i} w_{i,j} \hat{f}_j(t) + f_i(t+1) - f_i(t) \quad (13)$$

where  $N_i$  is the set of neighbors of player  $i$ . This update rule can be represented in more compact notation as,

$$\begin{pmatrix} \hat{f}_1(t+1) \\ \hat{f}_2(t+1) \\ \vdots \\ \hat{f}_n(t+1) \end{pmatrix} = W \begin{pmatrix} \hat{f}_1(t) \\ \hat{f}_2(t) \\ \vdots \\ \hat{f}_n(t) \end{pmatrix} + \begin{pmatrix} f_1(t+1) \\ f_2(t+1) \\ \vdots \\ f_n(t+1) \end{pmatrix} - \begin{pmatrix} f_1(t) \\ f_2(t) \\ \vdots \\ f_n(t) \end{pmatrix} \quad (14)$$

or equivalently,

$$\hat{f}(t+1) = W\hat{f}(t) + f(t+1) - f(t) \quad (15)$$

where  $W \in \mathbb{R}^{n \times n}$  is a weighting matrix defining the weight given to each neighbor.  $W$  must satisfy the following assumption

**A. 5.**  *$W$  is an  $n \times n$  matrix that is doubly stochastic, aperiodic, and irreducible.*

Note that given assumption 4 ( $G$  is a connected graph), it is always possible to find a matrix  $W$  satisfying these conditions.

iii) Player  $i$  computes the set of best responses using  $\hat{f}_i(t)$  as the assumed mixed strategy for each of the  $n - 1$  other players.<sup>6</sup> The next action

$$\Upsilon_i(t+1) \in \{\arg \max_{g_i \in Y_i} U(g_i, \hat{f}_{-i}(t))\} \quad (16)$$

is played according to the best response calculation. In the event of multiple pure strategy best responses, any may be chosen arbitrarily.  $f_i(t+1)$  is updated to reflect the action taken.

$$f_i(t+1) = f_i(t) + \frac{1}{t+1} (\Upsilon_i(t+1) - f_i(t)) \quad (17)$$

#### Terminate

iv) steps ii) - iii) are repeated until a termination condition is reached

### C. Main Result

We refer to any sequence of actions  $\{\Upsilon_i(t)\}_{t=1}^{\infty}$  which can be attained using the distributed MFP algorithm of section IV-B as a *distributed MFP process*.

**Theorem 2.** *Let  $\{\Upsilon_i(t)\}_{t=1}^{\infty}$  be a distributed MFP process such A. 1, 2, 4, and 5 hold. Then  $d(\hat{f}(t), C) \rightarrow 0$  as  $t \rightarrow \infty$ . In particular, the agent estimates  $\hat{f}_i(t)$ 's reach asymptotic consensus, i.e.  $d(\hat{f}_i(t), \hat{f}_j(t)) \rightarrow 0$  as  $t \rightarrow \infty$  for each pair  $(i, j)$  of agents. Moreover, the agents achieve asymptotic strategy learning, in the sense that  $d((\hat{f}_i(t))_n, C) \rightarrow 0$  as  $t \rightarrow \infty$  for all  $i = 1, \dots, n$ .*

**Proof Sketch.** *It can be shown that in a distributed MFP process, the maximum error in any players estimate of  $\bar{f}(t)$  decays as  $O(\frac{\log t}{t})$ . Applying theorem 1,  $d(\hat{f}_n(t), C) \rightarrow 0$ , from which it follows that  $d(\hat{f}(t), C) \rightarrow 0$ .*

Again, we emphasize that this is not the same as the more traditional definition of convergence in empirical frequencies given in (8). In words, theorem 2 says that joint strategy profile  $\hat{f}(t)$  where each player uses his estimate of  $\bar{f}(t)$  will converge to the set of consensus equilibria.

The notion of convergence to mixed equilibrium in empirical frequency has been the subject of criticism [5] since it does not actually imply that the strategies played will ever be optimal. Convergence in empirical frequency means only that the time average of the strategies played will converge to the Nash equilibrium. In many cases this can mean that play simply cycles through strategies as  $t \rightarrow \infty$  but, the actual

<sup>6</sup>According to (7),  $U(\cdot)$  should be maximized over the set  $\Delta_i$ . The set of mixed strategy best responses is the convex hull of all pure strategy best responses, so calculating the set of pure strategy best responses implicitly gives the set of mixed strategy best responses.

strategies played at any given  $t$  are never truly optimized. Our notion of convergence in average empirical frequency is not exempt from this criticism; play is never necessarily optimal. Even so, as play progresses in MFP, players do, in some sense, *learn* an optimal strategy. Each player  $i$  has direct access to  $\hat{f}_i(t)$ , which is their portion of the converging equilibrium strategy. This is an important point to note, since a distributed learning algorithm would be of limited use if it required much global knowledge, or if players did not learn the strategy relevant to their play.

## V. GENERALIZATIONS

The assumption A. 2, of identical permutation invariant utility functions can be relaxed in lieu of the following broader assumption:

**A. 6.** *The game  $\Gamma$  is an exact potential game with a permutation invariant potential function.*

Potential games were studied in [7]. A game which admits an exact potential function is known as an exact potential game. The class of potential games includes congestion games. Congestion have many useful applications in economics and engineering. We present an example of a congestion game in the distributed traffic routing example presented in the following section.

## VI. APPLICATIONS

### A. Distributed Traffic Routing

We simulated 25 cars traversing a traffic network of 6 roads. All vehicles start from the same location and have the same destination. The delay on road  $r$  for a joint strategy  $y$  is given by a cubic cost function. In this particular simulation, the weights  $w_{i,j}$  (see (13)) were computed to optimize the speed of distributed linear averaging [10]. A graph of the communication network used for simulations is shown in figure 2.

The key results in this paper are based around showing  $d(\hat{f}(t), C) \rightarrow 0$ . An important practical implication of this result is that  $\hat{f}(t) \in K_{\varepsilon_t}$  where  $\varepsilon_t \rightarrow 0$  and  $K_{\varepsilon_t}$  is defined in (4). Figure 1 shows a plot of the minimum  $\varepsilon_t$  such that  $f(t) \in \varepsilon_t$ . This plot shows that  $\varepsilon_t$  tends to zero which is suggestive of the fact that  $d(\hat{f}(t), C) \rightarrow 0$  holds as well. For the traffic routing application, this means that players concurrently learn a mixed strategy  $\varepsilon_t$  equilibrium, where  $\varepsilon_t$  can be made arbitrarily small. Once  $\varepsilon_t$  is sufficiently small for the game designers purposes, the algorithm can be terminated. The cost functions used to model road delay were specifically chosen as cubic polynomials in order to model a situation in which there may exist multiple consensus equilibria. Distributed MFP is particularly relevant to such situations since it can be used not only to compute a consensus equilibrium, but also to ensure that players agree on which consensus equilibrium is reached.

This example is an instance of a congestion game. Any congestion game can be shown to satisfy assumption 6, and therefore falls within the purview of MFP.

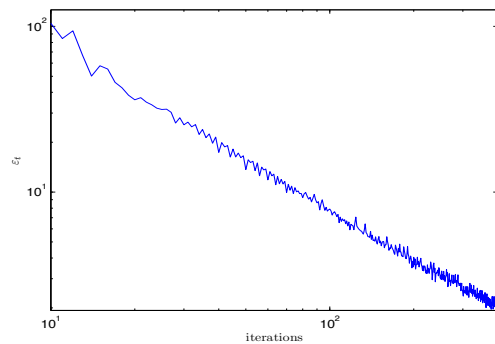


Fig. 1. The minimum  $\varepsilon_t$  on each iteration  $t$  such that  $\hat{f}(t) \in K_{\varepsilon_t}$ . The trend  $\varepsilon_t \rightarrow 0$  is consistent with convergence to the set of consensus equilibria, i.e.  $d(\hat{f}(t), C) \rightarrow 0$  as  $t \rightarrow \infty$ , as stated in theorem 2.

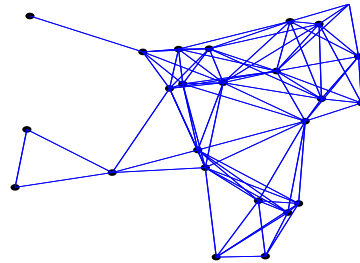


Fig. 2. Randomly generated sparse communication graph.

## VII. CONCLUSIONS

We have introduced a variant of fictitious play which we call modified fictitious play. MFP simplifies the communication and computational burdens inherent in FP, making it suitable to applications in large-scale games. MFP is shown to converge to the set of consensus equilibria for exact potential games with permutation-invariant potential functions. Our main result is a distributed implementation of MFP in which where players communicate via a preassigned, sparse communication graph. Future research may involve extending the scope of MFP to include finding general symmetric equilibria in both symmetric games, and more general games. It is also interesting to consider variants of MFP for learning non-consensus equilibria in broader classes of games.

## REFERENCES

- [1] Gurdal Arslan and Jeff S. Shamma. Distributed convergence to nash equilibria with local utility measurements. In *Conference on Decision and Control*, volume 2, 2004.
- [2] G. W. Brown. *Iterative Solutions of Games by Fictitious Play*. In *Activity Analysis of Production and Allocation*. Wiley, New York, 1951.
- [3] Shih-Fen Cheng, Daniel M. Reeves, Yevgeniy Vorobeychik, and Michael P. Wellman. Notes on Equilibria in Symmetric Games. *Autonomous Agents and Multi-agent Systems*, 2004.
- [4] L. Shapley D. Monderer. Fictitious play property for games with identical interests. *Journal of Economic Theory*, 68(1):258–265, 1996.
- [5] D Fudenberg. Learning Mixed Equilibria. *Games and Economic Behavior*, 5:320–367, 1993.
- [6] J. R. Marden, G. Arslan, and J. S. Shamma. Joint strategy fictitious play with inertia for potential games. *IEEE Transactions on Automatic Control*, 54:208–220, February 2009.
- [7] Dov Monderer and Lloyd S. Shapley. Potential Games. *Games and Economic Behavior*, 14:124–143, 1996.

- [8] John Nash. Equilibrium Points in n-person Games. *Proceedings of The National Academy of Sciences*, 36:48–49, 1950.
- [9] Jeff S. Shamma and Gurdal Arslan. Dynamic fictitious play, dynamic gradient play, and distributed convergence to Nash equilibria. *IEEE Transactions on Automatic Control*, 50:312–327, 2005.
- [10] Liii Xiao and Stephen Boyd. Fast linear iterations for distributed averaging. *Conference on Decision and Control*, 5, 2003.