MEAN-CENTRIC EQUILIBRIUM: AN EQUILIBRIUM CONCEPT FOR LEARNING IN LARGE-SCALE GAMES

BRIAN SWENSON[†*], SOUMMYA KAR[†] AND JOÃO XAVIER[*]

Abstract—The paper is concerned with learning in large-scale multi-agent games. The empirical centroid fictitious play (ECFP) algorithm is a variant of the well-known fictitious play algorithm that is practical and computationally tractable in large-scale games. ECFP has been shown to be an effective tool in learning consensus equilibria (a subset of the Nash equilibria) in certain games. However, the behavior of ECFP has only been characterized in terms of convergence of the *networked-average* empirical frequencies as opposed to the more traditional notion of learning mixed equilibria, namely the notion of convergence of individual empirical frequencies. The behavior of ECFP in terms of convergence in empirical frequencies is herein studied and the equilibrium concept of mean-centric equilibrium (MCE) is introduced. The concept of MCE is similar in spirit to that of Nash equilibrium (NE) but, in MCE each player is at equilibrium with respect to a centroid representing the aggregate behavior, as opposed to NE where players are at equilibrium with respect to the strategies of individual opponents. The MCE concept is well suited to large scale games where it is reflective of the fact that in many large scale games of interest, utilities are greatly affected by changes in the aggregate behavior but less susceptible to changes in the strategy of a particular opposing player. MCE is also well suited to large-scale games in that it can be learned using practical, low-information-overhead behavior rules (e.g. ECFP).

I. INTRODUCTION

Learning in multi-agent environments is an inherently difficult task - multiple interacting agents, all striving simultaneously to learn and adapt to one another tends to result in an environment which, from a single agent's perspective, may often be non-stationary. When the number of agents is large the difficulties are further compounded by large amounts of data, often spatially distributed amongst the agents, and an exponential increase in the size of the state space which may often result in commensurate increases in computational complexity.

Fictitious play (FP) is a game theoretic learning algorithm which has been proven, theoretically, to converge in certain classes of large-scale games [1],[2]. However, it is highly impractical in large-scale settings because of high computational complexity and the large amounts of data required. Empirical centroid fictitious play (ECFP) is a recently proposed variant of FP which is practical in large-scale environments [3],[4]. In ECFP, players are freed from the burden of tracking the empirical distributions of individual agents and instead track and respond to a single centroid distribution; in this way, the major issues associated with large-scale implementation of FP, namely those of large-scale data tracking and high computational complexity, are significantly mitigated. Moreover, the algorithm is demonstrated to be convergent in a distributed environment where inter-agent information exchange is restricted to a local neighborhood of each agent.[1]

The asymptotic behavior of ECFP has been studied in terms of *convergence in average empirical frequency*, where it was demonstrated to converge to a subset of the Nash equilibria (NE) called the consensus equilibria [4]. In this paper we wish to characterize the asymptotic behavior of ECFP in terms of the more traditional notion of *convergence in empirical frequency*.[2] By doing so, we are able to better characterize the limiting behavior of the algorithm and expand the scope of applicability by considering a larger class of games and a richer set of algorithmic limit points.

We introduce the notion of a mean-centric equilibrium (MCE) and show that ECFP converges to the set of MCE. A MCE is not necessarily a NE. Intuitively speaking, in a MCE players are not required to be in equilibrium with the precise strategies of other players, instead each player is in equilibrium with respect to a centroid distribution representing the aggregate strategy. This is a useful notion for large-scale games where it reflects the intuitive fact that, in such games, utilities tend to be greatly influenced by the actions of the aggregate and less dependent on strategies of individual opposing players. More importantly, the subset of consensus equilibria (introduced and studied in [3] and [4]) may be shown to be a subset of the set of MCE.

Our main contributions are the following. (i) We introduce the MCE concept as a simple and intuitive equilibrium notion for large-scale games which can be learned using computationally tractable, low-information-overhead behavior rules (e.g., ECFP). (ii) We supplement the work in [4] by demonstrating convergence of ECFP to the set of MCE in terms of empirical frequencies. The result is proved for a broader class of utility functions, i.e. satisfying **A.2**.

The remainder of the paper is organized as follows. Section II sets up the notation to be used in the subsequent development. In section III the ECFP is presented along with the basic ECFP convergence result. Section IV presents MCE as an equilibrium concept for large-scale games and our main result, convergence of ECFP to the set of MCE, is demonstrated in the same section. Section V concludes the paper.

II. PRELIMINARIES

A normal form game is given by the triple $\Gamma = (N, \{u_i(\cdot)\}_{i \in N}, Y^n)$, where $N = \{1, \ldots, n\}$ represents the

[1]For brevity and simplicity of presentation we omit the technical details related to the distributed implementation of ECFP. We note, however, that the results here presented also hold in a distributed setting (i.e. interagent information exchange is restricted to a local neighborhood of each agent) as given in [4]. More generally, the reader may refer to [5], [6], [7], [8], [4] and references therein for studies on various aspects of distributed interaction and learning in multi-agent networked games.

[2]Convergence of the the empirical frequency distribution to the set of Nash equilibria is one of the tradition ai convergence modes in which mixed equilibria are said to be learned (cf. [9],[10]). The concept is further described in section III.

set of players, Y^n indicates the set of actions, and $u_i(y)$: $Y^n \to \mathbb{R}$ indicates the payoff function for player i. We represent the individual action space of player i by Y_i and the joint action space of all players by $Y^n = \prod_{i=1}^n Y_i$.

The set of mixed strategies for player i is given by $\Delta_i = \{p \in \mathbb{R}^{m_i} : \sum_{k=1}^{m_i} p(k) = 1, \ p(k) \geq 0\}$, the m_i-simplex. A mixed strategy $p_i \in \Delta_i$ may be thought of as a probability distribution from which player i samples to choose an action. In this context, a pure strategy may be thought of as a vertex of the probability simplex. With a slight abuse of notation, we denote the set of actions, or pure strategies, in this context, for player i using the notation $A_i = \{e_1, e_2, \dots e_{m_i}\}$, where m_i is the number of strategies available to player i, and e_j is the jth canonical vector in $\mathbb{R}^{m_i}$. The set of joint mixed strategies is given by $\Delta^n = \prod_{i=1}^n \Delta_i$, and the set of joint actions is given by $A^n = \prod_{i=1}^n A_i$. A joint mixed strategy is given by the n-tuple $(p_1, p_2, \dots, p_n)$, $p_i \in \Delta_i$. In this paper we often make the assumption that players use identical action spaces, in which case we drop the subscript on individual action spaces and write $\Delta = \Delta_i \ \forall i$, $A = A_i \ \forall i$, and $Y = Y_i = \{1, 2, \dots, m\} \ \forall i$.

The mixed utility function for player i is given by the multilinear function $U_i(\cdot) : \Delta^n \to \mathbb{R}$.

$$U_i(p_1, \dots, p_n) := \sum_{y \in Y} u_i(y) p_1(y_1) \dots p_n(y_n).$$

Note that the mixed utility $U_i(p)$ may be interpreted as the expected value of $u_i(y)$ given that the players' mixed strategies are independent. For convenience the notation $U_i(p)$ will often be written as $U_i(p_i, p_{-i})$ where $p_i \in \Delta_i$ is the mixed strategy for player i, and p_{-i} indicates the joint mixed strategy for all players other than i.

Let $\{a_i(t)\}_{t=1}^\infty$ be a sequence of actions for player i, where $a_i(t) \in A_i \ \forall t$. Let $\{a(t)\}_{t=1}^\infty$ be the associated sequence of actions $a(t) \in A^n$. Note that $a_i(t) \in \mathbb{R}^m$; when necessary, we denote the kth element of the vector $a(t)$ by $a(t, k)$. Let $q_i(t)$ be the normalized histogram (empirical distribution) of the actions of player i up to time t, i.e., $q_i(t) = \frac{1}{t} \sum_{s=1}^t a_i(s)$. Similarly, $q(t) = \frac{1}{t} \sum_{s=1}^t a(s)$ is the joint empirical distribution corresponding to the joint actions of the players up to time t. The sequence of distributions $\{q(t)\}_{t=1}^\infty$ is often called a belief sequence.

A mixed strategy p is a Nash equilibrium of Γ if, for each player i, $U_i(p) \geq U_i(g_i, p_{-i}) \ \forall g_i \in \Delta_i$. We define the set of Nash equilibria as $K = \{p \in \Delta^n : U_i(p) \geq U_i(g_i, p_{-i}), \ \forall g_i \in \Delta_i, \ \forall i\}$. We define the set of consensus equilibria as $C = \{p \in K : p_1 = \dots = p_n\}$

The distance of a distribution $p \in \Delta^n$ from a set $S \subset \Delta^n$ is given by

$$d(p, S) = \inf\{\|p - g^\star\| : g^\star \in S\}.$$

Throughout the paper $\|\cdot\|$ denotes the standard $\mathcal{L}_2$ Euclidean norm. Given $\delta > 0$ and a set $S \subset \Delta^n$, we define the set

$$B_\delta(S) = \{p \in \Delta^m : \ d(p, S) < \delta\}.$$

We make the following assumptions unless otherwise stated,

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

A. 2. *The players' utility functions can be decomposed as $u_i(y) = f_i(y_i) + \phi(y)$ where $f_i(y_i)$ depends only on the action of player i and $\phi(y)$ is a permutation-invariant function, identical for all players.*

The mixed extensions of $f_i(\cdot)$ and $\phi(\cdot)$ are given by $F_i(p_i) : \Delta_i \to \mathbb{R}$ and $\Phi(p) : \Delta^n \to \mathbb{R}$ respectively.

Let

$$\bar{q}(t) = \frac{1}{n} \sum_{i=1}^n q_i(t) \tag{1}$$

be the average empirical distribution. Let $\bar{q}^n(t) = (\bar{q}(t), \bar{q}(t), \dots, \bar{q}(t)) \in \Delta^n$ denote the mixed strategy where all players use the empirical average as their individual strategy. For convenience in notation we sometimes write $U_i(\bar{q}_i(t), \bar{q}_{-i})$, where the subscripts indicate the strategy $\bar{q}(t)$ is being used by player i or by all players except i respectively.

We say a mixed strategy p is a mean-centric equilibrium of Γ if, for each player i, $U_i(p_i, \bar{p}_{-i}) \geq U_i(g_i, \bar{p}_{-i}) \ \forall g_i \in \Delta_i$. The set of mean-centric equilibria (MCE) is given by

$$M = \{p \in \Delta^n : \ U_i(p_i, \bar{p}_{-i}) \geq U_i(g_i, \bar{p}_{-i}), \ \forall g_i \in \Delta_i, \ \forall i\}, \tag{2}$$

and the set of ε-MCE is given by $M_\varepsilon = \{p \in \Delta^n : \ U_i(p_i, \bar{p}_{-i}) + \varepsilon \geq U_i(g_i, \bar{p}_{-i}), \ \forall g_i \in \Delta_i, \ \forall i\}$.

III. EMPIRICAL CENTROID FICTITIOUS PLAY

A. Empirical Centroid Fictitious Play

In traditional FP [11],[12] players operate under the naive assumption that others players' empirical distributions accurately represent their respective mixed strategies. This is a simplistic and inaccurate assumption, but allows players to pick strategies within the context of bounded rationality and frees them from the overwhelming computational tasks that would be required of a fully rational player. What is interesting is that despite this inaccurate assumption, players can learn a NE in certain classes of games. In ECFP we extend on this idea in such a way as to maintain computational difficulties at a tractable level when considering large-scale games; loosely speaking, we assume that player i makes the naive assumption that the centroid of the empirical distribution, $\bar{q}(t)$, accurately represents the mixed strategy played independently by each of the other players. Under this assumption, player i need only track a single vector, $\bar{q}(t)$ (see (1)), the centroid of the marginal empirical distributions, a vector whose size is invariant to the number of players in the game. In ECFP, the best response[3] for player i is given by

$$v_i^m(\bar{q}(t)) := \max_{\alpha_i \in A_i} U(\alpha_i, \bar{q}_{-i}(t)), \tag{3}$$

where $\bar{q}_{-i}(t) \in \Delta_{-i}$ is the $(n-1)$-tuple $\bar{q}_{-i}(t) = (\bar{q}_i(t), \dots, \bar{q}_i(t))$. Note that $\bar{q}_{-i}(t)$ essentially represents the assumption that each of the $-i$ (i.e., other) players is playing the strategy $\bar{q}(t)$ independently. A sequence of actions $\{a(t)\}_{t=1}^\infty$ is an *empirical centroid fictitious play process* if

$$v_i^m(\bar{q}(t)) = U(a_i(t+1), \bar{q}_{-i}(t)), \tag{4}$$

where the initial action $a_i(1)$ is chosen arbitrarily for all i. In [4] it was shown that under the following assumption,

A. 3. *All players use identical permutation invariant utility functions.*

---

[3]At first glance, the best response calculation appears to have the same computational complexity as FP. However, the symmetry inherent in tuple $\bar{q}_{-i}(t)$ tends to allow for substantial simplifications in solving the problem (3).

players asymptotically learn a consensus equilibrium[4]. The result is summarized in the following theorem.

Theorem 1. *Let $\{a(t)\}_{t \geq 1}$ be an ECFP process such that A.1 and A.3 hold. Then $\bar{d}(\bar{q}^n(t), C) \to 0$ where $C = \{p \in K : p_1 = \ldots = p_n\}$ is a non-empty subset of the Nash equilibria denoted as the consensus equilibria.*

Note that the mode of convergence in Theorem 1 considers the asymptotic behavior of the n-tuple $\bar{q}^n(t) = (\bar{q}(t), \ldots, \bar{q}(t))$; we call this mode of convergence, *convergence in average empirical frequency*. While Theorem 1 does provide a useful result, the mode of convergence is somewhat esoteric and only provides a limited characterization of the asymptotic behavior of the player empirical frequency distributions, $q_i(t)$. In this paper, we wish to consider the more traditional notion of convergence, *convergence in empirical frequencies* which considers the asymptotic behavior of the tuple of empirical distributions $q(t) = (q_1(t), \ldots, q_n(t))$.

By considering this mode of convergence, we are able to prove convergence results using weaker assumptions and consider a broader class of games. For our main result, Theorem 2, we relax **A.3** in lieu of the more general **A.2**.

IV. MEAN-CENTRIC EQUILIBRIUM

We say that a strategy p is a MCE if

$$U_i(p_i, \bar{p}_{-i}) \geq U_i(g_i, \bar{p}_{-i}), \ \forall g_i \in \Delta_i, \ \forall i. \quad (5)$$

The set of MCE is defined in (2). We demonstrate the set of MCE to be non-empty in the following proposition.

Proposition 1. *Let Γ be a finite normal form game. Then the set of MCE associated with Γ is non-empty.*

Proof: The proof follows similar fixed point type arguments as used in [13] for the existence of mixed Nash equilibria. We sketch the key details here. Define the set-valued mapping $r_i(p) = \arg\max_{g_i \in \Delta_i} U_i(g_i, \bar{p}_{-i}) : \Delta^n \rightrightarrows \Delta_i$ and $r(p) = (r_1(p), \ldots, r_n(p)) : \Delta^n \rightrightarrows \Delta^n$. Note that Δ^n is compact, convex, and nonempty. Also it may be shown that the set-valued mapping $r(p)$ is convex-valued, non-empty for all $p \in \Delta^n$, and upper hemicontinous. It then follows by Kakutani's fixed point theorem that $r(\cdot)$ has a fixed point in Δ^n. Since a strategy p that is a fixed point of $r(p)$ is necessarily in M, we may conclude that the set M is nonempty. ∎

The intuition behind (5) is that, in MCE players are at equilibrium with respect to the aggregate and not necessarily with respect to other individuals. Specifically, each player is at equilibrium with $\bar{q}(t)$, the centroid of all players' marginal mixed strategies.

In many large-scale games of interest, a player's utility tends to depend heavily on the behavior of the aggregate, and less so on the behavior of other individuals. MCE may be considered a simplification of the traditional NE concept such that dependency on any specific opposing players is extenuated.

In general, MCE are not necessarily NE, and NE are not necessarily MCE. However, it is interesting to note that any consensus equilibrium is a MCE. A very interesting future research direction would be to characterize the degree to which

---

[4]A consensus equilibrium is a NE such that $p_1 = p_2 = \ldots = p_n$, i.e., all players use the same strategy.

these two concepts approximate one another as well as the class of games for which the two concepts may be considered 'close'.

A. Main Result

The following theorem summarizes our main result, demonstrating that in an ECFP process, players asymptotically learn a MCE.

Theorem 2. *Let $\{a(t)\}_{t \geq 1}$ be an ECFP process such that assumptions A.1-A.2 hold. Then $d(q(t), M) \to 0$ as $t \to \infty$.*

Proof: Let $\bar{a}(t) = \frac{1}{n} \sum_{i=1}^{n} a_i(t)$, where $\bar{a}(t) \in \Delta, a_i(t) \in A$. Let $\bar{a}^n(t) \in \Delta^n$ be the n-tuple $(\bar{a}(t), \ldots, \bar{a}(t))$.

Note that for $t \geq 1$

$$\bar{q}^n(t+1) = \bar{q}^n(t) + \frac{1}{t+1}\left(\bar{a}^n(t+1) - \bar{q}^n(t)\right). \quad (6)$$

Using (6) we write

$$\Phi(\bar{q}^n(t+1)) = \Phi\left(\bar{q}^n(t) + \frac{1}{t+1}\left(\bar{a}^n(t+1) - \bar{q}^n(t)\right)\right).$$

Applying the multilinearity of $\Phi(\cdot)$, we obtain

$$\Phi(\bar{q}^n(t+1)) = \Phi(\bar{q}^n(t)) + \frac{1}{t+1}\sum_{i=1}^{n}\Phi\left(\bar{a}_i(t+1), \bar{q}_{-i}(t)\right)$$
$$- \frac{1}{t+1}\sum_{i=1}^{n}\Phi\left(\bar{q}_i(t), \bar{q}_{-i}(t)\right) + \zeta(t+1).$$

where we have explicitly written the first order terms of the expansion and collected the remaining terms in $\zeta(t+1)$. Note that the number of second order terms in the above expansion is finite and the terms are uniformly bounded since $\max_{p \in \Delta^n} |\Phi(p)| < \infty$. Hence, there exists a positive constant B (independent of t) large enough such that $|\zeta(t+1| \leq B(t+1)^{-2}$ for all t. Thus,

$$\Phi(\bar{q}^n(t+1)) \geq \Phi(\bar{q}^n(t)) + \frac{1}{t+1}\sum_{i=1}^{n}\Phi\left(\bar{a}_i(t+1), \bar{q}_{-i}(t)\right)$$
$$- \frac{1}{t+1}\sum_{i=1}^{n}\Phi\left(\bar{q}_i(t), \bar{q}_{-i}(t)\right) - \frac{B}{(t+1)^2}. \quad (7)$$

The permutation invariance and multilinearity of $\Phi(\cdot)$ permits a rearranging of terms. We use the notation $[a_j(t)]_i$ to indicate the action of player j at time t being played by player i.

$$\sum_{i=1}^{n}\Phi\left(\bar{a}_i(t+1), \bar{q}_{-i}(t)\right)$$
$$= \sum_{i=1}^{n}\Phi\left(\left[\frac{1}{n}\sum_{j=1}^{n}a_j(t+1)\right]_i, \bar{q}_{-i}(t)\right)$$
$$= \sum_{i=1}^{n}\frac{1}{n}\sum_{j=1}^{n}\Phi\left([a_j(t+1)]_i, \bar{q}_{-i}(t)\right)$$
$$= \sum_{i=1}^{n}\frac{1}{n}\sum_{j=1}^{n}\Phi\left([a_j(t+1)]_j, \bar{q}_{-j}(t)\right)$$
$$= \sum_{j=1}^{n}\Phi\left(a_j(t+1), \bar{q}_{-j}(t)\right).$$

By similar reasoning it also holds that $\sum_{i=1}^{n} \Phi\left(\bar{q}_i(t), \bar{q}_{-i}(t)\right) = \sum_{j=1}^{n} \Phi\left(q_j(t), \bar{q}_{-j}(t)\right)$. Thus (7) can be expressed as,

$$\Phi\left(\bar{q}^n(t+1)\right) - \Phi\left(\bar{q}^n(t)\right) + \frac{B}{(t+1)^2}$$
$$\geq \frac{1}{t+1} \sum_{i=1}^{n} \Phi\left(a_i(t+1), \bar{q}_{-i}(t)\right)$$
$$- \frac{1}{t+1} \sum_{i=1}^{n} \Phi(q_i(t), \bar{q}_{-i}(t)). \tag{8}$$

Using (6) and the linearity of $F_i(\cdot)$ in Δ_i, note that

$$F_i(q_i(t+1)) - F_i(q_i(t)) = \frac{1}{t+1} \left(F_i(a_i(t+1) - F_i(q_i(t))\right). \tag{9}$$

Combining (8) and (9) we get,

$$\sum_{i=1}^{n} \left[F_i(q_i(t+1)) - F_i(q_i(t))\right]$$
$$+ \Phi\left(\bar{q}^n(t+1)\right) - \Phi\left(\bar{q}^n(t)\right) + \frac{B}{(t+1)^2}$$
$$\geq \frac{1}{t+1} \sum_{i=1}^{n} F_i(a_i(t+1)) + \Phi\left(a_i(t+1), \bar{q}_{-i}(t)\right)$$
$$- \frac{1}{t+1} \sum_{i=1}^{n} F_i(q_i(t)) + \Phi(q_i(t), \bar{q}_{-i}(t))$$
$$= \frac{1}{t+1} \sum_{i=1}^{n} \left[U_i(a_i(t+1), \bar{q}_{-i}(t)) - U_i(q_i(t), \bar{q}_{-i}(t))\right]$$
$$= \frac{\alpha_{t+1}}{t+1}, \tag{10}$$

where the first equality comes from the definition of $u_i(\cdot)$ in **A.2**, and the second equality is obtained by defining α_{t+1} as

$$\alpha_{t+1} := \sum_{i=1}^{n} \left[U_i(a_i(t+1), \bar{q}_{-i}(t)) - U_i\left(q_i(t), \bar{q}_{-i}(t)\right)\right].$$

Summing over $1 \leq t \leq T$ in (10),

$$\sum_{i=1}^{n} \left[F_i(q_i(T)) - F_i(q_i(1))\right]$$
$$+ \Phi(\bar{q}^n(T+1)) - \Phi(\bar{q}^n(1)) + \sum_{t=1}^{T} \frac{B}{(t+1)^2}$$
$$\geq \sum_{t=1}^{T} \frac{\alpha_{t+1}}{t+1}.$$

Note that $\sum_{t=1}^{T} \frac{B}{(t+1)^2}$ is summable; therefore all terms on the left hand side are bounded above for all $T \geq 1$. Also note that by the (4), $\alpha_t \geq 0$ and hence it follows that $\sum_{t=2}^{T} \frac{\alpha_t}{t} < \infty$ converges as $T \to \infty$. Applying Kronnecker's lemma yields

$$\lim_{T \to \infty} \frac{\alpha_2 + \alpha_3 + \ldots + \alpha_T}{T} = 0.$$

Subsequently, following reasoning similar to [4] Lemma 6, we obtain for every $\varepsilon > 0$,

$$\lim_{T \to \infty} \frac{\#\{1 \leq t \leq T : q(t) \notin M_\varepsilon\}}{T} = 0.$$

Following reasoning similar to [4] Lemma 8, this is equivalent to

$$\lim_{T \to \infty} \frac{\#\{1 \leq t \leq T : q(t) \notin B_\delta(M)\}}{T} = 0$$

for every $\delta > 0$. Finally, following reasoning similar to [4] Lemma 9, we obtain

$$\lim_{t \to \infty} d(q(t), M) = 0.$$

∎

V. CONCLUSIONS

We introduced the concept of mean-centric equilibrium which we consider to be an intuitive and natural equilibrium concept for large-scale games. We demonstrated that the set of MCE is non-empty for the class of games considered, and provided simple and practical player behavior rules for learning MCE in large-scale games. Specifically, we demonstrated that the empirical centroid fictitious play algorithm converges in terms of empirical frequencies to the set of MCE. By considering MCE as limit points of the ECFP algorithm we were able to better characterize the asymptotic behavior of ECFP and expand the scope of applicablility of ECFP to include games with disparate action sets and non-permutation invariant potential functions. It would be interesting to study the relationship between MCE and NE, including a characterization of the games for which MCE approximate NE, either in terms of player utilities or some other metric.

REFERENCES

[1] D. Monderer and L. S. Shapley, "Fictitious play property for games with identical interests," *Journal of Economic Theory*, vol. 68, no. 1, pp. 258–265, Jan. 1996.
[2] ——, "Potential Games," *Games and Economic Behavior*, vol. 14, no. 1, pp. 124–143, May 1996.
[3] B. Swenson, S. Kar, and J. Xavier, "Distributed learning in large-scale multi-agent games: A modified fictitious play approach," in *46th Asilomar Conference on Signals, Systems, and Computers*, Pacifc Grove, CA, USA, Nov. 4 - 7 2012, pp. 1490 – 1495.
[4] ——, "Empirical centroid fictitious play: an approach for distributed learning in multi-agent games," 2012, submitted for journal publication. Online: http://arxiv.org/pdf/1304.4577v1.pdf.
[5] J. S. Shamma and G. Arslan, "Dynamic fictitious play, dynamic gradient play, and distributed convergence to Nash equilibria," *IEEE Transactions on Automatic Control*, vol. 50, no. 3, pp. 312–327, Mar. 2005.
[6] J. R. Marden, G. Arslan, and J. S. Shamma, "Joint strategy fictitious play with inertia for potential games," *IEEE Transactions on Automatic Control*, vol. 54, no. 2, pp. 208–220, February 2009.
[7] D. S. Leslie and E. J. Collins, "Generalised weakened fictitious play," *Games and Economic Behavior*, vol. 56, no. 2, pp. 285–298, Aug. 2006.
[8] C. Eksin, P. Molavi, A. Ribeiro, and A. Jadbabaie, "Learning in linear games over networks," in *Proceedings of the 50th Annual Allerton Conference on Communications, Control, and Computing*, 2012.
[9] D. Fudenberg, "Learning Mixed Equilibria," *Games and Economic Behavior*, vol. 5, pp. 320–367, 1993.
[10] D. Fudenberg and D. K. Levine, *The theory of learning in games*. MIT press, 1998, vol. 2.
[11] G. W. Brown, *"Iterative Solutions of Games by Fictitious Play" In Activity Analysis of Production and Allocation*, T. Coopmans, Ed. New York: Wiley, 1951.
[12] J. Robinson, "An iterative method of solving a game," *The Annals of Mathematics*, vol. 54, no. 2, pp. 296–301, Sep. 1951.
[13] J. Nash, "Equilibrium Points in n-person Games," *Proceedings of The National Academy of Sciences*, vol. 36, no. 1, pp. 48–49, 1950.